

MapReduce

Communications of the ACM

51, 107-113

DOI: [10.1145/1327452.1327492](https://doi.org/10.1145/1327452.1327492)

Citation Report

#	ARTICLE	IF	CITATIONS
1	Interpreting the Data: Parallel Analysis with Sawzall. Scientific Programming, 2005, 13, 277-298.	0.5	370
2	NetSolve/D: A Massively Parallel Grid Execution System for Scalable Data Intensive Collaboration. , 0, , .		3
3	Towards Scaling Fully Personalized PageRank: Algorithms, Lower Bounds, and Experiments. Internet Mathematics, 2005, 2, 333-358.	0.7	218
4	Using Many Machines to Handle an Enormous Error-Correcting Code. , 0, , .		1
5	Dryad. , 2007, , .		1,417
6	A Programming Model for Distributed Content-Based Image Retrieval. , 2007, , .		3
7	Concordance-Based Entity-Oriented Search. , 2007, , .		6
8	Design, Implementation, and Evaluation of Trellis-SDP for File-Level Data Parallelism. Parallel Processing (ICPP), Proceedings of the International Symposium, 2007, , .	0.0	0
9	Parallel Approximate Matrix Factorization for Kernel Methods. , 2007, , .		0
10	Google Books: making the public domain universally accessible. , 2007, , .		8
11	Detecting near-duplicates for web crawling. , 2007, , .		384
12	Google news personalization. , 2007, , .		1,050
13	High Performance Multi-paradigm Messaging Runtime Integrating Grids and Multicore Systems. , 2007, , .		11
14	Web-Scale Multimedia Data Management: Challenges and Remedies. , 2007, , .		0
15	SymGrid: A Framework for Symbolic Computations on the Grid. , 2007, , .		3
16	Minimum-cost delegation in service composition. Theoretical Computer Science, 2008, 409, 417-431.	0.5	8
18	An approach to a Cloud Computing network. , 2008, , .		115
19	VDM: Virtual Database Management for Distributed Databases and File Systems. , 2008, , .		3

#	ARTICLE	IF	CITATIONS
20	MRBench: A Benchmark for MapReduce Framework. , 2008, , .		48
21	Sindice.com: a document-oriented lookup index for open linked data. International Journal of Metadata, Semantics and Ontologies, 2008, 3, 37.	0.2	230
22	Fast support vector machine training and classification on graphics processors. , 2008, , .		251
23	DisCo: Distributed Co-clustering with Map-Reduce: A Case Study towards Petabyte-Scale End-to-End Mining. , 2008, , .		140
24	Map-reduce as a Programming Model for Custom Computing Machines. , 2008, , .		61
25	Introducing map-reduce to high end computing. , 2008, , .		47
26	MapReduce for Data Intensive Scientific Analyses. , 2008, , .		279
27	A Segmented Bloom Filter Algorithm for Efficient Predictors. , 2008, , .		4
28	Static Load Distribution for Communication Intensive Parallel Computing in Multiclusters. , 2008, , .		1
29	A Library for Asynchronous Concurrent Service Orchestration. , 2008, , .		0
30	Learning Causal Bayesian Network Structures From Experimental Data. Journal of the American Statistical Association, 2008, 103, 778-789.	1.8	123
31	A scalable parallel framework for analyzing terascale molecular dynamics simulation trajectories. , 2008, , .		2
32	Collective Intelligence in Distributed Systems and Semantic Data Visualization. , 2008, , .		0
33	Chinese Web Infrastructure Building: Challenges and Our Roadmap. , 2008, , .		3
34	Exploring data parallelism and locality in wide area networks. , 2008, , .		6
35	Concrete Partial Evaluation in Ruby. , 2008, , .		0
36	EOMT: A Master-Slave Task Scheduling Strategy for Grid Environment. , 2008, , .		1
37	Utilising Located Functions to Model and Optimise Distributed Computations. , 2008, , .		0

#	ARTICLE	IF	CITATIONS
38	An Overview of the Granules Runtime for Cloud Computing. , 2008, , .		14
39	A computational wireless network backplane: Performance in a distributed speaker identification application. , 2008, , .		2
40	Abstractions for concurrent programming in embedded systems. , 2008, , .		1
41	The data analysis computing hierarchy. , 2008, , .		0
42	Living with the Law: Can Automation give us Moore with Less?. , 2008, , .		0
43	The Dynamic Enterprise Bus. , 2008, , .		0
44	COSTES: Continuous spreadsheet-like computations. , 2008, , .		2
45	DECF: A Coarse-Grained Data-Parallel Programming Framework for Seismic Processing. , 2008, , .		2
46	A semantic web architecture for advocate agents to determine preferences and facilitate decision making. , 2008, , .		5
47	Resolving tag ambiguity. , 2008, , .		91
48	Better Scripts, Better Games. Queue, 2008, 6, 18-25.	0.8	6
49	Generating links by mining quotations. , 2008, , .		24
50	Fully distributed EM for very large datasets. , 2008, , .		56
51	Big Social Network Mining for "Following" Patterns. , 2008, , .		6
52	Cut-and-stitch. , 2008, , .		6
53	Firefox (In) security update dynamics exposed. Computer Communication Review, 2008, 39, 16-22.	1.5	27
54	Exploring a digital library through key ideas. , 2008, , .		8
55	p2pBIOFOCO: Proposing a Peer-to-Peer System for Distributed BLAST Execution. , 2008, , .		0

#	ARTICLE	IF	CITATIONS
56	A self-diagnosis approach for performance problem localization in component-based applications. , 2008, , .		0
57	ZettaDS: A Light-weight Distributed Storage System for Cluster. , 2008, , .		10
58	SQMD: Architecture for Scalable, Distributed Database System Built on Virtual Private Servers. , 2008, , .		2
59	Using Small Abstractions to Program Large Distributed Systems. , 2008, , .		1
60	GENI - global environment for network innovations. , 2008, , .		98
61	Low-pain, high-gain multicore programming in Haskell. , 2008, , .		4
62	Distributed online aggregations. Proceedings of the VLDB Endowment, 2009, 2, 443-454.	2.1	49
63	Mining document collections to facilitate accurate approximate entity matching. Proceedings of the VLDB Endowment, 2009, 2, 395-406.	2.1	26
64	Computing median values in a cloud environment using GridBatch and MapReduce. , 2009, , .		1
65	MITHRA: Multiple data independent tasks on a heterogeneous resource architecture. , 2009, , .		40
66	A Reference Model of Cloud Operating and Open Source Software Implementation Mapping. , 2009, , .		2
67	Analysis of hyperspectral data with diffusion maps and Fuzzy ART. , 2009, , .		6
68	A Web Service Framework for Distributed Multimedia Content Computing. , 2009, , .		0
69	Update Maps – A New Abstraction for High-Throughput Batch Processing. , 2009, , .		1
70	ppBLAST: A Computational Service over Peer-to-Peer Network for BLAST. , 2009, , .		0
71	Modeling MapReduce with CSP. , 2009, , .		7
72	Fault–Tolerant Master-Workers Framework for MapReduce Applications. , 2009, , .		0
73	Cloud Computing for Satellite Data Processing on High End Compute Clusters. , 2009, , .		52

#	ARTICLE	IF	CITATIONS
74	Supporting load balancing for distributed data-intensive applications. , 2009, , .		6
75	Design and implementation of an automatic parallelization of state space generation. , 2009, , .		0
76	Comparing and Optimising Parallel Haskell Implementations for Multicore Machines. , 2009, , .		2
77	Multi-dimensional characterization of temporal data mining on graphics processors. , 2009, , .		6
78	Towards a distributed execution framework for JUnit test cases. , 2009, , .		23
79	NIR: Content based image retrieval on cloud computing. , 2009, , .		22
80	Parallel and distributed approach for processing large-scale XML datasets. , 2009, , .		17
81	GrayWulf: Scalable Software Architecture for Data Intensive Computing. , 2009, , .		6
82	CloudBurst: highly sensitive read mapping with MapReduce. Bioinformatics, 2009, 25, 1363-1369.	1.8	551
83	Implementing Parallel Google Map-Reduce in Eden. Lecture Notes in Computer Science, 2009, , 990-1002.	1.0	13
84	A Quick Guide for Developing Effective Bioinformatics Programming Skills. PLoS Computational Biology, 2009, 5, e1000589.	1.5	44
85	Using Hadoop as a grid storage element. Journal of Physics: Conference Series, 2009, 180, 012047.	0.3	11
86	Applying Web Service to Distributed Multimedia Content Analysis. , 2009, , .		0
87	Distributed Multimedia Content Processing Based on Web Service. , 2009, , .		3
88	Large-scale multimedia semantic concept modeling using robust subspace bagging and MapReduce. , 2009, , .		31
89	Improving the responsiveness of internet services with automatic cache placement. , 2009, , .		3
91	Towards an Algebraic foundation for business planning. , 2009, , .		0
92	DiskReduce. , 2009, , .		107

#	ARTICLE	IF	CITATIONS
93	Quincy. , 2009, , .		598
94	Web service query service. , 2009, , .		1
95	TANGENT. , 2009, , .		83
96	Query processing of massive trajectory data based on mapreduce. , 2009, , .		62
97	Mixing Hadoop and HPC workloads on parallel filesystems. , 2009, , .		11
98	SPIDER. , 2009, , .		24
99	Toward visual analysis of ensemble data sets. , 2009, , .		20
100	Hadoop at home. SIGCSE Bulletin, 2009, 41, 106-110.	0.1	15
101	Spending Moore's dividend. Communications of the ACM, 2009, 52, 62-69.	3.3	54
102	Concordance-based entity-oriented search. Web Intelligence and Agent Systems, 2009, 7, 303-319.	0.4	5
103	A Data Accessing Method in Distributed Massive Computing. , 2009, , .		0
104	A Cloud Computing Platform for Large-Scale Forensic Computing. IFIP Advances in Information and Communication Technology, 2009, , 201-214.	0.5	31
105	Large-scale deep unsupervised learning using graphics processors. , 2009, , .		377
106	Using realistic simulation for performance analysis of mapreduce setups. , 2009, , .		62
107	The Changing Paradigm of Data-Intensive Computing. Computer, 2009, 42, 26-34.	1.2	105
108	Parallel Scripting for Applications at the Petascale and Beyond. Computer, 2009, 42, 50-60.	1.2	78
109	Graph Twiddling in a MapReduce World. Computing in Science and Engineering, 2009, 11, 29-41.	1.2	317
110	<i>Mining@home</i>: toward a public resource computing framework for distributed data mining. Concurrency Computation Practice and Experience, 2010, 22, 658-682.	1.4	3

#	ARTICLE	IF	CITATIONS
111	What the parallel-processing community has (failed) to offer the multi/many-core generation. Journal of Parallel and Distributed Computing, 2009, 69, 807-812.	2.7	1
112	Compute and storage clouds using wide area high performance networks. Future Generation Computer Systems, 2009, 25, 179-183.	4.9	142
113	Adaptive workload allocation in query processing in Autonomous heterogeneous environments. Distributed and Parallel Databases, 2009, 25, 125-164.	1.0	12
114	A Cloud Computing Platform Based on P2P. , 2009, , .		25
115	Situation awareness via abductive reasoning from Semantic Sensor data: A preliminary report. , 2009, , .		18
116	Computer vision on tap. , 2009, , .		3
117	Distributed Data Mining Tasks and Patterns as Services. Lecture Notes in Computer Science, 2009, , 415-422.	1.0	4
118	Understanding TCP incast throughput collapse in datacenter networks. , 2009, , .		274
119	The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines. Synthesis Lectures on Computer Architecture, 2009, 4, 1-108.	1.3	618
120	A Privacy Manager for Cloud Computing. Lecture Notes in Computer Science, 2009, , 90-106.	1.0	129
121	Stochastic gradient boosted distributed decision trees. , 2009, , .		197
122	Resource co-allocation for large-scale distributed environments. , 2009, , .		13
123	High-Performance Cloud Computing: A View of Scientific Applications. , 2009, , .		271
124	The Case for Cloud Computing. IT Professional, 2009, 11, 23-27.	1.4	353
125	A Taxonomy and Survey of Cloud Computing Systems. , 2009, , .		829
126	Optical switching for next generation data centers. , 2009, , .		5
127	Scaling Genetic Algorithms Using MapReduce. , 2009, , .		125
128	MapReduce as a general framework to support research in Mining Software Repositories (MSR). , 2009, , .		22

#	ARTICLE	IF	CITATIONS
129	Hadoop at home. , 2009, , .		27
130	Online Risk Analytics on the Cloud. , 2009, , .		26
131	XConveryer: Guarantee Hadoop Throughput via Lightweight OS-Level Virtualization. , 2009, , .		3
132	Collaborative filtering for orkut communities. , 2009, , .		113
133	Exploring many task computing in scientific workflows. , 2009, , .		24
134	Tour the world. , 2009, , .		79
135	An efficient multi-dimensional index for cloud data management. , 2009, , .		76
136	Large-scale behavioral targeting. , 2009, , .		143
137	Spatial Queries Evaluation with MapReduce. , 2009, , .		65
138	Calibrating Resource Allocation for Parallel Processing of Analytic Tasks. , 2009, , .		5
139	Better scripts, better games. Communications of the ACM, 2009, 52, 42-47.	3.3	63
140	Experiment setup for temporal distributed intrusion detection system on Amazon's elastic compute cloud. , 2009, , .		5
141	Distributed Algorithm for Computing Formal Concepts Using Map-Reduce Framework. Lecture Notes in Computer Science, 2009, , 333-344.	1.0	29
142	Resource Management Using Virtual Ontologies for Scientific Applications. , 2009, , .		0
143	Low-cost management of inverted files for online full-text search. , 2009, , .		12
144	What can visual content analysis do for text based image search?. , 2009, , .		8
145	Adaptive Parallelization of Queries over Dependent Web Service Calls. Proceedings - International Conference on Data Engineering, 2009, , .	0.0	14
146	A Conceptual Approach for Taxonomical Spectrum of Cloud Computing. , 2009, , .		9

#	ARTICLE	IF	CITATIONS
147	Dynamic load balancing for I/O-intensive applications on clusters. ACM Transactions on Storage, 2009, 5, 1-38.	1.4	63
148	Accelerating MapReduce with Distributed Memory Cache. , 2009, , .		29
149	DeLight: A Peer-to-Peer Storage and Processing System. , 2009, , .		0
150	DryadLINQ for Scientific Analyses. , 2009, , .		30
151	Automatic optimisation of MapReduce designs by geometric programming. , 2009, , .		8
152	Dynamically reconfigurable hybrid optical / electrical networks for high-bandwidth data centers. , 2009, , .		4
153	Networking Challenges and Resultant Approaches for Large Scale Cloud Construction. , 2009, , .		8
154	Horde: A parallel programming framework for clusters. , 2009, , .		0
155	A Cloud Computing Based Real Time Financial System. , 2009, , .		9
156	Investigating the use of autonomic cloudbursts for high-throughput medical image registration. , 2009, 2009, 34-41.		27
157	Speeding Up Distributed MapReduce Applications Using Hardware Accelerators. , 2009, , .		10
158	SolutionFinder: Intelligent Knowledge Integration and Dissemination for Solution Retrieval in IT Support Services. , 2009, , .		0
159	Challenges Towards Elastic Power Management in Internet Data Centers. , 2009, , .		77
160	Smart read/write for MPHIO. , 2009, , .		0
161	Refinements to task control in the Aspect-Oriented Programming model for computing grids. , 2009, , .		2
162	The Nornir Run-time System for Parallel Programs Using Kahn Process Networks. , 2009, , .		3
163	Cluster fault-tolerance: An experimental evaluation of checkpointing and MapReduce through simulation. , 2009, , .		7
164	A model for fast web mining prototyping. , 2009, , .		5

#	ARTICLE	IF	CITATIONS
165	Combinatorial Framework for Similarity Search. , 2009, , .		4
166	Cogset: A Unified Engine for Reliable Storage and Parallel Processing. , 2009, , .		9
167	Enabling Computational Steering with an Asynchronous-Iterative Computation Framework. , 2009, , .		2
168	Optimistic Synchronization of Parallel Simulations in Cloud Computing Environments. , 2009, , .		28
169	Performance optimizations for distributed real-time text indexing. , 2009, , .		0
170	Towards unlocking web video: Automatic people tracking and clustering. , 2009, , .		1
171	A Coordination Language for Programming Embedded Multi-Core Systems. , 2009, , .		4
172	The multikernel. , 2009, , .		532
173	GePSeA: A General-Purpose Software Acceleration Framework for Lightweight Task Offloading. , 2009, , .		9
174	Cagra Dependable Distributed Storage System for 3D Computer Graphics Rendering. , 2009, , .		3
175	Practical lessons of data mining at Yahoo!. , 2009, , .		5
176	Application Level Interoperability between Clouds and Grids. , 2009, , .		17
177	MBMF: A framework for macroprogramming data-centric sensor network applications using the Bird-Meertens Formalism. , 2009, , .		1
178	Server Designs for Warehouse-Computing Environments. IEEE Micro, 2009, 29, 41-49.	1.8	6
179	Granules: A lightweight, streaming runtime for cloud computing with support, for Map-Reduce. , 2009, , .		43
180	A simulation approach to evaluating design decisions in MapReduce setups. , 2009, , .		70
181	An Efficient Hierarchical Clustering Method for Large Datasets with Map-Reduce. , 2009, , .		29
182	SJMR: Parallelizing spatial join with MapReduce on clusters. , 2009, , .		93

#	ARTICLE	IF	CITATIONS
183	Designing efficient sorting algorithms for manycore GPUs. , 2009, , .		373
184	An improved staged event driven architecture for Master-Worker network computing. , 2009, , .		3
185	Parallel K-Means Clustering Based on MapReduce. Lecture Notes in Computer Science, 2009, , 674-679.	1.0	378
186	Parallel Large Scale Feature Selection for Logistic Regression. , 2009, , .		42
187	ACM SIGACT news distributed computing column 34. ACM SIGACT News, 2009, 40, 67-67.	0.1	0
188	Machine models for query processing. SIGMOD Record, 2009, 38, 18-28.	0.7	2
189	Font rendering on a GPU-based raster image processor. , 2010, , .		1
190	Implementing and Optimizing Multiple Group by Query in a MapReduce Approach. Journal of Algorithms and Computational Technology, 2010, 4, 183-205.	0.4	3
191	Knuckles: bringing the database to the data. International Journal of Computational Science and Engineering, 2010, 5, 214.	0.4	0
192	Adaptive method to support effective searching over large-scale web documents. , 2010, , .		0
193	A unified execution model for cloud computing. Operating Systems Review (ACM), 2010, 44, 12-17.	1.5	3
194	Understanding performance of distributed data-intensive applications. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 2010, 368, 4089-4102.	1.6	4
195	Automated web service query service. International Journal of Web and Grid Services, 2010, 6, 400.	0.4	4
196	High-Performance Parallel Computing with Cloud and Cloud Technologies. , 2010, , 275-308.		18
198	Query processing in a DBMS for cluster systems. Programming and Computer Software, 2010, 36, 205-215.	0.5	6
199	Trends in Trends in Functional Programming 1999/2000 versus 2007/2008. Higher-Order and Symbolic Computation, 2010, 23, 465-487.	0.3	2
200	Cognitive Radio Rides on the Cloud. , 2010, , .		17
201	Cloud Computing: a Perspective Study. New Generation Computing, 2010, 28, 137-146.	2.5	550

#	ARTICLE	IF	CITATIONS
202	Special Issue on Information Explosion. New Generation Computing, 2010, 28, 207-215.	2.5	6
203	Parallelizing XML data-streaming workflows via MapReduce. Journal of Computer and System Sciences, 2010, 76, 447-463.	0.9	23
204	The unique strengths and storage access characteristics ofÂdiscard-based search. Journal of Internet Services and Applications, 2010, 1, 31-44.	1.6	5
205	Multi-source geospatial information integration and sharing in Virtual Globes. Science China Technological Sciences, 2010, 53, 1-6.	2.0	15
206	An introduction to Tsinghua Cloud. Science China Information Sciences, 2010, 53, 1481-1486.	2.7	7
207	Weighted fair share scheduling for loosely controlled concurrent event systems. Service Oriented Computing and Applications, 2010, 4, 45-62.	1.3	1
208	Three challenges in data mining. Frontiers of Computer Science, 2010, 4, 324-333.	0.6	9
209	Geography and computational social science. Geo Journal, 2010, 75, 133-148.	1.7	53
210	Participatory sensing: applications and architecture [Internet Predictions]. IEEE Internet Computing, 2010, 14, 12-42.	3.2	124
211	Google-Wide Profiling: A Continuous Profiling Infrastructure for Data Centers. IEEE Micro, 2010, 30, 65-79.	1.8	133
212	Scale-Out Networking in the Data Center. IEEE Micro, 2010, 30, 29-41.	1.8	102
213	Server Engineering Insights for Large-Scale Online Services. IEEE Micro, 2010, 30, 8-19.	1.8	91
214	MrsRF: an efficient MapReduce algorithm for analyzing large collections of evolutionary trees. BMC Bioinformatics, 2010, 11, S15.	1.2	37
215	SeqWare Query Engine: storing and searching sequence data in the cloud. BMC Bioinformatics, 2010, 11, S2.	1.2	78
216	A data placement strategy in scientific cloud workflows. Future Generation Computer Systems, 2010, 26, 1200-1214.	4.9	293
217	APHID: An architecture for private, high-performance integrated data mining. Future Generation Computer Systems, 2010, 26, 891-904.	4.9	8
218	VDB-MR: MapReduce-based distributed data integration using virtual database. Future Generation Computer Systems, 2010, 26, 1418-1425.	4.9	8
219	A survey of algorithmic skeleton frameworks: high-level structured parallel programming enablers. Software - Practice and Experience, 2010, 40, 1135-1160.	2.5	137

#	ARTICLE	IF	CITATIONS
220	An efficient parallelization of phosphorylated peptide and protein identification. Rapid Communications in Mass Spectrometry, 2010, 24, 1791-1798.	0.7	10
221	From a stream of relational queries to distributed stream processing. Proceedings of the VLDB Endowment, 2010, 3, 1394-1405.	2.1	18
222	Training Phrase-Based Machine Translation Models on the Cloud: Open Source Machine Translation Toolkit Chaski. Prague Bulletin of Mathematical Linguistics, 2010, 93, 37-46.	0.7	9
223	MACH: Fast Randomized Tensor Decompositions. , 2010, , .		33
224	Parallel Collision Detection Algorithm Based on OBB Tree and MapReduce. Lecture Notes in Computer Science, 2010, , 610-620.	1.0	2
225	Distributed Biological Datasets Index Model. Advanced Materials Research, 2010, 143-144, 599-603.	0.3	0
226	laUD-MS: An Extensible System for Unstructured Data Management. , 2010, , .		1
227	SLA-Based Resource Co-Allocation in Multi-Cluster Grid. , 2010, , .		1
228	Concurrent Architecture for Automated Malware Classification. , 2010, , .		4
229	Attribute Reduction for Massive Data Based on Rough Set Theory and MapReduce. Lecture Notes in Computer Science, 2010, , 672-678.	1.0	23
230	Parallel K-Means Clustering of Remote Sensing Images Based on MapReduce. Lecture Notes in Computer Science, 2010, , 162-170.	1.0	79
231	Study on Service-Oriented Cloud Conferencing. , 2010, , .		5
232	An Interface Design for Future Cloud-Based Visualization Services. , 2010, , .		6
233	BlobSeer: Efficient data management for data-intensive applications distributed at large-scale. , 2010, , .		2
234	Nitrite Accumulation During Wastewater Denitrification. , 2010, , .		2
235	Towards a Theory of Cloud Computing in the Protocol-Composition Framework. , 2010, , .		1
236	A design of RESTful style digital gazetteer service in cloud computing environment. , 2010, , .		6
237	BlobSeer: Bringing high throughput under heavy concurrency to Hadoop Map-Reduce applications. , 2010, , .		44

#	ARTICLE	IF	CITATIONS
238	S4: Distributed Stream Computing Platform. , 2010, , .		655
239	Semi-supervised PLSA for Document Clustering. , 2010, , .		9
240	Under the Cloud: A Novel Content Addressable Data Framework for Cloud Parallelization to Create and Virtualize New Breeds of Cloud Applications. , 2010, , .		3
241	Parallelized Boosting with Map-Reduce. , 2010, , .		5
242	Scalable Analysis of Massive Graphs on a Parallel Data Flow System. , 2010, , .		0
243	Complex task query based on cloud computing Resource: Improved approximate skyline algorithm. , 2010, , .		1
244	Applying Twister to Scientific Applications. , 2010, , .		25
245	EST Clustering in Large Dataset with MapReduce. , 2010, , .		1
246	Machine learning and windowed subsecond event detection on PMU data via Hadoop and the openPDC. , 2010, , .		26
247	Scalable parallel co-clustering over multiple heterogeneous data types. , 2010, , .		4
248	Luth: Composing and Parallelizing Midpoint Inspection Devices. , 2010, , .		1
249	Experiments with Storage and Preservation of NASA's Planetary Data via the Cloud. IT Professional, 2010, 12, 28-35.	1.4	11
250	Towards a Virtual Research Environment for Language and Literature Researchers. , 2010, , .		2
251	Storage and index support for data intensive web applications. , 2010, , .		1
252	Construction of OSS-based E-Learning cloud in China. , 2010, , .		4
253	TMemCanal: A VM-oblivious Dynamic Memory Optimization Scheme for Virtual Machines in Cloud Computing. , 2010, , .		7
254	Improving query suggestion by utilizing user intent. , 2010, , .		5
255	Gossamer: A Lightweight Approach to Using Multicore Machines. , 2010, , .		4

#	ARTICLE	IF	CITATIONS
256	Coordination and Deployment of Mobile Agents on Dependable Systems. , 2010, , .		0
257	Scalable RDF store based on HBase and MapReduce. , 2010, , .		31
258	Watermarking Massive Remote Sensor Images in Parallel. , 2010, , .		2
259	Parallel Sequential Pattern Mining of Massive Trajectory Data. International Journal of Computational Intelligence Systems, 2010, 3, 343-356.	1.6	32
260	A prototype implementation of Map/Reduce Model on Message-Oriented Middleware. , 2010, , .		0
261	Using Index in the MapReduce Framework. , 2010, , .		11
262	Data-Intensive Text Processing with MapReduce. Synthesis Lectures on Human Language Technologies, 2010, 3, 1-177.	2.3	220
263	A Collectional Data Model for Scientific Workflow Composition. , 2010, , .		3
264	Applications in Data-Intensive Computing. Advances in Computers, 2010, , 1-70.	1.2	1
265	SecondNet. , 2010, , .		439
266	MOON. , 2010, , .		118
267	CAMEO: Enabling social networks for Massively Multiplayer Online Games through Continuous Analytics and cloud computing. , 2010, , .		9
268	A Capabilities-Aware Programming Model for Asymmetric High-End Systems. , 2010, , .		0
269	Executing multiple group-by query in a MapReduce approach. , 2010, , .		3
270	Application of FWT in Dielectric Loss Angle for On-line Measurement. , 2010, , .		0
271	Spatial indexing of global geographical data with HTM. , 2010, , .		2
272	A relaxed synchronization primitive for macroprogramming systems. , 2010, , .		1
273	Fluid evolutionary algorithms. , 2010, , .		9

#	ARTICLE	IF	CITATIONS
274	Cloud-MAQ: The cloud-enabled scalable whole genome reference Assembly application. , 2010, , .		2
275	Usage Patterns to Provision for Scientific Experimentation in Clouds. , 2010, , .		3
276	Exporting kernel page caching for efficient user-level I/O. , 2010, , .		2
277	Cloud-TM. Operating Systems Review (ACM), 2010, 44, 1-6.	1.5	42
278	Exploring application and infrastructure adaptation on hybrid grid-cloud infrastructure. , 2010, , .		19
279	Real-life performance of metric searching. SIGSPATIAL Special, 2010, 2, 28-31.	2.5	0
280	Clouds at the crossroads. Xrds, 2010, 16, 10-13.	0.2	6
281	Mining advertiser-specific user behavior using adfactors. , 2010, , .		18
282	Comet. , 2010, , .		114
283	Data parallelism in bioinformatics workflows using Hydra. , 2010, , .		15
284	Improving the Hadoop map/reduce framework to support concurrent appends through the BlobSeer BLOB management system. , 2010, , .		2
285	Generic multiset programming for language-integrated querying. , 2010, , .		3
286	A middleware for parallel processing of large graphs. , 2010, , .		1
288	Measuring the user experience on a large scale. , 2010, , .		147
289	Towards supporting the life cycle of large scale scientific experiments. International Journal of Business Process Integration and Management, 2010, 5, 79.	0.2	75
290	Social Services Computing: Concepts, Research Challenges, and Directions. , 2010, , .		4
291	A novel metadata management scheme in Cloud Computing. , 2010, , .		2
292	The impact of management operations on the virtualized datacenter. Computer Architecture News, 2010, 38, 326-337.	2.5	11

#	ARTICLE	IF	CITATIONS
293	Stateful bulk processing for incremental analytics. , 2010, , .		120
294	Tiled-MapReduce. , 2010, , .		92
295	Max-cover in map-reduce. , 2010, , .		62
296	A heterogeneous parallel system running open mpi on a broadband network of embedded set-top devices. , 2010, , .		5
297	The impact of management operations on the virtualized datacenter. , 2010, , .		61
298	XCo. , 2010, , .		2
299	A parallel Hyper-Surface Classifier for high dimensional data. , 2010, , .		3
301	Malstone. , 2010, , .		25
302	An overview of the Open Science Data Cloud. , 2010, , .		29
303	Cloud computing paradigms for pleasingly parallel biomedical applications. , 2010, , .		47
304	Parallelizing multiple group-by query in share-nothing environment. , 2010, , .		11
305	Data Dissemination supporting collaborative complex event processing. , 2010, , .		1
306	Distributed indexing for semantic search. , 2010, , .		7
308	Wisdom of the ages. , 2010, , .		18
309	Multi-view clustering with constraint propagation for learning with an incomplete mapping between views. , 2010, , .		21
310	Massive structured data management solution. , 2010, , .		0
311	Controlling your TV with gestures. , 2010, , .		33
312	Web search/browse log mining. , 2010, , .		3

#	ARTICLE	IF	CITATIONS
313	Distributed indexing of web scale datasets for the cloud. , 2010, , .		18
314	PeerWatch. , 2010, , .		27
315	Multicore education. , 2010, , .		3
316	Comparing the usability of library vs. language approaches to task parallelism. , 2010, , .		9
317	KAdvice. , 2010, , .		0
318	PH2. , 2010, , .		8
319	A marketplace for cloud resources. , 2010, , .		13
320	A service-oriented architecture for on-demand and dynamic ally reconfigurable computing. , 2010, , .		1
321	FlumeJava. , 2010, , .		210
322	Beyond online aggregation. , 2010, , .		21
323	A Comparative Study of Data Storage and Processing Architectures for the Smart Grid. , 2010, , .		32
324	Research of P2P architecture based on cloud computing. , 2010, , .		6
325	Exploring Social Contexts along the Time Dimension: Temporal Analysis of Named Entities. , 2010, , .		0
326	Resource allocation across multiple cloud data centres. , 2010, , .		38
327	A Real-Time Workload Driven Approach for the Cloud. , 2010, , .		0
328	XCo: Explicit Coordination for Preventing Congestion in Data Center Ethernet. , 2010, , .		6
329	Optimizing the utilization of virtual resources in Cloud environment. , 2010, , .		28
330	A Framework of Scientific Workflow Management Systems for Multi-tenant Cloud Orchestration Environment. , 2010, , .		8

#	ARTICLE	IF	CITATIONS
331	Rearchitecting MapReduce for Heterogeneous Multicore Processors with Explicitly Managed Memories. , 2010, , .		8
332	LEMO-MR: Low Overhead and Elastic MapReduce Implementation Optimized for Memory and CPU-Intensive Applications. , 2010, , .		25
333	MapReduce in the Clouds for Science. , 2010, , .		107
334	Integrating DBMSs as a Read-Only Execution Layer into Hadoop. , 2010, , .		4
335	Providing Scalable Database Services on the Cloud. Lecture Notes in Computer Science, 2010, , 1-19.	1.0	15
336	Parallel ISODATA Clustering of Remote Sensing Images Based on MapReduce. , 2010, , .		31
337	Improving MapReduce fault tolerance in the cloud. , 2010, , .		16
338	Statistics-driven workload modeling for the Cloud. , 2010, , .		134
339	A SVM-Based Audio Event Detection System. , 2010, , .		12
340	SciCumulus: A Lightweight Cloud Middleware to Explore Many Task Computing Paradigm in Scientific Workflows. , 2010, , .		89
341	Delay scheduling. , 2010, , .		1,041
342	A High Performance, Low-Cost Network Architecture for Data Centers. , 2010, , .		1
343	The Comparison of the Relative Entropy for Intrusion Detection on CPU and GPU. , 2010, , .		3
344	Self-Consistent MPI Performance Guidelines. IEEE Transactions on Parallel and Distributed Systems, 2010, 21, 698-709.	4.0	26
345	An efficient parallel storage scheme in data grid environment. , 2010, , .		0
346	Integrating MapReduce and RDBMSs. , 2010, , .		6
347	On the energy (in)efficiency of Hadoop clusters. Operating Systems Review (ACM), 2010, 44, 61-65.	1.5	248
348	Improving Many-Task computing in scientific workflows using P2P techniques. , 2010, , .		1

#	ARTICLE	IF	CITATIONS
349	Metabolic Flux Analysis in the Cloud. , 2010, , .		12
350	FlumeJava. ACM SIGPLAN Notices, 2010, 45, 363-375.	0.2	86
351	Constructing Classification Model with MapReduce. , 2010, , .		0
352	Using wordmap and score-based weight in opinion mining with mapreduce. , 2010, , .		1
353	An experience report on scaling tools for mining software repositories using MapReduce. , 2010, , .		24
354	Performance-driven task co-scheduling for MapReduce environments. , 2010, , .		125
355	Behavioral Targeting. ACM Transactions on Knowledge Discovery From Data, 2010, 4, 1-31.	2.5	3
356	Coupled semi-supervised learning for information extraction. , 2010, , .		257
357	Transfer data optimization in parallel computing. , 2010, , .		0
358	The Unreasonable Ineffectiveness of Security Engineering: An Overview. , 2010, , .		2
359	Retrospective detection of malware attacks by cloud computing. , 2010, , .		12
360	Speculative pipelining for compute cloud programming. , 2010, , .		4
361	Evolving agent-based simulations in the clouds. , 2010, , .		5
362	User-Based Collaborative-Filtering Recommendation Algorithms on Hadoop. , 2010, , .		139
363	Using Cloud Technologies to Optimize Data-Intensive Service Applications. , 2010, , .		9
364	Distributed community detection in social networks with genetic algorithms. , 2010, , .		10
365	Community Mining in Complex Network Based on Parallel Genetic Algorithm. , 2010, , .		0
366	The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 2010, 20, 1297-1303.	2.4	21,358

#	ARTICLE	IF	CITATIONS
367	A Taxonomy, Survey, and Issues of Cloud Computing Ecosystems. Computer Communications and Networks, 2010, , 21-46.	0.8	43
368	On Two-Dimensional Sparse Matrix Partitioning: Models, Methods, and a Recipe. SIAM Journal of Scientific Computing, 2010, 32, 656-683.	1.3	85
369	SCAREX: A framework for scalable, reliable, and extendable cluster computing. , 2010, , .		2
370	Evaluating Point-Based POMDP Solvers on Multicore Machines. IEEE Transactions on Systems, Man, and Cybernetics, 2010, 40, 1062-1074.	5.5	9
371	dSimpleGraph: A Novel Distributed Clustering Algorithm for Exploring Very Large Scale Unknown Data Sets. , 2010, , .		3
372	A MapReduce based distributed LSI. , 2010, , .		5
373	High Performance Parallel Computing with Clouds and Cloud Technologies. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2010, , 20-38.	0.2	86
374	Deployment models: Towards eliminating security concerns from cloud computing. , 2010, , .		24
375	Prototype of semantic search engine using ontology. , 2010, , .		3
376	Exploration of cloud computing technologies for geographic information services. , 2010, , .		9
377	Sustainable Network Resource Management System for Virtual Private Clouds. , 2010, , .		32
378	An optimized strategy for cloud computing architecture. , 2010, , .		2
379	Ontology management in cloud computing. , 2010, , .		0
380	Accelerating Spatial Data Processing with MapReduce. , 2010, , .		36
381	Predicting the Performance of Virtual Machine Migration. , 2010, , .		211
382	Visual, Log-Based Causal Tracing for Performance Debugging of MapReduce Systems. , 2010, , .		36
383	X-RIME: Hadoop-based large-scale Social Network Analysis. , 2010, , .		8
384	Distributed log information processing with Map-Reduce: A case study from raw data to final models. , 2010, , .		1

#	ARTICLE	IF	CITATIONS
385	Characterization of Hadoop Jobs Using Unsupervised Learning. , 2010, , .		18
386	Benchmarking a MapReduce Environment on a Full Virtualisation Platform. , 2010, , .		13
387	Exploiting multi-level parallelism for low-latency activity recognition in streaming video. , 2010, , .		18
388	Run-time workload driven structure in statistics for the cloud. , 2010, , .		2
389	Authority Control Approach of Run-Time Workload Driven Structure in Statistics for the Cloud. , 2010, , .		0
390	Private Searching on MapReduce. Lecture Notes in Computer Science, 2010, , 93-101.	1.0	3
391	A Mapreduce programming framework using message passing. , 2010, , .		5
392	Handbook of FPGA Design Security. , 2010, , .		33
393	Multidimensional Arrays for Warehousing Data on Clouds. , 2010, , 26-37.		9
394	Performance Considerations of Data Acquisition in Hadoop System. , 2010, , .		11
395	A 2-Tier Clustering Algorithm with Map-Reduce. , 2010, , .		11
396	Screening Data for Phylogenetic Analysis of Land Plants: A Parallel Approach. , 2010, , .		2
397	Howdah - A Flexible Pipeline Framework for Analyzing Genomic Data. , 2010, , .		0
398	Adaptation of Apriori to MapReduce to Build a Warehouse of Relations between Named Entities across the Web. , 2010, , .		12
399	Kahuna: Problem diagnosis for Mapreduce-based cloud computing environments. , 2010, , .		13
400	System anomaly detection in distributed systems through MapReduce-Based log analysis. , 2010, , .		2
401	Distributed, Scalable Clustering for Detecting Halos in Terascale Astronomy Datasets. , 2010, , .		1
402	Cogset vs. Hadoop: Measurements and Analysis. , 2010, , .		5

#	ARTICLE	IF	CITATIONS
403	AMREF: An Adaptive MapReduce Framework for Real Time Applications. , 2010, , .		8
404	Initial Findings for Provisioning Variation in Cloud Computing. , 2010, , .		44
405	Efficient Nash equilibrium based cloud resource allocation by using a continuous double auction. , 2010, , .		13
406	Large-scale distributed storage for highly concurrent Mapreduce applications. , 2010, , .		1
407	Large scale complex network analysis using the hybrid combination of a MapReduce cluster and a highly multithreaded system. , 2010, , .		3
408	Optimizing joins in a map-reduce environment. , 2010, , .		226
409	An Analysis of Traces from a Production MapReduce Cluster. , 2010, , .		226
410	Improving MapReduce performance through data placement in heterogeneous Hadoop clusters. , 2010, , .		126
411	Handbook of Cloud Computing. , 2010, , .		263
412	An improved schedule of MapReduce programming environment in cloud computing. , 2010, , .		0
413	Cassandra. Operating Systems Review (ACM), 2010, 44, 35-40.	1.5	1,896
414	Comparison of Map-Reduce and SQL on Large-Scale Data Processing. , 2010, , .		8
415	Evaluation of MapReduce for Gridding LIDAR Data. , 2010, , .		27
416	Formalizing MapReduce with CSP. , 2010, , .		12
417	Implementation of NOSQL for robotics. , 2010, , .		4
418	SSS: An Implementation of Key-Value Store Based MapReduce Framework. , 2010, , .		14
419	An Efficient Data Mining Framework on Hadoop using Java Persistence API. , 2010, , .		25
420	Service-oriented storage resource architecture for cloud computing. , 2010, , .		3

#	ARTICLE	IF	CITATIONS
421	Twister. , 2010, , .		589
422	Efficient parallel set-similarity joins using MapReduce. , 2010, , .		330
423	MapReduce as a programming model for association rules algorithm on Hadoop. , 2010, , .		86
424	Quake: quality-aware detection and correction of sequencing errors. <i>Genome Biology</i> , 2010, 11, R116.	13.9	470
425	Cloud-scale RNA-sequencing differential expression analysis with Myrna. <i>Genome Biology</i> , 2010, 11, R83.	13.9	268
426	A Peer-to-Peer Framework for Supporting MapReduce Applications in Dynamic Cloud Environments. <i>Computer Communications and Networks</i> , 2010, , 113-125.	0.8	15
427	Multiple-Job Optimization in MapReduce for Heterogeneous Workloads. , 2010, , .		20
428	Tools and Technologies for Building Clouds. <i>Computer Communications and Networks</i> , 2010, , 3-20.	0.8	32
429	A MapReduce approach to $G_{i,i} * (d_i)$ spatial statistic. , 2010, , .		6
430	Spatial scene similarity assessment on Hadoop. , 2010, , .		2
431	Network traffic characteristics of data centers in the wild. , 2010, , .		1,561
432	CloudBATCH: A Batch Job Queuing System on Clouds with Hadoop and HBase. , 2010, , .		12
433	Accelerating SQL database operations on a GPU with CUDA. , 2010, , .		181
434	BetterLife 2.0: Large-Scale Social Intelligence Reasoning on Cloud. , 2010, , .		5
435	Commit Protocols for a Distributed Transactional Memory. , 2010, , .		3
436	Towards peer-to-peer-based cryptanalysis. , 2010, , .		5
437	Efficient Metadata Generation to Enable Interactive Data Discovery over Large-Scale Scientific Data Collections. , 2010, , .		17
438	Parallel Simultaneous Co-clustering and Learning with Map-Reduce. , 2010, , .		7

#	ARTICLE	IF	CITATIONS
439	Trusted Data Sharing over Untrusted Cloud Storage Providers. , 2010, , .		52
440	Term Filtering with Bounded Error. , 2010, , .		0
441	The Interactive Computing Model for Cyber Physical Society. , 2010, , .		0
442	Moving Text Analysis Tools to the Cloud. , 2010, , .		10
443	Compute and data management strategies for grid deployment of high throughput protein structure studies. , 2010, , .		1
444	A MapReduce Style Framework for Computations on Trees. , 2010, , .		4
445	Research and implementation on spatial data storage and operation based on Hadoop platform. , 2010, , .		10
446	Taming complex bioinformatics workflows with weaver, makeflow, and starch. , 2010, , .		12
447	Integrating Domain-Specific Programming into Software Design. , 2010, , .		1
448	A Library-Based Tool to Improve CPU Assignment for Multicore Processor-Based Pervasive Servers. , 2010, , .		0
449	Bridging to the Cloud: Solution Design Trends Helping "Legacy" Systems Leverage Cloud Computing. , 2010, , .		2
450	CSAL: A Cloud Storage Abstraction Layer to Enable Portable Cloud Applications. , 2010, , .		30
451	Time-Series Linear Search for Video Copies Based on Compact Signature Manipulation and Containment Relation Modeling. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20, 1603-1613.	5.6	14
452	Development of Data Mining Platform MUSASHI Towards Service Computing. , 2010, , .		4
453	Statistical Approaches to Predicting and Diagnosing Performance Problems in Component-Based Distributed Systems: An Experimental Evaluation. , 2010, , .		5
454	Using Global Behavior Modeling to Improve QoS in Cloud Data Storage Services. , 2010, , .		6
455	Analyzing Electroencephalograms Using Cloud Computing Techniques. , 2010, , .		19
456	Implementation of a Cloud Computing Environment for Hiding Huge Amounts of Data. , 2010, , .		4

#	ARTICLE	IF	CITATIONS
457	Supporting multi-row distributed transactions with global snapshot isolation using bare-bones HBase. , 2010, , .		29
458	Online detection of utility cloud anomalies using metric distributions. , 2010, , .		28
459	Voronoi-Based Geospatial Query Processing with MapReduce. , 2010, , .		88
460	An Simulation of MapReduce as a Higher-Order Web Service in Cloud. , 2010, , .		0
461	User Demand Prediction from Application Usage Pattern in Virtual Smartphone. , 2010, , .		5
462	Dynamic Fractional Resource Scheduling for cluster platforms. , 2010, , .		4
463	Challenges in Scheduling Aggregation in Cyberphysical Information Processing Systems. , 2010, , .		2
464	High-efficiency model calibration engine based on parallel computing. , 2010, , .		1
465	Status and key techniques in cloud computing. , 2010, , .		5
466	Search and browse log mining for web information retrieval. , 2010, , .		1
467	AzureBlast. , 2010, , .		92
468	Benchmarking IP blacklists for financial botnet detection. , 2010, , .		9
469	Fast algorithms for time series mining. , 2010, , .		2
470	The design of a CCA framework with distribution, parallelism, and recursive composition. , 2010, , .		3
471	Flexible-Matching Data-Comparison Tool with the Cloud Applications. , 2010, , .		0
472	FlexPRICE: Flexible Provisioning of Resources in a Cloud Environment. , 2010, , .		28
473	Application Development: Fly to the Clouds or Stay In-house?. , 2010, , .		14
474	CEO a cloud epistasis computing model in GWAS. , 2010, , .		5

#	ARTICLE	IF	CITATIONS
475	Rapid scalability of complex and dynamic web-based systems: Challenges and recent approaches to mitigation. , 2010, , .		2
476	Social Tagging: Semantische Informationsverarbeitung am Beispiel der Profildaten des XING-Netzwerks. Hmd, 2010, 47, 26-34.	0.3	0
477	Parallel Techniques for Large Data Analysis in a Futures Trading Evaluation Service System. , 2010, , .		0
478	The model-summary problem and a solution for trees. , 2010, , .		5
479	CloudView: Describe and Maintain Resource View in Cloud. , 2010, , .		5
480	Materialized view maintenance in columnar storage for massive data analysis. , 2010, , .		2
481	Global iceberg detection over distributed data streams. , 2010, , .		21
482	A Novel Computing-Enhanced Cloud Storage Model Supporting Combined Service Aware. , 2010, , .		0
483	Parallel Collection of Live Data Using Hadoop. , 2010, , .		7
484	A maximum clique algorithm based on MapReduce. , 2010, , .		2
485	Building Open and Scalable Digital Forensic Tools. , 2011, , .		4
486	Cloud Computing Model without Resource Management Center. , 2011, , .		2
487	Implementation of a Scalable Next Generation Sequencing Business Cloud Platform--An Experience Report. , 2011, , .		7
488	DryadOpt: Branch-and-Bound on Distributed Data-Parallel Execution Engines. , 2011, , .		15
489	Application-centric Cloud management. , 2011, , .		8
490	A parallelization of orchard temperature predicting programs. , 2011, , .		1
491	Deliver Bioinformatics Services in Public Cloud: Challenges and Research Framework. , 2011, , .		5
492	QoS-Based Distributed Service Selection in Large-Scale Web Services. , 2011, , .		2

#	ARTICLE	IF	CITATIONS
493	Web Farm-inspired Computational Cluster in the Cloud. , 2011, , .		2
494	Reasoning about Global Clones: Scalable Semantic Clone Detection. , 2011, , .		3
495	Web Multimedia Object Clustering via Information Fusion. , 2011, , .		5
496	A Comprehensive Long-Term Evaluation on BGP Performance. , 2011, , .		3
497	Semantic XML Views Based on Geographical Context. , 2011, , .		0
498	Low-Overhead Fault Tolerance for High-Throughput Data Processing Systems. , 2011, , .		25
499	Cloud-based malware detection for evolving data streams. ACM Transactions on Management Information Systems, 2011, 2, 1-27.	2.1	35
500	Towards Reproducible eScience in the Cloud. , 2011, , .		10
501	Parallel Outlier Detection Using KD-Tree Based on MapReduce. , 2011, , .		5
502	Assessing Suitability of Cloud Oriented Platforms for Application Development. , 2011, , .		4
503	Beyond 'Caveman Communities': Hubs and Spokes for Graph Compression and Mining. , 2011, , .		90
504	Active Replication at (Almost) No Cost. , 2011, , .		35
505	Efficient Star Join for Column-oriented Data Store in the MapReduce Environment. , 2011, , .		2
506	Ad Hoc Aggregation Query Processing Algorithms Based on Bit-Store in Data Intensive Cloud. , 2011, , .		1
507	A Software Architecture for the Analysis of Large Sets of Data Streams in Cloud Infrastructures. , 2011, , .		13
508	Optimizing Multiple Machine Learning Jobs on MapReduce. , 2011, , .		6
509	Parallelization of the functional flow algorithm for prediction of protein function using protein-protein interaction networks. , 2011, , .		0
510	Private query processing on inverted index. , 2011, , .		0

#	ARTICLE	IF	CITATIONS
511	S3: An Efficient Shared Scan Scheduler on MapReduce Framework. , 2011, , .		7
512	On mitigating TCP Incast in Data Center Networks. , 2011, , .		41
513	A Cloud-unaware Programming Model for Easy Development of Composite Services. , 2011, , .		7
514	Large Scale Page-Based Book Similarity Clustering. , 2011, , .		6
515	An epidemic approach to dependable key-value substrates. , 2011, , .		3
516	Using Shared Memory to Accelerate MapReduce on Graphics Processing Units. , 2011, , .		39
517	Parallel Matching Algorithms of Publish/Subscribe System. , 2011, , .		3
518	An Exploratory Study of the Evolution of Communicated Information about the Execution of Large Software Systems. , 2011, , .		28
519	A new in-domain interconnected architecture based on tree-based topology. , 2011, , .		0
520	StreamWeb: Real-Time Web Monitoring with Stream Computing. , 2011, , .		7
521	New parallel support vector regression for predicting building energy consumption. , 2011, , .		9
522	Towards Delivering Analytical Solutions in Cloud: Business Models and Technical Challenges. , 2011, , .		7
523	A Power-Aware Scheduling of MapReduce Applications in the Cloud. , 2011, , .		14
524	Storage System Optimization in Data-Intensive Environment. , 2011, , .		2
525	Gemma in April: A matrix-like parallel programming architecture on OpenCL. , 2011, , .		0
526	Green Scheduling: A Scheduling Policy for Improving the Energy Efficiency of Fair Scheduler. , 2011, , .		8
527	TIFA: Enabling Real-Time Querying and Storage of Massive Stream Data. , 2011, , .		8
528	The DIEGO Lab Graph Based Gene Normalization System. , 2011, , .		1

#	ARTICLE	IF	CITATIONS
529	COMET: A Recipe for Learning and Using Large Ensembles on Massive Data. , 2011, , .		24
530	Scalable and Low-Latency Data Processing with Stream MapReduce. , 2011, , .		38
531	Fast n-point Correlation Function Approximation with Recursive Convolution for Scalar Fields. , 2011, , .		4
532	Accountable MapReduce in cloud computing. , 2011, , .		32
533	Esc: Towards an Elastic Stream Computing Platform for the Cloud. , 2011, , .		40
534	STEAMEngine: Driving MapReduce provisioning in the cloud. , 2011, , .		19
535	Data Clustering on a Network of Mobile Smartphones. , 2011, , .		14
536	Privacy-preserving Collaborative Filtering for the Cloud. , 2011, , .		16
537	OCSS: Ontology Cloud Storage System. , 2011, , .		14
538	Parallel visualization on large clusters using MapReduce. , 2011, , .		25
539	Open911: Experiences with the Mobile Plus Cloud Paradigm. , 2011, , .		7
540	Reducing communication overhead under parallel list processing in multicore clusters. , 2011, , .		1
541	Coordinating Computation and I/O in Massively Parallel Sequence Search. IEEE Transactions on Parallel and Distributed Systems, 2011, 22, 529-543.	4.0	41
542	A Hierarchical Approach to Maximizing MapReduce Efficiency. , 2011, , .		2
543	Parallelizing a machine translation decoder for multicore computer. , 2011, , .		0
544	The Saaz Framework for Turbulent Flow Queries. , 2011, , .		3
545	A Probabilistic Approach to Address TCP Incast in Data Center Networks. , 2011, , .		9
546	Parallel Metagenomic Sequence Clustering Via Sketching and Maximal Quasi-clique Enumeration on Map-Reduce Clouds. , 2011, , .		15

#	ARTICLE	IF	CITATIONS
547	Is the cloud the answer to scalability of ecologies? Using GAE to enable horizontal scalability. , 2011, , .		6
548	Galileo: A Framework for Distributed Storage of High-Throughput Data Streams. , 2011, , .		19
549	A Solution to Integrate Data for Phylogenetic Research. , 2011, , .		1
550	Interestingness -- Directing Analyst Focus to Significant Data. , 2011, , .		0
551	mrNA: The MPI Randomized Numerical Aligner. , 2011, , .		1
552	Parallelized network-driven analysis of network systems on commodity hardware. , 2011, , .		0
553	A Web-Based Personalized Mobility Service for Smartphone Applications. Computer Journal, 2011, 54, 800-814.	1.5	12
554	Teaching the cloud - experiences in designing and teaching an undergraduate-level course in cloud computing at the Carnegie Mellon University in Qatar. , 2011, , .		7
555	Large-Scale Semantic Concept Detection on Manycore Platforms for Multimedia Mining. , 2011, , .		4
556	bCloudBLAST: An efficient mapreduce program for bioinformatics applications. , 2011, , .		6
557	On the Evaluation of Space-Time Functions. , 2011, , .		1
558	Maintaining and evolving Service Level Agreements: Motivation and case study. , 2011, , .		2
559	IDEA -- An API for Parallel Computing with Large Spatial Datasets. , 2011, , .		2
560	Architecture for Component Library Retrieval on the Cloud. , 2011, , .		3
561	Latency Minimizing Tasking for Information Processing Systems. , 2011, , .		0
562	Building mega data center from heterogeneous containers. , 2011, , .		25
563	The economy of parallel and distributed computing in the cloud. , 2011, , .		2
564	Storage I/O generation and replay for datacenter applications. , 2011, , .		1

#	ARTICLE	IF	CITATIONS
565	Media cloud: When media revolution meets rise of cloud computing. , 2011, , .		28
566	A Hybrid Shared-Nothing/Shared-Data Storage Scheme for Large-Scale Data Processing. , 2011, , .		1
567	Creating a Cloud-based Life Science Gateway. , 2011, , .		8
568	Dax Toolkit: A proposed framework for data analysis and visualization at Extreme Scale. , 2011, , .		37
569	A fast, lock-free approach for efficient parallel counting of occurrences of <i>k</i> -mers. Bioinformatics, 2011, 27, 764-770.	1.8	3,146
570	K-means Clustering in the Cloud -- A Mahout Test. , 2011, , .		56
571	QoS Preference-Aware Replica Selection Strategy Using MapReduce-Based PGA in Data Grids. , 2011, , .		3
572	Scalable data center multicast using multi-class Bloom Filter. , 2011, , .		53
573	A research on storage performance optimization of virtual machine. , 2011, , .		2
574	Reduce, You Say: What NoSQL Can Do for Data Aggregation and BI in Large Repositories. , 2011, , .		29
575	Large-scale document similarity computation based on cloud computing platform. , 2011, , .		3
576	A distributed workflow management system with case study of real-life scientific applications. , 2011, , .		0
577	Map/Reduce in CBIR application. , 2011, , .		3
578	Odessa. , 2011, , .		273
579	Scaling-Up Item-Based Collaborative Filtering Recommendation Algorithm Based on Hadoop. , 2011, , .		73
580	A platform for scalable one-pass analytics using MapReduce. , 2011, , .		107
581	Load Balancing for Data-Parallel Applications on Network-on-Chip Enabled Multi-processor Platform. , 2011, , .		4
582	A rate control scheme for distributed high performance video encoding in cloud. , 2011, , .		4

#	ARTICLE	IF	CITATIONS
583	Managing data retention policies at scale. , 2011, , .		1
584	MapReduce framework based distributed NMS architecture. , 2011, , .		0
585	Improving MapReduce Performance via Heterogeneity-Load-Aware Partition Function. , 2011, , .		1
586	Auto-scaling to minimize cost and meet application deadlines in cloud workflows. , 2011, , .		366
587	Automated, Parallel Optimization Algorithms for Stochastic Functions. , 2011, , .		3
588	Simplified parallel domain traversal. , 2011, , .		24
589	A Load-Driven Task Scheduler with Adaptive DSC for MapReduce. , 2011, , .		11
590	Design and Implementation of a Web 2.0 Service Platform for DPWS-Based Home-Appliances in the Cloud Environment. , 2011, , .		7
591	Visualization of Urban Air Pollution with Cloud Computing. , 2011, , .		10
592	EthSpeeder: A High-performance Scalable Fault-Tolerant Ethernet Network Architecture for Data Center. , 2011, , .		0
593	Llama. , 2011, , .		93
594	SimMapReduce: A Simulator for Modeling MapReduce Framework. , 2011, , .		25
595	Optimization of parallel Genetic Algorithms for nVidia GPUs. , 2011, , .		14
596	Counting triangles and the curse of the last reducer. , 2011, , .		269
597	Sedic. , 2011, , .		115
598	Semantic video event search for surveillance video. , 2011, , .		5
599	Cooperative spectrum sensing in TV White Spaces: When Cognitive Radio meets Cloud. , 2011, , .		10
600	Channel allocation in wireless data center networks. , 2011, , .		38

#	ARTICLE	IF	CITATIONS
601	Implications of cloud computing for computer science education. , 2011, , .		1
602	A Scalable Communication Runtime for Clouds. , 2011, , .		9
603	A parallel cross-modal search engine over large-scale multimedia collections with interactive relevance feedback. , 2011, , .		7
604	More convenient more overhead. , 2011, , .		25
605	A workflow-oriented cloud computing framework and programming model for data intensive application. , 2011, , .		4
606	A Survey of Large Scale Data Management Approaches in Cloud Environments. IEEE Communications Surveys and Tutorials, 2011, 13, 311-336.	24.8	323
607	The Case for Evaluating MapReduce Performance Using Workload Suites. , 2011, , .		267
608	Adaptive Disk I/O Scheduling for MapReduce in Virtualized Environment. , 2011, , .		35
609	MARIANE: MApReduce Implementation Adapted for HPC Environments. , 2011, , .		23
610	Performance under Failures of MapReduce Applications. , 2011, , .		24
611	BAR: An Efficient Data Locality Driven Task Scheduling Algorithm for Cloud Computing. , 2011, , .		70
612	Energy-efficient Multi-task Scheduling Based on MapReduce for Cloud Computing. , 2011, , .		9
613	Fused State Machines for Fault Tolerance in Distributed Systems. Lecture Notes in Computer Science, 2011, , 266-282.	1.0	1
614	SecDM: Securing Data Migration between Cloud Storage Systems. , 2011, , .		15
615	Towards a Self-Adaptive Data Management System for Cloud Environments. , 2011, , .		7
616	A Distributed Real-time Database Index Algorithm Based on B+ Tree and Consistent Hashing. Procedia Engineering, 2011, 24, 171-176.	1.2	5
617	A Novel Solution of Distributed Memory NoSQL Database for Cloud Computing. , 2011, , .		28
618	An Evaluation of Power-Proportional Data Placement for Hadoop Distributed File Systems. , 2011, , .		7

#	ARTICLE	IF	CITATIONS
619	A Novel Cost-Effective Dynamic Data Replication Strategy for Reliability in Cloud Data Centres. , 2011, , .		67
620	Monte Carlo Linear System Solver using MapReduce. , 2011, , .		9
621	Linguistic Structure Prediction. Synthesis Lectures on Human Language Technologies, 2011, 4, 1-274.	2.3	18
622	Job Scheduling Optimization for Multi-user MapReduce Clusters. , 2011, , .		14
623	Opinion Mining in MapReduce Framework. Communications in Computer and Information Science, 2011, , 50-55.	0.4	0
624	Zone-based data striping for cloud storage. IBM Journal of Research and Development, 2011, 55, 1:1-1:10.	3.2	1
625	Query optimization using column statistics in hive. , 2011, , .		19
626	Performance, Power, and Thermal Analysis of Low-Power Processors for Scale-Out Systems. , 2011, , .		20
628	Load balancing in MapReduce environments for data intensive applications. , 2011, , .		10
629	SAND Join — A skew handling join algorithm for Google's MapReduce framework. , 2011, , .		20
630	Automatic Task Re-organization in MapReduce. , 2011, , .		13
631	Towards improved load balancing for data intensive distributed computing. , 2011, , .		2
632	Elastic complex event processing. , 2011, , .		3
633	SciHadoop. , 2011, , .		91
634	Medical Information Integration Based Cloud Computing. , 2011, , .		14
635	Byzantine Fault-Tolerant MapReduce: Faults are Not Just Crashes. , 2011, , .		31
636	A Hadoop solution for ballistic image analysis and recognition. , 2011, , .		21
637	Can MPI Benefit Hadoop and MapReduce Applications?. , 2011, , .		29

#	ARTICLE	IF	CITATIONS
638	ICAS: An inter-VM IDS Log Cloud Analysis System. , 2011, , .		15
639	Towards Synthesizing Realistic Workload Traces for Studying the Hadoop Ecosystem. , 2011, , .		24
640	Study on re-implement mechanism based on node-ability in Hadoop. , 2011, , .		0
641	Data management and analysis at the Large Scale Data Facility. , 2011, , .		4
642	Towards Deploying Elastic Hadoop in the Cloud. , 2011, , .		11
643	On the Performance of Distributed Clustering Algorithms in File and Streaming Processing Systems. , 2011, , .		4
644	DH-TRIE frequent pattern mining on Hadoop using JPA. , 2011, , .		9
645	LI-MR: A Local Iteration Map/Reduce Model and Its Application to Mine Community Structure in Large-Scale Networks. , 2011, , .		0
646	A Parallelized Network Traffic Classification Based on Hidden Markov Model. , 2011, , .		9
647	PSON: A Parallelized SON Algorithm with MapReduce for Mining Frequent Sets. , 2011, , .		6
648	DOFUR: DDoS Forensics Using MapReduce. , 2011, , .		11
649	Finding the 'Needle': Locating Interesting Nodes Using the K-shortest Paths Algorithm in MapReduce. , 2011, , .		5
650	Mining large graphs: Algorithms, inference, and discoveries. , 2011, , .		47
651	K-Means Clustering with Bagging and MapReduce. , 2011, , .		12
652	Modeling the Performance of the Hadoop Online Prototype. , 2011, , .		6
653	LARX: Large-Scale Anti-Phishing by Retrospective Data-Exploring Based on a Cloud Computing Platform. , 2011, , .		20
654	Efficient data distribution strategy for join query processing in the cloud. , 2011, , .		2
655	Improving Speculative Execution Performance with Coworker for Cloud Computing. , 2011, , .		6

#	ARTICLE	IF	CITATIONS
656	Analytics over large-scale multidimensional data. , 2011, , .		225
657	Cloud Technologies for Bioinformatics Applications. IEEE Transactions on Parallel and Distributed Systems, 2011, 22, 998-1011.	4.0	50
658	An Adaptive Framework for the Execution of Data-Intensive MapReduce Applications in the Cloud. , 2011, , .		10
659	Large-scale matrix factorization with distributed stochastic gradient descent. , 2011, , .		398
660	A Cloud Environment for Data-intensive Storage Services. , 2011, , .		34
661	Query optimization for massively parallel data processing. , 2011, , .		103
662	Cyclic Workflow Execution Mechanism on Top of MapReduce Framework. , 2011, , .		0
663	Parallelizing BLAST and SOM Algorithms with MapReduce-MPI Library. , 2011, , .		25
664	Scalable Service Containers. , 2011, , .		22
665	Large Volume Testing for Executable Formal Specification Using Hadoop. , 2011, , .		2
666	An Elastic OLAP Cloud Platform. , 2011, , .		10
667	Supporting dynamic parameter sweep in adaptive and user-steered workflow. , 2011, , .		15
668	Detecting Text Similarity over Chinese Research Papers Using MapReduce. , 2011, , .		2
669	ARIA. , 2011, , .		304
670	Query suggestion for E-commerce sites. , 2011, , .		18
671	Cloud application logging for forensics. , 2011, , .		120
672	The Application of MapReduce in the Cloud Computing. , 2011, , .		17
673	Simplifying MapReduce Data Processing. , 2011, , .		8

#	ARTICLE	IF	CITATIONS
674	On modeling MapReduce with granular computing. , 2011, , .		2
675	Applications and Evaluation of In-memory MapReduce. , 2011, , .		0
676	Multi-GPU MapReduce on GPU Clusters. , 2011, , .		153
677	iHadoop: Asynchronous Iterations for MapReduce. , 2011, , .		41
678	DELMA: Dynamically ELastic MapReduce Framework for CPU-Intensive Applications. , 2011, , .		29
679	VIAF: Verification-Based Integrity Assurance Framework for MapReduce. , 2011, , .		47
681	A Cloud Framework for Parameter Sweeping Data Mining Applications. , 2011, , .		22
682	Benchmarking MapReduce Implementations for Application Usage Scenarios. , 2011, , .		18
683	A Framework for Managing MapReduce Applications in Dynamic Distributed Environments. , 2011, , .		7
684	CREST: Towards Fast Speculation of Straggler Tasks in MapReduce. , 2011, , .		18
685	Dynamic Data Redistribution for MapReduce Joins. , 2011, , .		8
686	Job Aware Scheduling Algorithm for MapReduce Framework. , 2011, , .		39
687	CrowdForge. , 2011, , .		261
688	NoSQL databases. , 2011, , .		43
689	Performance Research on MapReduce Programming Model. , 2011, , .		2
690	Flexible Process-Based Applications in Hybrid Clouds. , 2011, , .		17
691	Shrinking MTU to Mitigate TCP Incast Throughput Collapse in Data Center Networks. , 2011, , .		13
692	An In-Memory Framework for Extended MapReduce. , 2011, , .		4

#	ARTICLE	IF	CITATIONS
693	Design and Implementation of MapReduce Using the PGAS Programming Model with UPC. , 2011, , .		5
694	Automatic task slots assignment in Hadoop MapReduce. , 2011, , .		10
695	A load-aware scheduler for MapReduce framework in heterogeneous cloud environments. , 2011, , .		30
696	Optimizing intermediate data management in MapReduce computations. , 2011, , .		16
697	New improvement of the Hadoop relevant data locality scheduling algorithm based on LATE. , 2011, , .		8
698	Dacoop: Accelerating Data-Iterative Applications on Map/Reduce Cluster. , 2011, , .		1
699	Application of Hadoop MapReduce technique to Virtual Database system design. , 2011, , .		10
700	Evaluation of Network Topology Inference in Opaque Compute Clouds through End-to-End Measurements. , 2011, , .		17
701	Inferring Network Topologies in Infrastructure as a Service Cloud. , 2011, , .		1
702	Master-Slave Parallel Genetic Algorithm Based on MapReduce Using Cloud Computing. Applied Mechanics and Materials, 2011, 121-126, 4023-4027.	0.2	2
703	XML Processing in the Cloud: Large-Scale Digital Preservation in Small Institutions. , 2011, , .		2
704	A MapReduce based parallel SVM for large scale spam filtering. , 2011, , .		55
705	Multi-layer graph-based semi-supervised learning for large-scale image datasets using mapreduce. , 2011, , .		2
706	Media Cloud: Sharing contents in the large. , 2011, , .		15
707	RanKloud: Scalable Multimedia Data Processing in Server Clusters. IEEE MultiMedia, 2011, 18, 64-77.	1.5	26
708	A System Call Analysis Method with MapReduce for Malware Detection. , 2011, , .		11
709	Matchmaking: A New MapReduce Scheduling Technique. , 2011, , .		79
710	Mobile Cloud Web-Service for U-City. , 2011, , .		3

#	ARTICLE	IF	CITATIONS
711	Optimizing Multiway Joins in a Map-Reduce Environment. IEEE Transactions on Knowledge and Data Engineering, 2011, 23, 1282-1298.	4.0	136
712	Parallel Processing Framework on a P2P System Using Map and Reduce Primitives. , 2011, , .		8
713	SAKU: A distributed system for data analysis in large-scale dataset based on cloud computing. , 2011, , .		3
714	A cloud queuing service with strong consistency and high availability. IBM Journal of Research and Development, 2011, 55, 10:1-10:12.	3.2	3
715	Portable Parallel Programming on Cloud and HPC: Scientific Applications of Twister4Azure. , 2011, , .		11
716	Leveraging on High-Performance Computing and Cloud Technologies in Digital Libraries: A Case Study. , 2011, , .		5
717	A Scalable Cloud-based Queuing Service with Improved Consistency Levels. , 2011, , .		1
718	A Survey on Cloud Storage. Journal of Computers, 2011, 6, .	0.4	20
719	SOLARE: Self-Organizing Latency-Aware Resource Ensemble. , 2011, , .		1
720	The Strategy of Mining Association Rule Based on Cloud Computing. , 2011, , .		73
721	A proofreading tool using brute force techniques. IEEE Potentials, 2011, 30, 18-22.	0.2	1
722	Media cloud: an open cloud computing middleware for content management. IEEE Transactions on Consumer Electronics, 2011, 57, 970-978.	3.0	71
723	Parallel Spectral Clustering in Distributed Systems. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33, 568-586.	9.7	443
724	Computational Engineering in the Cloud. Journal of Organizational and End User Computing, 2011, 23, 31-50.	1.6	4
725	Parallel Belief Propagation in Factor Graphs. , 0, , 190-216.		1
726	Distributed Transfer Learning via Cooperative Matrix Factorization. , 0, , 331-351.		0
727	Parallel Large-Scale Feature Selection. , 0, , 352-370.		3
728	Autonomic Management of Application Workflows on Hybrid Computing Infrastructure. Scientific Programming, 2011, 19, 75-89.	0.5	41

#	ARTICLE	IF	CITATIONS
729	Scalability of Parallel Scientific Applications on the Cloud. Scientific Programming, 2011, 19, 91-105.	0.5	23
730	Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. Frontiers in Neuroinformatics, 2011, 5, 13.	1.3	1,383
732	CORAL Server and CORAL Server Proxy: Scalable Access to Relational Databases from CORAL Applications. Journal of Physics: Conference Series, 2011, 331, 042025.	0.3	4
733	Image formation challenges in the MOSAIC platform. Proceedings of SPIE, 2011, , .	0.8	0
734	Performance evaluation of MapReduce using full virtualisation on a departmental cloud. International Journal of Applied Mathematics and Computer Science, 2011, 21, 275-284.	1.5	16
735	Physicists Get INSPIREd: INSPIRE Project and Grid Applications. Journal of Physics: Conference Series, 2011, 331, 082005.	0.3	1
736	Reducing dimensionality of hyperspectral data with diffusion maps and clustering with k-means and Fuzzy ART. International Journal of Systems, Control and Communications, 2011, 3, 232.	0.2	5
737	Centralities in Large Networks: Algorithms and Observations. , 2011, , .		86
739	The Challenge of Service Level Scalability for the Cloud. International Journal of Cloud Applications and Computing, 2011, 1, 34-44.	1.1	14
740	Ultrafast and scalable cone-beam CT reconstruction using MapReduce in a cloud computing environment. Medical Physics, 2011, 38, 6603-6609.	1.6	32
741	Large-Scale Spectral Clustering with Map Reduce and MPI. , 0, , 240-261.		1
742	Myriad. , 2011, , .		2
743	Social content matching in MapReduce. Proceedings of the VLDB Endowment, 2011, 4, 460-469.	2.1	26
744	Improving performance of MapReduce framework on InterCloud by avoiding transmission of unnecessary data. Performance Evaluation Review, 2011, 39, 113-115.	0.4	0
745	A parallel incremental extreme SVM classifier. Neurocomputing, 2011, 74, 2532-2540.	3.5	38
746	Brown Dwarf: A fully-distributed, fault-tolerant data warehousing system. Journal of Parallel and Distributed Computing, 2011, 71, 1434-1446.	2.7	9
747	MapReduce in MPI for Large-scale graph algorithms. Parallel Computing, 2011, 37, 610-632.	1.3	162
748	Swift: A language for distributed parallel scripting. Parallel Computing, 2011, 37, 633-652.	1.3	319

#	ARTICLE	IF	CITATIONS
749	Towards Identifying the Challenges Associated with Emerging Large Scale Social Networks. <i>Procedia Computer Science</i> , 2011, 5, 458-465.	1.2	15
750	Research of Hadoop-based data flow management system. <i>Journal of China Universities of Posts and Telecommunications</i> , 2011, 18, 164-168.	0.8	13
751	Computing Performance: Game Over or Next Level?. <i>Computer</i> , 2011, 44, 31-38.	1.2	120
752	Entrepreneurial Innovation at Google. <i>Computer</i> , 2011, 44, 56-61.	1.2	8
753	Extreme Data-Intensive Scientific Computing. <i>Computing in Science and Engineering</i> , 2011, 13, 34-41.	1.2	30
754	Principles of Elastic Processes. <i>IEEE Internet Computing</i> , 2011, 15, 66-71.	3.2	136
755	Robust Execution of Service Workflows Using Redundancy and Advance Reservations. <i>IEEE Transactions on Services Computing</i> , 2011, 4, 125-139.	3.2	20
756	Load-Balanced Parallel Streamline Generation on Large Scale Vector Fields. <i>IEEE Transactions on Visualization and Computer Graphics</i> , 2011, 17, 1785-1794.	2.9	48
757	CloudAligner: A fast and full-featured MapReduce based tool for sequence mapping. <i>BMC Research Notes</i> , 2011, 4, 171.	0.6	95
758	Algorithms and mechanisms for procuring services with uncertain durations using redundancy. <i>Artificial Intelligence</i> , 2011, 175, 2021-2060.	3.9	13
759	An investigative analysis into security in the clouds and the impact of virtualization on the security architecture. , 2011, , .		4
760	Application architecture considerations for cloud platforms. , 2011, , .		5
761	Architectural Requirements for Cloud Computing Systems: An Enterprise Cloud Approach. <i>Journal of Grid Computing</i> , 2011, 9, 3-26.	2.5	266
762	A unified representation of web logs for mining applications. <i>Information Retrieval</i> , 2011, 14, 215-236.	1.6	5
763	Semantic analysis and retrieval in personal and social photo collections. <i>Multimedia Tools and Applications</i> , 2011, 51, 5-33.	2.6	19
764	Automatic performance debugging of SPMD-style parallel programs. <i>Journal of Parallel and Distributed Computing</i> , 2011, 71, 925-937.	2.7	10
765	Special Section: Grid and pervasive computing (selected papers from 2010 International Conference on Tj ETQq0 0,0 rgBT /Overlock 10	4.9	1
766	Component Approach to Computational Applications on Clouds. <i>Procedia Computer Science</i> , 2011, 4, 432-441.	1.2	28

#	ARTICLE	IF	CITATIONS
767	Personalized News Recommendation: A Review and an Experimental Investigation. Journal of Computer Science and Technology, 2011, 26, 754-766.	0.9	64
768	Review of data storage and management technologies for massive remote sensing data. Science China Technological Sciences, 2011, 54, 3220-3232.	2.0	41
769	Green challenges to system software in data centers. Frontiers of Computer Science, 2011, 5, 353-368.	0.6	10
770	Research and implementation of scalable parallel computing based on Map-Reduce. Journal of Shanghai University, 2011, 15, 426-429.	0.1	2
771	Agile parallel bioinformatics workflow management using Pwrake. BMC Research Notes, 2011, 4, 331.	0.6	20
772	CloVR: A virtual machine for automated and portable sequence analysis from the desktop using cloud computing. BMC Bioinformatics, 2011, 12, 356.	1.2	246
773	A lightweight, flow-based toolkit for parallel and distributed bioinformatics pipelines. BMC Bioinformatics, 2011, 12, 61.	1.2	7
774	SAMQA: error classification and validation of high-throughput sequenced read data. BMC Genomics, 2011, 12, 419.	1.2	13
775	Cloud computing paradigms for pleasingly parallel biomedical applications. Concurrency Computation Practice and Experience, 2011, 23, 2338-2354.	1.4	36
776	Many task computing for orthologous genes identification in protozoan genomes using Hydra. Concurrency Computation Practice and Experience, 2011, 23, 2326-2337.	1.4	8
777	A material history of bits. Journal of the Association for Information Science and Technology, 2011, 62, 1042-1057.	2.6	140
778	Variable-sized map and locality-aware reduce on public-resource grids. Future Generation Computer Systems, 2011, 27, 843-849.	4.9	11
779	Universal connection architecture for interactive applications to achieve distributed computing. Journal of Network and Computer Applications, 2011, 34, 232-251.	5.8	4
780	BlobSeer: Next-generation data management for large scale infrastructures. Journal of Parallel and Distributed Computing, 2011, 71, 169-184.	2.7	101
781	Load and storage balanced posting file partitioning for parallel information retrieval. Journal of Systems and Software, 2011, 84, 864-884.	3.3	6
782	A Practical Tree Contraction Algorithm for Parallel Skeletons on Trees of Unbounded Degree. Procedia Computer Science, 2011, 4, 7-16.	1.2	5
783	I/O-Performance Prediction Method for Mission-critical Grid-batch Processing. Procedia Computer Science, 2011, 4, 2166-2175.	1.2	0
784	Parallelizing large-scale data processing applications with data skew. , 2011, , .		4

#	ARTICLE	IF	CITATIONS
785	Synoptic. , 2011, , .		28
786	Wireless link scheduling for data center networks. , 2011, , .		16
787	Parallelism and data movement characterization of contemporary application classes. , 2011, , .		10
788	Detecting adversarial advertisements in the wild. , 2011, , .		62
789	Prlter. , 2011, , .		92
790	Runtime model validation with parallel object constraint language. , 2011, , .		8
791	Efficient k-nearest neighbor graph construction for generic similarity measures. , 2011, , .		349
792	Architecture-based fault tolerance support for grid applications. , 2011, , .		2
793	Energy efficient scheduling of MapReduce workloads on heterogeneous clusters. , 2011, , .		39
794	A cloud-enabled regional climate model evaluation system. , 2011, , .		7
795	Brasil. , 2011, , .		0
796	A paradigm comparison for collecting TV channel statistics from high-volume channel zap events. , 2011, , .		4
797	Authentication of range query results in mapreduce environments. , 2011, , .		7
798	Distributed workflow-driven analysis of large-scale biological data using biokepler. , 2011, , .		7
799	Understanding and improving the diagnostic workflow of MapReduce users. , 2011, , .		1
800	Applying idealized lower-bound runtime models to understand inefficiencies in data-intensive computing. , 2011, , .		0
801	Efficient processing of RDF graph pattern matching on MapReduce platforms. , 2011, , .		10
802	CnC-Hadoop. , 2011, , .		0

#	ARTICLE	IF	CITATIONS
803	Incremental graph pattern matching. , 2011, , .		81
804	Garbage collection auto-tuning for Java mapreduce on multi-cores. , 2011, , .		17
805	OpenTopography. , 2011, , .		60
806	SpSJoin. , 2011, , .		13
807	Adapting MapReduce for HPC environments. , 2011, , .		7
808	Cache size in a cost model for heterogeneous skeletons. , 2011, , .		2
809	A hierarchical framework for cross-domain MapReduce execution. , 2011, , .		43
810	Strict serializability is harmless. , 2011, , .		5
811	Managing data transfers in computer clusters with orchestra. , 2011, , .		301
812	Utilizing green energy prediction to schedule mixed batch and service jobs in data centers. , 2011, , .		83
813	Map-reduce extensions and recursive queries. , 2011, , .		52
814	CrowdForge. , 2011, , .		99
815	A fast approach for parallel deduplication on multicore processors. , 2011, , .		19
816	Parallel evaluation of conjunctive queries. , 2011, , .		63
817	Large scale visual-based event matching. , 2011, , .		17
818	PigSPARQL. , 2011, , .		71
819	GBASE. , 2011, , .		90
820	Towards Haskell in the cloud. , 2011, , .		64

#	ARTICLE	IF	CITATIONS
821	Making standard ML a practical database programming language. , 2011, , .		17
822	Modeling and synthesizing task placement constraints in Google compute clusters. , 2011, , .		141
823	No one (cluster) size fits all. , 2011, , .		178
824	The jabberwocky programming environment for structured social computing. , 2011, , .		97
825	Scalable hashing for shared memory supercomputers. , 2011, , .		5
826	A distributed look-up architecture for text mining applications using MapReduce. , 2011, , .		4
827	Legal document clustering with built-in topic segmentation. , 2011, , .		34
828	A linear-time approximation of the earth mover's distance. , 2011, , .		12
829	Scientific data services. , 2011, , .		7
830	Geometric overpass extraction from vector road data and DSMs. , 2011, , .		4
831	A distributed look-up architecture for text mining applications using mapreduce. , 2011, 2011, .		7
832	Notice of Violation of IEEE Publication Principles: Pipelined-MapReduce: An Improved MapReduce Parallel Programing Model. , 2011, , .		2
833	On scheduling in map-reduce and flow-shops. , 2011, , .		79
834	HORT. , 2011, , .		4
835	Provenance for MapReduce-based data-intensive workflows. , 2011, , .		33
836	Investigating MapReduce framework extensions for efficient processing of geographically scattered datasets. Performance Evaluation Review, 2011, 39, 116-118.	0.4	8
838	Energy Prediction for MapReduce Workloads. , 2011, , .		13
839	A parallel programming framework with Markovian messaging for LC-MS peptide detection. , 2011, , .		0

#	ARTICLE	IF	CITATIONS
840	Towards a Sustainable Services Ecosystem for Traffic Management. , 2011, , .		1
841	An Automatic Approach for Duplicate Bibliographic Metadata Identification Using Classification. , 2011, , .		0
842	AUDR: An Advanced Unstructured Data Repository. , 2011, , .		3
843	Provisioning and Evaluating Multi-domain Networked Clouds for Hadoop-based Applications. , 2011, , .		22
844	Improvement of the MapReduce Model Based on Message Middleware Oriented Data Intensive Computing. , 2011, , .		0
845	Balancing load in stream processing with the cloud. , 2011, , .		25
846	A privacy aware media gateway for connecting private multimedia clouds to limited devices. , 2011, , .		6
847	Storage performance optimization of virtual machine. , 2011, , .		0
848	Improving MapReduce energy efficiency for computation intensive workloads. , 2011, , .		55
849	Translating Chapel to Use FREERIDE: A Case Study in Using an HPC Language for Data-Intensive Computing. , 2011, , .		1
850	Control Strategy of Group Behavior for Internet of Things. , 2011, , .		0
851	Accelerating SVMs by integrating GPUs into MapReduce clusters. , 2011, , .		12
852	Kangaroo: Reliable Execution of Scientific Applications with DAG Programming Model. , 2011, , .		0
853	Mapping the Blogosphere--Towards a Universal and Scalable Blog-Crawler. , 2011, , .		13
854	Study on Management and Dispatching of Virtual Computing Resources Based on Pool Theory. , 2011, , .		1
855	DisNet: A Framework for Distributed Graph Computation. , 2011, , .		15
856	RAFTing MapReduce: Fast recovery on the RAFT. , 2011, , .		43
857	Reverse hierarchical variable clustering on database accelerated algorithm. , 2011, , .		1

#	ARTICLE	IF	CITATIONS
858	Embedded Processor Virtualization for Broadband Grid Computing. , 2011, , .		4
859	Cloud Elasticity for Implementing an Agricultural Weather Service. , 2011, , .		0
860	ZoneFS: Stripe remodeling in cloud data centers. , 2011, , .		0
861	Activity Based Sector Synchronisation: Efficient Transfer of Disk-State for WAN Live Migration. , 2011, , .		7
862	A File-Based Approach for Recommender Systems in High-Performance Computing Environments. , 2011, , .		1
863	Smart Redundancy for Distributed Computation. , 2011, , .		30
864	The design and implementation of distributed mobile points of interest(POI) based on Mahout. , 2011, , .		0
865	A Cost-Based Resource Scheduling Paradigm in Cloud Computing. , 2011, , .		23
866	Community-based cloud for emergency management. , 2011, , .		20
867	Research on hadoop-based enterprise file cloud storage system. , 2011, , .		14
868	A MapReduce framework for on-road mobile fossil fuel combustion CO ₂ emission estimation. , 2011, , .		5
869	A cloud-based system for spatial analysis service. , 2011, , .		2
870	MapReduce Implementation of Prestack Kirchhoff Time Migration (PKTM) on Seismic Data. , 2011, , .		9
871	Scalable parallel building blocks for custom data analysis. , 2011, , .		47
872	Tri-training and MapReduce-based massive data learning. International Journal of General Systems, 2011, 40, 355-380.	1.2	0
873	Research on algorithm of point cloud MapReduce registration. , 2011, , .		4
874	Parameterized centrality metric for network analysis. Physical Review E, 2011, 83, 066118.	0.8	36
875	Distributed cube materialization on holistic measures. , 2011, , .		33

#	ARTICLE	IF	CITATIONS
876	Flexible use of cloud resources through profit maximization and price discrimination. , 2011, , .		71
877	Productive parallel programming with Pomelo. , 2011, , .		0
878	Error Correction and Clustering Algorithms for Next Generation Sequencing. , 2011, , .		2
879	Optimal sampling algorithms for frequency estimation in distributed data. , 2011, , .		12
880	Data Decomposition in Biomedical e-Science Applications. , 2011, , .		2
881	A parallel implementation of Singular Value Decomposition based on Map-Reduce and PARPACK. , 2011, , .		1
882	Consensus-based estimation protocol for decentralized dynamic load balancing over partially connected networks. , 2011, , .		3
883	Workload Adaptive Checkpoint Scheduling of Virtual Machine Replication. , 2011, , .		3
884	Data Aggregation Scheme for a Mashup Based Internet Application. , 2011, , .		1
885	SCDP: Scalable, cost-effective, distributed and parallel computing model for academics. , 2011, , .		5
886	QDFS: A quality-aware distributed file storage service based on HDFS. , 2011, , .		5
887	Distributed tuning of machine learning algorithms using MapReduce Clusters. , 2011, , .		16
889	Static type checking of Hadoop MapReduce programs. , 2011, , .		8
890	Forty data communications research questions. Computer Communication Review, 2011, 41, 24-35.	1.5	29
891	Column-oriented storage techniques for MapReduce. Proceedings of the VLDB Endowment, 2011, 4, 419-429.	2.1	95
892	Data-intensive service system: Model and analysis. , 2011, , .		0
893	Map reduction framework for parallel data mining: Multicore to distributed network systems. , 2011, , .		0
896	A Novel Approach in Improving I/O Performance of Small Meteorological Files on HDFS. Applied Mechanics and Materials, 0, 117-119, 1759-1765.	0.2	2

#	ARTICLE	IF	CITATIONS
897	MapReduce implementation of a hybrid spectral library-database search method for large-scale peptide identification. <i>Bioinformatics</i> , 2011, 27, 3072-3073.	1.8	23
898	Making standard ML a practical database programming language. <i>ACM SIGPLAN Notices</i> , 2011, 46, 307-319.	0.2	5
899	Clustering Weighted Slope One for distributed parallel computing. , 2011, , .		1
900	MR-LEGOS: a refined MapReduce model. <i>International Journal of Cloud Computing</i> , 2011, 1, 58.	0.3	0
901	Towards a privacy-aware stream data management system for cloud applications. <i>International Journal of Web and Grid Services</i> , 2011, 7, 246.	0.4	8
902	Monte Carlo simulation of photon migration in a cloud computing environment with MapReduce. <i>Journal of Biomedical Optics</i> , 2011, 16, 125003.	1.4	31
903	Efficient similarity search on massive gene data based on cloud computing. , 2011, , .		0
904	Better never than late. , 2011, , .		387
905	Alternative implementation techniques of parallelized Differential Evolution for multi-core processors. , 2011, , .		0
906	Leveraging User Diversity to Harvest Knowledge on the Social Web. , 2011, , .		4
907	Adaptive Failure Detection via Heartbeat under Hadoop. , 2011, , .		13
908	Hyracks: A flexible and extensible foundation for data-intensive computing. , 2011, , .		182
909	Watershed: A High Performance Distributed Stream Processing System. , 2011, , .		4
910	Scalable and Distributed Processing of Scientific XML Data. , 2011, , .		2
911	Process-in-Network for Image Providing Services. , 2011, , .		0
912	Optimal Client-Server Assignment for Internet Distributed Systems. , 2011, , .		17
913	The core of constructing the future power systems computation platform is cloud computing. , 2011, , .		4
914	A large-scale rendering system based on hadoop. , 2011, , .		3

#	ARTICLE	IF	CITATIONS
915	Performance Prediction in a Decentralized Environment for Peer-to-Peer Computing. , 2011, , .		7
916	A Distributed Reconfigurable Active SSD Platform for Data Intensive Applications. , 2011, , .		13
917	Application and Research of Large-Scale Parallel Computing in Analysis of Mobile Users' Online Behavior. , 2011, , .		0
918	An index model of global subdivision in cloud computing environment. , 2011, , .		4
919	Inferring Profile Elements from Publicly Available Social Network Data. , 2011, , .		3
920	EAPAC: An Enhanced Application Placement Framework for Data Centers. , 2011, , .		3
921	A method of power grid mass data preprocessing. , 2011, , .		0
922	SystemML: Declarative machine learning on MapReduce. , 2011, , .		205
923	An Integrated Home Financial Investment Learning Environment Applying Cloud Computing in Social Network Analysis. , 2011, , .		3
924	Simple Parallel Genetic Algorithm Using Cloud Computing. Applied Mechanics and Materials, 2011, 121-126, 4151-4155.	0.2	4
925	OOLAM. , 2011, , .		15
926	Enhancement of Xen's scheduler for MapReduce workloads. , 2011, , .		46
928	A New Scheduling Algorithm in Hadoop MapReduce. Communications in Computer and Information Science, 2011, , 537-543.	0.4	4
929	Mining large distributed log data in near real time. , 2011, , .		6
930	Prediction-based auto-scaling of scientific workflows. , 2011, , .		9
931	Multiclass Support Vector Machine. , 2011, , 293-311.		2
932	Scheduling large jobs by abstraction refinement. , 2011, , .		5
933	Scalable clone detection using description logic. , 2011, , .		10

#	ARTICLE	IF	CITATIONS
934	A latency and fault-tolerance optimizer for online parallel query plans. , 2011, , .		32
935	Enhancing web search by mining search and browse logs. , 2011, , .		1
936	Parallel programming. , 2011, , .		0
937	Semi-supervised ranking on very large graphs with rich metadata. , 2011, , .		45
938	Garbage collection auto-tuning for Java mapreduce on multi-cores. ACM SIGPLAN Notices, 2011, 46, 109-118.	0.2	16
939	An architecture for Concordia. , 2011, , .		0
940	tigre: Transcription factor inference through gaussian process reconstruction of expression for bioconductor. Bioinformatics, 2011, 27, 1026-1027.	1.8	4
941	A data model and database for high-resolution pathology analytical image informatics. Journal of Pathology Informatics, 2011, 2, 32.	0.8	56
942	Frequent Pattern Mining Based on Pattern Space Division in Map/Reduce Cluster. Advanced Materials Research, 0, 588-589, 2038-2041.	0.3	0
943	Energy-Efficient Multi-Job Scheduling Model for Cloud Computing and Its Genetic Algorithm. Mathematical Problems in Engineering, 2012, 2012, 1-16.	0.6	27
944	Research on Algorithm for Mining Frequent Itemsets Based on Cloud Computing. Applied Mechanics and Materials, 2012, 239-240, 1303-1307.	0.2	0
945	Shark. , 2012, , .		89
946	A Cloud-Computing-Based Data Placement Strategy in High-Speed Railway. Discrete Dynamics in Nature and Society, 2012, 2012, 1-15.	0.5	7
947	Optimizing Hardware Design by Composing Utility-Directed Transformations. IEEE Transactions on Computers, 2012, 61, 1800-1812.	2.4	3
948	Challenges of Cloud Computing Evaluation. Applied Mechanics and Materials, 0, 198-199, 396-401.	0.2	1
949	Parallel Implementation of Single-Source Shortest Path Algorithm Based on Haloop. Applied Mechanics and Materials, 0, 220-223, 2428-2432.	0.2	3
950	Coordination Method on Cloud Computing Clusters. Advanced Materials Research, 0, 605-607, 2160-2163.	0.3	0
952	High-performance processing of text queries with tunable pruned term and term pair indexes. ACM Transactions on Information Systems, 2012, 30, 1-32.	3.8	10

#	ARTICLE	IF	CITATIONS
953	Large scale microblog mining using distributed MB-LDA. , 2012, , .		22
954	Large-scale graph mining and learning for information retrieval. , 2012, , .		0
955	Constructing a data accessing layer for in-memory data grid. , 2012, , .		1
956	HyMR. , 2012, , .		4
957	Data-driven fault tolerance for work stealing computations. , 2012, , .		17
958	Jockey. , 2012, , .		211
959	Using TPIE for processing massive data sets in C++. SIGSPATIAL Special, 2012, 4, 24-27.	2.5	4
960	Performance guarantees for distributed reachability queries. Proceedings of the VLDB Endowment, 2012, 5, 1304-1316.	2.1	38
961	Outsourcing Encryption of Attribute-Based Encryption with MapReduce. Lecture Notes in Computer Science, 2012, , 191-201.	1.0	79
962	Utilizing green energy prediction to schedule mixed batch and service jobs in data centers. Operating Systems Review (ACM), 2012, 45, 53-57.	1.5	54
963	Research on Clustering Analysis of Big Data. Advanced Engineering Forum, 2012, 6-7, 82-87.	0.3	2
964	A case for dual stack virtualization. , 2012, , .		13
965	Cloud Computing Uncovered: A Research Landscape. Advances in Computers, 2012, 86, 41-85.	1.2	73
966	VMR. , 2012, , .		3
967	A cloud decision support system for the risk assessment of coronary heart disease. , 2012, , .		1
968	Cost-aware replication for dataflows. , 2012, , .		2
969	Automated profiling and resource management of pig programs for meeting service level objectives. , 2012, , .		34
970	Failure scenario as a service (FSaaS) for Hadoop clusters. , 2012, , .		22

#	ARTICLE	IF	CITATIONS
971	Mapping a data-flow programming model onto heterogeneous platforms. , 2012, , .		33
972	An algorithm for fuzzy-based sentence-level document clustering for micro-level contradiction analysis. , 2012, , .		2
973	An optimization framework for map-reduce queries. , 2012, , .		25
974	Parallel skyline queries. , 2012, , .		42
975	An Energy and Data Locality Aware Bi-level Multiobjective Task Scheduling Model Based on MapReduce for Cloud Computing. , 2012, , .		13
976	Matrix chain multiplication via multi-way join algorithms in MapReduce. , 2012, , .		6
977	Delta-SimRank computing on MapReduce. , 2012, , .		16
978	SciQL. , 2012, , .		2
979	ComSoc. , 2012, , .		88
980	DEMON. , 2012, , .		160
981	A distributed index for efficient parallel top-k keyword search on massive graphs. , 2012, , .		4
982	Factoring past exposure in display advertising targeting. , 2012, , .		5
983	Managing and mining large graphs. , 2012, , .		33
984	Recurring job optimization in scope. , 2012, , .		18
985	SofEA. , 2012, , .		4
986	How to translate a book within an hour. , 2012, , .		11
987	HEDC. , 2012, , .		8
988	Meeting service level objectives of Pig programs. , 2012, , .		2

#	ARTICLE	IF	CITATIONS
989	HadoopXML. , 2012, , .		24
990	Challenges and approaches for distributed workflow-driven analysis of large-scale biological data. , 2012, , .		19
991	Large scale data analytics on clouds. , 2012, , .		6
992	Enhancing TCP throughput of highly available virtual machines via speculative communication. , 2012, , .		7
993	Experiences teaching MapReduce in the cloud. , 2012, , .		20
994	Social networking in developing regions. , 2012, , .		0
995	Practical TDMA for datacenter ethernet. , 2012, , .		33
996	Applying traffic merging to datacenter networks. , 2012, , .		10
997	Clydesdale. , 2012, , .		13
998	To nest or not to nest, when and how much. , 2012, , .		4
999	Feasibility of a privacy preserving collaborative filtering scheme on the Google App Engine. , 2012, , .		8
1000	Transitive closure and recursive Datalog implemented on clusters. , 2012, , .		25
1001	Graph pattern matching revised for social network analysis. , 2012, , .		102
1002	Investigation of data locality and fairness in MapReduce. , 2012, , .		20
1003	Accelerate large-scale iterative computation through asynchronous accumulative updates. , 2012, , .		28
1004	PonD. , 2012, , .		2
1005	Massively-parallel stream processing under QoS constraints with Nephele. , 2012, , .		14
1006	Space-round tradeoffs for MapReduce computations. , 2012, , .		39

#	ARTICLE	IF	CITATIONS
1007	An iterative MapReduce approach to frequent subgraph mining in biological datasets. , 2012, , .		37
1008	Bridging the tenant-provider gap in cloud services. , 2012, , .		80
1009	Untangling cluster management with Helix. , 2012, , .		8
1010	True elasticity in multi-tenant data-intensive compute clusters. , 2012, , .		43
1011	An effective rule miner for instance matching in a web of data. , 2012, , .		27
1012	Elastic and effective spatio-temporal query processing scheme on Hadoop. , 2012, , .		7
1013	Intent-aware temporal query modeling for keyword suggestion. , 2012, , .		1
1014	A community based vaccination strategy over mobile phone records. , 2012, , .		6
1015	The DEBS 2012 grand challenge. , 2012, , .		33
1016	Development of a scalable image formation pipeline for multiscale gigapixel photography. Optics Express, 2012, 20, 22048.	1.7	27
1017	The implementation of an intelligent cloud service system for disease risk assessment - chronic kidney disease as an example. , 2012, , .		0
1018	Design and Implementation of a Scalable and QoS-aware Stream Processing Framework: The Quasit Prototype. , 2012, , .		4
1019	A Fault-Tolerant Distributed Solution for the Parallel Gaussian Reduction of Huge Arrays. , 2012, , .		0
1020	A self-evolving anomaly detection framework for developing highly dependable utility clouds. , 2012, , .		20
1021	Result verification mechanism for MapReduce computation integrity in cloud computing. , 2012, , .		6
1022	Cloud-GSQCT:a parallel approach to screen gene sequences for phylogenetics analysis. , 2012, , .		2
1023	VM Economics for Java Cloud Computing: An Adaptive and Resource-Aware Java Runtime with Quality-of-Execution. , 2012, , .		7
1024	Dynamic heterogeneous data processing in IaaS cloud. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1025	Parallel Text Clustering Based on MapReduce. , 2012, , .		8
1026	A service integrity assurance framework for cloud computing based on MapReduce. , 2012, , .		6
1027	MySQL to NoSQL. , 2012, , .		57
1028	Towards Parallel Spatial Query Processing for Big Spatial Data. , 2012, , .		65
1029	A Technique for Self-Optimizing Scalable and Dependable Server Clusters under QoS Constraints. , 2012, , .		2
1030	Sketching and streaming algorithms for processing massive data. Xrds, 2012, 19, 14-19.	0.2	6
1031	Acceleration of Data-Intensive Workflow Applications by Using File Access History. , 2012, , .		2
1032	Earlybird: Real-Time Search at Twitter. , 2012, , .		119
1033	MapReduce Framework Optimization via Performance Modeling. , 2012, , .		7
1034	Efficient Updates for Web-Scale Indexes over the Cloud. , 2012, , .		1
1035	Parallelized Near-Duplicate Document Detection Algorithm for Large Scale Chinese Web Pages. , 2012, , .		0
1036	Hardware-Based Task Dependency Resolution for the StarSs Programming Model. , 2012, , .		10
1037	High-performance implementations of a clustering algorithm for finding network communities. , 2012, , .		0
1038	Scalability and Bandwidth Optimization for Data Center Networks. , 2012, , .		0
1039	FRIEDA: Flexible Robust Intelligent Elastic Data Management in Cloud Environments. , 2012, , .		10
1040	Balancing reducer skew in MapReduce workloads using progressive sampling. , 2012, , .		59
1041	Trust based interoperability security protocol for grid and Cloud computing. , 2012, , .		3
1042	Opportunistic Data-driven Execution of Parallel Programs for Efficient I/O Services. , 2012, , .		18

#	ARTICLE	IF	CITATIONS
1043	Mass data storage system for campus network based on cloud computing. , 2012, , .		1
1044	Dynamic Active Storage for High Performance I/O. , 2012, , .		7
1045	Curvilinearly T-Stiffened Panel Optimization Framework under Multiple Load Cases Using Parallel Processing. , 2012, , .		2
1046	MapReduce model-based optimization of range queries. , 2012, , .		3
1047	Cloud Based Test Coverage Service. , 2012, , .		0
1048	Distributed storage and prediction method for mooring monitoring system. , 2012, , .		0
1049	Towards Trusted Services: Result Verification Schemes for MapReduce. , 2012, , .		22
1050	Symbolic State Space Exploration of RT Systems in the Cloud. , 2012, , .		11
1051	Deconstructing the overhead in parallel applications. , 2012, , .		16
1052	Tagger: Bringing Real World Graffiti Social Interaction to Virtual San Francisco. , 2012, , .		1
1053	Parallelized incremental support vector machines based on MapReduce and Bagging technique. , 2012, , .		4
1054	DISco: A distributed information store for network challenges and their outcome. , 2012, , .		1
1055	Degree-guided map-reduce task assignment with data locality constraint. , 2012, , .		7
1056	CHAIO: Enabling HPC Applications on Data-Intensive File Systems. , 2012, , .		8
1057	MapReduce Skyline Query Processing with a New Angular Partitioning Approach. , 2012, , .		22
1058	View of MapReduce: Programming model, methods, and its applications. IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India), 2012, 29, 380.	2.1	3
1059	On superposition of heterogeneous edge processes in dynamic random graphs. , 2012, , .		3
1060	Automatic generation of software pipelines for heterogeneous parallel systems. , 2012, , .		9

#	ARTICLE	IF	CITATIONS
1061	Performance Evaluation of Yahoo! S4: A First Look. , 2012, , .		22
1062	MoCCA: A mobile cellular cloud architecture. , 2012, , .		4
1063	Research based on large-scale data query with mapreduce technology in cloud computing. , 2012, , .		0
1064	A Blog Personality Recommender System Based on Cloud Computing Infrastructure. , 2012, , .		6
1065	VirtualKnotter: Online Virtual Machine Shuffling for Congestion Resolving in Virtualized Datacenter. , 2012, , .		23
1066	Policy-based tuning for performance portability and library co-optimization. , 2012, , .		17
1067	Holistic View Angles of Cloud Computing Services Provisions. , 2012, , .		1
1068	A broadband embedded computing system for MapReduce utilizing Hadoop. , 2012, , .		8
1069	Adapting data-intensive workloads to generic allocation policies in cloud infrastructures. , 2012, , .		4
1070	Large-scale machine learning at twitter. , 2012, , .		143
1071	Opening the black boxes in data flow optimization. Proceedings of the VLDB Endowment, 2012, 5, 1256-1267.	2.1	73
1072	A MapReduce Framework for Mining Maximal Contiguous Frequent Patterns in Large DNA Sequence Datasets. IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India), 2012, 29, 162.	2.1	15
1073	GLADE. , 2012, , .		48
1074	Multi-core and many-core shared-memory parallel raycasting volume rendering optimization and tuning. International Journal of High Performance Computing Applications, 2012, 26, 399-412.	2.4	21
1075	Hadoop-BAM: directly manipulating next generation sequencing data in the cloud. Bioinformatics, 2012, 28, 876-877.	1.8	118
1076	Implementation of Space Optimized Bisecting K-Means (BKM) Based on Hadoop. , 2012, , .		1
1077	Applications Monitoring for Self-Optimization in GridGain. , 2012, , .		1
1078	A Scalable Architecture for Real-Time Monitoring of Large Information Systems. , 2012, , .		22

#	ARTICLE	IF	CITATIONS
1079	Dynamic energy-aware capacity provisioning for cloud computing environments. , 2012, , .		164
1080	Towards Optimal Bidding Strategy for Amazon EC2 Cloud Spot Instance. , 2012, , .		61
1081	Guess Again (and Again and Again): Measuring Password Strength by Simulating Password-Cracking Algorithms. , 2012, , .		215
1082	Performance analysis of work-conserving schedulers for minimizing total flow-time with phase precedence. , 2012, , .		3
1083	Distributed acoustic modeling with back-off n-grams. , 2012, , .		4
1084	MapReduce-Based Bayesian Automatic Text Classifier Used in Digital Library. Communications in Computer and Information Science, 2012, , 121-126.	0.4	0
1085	A MapReduce-Enabled Scientific Workflow Framework with Optimization Scheduling Algorithm. , 2012, , .		4
1086	A Comparative Study of Data Processing Approaches for Text Processing Workflows. , 2012, , .		0
1087	A Model-Based Deployment Framework of Integrated Public Cloud Service. , 2012, , .		1
1088	Towards an agent-based symbiotic architecture for autonomic management of virtualized data centers. , 2012, , .		4
1089	Life logging practice for human behavior modeling. , 2012, , .		5
1090	Grey-Box Approach for Performance Prediction in Map-Reduce Based Platforms. , 2012, , .		24
1091	A Structured Overlay Network for Aggregating Sensor Data. , 2012, , .		4
1092	Scheduling Data on Data-Driven Master/Worker Platform. , 2012, , .		3
1093	Minimizing the Energy Cost of Servers in Data Center Networks. , 2012, , .		2
1094	A Highly Efficient Consolidated Platform for Stream Computing and Hadoop. , 2012, , .		2
1095	A generic robot database and its application in fault analysis and performance evaluation. , 2012, , .		27
1096	A technical infrastructure to support personalized medicine. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1097	On Packing Very Large R-trees. , 2012, , .		3
1098	A Privacy Preserving Scalable Architecture for Collaborative Event Correlation. , 2012, , .		3
1099	Configuring a MapReduce Framework for Dynamic and Efficient Energy Adaptation. , 2012, , .		9
1100	Technique of Constructing Cloud Computing Platform Based on Ubuntu Enterprise Cloud. Advanced Materials Research, 2012, 482-484, 713-716.	0.3	0
1101	MapReduce Applications in the Cloud: A Cost Evaluation of Computation and Storage. Lecture Notes in Computer Science, 2012, , 37-48.	1.0	4
1102	Cost models for view materialization in the cloud. , 2012, , .		25
1103	A Bigtable/MapReduce-Based Cloud Infrastructure for Effectively and Efficiently Managing Large-Scale Sensor Networks. Lecture Notes in Computer Science, 2012, , 25-36.	1.0	9
1104	Haskell vs. f# vs. scala. , 2012, , .		3
1105	A framework for readapting and running bioinformatics applications in the cloud. , 2012, , .		0
1107	Design of Electric Vehicleâ€™s Operation Management Platform Based on Cloud Computing and Internet of Vehicle. Advanced Materials Research, 0, 608-609, 1647-1655.	0.3	0
1108	Future Trends in Similarity Searching. Lecture Notes in Computer Science, 2012, , 8-24.	1.0	7
1109	Construction of Tag-Based Dynamic Data Catalog (TaDDCat) Using Ontology. , 2012, , .		2
1110	Towards Haskell in the cloud. ACM SIGPLAN Notices, 2012, 46, 118-129.	0.2	30
1111	Towards mobile cloud applications. International Journal of Pervasive Computing and Communications, 2012, 8, 344-367.	1.1	27
1112	M3R. Proceedings of the VLDB Endowment, 2012, 5, 1736-1747.	2.1	100
1113	Consensus-based wireless tomography. , 2012, , .		2
1114	Distributed storage of network measurement data on HBase. , 2012, , .		9
1115	Turbine. , 2012, , .		19

#	ARTICLE	IF	CITATIONS
1116	CIRCUMFLEX. Operating Systems Review (ACM), 2012, 46, 26-32.	1.5	9
1117	More Parts Than Elements: How Databases Multiply. Environment and Planning D: Society and Space, 2012, 30, 335-350.	2.3	37
1118	Cloudpress 2.0. , 2012, , .		0
1119	Skew-Tolerant Key Distribution for Load Balancing in MapReduce. IEICE Transactions on Information and Systems, 2012, E95-D, 677-680.	0.4	3
1120	Delay tails in MapReduce scheduling. Performance Evaluation Review, 2012, 40, 5-16.	0.4	27
1121	COMPUTATIONAL BIOLOGY IN THE CLOUD: METHODS AND NEW INSIGHTS FROM COMPUTING AT SCALE. , 2012, , .		5
1122	E3: an Elastic Execution Engine for Scalable Data Processing. Journal of Information Processing, 2012, 20, 65-76.	0.3	8
1123	Performance Evaluation of A Testing Framework Using QuickCheck and Hadoop. Journal of Information Processing, 2012, 20, 340-346.	0.3	6
1124	Massively Parallel Databases and MapReduce Systems. Foundations and Trends in Databases, 2012, 5, 1-104.	4.1	28
1125	Simple and Efficient Model Filtering in Statistical Machine Translation. Prague Bulletin of Mathematical Linguistics, 2012, 98, 5-24.	0.7	2
1126	Time Series Analysis of Dynamic Networks. , 2012, , .		0
1127	MICE Online Data QualityJournal of Physics: Conference Series. Journal of Physics: Conference Series, 2012, 396, 012025.	0.3	0
1128	Enhancing TCP throughput of highly available virtual machines via speculative communication. ACM SIGPLAN Notices, 2012, 47, 87-96.	0.2	0
1129	A Flexible Open-Source Toolbox for Scalable Complex Graph Analysis. , 2012, , .		56
1130	A Survey on Mining Software Repositories. IEICE Transactions on Information and Systems, 2012, E95.D, 1384-1406.	0.4	7
1131	Applying the Model-View-Controller Paradigm to Adaptive Test. IEEE Design and Test of Computers, 2012, 29, 28-35.	1.4	8
1132	Recent Advances of Large-Scale Linear Classification. Proceedings of the IEEE, 2012, 100, 2584-2603.	16.4	206
1133	Paradigms for Developing Cloud Applications. , 2012, , 205-253.		1

#	ARTICLE	IF	CITATIONS
1134	Cloud Guided Stream Classification Using Class-Based Ensemble. , 2012, , .		14
1135	A survey of emerging approaches to spam filtering. ACM Computing Surveys, 2012, 44, 1-27.	16.1	71
1136	A prediction-based active queue management for TCP networks. , 2012, , .		0
1137	Cloud Computing and Big Data Analytics: What Is New from Databases Perspective?. Lecture Notes in Computer Science, 2012, , 42-61.	1.0	47
1138	P-Tracer: Path-Based Performance Profiling in Cloud Computing Systems. , 2012, , .		11
1139	A Modeling Method of Scientific Workflow Based on Cloud Environment. , 2012, , .		0
1140	P2P-MapReduce: Parallel data processing in dynamic Cloud environments. Journal of Computer and System Sciences, 2012, 78, 1382-1402.	0.9	81
1141	Cloud Workload Analysis with SWAT. , 2012, , .		4
1142	Implementing Hadoop Platform on Eucalyptus Cloud Infrastructure. , 2012, , .		4
1143	A Dispatching-Rule-Based Task Scheduling Policy for MapReduce with Multi-type Jobs in Heterogeneous Environments. , 2012, , .		8
1144	Supporting placement and data consistency strategies using hybrid clouds. , 2012, , .		1
1145	The Application of Cloud Computing in Smart Grid Status Monitoring. Communications in Computer and Information Science, 2012, , 460-465.	0.4	19
1146	GreenHadoop. , 2012, , .		217
1147	The F&amp;A Methodology and Its Experimental Validation on a Real-Life Parallel Processing Database System. , 2012, , .		2
1148	Energy-efficient Resource Management for QoS-guaranteed Computing Clusters. , 2012, , .		0
1149	Cloud intelligent track - Risk analysis and privacy data management in the cloud computing. , 2012, , .		2
1150	RDB-KV: A Cloud Database Framework for Managing Massive Heterogeneous Sensor Stream Data. , 2012, , .		2
1151	A Workflow-Aware Storage System: An Opportunity Study. , 2012, , .		22

#	ARTICLE	IF	CITATIONS
1152	Trust and Privacy. , 2012, , 117-138.		0
1153	Social group formation with mobile cloud services. Service Oriented Computing and Applications, 2012, 6, 351-362.	1.3	14
1154	A Provenance-based Adaptive Scheduling Heuristic for Parallel Scientific Workflows in Clouds. Journal of Grid Computing, 2012, 10, 521-552.	2.5	79
1155	Energy-efficient data centers. Computing (Vienna/New York), 2012, 94, 973-994.	3.2	78
1156	Lumping algorithms for computing Googleâ€™s PageRank and its derivative, with attention to unreferenced nodes. Information Retrieval, 2012, 15, 503-526.	1.6	23
1157	Enhancing genetic algorithms for dependent job scheduling in grid computing environments. Journal of Supercomputing, 2012, 62, 290-314.	2.4	25
1158	Analysis and improvement of map-reduce data distribution in read mapping applications. Journal of Supercomputing, 2012, 62, 1305-1317.	2.4	5
1159	PPModel: a modeling tool for source code maintenance and optimization of parallel programs. Journal of Supercomputing, 2012, 62, 1560-1582.	2.4	2
1160	Tools and techniques for non-invasive explicit parallelization. Journal of Supercomputing, 2012, 62, 1583-1608.	2.4	2
1161	Home-Appliance Control using Mobile Cloud Technology in Web2.0 Platform. Procedia Engineering, 2012, 38, 3587-3595.	1.2	4
1162	Data Management Challenges in Next Generation Sequencing. Datenbank-Spektrum, 2012, 12, 161-171.	1.2	25
1163	Scientific Workflows and Provenance: Introduction and Research Opportunities. Datenbank-Spektrum, 2012, 12, 193-203.	1.2	17
1164	Mrs: MapReduce for Scientific Computing in Python. , 2012, , .		1
1165	Dimensioning Scientific Computing Systems to Improve Performance of Map-Reduce based Applications. Procedia Computer Science, 2012, 9, 226-235.	1.2	6
1166	LBSNRank. , 2012, , .		15
1167	Cloud Parallel Processing of Tandem Mass Spectrometry Based Proteomics Data. Journal of Proteome Research, 2012, 11, 5101-5108.	1.8	38
1168	Kineograph. , 2012, , .		183
1169	PARMA. , 2012, , .		83

#	ARTICLE	IF	CITATIONS
1170	A "Minimal Disruption" Skeleton Experiment: Seamless Map & Reduce Embedding in OCaml. <i>Procedia Computer Science</i> , 2012, 9, 1837-1846.	1.2	20
1171	Parallel Tree Reduction on MapReduce. <i>Procedia Computer Science</i> , 2012, 9, 1827-1836.	1.2	9
1172	An Improvement of Choosing Map-join Candidates in Hive. <i>Procedia Computer Science</i> , 2012, 9, 2012-2015.	1.2	1
1173	A Framework for Distributed Data-Parallel Execution in the Kepler Scientific Workflow System. <i>Procedia Computer Science</i> , 2012, 9, 1620-1629.	1.2	18
1174	Mobile Sensor Data Classification for Human Activity Recognition using MapReduce on Cloud. <i>Procedia Computer Science</i> , 2012, 10, 585-592.	1.2	22
1175	Optimizing the Trade-offs Between Cost and Performance in Scientific Computing. <i>Procedia Computer Science</i> , 2012, 9, 498-507.	1.2	3
1176	Debugging Scientific Applications With Statistical Assertions. <i>Procedia Computer Science</i> , 2012, 9, 1940-1949.	1.2	2
1177	Adaptive Interpolation of Multidimensional Scaling. <i>Procedia Computer Science</i> , 2012, 9, 393-402.	1.2	9
1178	Tarazu. <i>Computer Architecture News</i> , 2012, 40, 61-74.	2.5	66
1179	Big Data Challenges: A Program Optimization Perspective. , 2012, , .		10
1180	Location Extraction from Social Networks with Commodity Software and Online Data. , 2012, , .		19
1181	A Simulation Study on Urban Water Threat Detection in Modern Cyberinfrastructures. , 2012, , .		0
1182	Towards building large-scale distributed systems for twitter sentiment analysis. , 2012, , .		45
1183	The structure of online diffusion networks. , 2012, , .		259
1184	Efficient multi-way theta-join processing using MapReduce. <i>Proceedings of the VLDB Endowment</i> , 2012, 5, 1184-1195.	2.1	67
1185	Managing Data Retention Policies at Scale. <i>IEEE Transactions on Network and Service Management</i> , 2012, 9, 393-406.	3.2	10
1186	Cloudqueue: An Internet-scale messaging infrastructure based on Hadoop. , 2012, , .		0
1187	Sedna: A Memory Based Key-Value Storage System for Realtime Processing in Cloud. , 2012, , .		14

#	ARTICLE	IF	CITATIONS
1188	Computaion integrity mechanism for MapReduce in cloud computing system. , 2012, , .		0
1189	Flexible Analysis Software for Emerging Architectures. , 2012, , .		5
1190	Executing Data-Intensive Workloads in a Cloud. , 2012, , .		7
1191	A Parallel Genetic Algorithm Based on Hadoop MapReduce for the Automatic Generation of JUnit Test Suites. , 2012, , .		53
1192	Parallel Exact Inference on Multicore Using MapReduce. , 2012, , .		2
1193	Image retrieval in the unstructured data management system AUDR. , 2012, , .		4
1194	A short survey on the state of the art in architectures and platforms for large scale data analysis and knowledge discovery from data. , 2012, , .		10
1195	Energy-efficient synchronization through program patterns. , 2012, , .		13
1196	Petri nets state space analysis in the cloud. , 2012, , .		8
1197	The Provenance of WINE. , 2012, , .		3
1198	Designing a Collaborative Filtering Recommender on the Single Chip Cloud Computer. , 2012, , .		2
1199	Practical Optimizations for Perceptron Algorithms in Large Malware Dataset. , 2012, , .		2
1200	WarpNet: A Novel Network for IaaS Data Center. , 2012, , .		0
1201	A New Two-Layer Topology for Data Center Network. , 2012, , .		0
1202	Large Scale KNN-Graph Approximation. , 2012, , .		3
1203	Parallel Concept Drift Detection with Online Map-Reduce. , 2012, , .		6
1204	Overview of Time Series Storage and Processing in a Cloud Environment. , 2012, , .		26
1205	Architecting security into the clouds: An enterprise security model. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1206	Cloud Services Gateway: A Tool for Exposing Private Services to the Public Cloud with Fine-grained Control. , 2012, , .		4
1207	Horton: Online Query Execution Engine for Large Distributed Graphs. , 2012, , .		26
1208	The Impact of Capacity Scheduler Configuration Settings on MapReduce Jobs. , 2012, , .		5
1209	GSQCT: A solution to screening gene sequences for phylogenetics analysis. , 2012, , .		2
1210	A cost aware adaptive multiple table join evaluation in MapReduce. , 2012, , .		0
1211	Hercules: Integrated control framework for datacenter traffic management. , 2012, , .		0
1212	User-driven cloud transportation system for smart driving. , 2012, , .		12
1213	On distributed file tree walk of parallel file systems. , 2012, , .		7
1214	4-4, 1-4: Architecture for Data Center Network Based on IP Address Hierarchy for Efficient Routing. , 2012, , .		4
1215	A distributed abnormal packet generation engine based on MapReduce. , 2012, , .		0
1216	Inverted Grid-Based kNN Query Processing with MapReduce. , 2012, , .		24
1217	Enhancing File Transfer Scheduling and Server Utilization in Data Distribution Infrastructures. , 2012, , .		3
1218	HDFSbench: Understanding the Efficiency and Bottlenecks of Cloud File Systems. , 2012, , .		0
1219	TMR: Towards a Trusted MapReduce Infrastructure. , 2012, , .		17
1220	Switching Optically-Connected Memories in a Large-Scale System. , 2012, , .		1
1221	Trend-Aware User Modeling with Location-Aware Trends on Twitter. , 2012, , .		3
1222	Accelerating Biomedical Data-Intensive Applications Using MapReduce. , 2012, , .		3
1223	Hadoop-Based Genome Comparisons. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1224	Executing a biological sequence comparison application on a federated cloud environment. , 2012, , .		6
1225	Modeling I/O Interference in Data Intensive Map-Reduce Applications. , 2012, , .		6
1226	Node capability aware resource provisioning in a heterogeneous cloud. , 2012, , .		4
1227	Generic Parallel Programming for Massive Remote Sensing Data Processing. , 2012, , .		13
1228	MARISSA: MAPReduce Implementation for Streaming Science Applications. , 2012, , .		13
1229	DOSAS: Mitigating the Resource Contention in Active Storage Systems. , 2012, , .		12
1230	Parameter estimation of Conditional Random Fields model based on cloud computing. , 2012, , .		0
1231	Email marketing and scalability using Hadoop. , 2012, , .		0
1232	Visual analysis of massive web session data. , 2012, , .		16
1233	Data-Intensive Services for Large-Scale Archive Access. , 2012, , .		4
1234	Design fast matrix algorithms on high-performance cloud platforms. , 2012, , .		1
1235	A Coding-based Approach to Mitigate TCP Incast in Data Center Networks. , 2012, , .		6
1236	Making sense of healthcare benefits. , 2012, , .		0
1237	Distributed high-dimensional index creation using Hadoop, HDFS and C++. , 2012, , .		4
1238	On the Performance of Virtualized Infrastructures for Processing Realtime Streaming Data. , 2012, , .		0
1239	Cost Optimization of Cloud-Based Data Integration System. , 2012, , .		6
1240	From High-School Algebra to Computing through Functional Programming. , 2012, , .		0
1241	Virtual Machine and Bytecode for Optimization on Heterogeneous Systems. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1242	Improving resource utilization in a heterogeneous cloud environment. , 2012, , .		1
1243	Research and application of MapReduce-based MST text clustering algorithm. , 2012, , .		2
1244	An active storage framework for object storage devices. , 2012, , .		10
1245	An Analysis of SMP Memory Allocators: MapReduce on Large Shared-Memory Systems. , 2012, , .		1
1246	An Analytic Goodput Model for TCP Incast. , 2012, , .		2
1247	Stratus: A Distributed Computing Framework for Scientific Simulations on the Cloud. , 2012, , .		8
1248	KNOWAC: I/O Prefetch via Accumulated Knowledge. , 2012, , .		21
1249	ORCA: An offloading framework for I/O-intensive applications on clusters. , 2012, , .		0
1250	AFD: Adaptive failure detection system for cloud computing infrastructures. , 2012, , .		12
1251	MapReduce on opportunistic resources leveraging resource availability. , 2012, , .		9
1252	H2RDF. , 2012, , .		78
1253	Parallel code clone detection using MapReduce. , 2012, , .		8
1254	Maestro: A cloud computing framework with automated locking. , 2012, , .		2
1255	Fault Tolerant Implementation of Peer-to-peer Distributed Iterative Algorithms. , 2012, , .		2
1256	GHTorrent: Github's data from a firehose. , 2012, , .		139
1257	AROM: Processing big data with Data Flow Graphs and functional programming. , 2012, , .		5
1258	A distributed transcoding system for mobile video delivery. , 2012, , .		2
1259	PlugCloud - Scaling by Plugging a Personal Cloud Infrastructure. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1260	Data-Intensive Workload Consolidation for the Hadoop Distributed File System. , 2012, , .		14
1261	Implement a reliable and secure cloud distributed file system. , 2012, , .		6
1262	Improved Solution to TCP Incast Problem in Data Center Networks. , 2012, , .		4
1263	ScalaPipe: A Streaming Application Generator. , 2012, , .		5
1264	Autotuning Skeleton-Driven Optimizations for Transactional Worklist Applications. IEEE Transactions on Parallel and Distributed Systems, 2012, 23, 2205-2218.	4.0	2
1265	Measuring Distributed Applications through MapReduce and Traffic Analysis. , 2012, , .		5
1266	Scalable Link Prediction in Social Networks Based on Local Graph Characteristics. , 2012, , .		6
1267	A H.264 SVC distributed content protection system with flexible key stream generation. , 2012, , .		4
1268	Low-Level Scheduling Implications for Data-Intensive Cyclic Workloads on Modern Microarchitectures. , 2012, , .		0
1269	Coding for fast content download. , 2012, , .		73
1270	Towards an Adaptable Middleware for Parallel Computing in Heterogeneous Environments. , 2012, , .		2
1271	Design of MAP-REDUCE and K-MEANS based network clustering protocol for sensor networks. , 2012, , .		2
1272	Optimizing a Collaborative Filtering Recommender for Many-Core Processors. , 2012, , .		1
1273	Adapting MPI to MapReduce PaaS Clouds: An Experiment in Cross-Paradigm Execution. , 2012, , .		3
1274	Combining in-situ and in-transit processing to enable extreme-scale scientific analysis. , 2012, , .		104
1275	Single Operation Multiple Data - Data Parallelism at Subroutine Level. , 2012, , .		0
1276	A MapReduce based ISODATA algorithm. , 2012, , .		2
1277	Implementing short-term traffic flow forecasting based on multipoint WPRA with MapReduce. , 2012, , .		1

#	ARTICLE	IF	CITATIONS
1278	Research on visualization platform for the production process in generic manufacturing. , 2012, , .		1
1279	Heterogeneous Task Scheduling for Accelerated OpenMP. , 2012, , .		45
1280	Assessing MapReduce for Internet Computing: A Comparison of Hadoop and BitDew-MapReduce. , 2012, , .		14
1281	pGraph: Efficient Parallel Construction of Large-Scale Protein Sequence Homology Graphs. IEEE Transactions on Parallel and Distributed Systems, 2012, 23, 1923-1933.	4.0	28
1282	Several public commercial clouds and open source cloud computing software. , 2012, , .		6
1283	Accelerating Expectation-Maximization Algorithms with Frequent Updates. , 2012, , .		24
1284	Hybrid cloud computing platform: The next generation IT backbone for smart grid. , 2012, , .		8
1285	Managing service performance in NoSQL distributed storage systems. , 2012, , .		7
1286	Delay tails in MapReduce scheduling. , 2012, , .		49
1287	Cloud Based Short Read Mapping Service. , 2012, , .		4
1288	Heterogeneity and dynamicity of clouds at scale. , 2012, , .		723
1289	Energy efficiency for large-scale MapReduce workloads with significant interactive analysis. , 2012, , .		129
1290	Energy Efficient Hadoop Using Mirrored Data Block Replication Policy. , 2012, , .		1
1291	An Efficient Geometry Data Allocation Algorithm in Cloud Computing Environments. , 2012, , .		1
1292	Implementation and evaluation of scalable data structure over HBase. , 2012, , .		6
1293	CCSA: A Cloud Computing Service Architecture for Sensor Networks. , 2012, , .		5
1294	Workload Characteristic Oriented Scheduler for MapReduce. , 2012, , .		20
1295	On Managing Very Large Sensor-Network Data Using Bigtable. , 2012, , .		37

#	ARTICLE	IF	CITATIONS
1296	Optimizing the storage of massive electronic pedigrees in HDFS. , 2012, , .		11
1297	Multicore-Enabled Smart Storage for Clusters. , 2012, , .		1
1298	Scalable similarity-based neighborhood methods with MapReduce. , 2012, , .		31
1299	The Algorithm for Constructing Extremal Graphs Based on MapReduce. , 2012, , .		1
1300	A preliminary study of incorporating GPUs in the Hadoop framework. , 2012, , .		7
1301	ADAPT: Availability-Aware MapReduce Data Placement for Non-dedicated Distributed Computing. , 2012, , .		50
1302	A novel parallel clustering algorithm PXM based on FP-Tree. , 2012, , .		0
1303	Building an Information System for Reality Mining Based on Communication Traces. , 2012, , .		5
1304	Energy-Aware Replica Selection for Data-Intensive Services in Cloud. , 2012, , .		10
1305	A Content-Based Image Retrieval System Based on Hadoop and Lucene. , 2012, , .		15
1306	An Efficient Map-Reduce Algorithm for the Incremental Computation of All-Pairs Shortest Paths in Social Networks. , 2012, , .		2
1307	Risk Aware Provisioning and Resource Aggregation Based Consolidation of Virtual Machines. , 2012, , .		25
1308	MTSD: A Task Scheduling Algorithm for MapReduce Base on Deadline Constraints. , 2012, , .		23
1309	Elastic computing: A portable optimization framework for hybrid computers. Parallel Computing, 2012, 38, 438-464.	1.3	11
1310	Memory-restricted latent semantic analysis to accumulate term-document co-occurrence events. Pattern Recognition Letters, 2012, 33, 1623-1631.	2.6	1
1311	A Novel Mass Data Processing Framework Based on Hadoop for Electrical Power Monitoring System. , 2012, , .		3
1312	A Novel Reputation Management System for volunteer clouds. , 2012, , .		10
1313	Deploying and researching Hadoop in virtual machines. , 2012, , .		27

#	ARTICLE	IF	CITATIONS
1314	Learning by expansion: Exploiting social media for image classification with few training examples. <i>Neurocomputing</i> , 2012, 95, 117-125.	3.5	8
1315	Cooperative private searching in clouds. <i>Journal of Parallel and Distributed Computing</i> , 2012, 72, 1019-1031.	2.7	28
1316	A scalable and accurate method for classifying protein-ligand binding geometries using a MapReduce approach. <i>Computers in Biology and Medicine</i> , 2012, 42, 758-771.	3.9	27
1317	An Integrated Approach for Specifying and Enforcing SLAs for Cloud Services. , 2012, , .		18
1318	Towards Quality Aware Collaborative Video Analytic Cloud. , 2012, , .		9
1319	Traffic anomaly detection based on the IP size distribution. , 2012, , .		18
1320	EasyPDP: An Efficient Parallel Dynamic Programming Runtime System for Computational Biology. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2012, 23, 862-872.	4.0	21
1321	FCFA: A semantic-based federated cloud framework architecture. , 2012, , .		12
1322	Mitigating High Latency Outliers for Cloud-Based Telecommunication Services. <i>Bell Labs Technical Journal</i> , 2012, 17, 121-142.	0.7	0
1323	Cloudgene: A graphical execution platform for MapReduce programs on private and public clouds. <i>BMC Bioinformatics</i> , 2012, 13, 200.	1.2	43
1324	Hydra: a scalable proteomic search engine which utilizes the Hadoop distributed computing framework. <i>BMC Bioinformatics</i> , 2012, 13, 324.	1.2	48
1325	A user-friendly tool to transform large scale administrative data into wide table format using a mapreduce program with a pig latin based script. <i>BMC Medical Informatics and Decision Making</i> , 2012, 12, 151.	1.5	8
1326	Privacy preserving collaborative filtering for SaaS enabling PaaS clouds. <i>Journal of Cloud Computing: Advances, Systems and Applications</i> , 2012, 1, 8.	2.1	13
1327	Scalable and Parallel Boosting with MapReduce. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2012, 24, 1904-1916.	4.0	79
1328	Towards Description Logics Based Cloud Service for UML Verification. <i>Lecture Notes in Electrical Engineering</i> , 2012, , 1184-1191.	0.3	1
1329	Big Data Processing in Cloud Computing Environments. , 2012, , .		193
1330	Research on cloud computing security problem and strategy. , 2012, , .		127
1331	Research on cloud computing data security model based on multi-dimension. , 2012, , .		14

#	ARTICLE	IF	CITATIONS
1332	A Cloud Service Design for Quality Evaluation of Design Patterns. , 2012, , .		2
1333	Parallel ABox Reasoning of $\{EL\}$ Ontologies. Lecture Notes in Computer Science, 2012, , 17-32.	1.0	12
1334	Analysis and Detection of Web Spam by Means of Web Content. Lecture Notes in Computer Science, 2012, , 43-57.	1.0	11
1335	A Parallel Framework for Processing Massive Spatial Data with a Split-and-Merge Paradigm. Transactions in GIS, 2012, 16, 829-843.	1.0	9
1336	A Multi-Granularity Parallel Model for Unified Remote Sensing Image Processing <sc>W</sc>eb<sc>S</sc>ervices. Transactions in GIS, 2012, 16, 845-866.	1.0	4
1337	From Semantic Grid to knowledge service cloud. Journal of Zhejiang University: Science C, 2012, 13, 253-256.	0.7	4
1338	Small World Asynchronous Parallel Model for Genome Assembly. Lecture Notes in Computer Science, 2012, , 145-155.	1.0	10
1339	SOAP Processing Performance and Enhancement. IEEE Transactions on Services Computing, 2012, 5, 387-403.	3.2	35
1340	Biological Databases for Behavioral Neurobiology. International Review of Neurobiology, 2012, 103, 19-38.	0.9	2
1341	Non-structure Data Storage Technology: A Discussion. , 2012, , .		7
1342	CDA: A Cloud Dependability Analysis Framework for Characterizing System Dependability in Cloud Computing Infrastructures. , 2012, , .		20
1343	MapReduce across Distributed Clusters for Data-intensive Applications. , 2012, , .		25
1344	A tile-based scalable raster data management system based on HDFS. , 2012, , .		2
1345	A distributed geospatial data storage and processing framework for large-scale WebGIS. , 2012, , .		11
1346	LBVP: A load balance algorithm based on Virtual Partition in Hadoop cluster. , 2012, , .		7
1347	Load Balancing in MapReduce Based on Scalable Cardinality Estimates. , 2012, , .		88
1348	A Parameter Dynamic-Tuning Scheduling Algorithm Based on History in Heterogeneous Environments. , 2012, , .		2
1349	Tracking Users on the Internet with Behavioral Patterns: Evaluation of Its Practical Feasibility. International Federation for Information Processing, 2012, , 235-248.	0.4	19

#	ARTICLE	IF	CITATIONS
1350	Use of permutation prefixes for efficient and scalable approximate similarity search. Information Processing and Management, 2012, 48, 889-902.	5.4	46
1351	An efficient global/multi-local stress analysis of complicated engineering composite structures using multi-goal MapReduce. Computational Materials Science, 2012, 65, 149-156.	1.4	0
1352	Improving resource efficiency in data centers using reputation-based resource selection. Sustainable Computing: Informatics and Systems, 2012, 2, 138-150.	1.6	5
1353	9.6 New Technologies for Molecular Dynamics Simulations. , 2012, , 86-104.		8
1354	Design Principles for Effective Knowledge Discovery from Big Data. , 2012, , .		94
1356	vHadoop: A Scalable Hadoop Virtual Cluster Platform for MapReduce-Based Parallel Machine Learning with Performance Consideration. , 2012, , .		17
1357	Scheduling MapReduce Jobs in HPC Clusters. Lecture Notes in Computer Science, 2012, , 179-190.	1.0	5
1358	Genotyping in the Cloud with Crossbow. Current Protocols in Bioinformatics, 2012, 39, Unit15.3.	25.8	30
1359	Banking on Decoupling: Budget-Driven Sustainability for HPC Applications on EC2 Spot Instances. , 2012, , .		2
1360	Data Management in the Cloud: Challenges and Opportunities. Synthesis Lectures on Data Management, 2012, 4, 1-138.	0.6	31
1361	Remote sensing algorithm platform in Windows Azure. , 2012, , .		1
1362	Key Research Issues for Privacy Protection and Preservation in Cloud Computing. , 2012, , .		5
1363	Using Performance Measurements to Improve MapReduce Algorithms. Procedia Computer Science, 2012, 9, 1920-1929.	1.2	8
1364	Resource Management for Dynamic MapReduce Clusters in Multicloud Systems. , 2012, , .		12
1365	A MapReduce-based Algorithm for Motif Search. , 2012, , .		3
1366	Distributed Computing with RESTful Web Services. , 2012, , .		5
1367	Data Warehouses: Next Challenges. Lecture Notes in Business Information Processing, 2012, , 1-26.	0.8	8
1368	A novel development and analysis solution to PaaS log by using CouchDB. , 2012, , .		2

#	ARTICLE	IF	CITATIONS
1369	Inside "Big Data management". , 2012, , .		91
1370	Performance evaluation of a massively parallel ESB-oriented architecture. , 2012, , .		4
1371	Shifting the bioinformatics computing paradigm: A case study in parallelizing genome annotation using MAKER and Work Queue. , 2012, , .		3
1372	MAPCloud: Mobile Applications on an Elastic and Scalable 2-Tier Cloud Architecture. , 2012, , .		110
1373	Data-intensive computing with map-reduce and hadoop. , 2012, , .		12
1374	Component-based approach for programming and running scientific applications on grids and clouds. International Journal of High Performance Computing Applications, 2012, 26, 275-295.	2.4	12
1375	An integrated runtime monitoring method for Internet-Based Virtual Computing Environment. , 2012, , .		0
1376	Notice of Violation of IEEE Publication Principles: On medical informatics for pervasive and ubiquitous computing in eHealth. , 2012, , .		2
1377	Next challenges for adaptive learning systems. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2012, 14, 48-55.	3.2	67
1378	Construction of a Taiwan Biodiversity Genetic Database in Cloud Environments. , 2012, , .		0
1379	Efficient and Accurate Anomaly Identification Using Reduced Metric Space in Utility Clouds. , 2012, , .		9
1380	Parallel data processing with MapReduce. SIGMOD Record, 2012, 40, 11-20.	0.7	448
1381	A Storage Infrastructure for Heterogeneous and Multimedia Data in the Internet of Things. , 2012, , .		18
1382	VC-Migration: Live Migration of Virtual Clusters in the Cloud. , 2012, , .		30
1383	An Implementation of GPU Accelerated MapReduce: Using Hadoop with OpenCL for Data- and Compute-Intensive Jobs. , 2012, , .		11
1384	Analysis and Optimization of Data Import with Hadoop. , 2012, , .		7
1385	CloudRank-D: benchmarking and ranking cloud computing systems for data processing applications. Frontiers of Computer Science, 2012, 6, 347-362.	1.6	70
1386	A parallel graph partitioning algorithm to speed up the large-scale distributed graph mining. , 2012, , .		5

#	ARTICLE	IF	CITATIONS
1387	De Novo Assembly of High-Throughput Sequencing Data with Cloud Computing and New Operations on String Graphs. , 2012, , .		7
1388	ISOBAR Preconditioner for Effective and High-throughput Lossless Data Compression. , 2012, , .		47
1389	A Big Data Model Supporting Information Recommendation in Social Networks. , 2012, , .		23
1390	Efficient alignment of next generation sequencing data using MapReduce on the cloud. , 2012, , .		0
1391	Energy-Efficient Datacenters. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2012, 31, 1465-1484.	1.9	100
1392	Distributed PrefixSpan algorithm based on MapReduce. , 2012, , .		15
1393	Large-scale multimedia data mining using MapReduce framework. , 2012, , .		16
1394	Improving large graph processing on partitioned graphs in the cloud. , 2012, , .		67
1395	Parallel K-PSO based on MapReduce. , 2012, , .		10
1396	A Hybrid Scheduling Approach for Scalable Heterogeneous Hadoop Systems. , 2012, , .		42
1397	Dynamic Processing Slots Scheduling for I/O Intensive Jobs of Hadoop MapReduce. , 2012, , .		13
1398	HOG: Distributed Hadoop MapReduce on the Grid. , 2012, , .		23
1399	Speculative Genetic Scheduling Method for Hadoop Environments. , 2012, , .		3
1401	SAHAD: Subgraph Analysis in Massive Networks Using Hadoop. , 2012, , .		44
1402	A system of job log analyzing for Hadoop. , 2012, , .		1
1404	Analysis of resource usage profile for MapReduce applications using Hadoop on cloud. , 2012, , .		3
1405	Statistics-based Workload Modeling for MapReduce. , 2012, , .		15
1406	A Distributed Network Measurement System Based on Hadoop. , 2012, , .		3

#	ARTICLE	IF	CITATIONS
1407	Molecular dynamics simulation: Implementation and optimization based on Hadoop. , 2012, , .		3
1408	Extending Map-Reduce for Efficient Predicate-Based Sampling. , 2012, , .		31
1409	RPig: A scalable framework for machine learning and advanced statistical functionalities. , 2012, , .		7
1410	Reengineering High-throughput Molecular Datasets for Scalable Clustering Using MapReduce. , 2012, , .		5
1411	Providing Dependability and Resilience in the Cloud: Challenges and Opportunities. , 2012, , 65-81.		17
1412	MapReduce neural network framework for efficient content based image retrieval from large datasets in the cloud. , 2012, , .		13
1413	Design of an Adaptive Framework for Utility-Based Optimization of Scientific Applications in the Cloud. , 2012, , .		8
1414	Computing large-scale distance matrices on GPU. , 2012, , .		5
1415	Extracting weblog of Siam University for learning user behavior on MapReduce. , 2012, , .		12
1416	Optimizing Completion Time and Resource Provisioning of Pig Programs. , 2012, , .		6
1417	A Distributed Methodology for Imbalanced Classification Problems. , 2012, , .		4
1418	Design and Implementation of Parallel Term Contribution Algorithm Based on Mapreduce Model. , 2012, , .		0
1419	PIC: Partitioned Iterative Convergence for Clusters. , 2012, , .		5
1420	Scaling in distributed evolutionary algorithms with persistent population. , 2012, , .		5
1421	Pool vs. Island Based Evolutionary Algorithms: An Initial Exploration. , 2012, , .		6
1422	Hatman: Intra-cloud Trust Management for Hadoop. , 2012, , .		31
1423	Evaluating Hadoop for Data-Intensive Scientific Operations. , 2012, , .		27
1424	Dash: A Novel Search Engine for Database-Generated Dynamic Web Pages. , 2012, , .		1

#	ARTICLE	IF	CITATIONS
1425	Towards matching food metadata in emergency decision-making using ontology and MapReduce. , 2012, , .		1
1426	A Cloud Oriented Account Service Mechanism for SME SaaS Ecosystem. , 2012, , .		3
1427	A GPU Implementation of Generalized Graph Processing Algorithm GIM-V. , 2012, , .		3
1428	A mapreduce approach for processing large-scale remote sensing images. , 2012, , .		2
1429	A Large-Scale Graph Learning Framework of Technological Gatekeepers by MapReduce. , 2012, , .		0
1430	Research on dynamic resource allocation with cooperation strategy in cloud computing. , 2012, , .		7
1431	Parallel Processing of Massive EEG Data with MapReduce. , 2012, , .		16
1432	Data Analytics in Ubiquitous Sensor-Based Health Information Systems. , 2012, , .		24
1433	Two-stage sparse graph construction using MinHash on MapReduce. , 2012, , .		1
1434	Scan-Sharing for Optimizing RDF Graph Pattern Matching on MapReduce. , 2012, , .		9
1435	Differentially private top-k query over MapReduce. , 2012, , .		6
1436	Parallelization with Multiplicative Algorithms for Big Data Mining. , 2012, , .		14
1437	Efficient Distributed Parallel Top-Down Computation of ROLAP Data Cube Using MapReduce. Lecture Notes in Computer Science, 2012, , 168-179.	1.0	9
1438	Dynamic Cloud Deployment of a MapReduce Architecture. IEEE Internet Computing, 2012, 16, 40-50.	3.2	18
1439	A cloud model and concept prototype for cognitive radio networks. IEEE Wireless Communications, 2012, 19, 49-58.	6.6	30
1440	Analyzing runtime adaptability of collaboration patterns. , 2012, , .		4
1441	Accelerating Bayesian network parameter learning using Hadoop and MapReduce. , 2012, , .		21
1442	Clydesdale. , 2012, , .		36

#	ARTICLE	IF	CITATIONS
1443	BIG DATA PROCESSING: BIG CHALLENGES AND OPPORTUNITIES. Journal of Interconnection Networks, 2012, 13, 1250009.	0.6	26
1444	JRBridge: A Framework of Large-Scale Statistical Computing for R. , 2012, , .		1
1445	K%-Fair scheduling: A flexible task scheduling strategy for balancing fairness and efficiency in MapReduce systems. , 2012, , .		0
1446	Multi-file queries performance improvement through data placement in Hadoop. , 2012, , .		2
1447	Optimization of task assignment strategy for map-reduce. , 2012, , .		5
1448	Dual-JT: Toward the high availability of JobTracker in Hadoop. , 2012, , .		4
1449	BIDE-Based Parallel Mining of Frequent Closed Sequences with MapReduce. Lecture Notes in Computer Science, 2012, , 177-186.	1.0	9
1450	SkewTune. , 2012, , .		304
1451	Constructing Virtual Documents for Ontology Matching Using MapReduce. Lecture Notes in Computer Science, 2012, , 48-63.	1.0	8
1452	Smart Traffic Cloud: An Infrastructure for Traffic Applications. , 2012, , .		14
1453	Layered Models for General Parallel Computation Based on Heterogeneous System. , 2012, , .		0
1454	Big data @ facebook. , 2012, , .		23
1455	Harnessing user library statistics for research evaluation and knowledge domain visualization. , 2012, , .		10
1456	Distributed Ontology Cloud Storage System. , 2012, , .		3
1457	Enabling Active Data Archival over Cloud. , 2012, , .		4
1458	Exploiting and Evaluating MapReduce for Large-Scale Graph Mining. , 2012, , .		0
1459	Affinity-aware Virtual Cluster Optimization for MapReduce Applications. , 2012, , .		5
1460	Scalable Clustering Using PACT Programming Model. , 2012, , .		0

#	ARTICLE	IF	CITATIONS
1461	BSPCloud: A Programming Model for Cloud Computing. , 2012, , .		5
1462	Green-aware workload scheduling in geographically distributed data centers. , 2012, , .		59
1463	CMQ - A lightweight, asynchronous high-performance messaging queue for the cloud. Journal of Cloud Computing: Advances, Systems and Applications, 2012, 1, 20.	2.1	3
1464	MR-CloudSim: Designing and implementing MapReduce computing model on CloudSim. , 2012, , .		34
1465	Design Dynamic Data Allocation Scheduler to Improve MapReduce Performance in Heterogeneous Clouds. , 2012, , .		7
1466	Hierarchical MapReduce Programming Model and Scheduling Algorithms. , 2012, , .		28
1467	IncMR: Incremental Data Processing Based on MapReduce. , 2012, , .		26
1468	Maestro: Replica-Aware Map Scheduling for MapReduce. , 2012, , .		70
1469	Minimizing Cost of Virtual Machines for Deadline-Constrained MapReduce Applications in the Cloud. , 2012, , .		52
1470	A conceptual model and prototype of Cognitive Radio Cloud Networks in TV White Spaces. , 2012, , .		1
1471	A new mechanism to ensure integrity for MapReduce in cloud computing. , 2012, , .		1
1472	The Partition Cost Model for Load Balancing in MapReduce. Service Science: Research and Innovations in the Service Economy, 2012, , 371-387.	1.1	6
1473	Distributed Simulated Annealing with MapReduce. Lecture Notes in Computer Science, 2012, , 466-476.	1.0	15
1474	Anomaly Teletraffic Intrusion Detection Systems on Hadoop-Based Platforms: A Survey of Some Problems and Solutions. , 2012, , .		15
1475	MARLA: MapReduce for Heterogeneous Clusters. , 2012, , .		33
1476	Modeling and Analyzing Key Performance Factors of Shared Memory MapReduce. , 2012, , .		4
1477	Performance Optimization for Short MapReduce Job Execution in Hadoop. , 2012, , .		16
1478	Dynamic slot-based task scheduling based on node workload in a MapReduce computation model. , 2012, , .		3

#	ARTICLE	IF	CITATIONS
1479	LoadAtomizer: A locality and I/O load aware task scheduler for MapReduce. , 2012, , .		5
1480	Improve availability of fault-tolerant computing: Optimal multi-task allocation in MapReduce. , 2012, , .		3
1481	Tuple MapReduce: Beyond Classic MapReduce. , 2012, , .		7
1482	An locality-aware scheduling based on a novel scheduling model to improve system throughput of MapReduce cluster. , 2012, , .		4
1483	ESAMR: An Enhanced Self-Adaptive MapReduce Scheduling Algorithm. , 2012, , .		49
1484	A de novo next generation genomic sequence assembler based on string graph and MapReduce cloud computing framework. BMC Genomics, 2012, 13, S28.	1.2	30
1485	Accelerating MapReduce Analytics Using CometCloud. , 2012, , .		8
1486	Hadoop plugin for distributed and parallel image processing. , 2012, , .		4
1487	Parallel rough set based knowledge acquisition using MapReduce from big data. , 2012, , .		26
1488	Task Scheduling Algorithm with Fault Tolerance for Cloud. , 2012, , .		12
1489	Fractal MapReduce decomposition of sequence alignment. Algorithms for Molecular Biology, 2012, 7, 12.	0.3	21
1490	Tarazu. , 2012, , .		133
1491	AROMA. , 2012, , .		155
1492	Coupling scheduler for MapReduce/Hadoop. , 2012, , .		8
1493	An In-depth Study of Map Reduce in Cloud Environment. , 2012, , .		3
1494	Large-scale volunteer computing over the Internet. Journal of Internet Services and Applications, 2012, 3, 329-346.	1.6	10
1495	Computational challenges in nanoparticle partition function calculation. , 2012, , .		0
1496	Scheduling Data Intensive Workloads through Virtualization on MapReduce based Clouds. International Journal of Distributed and Parallel Systems, 2012, 3, 99-110.	0.2	6

#	ARTICLE	IF	CITATIONS
1497	Towards Efficient Query Processing on Massive Time-Evolving Graphs. , 2012, , .		19
1499	WIKI::SCORE A collaborative environment for music transcription and publishing1. Information Services and Use, 2012, 31, 177-187.	0.1	0
1500	Large-Scale Data Processing Using MapReduce in Cloud Computing Environment. International Journal on Web Service Computing, 2012, 3, 1-13.	0.2	9
1501	Data-Intensive Computing: A Challenge for the 21st Century. , 0, , 1-11.		1
1502	Large-Scale Data Management Techniques in Cloud Computing Platforms. , 0, , 85-123.		0
1503	Parallel Processing of cluster by Map Reduce. International Journal of Distributed and Parallel Systems, 2012, 3, 167-179.	0.2	12
1504	Evaluation of Data Processing Using MapReduce Framework in Cloud and Stand - Alone Computing. International Journal of Distributed and Parallel Systems, 2012, 3, 51-63.	0.2	3
1505	Development of an intelligent distributed news retrieval system. International Journal of Knowledge-Based and Intelligent Engineering Systems, 2012, 16, 129-140.	0.7	2
1506	GPU Scripting and Code Generation with PyCUDA. , 2012, , 373-385.		4
1507	Scalable Computing for Evolutionary Genomics. Methods in Molecular Biology, 2012, 856, 529-545.	0.4	5
1508	On Cloud computational models and the heterogeneity challenge. Journal of Internet Services and Applications, 2012, 3, 77-86.	1.6	34
1509	Explicit coordination to prevent congestion in data center networks. Cluster Computing, 2012, 15, 183-200.	3.5	4
1510	Reliable MapReduce computing on opportunistic resources. Cluster Computing, 2012, 15, 145-161.	3.5	15
1511	MapReduce Workload Modeling with Statistical Approach. Journal of Grid Computing, 2012, 10, 279-310.	2.5	46
1512	iMapReduce: A Distributed Computing Framework for Iterative Computation. Journal of Grid Computing, 2012, 10, 47-68.	2.5	139
1513	Enhancing privacy in cloud computing via policy-based obfuscation. Journal of Supercomputing, 2012, 61, 267-291.	2.4	56
1514	Reference deployment models for eliminating user concerns on cloud security. Journal of Supercomputing, 2012, 61, 337-352.	2.4	16
1515	Filter-embedding semiring fusion for programming with MapReduce. Formal Aspects of Computing, 2012, 24, 623-645.	1.4	10

#	ARTICLE	IF	CITATIONS
1516	Costs of authentic pharmaceuticals: research on qualitative and quantitative aspects of enabling anti-counterfeiting in RFID-aided supply chains. <i>Personal and Ubiquitous Computing</i> , 2012, 16, 271-289.	1.9	17
1517	A vision for a biomedical cloud. <i>Journal of Internal Medicine</i> , 2012, 271, 122-130.	2.7	38
1518	Towards efficient data search and subsetting of large-scale atmospheric datasets. <i>Future Generation Computer Systems</i> , 2012, 28, 112-118.	4.9	17
1519	Adaptive heterogeneous language support within a cloud runtime. <i>Future Generation Computer Systems</i> , 2012, 28, 128-135.	4.9	5
1520	OPTIMIS: A holistic approach to cloud service provisioning. <i>Future Generation Computer Systems</i> , 2012, 28, 66-77.	4.9	307
1521	Dynamic energy efficient data placement and cluster reconfiguration algorithm for MapReduce framework. <i>Future Generation Computer Systems</i> , 2012, 28, 119-127.	4.9	82
1522	The Aneka platform and QoS-driven resource provisioning for elastic applications on hybrid Clouds. <i>Future Generation Computer Systems</i> , 2012, 28, 861-870.	4.9	144
1523	Performance evaluation of hybrid programming patterns for large CPU/GPU heterogeneous clusters. <i>Computer Physics Communications</i> , 2012, 183, 1172-1181.	3.0	25
1524	A parallel method for computing rough set approximations. <i>Information Sciences</i> , 2012, 194, 209-223.	4.0	100
1525	MapReduce based Betweenness Approximation Engineering in Large Scale Graph. <i>Systems Engineering Procedia</i> , 2012, 5, 162-167.	0.3	2
1526	Program Ultra-Dispatcher for launching applications in a customization manner on cloud computing. <i>Journal of Network and Computer Applications</i> , 2012, 35, 423-446.	5.8	4
1527	Online optimization for scheduling preemptable tasks on IaaS cloud systems. <i>Journal of Parallel and Distributed Computing</i> , 2012, 72, 666-677.	2.7	327
1528	Fast and accurate link prediction in social networking systems. <i>Journal of Systems and Software</i> , 2012, 85, 2119-2132.	3.3	154
1529	Reasoning with Large Scale Ontologies in Fuzzy pD* Using MapReduce. <i>IEEE Computational Intelligence Magazine</i> , 2012, 7, 54-66.	3.4	16
1530	Evaluating High-Performance Computing on Google App Engine. <i>IEEE Software</i> , 2012, 29, 52-58.	2.1	25
1531	Modeling Energy-Time Trade-Offs in VLSI Computation. <i>IEEE Transactions on Computers</i> , 2012, 61, 530-547.	2.4	4
1532	VDoc+: a virtual document based approach for matching large ontologies using MapReduce. <i>Journal of Zhejiang University: Science C</i> , 2012, 13, 257-267.	0.7	9
1533	Descriptive matrix factorization for sustainability Adopting the principle of opposites. <i>Data Mining and Knowledge Discovery</i> , 2012, 24, 325-354.	2.4	53

#	ARTICLE	IF	CITATIONS
1534	Data and task parallelism in ILP using MapReduce. Machine Learning, 2012, 86, 141-168.	3.4	44
1535	Multi-pass sorted neighborhood blocking with MapReduce. Computer Science - Research and Development, 2012, 27, 45-63.	2.7	58
1536	An Efficient Grouped Virtual Mapreduce Cluster. , 2013, , .		0
1537	A throughput optimal algorithm for map task scheduling in mapreduce with data locality. Performance Evaluation Review, 2013, 40, 33-42.	0.4	19
1538	Improving ReduceTask data locality for sequential MapReduce jobs. , 2013, , .		47
1539	Towards an Optimized Big Data Processing System. , 2013, , .		5
1540	Adaptive data transmission in the cloud. , 2013, , .		2
1541	Cloud Services for Improved User Experience in Sharing Mobile Videos. , 2013, , .		5
1542	A task routing approach to large-scale scheduling. Future Generation Computer Systems, 2013, 29, 1097-1111.	4.9	13
1543	Exploiting symmetries for scaling loopy belief propagation and relational training. Machine Learning, 2013, 92, 91-132.	3.4	50
1544	Designing and testing a pool-based evolutionary algorithm. Natural Computing, 2013, 12, 149-162.	1.8	6
1545	The grid, the load and the gradient. Natural Computing, 2013, 12, 69-85.	1.8	2
1546	Scientific computing with Google App Engine. Future Generation Computer Systems, 2013, 29, 1851-1859.	4.9	16
1547	Performance evaluation of a MongoDB and hadoop platform for scientific data analysis. , 2013, , .		100
1548	Integrating remotely sensed and ground observations for modeling, analysis, and decision support. , 2013, , .		1
1549	ComMapReduce: An improvement of MapReduce with lightweight communication mechanisms. Data and Knowledge Engineering, 2013, 88, 224-247.	2.1	16
1550	Graph-based semi-supervised learning with multi-modality propagation for large-scale image datasets. Journal of Visual Communication and Image Representation, 2013, 24, 295-302.	1.7	15
1551	An open-source toolkit for mining Wikipedia. Artificial Intelligence, 2013, 194, 222-239.	3.9	229

#	ARTICLE	IF	CITATIONS
1552	MatchTree: Flexible, scalable, and fault-tolerant wide-area resource discovery with distributed matchmaking and aggregation. <i>Future Generation Computer Systems</i> , 2013, 29, 1596-1610.	4.9	8
1553	Managing Data and Processes in Cloud-Enabled Large-Scale Sensor Networks: State-of-the-Art and Future Research Directions. , 2013, , .		55
1554	Cost-minimizing preemptive scheduling of mapreduce workloads on hybrid clouds. , 2013, , .		3
1555	A Lightweight Continuous Jobs Mechanism for MapReduce Frameworks. , 2013, , .		2
1556	Bi-Hadoop: Extending Hadoop to Improve Support for Binary-Input Applications. , 2013, , .		5
1557	An efficient and compact indexing scheme for large-scale data store. , 2013, , .		8
1559	Join Optimization for Large-Scale Data Analysis in MapReduce. <i>Lecture Notes in Electrical Engineering</i> , 2013, , 651-657.	0.3	0
1560	A MapReduce task scheduling algorithm for deadline constraints. <i>Cluster Computing</i> , 2013, 16, 651-662.	3.5	55
1561	h-MapReduce: A Framework for Workload Balancing in MapReduce. , 2013, , .		23
1562	Deadline-Based MapReduce Workload Management. <i>IEEE Transactions on Network and Service Management</i> , 2013, 10, 231-244.	3.2	37
1563	LaSA: A locality-aware scheduling algorithm for Hadoop-MapReduce resource assignment. , 2013, , .		22
1564	Map task scheduling in MapReduce with data locality: Throughput and heavy-traffic optimality. , 2013, , .		49
1565	Disk-based k-mer counting on a PC. <i>BMC Bioinformatics</i> , 2013, 14, 160.	1.2	61
1566	Applying representational state transfer (REST) architecture to archetype-based electronic health record systems. <i>BMC Medical Informatics and Decision Making</i> , 2013, 13, 57.	1.5	26
1567	SAAD, a content based Web Spam Analyzer and Detector. <i>Journal of Systems and Software</i> , 2013, 86, 2906-2918.	3.3	16
1568	An improved partitioning mechanism for optimizing massive data analysis using MapReduce. <i>Journal of Supercomputing</i> , 2013, 66, 539-555.	2.4	39
1569	Mining and generating large-scaled social networks via MapReduce. <i>Social Network Analysis and Mining</i> , 2013, 3, 1449-1469.	1.9	4
1570	Exploiting geospatial and chronological characteristics in data streams to enable efficient storage and retrievals. <i>Future Generation Computer Systems</i> , 2013, 29, 1049-1061.	4.9	28

#	ARTICLE	IF	CITATIONS
1571	Handling partitioning skew in MapReduce using LEEN. Peer-to-Peer Networking and Applications, 2013, 6, 409-424.	2.6	40
1572	Tiled-MapReduce. Transactions on Architecture and Code Optimization, 2013, 10, 1-30.	1.6	12
1573	Choosy. , 2013, , .		126
1574	Artificial Intelligence and Big Data. IEEE Intelligent Systems, 2013, 28, 96-99.	4.0	340
1575	Anomalies in network traffic. , 2013, , .		5
1576	Skew-Aware Task Scheduling in Clouds. , 2013, , .		0
1577	Investigating hybrid SSD FTL schemes for Hadoop workloads. , 2013, , .		11
1578	Exploiting inter-operation parallelism for matrix chain multiplication using MapReduce. Journal of Supercomputing, 2013, 66, 594-609.	2.4	2
1579	A compound OpenMP/MPI program development toolkit for hybrid CPU/GPU clusters. Journal of Supercomputing, 2013, 66, 381-405.	2.4	9
1580	Efficient programming paradigm for video streaming processing on TILE64 platform. Journal of Supercomputing, 2013, 65, 823-847.	2.4	0
1581	Development of a virtualized supercomputing environment for genomic analysis. Journal of Supercomputing, 2013, 65, 71-85.	2.4	2
1582	Parallelizing Multiple Group-by queries using MapReduce: optimization and cost estimation. Telecommunication Systems, 2013, 52, 635-645.	1.6	8
1583	Parallel computation of determinants of matrices with multivariate polynomial entries. Science China Information Sciences, 2013, 56, 1-16.	2.7	3
1584	Indexing and searching 100M images with map-reduce. , 2013, , .		59
1585	Mining Frequent Patterns from Uncertain Data with MapReduce for Big Data Analytics. Lecture Notes in Computer Science, 2013, , 440-455.	1.0	47
1586	Applying MapReduce to Learning User Preferences in Near Real-Time. Lecture Notes in Computer Science, 2013, , 15-28.	1.0	3
1587	An Efficient Data Access Method Exploiting Quadrees on MapReduce Frameworks. Lecture Notes in Computer Science, 2013, , 86-100.	1.0	4
1588	Machine learning for big data. , 2013, , .		26

#	ARTICLE	IF	CITATIONS
1589	The emerging role of cloud computing in molecular modelling. Journal of Molecular Graphics and Modelling, 2013, 44, 177-187.	1.3	29
1590	Big Data Challenges in Industrial Automation. Lecture Notes in Computer Science, 2013, , 305-316.	1.0	23
1591	ELM-MapReduce: MapReduce accelerated extreme learning machine for big spatial data analysis. , 2013, , .		11
1592	Scalable maximum clique computation using MapReduce. , 2013, , .		38
1593	Research and application of massive data processing technology. , 2013, , .		1
1594	An online service-oriented performance profiling tool for cloud computing systems. Frontiers of Computer Science, 2013, 7, 431-445.	1.6	8
1595	MapReduce optimization algorithm based on machine learning in heterogeneous cloud environment. Journal of China Universities of Posts and Telecommunications, 2013, 20, 77-121.	0.8	6
1596	Improving MapReduce Performance by Streaming Input Data from Multiple Replicas. , 2013, , .		0
1597	Dynamic Data Partitioning and Virtual Machine Mapping: Efficient Data Intensive Computation. , 2013, , .		0
1598	Introducing PRECIP: An API for Managing Repeatable Experiments in the Cloud. , 2013, , .		16
1599	Map-Reduce for Calibrating Massive Bus Trajectory Data. , 2013, , .		0
1600	Modeling the Performance of MapReduce under Resource Contentions and Task Failures. , 2013, , .		6
1601	Real-Time Handling of Network Monitoring Data Using a Data-Intensive Framework. , 2013, , .		5
1602	Map-Side Merge Joins for Scalable SPARQL BGP Processing. , 2013, , .		12
1603	Insight and reduction of MapReduce stragglers in heterogeneous environment. , 2013, , .		6
1604	Visual Analytics Framework for Cloud Infrastructure Data. , 2013, , .		1
1605	Concurrency Optimized Task Scheduling for Workflows in Cloud. , 2013, , .		7
1606	Parallelized Jaccard-based learning method and MapReduce implementation for mobile devices recognition from massive network data. China Communications, 2013, 10, 71-84.	2.0	7

#	ARTICLE	IF	CITATIONS
1607	Matrix-Query: A Distributed SQL-Like Query Processing Model for Large Database Clusters. , 2013, , .		0
1608	WARP: Workload-aware replication and partitioning for RDF. , 2013, , .		62
1609	Improving HDFS performance using local caching system. , 2013, , .		3
1610	Enabling Proactive Data Management in Virtualized Hadoop Clusters Based on Predicted Data Activity Patterns. , 2013, , .		7
1611	Meta-learning for large scale machine learning with MapReduce. , 2013, , .		3
1612	A New Hybrid Clustering Algorithm Based on Stimulated Annealing. , 2013, , .		1
1613	Verification and validation of MapReduce program model for parallel K-means algorithm on Hadoop cluster. , 2013, , .		10
1614	Resource aware scheduling in Hadoop for heterogeneous workloads based on load estimation. , 2013, , .		7
1615	The Use of Locality Information on Data Intensive Parallel File Systems. , 2013, , .		1
1616	A Deadline Scheduler for Jobs in Distributed Systems. , 2013, , .		14
1617	An adaptive machine learning on Map-Reduce framework for improving performance of large-scale data analysis on EC2. , 2013, , .		7
1618	Scheduling real-time workflow on MapReduce-based cloud. , 2013, , .		8
1619	A Throughput Driven Task Scheduler for Improving MapReduce Performance in Job-Intensive Environments. , 2013, , .		7
1620	A self-tuning system based on application Profiling and Performance Analysis for optimizing Hadoop MapReduce cluster configuration. , 2013, , .		36
1621	Resource optimization for speculative execution in a MapReduce Cluster. , 2013, , .		1
1622	Minimizing thermal residual stresses in ceramic matrix composites by using Iterative MapReduce guided particle swarm optimization algorithm. Composite Structures, 2013, 99, 388-396.	3.1	14
1623	Mining search and browse logs for web search. ACM Transactions on Intelligent Systems and Technology, 2013, 4, 1-37.	2.9	44
1624	Clustering on the cloud. , 2013, , .		6

#	ARTICLE	IF	CITATIONS
1625	The Cloud Streaming Service Migration in Cloud Video Storage System. , 2013, , .		3
1626	Adapting Scientific Applications to Cloud by Using Distributed Computing Frameworks. , 2013, , .		4
1627	Parallel Hierarchical Affinity Propagation with MapReduce. , 2013, , .		0
1628	MobSched: Customizable scheduler for mobile cloud computing. , 2013, , .		7
1629	Learning to crawl deep web. Information Systems, 2013, 38, 801-819.	2.4	32
1630	Learning-Based Interactive Retrieval in Large-Scale Multimedia Collections. Lecture Notes in Computer Science, 2013, , 1-17.	1.0	0
1631	Multi-resolution Social Network Community Identification and Maintenance on Big Data Platform. , 2013, , .		4
1632	Building OLAP cubes on a Cloud Computing environment with MapReduce. , 2013, , .		13
1633	Analytics over Big Data: Exploring the Convergence of DataWarehousing, OLAP and Data-Intensive Cloud Infrastructures. , 2013, , .		61
1634	Scaling MapReduce Applications Across Hybrid Clouds to Meet Soft Deadlines. , 2013, , .		56
1635	Cloud Federation: Effects of Federated Compute Resources on Quality of Service and Cost*. , 2013, , .		23
1636	Data warehousing and OLAP over big data. , 2013, , .		90
1637	Collaborative pseudo-relevance feedback. Expert Systems With Applications, 2013, 40, 6805-6812.	4.4	12
1638	An Economical Query Cost Model in the Cloud. Lecture Notes in Computer Science, 2013, , 16-30.	1.0	3
1639	BioPig: a Hadoop-based analytic toolkit for large-scale sequence data. Bioinformatics, 2013, 29, 3014-3019.	1.8	101
1641	Efficient Data Analytics Over Cloud. Advances in Computers, 2013, , 367-401.	1.2	0
1642	Model-Driven Engineering of Reliable Fault-Tolerant Systemsâ€™A State-of-the-Art Survey. Advances in Computers, 2013, , 119-205.	1.2	7
1643	Data-intensive Spatial Indexing on the Clouds. Procedia Computer Science, 2013, 18, 2615-2618.	1.2	6

#	ARTICLE	IF	CITATIONS
1644	CAP. , 2013, , .		13
1645	Data Prospecting—A Step Towards Data Intensive Science. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2013, 6, 1233-1241.	2.3	0
1646	Katana: Generalized Data Processing on Peer-to-Peer Overlays. , 2013, , .		1
1647	Performance and Reliability Effects of Multi-tier Bidding on MapReduce in Auction-Based Clouds. , 2013, , .		2
1648	System G Data Store: Big, Rich Graph Data Analytics in the Cloud. , 2013, , .		5
1649	Trends and outlook for the massive-scale analytics stack. IBM Journal of Research and Development, 2013, 57, 2:1-2:11.	3.2	0
1650	Automatic tag recommendation for metadata annotation using probabilistic topic modeling. , 2013, , .		52
1651	Random forests on Hadoop for genome-wide association studies of multivariate neuroimaging phenotypes. BMC Bioinformatics, 2013, 14, S6.	1.2	34
1652	MapReduce Operations with WS-VLAM Workflow Management System. Procedia Computer Science, 2013, 18, 2599-2602.	1.2	4
1653	Accelerating iterative algorithms with asynchronous accumulative updates on FPGAs. , 2013, , .		1
1654	Cloud Based Predictive Analytics: Text Classification, Recommender Systems and Decision Support. , 2013, , .		8
1655	BBS opinion leader mining based on an improved PageRank algorithm using MapReduce. , 2013, , .		7
1656	A New Schedule Strategy for Heterogenous Workload-aware in Hadoop. , 2013, , .		1
1657	H-PFSP: Efficient Hybrid Parallel PFSP Protected Scheduling for MapReduce System. , 2013, , .		1
1658	MapReduce Design of K-Means Clustering Algorithm. , 2013, , .		36
1659	Implementing quasi-parallel breadth-first search in MapReduce for large-scale social network mining. , 2013, , .		0
1660	Towards Real-Time Analytics in the Cloud. , 2013, , .		23
1661	Evaluating task scheduling in hadoop-based cloud systems. , 2013, , .		7

#	ARTICLE	IF	CITATIONS
1662	Index-based join operations in Hive. , 2013, , .		9
1663	Utilising condor for data parallel analytics in an IoT context — An experience report. , 2013, , .		8
1664	MapReducing GEPETO or Towards Conducting a Privacy Analysis on Millions of Mobility Traces. , 2013, , .		4
1665	A distributed approach for graph-oriented multidimensional analysis. , 2013, , .		10
1666	Distributed classification of Persian News (Case study: Hamshahri News dataset). , 2013, , .		1
1667	SAFAL: A MapReduce Spatio-temporal Analyzer for UNAVCO FTP Logs. , 2013, , .		2
1668	A MapReduce based glowworm swarm optimization approach for multimodal functions. , 2013, , .		9
1669	Mining repetitive sequences using a big data ecosystem. , 2013, , .		0
1670	H2T: A Simple Hadoop-to-Twister Translator for Cloud Computing. , 2013, , .		1
1671	Implementation and Evaluation of the JobTracker Initiative Task Scheduling on Hadoop. , 2013, , .		6
1672	A CUDA-enabled Hadoop cluster for fast distributed image processing. , 2013, , .		11
1673	An MPI-enabled MapReduce framework for molecular dynamics simulation applications. , 2013, , .		1
1674	MC framework: High-performance distributed framework for standalone data analysis packages over Hadoop-based cloud. , 2013, , .		0
1676	Dynamic Scheduling for Wireless Data Center Networks. IEEE Transactions on Parallel and Distributed Systems, 2013, 24, 2365-2374.	4.0	32
1677	Scalable intrusion detection systems log analysis using cloud computing infrastructure. , 2013, , .		15
1678	A MapReduce and Information Compression Based Social Community Structure Mining Method. , 2013, , .		0
1679	The Convergence of Big Data and Mobile Computing. , 2013, , .		2
1680	Techniques for Graph Analytics on Big Data. , 2013, , .		26

#	ARTICLE	IF	CITATIONS
1681	Rapid processing of remote sensing images based on cloud computing. Future Generation Computer Systems, 2013, 29, 1963-1968.	4.9	53
1682	Tractography density and network measures in Alzheimer'S disease. , 2013, 2013, 692-695.		32
1683	The CPS with the Hadoop ecosystems. , 2013, , .		1
1684	Distributed Clique Percolation based community detection on social networks using MapReduce. , 2013, , .		8
1685	Quantitative reactive modeling and verification. Computer Science - Research and Development, 2013, 28, 331-344.	2.7	29
1686	Scalable model of parallel computations for applications with intensive input-output. Journal of Computer and Systems Sciences International, 2013, 52, 831-836.	0.2	0
1687	Power consumption evaluation of all-optical data center networks. Cluster Computing, 2013, 16, 611-623.	3.5	42
1688	Lit: A high performance massive data computing framework based on CPU/GPU cluster. , 2013, , .		6
1689	TSHC: Trusted Scheme for Hadoop Cluster. , 2013, , .		7
1690	MapReduce: Limitations, Optimizations and Open Issues. , 2013, , .		33
1691	Elephant, Do Not Forget Everything! Efficient Processing of Growing Datasets. , 2013, , .		3
1692	Scalable Parallel Join for Huge Tables. , 2013, , .		1
1693	SciFlow: A dataflow-driven model architecture for scientific computing using Hadoop. , 2013, , .		7
1694	Assisting developers of Big Data Analytics Applications when deploying on Hadoop clouds. , 2013, , .		84
1695	Block Sampling: Efficient Accurate Online Aggregation in MapReduce. , 2013, , .		14
1696	Data Pipeline in MapReduce. , 2013, , .		3
1697	A simple implementation of distributed vertical search and information integration technology. Wuhan University Journal of Natural Sciences, 2013, 18, 511-516.	0.2	1
1698	Internet-scale support for map-reduce processing. Journal of Internet Services and Applications, 2013, 4, 18.	1.6	17

#	ARTICLE	IF	CITATIONS
1699	Partition-Based Online Aggregation with Shared Sampling in the Cloud. Journal of Computer Science and Technology, 2013, 28, 989-1011.	0.9	13
1700	Analysis of I/O Performance on an Amazon EC2 Cluster Compute and High I/O Platform. Journal of Grid Computing, 2013, 11, 613-631.	2.5	15
1701	Paragon. , 2013, , .		375
1702	Dache: A data aware caching for big-data applications using the MapReduce framework. , 2013, , .		21
1703	Implementing BFS-based Traversals of RDF Graphs over MapReduce Efficiently. , 2013, , .		4
1704	A parallel optical implementation of arithmetic operations. Optics and Laser Technology, 2013, 49, 173-182.	2.2	3
1705	A Review on Distributed Application Processing Frameworks in Smart Mobile Devices for Mobile Cloud Computing. IEEE Communications Surveys and Tutorials, 2013, 15, 1294-1313.	24.8	244
1706	Big graph mining. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2013, 14, 29-36.	3.2	47
1707	ShmStreaming: A Shared Memory Approach for Improving Hadoop Streaming Performance. , 2013, , .		7
1708	Scalable high-dimensional indexing with Hadoop. , 2013, , .		4
1709	Non-intrusive Slot Layering in Hadoop. , 2013, , .		4
1710	Job scheduling for optimizing data locality in Hadoop clusters. , 2013, , .		9
1712	A cloud computing based framework for general 2D and 3D cellular automata simulation. Advances in Engineering Software, 2013, 65, 78-89.	1.8	6
1713	Scalable all-pairs similarity search in metric spaces. , 2013, , .		40
1714	On interference-aware provisioning for cloud-based big data processing. , 2013, , .		18
1715	Big Data Security Hardening Methodology Using Attributes Relationship. , 2013, , .		11
1716	High performance risk aggregation. , 2013, , .		0
1717	Towards Improving Query Performance of Web Feature Services (WFS) for Disaster Response. ISPRS International Journal of Geo-Information, 2013, 2, 67-81.	1.4	13

#	ARTICLE	IF	CITATIONS
1718	Web Multimedia Object Classification Using Cross-Domain Correlation Knowledge. IEEE Transactions on Multimedia, 2013, 15, 1920-1929.	5.2	14
1719	Text Categorization of Biomedical Data Sets Using Graph Kernels and a Controlled Vocabulary. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2013, 10, 1211-1217.	1.9	26
1720	Differential Power Processing for DC Systems. IEEE Transactions on Power Electronics, 2013, 28, 1795-1806.	5.4	83
1721	Power and energy characteristics of MapReduce data movements. , 2013, , .		8
1722	Efficient Geo-distributed Data Processing with Rout. , 2013, , .		5
1723	It's you on photo?: Automatic detection of Twitter accounts infected with the Blackhole Exploit Kit. , 2013, , .		1
1724	Algorithm Parallelization Using Software Design Patterns, an Embedded Case Study Approach. , 2013, , .		1
1725	Complete storm identification algorithms from big raw rainfall data using MapReduce framework. , 2013, , .		4
1726	Kanthaka: Big Data Caller Detail Record (CDR) Analyzer for Near Real Time Telecom Promotions. , 2013, , .		6
1727	Multi-core Supported High Performance Security Analytics. , 2013, , .		0
1728	A managed distributed processing pipeline with Storm and Mesos. , 2013, , .		2
1729	A reconfigurable computing architecture for semantic information filtering. , 2013, , .		1
1730	A framework for cloud-based POI search and trip planning systems. , 2013, , .		1
1731	Tango. , 2013, , .		72
1732	Cloud Computing: A Novel Approach for Improving Service Capability of Wireless Mesh Networks. Advanced Materials Research, 2013, 765-767, 1142-1145.	0.3	0
1733	Large Scale predictive analytics for real-time energy management. , 2013, , .		22
1734	Web Monitoring " eine EinfÄ¼hrung. Hmd, 2013, 50, 7-12.	0.3	0
1735	Tag-Based Integrated Semantic Ontology Construction and Evolution. , 2013, , .		2

#	ARTICLE	IF	CITATIONS
1736	Towards Machine Learning-Based Auto-tuning of MapReduce. , 2013, , .		71
1737	Scheduler for data-intensive workflows in public clouds. , 2013, , .		7
1738	A scalable framework for cloud powered workflow execution. , 2013, , .		1
1739	Performance Modeling of MapReduce Jobs in Heterogeneous Cloud Environments. , 2013, , .		51
1740	Comparing the Performance and Power Usage of GPU and ARM Clusters for Map-Reduce. , 2013, , .		2
1741	Optimizations and Analysis of BSP Graph Processing Models on Public Clouds. , 2013, , .		27
1742	COLA: A cloud-based system for online aggregation. , 2013, , .		2
1743	Triples in the clouds. , 2013, , .		3
1744	A scalable data repository for recording self-managed longitudinal health data of digital patients. , 2013, , .		0
1745	A novel decentralized asynchronous scheduler for Hadoop. , 2013, , .		1
1746	Scaling SeerSuite in the Cloud. , 2013, , .		3
1747	ExoApp: Performance Evaluation of Data-Intensive Applications on ExoGENI. , 2013, , .		0
1748	Adapting CakeDB to Integrate High-Pressure Big Data Streams with Low-Pressure Systems. , 2013, , .		0
1749	Non Linear Divisible Loads: There is No Free Lunch. , 2013, , .		7
1750	An event-processing system alerting analytically to networked vehicles. , 2013, , .		0
1751	A characteristic study on failures of production distributed data-parallel programs. , 2013, , .		21
1752	MapReduce-based frequent itemset mining for analysis of electronic evidence. , 2013, , .		3
1753	Coupling task progress for MapReduce resource-aware scheduling. , 2013, , .		61

#	ARTICLE	IF	CITATIONS
1754	Obtaining ground-truth software architectures. , 2013, , .		54
1755	Ray tracing in the cloud using MapReduce. , 2013, , .		3
1756	Multi-query Unification for Generating Efficient Big Data Processing Components from a DFD. , 2013, , .		1
1757	Distributed Column Subset Selection on MapReduce. , 2013, , .		22
1758	Facilitating Scientific Workflow Configuration with Parameterized Workflow Skeletons. , 2013, , .		2
1759	Expediting search trend detection via prediction of query counts. , 2013, , .		21
1760	StreamMine3G OneClick – Deploy and Monitor ESP Applications with a Single Click. , 2013, , .		2
1761	RapidRAID: Pipelined erasure codes for fast data archival in distributed storage systems. , 2013, , .		27
1762	Data Intensive Computing of X-Ray Computed Tomography Reconstruction at the LSDF. , 2013, , .		1
1763	Correlation-based performance analysis for full-system MapReduce optimization. , 2013, , .		5
1764	Unraveling the Complexities of Life Sciences Data. Big Data, 2013, 1, 42-50.	2.1	46
1765	Dataset Scaling and MapReduce Performance. , 2013, , .		7
1766	Session 14: Information security. , 2013, , .		1
1767	Research of Video Data Store Based on Cloud Storage. , 2013, , .		0
1768	Fast scalable selection algorithms for large scale data. , 2013, , .		4
1769	Practical distributed classification using the Alternating Direction Method of Multipliers algorithm. , 2013, , .		13
1770	MapReduce algorithm for finding st-connectivity. , 2013, , .		2
1771	A concept of Generic Workspace for Big Data Processing in Humanities. , 2013, , .		1

#	ARTICLE	IF	CITATIONS
1772	Distributed Collaborative Filtering on a Single Chip Cloud Computer*. , 2013, , .		0
1773	Reality mining in urban space. , 2013, , .		1
1774	Can we analyze big data inside a DBMS?. , 2013, , .		19
1775	Design and Evaluation of Lifelog Mashup Platform with NoSQL Database. , 2013, , .		0
1776	Extracting Knowledge from Wikipedia Articles through Distributed Semantic Analysis. , 2013, , .		4
1777	A simple cluster-scaling policy for MapReduce clouds. , 2013, , .		0
1778	Priority Scheduling in MapReduce Based on Queuing Theory. , 2013, , .		1
1780	The DFrame: Parallel Programming Using a Distributed Framework Implemented in MPI. , 2013, , .		0
1781	Optical path scheduling methods considering host bandwidth in data center networks. , 2013, , .		3
1782	Cluster-Size Scaling and MapReduce Execution Times. , 2013, , .		10
1783	Scalable and Real-Time Deep Packet Inspection. , 2013, , .		9
1784	Large scale ad latency analysis. , 2013, , .		0
1785	Context-Aware Service Recommendation on High Performance Cluster. , 2013, , .		0
1786	Implementing Materialized View of Large-Scale Power Consumption Log Using MapReduce. , 2013, , .		2
1787	Distributed confidence-weighted classification on MapReduce. , 2013, , .		4
1788	Participatory networking. Computer Communication Review, 2013, 43, 327-338.	1.5	78
1789	Fair Quantized Congestion Notification in Data Center Networks. IEEE Transactions on Communications, 2013, 61, 4690-4699.	4.9	35
1790	Adapting MapReduce framework for genetic algorithm with large population. , 2013, , .		5

#	ARTICLE	IF	CITATIONS
1791	Co-processing SPMD computation on CPUs and GPUs cluster. , 2013, , .		1
1792	GPU-Accelerated Query by Humming Using Modified SPRING Algorithm. , 2013, , .		0
1793	Remote Management: An Architecture for Mechatronic Systems. , 2013, , .		0
1794	Exploring heterogeneity of unreliable machines for P2P backup. , 2013, , .		2
1795	Viability of the bulk synchronous parallel model for science on cloud. , 2013, , .		5
1796	SESM: Semantic enrichment of schema mappings. , 2013, , .		3
1797	The Design and Realization of a Science and Technology Dissertations Retrieval System Based on Vector Space Model under Cloud Computing Environment. , 2013, , .		0
1798	Paradigms for Realizing Machine Learning Algorithms. Big Data, 2013, 1, 207-214.	2.1	22
1799	HcBench: Methodology, development, and characterization of a customer usage representative big data/Hadoop benchmark. , 2013, , .		7
1800	QuPARA: Query-driven large-scale portfolio aggregate risk analysis on MapReduce. , 2013, , .		6
1801	New cloud consolidation architecture for electrical energy consumption management. , 2013, , .		4
1802	PLAR: Parallel Large-Scale Attribute Reduction on Cloud Systems. , 2013, , .		9
1803	The BTWorld use case for big data analytics: Description, MapReduce logical workflow, and empirical evaluation. , 2013, , .		13
1804	A Parallel Comparator of Documents. , 2013, , .		3
1805	Sensing as a Service: Challenges, Solutions and Future Directions. IEEE Sensors Journal, 2013, 13, 3733-3741.	2.4	168
1806	Application of the cloud computing technology in the sports training. , 2013, , .		8
1807	Improving load balancing for MapReduce-based entity matching. , 2013, , .		12
1808	Using Traditional Data Analysis Algorithms to Detect Access Patterns for Massive Data Processing. , 2013, , .		1

#	ARTICLE	IF	CITATIONS
1809	Efficient Local Statistical Analysis via Integral Histograms with Discrete Wavelet Transform. IEEE Transactions on Visualization and Computer Graphics, 2013, 19, 2693-2702.	2.9	38
1810	X10-PM: Transparent and Efficient Place Migration for APGAS Language. , 2013, , .		0
1811	Parallelizing Probabilistic Streaming Skyline Operator in Cloud Computing Environments. , 2013, , .		2
1812	Managing Network Reservation for Tenants in Oversubscribed Clouds. , 2013, , .		1
1813	Terabyte-scale image similarity search: Experience and best practice. , 2013, , .		15
1814	Simfrastructure: A Flexible and Adaptable Middleware Platform for Modeling and Analysis of Socially Coupled Systems. , 2013, , .		5
1815	GPS: A General Framework for Parallel Queries over Data Streams in Cloud. , 2013, , .		1
1816	Finding connected components in map-reduce in logarithmic rounds. , 2013, , .		74
1817	Result Integrity Check for MapReduce Computation on Hybrid Clouds. , 2013, , .		17
1818	Towards 60GHz wireless switching interconnect. , 2013, , .		1
1819	Ripple: Improved Architecture and Programming Model for Bulk Synchronous Parallel Style of Analytics. , 2013, , .		4
1820	Provisioning Legacy Simulation Applications in Product Lifecycle Management via a Cloud Platform. , 2013, , .		1
1821	Machine learning on Big Data. , 2013, , .		17
1822	Locality Based Data Partitioning in MapReduce. , 2013, , .		5
1823	Arbor: Efficient Large-Scale Graph Data Computing Model. , 2013, , .		1
1824	Optimizing Multi-join in Cloud Environment. , 2013, , .		2
1825	Mechanisms of Optimizing MapReduce Framework on High Performance Computer. , 2013, , .		3
1826	JUMMP: Job Uninterrupted Maneuverable MapReduce Platform. , 2013, , .		10

#	ARTICLE	IF	CITATIONS
1827	Savant: Automatic parallelization of a scheduling heuristic with machine learning. , 2013, , .		4
1828	A comparative study of the semi-elastic and fully-elastic mapreduce models. , 2013, , .		2
1829	Research of MapReduce Oriented Graphical Programming. , 2013, , .		0
1830	Map-reduce for calibrating massive bus trajectory data. , 2013, , .		0
1831	Best-fit mobile recharge pack recommendation. , 2013, , .		3
1832	Mass flow logs analysis system based on Hadoop. , 2013, , .		5
1833	A Scalable Implementation of Malware Detection Based on Network Connection Behaviors. , 2013, , .		4
1834	A BSP-Based Parallel Iterative Processing System with Multiple Partition Strategies for Big Graphs. , 2013, , .		7
1835	Improving Multi-job MapReduce Scheduling in an Opportunistic Environment. , 2013, , .		6
1836	Large Scale Distributed Acoustic Modeling With Back-Off $\{m, N\}$ -Grams. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 1158-1169.	3.8	5
1837	A Scalable Heterogeneous Parallelization Framework for Iterative Local Searches. , 2013, , .		10
1838	Summarizing Data Center Network Traffic by Partitioned Conservative Update. IEEE Communications Letters, 2013, 17, 2168-2171.	2.5	4
1839	Real-Time Scheduling in MapReduce Clusters. , 2013, , .		6
1840	HCube: A Server-centric Data Center Structure for Similarity Search. , 2013, , .		3
1841	Ease the Queue Oscillation: Analysis and Enhancement of DCTCP. , 2013, , .		10
1842	Data Compression Techniques and Algorithms for Effectively and Efficiently Managing Multidimensional Stream Cubes over Grids. , 2013, , .		0
1843	Efficiently integrating MapReduce-based computing into a Hurricane Loss Projection model. , 2013, , .		5
1844	Domain-Based MapReduce Programming Model for Complex Scientific Applications. , 2013, , .		2

#	ARTICLE	IF	CITATIONS
1845	Semantic characterization of MapReduce workloads. , 2013, , .		9
1846	Greening data center networks with flow preemption and energy-aware routing. , 2013, , .		6
1847	SksOpen: Efficient Indexing, Querying, and Visualization of Geo-spatial Big Data. , 2013, , .		9
1848	A scalable random forest algorithm based on MapReduce. , 2013, , .		7
1849	Parallel Map Matching on Massive Vehicle GPS Data Using MapReduce. , 2013, , .		8
1850	On the Performance of Apache Hadoop in a Tiny Private IaaS Cloud. , 2013, , .		3
1851	Energy-optimized dynamic deferral of workload for capacity provisioning in data centers. , 2013, , .		4
1852	Parallel subgroup discovery on computing clusters — First results. , 2013, , .		3
1853	Using clouds for MapReduce measurement assignments. ACM Transactions on Computing Education, 2013, 13, 1-18.	2.9	6
1854	Large-scale feature matching with distributed and heterogeneous computing. , 2013, , .		2
1855	Analysis and Improvement of Makespan and Utilization for MapReduce. , 2013, , .		1
1856	Eywa: Crowdsourced and cloudsourced omniscience. , 2013, , .		1
1857	Utilizing Human Intelligence in a Crowdsourcing Marketplace for Big Data Processing. , 2013, , .		1
1858	LEIA: The Live Evidence Information Aggregator: Towards efficient cyber-law enforcement. , 2013, , .		5
1859	A Failure Recovery Solution for Transplanting High-Performance Data-Intensive Algorithms from the Cluster to the Cloud. , 2013, , .		3
1860	Dynamic Creation of BSP/CGM Clusters on Cloud Computing Platforms. , 2013, , .		1
1861	RUBA: Real-time unstructured big data analysis framework. , 2013, , .		3
1862	An Iterative Divide-and-Merge-Based Approach for Solving Large-Scale Least Squares Problems. IEEE Transactions on Parallel and Distributed Systems, 2013, 24, 428-438.	4.0	6

#	ARTICLE	IF	CITATIONS
1863	vGreenNet: Managing Server and Networking Resources of Co-Located Heterogeneous VMs. , 2013, , .		0
1864	Minimizing Remote Accesses in MapReduce Clusters. , 2013, , .		6
1865	Optimizing data access latencies in cloud systems by intelligent virtual machine placement. , 2013, , .		87
1866	RPC automation: Making legacy code relevant. , 2013, , .		1
1867	A lightweight approach to simulate a 2D radar coverage for virtual maritime environments. , 2013, , .		1
1868	Distributed storage and index of vector spatial data based on HBase. , 2013, , .		14
1869	A cloud computing based architecture for cyber security situation awareness. , 2013, , .		41
1870	HadoopCL: MapReduce on Distributed Heterogeneous Platforms through Seamless Integration of Hadoop and OpenCL. , 2013, , .		33
1871	On Access Control Schemes for Hadoop Data Storage. , 2013, , .		2
1872	Efficient Similarity Join for Time Sequences Using Locality Sensitive Hash and Mapreduce. , 2013, , .		2
1873	Design and Optimization of a Big Data Computing Framework Based on CPU/GPU Cluster. , 2013, , .		3
1874	Interval indexing and querying on key-value cloud stores. , 2013, , .		16
1875	Rake: Semantics Assisted Network-Based Tracing Framework. IEEE Transactions on Network and Service Management, 2013, 10, 3-14.	3.2	2
1876	VegalIndexer: A Distributed composite index scheme for big spatio-temporal sensor data on cloud. , 2013, , .		10
1877	An Energy-Aware Optimization Model Based on Data Placement and Task Scheduling. , 2013, , .		0
1878	A novel algorithm for online exact string matching. , 2013, , .		0
1879	Speeding Up Batch Alignment of Large Ontologies Using MapReduce. , 2013, 2013, 110-113.		5
1880	An improved referrer-based session identification algorithm using MapReduce. , 2013, , .		4

#	ARTICLE	IF	CITATIONS
1881	An Unsupervised Algorithm for Learning Blocking Schemes. , 2013, , .		32
1882	Mobile cloud computing: A survey. Future Generation Computer Systems, 2013, 29, 84-106.	4.9	1,349
1883	Energy-Efficient Virtual Machine Consolidation. IT Professional, 2013, 15, 28-34.	1.4	48
1884	The Future of Virtual Compound Screening. Chemical Biology and Drug Design, 2013, 81, 33-40.	1.5	74
1885	Adaptive atomic capture of multiple molecules. Journal of Parallel and Distributed Computing, 2013, 73, 1251-1266.	2.7	3
1886	AlmaNebula: A Computer Forensics Framework for the Cloud. Procedia Computer Science, 2013, 19, 139-146.	1.2	9
1887	Offline traffic analysis system based on Hadoop. Journal of China Universities of Posts and Telecommunications, 2013, 20, 97-103.	0.8	9
1888	Identifying and understanding urban sport areas using Nokia Sports Tracker. Pervasive and Mobile Computing, 2013, 9, 616-628.	2.1	15
1889	A MapReduce Framework for Analysing Portfolios of Catastrophic Risk with Secondary Uncertainty. Procedia Computer Science, 2013, 18, 2317-2326.	1.2	3
1890	Neighborhood Preserving Codes for Assigning Point Labels: Applications to Stochastic Search. Procedia Computer Science, 2013, 18, 956-965.	1.2	3
1891	Joint optimization of overlapping phases in MapReduce. Performance Evaluation, 2013, 70, 720-735.	0.9	51
1892	DBridges: Flexible floodless frame forwarding. Computer Networks, 2013, 57, 3601-3616.	3.2	4
1893	A study of unpredictability in fault-tolerant middleware. Computer Networks, 2013, 57, 682-698.	3.2	3
1894	Detecting near-duplicate documents using sentence-level features and supervised learning. Expert Systems With Applications, 2013, 40, 1467-1476.	4.4	21
1895	An Optimization for MapReduce Frameworks in Multi-core Architectures. Procedia Computer Science, 2013, 18, 2587-2590.	1.2	2
1896	Apache Hadoop YARN. , 2013, , .		1,314
1897	Person-of-interest detection system using cloud-supported computerized-eyewear. , 2013, , .		5
1898	Building Green Systems with Green Students: An Educational Experiment with GENI Infrastructure. , 2013, , .		3

#	ARTICLE	IF	CITATIONS
1899	A Framework for Multi-cloud Cooperation with Hardware Reconfiguration Support. , 2013, , .		2
1900	Cloud providers collaboration for a higher service level in cloud computing. , 2013, , .		4
1901	Self-Adaptive, Deadline-Aware Resource Control in Cloud Computing. , 2013, , .		6
1902	OWL reasoning over big biomedical data. , 2013, , .		6
1903	DM-Middleware: A Middleware to Enable High Performance Data Mining in Heterogeneous Cloud. , 2013, , .		0
1904	WHEN CLOUD COMPUTING MEETS BIOINFORMATICS: A REVIEW. Journal of Bioinformatics and Computational Biology, 2013, 11, 1330002.	0.3	13
1905	Complex statistical analysis of big data: Implementation and application of Apriori and FP-Growth algorithm based on MapReduce. , 2013, , .		10
1906	CPI ² . , 2013, , .		206
1907	Compilers for Low Power with Design Patterns on Embedded Multicore Systems. , 2013, , .		1
1908	A distributed tree data structure for real-time OLAP on cloud architectures. , 2013, , .		16
1909	Parallel Matrix Multiplication Algorithm Based on Vector Linear Combination Using MapReduce. , 2013, , .		6
1910	Assessment of machine learning algorithms in cloud computing frameworks. , 2013, , .		15
1911	Revisiting flow-based load balancing: Stateless path selection in data center networks. Computer Networks, 2013, 57, 1204-1216.	3.2	14
1912	A self-adapted method for the categorization of social resources. Expert Systems With Applications, 2013, 40, 3696-3714.	4.4	3
1913	Generating storylines from sensor data. Pervasive and Mobile Computing, 2013, 9, 838-847.	2.1	7
1914	Producer-Consumer: The Programming Model for Future Many-Core Processors. Lecture Notes in Computer Science, 2013, , 110-121.	1.0	2
1915	Performance comparison under failures of MPI and MapReduce: An analytical approach. Future Generation Computer Systems, 2013, 29, 1808-1815.	4.9	10
1916	Scalable parallel computing on clouds using Twister4Azure iterative MapReduce. Future Generation Computer Systems, 2013, 29, 1035-1048.	4.9	46

#	ARTICLE	IF	CITATIONS
1917	NoSQL databases: a step to database scalability in web environment. International Journal of Web Information Systems, 2013, 9, 69-82.	1.3	165
1918	Comparison of MapReduce and SQL on large-scale data processing. Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers, Series A/Chung-kuo Kung Ch'eng Hsueh K'an, 2013, 36, 27-34.	0.6	2
1919	GreX: An efficient MapReduce framework for graphics processing units. Journal of Parallel and Distributed Computing, 2013, 73, 522-533.	2.7	28
1920	Towards an Integrated Platform for Big Data Analysis. , 2013, , 47-56.		17
1921	Introduction to Information Visualisation. Lecture Notes in Computer Science, 2013, , 1-27.	1.0	7
1922	An ontology enhanced parallel SVM for scalable spam filter training. Neurocomputing, 2013, 108, 45-57.	3.5	37
1923	Behavior-based tracking: Exploiting characteristic patterns in DNS traffic. Computers and Security, 2013, 39, 17-33.	4.0	56
1924	Dynamic Task Scheduling in Cloud Computing Based on Greedy Strategy. Communications in Computer and Information Science, 2013, , 156-162.	0.4	11
1925	Algebraic Optimization for Processing Graph Pattern Queries in the Cloud. IEEE Internet Computing, 2013, 17, 52-61.	3.2	9
1926	On the performance of high dimensional data clustering and classification algorithms. Future Generation Computer Systems, 2013, 29, 1024-1034.	4.9	39
1927	Impacts of Virtualization Technologies on Hadoop. , 2013, , .		10
1928	An Overview on the Structure and Applications for Business Intelligence and Data Mining in Cloud Computing. Advances in Intelligent Systems and Computing, 2013, , 559-570.	0.5	4
1930	An efficient quasi-identifier index based approach for privacy preservation over incremental data sets on cloud. Journal of Computer and System Sciences, 2013, 79, 542-555.	0.9	68
1931	Real-Time Mental Arithmetic Task Recognition From EEG Signals. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2013, 21, 225-232.	2.7	150
1932	Analytical Performance Models for MapReduce Workloads. International Journal of Parallel Programming, 2013, 41, 495-525.	1.1	50
1933	A data infrastructure reference model with applications: towards realization of a ScienceTube vision with a data replication service. Journal of Internet Services and Applications, 2013, 4, 1.	1.6	52
1934	Managing node disappearance based on information flow in social networks. Social Network Analysis and Mining, 2013, 3, 299-311.	1.9	8
1935	Scalable community detection in massive social networks using MapReduce. IBM Journal of Research and Development, 2013, 57, 12:1-12:14.	3.2	14

#	ARTICLE	IF	CITATIONS
1936	Optical Interconnects in Next Generation Data Centers: An End to End View. Optical Networks Series, 2013, , 31-46.	1.1	1
1937	On Architecture Design, Congestion Notification, TCP Incast and Power Consumption in Data Centers. IEEE Communications Surveys and Tutorials, 2013, 15, 39-64.	24.8	127
1938	Security and Privacy in Cloud Computing. IEEE Communications Surveys and Tutorials, 2013, 15, 843-859.	24.8	425
1939	Moving code â€“ Sharing geoprocessing logic on the Web. ISPRS Journal of Photogrammetry and Remote Sensing, 2013, 83, 193-203.	4.9	20
1940	Scaling up Data Mining Techniques to Large Datasets Using Parallel and Distributed Processing. Advanced Information and Knowledge Processing, 2013, , 243-259.	0.2	7
1941	MGMR: Multi-GPU Based MapReduce. Lecture Notes in Computer Science, 2013, , 433-442.	1.0	16
1942	A Data Management and Analysis System in Healthcare Cloud. , 2013, , .		2
1943	HSim: A MapReduce simulator in enabling Cloud Computing. Future Generation Computer Systems, 2013, 29, 300-308.	4.9	54
1944	Design and implementation of GXP make â€” A workflow system based on make. Future Generation Computer Systems, 2013, 29, 662-672.	4.9	11
1945	Internet-based Virtual Computing Environment: Beyond the data center as a computer. Future Generation Computer Systems, 2013, 29, 309-322.	4.9	74
1946	SubSift web services and workflows for profiling and comparing scientists and their published works. Future Generation Computer Systems, 2013, 29, 569-581.	4.9	4
1947	Towards a virtual research environment for language and literature researchers. Future Generation Computer Systems, 2013, 29, 549-559.	4.9	7
1948	G-Hadoop: MapReduce across distributed data centers for data-intensive computing. Future Generation Computer Systems, 2013, 29, 739-750.	4.9	292
1949	Monte Carlo simulation on heterogeneous distributed systems: A computing framework with parallel merging and checkpointing strategies. Future Generation Computer Systems, 2013, 29, 728-738.	4.9	26
1950	Efficient and Scalable Processing of String Similarity Join. IEEE Transactions on Knowledge and Data Engineering, 2013, 25, 2217-2230.	4.0	46
1951	Sketch-Based Retrieval of Document Illustrations and Regions of Interest. , 2013, , .		5
1952	Exploiting compression and approximation paradigms for effective and efficient online analytical processing over sensor network readings in data grid environments. Concurrency Computation Practice and Experience, 2013, 25, 2016-2035.	1.4	17
1953	Map-based graph analysis on MapReduce. , 2013, , .		4

#	ARTICLE	IF	CITATIONS
1954	MyCloud: On-Demand Virtual Cluster Provisioning on HPC Resources. , 2013, , .		2
1955	Smart Task Distributor for MapReduce on Cloud Computing. , 2013, , .		0
1956	Orchestrating an Ensemble of MapReduce Jobs for Minimizing Their Makespan. IEEE Transactions on Dependable and Secure Computing, 2013, 10, 314-327.	3.7	56
1957	A Queueing Model to Achieve Proper Elasticity for Cloud Cluster Jobs. , 2013, , .		37
1958	Approximate triangle counting algorithms on multi-cores. , 2013, , .		30
1959	Breaking the boundary for whole-system performance optimization of big data. , 2013, , .		5
1960	Trusted Sampling-Based Result Verification on Mass Data Processing. , 2013, , .		3
1961	Point representation for local optimization. , 2013, , .		0
1962	A Delay Scheduling Algorithm Based on History Time in Heterogeneous Environments. , 2013, , .		5
1963	Traffic measurement and analysis in an organic enterprise data center. , 2013, , .		7
1964	An Automatic Tag Recommendation Algorithm for Micro-blogging Users. , 2013, , .		3
1965	Architecture design of mobile cloud and prototype test. , 2013, , .		1
1966	Overlapping correlation clustering. Knowledge and Information Systems, 2013, 35, 1-32.	2.1	36
1967	WHAD: Wikipedia historical attributes data. Language Resources and Evaluation, 2013, 47, 1163-1190.	1.8	13
1968	An approach for log analysis based failure monitoring in Hadoop cluster. , 2013, , .		3
1969	Vehicle-sharing: Technological infrastructure, vehicles, and user-side devices - Technological review. , 2013, , .		7
1970	PVTCP: Towards practical and effective congestion control in virtualized datacenters. , 2013, , .		13
1971	Distance-aware virtual cluster performance optimization: A hadoop case study. , 2013, , .		3

#	ARTICLE	IF	CITATIONS
1972	Algorithms for high-throughput disk-to-disk sorting. , 2013, , .		8
1973	Presto. , 2013, , .		57
1974	Comparative performance analysis of a Big Data NORA problem on a variety of architectures. , 2013, , .		6
1975	Mining Social Influence in Microblogging via Tensor Factorization Approach. , 2013, , .		2
1976	Massive distributed and parallel log analysis for organizational security. , 2013, , .		8
1977	How to Use Google App Engine for Free Computing. IEEE Internet Computing, 2013, 17, 50-59.	3.2	24
1978	Construct rough approximation based on GAE. , 2013, , .		1
1979	Modeling Cognitive Radio Networks for Efficient Data Transfer Using Cloud Link. , 2013, , .		6
1980	Big data platform development with a domain specific language for telecom industries. , 2013, , .		7
1981	Scalable sentiment classification for Big Data analysis using Naïve Bayes Classifier. , 2013, , .		108
1982	A Measurement Study of Data-Intensive Network Traffic Patterns in a Private Cloud. , 2013, , .		7
1983	Omega. , 2013, , .		470
1984	Nearest group queries. , 2013, , .		20
1985	Job Classification for MapReduce Scheduler in Heterogeneous Environment. , 2013, , .		4
1986	Strategic Alignment of Cloud-Based Architectures for Big Data. , 2013, , .		37
1987	Opportunities of innovation under challenges of big data. , 2013, , .		9
1988	Attribute Relationship Evaluation Methodology for Big Data Security. , 2013, , .		37
1989	SeaRum: A Cloud-Based Service for Association Rule Mining. , 2013, , .		13

#	ARTICLE	IF	CITATIONS
1990	Event-Driven Fault Tolerance for Building Nonstop Active Message Programs. , 2013, , .		2
1991	Bandwidth-aware data placement scheme for Hadoop. , 2013, , .		12
1992	Error Tolerant Multimedia Stream Processing: There's Plenty of Room at the Top (of the System Stack). IEEE Transactions on Multimedia, 2013, 15, 291-303.	5.2	8
1993	A MapReduce Computing Framework Based on GPU Cluster. , 2013, , .		2
1994	Milieu: Lightweight and Configurable Big Data Provenance for Science. , 2013, , .		21
1995	Towards a Quality-centric Big Data Architecture for Federated Sensor Services. , 2013, , .		49
1996	SAGE: Geo-Distributed Streaming Data Analysis in Clouds. , 2013, , .		10
1997	Evaluating Cloud Computing Techniques for Smart Power Grid Design Using Parallel Scripting. , 2013, , .		19
1998	Solving the NP-hard computational problem in Bayesian networks using apache hadoop MapReduce. , 2013, , .		0
1999	On Efficiently Capturing Scientific Properties in Distributed Big Data without Moving the Data: A Case Study in Distributed Structural Biology Using MapReduce. , 2013, , .		6
2000	A Language Based Security Approach for Securing Map-Reduce Computations in the Cloud. , 2013, , .		2
2001	NEWT - A Fault Tolerant BSP Framework on Hadoop YARN. , 2013, , .		1
2002	Evaluating parallel logistic regression models. , 2013, , .		14
2003	The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, Second edition. Synthesis Lectures on Computer Architecture, 2013, 8, 1-154.	1.3	356
2004	HOSTED SCIENCE: MANAGING COMPUTATIONAL WORKFLOWS IN THE CLOUD. Parallel Processing Letters, 2013, 23, 1340004.	0.4	10
2005	Optimal Client-Server Assignment for Internet Distributed Systems. IEEE Transactions on Parallel and Distributed Systems, 2013, 24, 565-575.	4.0	17
2006	A Sketch of Big Data Technologies. , 2013, , .		10
2007	A federated approach on heterogeneous NoSQL data stores. , 2013, , .		16

#	ARTICLE	IF	CITATIONS
2008	Performance analysis of current parallel programming models for many-core systems. , 2013, , .		0
2009	MREclat: An Algorithm for Parallel Mining Frequent Itemsets. , 2013, , .		19
2010	Computation Certification as a Service in the Cloud. , 2013, , .		3
2011	Sketching the data center network traffic. IEEE Network, 2013, 27, 33-39.	4.9	16
2012	Research on Improved A Priori Algorithm Based on Coding and MapReduce. , 2013, , .		13
2013	Distributed MapReduce framework using distributed hash table. , 2013, , .		6
2014	MillWheel. Proceedings of the VLDB Endowment, 2013, 6, 1033-1044.	2.1	409
2015	Big Data Infrastructure for Active Situation Awareness on Social Network Services. , 2013, , .		4
2016	Scheduling associative reductions with homogeneous costs when overlapping communications and computations. , 2013, , .		2
2017	Optimization strategies for A/B testing on HADOOP. Proceedings of the VLDB Endowment, 2013, 6, 973-984.	2.1	3
2018	Towards Dependability Testing of MapReduce Systems. , 2013, , .		1
2019	Modeling and querying data in NoSQL databases. , 2013, , .		67
2020	Security Event Correlation Supported by Multi-Core Architecture. , 2013, , .		4
2021	Forecasting consumer behavior with innovative value proposition for organizations using big data analytics. , 2013, , .		15
2022	A parallel Probabilistic Latent Semantic Analysis method on MapReduce platform. , 2013, , .		6
2023	A Higher-order data flow model for heterogeneous Big Data. , 2013, , .		3
2024	ZCluster: A Zynq-based Hadoop cluster. , 2013, , .		35
2025	A case for MapReduce over the internet. , 2013, , .		2

#	ARTICLE	IF	CITATIONS
2026	The family of mapreduce and large-scale data processing systems. ACM Computing Surveys, 2013, 46, 1-44.	16.1	127
2027	Prefix filtering with data partitioning for similarity join. , 2013, , .		0
2028	Cloud-Based Big Data Mining & Analyzing Services Platform Integrating R. , 2013, , .		16
2029	Verification Requirements for Secure and Reliable Cloud Computing. , 2013, , .		3
2030	A Survey of Visualization Pipelines. IEEE Transactions on Visualization and Computer Graphics, 2013, 19, 367-378.	2.9	57
2031	Distributed computing practice for large-scale science and engineering applications. Concurrency Computation Practice and Experience, 2013, 25, 1559-1585.	1.4	15
2032	Efficient and exact duplicate detection on cloud. Concurrency Computation Practice and Experience, 2013, 25, 2187-2206.	1.4	2
2033	High performance parallel evolutionary algorithm model based on MapReduce framework. International Journal of Computer Applications in Technology, 2013, 46, 290.	0.3	17
2034	An architecture for efficiently establishing the classroom cloud. , 2013, , .		2
2035	Scorpion. Proceedings of the VLDB Endowment, 2013, 6, 553-564.	2.1	154
2036	Data management in cloud environments: NoSQL and NewSQL data stores. Journal of Cloud Computing: Advances, Systems and Applications, 2013, 2, .	2.1	183
2037	An ad-hoc distributed execution environment for multi-agent systems. , 2013, , .		2
2038	Towards reproducible performance studies of datacenter network architectures using an open-source simulation approach. , 2013, , .		13
2039	Parameterised architectural patterns for providing cloud service fault tolerance with accurate costings. , 2013, , .		7
2040	Towards an Expressive and Scalable Twitter's Users Profiles. , 2013, , .		3
2041	Why high performance visual data analytics is both relevant and difficult. Proceedings of SPIE, 2013, , .	0.8	1
2042	Petapixel photography and the limits of camera information capacity. , 2013, , .		1
2043	Distributed Online Test Generation for Model-Based Testing. , 2013, , .		3

#	ARTICLE	IF	CITATIONS
2044	Efficient multicast delivery for wireless data center networks. , 2013, , .		1
2045	Toward lightweight transparent data middleware in support of document stores. , 2013, , .		7
2046	Tracking international migration from RFID data using map reduce method. , 2013, , .		2
2047	Exploring GPU architectures to accelerate semantic comparison for intention-based search. , 2013, , .		1
2048	Audience segment expansion using distributed in-database k-means clustering. , 2013, , .		4
2049	Design of an active storage cluster file system for DAG workflows. , 2013, , .		2
2050	Designing a database system for modern processing architectures. , 2013, , .		0
2051	Efficient social network data query processing on MapReduce. , 2013, , .		7
2052	REEF. Proceedings of the VLDB Endowment, 2013, 6, 1370-1373.	2.1	8
2053	Eagle-eyed elephant. , 2013, , .		32
2054	EBM. , 2013, , .		112
2055	Incremental stream processing using computational conflict-free replicated data types. , 2013, , .		6
2056	Speeding up model building for ECGA on CUDA platform. , 2013, , .		2
2058	Scalable and incremental software bug detection. , 2013, , .		34
2059	MRPacker. , 2013, , .		1
2060	Distributed community detection in web-scale networks. , 2013, , .		17
2061	A novel KNN join algorithms based on Hilbert R-tree in MapReduce. , 2013, , .		5
2062	Towards privacy-preserving computing on distributed electronic health record data. , 2013, , .		7

#	ARTICLE	IF	CITATIONS
2063	RadixVM. , 2013, , .		61
2064	Evaluating MapReduce for profiling application traffic. , 2013, , .		4
2065	Does RDMA-based enhanced Hadoop MapReduce need a new performance model?. , 2013, , .		1
2066	Data Processing on FPGAs. Synthesis Lectures on Data Management, 2013, 5, 1-118.	0.6	23
2067	Natjam. , 2013, , .		53
2068	GRAPHICAL INDUCTION OF QUALIFIED MEDICAL KNOWLEDGE. International Journal of Semantic Computing, 2013, 07, 377-405.	0.4	2
2069	Benchmarking approach for designing a mapreduce performance model. , 2013, , .		31
2070	Octopus. , 2013, , .		0
2071	Channel reservation protocol for over-subscribed channels and destinations. , 2013, , .		19
2072	River trail. , 2013, , .		25
2073	Matching bounds for the all-pairs MapReduce problem. , 2013, , .		8
2074	MapReduce performance evaluation for knowledge-based recommendation of context-tagged photos. , 2013, , .		3
2075	BDMPI. , 2013, , .		1
2076	A platform for eXtreme Analytics. IBM Journal of Research and Development, 2013, 57, 4:1-4:11.	3.2	8
2077	A Large-scale Images Processing Model Based on Hadoop Platform. , 2013, , .		4
2078	A Framework of the Civil Engineering CAD Experimental Platform Based on Cloud Computing. , 2013, , .		3
2079	Wavelet-based multi-scale anomaly identification in cloud computing systems. , 2013, , .		4
2080	Automatic failure recovery for software-defined networks. , 2013, , .		28

#	ARTICLE	IF	CITATIONS
2081	Towards systematic parallel programming of graph problems via tree decomposition and tree parallelism. , 2013, , .		0
2082	XDB. , 2013, , .		2
2083	CRUCIBLE. , 2013, , .		1
2084	Distributed Risk Aversion Parameter Estimation for First-Price Auction in Sensor Networks. International Journal of Distributed Sensor Networks, 2013, 9, 795630.	1.3	0
2085	Efficient query evaluation on distributed graphs with Hadoop environment. , 2013, , .		12
2086	Parallel graph processing on graphics processors made easy. Proceedings of the VLDB Endowment, 2013, 6, 1270-1273.	2.1	12
2087	Actor scheduling for multicore hierarchical memory platforms. , 2013, , .		7
2088	MRM. , 2013, , .		0
2089	Monte Carlo Based Personalized PageRank on Dynamic Networks. International Journal of Distributed Sensor Networks, 2013, 9, 829804.	1.3	3
2090	X10-FT. , 2013, , .		8
2091	A more formal approach to "computer science: principles". , 2013, , .		1
2092	Locating executable fragments with Concordia, a scalable, semantics-based architecture. , 2013, , .		1
2093	Component-based scalability for cloud applications. , 2013, , .		17
2094	Building an on-demand virtual computing market in non-commercial communities. , 2013, , .		2
2095	Input data organization for batch processing in time window based computations. , 2013, , .		1
2096	Large-scale bisimulation of RDF graphs. , 2013, , .		18
2097	Participatory networking. , 2013, , .		172
2098	Parallel stepwise stochastic simulation. , 2013, , .		0

#	ARTICLE	IF	CITATIONS
2099	Adaptive online scheduling in storm. , 2013, , .		178
2100	Direct out-of-memory distributed parallel frequent pattern mining. , 2013, , .		2
2101	PATRIC. , 2013, , .		85
2102	Enforcing Minimum Necessary Access in Healthcare Through Integrated Audit and Access Control. , 2013, , .		6
2103	HJ-Hadoop. , 2013, , .		6
2104	Large-scale exploratory analysis, cleaning, and modeling for event detection in real-world power systems data. , 2013, , .		0
2105	Distributed and Parallel Big Textual Data Parsing for Social Sensor Network. International Journal of Distributed Sensor Networks, 2013, 9, 525687.	1.3	4
2106	TigerQuoll. , 2013, , .		2
2107	Sparkler. , 2013, , .		24
2108	A data-centric heuristic for Hadoop provisioning in the cloud. , 2013, , .		0
2109	Minimizing data transfers for regular reachability queries on distributed graphs. , 2013, , .		4
2110	On implementing provenance-aware regular path queries with relational query engines. , 2013, , .		24
2111	Prolog programming with a map-reduce parallel construct. , 2013, , .		3
2112	Modeling I/O interference for data intensive distributed applications. , 2013, , .		12
2113	Performance Modeling and Optimization of Deadline-Driven Pig Programs. ACM Transactions on Autonomous and Adaptive Systems, 2013, 8, 1-28.	0.4	11
2114	MR-runner. , 2013, , .		0
2115	Sharing data and work across concurrent analytical queries. Proceedings of the VLDB Endowment, 2013, 6, 637-648.	2.1	25
2116	Instant loading for main memory databases. Proceedings of the VLDB Endowment, 2013, 6, 1702-1713.	2.1	51

#	ARTICLE	IF	CITATIONS
2117	Efficient query processing in distributed search engines. ACM SIGIR Forum, 2013, 47, 60-61.	0.4	0
2118	News vertical search using user-generated content. ACM SIGIR Forum, 2013, 47, 62-63.	0.4	1
2119	TigerQuoll. ACM SIGPLAN Notices, 2013, 48, 251-260.	0.2	3
2120	A self-updating road map of The Cancer Genome Atlas. Bioinformatics, 2013, 29, 1333-1340.	1.8	35
2121	An Abstract Description Method of Map-Reduce-Merge Using Haskell. Mathematical Problems in Engineering, 2013, 2013, 1-12.	0.6	3
2122	Translational Biomedical Informatics in the Cloud: Present and Future. BioMed Research International, 2013, 2013, 1-8.	0.9	45
2123	One-to-One Disjoint Path Covers in DCell. Lecture Notes in Computer Science, 2013, , 61-70.	1.0	5
2124	Verification and validation of Parallel Support Vector Machine algorithm based on MapReduce Program model on Hadoop cluster. , 2013, , .		8
2125	Optimization of cloud task processing with checkpoint-restart mechanism. , 2013, , .		55
2126	Parallel Execution of ATL Transformation Rules. Lecture Notes in Computer Science, 2013, , 656-672.	1.0	19
2127	Banking on decoupling. Operating Systems Review (ACM), 2013, 47, 41-50.	1.5	4
2128	Data Mining with BP Neural Network Algorithm Based MapReduce. Applied Mechanics and Materials, 2013, 380-384, 2915-2919.	0.2	3
2129	Evaluation of Parallel Algorithms of Traffic Volume Statistics per Cell in Cellular Network Based on MapReduce Programming Model. Applied Mechanics and Materials, 0, 401-403, 1859-1863.	0.2	0
2130	Building Quotient Cube with MapReduce in Hadoop. Advanced Materials Research, 2013, 765-767, 1031-1035.	0.3	1
2131	Hone. Proceedings of the VLDB Endowment, 2013, 6, 1354-1357.	2.1	15
2132	Enabling Large-Scale Biomedical Analysis in the Cloud. BioMed Research International, 2013, 2013, 1-6.	0.9	15
2133	Analyzing job completion reliability and job energy consumption for a general MapReduce infrastructure. Journal of High Speed Networks, 2013, 19, 203-214.	0.6	9
2134	An Improved Cop-Kmeans Clustering for Solving Constraint Violation Based on MapReduce Framework. Fundamenta Informaticae, 2013, 126, 301-318.	0.3	17

#	ARTICLE	IF	CITATIONS
2135	Turbine: A Distributed-memory Dataflow Engine for High Performance Many-task Applications. Fundamenta Informaticae, 2013, 128, 337-366.	0.3	23
2136	A Distributed Algorithm for Skyline Query Based on Pre-Clustering. Advanced Materials Research, 2013, 756-759, 3982-3986.	0.3	1
2137	The Research and Application of MapReduce Based Neighbor Model in Personalized Recommendation. Applied Mechanics and Materials, 0, 373-375, 1674-1677.	0.2	0
2138	Stream-monitoring with blockmon. Computer Communication Review, 2013, 43, 29-36.	1.5	16
2139	Approaches to Distributed Execution of Scientific Workflows in Kepler. Fundamenta Informaticae, 2013, 128, 281-302.	0.3	13
2140	A Research on MapReduce-Based Redundancy Pruning for Record Linkage. Advanced Materials Research, 0, 753-755, 3009-3013.	0.3	0
2141	An Improved MapReduce Model for Computation-Intensive Task. Advanced Materials Research, 2013, 756-759, 1701-1705.	0.3	0
2142	Research on Cloud Platform for Wind Turbine Fault Diagnosis. Advanced Materials Research, 2013, 765-767, 2255-2258.	0.3	1
2143	Making queries tractable on big data with preprocessing. Proceedings of the VLDB Endowment, 2013, 6, 685-696.	2.1	22
2144	Design and Implementation of Massive Data Retrieving Based on Cloud Computing Platform. Applied Mechanics and Materials, 0, 303-306, 2235-2240.	0.2	1
2145	CG_Hadoop. , 2013, , .		81
2146	Optimus. , 2013, , .		25
2147	PESC: A parallel system for clustering ECG streams based on MapReduce. , 2013, , .		1
2148	An Improved Association Rules Mining Algorithm Based on Power Set and Hadoop. , 2013, , .		3
2149	Incentives for Dynamic and Energy-Aware Capacity Allocation for Multi-tenant Clusters. Lecture Notes in Computer Science, 2013, , 106-121.	1.0	1
2150	Searching for truss alpha users in mobile telecommunications social networks. , 2013, , .		0
2151	Processing multi-way spatial joins on map-reduce. , 2013, , .		29
2152	Product Documentation Management Through REST-Based Web Service. Intelligent Systems, Control and Automation: Science and Engineering, 2013, , 179-189.	0.3	0

#	ARTICLE	IF	CITATIONS
2153	A MRACO Algorithm for Structural Multi-Damage Detection. Applied Mechanics and Materials, 0, 405-408, 2443-2447.	0.2	0
2154	The Research of Cloud Platform Based on KNN Collaborative Filtering Algorithm in Fruits Marketing System. Applied Mechanics and Materials, 2013, 373-375, 1826-1829.	0.2	0
2155	Graph Calculus: Scalable Shortest Path Analytics for Large Social Graphs through Core Net. , 2013, , .		3
2156	Design of Distributed Parallel Computing Using by MapReduce/MPI Technology. Lecture Notes in Computer Science, 2013, , 139-148.	1.0	1
2157	The Application of Cloud Based on Latent Factor Algorithm in Personalized Recommendation. Applied Mechanics and Materials, 2013, 411-414, 2288-2291.	0.2	0
2158	Study on Cloud Computing Clusters Based on Coordination Method. Applied Mechanics and Materials, 2013, 411-414, 182-185.	0.2	0
2159	Parallel computation of skyline and reverse skyline queries using mapreduce. Proceedings of the VLDB Endowment, 2013, 6, 2002-2013.	2.1	80
2160	Near Duplicated Text Detection Based on MapReduce. Applied Mechanics and Materials, 2013, 427-429, 2618-2621.	0.2	1
2161	Multilevel Active Storage for big data applications in high performance computing. , 2013, , .		3
2162	Integrating Online Compression to Accelerate Large-Scale Data Analytics Applications. , 2013, , .		42
2163	An Approach for Characterizing Workloads in Google Cloud to Derive Realistic Resource Utilization Models. , 2013, , .		82
2164	Cumulon. , 2013, , .		55
2165	Algebraic dataflows for big data analysis. , 2013, , .		13
2166	Exploring Time and Frequency Domains for Accurate and Automated Anomaly Detection in Cloud Computing Systems. , 2013, , .		14
2167	Divide-and-Conquer Anchoring for Near-Separable Nonnegative Matrix Factorization and Completion in High Dimensions. , 2013, , .		9
2168	Memorization of Materialization Points. , 2013, , .		0
2169	Achieving elasticity for cloud MapReduce jobs. , 2013, , .		9
2170	Highly Scalable Sequential Pattern Mining Based on MapReduce Model on the Cloud. , 2013, , .		25

#	ARTICLE	IF	CITATIONS
2171	Fast Quasi-biclique Mining with Giraph. , 2013, , .		4
2172	Scaling Genetic Programming for data classification using MapReduce methodology. , 2013, , .		6
2173	Smart Intermediate Data Transfer for MapReduce on Cloud Computing. , 2013, , .		0
2174	Labeling Instances in Evolving Data Streams with MapReduce. , 2013, , .		10
2175	Building an Elastic Cloud out of Small Datacenters. , 2013, , .		4
2176	Distributed k-Core Decomposition. IEEE Transactions on Parallel and Distributed Systems, 2013, 24, 288-300.	4.0	122
2177	Logic-in-memory based big-data computing by nonvolatile domain-wall nanowire devices. , 2013, , .		5
2178	Scalable Data Center Network with distributed placement of component sets in optical networks. , 2013, , .		2
2179	HybridMR: A Hierarchical MapReduce Scheduler for Hybrid Data Centers. , 2013, , .		52
2180	A recommendation model based on collaborative filtering and factorization machines for social networks. , 2013, , .		2
2181	Improving the Network Energy Efficiency in MapReduce Systems. , 2013, , .		2
2182	Study on QC-Tree with MapReduce and Hbase in Hadoop. , 2013, , .		1
2183	AzureBOT: A Framework for Bag-of-Tasks Applications on the Azure Cloud Platform. , 2013, , .		11
2184	When big data meets big smog. , 2013, , .		9
2185	Elastic framework for augmenting the performance of mobile applications using cloud computing. , 2013, , .		18
2186	Identifying abnormal patterns in cellular communication flows. , 2013, , .		2
2187	Predicting Ship Behavior Navigating through Heavily Trafficked Fairways by Analyzing AIS Data on Apache HBase. , 2013, , .		18
2188	Using MapReduce for High Energy Physics Data Analysis. , 2013, , .		5

#	ARTICLE	IF	CITATIONS
2189	Robust graph traversal: Resiliency techniques for data intensive supercomputing. , 2013, , .		2
2190	Risk adjustment of patient expenditures: A big data analytics approach. , 2013, , .		28
2191	An Indexing Network Model for Information Services and Its Applications. , 2013, , .		2
2192	A NUMA-Aware Runtime Environment for the Actor Model. , 2013, , .		3
2193	MapReduce based log file analysis for system threats and problem identification. , 2013, , .		6
2194	Deriving Job Completion Reliability and Job Energy Consumption for a General MapReduce Infrastructure from Single-Job Perspective. , 2013, , .		3
2195	A framework of spatial co-location mining on MapReduce. , 2013, , .		2
2196	Toward an Ecosystem for Precision Sharing of Segmented Big Data. , 2013, , .		8
2197	A distributed clustering method to segment micro-blog users on cloud environments. , 2013, , .		1
2198	P-Mine: Parallel itemset mining on large datasets. , 2013, , .		12
2199	A Method of Data Distribution for Distributed Cross Join. , 2013, , .		0
2200	Bisimulation Reduction of Big Graphs on MapReduce. Lecture Notes in Computer Science, 2013, , 189-203.	1.0	12
2201	MapReduce Based Frameworks for Classifying Evolving Data Stream. , 2013, , .		0
2202	Evaluating data storage structures of MapReduce. , 2013, , .		1
2203	Data chaos: An entropy based MapReduce framework for scalable learning. , 2013, , .		2
2204	Experimental Study on the Energy Consumption in IaaS Cloud Environments. , 2013, , .		2
2205	Theoretical Analysis of Wavelet Synopsis on Partitioned Data Sets. , 2013, , .		0
2206	An efficient MapReduce algorithm for counting triangles in a very large graph. , 2013, , .		66

#	ARTICLE	IF	CITATIONS
2207	Research and Implementation of MapReduce Programming Oriented Graphical Modeling System. , 2013, , .		2
2208	Moim: A Multi-GPU MapReduce Framework. , 2013, , .		4
2209	Scalable, incremental learning with MapReduce parallelization for cell detection in high-resolution 3D microscopy data. , 2013, , .		3
2210	Rethinking data management for big data scientific workflows. , 2013, , .		25
2211	Elastic stream processing in the Cloud. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2013, 3, 333-345.	4.6	34
2212	PER-MARE: Adaptive Deployment of MapReduce over Pervasive Grids. , 2013, , .		8
2213	Optimizing a MapReduce module of preprocessing high-throughput DNA sequencing data. , 2013, , .		1
2214	A Distributed Real-Time Storage Method for Stream Data. , 2013, , .		2
2215	Scheduling Divisible Loads from Multiple Input Sources in MapReduce. , 2013, , .		0
2216	NativeTask: A Hadoop compatible framework for high performance. , 2013, , .		9
2217	Massive image data management using HBase and MapReduce. , 2013, , .		12
2218	Designing Hybrid Architectures for Massive-Scale Graph Analysis. , 2013, , .		3
2219	Distributed storage evaluation on a three-wide inter-data center deployment. , 2013, , .		5
2220	IntegrityMR: Integrity assurance framework for big data analytics and management applications. , 2013, , .		14
2221	A performance evaluation of Hive for scientific data management. , 2013, , .		7
2222	Load Balancing in Heterogeneous MapReduce Environments. , 2013, , .		2
2223	Enabling an Enhanced Data-as-a-Service Ecosystem. , 2013, , .		2
2224	Flexible Computing for personal electronic devices. , 2013, , .		8

#	ARTICLE	IF	CITATIONS
2225	Feliss: Flexible distributed computing framework with light-weight checkpointing. , 2013, , .		1
2226	Diagnosis Rule Mining of Airborne Avionics Using Formal Concept Analysis. , 2013, , .		1
2227	Risk Intelligence: Profiting from Uncertainty in Data Processing System. , 2013, , .		0
2228	Briareus: Accelerating Python Applications with Cloud. , 2013, , .		0
2229	CloudRS: An error correction algorithm of high-throughput sequencing data based on scalable framework. , 2013, , .		19
2230	Scalable power grid transient analysis via MOR-assisted time-domain simulations. , 2013, , .		5
2231	Design and implementation of a massively parallel ESB. , 2013, , .		1
2232	TimeStream. , 2013, , .		169
2233	Toward scalable and high performance data processing for Cellular Data Network. , 2013, , .		0
2234	A Multi-GPU Framework for In-Memory Text Data Analytics. , 2013, , .		3
2235	VOLUME: Enable Large-Scale In-Memory Computation on Commodity Clusters. , 2013, , .		2
2236	A Real Time Index Model for Big Data Based on DC-Tree. , 2013, , .		3
2237	Scalable context-aware role mining with MapReduce. , 2013, , .		0
2238	Clustering remote RDF data using SPARQL update queries. , 2013, , .		5
2239	The state of peer-to-peer network simulators. ACM Computing Surveys, 2013, 45, 1-25.	16.1	16
2240	Investigating the Feasibility of Making Contexts Explicit in Designing Cloud Workflow. , 2013, , .		3
2241	UCMS: User-side cloud management system. , 2013, , .		5
2242	Leveraging Performance Counters and Execution Logs to Diagnose Memory-Related Performance Issues. , 2013, , .		39

#	ARTICLE	IF	CITATIONS
2243	A Preference-Aware Service Recommendation Method on Map-Reduce. , 2013, , .		4
2244	A DT-SVM Strategy for Stock Futures Prediction with Big Data. , 2013, , .		23
2245	Leveraging real-world data while protecting privacy Framework using personal agent system for distributed data. , 2013, , .		0
2246	A Practical Approach for Scalable Record Linkage on Hadoop. Advanced Materials Research, 2013, 753-755, 3018-3024.	0.3	1
2247	Secure Cache Provision: Provable DDOS Prevention for Randomly Partitioned Services with Replication. , 2013, , .		0
2248	A Parallel Algorithm for Bayesian Network Parameter Learning Based on Factor Graph. , 2013, , .		2
2249	Stochastic modeling and analysis of opportunistic computing in intermittent mobile cloud. , 2013, , .		2
2250	A Compression Framework for Multidimensional Scientific Datasets. , 2013, , .		4
2251	A scalable data analysis platform for metagenomics. , 2013, , .		25
2252	On the use of shared storage in shared-nothing environments. , 2013, , .		6
2253	Platform for General-Purpose Distributed Data-Mining on Large Dynamic Graphs. , 2013, , .		1
2254	Suitable Database Development Framework for Business Component Migration in SaaS Multi-tenant Model. , 2013, , .		1
2255	Pipeline-Based Parallel Framework for Mass File Processing. , 2013, , .		1
2256	Information Centric Networks for Parallel Processing in the Datacenter. , 2013, , .		4
2257	Improvement of the Data Mining Algorithm of Rough Set under the Framework of Map/Reduce. , 2013, , .		1
2258	BigBench. , 2013, , .		258
2259	"All roads lead to Rome". , 2013, , .		40
2260	Distributed matrix factorization with mapreduce using a series of broadcast-joins. , 2013, , .		28

#	ARTICLE	IF	CITATIONS
2262	Curvilinearly T-Stiffened Panel-Optimization Framework Under Multiple Load Cases Using Parallel Processing. <i>Journal of Aircraft</i> , 2013, 50, 1540-1554.	1.7	11
2263	Simulation process support for climate data analysis. , 2013, , .		0
2264	Ghostbusters: A Parts-based NMF algorithm. , 2013, , .		8
2265	IBM Streams Processing Language: Analyzing Big Data in motion. <i>IBM Journal of Research and Development</i> , 2013, 57, 7:1-7:11.	3.2	92
2266	Design of Intelligent System for Large Scale Database. , 2013, , .		0
2267	Pipelined Multi-GPU MapReduce for Big-Data Processing. <i>Studies in Computational Intelligence</i> , 2013, , 231-246.	0.7	10
2268	Extending modern PaaS clouds with BSP to execute legacy MPI applications. , 2013, , .		1
2269	TST: Threshold based similarity transitivity method in collaborative filtering with cloud computing. <i>Tsinghua Science and Technology</i> , 2013, 18, 318-327.	4.1	17
2270	Energy and locality aware load balancing in cloud computing. <i>Integrated Computer-Aided Engineering</i> , 2013, 20, 361-374.	2.5	10
2271	The quantcast file system. <i>Proceedings of the VLDB Endowment</i> , 2013, 6, 1092-1101.	2.1	124
2272	MASSIVELY DISTRIBUTED CONCEPT DRIFT HANDLING IN LARGE NETWORKS. <i>International Journal of Modeling, Simulation, and Scientific Computing</i> , 2013, 16, 1350021.	0.9	2
2273	Order matters! Harnessing a world of orderings for reasoning over massive data. <i>Semantic Web</i> , 2013, 4, 219-231.	1.1	24
2274	River trail. <i>ACM SIGPLAN Notices</i> , 2013, 48, 729-744.	0.2	7
2275	Power-aware dynamic memory management on many-core platforms utilizing DVFS. <i>Transactions on Embedded Computing Systems</i> , 2013, 13, 1-25.	2.1	66
2276	SMINER - a platform for data mining based on service-oriented architecture. <i>International Journal of Business Intelligence and Data Mining</i> , 2013, 8, 1.	0.2	3
2277	TeachCloud: a cloud computing educational toolkit. <i>International Journal of Cloud Computing</i> , 2013, 2, 237.	0.3	59
2278	Vision Cloud: A Cloud Storage Solution Supporting Modern Media Production. <i>Smpte Motion Imaging Journal</i> , 2013, 122, 30-37.	0.2	1
2279	Scalable and Adaptive Graph Querying with MapReduce. <i>IEICE Transactions on Information and Systems</i> , 2013, E96.D, 2126-2130.	0.4	0

#	ARTICLE	IF	CITATIONS
2281	Autonomous, failure-resilient orchestration of distributed discrete event simulations. , 2013, , .		3
2282	Hardware-oblivious parallelism for in-memory column-stores. Proceedings of the VLDB Endowment, 2013, 6, 709-720.	2.1	82
2284	BFT-TO: Intrusion Tolerance with Less Replicas. Computer Journal, 2013, 56, 693-715.	1.5	11
2285	Scalable Graph Clustering with Pregel. Studies in Computational Intelligence, 2013, , 133-144.	0.7	3
2286	Design and Implement a MapReduce Framework for Executing Standalone Software Packages in Hadoop-based Distributed Environments. Smart Science, 2013, 1, 99-107.	1.9	1
2287	Real-Time Collaborative Planning with Big Data: Technical Challenges and In-Place Computing. , 2013, , .		1
2288	Graph-based malware distributors detection. , 2013, , .		6
2289	Agents' Monitoring Approach for Big Data. , 2013, , .		0
2290	Boosting energy efficiency with mirrored data block replication policy and energy scheduler. Operating Systems Review (ACM), 2013, 47, 33-40.	1.5	9
2291	SPGM: an efficient algorithm for mapping MapReduce-like data-intensive applications in data centre network. International Journal of Web and Grid Services, 2013, 9, 172.	0.4	6
2292	The Role and Challenges of Exome Sequencing in Studies of Human Diseases. Frontiers in Genetics, 2013, 4, 160.	1.1	79
2293	Enabling actionable analytics for mobile devices: performance issues of distributed analytics on Hadoop mobile clusters. Journal of Cloud Computing: Advances, Systems and Applications, 2013, 2, 15.	2.1	7
2294	Monitoring social relationship among Twitter users by using NodeXL. , 2013, , .		5
2295	A high-level framework for parallelizing legacy applications for multiple platforms. , 2013, , .		2
2296	A survey of web archive search architectures. , 2013, , .		17
2297	Butterfly Mixing: Accelerating Incremental-Update Algorithms on Clusters. , 2013, , .		6
2298	Mammoth. , 2013, , .		5
2299	An Overview of Multiple Sequence Alignments and Cloud Computing in Bioinformatics. , 2013, 2013, 1-14.		59

#	ARTICLE	IF	CITATIONS
2300	RFID Application of Smart Grid for Asset Management. International Journal of Antennas and Propagation, 2013, 2013, 1-6.	0.7	3
2301	Astronomical data processing in EXTASCID. , 2013, , .		12
2302	Scalable processing of flexible graph pattern queries on the cloud. , 2013, , .		3
2303	Identifying abnormal patterns in cellular communication flows. , 2013, , .		4
2304	Design and Realization of user Behaviors Recommendation System Based on Association rules under Cloud Environment. Research Journal of Applied Sciences, Engineering and Technology, 2013, 6, 1669-1673.	0.1	0
2305	Document-Based Databases for Medical Information Systems and Crisis Management. International Journal of Information Systems for Crisis Response and Management, 2013, 5, 63-80.	0.7	6
2306	Evaluation and Analysis of Distributed Graph-Parallel Processing Frameworks. Journal of Cyber Security and Mobility, 2014, 3, 289-316.	0.7	19
2307	CloudDOE: A User-Friendly Tool for Deploying Hadoop Clouds and Analyzing High-Throughput Sequencing Data with MapReduce. PLoS ONE, 2014, 9, e98146.	1.1	24
2308	Benchmarking Undedicated Cloud Computing Providers for Analysis of Genomic Datasets. PLoS ONE, 2014, 9, e108490.	1.1	10
2309	Applying MapReduce Framework to Peer-to-Peer Overlay Network. , 2014, , .		2
2310	MOCgraph. Proceedings of the VLDB Endowment, 2014, 8, 377-388.	2.1	32
2311	PhysioDroid: Combining Wearable Health Sensors and Mobile Devices for a Ubiquitous, Continuous, and Personal Monitoring. Scientific World Journal, The, 2014, 2014, 1-11.	0.8	73
2312	Efficient and Scalable Graph Similarity Joins in MapReduce. Scientific World Journal, The, 2014, 2014, 1-11.	0.8	3
2313	An Improved Clustering Algorithm of Tunnel Monitoring Data for Cloud Computing. Scientific World Journal, The, 2014, 2014, 1-6.	0.8	6
2314	Context-Aware and Locality-Constrained Coding for Image Categorization. Scientific World Journal, The, 2014, 2014, 1-14.	0.8	0
2315	Shortest Paths Based Web Service Selection in Internet of Things. Journal of Sensors, 2014, 2014, 1-10.	0.6	8
2316	Meteorological Data Analysis Using MapReduce. Scientific World Journal, The, 2014, 2014, 1-10.	0.8	10
2317	A Cloud-Based X73 Ubiquitous Mobile Healthcare System: Design and Implementation. Scientific World Journal, The, 2014, 2014, 1-14.	0.8	19

#	ARTICLE	IF	CITATIONS
2318	Risk Intelligence: Making Profit from Uncertainty in Data Processing System. Scientific World Journal, The, 2014, 2014, 1-16.	0.8	1
2319	Neuroimaging explanations of age-related differences in task performance. Frontiers in Aging Neuroscience, 2014, 6, 46.	1.7	21
2320	Machine learning patterns for neuroimaging-genetic studies in the cloud. Frontiers in Neuroinformatics, 2014, 8, 31.	1.3	11
2321	Streaming Progressive TIN Densification Filter for Airborne LiDAR Point Clouds Using Multi-Core Architectures. Remote Sensing, 2014, 6, 7212-7232.	1.8	17
2322	DGFIIndex for smart grid. Proceedings of the VLDB Endowment, 2014, 7, 1496-1507.	2.1	18
2323	Educational Data Mining and Learning Analytics. , 2014, , 253-272.		217
2324	Scaling Distributed Machine Learning with the Parameter Server. , 2014, , .		552
2325	Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques. Nonlinear Processes in Geophysics, 2014, 21, 777-795.	0.6	40
2326	Challenges and Potential Solutions for Big Data Implementations in Developing Countries. Yearbook of Medical Informatics, 2014, 23, 36-41.	0.8	25
2327	Sequence Alignment Tools: One Parallel Pattern to Rule Them All?. BioMed Research International, 2014, 2014, 1-12.	0.9	10
2328	Nesting Strategies for Enabling Nimble MapReduce Dataflows for Large RDF Data. International Journal on Semantic Web and Information Systems, 2014, 10, 1-26.	2.2	7
2329	How In-memory Technology Can Create Business Value: Insights from the Hilti Case. Communications of the Association for Information Systems, 2014, 34, .	0.7	15
2330	HEPDOOP: High-Energy Physics Analysis using Hadoop. Journal of Physics: Conference Series, 2014, 513, 022004.	0.3	1
2331	MAPREDUCE CHALLENGES ON PERVASIVE GRIDS. Journal of Computer Science, 2014, 10, 2194-2210.	0.5	8
2332	Improving MapReduce Performance in a Heterogeneous Cloud: A Measurement Study. , 2014, , .		7
2333	R-HBase: A Multi-dimensional Indexing Framework for Cloud Computing Environment. , 2014, , .		9
2334	Scalable and Reliable Data Broadcast with Cascade. , 2014, , .		0
2335	Stream Bench: Towards Benchmarking Modern Distributed Stream Computing Frameworks. , 2014, , .		65

#	ARTICLE	IF	CITATIONS
2336	An efficient motif finding algorithm for large DNA data sets. , 2014, , .		5
2337	Cloud-Niagara: A high availability and low overhead fault tolerance middleware for the cloud. , 2014, , .		6
2338	Rubato DB. , 2014, , .		10
2339	Execution Drafting: Energy Efficiency through Computation Deduplication. , 2014, , .		14
2340	A Parallel Algorithm for Datacleansing in Incomplete Information Systems Using MapReduce. , 2014, , .		5
2341	MapReduce for Large-Scale Monitor Data Analyses. , 2014, , .		1
2342	AlignMR. , 2014, , .		1
2343	Fuzzy K-mean clustering in MapReduce on cloud based hadoop. , 2014, , .		14
2344	Acceleration of Naive-Bayes algorithm on multicore processor for massive text classification. , 2014, , .		6
2345	PAD: Performance Anomaly Detection in Multi-server Distributed Systems. , 2014, , .		19
2346	Parallel Web server load balancing technology of cloud computing environments. , 2014, , .		0
2347	A Cloud-Based SOA for Enhancing Information Exchange and Decision Support in ITT Operations. Lecture Notes in Computer Science, 2014, , 112-131.	1.0	7
2348	Dynamic Colocation Algorithm for Hadoop. , 2014, , .		4
2349	Distributed in-memory cluster computing approach in scala for solving graph data applications. , 2014, , .		7
2350	Optimizing the Join Operation on Hive to Accelerate Cross-Matching in Astronomy. , 2014, , .		6
2351	Parallel feature selection using positive approximation based on MapReduce. , 2014, , .		8
2352	ABC2: Adaptively Balancing Computation and Communication in a DSM Cluster of Multicores for Irregular Applications. , 2014, , .		2
2353	A study on specialist or special disease clinics based on big data. Frontiers of Medicine, 2014, 8, 376-381.	1.5	8

#	ARTICLE	IF	CITATIONS
2354	TwigMRR: Distributed XML Twig Query Processing. , 2014, , .		0
2355	A performance analysis of MapReduce applications on big data in cloud based Hadoop. , 2014, , .		6
2356	Semantic HMC for big data analysis. , 2014, , .		4
2357	Spark-based anomaly detection over multi-source VMware performance data in real-time. , 2014, , .		29
2358	A Data Science Solution for Mining Interesting Patterns from Uncertain Big Data. , 2014, , .		50
2359	Performance Variations in Resource Scaling for MapReduce Applications on Private and Public Clouds. , 2014, , .		4
2360	Robust scheduling in a flexible fork-join network. , 2014, , .		6
2361	Non-iteration Parallel Algorithm for Frequent Pattern Discovery. , 2014, , .		1
2362	Predicting energy consumption with StreamMine3G. , 2014, , .		8
2363	Progressive Visual Analytics: User-Driven Visual Exploration of In-Progress Analytics. IEEE Transactions on Visualization and Computer Graphics, 2014, 20, 1653-1662.	2.9	146
2364	A mitral annulus tracking approach for navigation of offâ€pump beating heart mitral valve repair. Medical Physics, 2015, 42, 456-468.	1.6	10
2365	A distributed architecture for intra- and inter-cloud data management. , 2014, , .		1
2366	Scheduling tasks with precedence constraints on multiple servers. , 2014, , .		28
2367	A cross-job framework for MapReduce scheduling. , 2014, , .		0
2368	DRAW: A New Data-gRouping-AWare Data Placement Scheme for Data Intensive Applications with Interest Locality. , 2014, , 149-174.		17
2369	Optimizing Power and Performance Trade-offs of MapReduce Job Processing with Heterogeneous Multi-core Processors. , 2014, , .		8
2370	A Hybrid Optimization Model for Green Cloud Computing. , 2014, , .		14
2371	Scalable Computation of Stream Surfaces on Large Scale Vector Fields. , 2014, , .		14

#	ARTICLE	IF	CITATIONS
2372	Increasing the accessibility to Big Data systems via a common services API. , 2014, , .		5
2373	Big Data for Cyber Physical Systems: An Analysis of Challenges, Solutions and Opportunities. , 2014, , .		28
2374	SE-CDA: A scalable and efficient community detection algorithm. , 2014, , .		3
2375	A Hadoop Extension to Process Mail Folders and its Application to a Spam Dataset. , 2014, , .		1
2376	BigExcel: A web-based framework for exploring big data in social sciences. , 2014, , .		3
2377	Multiple Two-Phase Data Processing with MapReduce. , 2014, , .		3
2378	Big data analytics on PMU measurements. , 2014, , .		23
2379	Survey on clustering methods: Towards fuzzy clustering for big data. , 2014, , .		39
2380	A layer based architecture for provenance in big data. , 2014, , .		15
2381	A Big Data Financial Information Management Architecture for Global Banking. , 2014, , .		14
2382	Preventing passive TCP timeouts in data center networks with packet drop notification. , 2014, , .		3
2383	Big data implementation and visualization. , 2014, , .		4
2384	Visual analysis of large dental imaging data in caries research. , 2014, , .		3
2385	Synergistic partitioning in multiple large scale social networks. , 2014, , .		21
2386	Cloud based system for measurement data management in large scale electronic production. , 2014, , .		3
2387	BigCache for big-data systems. , 2014, , .		3
2388	A framework for Internet data real-time processing: A machine-learning approach. , 2014, , .		11
2389	Priority-Based Resource Scheduling in Distributed Stream Processing Systems for Big Data Applications. , 2014, , .		21

#	ARTICLE	IF	CITATIONS
2390	MR-microT. , 2014, , .		9
2391	Big Data Analysis Using Apache Hadoop. , 2014, , .		45
2392	Cloud-scale genomic signals processing classification analysis for gene expression microarray data. , 2014, 2014, 1843-6.		2
2393	Big data based retail recommender system of non E-commerce. , 2014, , .		12
2394	Do Search-Based Approaches Improve the Design of Self-Adaptive Systems? A Controlled Experiment. , 2014, , .		2
2395	Performance analysis model for big data applications in cloud computing. Journal of Cloud Computing: Advances, Systems and Applications, 2014, 3, .	2.1	19
2396	Joint scheduling of MapReduce jobs with servers: Performance bounds and experiments. , 2014, , .		20
2397	Scheduling Decisions in Stream Processing on Heterogeneous Clusters. , 2014, , .		30
2398	A MapReduce based parallel algorithm for CIM data verification. , 2014, , .		0
2399	Correlation clustering in MapReduce. , 2014, , .		36
2400	SAMURAI: A Streaming Multi-tenant Context-Management Architecture for Intelligent and Scalable Internet of Things Applications. , 2014, , .		14
2401	Congestion pattern model for predicting short-term traffic decongestion times. , 2014, , .		5
2402	Writing Parallel Embedded Software Effectively. , 2014, , .		0
2403	Distributed MapReduce engine with fault tolerance. , 2014, , .		2
2404	Distributed video transcoding based on MapReduce. , 2014, , .		11
2405	High level programming framework for FPGAs in the data center. , 2014, , .		12
2406	PIIM: Method of Identifying Malicious Workers in the MapReduce System with an Open Environment. , 2014, , .		2
2407	AJIRA: A Lightweight Distributed Middleware for MapReduce and Stream Processing. , 2014, , .		16

#	ARTICLE	IF	CITATIONS
2408	Performance Evaluation of the SSD-Based Swap System for Big Data Processing. , 2014, , .		4
2409	Snore related signals processing in a private cloud computing system. Interdisciplinary Sciences, Computational Life Sciences, 2014, 6, 216-221.	2.2	3
2410	Statistical technique for online anomaly detection using Spark over heterogeneous data from multi-source VMware performance data. , 2014, , .		22
2411	A micropartitioning technique for massive data analysis using MapReduce. , 2014, , .		1
2412	Understanding the behavior of in-memory computing workloads. , 2014, , .		35
2413	An improved parallel association rules algorithm based on MapReduce framework for big data. , 2014, , .		14
2414	Design and implementation of data mining platform based on the cloud computing. , 2014, , .		0
2415	Enhancing Big Data Security with Collaborative Intrusion Detection. IEEE Cloud Computing, 2014, 1, 27-33.	5.3	63
2416	Energy-Efficient Data Center Networks Planning with Virtual Machine Placement and Traffic Configuration. , 2014, , .		21
2417	Optimal Capacity Allocation for Executing MapReduce Jobs in Cloud Systems. , 2014, , .		7
2418	Data Trasfer From MySQL To Hadoop. , 2014, , .		2
2419	RuleMR: Classification rule discovery with MapReduce. , 2014, , .		4
2420	A contention aware hybrid evaluator for schedulers of big data applications in computer clusters. , 2014, , .		2
2421	LUT based secure cloud computing — An implementation using FPGAs. , 2014, , .		0
2422	Efficient UDP-based congestion aware transport for data center traffic. , 2014, , .		1
2423	MapReduce approach to relational influence propagation in complex networks. Pattern Analysis and Applications, 2014, 17, 739-746.	3.1	11
2424	Allocating Bandwidth in Datacenter Networks: A Survey. Journal of Computer Science and Technology, 2014, 29, 910-917.	0.9	13
2425	Incremental and parallel spatial association mining. , 2014, , .		12

#	ARTICLE	IF	CITATIONS
2426	Server farms with batch arrival and staggered setup. , 2014, , .		14
2427	On the Application and Performance of MongoDB for Climate Satellite Data. , 2014, , .		13
2428	Information retrieval using dynamic indexing. , 2014, , .		2
2429	Self adaptive Hadoop scheduler for heterogeneous resources. , 2014, , .		4
2430	A Parallel Implementation of Idea Graph to Extract Rare Chances from Big Data. , 2014, , .		3
2431	Concepts and requirements for a cloud-based optimization service. , 2014, , .		3
2432	Astro: A predictive model for anomaly detection and feedback-based scheduling on Hadoop. , 2014, , .		4
2433	Scheduling divisible reduce tasks in MapReduce. , 2014, , .		0
2434	SOM Clustering Using Spark-MapReduce. , 2014, , .		16
2435	Scalable and elastic realtime click stream analysis using StreamMine3G. , 2014, , .		17
2436	On Building Decision Trees from Large-scale Data in Applications of On-line Advertising. , 2014, , .		4
2437	An improvement of weighted PageRank to handle the zero link similarity. , 2014, , .		0
2438	Perladoop: Efficient execution of Perl scripts on Hadoop clusters. , 2014, , .		5
2439	Hot topic analysis and content mining in social media. , 2014, , .		6
2440	SARAH - Statistical Analysis for Resource Allocation in Hadoop. , 2014, , .		2
2441	Optimising parallel R correlation matrix calculations on gene expression data using MapReduce. BMC Bioinformatics, 2014, 15, 351.	1.2	19
2442	FedLoop: Looping on Federated MapReduce. , 2014, , .		5
2443	Towards Energy Awareness in Hadoop. , 2014, , .		5

#	ARTICLE	IF	CITATIONS
2444	Parallel algorithm for Hill Cipher on MapReduce. , 2014, , .		3
2445	Low Cost Energy Forecasting for Smart Grids Using Stream Mine 3G and Amazon EC2. , 2014, , .		3
2446	SiftD: A CPU & GPU distributed hybrid system for SIFT. , 2014, , .		0
2447	Palantir: Reseizing Network Proximity in Large-Scale Distributed Computing Frameworks Using SDN. , 2014, , .		5
2448	Big Data Processing with Probabilistic Latent Semantic Analysis on MapReduce. , 2014, , .		2
2449	The Use of Distributed Processing and Cloud Computing in Agricultural Decision-Making Support Systems. , 2014, , .		5
2450	MR-TRIAGE: Scalable multi-criteria clustering for big data security intelligence applications. , 2014, , .		3
2451	Partitioner: A Distributed HDFS Metadata Server Cluster. , 2014, , .		6
2452	A system architecture for manufacturing process analysis based on big data and process mining techniques. , 2014, , .		27
2454	A cooperative sensing and mining system for transportation activity survey. , 2014, , .		2
2455	Use of Big Data technology in Vehicular Ad-hoc Networks. , 2014, , .		32
2456	A Task Scheduling Algorithm Based on Genetic Algorithm and Ant Colony Optimization in Cloud Computing. , 2014, , .		57
2457	Frame-Semantic Parsing. Computational Linguistics, 2014, 40, 9-56.	2.5	160
2458	A study of big data processing constraints on a low-power Hadoop cluster. , 2014, , .		22
2459	PFC: Privacy Preserving FPGA Cloud - A Case Study of MapReduce. , 2014, , .		21
2460	Scalable Collaborative Filtering Recommendation Algorithm with MapReduce. , 2014, , .		5
2461	A Hadoop-Galaxy adapter for user-friendly and scalable data-intensive bioinformatics in Galaxy. , 2014, , .		5
2462	Big data machine learning and graph analytics: Current state and future challenges. , 2014, , .		18

#	ARTICLE	IF	CITATIONS
2463	Fast classification in industrial Big Data environments. , 2014, , .		5
2464	A round robin with multiple feedback job scheduler in Hadoop. , 2014, , .		8
2465	Parallelizing K-Means Algorithm for 1-D Data Using MPI. , 2014, , .		8
2466	GISQF: An Efficient Spatial Query Processing System. , 2014, , .		19
2467	Web oriented learning strategies of relationships among domain ontology concepts. , 2014, , .		0
2468	A VoLTE traffic classification method in LTE network. , 2014, , .		7
2469	GPFS-based implementation of a hyperconverged system for software defined infrastructure. IBM Journal of Research and Development, 2014, 58, 6:1-6:12.	3.2	7
2470	A novel approach to implement a shop bot on distributed web crawler. , 2014, , .		1
2471	A Platform to Deploy Customized Scientific Virtual Infrastructures on the Cloud. , 2014, , .		1
2473	Automata-Based Dynamic Data Processing for Clouds. Lecture Notes in Computer Science, 2014, , 93-104.	1.0	3
2474	Big Data architecture for IT incident management. , 2014, , .		6
2475	Reliability and Availability Properties of Distributed Database Systems. , 2014, , .		15
2476	An efficient PAM spatial clustering algorithm based on MapReduce. , 2014, , .		1
2477	Rethinking the Data Center Networking: Architecture, Network Protocols, and Resource Sharing. IEEE Access, 2014, 2, 1481-1496.	2.6	72
2478	Statistical Frameworks for Detecting Tunnelling in Cyber Defence Using Big Data. , 2014, , .		4
2480	In unity there is strength: Showcasing a unified big data platform with MapReduce Over both object and file storage. , 2014, , .		2
2481	A survey on security of Hadoop. , 2014, , .		22
2482	Research on MapReduce-Based Rocchio Relevance Feedback in Massive Information Filtering. Advanced Materials Research, 0, 962-965, 2712-2715.	0.3	0

#	ARTICLE	IF	CITATIONS
2483	Towards Realizing the Secured Multilateral Co-Operative Computing Architectural Framework. , 2014, , .		3
2484	Incremental Text Indexing for Fast Disk-Based Search. ACM Transactions on the Web, 2014, 8, 1-31.	2.0	2
2485	CloudETL. , 2014, , .		24
2486	University Employment Information Integration Model Based on Cloud Computing. Applied Mechanics and Materials, 0, 599-601, 1807-1811.	0.2	0
2487	Research on Parallel Computation of Semantic Similarity in Linked Data. Advanced Materials Research, 0, 1049-1050, 1320-1326.	0.3	0
2488	Global Disease Monitoring and Forecasting with Wikipedia. PLoS Computational Biology, 2014, 10, e1003892.	1.5	161
2489	Ceph CRUSH Data Distribution Algorithms. Applied Mechanics and Materials, 0, 596, 196-199.	0.2	0
2490	MRScheduling: An Effective Technique for Multi-Tenant Meeting Deadline in MapReduce. Applied Mechanics and Materials, 0, 644-650, 4482-4486.	0.2	0
2491	MapBeing: An architecture for manipulating and publishing vector data in web based Geographic Information System. , 2014, , .		1
2492	Congestion Score Computation of Big Traffic Data. , 2014, , .		6
2493	Research on Data Mining Optimization and Security Based on MapReduce. Applied Mechanics and Materials, 0, 631-632, 1053-1056.	0.2	0
2494	A Hadoop-Based Output Analyzer for Large-Scale Simulation Data. , 2014, , .		2
2495	SDNs, Clouds, and Big Data: New Opportunities. , 2014, , .		14
2496	Stream computing for large-scale, multi-channel cyber threat analytics. , 2014, , .		4
2497	Emerging Issues in Cloud Storage Security: Encryption, Key Management, Data Redundancy, Trust Mechanism. Communications in Computer and Information Science, 2014, , 297-310.	0.4	9
2498	Bioinformatics for precision medicine in oncology: principles and application to the SHIVA clinical trial. Frontiers in Genetics, 2014, 5, 152.	1.1	72
2499	Hadoop Extensions for Distributed Computing on Reconfigurable Active SSD Clusters. Transactions on Architecture and Code Optimization, 2014, 11, 1-26.	1.6	11
2500	Reconfigurable network testbed for evaluation of datacenter topologies. , 2014, , .		6

#	ARTICLE	IF	CITATIONS
2501	Distributed REScala. , 2014, , .		26
2502	SemStore. , 2014, , .		23
2503	Analysis of large call data records with big data. , 2014, , .		2
2504	Analysis of dynamic scheduling strategies for matrix multiplication on heterogeneous platforms. , 2014, , .		3
2505	Secure streaming forensic data transmission for trusted cloud. , 2014, , .		5
2506	ScadiBino. , 2014, , .		11
2507	Taking Authenticated Range Queries to Arbitrary Dimensions. , 2014, , .		32
2508	MaPLE: A MapReduce Pipeline for Lattice-based Evaluation and its application to SNOMED CT. , 2014, 2014, 754-759.		18
2509	Notice of Violation of IEEE Publication Principles: Hadoop Preemptive Deadline Constraint Scheduler. , 2014, , .		2
2510	SQL-MapReduce hybrid approach towards distributed projected clustering. , 2014, , .		1
2511	Distributed file load rebalancing methodology for map reduce system. , 2014, , .		0
2512	The Application of Mobile Cloud in Heterogeneous Data Storage in Web of Things System. , 2014, , .		1
2513	OWL Reasoning Framework over Big Biological Knowledge Network. BioMed Research International, 2014, 2014, 1-16.	0.9	3
2514	Improving Accuracy and Scalability of Personal Recommendation Based on Bipartite Network Projection. Mathematical Problems in Engineering, 2014, 2014, 1-11.	0.6	3
2516	The Optimization of Hadoop Scheduling Algorithms on Distributed System for Processing Traffic Information. Advances in Intelligent Systems and Computing, 2014, , 389-396.	0.5	0
2517	Taming Computation Skews of Block-Oriented Iterative Scientific Applications in MapReduce Systems. , 2014, , .		0
2518	Natural Language Understanding and Prediction: from Formal Grammars to Large Scale Machine Learning. Fundamenta Informaticae, 2014, 131, 425-440.	0.3	2
2519	Study of Map-Reduce over Hadoop Based Cloud Computing Environment. Applied Mechanics and Materials, 2014, 509, 175-181.	0.2	0

#	ARTICLE	IF	CITATIONS
2520	An Optimization Scheme for Bank Batch Processing Based on Cloud Computation. Applied Mechanics and Materials, 0, 539, 339-344.	0.2	0
2521	Medusa. SIGMOD Record, 2014, 43, 35-40.	0.7	27
2522	Exploring new privacy approaches in a scalable classification framework. , 2014, , .		4
2523	A Load Balancing Mechanism Using Bloom Filter in Storm System. Applied Mechanics and Materials, 0, 543-547, 1972-1976.	0.2	0
2524	The Design and Implementation of Astronomical Multi-Catalog Storage and Cross-Match Based on Hadoop. Advanced Materials Research, 2014, 1042, 121-125.	0.3	0
2525	E-Commerce Recommendation System Based on Cloud Computing. Advanced Materials Research, 0, 926-930, 2570-2573.	0.3	1
2526	Distributed Collaborative Filtering Recommendation Model Based on Expand-Vector. Advanced Materials Research, 2014, 989-994, 2188-2191.	0.3	2
2527	CUDA on Hadoop: A Mixed Computing Framework for Massive Data Processing. Advances in Intelligent Systems and Computing, 2014, , 253-260.	0.5	2
2528	Parallel Attribute Reduction Based on MapReduce. Lecture Notes in Computer Science, 2014, , 631-641.	1.0	1
2529	Savant: Automatic generation of a parallel scheduling heuristic for map-reduce. International Journal of Hybrid Intelligent Systems, 2014, 11, 287-302.	0.9	7
2530	On Designing Multicore-Aware Simulators for Systems Biology Endowed with OnLine Statistics. BioMed Research International, 2014, 2014, 1-14.	0.9	2
2531	epiC. Proceedings of the VLDB Endowment, 2014, 7, 541-552.	2.1	47
2532	Semantic Web technologies for the big data in life sciences. BioScience Trends, 2014, 8, 192-201.	1.1	11
2533	Leveraging cloud computing for spatial association mining. , 2014, , .		1
2534	An efficient parallel method for canopy clusteing with multi-core platform. , 2014, , .		0
2535	Mesa. Proceedings of the VLDB Endowment, 2014, 7, 1259-1270.	2.1	56
2536	Efficient Spatio-textual Similarity Join Using MapReduce. , 2014, , .		17
2537	Parallelization of ontology construction and fusion based on MapReduce. , 2014, , .		2

#	ARTICLE	IF	CITATIONS
2538	Learning Based Distributed Orchestration of Stochastic Discrete Event Simulations. , 2014, , .		0
2539	ParSim: A Tool for Workload Modeling and Reproduction of Parallel Applications. , 2014, , .		1
2540	An Analysis of Failure-Related Energy Waste in a Large-Scale Cloud Environment. IEEE Transactions on Emerging Topics in Computing, 2014, 2, 166-180.	3.2	48
2541	Optimizing performance and power consumption for an ARM-based big data cluster. , 2014, , .		4
2542	Shared I/O Scheduling in Cloud for Structured Data Processing. , 2014, , .		1
2543	Graph processing on GPUs: Where are the bottlenecks?. , 2014, , .		55
2544	An ontology based digital preservation system for enterprise collaboration. , 2014, , .		3
2545	Computational Models for Big Data Processing. , 2014, , .		0
2546	MyBSP: An Iterative Processing Framework Based on the Cloud Platform for Graph Data. , 2014, , .		1
2547	Querying Road Traffic Data from a Document Store. , 2014, , .		4
2548	Workflow Partitioning and Deployment on the Cloud Using Orchestra. , 2014, , .		11
2549	Evaluating Grasp-based cloud dimensioning for comparative genomics: A practical approach. , 2014, , .		8
2550	A test-suite generator for database systems. , 2014, , .		4
2551	Deadline-Constrained MapReduce Scheduling Based on Graph Modelling. , 2014, , .		5
2552	Dynamic taxonomy composition via keyqueries. , 2014, , .		6
2553	Distributed evaluation of XPath queries over large integrated XML data. , 2014, , .		2
2554	Analysis of network node behavior based on MapReduce. , 2014, , .		0
2555	Smart MapReduce cloud: Applying extra processing to intermediate data on demand. , 2014, , .		2

#	ARTICLE	IF	CITATIONS
2556	A distributed gaussian-means clustering algorithm for forecasting domestic energy usage. , 2014, , .		4
2557	WORMStore: A Specialized Object Store for Write-Once Read-Many Workloads. , 2014, , .		0
2558	Greening 4-4, 1-4 Data Center Network: A Greedy Approach for Finding an Energy Efficient Sub-network. , 2014, , .		1
2559	Factorized Similarity Learning in Networks. , 2014, , .		43
2560	JUNIPER. , 2014, , .		6
2561	Peer-Comparison Based Fault Diagnosis for Hadoop Systems. Applied Mechanics and Materials, 0, 621, 235-240.	0.2	0
2562	NITOS testbed: A cloud based wireless experimentation facility. , 2014, , .		22
2563	Scalable Security Analysis Using a Partition and Merge Approach in an Infrastructure as a Service Cloud. , 2014, , .		3
2564	Advanced architectures distributed systems for the implementation of neural networks. , 2014, , .		0
2565	VENU: Orchestrating SSDs in hadoop storage. , 2014, , .		17
2566	A Novel Top-K Automobiles Probabilistic Recommendation Model Using User Preference and User Community. , 2014, , .		2
2567	An Image Processing System Based on Heterogeneous Embedded Multicore Processors. , 2014, , .		2
2568	An efficient cloud-assisted best-parking algorithm for BikeNet. International Journal of Ad Hoc and Ubiquitous Computing, 2014, 16, 136.	0.3	1
2569	A Covering Classification Rule Induction Approach for Big Datasets. , 2014, , .		2
2570	Sparse matrix partitioning for parallel eigenanalysis of large static and dynamic graphs. , 2014, , .		3
2571	Energy Consumption Data Based Machine Anomaly Detection. , 2014, , .		8
2572	Toward Enhancing Block I/O Performance for Virtualized Hadoop Cluster. , 2014, , .		3
2573	A Computing Model for Real-Time Stream Processing. , 2014, , .		4

#	ARTICLE	IF	CITATIONS
2574	R2Time: A Framework to Analyse Open TSDB Time-Series Data in HBase. , 2014, , .		3
2575	A Comparison Study of Graph Data Processing Based on MyBSP and MapReduce. , 2014, , .		0
2576	Scalability Analysis and Improvement of Hadoop Virtual Cluster with Cost Consideration. , 2014, , .		5
2577	Fundamentals of a graph transformation based web data processing system. , 2014, , .		0
2578	Distributed Evaluation of XPath Axes Queries over Large XML Documents Stored in MapReduce Clusters. , 2014, , .		2
2579	Using an Object-Based Active Storage Framework to Improve Parallel Storage Systems. , 2014, , .		4
2580	TS-Hadoop: Handling access skew in MapReduce by using tiered storage infrastructure. , 2014, , .		2
2581	A Portable Tool for Running MPI Applications in the Cloud. , 2014, , .		2
2582	Computation and communication efficient graph processing with distributed immutable view. , 2014, , .		51
2583	Bundle recommendation in ecommerce. , 2014, , .		73
2584	Scalable Indexing and Adaptive Querying of RDF Data in the cloud. , 2014, , .		1
2585	Home Appliance Management System for Monitoring Digitized Devices Using Cloud Computing Technology in Ubiquitous Sensor Network Environment. International Journal of Distributed Sensor Networks, 2014, 10, 174097.	1.3	9
2586	Enabling Prioritized Cloud I/O Service in Hadoop Distributed File System. , 2014, , .		0
2587	[phi]Sched: A Heterogeneity-Aware Hadoop Workflow Scheduler. , 2014, , .		23
2588	Reduce and aggregate. , 2014, , .		4
2589	Combining Hadoop and GPU to preprocess large Affymetrix microarray data. , 2014, , .		6
2590	SLAM. , 2014, , .		7
2591	1-to-N and N-to-1 Communication for Optical Networks. , 2014, , .		0

#	ARTICLE	IF	CITATIONS
2592	Taking a walk on the wild side. , 2014, , .		8
2593	Data-parallel finite-state machines. , 2014, , .		71
2594	Continuous validation of load test suites. , 2014, , .		21
2595	Stratified-sampling over social networks using mapreduce. , 2014, , .		8
2596	TriAD. , 2014, , .		119
2597	Efficient task replication for fast response times in parallel computation. , 2014, , .		48
2598	FileMap. , 2014, , .		1
2599	Adapting our view of software adaptation: an architectural perspective (keynote). , 2014, , .		0
2600	Seraph. , 2014, , .		40
2601	Data stream processing in dynamic and decentralized peer-to-peer networks. , 2014, , .		5
2602	Reusing constraint proofs for scalable program analysis. , 2014, , .		1
2603	Auto-scaling techniques for elastic data stream processing. , 2014, , .		19
2604	Database Queries that Explain their Work. , 2014, , .		11
2605	DiFS. , 2014, , .		25
2606	MapReduce. , 2014, , .		7
2607	CheckCell. , 2014, , .		18
2608	Software engineering for multi-tenancy computing challenges and implications. , 2014, , .		5
2609	Agility and Performance in Elastic Distributed Storage. ACM Transactions on Storage, 2014, 10, 1-27.	1.4	9

#	ARTICLE	IF	CITATIONS
2610	Adaptive Stream Processing using Dynamic Batch Sizing. , 2014, , .		81
2611	The X-flex cross-platform scheduler. , 2014, , .		3
2612	A Data Mining-based Blocks Placement Optimization for Distributed Data Warehouses. , 2014, , .		0
2613	An Intrusive Analyzer for Hadoop Systems Based on Wireless Sensor Networks. International Journal of Distributed Sensor Networks, 2014, 10, 196040.	1.3	1
2614	A Distributed Weighted Possibilistic c-Means Algorithm for Clustering Incomplete Big Sensor Data. International Journal of Distributed Sensor Networks, 2014, 10, 430814.	1.3	27
2615	DSMC: A Novel Distributed Store-Retrieve Approach of Internet Data Using MapReduce Model and Community Detection in Big Data. International Journal of Distributed Sensor Networks, 2014, 10, 430848.	1.3	0
2616	A Shared Interest Discovery Model for Coauthor Relationship in SNS. International Journal of Distributed Sensor Networks, 2014, 10, 820715.	1.3	3
2617	Performance analysis of MapReduce on OpenStack-based hadoop virtual cluster. , 2014, , .		2
2618	Big data clustering via random sketching and validation. , 2014, , .		1
2619	To Overlap or Not to Overlap: Optimizing Incremental MapReduce Computations for On-Demand Data Upload. , 2014, , .		2
2620	A Spatio-temporal Parallel Processing System for Traffic Sensory Data. , 2014, , .		3
2621	A Data Processing Algorithm in EPC Internet of Things. , 2014, , .		4
2622	PAC: Taming TCP Incast Congestion Using Proactive ACK Control. , 2014, , .		30
2623	Modeling and Understanding TCP's Fairness Problem in Data Center Networks. , 2014, , .		0
2624	Parallel collaborative filtering recommendation model based on expand-vector. , 2014, , .		5
2625	Workflow Skeletons: A Non-intrusive Approach for Facilitating Scientific Workflow Modeling. , 2014, , .		1
2626	A load balance algorithm based on nodes performance in Hadoop cluster. , 2014, , .		7
2627	MapReduce-Based Data Processing on IoT. , 2014, , .		7

#	ARTICLE	IF	CITATIONS
2628	How to read the web in portuguese using the never-ending language learner's principles. , 2014, , .		4
2629	SkewControl: Gini Out of the Bottle. , 2014, , .		2
2630	Data locality in MapReduce: A network perspective. , 2014, , .		8
2631	Scaling the Real-Time Traffic Sensing with GPS Equipped Probe Vehicles. , 2014, , .		5
2632	Efficient and robust allocation algorithms in clouds under memory constraints. , 2014, , .		4
2633	Data-centric HPC for Numerical Weather Forecasting. , 2014, , .		2
2634	A Distributed Dataflow Model for Task-Uncoordinated Parallel Program Execution. , 2014, , .		3
2635	An approach to discover the best-fit factors for the optimal performance of Hadoop map reduce in virtualized environment. , 2014, , .		0
2636	Tradeoffs for task parallelization in distributed optimization. , 2014, , .		0
2637	Quartz. , 2014, , .		25
2638	Topic similarity networks: Visual analytics for large document sets. , 2014, , .		11
2639	Distributed index mechanism based on Hadoop. , 2014, , .		0
2640	Orion: Scaling Genomic Sequence Matching with Fine-Grained Parallelization. , 2014, , .		9
2641	A fast big data collection system using MapReduce framework. , 2014, , .		4
2642	A general framework for performance guaranteed green data center networking. , 2014, , .		11
2643	MoDisSENSE: A distributed platform for social networking services over mobile devices. , 2014, , .		3
2644	Evaluation of hadoop/mapreduce framework migration tools. , 2014, , .		4
2645	Synchronizing Small Data in a big world. , 2014, , .		0

#	ARTICLE	IF	CITATIONS
2646	ClusterJoin. Proceedings of the VLDB Endowment, 2014, 7, 1059-1070.	2.1	52
2647	A Novel Framework for Application of Cloud Computing in Wireless Mesh Networks. , 2014, , .		1
2648	Balanced resource allocations across multiple dynamic MapReduce clusters. Performance Evaluation Review, 2014, 42, 329-341.	0.4	13
2649	Collaborative Filtering of Web Service Based on MapReduce. , 2014, , .		4
2650	A Realtime Framework for Video Object Detection with Storm. , 2014, , .		5
2651	Controlling the Processing of Smart City Data in the Cloud with Domain-Specific Languages. , 2014, , .		6
2652	Extracting Trends of Traffic Congestion Using a NoSQL Database. , 2014, , .		11
2653	Using geometric structures to improve the error correction algorithm of high-throughput sequencing data on MapReduce framework. , 2014, , .		3
2654	An efficient framework for distributed computing in heterogeneous Beowulf clusters and cluster-management. , 2014, , .		5
2655	Job scheduling in Hadoop with Shared Input Policy and RAMDISK. , 2014, , .		8
2656	Big data analysis for sensor time-series in automation. , 2014, , .		11
2657	Modelling on classification and retrieval strategy in map-reduce based IR system. , 2014, , .		1
2658	Multi-touch Attribution in Online Advertising with Survival Theory. , 2014, , .		23
2659	Reservation-based Scheduling. , 2014, , .		97
2660	Study on the data integration platform of multi-regional SCADA system based on cloud computing. , 2014, , .		0
2661	Fine-grain task aggregation and coordination on GPUs. Computer Architecture News, 2014, 42, 181-192.	2.5	11
2662	Improve Parallelism of Task Execution to Optimize Utilization of MapReduce Cluster Resources. , 2014, , .		2
2663	F<sc>lexi</sc>F<sc>a</sc>CT: Scalable Flexible Factorization of Coupled Tensors on Hadoop. , 2014, , .		71

#	ARTICLE	IF	CITATIONS
2664	SLDP: A Novel Data Placement Strategy for Large-Scale Heterogeneous Hadoop Cluster. , 2014, , .		9
2665	GraphIVE: Heterogeneity-Aware Adaptive Graph Partitioning in GraphLab. , 2014, , .		6
2666	Cloud computing based localization for mobile robot systems. , 2014, , .		4
2667	freeCycles. , 2014, , .		8
2668	ASPIRE. , 2014, , .		40
2669	Parallel algorithms for geometric graph problems. , 2014, , .		76
2670	Improved GEP Algorithm for Task Scheduling in Cloud Computing. , 2014, , .		3
2671	Balanced resource allocations across multiple dynamic MapReduce clusters. , 2014, , .		17
2672	GatorCloud. , 2014, , .		2
2673	Haggis. , 2014, , .		16
2674	Partitioning strategies for spatio-textual similarity join. , 2014, , .		12
2675	Performance analysis of graph based iterative algorithms on MapReduce framework. , 2014, , .		0
2676	A hybrid mapreduce model for prolog. , 2014, , .		0
2677	Functional programming for dynamic and large data with self-adjusting computation. , 2014, , .		5
2678	Parallel in-situ data processing with speculative loading. , 2014, , .		35
2679	Using MapReduce to Compare Large Collections of Phylogenetic Trees. , 2014, , .		0
2680	MARS: Scheduling non-local tasks in mapreduce. , 2014, , .		1
2681	Towards efficient closed pattern mining from distributed multi-relational data. , 2014, , .		1

#	ARTICLE	IF	CITATIONS
2682	Multi-Q: Multiple Queries Optimization Based on MapReduce in Cloud. , 2014, , .		6
2683	GasCL: A vertex-centric graph model for GPUs. , 2014, , .		17
2684	Study in Usefulness of Middleware-Only Provenance. , 2014, , .		3
2685	Rolling window time series prediction using MapReduce. , 2014, , .		13
2686	Resilient X10. ACM SIGPLAN Notices, 2014, 49, 67-80.	0.2	8
2687	Optimization Techniques within the Hadoop Eco-system: A Survey. , 2014, , .		5
2688	Improving utilization through dynamic VM resource allocation in hybrid cloud environment. , 2014, , .		7
2689	A performance/cost model for a CUDA drug discovery application on physical and public cloud infrastructures. Concurrency Computation Practice and Experience, 2014, 26, 1787-1798.	1.4	11
2690	Evolving Big Data Stream Classification with MapReduce. , 2014, , .		8
2691	Efficient Random Network Coding for Distributed Storage Systems. Lecture Notes in Computer Science, 2014, , 385-394.	1.0	0
2692	MPI for Big Data: New tricks for an old dog. Parallel Computing, 2014, 40, 754-767.	1.3	11
2693	Computational challenges in the analysis of large, sparse, spatiotemporal data. , 2014, , .		2
2694	Automatic extraction of topics on big data streams through scalable advanced analysis. , 2014, , .		5
2695	Social relation extraction of large-scale logistics network based on mapreduce. , 2014, , .		2
2696	Columnar NoSQL CUBE: Agregation operator for columnar NoSQL data warehouse. , 2014, , .		21
2697	Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends. BioData Mining, 2014, 7, 22.	2.2	102
2698	A new approach to intelligent search in electrical engineering literature. , 2014, , .		0
2699	Addressing failures in exascale computing. International Journal of High Performance Computing Applications, 2014, 28, 129-173.	2.4	265

#	ARTICLE	IF	CITATIONS
2700	Burstiness-aware I/O scheduler for MapReduce framework on virtualized environments. , 2014, , .		1
2701	Federated MapReduce to Transparently Run Applications on Multicluster Environment. , 2014, , .		7
2702	Designing a new TCP based on FAST TCP for datacenter. , 2014, , .		3
2703	Hashdoop: A MapReduce framework for network anomaly detection. , 2014, , .		37
2704	A usage-aware scheduler for improving MapReduce performance in heterogeneous environments. , 2014, , .		7
2705	A review of adaptive approaches to MapReduce scheduling in heterogeneous environments. , 2014, , .		7
2706	The establishment and data mining of meteorological data warehouse. , 2014, , .		2
2707	A Scalable Approach to Source Camera Identification over Hadoop. , 2014, , .		14
2708	Automated design of self-adaptive software with control-theoretical formal guarantees. , 2014, , .		104
2709	Spam filtering techniques and MapReduce with SVM: A study. , 2014, , .		4
2710	Performance Characterization of Hadoop and Data MPI Based on Amdahl's Second Law. , 2014, , .		12
2711	An Architecture for Orchestrating Hadoop Applications in Hybrid Cloud. , 2014, , .		1
2712	A Framework for Amazon EC2 Bidding Strategy under SLA Constraints. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 2-11.	4.0	39
2713	DDDAS-Based Parallel Simulation of Threat Management for Urban Water Distribution Systems. Computing in Science and Engineering, 2014, 16, 8-17.	1.2	23
2714	Enabling Location-Based Services 2.0: Challenges and Opportunities. , 2014, , .		1
2715	Architecture-Awareness for Real-Time Big Data Systems. , 2014, , .		5
2716	Towards an MPI-Like Framework for the Azure Cloud Platform. , 2014, , .		3
2717	Evaluating Streaming Strategies for Event Processing Across Infrastructure Clouds. , 2014, , .		3

#	ARTICLE	IF	CITATIONS
2718	Performance of the most common non-cryptographic hash functions. Software - Practice and Experience, 2014, 44, 681-698.	2.5	28
2719	Parameterizable benchmarking framework for designing a MapReduce performance model. Concurrency Computation Practice and Experience, 2014, 26, 2005-2026.	1.4	14
2720	Skyline Query Processing over Encrypted Data. , 2014, , .		15
2721	MiCA: Real-Time Mixed Compression Scheme for Large-Scale Distributed Monitoring. , 2014, , .		1
2722	Reducing MapReduce Abstraction Costs for Text-centric Applications. , 2014, , .		1
2723	MapReduce-based web mining for prediction of web-user navigation. Journal of Information Science, 2014, 40, 557-567.	2.0	17
2724	Druid. , 2014, , .		122
2725	HPC in Big Data Age. , 2014, , .		6
2726	Contextual crowd intelligence. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2014, 16, 39-46.	3.2	13
2727	On the Locality of Java 8 Streams in Real-Time Big Data Applications. , 2014, , .		5
2728	How Many Folders Do You Really Need?. , 2014, , .		39
2729	Using extreme learning machine for intrusion detection in a big data environment. , 2014, , .		23
2730	FlexSlot: Moving Hadoop Into the Cloud with Flexible Slot Management. , 2014, , .		15
2731	Misleading Generalized Itemset Mining in the Cloud. , 2014, , .		4
2732	Towards a Community Cloud Storage. , 2014, , .		6
2733	Prediction of Rare Single-Nucleotide Causative Mutations for Muscular Diseases in Pooled Next-Generation Sequencing Experiments. Journal of Computational Biology, 2014, 21, 665-675.	0.8	4
2734	A K-means clustering with optimized initial center based on Hadoop platform. , 2014, , .		3
2735	Architectural solution for virtualized processing of big earth data. , 2014, , .		2

#	ARTICLE	IF	CITATIONS
2736	Discovering biological knowledge by integrating high-throughput data and scientific literature on the cloud. <i>Concurrency Computation Practice and Experience</i> , 2014, 26, 1771-1786.	1.4	3
2737	On the design and analysis of Data Center Network architectures for interconnecting dual-port servers. , 2014, , .		21
2738	Three-way joins on MapReduce: An experimental study. , 2014, , .		3
2739	MIMP: Deadline and Interference Aware Scheduling of Hadoop Virtual Machines. , 2014, , .		41
2740	An exploratory study of the evolution of communicated information about the execution of large software systems. <i>Journal of Software: Evolution and Process</i> , 2014, 26, 3-26.	1.2	41
2741	Replication-Based Fault-Tolerance for Large-Scale Graph Processing. , 2014, , .		37
2742	Near Real-Time Processing of Proteomics Data Using Hadoop. <i>Big Data</i> , 2014, 2, 44-49.	2.1	5
2743	Multi-agent Architecture for Real-Time Big Data Processing. , 2014, , .		41
2744	Exploiting cloud heterogeneity for optimized cost/performance MapReduce processing. , 2014, , .		18
2745	A weighted kernel possibilistic <i>c</i> -means algorithm based on cloud computing for clustering big data. <i>International Journal of Communication Systems</i> , 2014, 27, 1378-1391.	1.6	43
2746	A Wearable Device-Based Personalized Big Data Analysis Model. <i>Lecture Notes in Computer Science</i> , 2014, , 236-242.	1.0	8
2747	Virtual Network Reconfiguration for Reducing Energy Consumption in Optical Data Centers. <i>Journal of Optical Communications and Networking</i> , 2014, 6, 925.	3.3	16
2748	Disk cache-aware task scheduling for data-intensive and many-task workflow. , 2014, , .		7
2750	hatS: A Heterogeneity-Aware Tiered Storage for Hadoop. , 2014, , .		52
2751	SCADAMAR. , 2014, , .		5
2752	Smart data analytics methods for remote sensing applications. , 2014, , .		8
2753	Achieving Absolutely Optimal Block Pipelining in Organized Network Coded Gossip. , 2014, , .		1
2754	A Cloud-Based Framework for Supporting Effective and Efficient OLAP in Big Data Environments. , 2014, , .		9

#	ARTICLE	IF	CITATIONS
2755	Answering the demands of digital genomics. Concurrency Computation Practice and Experience, 2014, 26, 917-928.	1.4	2
2756	Querying Big Data: Bridging Theory and Practice. Journal of Computer Science and Technology, 2014, 29, 849-869.	0.9	23
2757	Energy efficient in-memory machine learning for data intensive image-processing by non-volatile domain-wall memory. , 2014, , .		22
2758	Accelerating Comparative Genomics Workflows in a Distributed Environment with Optimized Data Partitioning. , 2014, , .		4
2759	On the use of MapReduce to build linguistic fuzzy rule based classification systems for big data. , 2014, , .		15
2760	Logic Programming. Handbook of the History of Logic, 2014, 9, 523-569.	0.5	11
2761	Sampling based image splitting in large scale distributed computing of earth observation data. , 2014, , .		1
2762	Adaptive Resource Provisioning for the Cloud Using Online Bin Packing. IEEE Transactions on Computers, 2014, 63, 2647-2660.	2.4	178
2763	The challenges of image segmentation in big remotely sensed imagery data. Annals of GIS, 2014, 20, 233-244.	1.4	26
2764	Will They Blend?: Exploring Big Data Computation Atop Traditional HPC NAS Storage. , 2014, , .		8
2765	An energy-efficient parallel-processing method based on master-hibernating DVFS. , 2014, , .		1
2766	Equinox: Adaptive network reservation in the Cloud. , 2014, , .		2
2767	Using stop-and-wait to improve TCP throughput in fast optical switching (FOS) networks over short physical distances. , 2014, , .		6
2768	Optimizing an index with spatiotemporal patterns to support GEOSS Clearinghouse. International Journal of Geographical Information Science, 2014, 28, 1459-1481.	2.2	15
2769	MapCheckReduce: An Improved MapReduce Computing Model for Imprecise Applications. , 2014, , .		4
2770	A cloud-based service recommendation system for use in UCWW. , 2014, , .		5
2771	Research of Hadoop Task Speculative Execution Based on Process Migration. Applied Mechanics and Materials, 0, 607, 877-880.	0.2	0
2772	MapReduce delay scheduling with deadline constraint. Concurrency Computation Practice and Experience, 2014, 26, 766-778.	1.4	19

#	ARTICLE	IF	CITATIONS
2773	An approximation algorithm for client assignment in client/server systems. , 2014, , .		2
2774	Towards Fast Multimedia Feature Extraction: Hadoop or Storm. , 2014, , .		17
2775	On the performance of MapReduce: A stochastic approach. , 2014, , .		7
2776	Large-Scale Deep Belief Nets With MapReduce. IEEE Access, 2014, 2, 395-403.	2.6	47
2777	A Novel Clustering Algorithm on Large-Scale Graph Data. , 2014, , .		0
2778	Bloom filter based optimization on HBase with MapReduce. , 2014, , .		8
2779	HadoopMutator: A Cloud-Based Mutation Testing Framework. Lecture Notes in Computer Science, 2014, , 172-187.	1.0	8
2780	MapReduce Performance in Federated Cloud Computing Environments. , 2014, , 301-322.		0
2781	HaSTE: Hadoop YARN Scheduling Based on Task-Dependency and Resource-Demand. , 2014, , .		43
2782	HyRec. , 2014, , .		22
2783	An Efficient Intermediate Data Fault-Tolerance Approach in the Cloud. , 2014, , .		4
2784	Scaling up analogy-based software effort estimation: a comparison of multiple hadoop implementation schemes. , 2014, , .		1
2785	Training Large Scale Deep Neural Networks on the Intel Xeon Phi Many-Core Coprocessor. , 2014, , .		18
2786	MRTune: A simulator for performance tuning of MapReduce jobs with skewed data. , 2014, , .		2
2787	A Sensor Data Processing and Access Platform Based on Hadoop for Smart Environments. , 2014, , .		6
2788	Neighbourhood systems based knowledge acquisition using MapReduce from Big Data over cloud computing. , 2014, , .		3
2789	Big Data - Characterizing an Emerging Research Field Using Topic Models. , 2014, , .		15
2790	Hadoop and PaaS Collaborative Practice Research in Video Monitoring Platform. Applied Mechanics and Materials, 2014, 543-547, 3549-3554.	0.2	2

#	ARTICLE	IF	CITATIONS
2791	In-memory I/O and replication for HDFS with Memcached: Early experiences. , 2014, , .		9
2792	Data locality in Hadoop cluster systems. , 2014, , .		5
2793	A Parallel Method for Rough Entropy Computation Using MapReduce. , 2014, , .		2
2794	Opening the Black Box: Strategies for Increased User Involvement in Existing Algorithm Implementations. IEEE Transactions on Visualization and Computer Graphics, 2014, 20, 1643-1652.	2.9	99
2795	Network traffic analysis based on Hadoop. , 2014, , .		3
2796	Performance Study of Spindle, A Web Analytics Query Engine Implemented in Spark. , 2014, , .		2
2797	Clustering high-dimensional data via random sampling and consensus. , 2014, , .		2
2798	A scalable cloud computing deployment framework for efficient MapReduce operations using Apache YARN. , 2014, , .		1
2799	Big Data: A Strategic Perspective. International Journal of Semantic Computing, 2014, 08, 319-333.	0.4	1
2800	SDPMN: Privacy Preserving MapReduce Network Using SDN. , 2014, , .		2
2801	A parallel sampling-PSO-multi-core-K-means algorithm using mapreduce. , 2014, , .		5
2802	A Massively Parallel Processing for the Multiple Linear Regression. , 2014, , .		7
2803	A MapReduce Enabled Simulated Annealing Genetic Algorithm. , 2014, , .		1
2804	Teaching HDFS/MapReduce Systems Concepts to Undergraduates. , 2014, , .		13
2805	FlexDAS: A flexible direct attached storage for I/O intensive applications. , 2014, , .		5
2806	TIDE: Inter-chromosomal translocation and insertion detection using embeddings. , 2014, , .		3
2807	A distributed instance-weighted SVM algorithm on large-scale imbalanced datasets. , 2014, , .		11
2808	Large Imbalance Data Classification Based on MapReduce for Traffic Accident Prediction. , 2014, , .		27

#	ARTICLE	IF	CITATIONS
2809	Group tenant based SaaS application. , 2014, , .		0
2810	Enabling Cloud Computing in Emergency Management Systems. IEEE Cloud Computing, 2014, 1, 60-67.	5.3	106
2811	HTSeq-Hadoop: Extending HTSeq for Massively Parallel Sequencing Data Analysis Using Hadoop. , 2014, , .		4
2812	Space Shuffle: A Scalable, Flexible, and High-Bandwidth Data Center Network. , 2014, , .		24
2813	Optimizing the Topologies of Virtual Networks for Cloud-Based Big Data Processing. , 2014, , .		0
2814	Cutting the cord. , 2014, , .		58
2815	Evaluating GPU Passthrough in Xen for High Performance Cloud Computing. , 2014, , .		24
2816	Handling big data on agent-based modeling of Online Social Networks with MapReduce. , 2014, , .		0
2817	Scheduling MapReduce tasks on virtual MapReduce clusters from a tenant's perspective. , 2014, , .		2
2818	Big Data Management in Digital Forensics. , 2014, , .		5
2819	A cache-based method for SPARQL query processing. , 2014, , .		0
2820	Using Accumulo for Graph Twiddling. , 2014, , .		0
2821	Data anonymization through generalization using map reduce on cloud. , 2014, , .		8
2822	A Time Based Analysis of Data Processing on Hadoop Cluster. , 2014, , .		0
2823	Large-scale logistic regression and linear support vector machines using spark. , 2014, , .		44
2824	Spatial indexing and analytics on Hadoop. , 2014, , .		68
2825	Hierarchical Parallelization and Runtime Scheduling for Pregel-Like Graph Processing Systems. , 2014, , .		1
2826	On the efficiency of automatically generated accelerators for reconfigurable active SSDs. , 2014, , .		5

#	ARTICLE	IF	CITATIONS
2827	Big Data Sensing Information Processing Platform for Intelligent Traffic. Applied Mechanics and Materials, 0, 667, 324-327.	0.2	5
2828	Real-time anomaly detection over VMware performance data using storm. , 2014, , .		9
2829	Mining Unit Tests for Discovery and Migration of Math APIs. ACM Transactions on Software Engineering and Methodology, 2014, 24, 1-33.	4.8	5
2830	Research on Recommendation Algorithm of Matrix Factorization Method Based on MapReduce. Applied Mechanics and Materials, 0, 631-632, 138-141.	0.2	0
2831	MapReduce-Based Parallel Linear Regression for Face Recognition. Applied Mechanics and Materials, 2014, 556-562, 2628-2632.	0.2	1
2832	A framework for management of massive knowledge in cloud environment. , 2014, , .		2
2833	Constructing Sentiment Lexicons in Norwegian from a Large Text Corpus. , 2014, , .		11
2834	MRSMRS: Mining repetitive sequences in a MapReduce setting. , 2014, , .		2
2835	Software-Defined Cloud Computing: Architectural elements and open challenges. , 2014, , .		42
2836	Towards unified secure on- and off-line analytics at scale. Parallel Computing, 2014, 40, 738-753.	1.3	4
2837	A data reusing strategy based on hive. , 2014, , .		0
2838	A metrics based analysis of cloud resource management techniques. , 2014, , .		2
2839	Crowdsourced query augmentation through semantic discovery of domain-specific jargon. , 2014, , .		12
2840	Spores: A Type-Based Foundation for Closures in the Age of Concurrency and Distribution. Lecture Notes in Computer Science, 2014, , 308-333.	1.0	26
2841	Enabling Non-repudiable Data Possession Verification in Cloud Storage Systems. , 2014, , .		9
2842	Big Data with Cloud Computing: an insight on the computing environment, <scp>MapReduce</scp>, and programming frameworks. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2014, 4, 380-409.	4.6	175
2843	A Cloudification Methodology for Numerical Simulations. Lecture Notes in Computer Science, 2014, , 375-386.	1.0	3
2844	Dependency-Aware Data Locality for MapReduce. , 2014, , .		5

#	ARTICLE	IF	CITATIONS
2845	Improving robustness of complex networks via the effective graph resistance. European Physical Journal B, 2014, 87, 1.	0.6	63
2846	Business Intelligence and Big Data in the Cloud: Opportunities for Design-Science Researchers. Lecture Notes in Computer Science, 2014, , 75-84.	1.0	4
2847	Adopting the MapReduce framework to pre-train 1-D and 2-D protein structure predictors with large protein datasets. , 2014, , .		0
2848	Benchmarking Performance for Migrating a Relational Application to a Parallel Implementation. Lecture Notes in Computer Science, 2014, , 55-64.	1.0	1
2849	Access Control & Temperature and Humidity Systems Based on Distributed Parallel Processing. Advanced Materials Research, 0, 912-914, 1127-1130.	0.3	0
2850	Diagnosing Cloud Performance Anomalies Using Large Time Series Dataset Analysis. , 2014, , .		11
2851	DCP: An efficient and distributed data center cache protocol with Fat-Tree topology. , 2014, , .		2
2852	Bridging the gap between real world repositories and Scalable Preservation Environments. , 2014, , .		1
2853	Massive Data Hybrid Encryption Algorithm Based on Cloud Computing. Applied Mechanics and Materials, 0, 651-653, 1864-1867.	0.2	0
2854	Fault-Tolerant Dynamic Task Graph Scheduling. , 2014, , .		20
2855	Congestion integrated control in virtualized clouds. , 2014, , .		0
2856	ZeroVM: secure distributed processing for big data analytics. , 2014, , .		13
2857	Incomplete Big Data Distributed Clustering. Applied Mechanics and Materials, 2014, 687-691, 1496-1499.	0.2	1
2858	An Efficient Patent Storing Mechanism Based on SQLite on Hadoop Platform. Lecture Notes in Computer Science, 2014, , 382-392.	1.0	1
2859	BioTCM-SE: A Semantic Search Engine for the Information Retrieval of Modern Biology and Traditional Chinese Medicine. Computational and Mathematical Methods in Medicine, 2014, 2014, 1-13.	0.7	4
2860	The Application of Big Geospatial Data in Smart Agriculture. Advanced Materials Research, 0, 955-959, 3835-3839.	0.3	0
2861	An adaptive framework for utility-based optimization of scientific applications in the cloud. Journal of Cloud Computing: Advances, Systems and Applications, 2014, 3, 4.	2.1	11
2862	Smart Sensor-based Interface Control in the Mobile Cloud Environment. International Journal of Control and Automation, 2014, 7, 419-432.	0.3	0

#	ARTICLE	IF	CITATIONS
2863	Data-Intensive Workflow Optimization Based on Application Task Graph Partitioning in Heterogeneous Computing Systems. , 2014, , .		21
2864	STORE: Data recovery with approximate minimum network bandwidth and disk I/O in distributed storage systems. , 2014, , .		6
2865	Performance Analysis of R-DCN Architecture for Next Generation Web Application Integration. , 2014, , .		4
2866	Embed and Conquer: Scalable Embeddings for Kernel <i>k</i> -Means on MapReduce. , 2014, , .		6
2867	Research on MapReduce Task Dynamic Balancing Strategy Based on File Label. Applied Mechanics and Materials, 0, 571-572, 17-21.	0.2	0
2868	iMR Aproiri: A Parallel Aproiri Algorithm Based on Twister. Applied Mechanics and Materials, 0, 599-601, 1863-1867.	0.2	0
2869	Parallel Bayesian Additive Regression Trees. Journal of Computational and Graphical Statistics, 2014, 23, 830-852.	0.9	39
2870	Low frequency ambient noise variation of underwater acoustic data from MACHO hydrophone. , 2014, , .		0
2871	Distributed pattern matching and document analysis in big data using Hadoop MapReduce model. , 2014, , .		3
2872	In-Map/In-Reduce: Concurrent Job Execution in MapReduce. , 2014, , .		1
2873	CLOUDFARM: An Elastic Cloud Platform with Flexible and Adaptive Resource Management. , 2014, , .		5
2874	Random Prism: a noise-tolerant alternative to Random Forests. Expert Systems, 2014, 31, 411-420.	2.9	13
2875	Mining Mobile Internet Packets for Malware Detection. , 2014, , .		2
2876	Efficient traffic speed forecasting based on massive heterogenous historical data. , 2014, , .		15
2877	A mahout based image classification framework for very large dataset. , 2014, , .		2
2878	A combined MapReduce-windowing two-level parallel scheme for evolutionary prototype generation. , 2014, , .		5
2879	An Improvement to Feature Selection of Random Forests on Spark. , 2014, , .		6
2880	Evolutionary data reorganization for efficient workload processing. , 2014, , .		4

#	ARTICLE	IF	CITATIONS
2881	A MapReduce framework implementation for Network-on-Chip platforms. , 2014, , .		1
2882	Parallel K-Medoids clustering algorithm based on Hadoop. , 2014, , .		11
2883	Benchmarking Different MapReduce Implementations for Computer-Aided Hardware Development. , 2014, , .		0
2884	Hierarchical MapReduce: towards simplified cross-domain data processing. Concurrency Computation Practice and Experience, 2014, 26, 878-893.	1.4	5
2885	Multi-view constrained clustering with an incomplete mapping between views. Knowledge and Information Systems, 2014, 38, 231-257.	2.1	28
2886	A formal framework to analyze cost and performance in Map-Reduce based applications. Journal of Computational Science, 2014, 5, 106-118.	1.5	10
2887	A comparison of parallel large-scale knowledge acquisition using rough set theory on different MapReduce runtime systems. International Journal of Approximate Reasoning, 2014, 55, 896-907.	1.9	67
2888	Achieving Accountable MapReduce in cloud computing. Future Generation Computer Systems, 2014, 30, 1-13.	4.9	51
2889	Towards a framework for energy-aware information systems in manufacturing. Computers in Industry, 2014, 65, 419-433.	5.7	40
2890	CodeCloud: A platform to enable execution of programming models on the Clouds. Journal of Systems and Software, 2014, 93, 187-198.	3.3	17
2891	Parallel attribute reduction algorithms using MapReduce. Information Sciences, 2014, 279, 671-690.	4.0	89
2892	Lexicon expansion for latent variable grammars. Pattern Recognition Letters, 2014, 42, 47-55.	2.6	7
2893	Streaming data analytics via message passing with application to graph algorithms. Journal of Parallel and Distributed Computing, 2014, 74, 2687-2698.	2.7	12
2895	A Generate-Test-Aggregate parallel programming library for systematic parallel programming. Parallel Computing, 2014, 40, 116-135.	1.3	0
2896	A Survey of Mobile Cloud Computing Application Models. IEEE Communications Surveys and Tutorials, 2014, 16, 393-413.	24.8	548
2897	Indexing spatial data in cloud data managements. Pervasive and Mobile Computing, 2014, 15, 48-61.	2.1	25
2898	Medusa: Simplified Graph Processing on GPUs. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 1543-1552.	4.0	186
2899	A survey of large-scale analytical query processing in MapReduce. VLDB Journal, 2014, 23, 355-380.	2.7	180

#	ARTICLE	IF	CITATIONS
2900	Design and Evaluation of Network-Levitated Merge for Hadoop Acceleration. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 602-611.	4.0	17
2901	Nephele streaming: stream processing under QoS constraints at scale. Cluster Computing, 2014, 17, 61-78.	3.5	25
2902	Parallel Programming Paradigms and Frameworks in Big Data Era. International Journal of Parallel Programming, 2014, 42, 710-738.	1.1	53
2903	ELM: distributed extreme learning machine with MapReduce. World Wide Web, 2014, 17, 1189-1204.	4.7	55
2904	A tool for mapping Single Nucleotide Polymorphisms using Graphics Processing Units. BMC Bioinformatics, 2014, 15, S10.	1.2	5
2905	Sparse optimization in feature selection: application in neuroimaging. Journal of Global Optimization, 2014, 59, 439-457.	1.1	30
2906	Efficient adaptive QoS-based service selection. Service Oriented Computing and Applications, 2014, 8, 261-276.	1.3	7
2907	Towards a framework for large-scale multimedia data storage and processing on Hadoop platform. Journal of Supercomputing, 2014, 68, 488-507.	2.4	26
2908	Improvement of path analysis algorithm in social networks based on HBase. Journal of Combinatorial Optimization, 2014, 28, 588-599.	0.8	2
2909	A survey on architectures and energy efficiency in Data Center Networks. Computer Communications, 2014, 40, 1-21.	3.1	234
2910	Design of reliable virtual infrastructure with resource sharing. Computer Networks, 2014, 62, 137-151.	3.2	2
2911	Network monitoring: Present and future. Computer Networks, 2014, 65, 84-98.	3.2	74
2912	A security framework in G-Hadoop for big data computing across distributed Cloud data centres. Journal of Computer and System Sciences, 2014, 80, 994-1007.	0.9	96
2913	Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences, 2014, 275, 314-347.	4.0	2,076
2914	Towards Data Confidentiality and a Vulnerability Analysis Framework for Cloud Computing. , 2014, , 213-238.		5
2915	A new multi-objective bi-level programming model for energy and locality aware multi-job scheduling in cloud computing. Future Generation Computer Systems, 2014, 36, 91-101.	4.9	58
2916	On the use of MapReduce for imbalanced big data using Random Forest. Information Sciences, 2014, 285, 112-137.	4.0	236
2917	CPU + GPU scheduling with asymptotic profiling. Parallel Computing, 2014, 40, 107-115.	1.3	10

#	ARTICLE	IF	CITATIONS
2918	Parallel patterns for heterogeneous CPU/GPU architectures: Structured parallelism from cluster to cloud. <i>Future Generation Computer Systems</i> , 2014, 37, 354-366.	4.9	23
2919	Scalable entity-based summarization of web search results using MapReduce. <i>Distributed and Parallel Databases</i> , 2014, 32, 405-446.	1.0	16
2920	Tuple MapReduce and Pangool: an associated implementation. <i>Knowledge and Information Systems</i> , 2014, 41, 531-557.	2.1	3
2921	SmartJoin: a network-aware multiway join for MapReduce. <i>Cluster Computing</i> , 2014, 17, 629-641.	3.5	13
2922	Introduction to cloud computing: platforms and solutions. <i>Cluster Computing</i> , 2014, 17, 1225-1229.	3.5	24
2923	Information Systems for "Wicked Problems". <i>Business and Information Systems Engineering</i> , 2014, 6, 3-10.	4.0	15
2924	Fault tolerance in distributed systems using fused state machines. <i>Distributed Computing</i> , 2014, 27, 287-311.	0.7	2
2926	Querying dynamic communities in online social networks. <i>Journal of Zhejiang University: Science C</i> , 2014, 15, 81-90.	0.7	8
2927	A novel real-time scheduling algorithm and performance analysis of a MapReduce-based cloud. <i>Journal of Supercomputing</i> , 2014, 69, 739-765.	2.4	14
2928	Parallel labeling of massive XML data with MapReduce. <i>Journal of Supercomputing</i> , 2014, 67, 408-437.	2.4	12
2929	Designing a parallel evolutionary algorithm for inferring gene networks on the cloud computing environment. <i>BMC Systems Biology</i> , 2014, 8, 5.	3.0	35
2930	SPHeRe. <i>Journal of Supercomputing</i> , 2014, 68, 274-301.	2.4	11
2931	Design patterns percolating to parallel programming framework implementation. <i>International Journal of Parallel Programming</i> , 2014, 42, 1012-1031.	1.1	16
2932	In-memory, distributed content-based recommender system. <i>Journal of Intelligent Information Systems</i> , 2014, 42, 645-669.	2.8	18
2933	Engineering Energy Efficient Visual Sensor Network Applications Using Skeletons. <i>International Journal of Parallel Programming</i> , 2014, 42, 663-680.	1.1	3
2934	MapReduce framework energy adaptation via temperature awareness. <i>Cluster Computing</i> , 2014, 17, 111-127.	3.5	8
2935	An Overview of Issues and Recent Developments in Cloud Computing and Storage Security. , 2014, , 3-33.		26
2936	Chunks and Tasks: A programming model for parallelization of dynamic algorithms. <i>Parallel Computing</i> , 2014, 40, 328-343.	1.3	16

#	ARTICLE	IF	CITATIONS
2937	GreenDCN: A General Framework for Achieving Energy Efficiency in Data Center Networks. IEEE Journal on Selected Areas in Communications, 2014, 32, 4-15.	9.7	135
2938	TCP PLATO: Packet Labelling to Alleviate Time-Out. IEEE Journal on Selected Areas in Communications, 2014, 32, 65-76.	9.7	23
2939	Balancing scalability, performance and fault tolerance for structured data (BSPF). , 2014, , .		1
2940	Cost Minimization for Big Data Processing in Geo-Distributed Data Centers. IEEE Transactions on Emerging Topics in Computing, 2014, 2, 314-323.	3.2	91
2941	Cloud-Based Augmentation for Mobile Devices: Motivation, Taxonomies, and Open Challenges. IEEE Communications Surveys and Tutorials, 2014, 16, 337-368.	24.8	342
2942	Reliability-Based Design Optimization for Cloud Migration. IEEE Transactions on Services Computing, 2014, 7, 223-236.	3.2	44
2943	A framework for building hypercubes using MapReduce. Computer Physics Communications, 2014, 185, 1429-1438.	3.0	8
2944	On the Delay-Storage Trade-Off in Content Download from Coded Distributed Storage Systems. IEEE Journal on Selected Areas in Communications, 2014, 32, 989-997.	9.7	138
2945	Process Mining in the Large: A Tutorial. Lecture Notes in Business Information Processing, 2014, , 33-76.	0.8	40
2946	A novel image denoising algorithm using linear Bayesian MAP estimation based on sparse representation. Signal Processing, 2014, 100, 132-145.	2.1	32
2947	Cost-based join processing scheme in a hybrid RDBMS and hive system. , 2014, , .		1
2948	Parallel Processing of Big Data Using Power Iteration Clustering over MapReduce. , 2014, , .		4
2949	Efficient k -Means++ Approximation with MapReduce. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 3135-3144.	4.0	44
2950	High performance parallel k -means clustering for disk-resident datasets on multi-core CPUs. Journal of Supercomputing, 2014, 69, 845-863.	2.4	37
2951	Multimodal and multimedia image analysis and collaborative networking for digestive endoscopy. Irbm, 2014, 35, 88-93.	3.7	1
2952	False alarm minimization techniques in signature-based intrusion detection systems: A survey. Computer Communications, 2014, 49, 1-17.	3.1	144
2953	A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using MapReduce on Cloud. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 363-373.	4.0	185
2954	Software-as-a-service (SaaS): perspectives and challenges. Science China Information Sciences, 2014, 57, 1-15.	2.7	73

#	ARTICLE	IF	CITATIONS
2955	Towards optimization of RDF analytical queries on MapReduce. , 2014, , .		2
2956	Cognitive Internet of Things: A New Paradigm Beyond Connection. IEEE Internet of Things Journal, 2014, 1, 129-143.	5.5	434
2957	KASR: A Keyword-Aware Service Recommendation Method on MapReduce for Big Data Applications. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 3221-3231.	4.0	127
2958	SlashBurn: Graph Compression and Mining beyond Caveman Communities. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 3077-3089.	4.0	123
2959	Sierpinski triangle based data center architecture in cloud computing. Journal of Supercomputing, 2014, 69, 887-907.	2.4	48
2960	Integrated Resiliency Planning in Storage Clouds. IEEE Transactions on Network and Service Management, 2014, 11, 3-14.	3.2	6
2961	Supercomputing for the parallelization of whole genome analysis. Bioinformatics, 2014, 30, 1508-1513.	1.8	52
2962	SHadoop: Improving MapReduce performance by optimizing job execution mechanism in Hadoop clusters. Journal of Parallel and Distributed Computing, 2014, 74, 2166-2179.	2.7	80
2963	Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways. Nature Chemistry, 2014, 6, 15-21.	6.6	388
2964	JackHare: a framework for SQL to NoSQL translation using MapReduce. Automated Software Engineering, 2014, 21, 489-508.	2.2	31
2965	PCA for large data sets with parallel data summarization. Distributed and Parallel Databases, 2014, 32, 377-403.	1.0	20
2966	Big data as the new enabler in business and other intelligence. Vietnam Journal of Computer Science, 2014, 1, 3-14.	1.0	35
2967	Mobile Cloud Computing: A Survey, State of Art and Future Directions. Mobile Networks and Applications, 2014, 19, 133-143.	2.2	315
2968	Big Data: A Survey. Mobile Networks and Applications, 2014, 19, 171-209.	2.2	2,248
2969	Neural Circuits and Parallel Implementation. , 2014, , 705-725.		0
2970	Load Balancing Task Scheduling Based on Genetic Algorithm in Cloud Computing. , 2014, , .		83
2971	PAGE: A Framework for Easy PARallelization of GENomic Applications. , 2014, , .		4
2972	Enumerating Maximal Bicliques from a Large Graph Using MapReduce. , 2014, , .		8

#	ARTICLE	IF	CITATIONS
2973	Distributed malware detection based on binary file features in cloud computing environment. , 2014, , .		3
2974	Parallel health tourism information extraction and ontology storage. , 2014, , .		3
2975	Privacy and Security of Big Data. , 2014, , .		90
2976	Distributed Implementation of Latent Rating Pattern Sharing Based Cross-domain Recommender System Approach. , 2014, , .		2
2977	Optimizing mapreduce with low memory requirements for shared-memory systems. , 2014, , .		1
2978	Silver Lining: Enforcing Secure Information Flow at the Cloud Edge. , 2014, , .		9
2979	Diagnosing Virtualized Hadoop Performance from Benchmark Results: An Exploratory Study. , 2014, , .		3
2980	Towards an Evidence-Based Understanding of Emergence of Architecture through Continuous Refactoring in Agile Software Development. , 2014, , .		14
2981	A Scalable K-Anonymization Solution for Preserving Privacy in an Aging-in-Place Welfare Intercloud. , 2014, , .		5
2982	SLA guarantee real-time monitoring system with soft deadline constraint. , 2014, , .		2
2983	CDRDN: Content Driven Routing in Datacenter Network. , 2014, , .		2
2984	MapReduce join strategies for key-value storage. , 2014, , .		10
2985	FSM-H: Frequent Subgraph Mining Algorithm in Hadoop. , 2014, , .		4
2986	Practising Scalable Graph Similarity Joins in MapReduce. , 2014, , .		2
2987	Pareto-Optimal Cloud Bursting. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 2670-2682.	4.0	43
2988	Parallelizing generalized one-dimensional bin packing problem using MapReduce. , 2014, , .		1
2989	An Architecture for Cloud-Based Consumer Support Software-as-a-Service. , 2014, , .		1
2990	GPU-in-Hadoop: Enabling MapReduce across distributed heterogeneous platforms. , 2014, , .		15

#	ARTICLE	IF	CITATIONS
2991	Big data analysis and query optimization improve HadoopDB performance. , 2014, , .		3
2992	Dynamic core affinity for high-performance file upload on Hadoop Distributed File System. Parallel Computing, 2014, 40, 722-737.	1.3	8
2993	Change detection in streaming data in the era of big data. SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining, 2014, 16, 30-38.	3.2	19
2994	User behavior learning and transfer in composite social networks. ACM Transactions on Knowledge Discovery From Data, 2014, 8, 1-32.	2.5	31
2995	Afluentes Concurrent I/O Made Easy with Lazy Evaluation. , 2014, , .		1
2996	The Research of Web Parallel Information Extraction Based on Hadoop. Advances in Intelligent Systems and Computing, 2014, , 341-348.	0.5	0
2997	SICSA multicore challenge editorial preface. Concurrency Computation Practice and Experience, 2014, 26, 929-934.	1.4	0
2998	Provisioning, Placement and Pipelining Strategies for Data-Intensive Applications in Cloud Environments. , 2014, , .		5
2999	Online stream processing of machine-to-machine communications traffic: A platform comparison. , 2014, , .		3
3000	Towards unified ad-hoc data processing. , 2014, , .		5
3001	A Tool for Simple and Efficient Clinical Protocol Evaluation. , 2014, , .		1
3002	Resilient X10. , 2014, , .		32
3003	On the Organization of Cluster Voting with Massive Distributed Streams. , 2014, , .		5
3004	DETS: A Dynamic and Elastic Task Scheduler Supporting Multiple Parallel Schemes. , 2014, , .		1
3005	Parallel Haskell implementations of the <i>N-body</i> problem. Concurrency Computation Practice and Experience, 2014, 26, 987-1019.	1.4	9
3006	A novel monitoring mechanism by event trigger for Hadoop system performance analysis. Software - Practice and Experience, 2014, 44, 823-834.	2.5	4
3007	Guidelines for Selecting Hadoop Schedulers Based on System Heterogeneity. Journal of Grid Computing, 2014, 12, 499-519.	2.5	19
3008	Organizing and Storing Method for Large-Scale Unstructured Data Set with Complex Content. , 2014, , .		3

#	ARTICLE	IF	CITATIONS
3009	Cloud-Aware Development of Scientific Applications. , 2014, , .		1
3010	WOHA: Deadline-Aware Map-Reduce Workflow Scheduling Framework over Hadoop Clusters. , 2014, , .		41
3011	Combining technical trading rules using parallel particle swarm optimization based on Hadoop. , 2014, , .		11
3012	HAWQ. , 2014, , .		37
3013	Adaptive Indexing for Distributed Array Processing. , 2014, , .		1
3014	Scalable and Efficient Diagnosis for 5G Data Center Network Traffic. IEEE Access, 2014, 2, 841-855.	2.6	10
3015	Towards a User-Friendly Loading System for the Analysis of Big Data in the Internet of Things. , 2014, , .		13
3016	RABID: A Distributed Parallel R for Large Datasets. , 2014, , .		3
3017	Scalable matrix inversion using MapReduce. , 2014, , .		22
3018	BigObject Store: In-Place Computing for Interactive Analytics. , 2014, , .		1
3019	NEWT - A resilient BSP framework for Iterative algorithms on hadoop YARN. , 2014, , .		1
3020	Using client-side access partitioning for data clustering in big data applications. , 2014, , .		1
3021	Scanning Files for Signatures with a MapReduce System on Cloud Computing. , 2014, , .		0
3022	Software system for broadcasting and monitoring traffic information. , 2014, , .		1
3023	MapReduce Analysis for Cloud-Archived Data. , 2014, , .		3
3024	Temporal Event Tracing on Big Healthcare Data Analytics. , 2014, , .		25
3025	Integration of Clustering and Multidimensional Scaling to Determine Phylogenetic Trees as Spherical Phylograms Visualized in 3 Dimensions. , 2014, , .		7
3026	Data Vitalization's Perspective Towards Smart City: A Reference Model for Data Service Oriented Architecture. , 2014, , .		7

#	ARTICLE	IF	CITATIONS
3027	Vigiles: Fine-Grained Access Control for MapReduce Systems. , 2014, , .		36
3028	A Parallel Spatial Co-location Mining Algorithm Based on MapReduce. , 2014, , .		27
3029	Large-scale neural modeling in MapReduce and Giraph. , 2014, , .		0
3030	Heterogeneous cores for MapReduce processing: Opportunity or challenge?. , 2014, , .		10
3031	Near neighbor join. , 2014, , .		4
3032	Decentralized NIC-Switching Architecture Using SR-IOV PCI Express Network Device. IEEE Micro, 2014, 34, 42-50.	1.8	16
3033	ScalScheduling: A Scalable Scheduling Architecture for MPI-based interactive analysis programs. , 2014, , .		2
3034	A Survey of Cloud-Based Service Computing Solutions for Mammalian Genomics. IEEE Transactions on Services Computing, 2014, 7, 726-740.	3.2	10
3035	Obtaining Technology Insights from Large and Heterogeneous Document Collections. , 2014, , .		3
3036	Contextual Anomaly Detection in Big Sensor Data. , 2014, , .		52
3037	From news to facts: An Hadoop-based social graphs analysis. , 2014, , .		1
3038	Big data analysis: Trends & challenges. , 2014, , .		5
3039	An Optimal Task Selection Scheme for Hadoop Scheduling. IERI Procedia, 2014, 10, 70-75.	0.3	14
3040	A MapReduce based parallel SVM for large-scale predicting protein-protein interactions. Neurocomputing, 2014, 145, 37-43.	3.5	109
3041	Ubicon and its applications for ubiquitous social computing. New Review of Hypermedia and Multimedia, 2014, 20, 53-77.	0.9	35
3042	Elastic MapReduce Execution. , 2014, , .		3
3043	DBStream: An online aggregation, filtering and processing system for network traffic monitoring. , 2014, , .		17
3044	BigGIS: How big data can shape next-generation GIS. , 2014, , .		19

#	ARTICLE	IF	CITATIONS
3045	Quatrain: Accelerating Data Aggregation between Multiple Layers. IEEE Transactions on Computers, 2014, 63, 1207-1219.	2.4	1
3046	Impact of MapReduce Task Re-execution Policy on Job Completion Reliability and Job Completion Time. , 2014, , .		2
3047	A Heuristic Speculative Execution Strategy in Heterogeneous Distributed Environments. , 2014, , .		10
3048	Big R: Large-Scale Analytics on Hadoop Using R. , 2014, , .		5
3049	Dynamic Scheduling for Speculative Execution to Improve MapReduce Performance in Heterogeneous Environment. , 2014, , .		3
3050	Advancing M2M communications management: A cloud-based system for cellular traffic analysis. , 2014, , .		4
3051	TideWatch: Fingerprinting the cyclicity of big data workloads. , 2014, , .		8
3052	Online algorithms for uploading deferrable big data to the cloud. , 2014, , .		17
3053	ELTA: New Approach in Designing Business Intelligence Solutions in Era of Big Data. Procedia Technology, 2014, 16, 667-674.	1.1	39
3054	Counting Triangles in Massive Graphs with MapReduce. SIAM Journal of Scientific Computing, 2014, 36, S48-S77.	1.3	61
3055	Robust scheduling and congestion control for flexible queueing networks. , 2014, , .		4
3056	An automated bot detection system through honeypots for large-scale. , 2014, , .		8
3057	Capturing Long-Tail Distributions of Object Subcategories. , 2014, , .		75
3058	Fine-grain task aggregation and coordination on GPUs. , 2014, , .		16
3059	Optimal approximation algorithm of virtual machine placement for data latency minimization in cloud systems. , 2014, , .		42
3060	A Real-Time Anomalies Detection System Based on Streaming Technology. , 2014, , .		6
3061	Management of Big Data in the Internet of Things in Agriculture Based on Cloud Computing. Applied Mechanics and Materials, 0, 548-549, 1438-1444.	0.2	20
3062	Study of an Improved Hadoop Speculative Execution Algorithm. Applied Mechanics and Materials, 0, 513-517, 2281-2284.	0.2	1

#	ARTICLE	IF	CITATIONS
3063	The research of recommendation system based on Hadoop cloud platform. , 2014, , .		2
3064	Big Data in the Public Cloud. IEEE Cloud Computing, 2014, 1, 13-15.	5.3	7
3065	Hippo: An enhancement of pipeline-aware in-memory caching for HDFS. , 2014, , .		5
3066	Correlation Aware Technique for SQL to NoSQL Transformation. , 2014, , .		15
3067	Toward Detecting Compromised MapReduce Workers through Log Analysis. , 2014, , .		14
3068	Towards a Collective Layer in the Big Data Stack. , 2014, , .		6
3069	A Google approach for computational intelligence in big data. , 2014, , .		2
3070	Towards a big data exploration framework for astronomical archives. , 2014, , .		4
3071	Applying MapReduce Programming Model for Handling Scientific Problems. , 2014, , .		0
3072	Characterization and Optimization of Memory-Resident MapReduce on HPC Systems. , 2014, , .		40
3073	Auto-scaling techniques for elastic data stream processing. , 2014, , .		40
3074	Scalable community detection from networks by computing edge betweenness on MapReduce. , 2014, , .		9
3075	Processing universal quantification queries using MapReduce. , 2014, , .		2
3076	A parallel algorithm for approximate frequent itemset mining using MapReduce. , 2014, , .		5
3077	Comparing Implementations of Near-Data Computing with In-Memory MapReduce Workloads. IEEE Micro, 2014, 34, 44-52.	1.8	99
3078	Relational on demand data management for IT-services. , 2014, , .		1
3079	ToMaR – A Data Generator for Large Volumes of Content. , 2014, , .		1
3080	Speculative Software Modification and its Use in Securing SOUP. , 2014, , .		2

#	ARTICLE	IF	CITATIONS
3081	An Integrated Processing Platform for Traffic Sensor Data and Its Applications in Intelligent Transportation Systems. , 2014, , .		6
3082	A virtual machine based task scheduling approach to improving data locality for virtualized Hadoop. , 2014, , .		4
3083	Toward Scalable Systems for Big Data Analytics: A Technology Tutorial. IEEE Access, 2014, 2, 652-687.	2.6	965
3084	DynamicMR: A Dynamic Slot Allocation Optimization Framework for MapReduce Clusters. IEEE Transactions on Cloud Computing, 2014, 2, 333-347.	3.1	57
3085	A MapReduce Technique to Mosaic Continental-Scale Weather Radar Data in Real-Time. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7, 721-732.	2.3	19
3086	Towards performance-centric fairness in datacenter networks. , 2014, , .		19
3087	Dynamic assembly simplification for virtual assembly process of complex product in cloud computing environment. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2014, 228, 1198-1213.	1.5	5
3088	Exploring mobile data on smartphones from collection to analysis. , 2014, , .		2
3089	The Stream Software Connector Design Space: Frameworks and Languages for Distributed Stream Processing. , 2014, , .		8
3090	Abnormality analysis of streamed log data. , 2014, , .		7
3091	MapReduce Design Patterns for Social Networking Analysis. , 2014, , .		7
3092	Process data analytics in the era of big data. AIChE Journal, 2014, 60, 3092-3100.	1.8	309
3093	Big Data Applications Using Workflows for Data Parallel Computing. Computing in Science and Engineering, 2014, 16, 11-21.	1.2	30
3094	Data center network architecture in cloud computing: review, taxonomy, and open research issues. Journal of Zhejiang University: Science C, 2014, 15, 776-793.	0.7	24
3095	Efficient Response Time Approximations for Multiclass Fork and Join Queues in Open and Closed Queuing Networks. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 1437-1446.	4.0	23
3096	Cloud-based Manufacturing: Old Wine in New Bottles?. Procedia CIRP, 2014, 17, 94-99.	1.0	52
3097	GEA: A Goal-Driven Approach to Discovering Early Aspects. IEEE Transactions on Software Engineering, 2014, 40, 584-602.	4.3	4
3098	Advanced communication systems for enhanced big data technology and applications. International Journal of Communication Systems, 2014, 27, 825-827.	1.6	3

#	ARTICLE	IF	CITATIONS
3099	MR-tree: an efficient index for MapReduce. International Journal of Communication Systems, 2014, 27, 828-838.	1.6	6
3100	Mapping brain activity at scale with cluster computing. Nature Methods, 2014, 11, 941-950.	9.0	257
3101	A Dynamic Data Placement Strategy for Hadoop in Heterogeneous Environments. Big Data Research, 2014, 1, 14-22.	2.6	58
3102	Efficient Indexing and Query Processing of Model-View Sensor Data in the Cloud. Big Data Research, 2014, 1, 52-65.	2.6	24
3103	Cloud Computing Survey. Applied Mechanics and Materials, 0, 530-531, 650-661.	0.2	32
3104	Parallel Randomly Compressed Cubes : A scalable distributed architecture for big tensor decomposition. IEEE Signal Processing Magazine, 2014, 31, 57-70.	4.6	53
3105	An Intelligent Touring System Based on Mobile Social Network and Cloud Computing for Travel Recommendation. , 2014, , .		4
3106	Survey of real-time processing systems for big data. , 2014, , .		78
3107	A Performance Analysis of MapReduce Task with Large Number of Files Dataset in Big Data Using Hadoop. , 2014, , .		12
3108	Semantic analysis of web documents for the generation of optimal content. Engineering Applications of Artificial Intelligence, 2014, 35, 114-130.	4.3	23
3109	Parallelizing Fully Homomorphic Encryption. , 2014, , .		4
3110	High productivity data processing analytics methods with applications. , 2014, , .		4
3111	Power Grid Data Analysis with R and Hadoop. , 2014, , 1-34.		1
3112	Monitoring and analyzing big traffic data of a large-scale cellular network with Hadoop. IEEE Network, 2014, 28, 32-39.	4.9	117
3113	Toward integrating overlay and physical networks for robust parallel processing architecture. IEEE Network, 2014, 28, 40-45.	4.9	6
3114	Multi-threading based Map Reduce tasks scheduling. , 2014, , .		11
3115	A comprehensive view of Hadoop research—A systematic literature review. Journal of Network and Computer Applications, 2014, 46, 1-25.	5.8	106
3116	ANTELOPE: A Semantic-Aware Data Cube Scheme for Cloud Data Center Networks. IEEE Transactions on Computers, 2014, 63, 2146-2159.	2.4	18

#	ARTICLE	IF	CITATIONS
3117	Big Data Frequent Pattern Mining. , 2014, , 225-259.		21
3118	Analyzing neural data at huge scale. Nature Methods, 2014, 11, 911-912.	9.0	4
3119	Challenges for MapReduce in Big Data. , 2014, , .		97
3120	T-Storm: Traffic-Aware Online Scheduling in Storm. , 2014, , .		149
3121	Intelligent Data Analysis and its challenges in big data environment. , 2014, , .		5
3122	Cloud Computing education strategies. , 2014, , .		18
3123	Improving the performance of Hadoop Hive by sharing scan and computation tasks. Journal of Cloud Computing: Advances, Systems and Applications, 2014, 3, .	2.1	18
3124	Universal Cross-Cloud Communication. IEEE Transactions on Cloud Computing, 2014, 2, 103-116.	3.1	9
3125	Exploiting Hadoop Topology in Virtualized Environments. , 2014, , .		0
3126	Reducing the Search Space for Big Data Mining for Interesting Patterns from Uncertain Data. , 2014, , .		45
3127	MRPrePost—A parallel algorithm adapted for mining big data. , 2014, , .		8
3128	Cost-Based Optimization of Logical Partitions for a Query Workload in a Hadoop Data Warehouse. Lecture Notes in Computer Science, 2014, , 559-567.	1.0	2
3129	Map-reduce processing of k-means algorithm with FPGA-accelerated computer cluster. , 2014, , .		28
3130	Graph-Based Semi-Supervised Learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2014, 8, 1-125.	0.6	60
3131	Asyler: tolerating computational skew of synchronous iterative applications via computing decomposition. Knowledge and Information Systems, 2014, 41, 379-400.	2.1	3
3132	A dynamic block device reconfiguration algorithm in virtual MapReduce cluster. Cluster Computing, 2014, 17, 1171-1183.	3.5	2
3133	CloudDMSS: robust Hadoop-based multimedia streaming service architecture for a cloud computing environment. Cluster Computing, 2014, 17, 605-628.	3.5	16
3134	The SAMS: Smartphone Addiction Management System and Verification. Journal of Medical Systems, 2014, 38, 1.	2.2	491

#	ARTICLE	IF	CITATIONS
3135	Extending MapReduce across Clouds with BStream. IEEE Transactions on Cloud Computing, 2014, 2, 362-376.	3.1	21
3136	Computational Intelligence Challenges and Applications on Large-Scale Astronomical Time Series Databases. IEEE Computational Intelligence Magazine, 2014, 9, 27-39.	3.4	55
3137	Bwasw-Cloud: Efficient sequence alignment algorithm for two big data with MapReduce. , 2014, , .		4
3138	Budget-Driven Scheduling Algorithms for Batches of MapReduce Jobs in Heterogeneous Clouds. IEEE Transactions on Cloud Computing, 2014, 2, 306-319.	3.1	71
3139	An Overlay-Based Data Mining Architecture Tolerant to Physical Network Disruptions. IEEE Transactions on Emerging Topics in Computing, 2014, 2, 292-301.	3.2	8
3140	Energy-saving cloud computing platform based on micro-embedded system. , 2014, , .		2
3141	Speeding-up codon analysis on the cloud with local MapReduce aggregation. Information Sciences, 2014, 263, 175-185.	4.0	16
3142	Scalable hybrid stream and hadoop network analysis system. , 2014, , .		20
3143	FENNEL. , 2014, , .		227
3144	Towards Model-Driven Engineering for Big Data Analytics -- An Exploratory Analysis of Domain-Specific Languages for Machine Learning. , 2014, , .		19
3145	Workload characteristics of DNA sequence analysis. , 2014, , .		1
3146	Think Big with Big Data: Identifying Suitable Big Data Strategies in Corporate Environments. , 2014, , .		29
3147	Optimal metadata replications and request balancing strategy on cloud data centers. Journal of Parallel and Distributed Computing, 2014, 74, 2934-2940.	2.7	24
3148	Real-Time Tasks Oriented Energy-Aware Scheduling in Virtualized Clouds. IEEE Transactions on Cloud Computing, 2014, 2, 168-180.	3.1	181
3149	Cloud computing for geodetic imaging data processing, analysis, and modeling. , 2014, , .		0
3150	Sharing across Multiple MapReduce Jobs. ACM Transactions on Database Systems, 2014, 39, 1-46.	1.5	7
3151	Efficient way of searching data in MapReduce paradigm. , 2014, , .		1
3152	A personalized IPTV channel-recommendation mechanism based on the MapReduce framework. Journal of Supercomputing, 2014, 69, 225-247.	2.4	9

#	ARTICLE	IF	CITATIONS
3153	Strategic and suave processing for performing similarity joins using MapReduce. Journal of Supercomputing, 2014, 69, 930-954.	2.4	1
3154	Optimized big data K-means clustering using MapReduce. Journal of Supercomputing, 2014, 70, 1249-1259.	2.4	125
3155	Cloud-assisted analysis for energy efficiency in intelligent video systems. Journal of Supercomputing, 2014, 70, 1345-1364.	2.4	6
3156	Peacock: a customizable MapReduce for multicore platform. Journal of Supercomputing, 2014, 70, 1496-1513.	2.4	1
3157	CloudThinking as an Intelligent Infrastructure for Mobile Robotics. Wireless Personal Communications, 2014, 76, 231-244.	1.8	15
3158	FIGG: Simulating populations of whole genome sequences for heterogeneous data analyses. BMC Bioinformatics, 2014, 15, 149.	1.2	9
3159	Resource Allocation and Scheduling in Cloud Computing: Policy and Algorithm. IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India), 2014, 31, 4-16.	2.1	40
3160	Scholarly big data information extraction and integration in the CiteSeer<sup>χ</sup> digital library. , 2014, , .		16
3161	Federation in Cloud Data Management: Challenges and Opportunities. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 1670-1678.	4.0	17
3162	Finding nearest facility location with open box query using Geohashing and MapReduce. , 2014, , .		0
3163	Workload-Aware Provisioning in Public Clouds. IEEE Internet Computing, 2014, 18, 15-21.	3.2	14
3164	A highly parallelized control system platform architecture using multicore CPU and FPGA for multi-DoF robots. , 2014, , .		7
3165	Evaluating MapReduce frameworks for iterative Scientific Computing applications. , 2014, , .		15
3166	Rule Based Classification on a Multi Node Scalable Hadoop Cluster. Lecture Notes in Computer Science, 2014, , 174-183.	1.0	4
3167	DataMPI: Extending MPI to Hadoop-Like Big Data Computing. , 2014, , .		42
3168	Cloud computing in e-Science: research challenges and opportunities. Journal of Supercomputing, 2014, 70, 408-464.	2.4	34
3169	Profiling and evaluating hardware choices for MapReduce environments: An application-aware approach. Performance Evaluation, 2014, 79, 328-344.	0.9	14
3170	Hybrid address spaces: A methodology for implementing scalable high-level programming models on non-coherent many-core architectures. Journal of Systems and Software, 2014, 97, 47-64.	3.3	2

#	ARTICLE	IF	CITATIONS
3171	A vectorial DEVS extension for large scale system modeling and parallel simulation. Simulation, 2014, 90, 522-546.	1.1	16
3172	New Scheduling Algorithm in Hadoop Based on Resource Aware. Advances in Intelligent Systems and Computing, 2014, , 1011-1020.	0.5	0
3173	Modeling and Optimization for Big Data Analytics: (Statistical) learning tools for our era of data deluge. IEEE Signal Processing Magazine, 2014, 31, 18-31.	4.6	185
3174	An event processing approach to text stream analysis. , 2014, , .		2
3175	Performance models and dynamic characteristics analysis for HDFS write and read operations: A systematic view. Journal of Systems and Software, 2014, 93, 132-151.	3.3	17
3176	MARIANE: Using MAppReduce in HPC environments. Future Generation Computer Systems, 2014, 36, 379-388.	4.9	14
3177	Temporal representation for mining scientific data provenance. Future Generation Computer Systems, 2014, 36, 363-378.	4.9	34
3178	Dynamic tuning of the workload partition factor and the resource utilization in data-intensive applications. Future Generation Computer Systems, 2014, 37, 162-177.	4.9	4
3179	VirtualKnotter: Online virtual machine shuffling for congestion resolving in virtualized datacenter. Computer Networks, 2014, 67, 141-153.	3.2	19
3180	X10-FT: Transparent fault tolerance for APGAS language and runtime. Parallel Computing, 2014, 40, 136-156.	1.3	1
3181	Benchmarking MapReduce implementations under different application scenarios. Future Generation Computer Systems, 2014, 36, 389-399.	4.9	24
3182	Reliable scalable symbolic computation: The design of SymGridPar2. Computer Languages, Systems and Structures, 2014, 40, 19-35.	1.4	7
3183	Towards minimizing disk I/O contention: A partitioned file assignment approach. Future Generation Computer Systems, 2014, 37, 178-190.	4.9	7
3184	Sensor data compression based on MapReduce. Journal of China Universities of Posts and Telecommunications, 2014, 21, 60-66.	0.8	5
3185	A spatiotemporal compression based approach for efficient big data processing on Cloud. Journal of Computer and System Sciences, 2014, 80, 1563-1583.	0.9	67
3186	Partial-update dimensionality reduction for accumulating co-occurrence events. Pattern Recognition Letters, 2014, 36, 62-73.	2.6	1
3187	Incremental Detection of Inconsistencies in Distributed Data. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 1367-1383.	4.0	22
3188	Managing Performance Overhead of Virtual Machines in Cloud Computing: A Survey, State of the Art, and Future Directions. Proceedings of the IEEE, 2014, 102, 11-31.	16.4	215

#	ARTICLE	IF	CITATIONS
3189	A Parallel Algorithm for the Best k-Mismatches Alignment Problem. , 2014, , .		0
3190	Towards software performance engineering for multicore and manycore systems. Performance Evaluation Review, 2014, 41, 2-11.	0.4	8
3191	Twoâ€phase groupingâ€based resource management for big data processing in mobile cloud computing. International Journal of Communication Systems, 2014, 27, 839-851.	1.6	19
3192	Scalable graph clustering with parallel approximate PageRank. Social Network Analysis and Mining, 2014, 4, 1.	1.9	6
3193	Cloud computing simulators: A detailed survey and future direction. , 2014, , .		53
3194	LSShare: an efficient multiple query optimization system in the cloud. Distributed and Parallel Databases, 2014, 32, 583-605.	1.0	10
3195	Scalable multimedia content analysis on parallel platforms using python. ACM Transactions on Multimedia Computing, Communications and Applications, 2014, 10, 1-22.	3.0	38
3196	Identifying Information Requirement for Scheduling Kepler Workflow in the Cloud. Procedia Computer Science, 2014, 29, 1762-1769.	1.2	2
3197	Efficient Data Structures for Risk Modelling in Portfolios of Catastrophic Risk Using MapReduce. Procedia Computer Science, 2014, 29, 1557-1568.	1.2	1
3198	Scalable nearest neighbor query processing based on Inverted Grid Index. Journal of Network and Computer Applications, 2014, 44, 172-182.	5.8	14
3199	Formal verification problems in a big data world: towards a mighty synergy. , 2014, , .		14
3200	Dynamic workload management in heterogeneous Cloud computing environments. , 2014, , .		13
3201	An Empirical Study of Hadoop's Energy Efficiency on a HPC Cluster. Procedia Computer Science, 2014, 29, 62-72.	1.2	15
3202	Privacy preserving growing neural gas over arbitrarily partitioned data. Neurocomputing, 2014, 144, 427-435.	3.5	9
3203	A distributed approach for large-scale classifier training and image classification. Neurocomputing, 2014, 144, 304-317.	3.5	11
3204	Productivity Frameworks in Big Data Image Processing Computations - Creating Photographic Mosaics with Hadoop and Scalding. Procedia Computer Science, 2014, 29, 2306-2314.	1.2	8
3205	Masterâ€worker model for MapReduce paradigm on the TILE64 many-core platform. Future Generation Computer Systems, 2014, 36, 19-30.	4.9	5
3206	A verified algebra for readâ€write Linked Data. Science of Computer Programming, 2014, 89, 2-22.	1.5	5

#	ARTICLE	IF	CITATIONS
3207	Accelerating MapReduce framework on multi-GPU systems. Cluster Computing, 2014, 17, 293-301.	3.5	15
3208	VM consolidation: A real case based on OpenStack Cloud. Future Generation Computer Systems, 2014, 32, 118-127.	4.9	191
3209	Improving MapReduce Performance Using Smart Speculative Execution Strategy. IEEE Transactions on Computers, 2014, 63, 954-967.	2.4	139
3210	Security and privacy for storage and computation in cloud computing. Information Sciences, 2014, 258, 371-386.	4.0	405
3211	Research on Distributed Data Stream Mining in Internet of Things. , 0, , .		0
3212	Building an Experimental Platform for Cloud and Big Data Education. , 2014, , .		1
3213	Scalable Data Mining, Archiving, and Big Data Management for the Next Generation Astronomical Telescopes. Advances in Data Mining and Database Management Book Series, 0, , 196-221.	0.4	4
3214	The ARCOMEM Architecture for Social- and Semantic-Driven Web Archiving. Future Internet, 2014, 6, 688-716.	2.4	20
3215	Once you SCOOP, no need to fork. , 2014, , .		38
3216	Indoor air quality analysis based on Hadoop. IOP Conference Series: Earth and Environmental Science, 2014, 17, 012260.	0.2	0
3217	Online Data Processing on S4 Engine: A Study Case on Natural Disasters. , 2014, , .		0
3218	A Control Approach for Performance of Big Data Systems. IFAC Postprint Volumes IPPV / International Federation of Automatic Control, 2014, 47, 152-157.	0.4	18
3219	Performance Modeling of Computation and Communication Tradeoffs in Vertex-Centric Graph Processing Clusters. , 2014, , .		3
3220	Scaling up genome annotation using MAKER and work queue. International Journal of Bioinformatics Research and Applications, 2014, 10, 447.	0.1	11
3221	E-learning and educational data mining in cloud computing: an overview. International Journal of Learning Technology, 2014, 9, 25.	0.2	33
3222	A Survey on Parallel Computing and its Applications in Data-Parallel Problems Using GPU Architectures. Communications in Computational Physics, 2014, 15, 285-329.	0.7	126
3223	ParaLite: A Parallel Database System for Data-Intensive Workflows. IEICE Transactions on Information and Systems, 2014, E97.D, 1211-1224.	0.4	0
3224	NOMAD. Proceedings of the VLDB Endowment, 2014, 7, 975-986.	2.1	75

#	ARTICLE	IF	CITATIONS
3225	Efficient K-Nearest Neighbor Graph Construction Using MapReduce for Large-Scale Data Sets. IEICE Transactions on Information and Systems, 2014, E97.D, 3142-3154.	0.4	6
3226	Next Generation Distributed Computing for Cancer Research. Cancer Informatics, 2014, 13s7, CIN.S16344.	0.9	7
3227	Parallel algorithms for clustering biological graphs on distributed and shared memory architectures. International Journal of High Performance Computing and Networking, 2014, 7, 241.	0.4	21
3228	Fault-Tolerant Streaming Computation with BlockMon. , 2014, , .		0
3229	Parallel Post-Processing with MPI-Bash. , 2014, , .		0
3230	Waiting for Godot? the right language abstractions for parallel programming should be here soon. Ubiquity, 2014, 2014, 1-12.	0.2	1
3231	Utility-Based VM Assignment in Cloud Scenarios. , 2014, , .		0
3232	SWAP-Assembler: scalable and efficient genome assembly towards thousands of cores. BMC Bioinformatics, 2014, 15, S2.	1.2	35
3233	Operational real-time GPS-enhanced earthquake early warning. Journal of Geophysical Research: Solid Earth, 2014, 119, 7944-7965.	1.4	100
3234	Early Prediction of the Cost of HPC Application Execution in the Cloud. , 2014, , .		3
3235	Cress: Dynamic Scheduling for Resource Constrained Jobs. , 2014, , .		3
3236	Incast-Aware Switch-Assisted TCP Congestion Control for Data Centers. , 2014, , .		1
3237	The Problem with Data. , 2014, , .		3
3238	Efficient filtering processes for machine-to-machine data based on automation modules and data-agnostic algorithms. International Journal of Business Process Integration and Management, 2014, 7, 73.	0.2	3
3239	Joint optimization of overlapping phases in MapReduce. Performance Evaluation Review, 2014, 41, 16-18.	0.4	11
3240	Using Distributed Analytics to Enable Real-Time Exploration of Discrete Event Simulations. , 2014, , .		4
3241	Incentive-Driven QoS for Ephemeral Virtual Clouds. , 2014, , .		0
3242	A MapReduce based parallel algorithm for CIM data verification. , 2014, , .		1

#	ARTICLE	IF	CITATIONS
3243	Scalable Mobile Data Streaming with Trajectory Preserving Partitioning. , 2014, , .		3
3244	Novel approach for perception analysis in a learning environment. , 2014, , .		1
3245	Dust: Real-Time Code Offloading System for Wearable Computing. , 2014, , .		0
3246	Watershed reengineering: Making Streams Programmable. , 2014, , .		1
3247	Toward Optimal Resource Provisioning for Cloud MapReduce and Hybrid Cloud Applications. , 2014, , .		8
3248	Improving Hadoop Monetary Efficiency in the Cloud Using Spot Instances. , 2014, , .		6
3249	Resource Centered Computing Delivering High Parallel Performance. , 2014, , .		0
3250	The k-Nearest Neighbor Algorithm Using MapReduce Paradigm. , 2014, , .		24
3251	Managing the academic data lifecycle: A case study of HPCC. , 2014, , .		4
3252	A Reconfigurable MapReduce accelerator for multi-core all-programmable SoCs. , 2014, , .		6
3253	NDN-Q: An NDN query mechanism for efficient V2X data collection. , 2014, , .		17
3254	ATLAS: Big Data in a Small Package?. Proceedings of the International Astronomical Union, 2015, 10, 299-305.	0.0	0
3255	Scalability of Stochastic Gradient Descent based on "Smart" Sampling Techniques. Procedia Computer Science, 2015, 53, 308-315.	1.2	3
3256	Pervasive Data Processing. Procedia Computer Science, 2015, 63, 16-23.	1.2	0
3257	Lessons from the Congested Clique applied to MapReduce. Theoretical Computer Science, 2015, 608, 268-281.	0.5	33
3258	Federated Big Data for Resource Aggregation and Load Balancing with DIRAC. Procedia Computer Science, 2015, 51, 2769-2773.	1.2	3
3259	SparkRDF: Elastic Discreted RDF Graph Processing Engine With Distributed Memory. , 2015, , .		16
3260	Scalable multipartite subgraph enumeration for integrative analysis of heterogeneous experimental functional genomics data. , 2015, , .		3

#	ARTICLE	IF	CITATIONS
3261	Assessing Process Discovery Scalability in Data Intensive Environments. , 2015, , .		8
3262	High performance CDR processing with MapReduce. , 2015, , .		0
3263	Providing anonymity using top down specialization on Big Data using hadoop framework. , 2015, , .		0
3264	Optimal VM placement in data centres with architectural and resource constraints. International Journal of Autonomous and Adaptive Communications Systems, 2015, 8, 392.	0.2	9
3265	Ontology-based metadata matching for emergency decision-making using MapReduce. International Journal of Metadata, Semantics and Ontologies, 2015, 10, 75.	0.2	0
3266	Introducing extreme data storage middleware of schema-free document stores using MapReduce. International Journal of Ad Hoc and Ubiquitous Computing, 2015, 20, 274.	0.3	10
3267	Fault-Tolerant Streaming Computation with BlockMon. , 2015, , .		1
3268	Conformal Prediction in Spark: Large-Scale Machine Learning with Confidence. , 2015, , .		5
3269	NBC: An Efficient Hierarchical Clustering Algorithm for Large Datasets. International Journal of Semantic Computing, 2015, 09, 307-331.	0.4	4
3270	Crowd Fraud Detection in Internet Advertising. , 2015, , .		43
3271	VAT: A System for Visualizing, Analyzing and Transforming Spatial Data in Science. Datenbank-Spektrum, 2015, 15, 175-184.	1.2	10
3272	Range-Based Clustering Supporting Similarity Search in Big Data. , 2015, , .		2
3273	MapReduce: A critical analysis of existing authentication methods. , 2015, , .		1
3274	Reducing Job Slowdown Variability for Data-Intensive Workloads. , 2015, , .		1
3275	Parallelizing user-defined aggregations using symbolic execution. , 2015, , .		36
3276	Dynamic reduce task adjustment for hadoop workloads. , 2015, , .		3
3277	A document-based data model for large scale computational maritime situational awareness. , 2015, , .		9
3278	How Hard Can It Be?: Understanding the Complexity of Replica Aware Virtual Cluster Embeddings. , 2015, , .		4

#	ARTICLE	IF	CITATIONS
3279	Parallel genome-wide analysis with central and graphic processing units. , 2015, , .		1
3280	Scaling Computation on GPUs Using Powerlists. , 2015, , .		4
3281	SciSpark: Applying in-memory distributed computing to weather event detection and tracking. , 2015, , .		31
3282	Network-aware resource management for scalable data analytics frameworks. , 2015, , .		7
3283	A prototype of healthcare big data processing system based on Spark. , 2015, , .		6
3284	Scalable extraction of timeline information from road traffic data using MapReduce. , 2015, , .		2
3285	Bridging the gap for retrieving DBpedia data. , 2015, , .		1
3286	Parallel knowledge acquisition algorithm for big data using MapReduce. , 2015, , .		0
3287	Optimised data structures for large scale content-based geo-indexing. , 2015, , .		0
3288	Toward Scalable Emulation of Future Internet Applications with Simulation Symbiosis. , 2015, , .		2
3289	SLA-Aware Energy-Efficient Scheduling Scheme for Hadoop YARN. , 2015, , .		7
3290	Hardware accelerator design for data centers. , 2015, , .		4
3291	Design and Implementation of Clinical Data Center Based on Hadoop. , 2015, , .		0
3292	CIRRUS: Increased image dissemination speed using cloud resources. , 2015, , .		0
3293	Efficient large scale distributed matrix computation with spark. , 2015, , .		15
3294	Drug discovery for breast cancer based on big data analytics techniques. , 2015, , .		2
3295	A New Approach for Scholars Matching Using Universal Quantifier Queries. , 2015, , .		1
3296	Spark-based Monte Carlo parallel simulation approach for FTA. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
3297	Scalable and reliable monitoring for power systems. , 2015, , .		1
3298	Inferring Risk Aversion for Decentralized Investment Portfolios. , 2015, , .		0
3299	Balancing Thread-Level and Task-Level Parallelism for Data-Intensive Workloads on Clusters and Clouds. , 2015, , .		2
3300	FP-Hadoop. Proceedings of the VLDB Endowment, 2015, 8, 1856-1859.	2.1	6
3301	High-performance computing: to boldly go where no human has gone before. Concurrency Computation Practice and Experience, 2015, 27, 3145-3165.	1.4	7
3302	XHAMI – Extended HDFS and MapReduce Interface for Image Processing Applications. , 2015, , .		3
3303	A novel data security framework using E-MOD for big data. , 2015, , .		6
3304	IOSIG+: On the Role of I/O Tracing and Analysis for Hadoop Systems. , 2015, , .		1
3305	Isolates, channels, and event streams for composable distributed programming. , 2015, , .		20
3306	An improved Fuzzy C-Means algorithm based on MapReduce. , 2015, , .		10
3307	Top-k computations in MapReduce: A case study on recommendations. , 2015, , .		1
3308	A Case Study of Optimizing Big Data Analytical Stacks Using Structured Data Shuffling. , 2015, , .		1
3309	Parallel algorithms for large scale constrained tensor decomposition. , 2015, , .		3
3310	An NMF Perspective on Binary Hashing. , 2015, , .		7
3311	Spatial data processing with MapReduce. , 2015, , .		1
3312	Study and Implementation of Elastic Stream Computing In The Cloud. , 2015, , .		0
3313	Exploiting large system dynamics for designing simple data center schedulers. , 2015, , .		3
3314	A Resilient Framework for Iterative Linear Algebra Applications in X10. , 2015, , .		6

#	ARTICLE	IF	CITATIONS
3315	Hunting attacks in the dark: clustering and correlation analysis for unsupervised anomaly detection. International Journal of Network Management, 2015, 25, 283-305.	1.4	25
3316	A novel run-time load balancing method for MapReduce. , 2015, , .		1
3317	MapReduce Functions to Analyze Sentiment Information from Social Big Data. International Journal of Distributed Sensor Networks, 2015, 11, 417502.	1.3	17
3318	The K-clique Densest Subgraph Problem. , 2015, , .		145
3319	Database technologies in the world of big data. , 2015, , .		10
3320	CLAMShell. Proceedings of the VLDB Endowment, 2015, 9, 372-383.	2.1	34
3321	Automated multi-objective control for self-adaptive software design. , 2015, , .		62
3322	Enhancing Performance of MapReduce Framework in Heterogeneous Environments. , 2015, , .		1
3323	Social intelligence framework: Extracting and analyzing opinions for social CRM. , 2015, , .		7
3324	Design and Architecture of a Software Defined Proximity Cloud. , 2015, , 123-164.		1
3325	Studying the Effect of Multi-Query Functionality on a Correlation-Aware SQL-To-MapReduce Translator. , 2015, , .		1
3326	Multi-Player Gaming for Public Transport Crowd. , 2015, , 217-232.		0
3327	TWIN: Personality-based Intelligent Recommender System. Journal of Intelligent and Fuzzy Systems, 2015, 28, 2059-2071.	0.8	28
3328	An abstract machine for integrating heterogeneous web applications. , 2015, , .		0
3329	Super Rack: Reusing the Results of Queries in MapReduce Systems. , 2015, , .		1
3330	A GPU-based MapReduce framework for MSR-Bing Image Retrieval Challenge. , 2015, , .		0
3331	Minimizing Skew in MapReduce Applications Using Node Clustering in Heterogeneous Environment. , 2015, , .		3
3332	Management of Unstructured Geological Data Based on Hadoop. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
3333	MapReduce based scalable range query architecture for big spatial data. , 2015, , .		0
3334	Taxonomy on the Integration of Hadoop and Rapid Miner for Big Data Analytics. , 2015, , .		1
3337	Optimising virtual machine allocation in MapReduce cloud for improved data locality. International Journal of Big Data Intelligence, 2015, 2, 2.	0.4	23
3338	A platform for big data analytics on distributed scale-out storage system. International Journal of Big Data Intelligence, 2015, 2, 127.	0.4	34
3339	Heterogeneity-aware scheduler for stream processing frameworks. International Journal of Big Data Intelligence, 2015, 2, 70.	0.4	12
3340	Potentials of Online Media and Location-Based Big Data for Urban Transit Networks in Developing Countries. Transportation Research Record, 2015, 2537, 52-61.	1.0	7
3341	Large-Scale Merging of Histograms using Distributed In-Memory Computing. Journal of Physics: Conference Series, 2015, 664, 092003.	0.3	4
3342	Distributed multinomial regression. Annals of Applied Statistics, 2015, 9, .	0.5	48
3343	On data skewness, stragglers, and MapReduce progress indicators. , 2015, , .		20
3344	A framework for scheduling and managing big data applications in a distributed infrastructure. , 2015, , .		2
3345	Mitigate I/O access pattern divergence with heterogeneous architecture in HDFS. , 2015, , .		2
3346	Exploring network optimizations for large-scale graph analytics. , 2015, , .		5
3347	Combination of OSELM classifiers with fuzzy integral for large scale classification. Journal of Intelligent and Fuzzy Systems, 2015, 28, 2257-2268.	0.8	2
3348	Data partitioning strategies for graph workloads on heterogeneous clusters. , 2015, , .		29
3349	DisasterMapper. , 2015, , .		25
3350	Development and Evaluation of Congestion Detection System Using Complex Event Processing for Mobile Networks. Journal of Information Processing, 2015, 23, 430-440.	0.3	0
3351	DBalancer: A tool for dynamic changing of workers number in storm. , 2015, , .		0
3352	What is big data? A consensual definition and a review of key research topics. AIP Conference Proceedings, 2015, , .	0.3	257

#	ARTICLE	IF	CITATIONS
3353	IOT Big Data Process Based on Context Temporal-Spatial Characteristics in Petroleum Products Distribution. , 2015, , .		2
3354	Adaptive task scheduling in storm. , 2015, , .		4
3355	ATL-MR: model transformation on MapReduce. , 2015, , .		2
3356	Sparse Kernel Clustering of Massive High-Dimensional Data sets with Large Number of Clusters. , 2015, , .		3
3357	Building Reliable Storage Clouds: Models, Fundamental Tradeoffs, and Solutions. Foundations and Trends in Networking, 2015, 9, 219-315.	10.2	1
3358	Application of an AIS to the problem of through life health management of remotely piloted aircraft. , 2015, , .		0
3359	Leveraging Data Intensive Applications on a Pervasive Computing Platform: The Case of MapReduce. Procedia Computer Science, 2015, 52, 1034-1039.	1.2	5
3360	Subset selection of high-depth next generation sequencing reads for de novo genome assembly using MapReduce framework. BMC Genomics, 2015, 16, S9.	1.2	6
3361	An alignment-free approach to cluster proteins using frequency of conserved k-mers. , 2015, , .		3
3362	Social Trove: A Self-Summarizing Storage Service for Social Sensing. , 2015, , .		11
3363	Tackling the Achilles Heel of Social Networks. , 2015, , .		10
3364	Efficient and Scalable Collection of Dynamic Metrics Using MapReduce. , 2015, , .		4
3365	An Optimized Strategy for Speculative Execution in a Heterogeneous Environment. , 2015, , .		1
3366	Towards a Load-Aware Scheduling Framework for Realtime Video Cloud. , 2015, , .		0
3367	Reducing Communication and Merging Overheads for Distributed Clustering Algorithms on the Cloud. , 2015, , .		2
3368	Cloud-assisted privacy-preserving genetic paternity test. , 2015, , .		1
3369	Telemetry for Elastic Data (TED): Middleware for MapReduce Job Metering and Rating. , 2015, , .		0
3370	Enabling Big Data Analytics in the Hybrid Cloud Using Iterative MapReduce. , 2015, , .		17

#	ARTICLE	IF	CITATIONS
3371	DRAT: An Unobtrusive, Scalable Approach to Large Scale Software License Analysis. , 2015, , .		0
3372	The Dynamic Delay Scheduling Algorithm Based on Task Classification. , 2015, , .		2
3373	Does chunk size matter in distributed video transcoding?. , 2015, , .		7
3374	An Inter-framework Cache for Diverse Data-Intensive Computing Environments. , 2015, , .		0
3375	A New Label Propagation With Dams. , 2015, , .		5
3376	VPCB: A Consistent Hashing Algorithm for Better Load Balancing in a Hadoop Environment. , 2015, , .		4
3377	Dust: Real-Time Code Offloading System for Wearable Computing. , 2015, , .		7
3378	Web users browsing behavior prediction by implementing support vector machines in MapReduce using cloud based Hadoop. , 2015, , .		0
3379	A Data-Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition. Journal of Nonlinear Science, 2015, 25, 1307-1346.	1.0	1,044
3380	The Research on the Application of Qualitative Mapping in MapReduce. Annals of Data Science, 2015, 2, 271-280.	1.7	3
3381	A Stable and Distributed Community Detection Algorithm Based on Maximal Cliques. , 2015, , .		1
3382	MapReduce SVM Game. Procedia Computer Science, 2015, 53, 298-307.	1.2	1
3383	An Analytic Model to Optimize Search Results Using Collaborative Searching Methods. Procedia Computer Science, 2015, 57, 1255-1264.	1.2	0
3384	Spark-gram: Mining frequent N-grams using parallel processing in Spark. , 2015, , .		2
3385	Bandit-based Monte-Carlo structure learning of probabilistic logic programs. Machine Learning, 2015, 100, 127-156.	3.4	8
3386	D3-MapReduce: Towards MapReduce for Distributed and Dynamic Data Sets. , 2015, , .		0
3387	Truth Discovery on Crowd Sensing of Correlated Entities. , 2015, , .		123
3388	Applying semi-synchronised task farming to large-scale computer vision problems. International Journal of High Performance Computing Applications, 2015, 29, 437-460.	2.4	0

#	ARTICLE	IF	CITATIONS
3389	A Parallel Version of Differential Evolution Based on Resilient Distributed Datasets Model. Communications in Computer and Information Science, 2015, , 84-93.	0.4	7
3390	Bridging the gap between real world repositories and scalable preservation environments. International Journal on Digital Libraries, 2015, 16, 267-282.	1.1	4
3391	SigMR: MapReduce-based SPARQL query processing by signature encoding and multi-way join. Journal of Supercomputing, 2015, 71, 3695-3725.	2.4	5
3392	Slowing Little Quickens More: Improving DCTCP for Massive Concurrent Flows. , 2015, , .		6
3393	High-Speed Classification of Financial Network Public Opinion Based on Hadoop. , 2015, , .		1
3394	A Task Decomposition Framework for Surveying the Crowd Contextual Insights. , 2015, , .		1
3395	Semi-Supervised Multiple Disambiguation. , 2015, , .		0
3396	SeedsGraph: an efficient assembler for next-generation sequencing data. BMC Medical Genomics, 2015, 8, S13.	0.7	2
3397	Parallel Optimization of Pixel Purity Index Algorithm for Hyperspectral Unmixing Based on Spark. , 2015, , .		4
3398	MapReduce Job Optimization: A Mapping Study. , 2015, , .		2
3399	A Hybrid Algorithm for Frequent Pattern Mining Using MapReduce Framework. , 2015, , .		4
3400	Empowering Fast Incremental Computation over Large Scale Dynamic Graphs. , 2015, , .		3
3401	Cracker: Crumbling large graphs into connected components. , 2015, , .		18
3402	Scalable Spatial Crowdsourcing: A Study of Distributed Algorithms. , 2015, , .		29
3403	Context-aware scheduling in MapReduce: a compact review. Concurrency Computation Practice and Experience, 2015, 27, 5332-5349.	1.4	10
3404	Decentralized scheduling with data locality for data-parallel computation on peer-to-peer networks. , 2015, , .		7
3405	Ensemble prediction of vascular injury in Trauma care: Initial efforts towards data-driven, low-cost screening. , 2015, , .		0
3406	Knowledge structures: Which one to use for the query disambiguation?. , 2015, , .		5

#	ARTICLE	IF	CITATIONS
3407	SYNC or ASYNC: time to fuse for distributed graph-parallel computation. ACM SIGPLAN Notices, 2015, 50, 194-204.	0.2	27
3408	Efficient replication of queued tasks for latency reduction in cloud systems. , 2015, , .		35
3409	A Parallel Clustering Algorithm Based on Grid Index for Spatio-temporal Trajectories. , 2015, , .		1
3410	BF-MapReduce: A Bloom Filter Based Efficient Lightweight Search. , 2015, , .		0
3411	STAC: Simultaneous transmitting and air computing in wireless data center networks. , 2015, , .		2
3412	Segmentation of Chinese Web Text Based on Spark. , 2015, , .		0
3413	A partial-decentralized coflow scheduling scheme in data center networks. , 2015, , .		0
3414	Unique Sense: Smart Computing Prototype. Procedia Computer Science, 2015, 50, 223-228.	1.2	12
3415	Scalable and cost-effective NGS genotyping in the cloud. BMC Medical Genomics, 2015, 8, 64.	0.7	19
3416	Low-latency Java communication devices on RDMA-enabled networks. Concurrency Computation Practice and Experience, 2015, 27, 4852-4879.	1.4	0
3417	Failure-resilient real-time processing of health streams. Concurrency Computation Practice and Experience, 2015, 27, 1695-1717.	1.4	0
3418	Nonclairvoyant reduction algorithms for heterogeneous platforms. Concurrency Computation Practice and Experience, 2015, 27, 1612-1624.	1.4	1
3419	Automatic Leaf Recognition from a Big Hierarchical Image Database. International Journal of Intelligent Systems, 2015, 30, 871-886.	3.3	9
3420	Blast: Accelerating high-performance data analytics applications by optical multicast. , 2015, , .		20
3421	BiFennel: Fast Bipartite Graph Partitioning Algorithm for Big Data. , 2015, , .		0
3422	Alleviation of Disk I/O Contention in Virtualized Settings for Data-Intensive Computing. , 2015, , .		4
3423	System Status Aware Hadoop Scheduling Methods for Job Performance Improvement. IEICE Transactions on Information and Systems, 2015, E98.D, 1275-1285.	0.4	3
3424	3D Objects Tracking by MapReduce GPGPU-Enhanced Particle Filter. IEICE Transactions on Information and Systems, 2015, E98.D, 1035-1044.	0.4	1

#	ARTICLE	IF	CITATIONS
3425	Parallel Markov chain Monte Carlo for non-Gaussian posterior distributions. Stat, 2015, 4, 304-319.	0.3	7
3426	Identifying the most appropriate expansion of acronyms used in wikipedia text. Software - Practice and Experience, 2015, 45, 1073-1086.	2.5	4
3427	Auto associative Extreme Learning Machine based non-linear principal component regression for big data applications. , 2015, , .		5
3428	Inter-cloud data persistence in DIL networks. , 2015, , .		0
3429	Experiences on File Systems: Which is the best file system for you?. Journal of Physics: Conference Series, 2015, 664, 042004.	0.3	3
3430	HybridMR: a new approach for hybrid MapReduce combining desktop grid and cloud infrastructures. Concurrency Computation Practice and Experience, 2015, 27, 4140-4155.	1.4	19
3431	SONAR: A scalable stream-oriented system for real-time network traffic measurements. , 2015, , .		0
3432	Ant colony optimization of virtual machine placement for data latency minimization in cloud systems. , 2015, , .		0
3433	Utility-Based VM Assignment in Cloud Scenarios. , 2015, , .		0
3434	Scalable multi-dimensional RNN query processing. Concurrency Computation Practice and Experience, 2015, 27, 4156-4171.	1.4	4
3435	A platform to deploy customized scientific virtual infrastructures on the cloud. Concurrency Computation Practice and Experience, 2015, 27, 4318-4329.	1.4	13
3436	Distributed forests for MapReduce-based machine learning. , 2015, , .		6
3437	Congestion-free routing strategy in software defined data center networks. Concurrency Computation Practice and Experience, 2015, 27, 5735-5748.	1.4	3
3438	Recognizing and characterizing dynamics of cellular devices in cellular data network through massive data analysis. International Journal of Communication Systems, 2015, 28, 1884-1897.	1.6	9
3439	Applying Big Data Analytics to Logistics Processes of Oil and Gas Exploration and Production Through a Hybrid Modeling and Simulation Approach. , 2015, , .		4
3440	An architecture of Distributed Beta Wavelet Networks for large image classification in MapReduce. , 2015, , .		3
3441	Concurrent software architectures for exploratory data analysis. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2015, 5, 165-180.	4.6	1
3442	An urgency and congestion control scheme for larger-scale TCP incast problem in data center. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
3443	MapReduce scheduling in hybrid cloud with multi-level privacy. , 2015, , .		0
3444	Data, Big Data, and Metadata in Anesthesiology. Anesthesia and Analgesia, 2015, 121, 1661-1667.	1.1	31
3445	Efficient ELM-Based Two Stages Query Processing Optimization for Big Data. Mathematical Problems in Engineering, 2015, 2015, 1-12.	0.6	4
3446	Partitioning Uncertain Workflows. SSRN Electronic Journal, 0, , .	0.4	0
3447	Faster single-end alignment generation utilizing multi-thread for BWA. Bio-Medical Materials and Engineering, 2015, 26, S1791-S1796.	0.4	46
3448	Efficient Multicriteria Protein Structure Comparison on Modern Processor Architectures. BioMed Research International, 2015, 2015, 1-13.	0.9	3
3450	Hadoop Mapreduce Performance Enhancement Using In-Node Combiners. International Journal of Computer Science and Information Technology, 2015, 7, 1-17.	0.3	8
3451	Plato: A Selective Context Model for Entity Resolution. Transactions of the Association for Computational Linguistics, 2015, 3, 503-515.	3.2	38
3452	A Modified Key Partitioning for BigData Using MapReduce in Hadoop. Journal of Computer Science, 2015, 11, 490-497.	0.5	0
3453	Toward High Utilization of Heterogeneous Computing Resources in SNP Detection. ETRI Journal, 2015, 37, 212-221.	1.2	0
3454	Shifting Perspectives. International Journal of Strategic Decision Sciences, 2015, 6, 33-52.	0.0	6
3455	Implementation of a Parallel Algorithm Based on a Spark Cloud Computing Platform. Algorithms, 2015, 8, 407-414.	1.2	14
3456	A Data Analytic Algorithm for Managing, Querying, and Processing Uncertain Big Data in Cloud Environments. Algorithms, 2015, 8, 1175-1194.	1.2	63
3457	Can COTS Ethernet Switches Handle Uncompressed Video?. Smppte Motion Imaging Journal, 2015, 124, 21-28.	0.2	0
3458	Big Data in Design and Manufacturing Engineering. American Journal of Engineering and Applied Sciences, 2015, 8, 223-232.	0.3	24
3459	Accelerated Simplified Swarm Optimization with Exploitation Search Scheme for Data Clustering. PLoS ONE, 2015, 10, e0137246.	1.1	21
3460	Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. , 2015, , 577-602.		29
3461	Big Data Analytics in Healthcare. BioMed Research International, 2015, 2015, 1-16.	0.9	332

#	ARTICLE	IF	CITATIONS
3462	Performance Comparison of OpenMP, MPI, and MapReduce in Practical Problems. <i>Advances in Multimedia</i> , 2015, 2015, 1-9.	0.2	51
3463	An Effective Big Data Supervised Imbalanced Classification Approach for Ortholog Detection in Related Yeast Species. <i>BioMed Research International</i> , 2015, 2015, 1-12.	0.9	17
3464	A Parallel Community Structure Mining Method in Big Social Networks. <i>Mathematical Problems in Engineering</i> , 2015, 2015, 1-13.	0.6	5
3465	A Replication-Based Mechanism for Fault Tolerance in MapReduce Framework. <i>Mathematical Problems in Engineering</i> , 2015, 2015, 1-7.	0.6	7
3466	Performances Evaluation of a Novel Hadoop and Spark Based System of Image Retrieval for Huge Collections. <i>Advances in Multimedia</i> , 2015, 2015, 1-7.	0.2	9
3467	Implementation of Secondary Index on Cloud Computing NoSQL Database in Big Data Environment. <i>Scientific Programming</i> , 2015, 2015, 1-10.	0.5	5
3468	Evolutionary Feature Selection for Big Data Classification: A MapReduce Approach. <i>Mathematical Problems in Engineering</i> , 2015, 2015, 1-11.	0.6	97
3469	MapReduce Based Parallel Neural Networks in Enabling Large Scale Machine Learning. <i>Computational Intelligence and Neuroscience</i> , 2015, 2015, 1-13.	1.1	33
3470	An Efficient MapReduce-Based Parallel Clustering Algorithm for Distributed Traffic Subarea Division. <i>Discrete Dynamics in Nature and Society</i> , 2015, 2015, 1-18.	0.5	17
3471	Architecture and Implementation of a Scalable Sensor Data Storage and Analysis System Using Cloud Computing and Big Data Technologies. <i>Journal of Sensors</i> , 2015, 2015, 1-11.	0.6	41
3472	Supporting online analytics with user-defined estimation and early termination in a MapReduce-like framework. , 2015, , .		2
3473	MapReduce Accounting System Integrated with High-Performance Computing Infrastructure. , 2015, , .		0
3474	Distributed and parallel approach for handle and perform huge datasets. <i>AIP Conference Proceedings</i> , 2015, , .	0.3	1
3475	Random Walk TripleRush. , 2015, , .		7
3476	Improving Throughput of a Pipeline Model Indexer. , 2015, , .		1
3477	A novel similarity comparison approach for dynamic ECG series. <i>Bio-Medical Materials and Engineering</i> , 2015, 26, S1095-S1105.	0.4	2
3478	The production of salary profiles of ICT professionals: Moving from structured database to big data analytics. <i>Statistical Journal of the IAOS</i> , 2015, 31, 177-191.	0.2	6
3479	Energy Efficient Scale-In Clusters with In-Storage Processing for Big-Data Analytics. , 2015, , .		10

#	ARTICLE	IF	CITATIONS
3480	Performance evaluation and tuning of BioPig for genomic analysis. , 2015, , .		1
3481	A Novel Multidimensional Approach to Integrate Big Data in Business Intelligence. Journal of Database Management, 2015, 26, 14-31.	1.0	11
3482	Distributed and collaborative Web Change Detection system. Computer Science and Information Systems, 2015, 12, 91-114.	0.7	10
3483	Experimental evaluation of a flexible I/O architecture for accelerating workflow engines in cloud environments. , 2015, , .		1
3484	Model of Cloud-Based Services for Data Mining Analysis. Journal of Computer and Information Science, 2015, 8, 40.	0.2	0
3485	Scalable Metering for an Affordable IT Cloud Service Management. , 2015, , .		5
3486	NextSearch: A Search Engine for Mass Spectrometry Data against a Compact Nucleotide Exon Graph. Journal of Proteome Research, 2015, 14, 2784-2791.	1.8	12
3487	In-Memory Big Data Management and Processing: A Survey. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 1920-1948.	4.0	297
3488	MrPhi: An Optimized MapReduce Framework on Intel Xeon Phi Coprocessors. IEEE Transactions on Parallel and Distributed Systems, 2015, 26, 3066-3078.	4.0	23
3489	Rethinking Data-Intensive Science Using Scalable Analytics Systems. , 2015, , .		67
3490	Modified Soft Rough Set Based ECG Signal Classification for Cardiac Arrhythmias. Studies in Big Data, 2015, , 445-470.	0.8	4
3491	A Hadoop-Based Framework for Large-Scale Landmine Detection Using Ubiquitous Big Satellite Imaging Data. , 2015, , .		2
3492	A Course on Advanced Software Tools for Operations Research and Analytics. INFORMS Transactions on Education, 2015, 15, 169-179.	0.4	4
3493	Cloud computing-based map-matching for transportation data center. Electronic Commerce Research and Applications, 2015, 14, 431-443.	2.5	8
3494	Community structure mining in big data social media networks with MapReduce. Cluster Computing, 2015, 18, 999-1010.	3.5	36
3495	Distributed evolutionary algorithms and their models: A survey of the state-of-the-art. Applied Soft Computing Journal, 2015, 34, 286-300.	4.1	361
3496	Enhancing speed of SQL database operations using GPU. , 2015, , .		4
3497	Modeling, Enacting, and Integrating Custom Crowdsourcing Processes. ACM Transactions on the Web, 2015, 9, 1-43.	2.0	29

#	ARTICLE	IF	CITATIONS
3498	Big data analytics and business analytics. Journal of Management Analytics, 2015, 2, 1-21.	1.6	116
3500	DaSH: A benchmark suite for hybrid dataflow and shared memory programming models. Parallel Computing, 2015, 45, 18-48.	1.3	2
3501	Real-Time Big Data Analytical Architecture for Remote Sensing Application. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8, 4610-4621.	2.3	138
3502	On Fast and Coordinated Data Backup in Geo-Distributed Optical Inter-Datacenter Networks. Journal of Lightwave Technology, 2015, , 1-1.	2.7	47
3503	A parallel heuristic reduction based approach for distribution network fault diagnosis. International Journal of Electrical Power and Energy Systems, 2015, 73, 548-559.	3.3	10
3504	Ontology Based Document Clustering Using MapReduce. International Journal of Database Management Systems, 2015, 7, 01-12.	0.2	7
3505	Semi-Supervised Approach to Named Entity Recognition in Spanish Applied to a Real-World Conversational System. Lecture Notes in Computer Science, 2015, , 224-235.	1.0	0
3508	Halvade: scalable sequence analysis with MapReduce. Bioinformatics, 2015, 31, 2482-2488.	1.8	69
3509	Discovering Big Data Modelling for Educational World. Procedia, Social and Behavioral Sciences, 2015, 176, 642-649.	0.5	23
3510	ParSA: High-throughput scientific data analysis framework with distributed file system. Future Generation Computer Systems, 2015, 51, 111-119.	4.9	6
3511	Performance Improvement of MapReduce Framework in Heterogeneous Context using Reinforcement Learning. Procedia Computer Science, 2015, 50, 169-175.	1.2	14
3512	An Enhanced DACHE Model for the MapReduce Environment. Procedia Computer Science, 2015, 50, 579-584.	1.2	5
3513	A Two Pass Scheduling Policy Based Resource Allocation for MapReduce. Procedia Computer Science, 2015, 46, 627-634.	1.2	3
3514	G2P: A Partitioning Approach for Processing DBSCAN with MapReduce. Lecture Notes in Computer Science, 2015, , 191-202.	1.0	3
3515	Adopting cloud computing to optimize spatial web portals for better performance to support Digital Earth and other global geospatial initiatives. International Journal of Digital Earth, 2015, 8, 451-475.	1.6	18
3516	Fast K-Means Clustering for Very Large Datasets Based on MapReduce Combined with a New Cutting Method. Advances in Intelligent Systems and Computing, 2015, , 287-298.	0.5	11
3517	Greedy column subset selection for large-scale data sets. Knowledge and Information Systems, 2015, 45, 1-34.	2.1	41
3518	Analyzing job completion reliability and job energy consumption for a heterogeneous MapReduce cluster under different intermediate-data replication policies. Journal of Supercomputing, 2015, 71, 1657-1677.	2.4	54

#	ARTICLE	IF	CITATIONS
3519	An integrated visualization environment for the virtual observatory: Current status and future directions. <i>Astronomy and Computing</i> , 2015, 11, 146-154.	0.8	7
3520	A modular software architecture for processing of big geospatial data in the cloud. <i>Computers and Graphics</i> , 2015, 49, 69-81.	1.4	58
3521	Design of a web-based application of the coupled multi-agent system model and environmental model for watershed management analysis using Hadoop. <i>Environmental Modelling and Software</i> , 2015, 70, 149-162.	1.9	31
3522	INDREX: In-database relation extraction. <i>Information Systems</i> , 2015, 53, 124-144.	2.4	4
3523	Mining maximal cliques from a large graph using MapReduce: Tackling highly uneven subproblem sizes. <i>Journal of Parallel and Distributed Computing</i> , 2015, 79-80, 104-114.	2.7	24
3524	Recent advances and emerging challenges of feature selection in the context of big data. <i>Knowledge-Based Systems</i> , 2015, 86, 33-45.	4.0	219
3525	A distributed framework for trimmed Kernel k-Means clustering. <i>Pattern Recognition</i> , 2015, 48, 2685-2698.	5.1	31
3526	OntoMate: a text-mining tool aiding curation at the Rat Genome Database. <i>Database: the Journal of Biological Databases and Curation</i> , 2015, 2015, .	1.4	21
3527	The Construction of the Knowledge Management System in the Cloud Computing Environment. <i>Applied Mechanics and Materials</i> , 2015, 713-715, 2366-2369.	0.2	0
3528	Simple and Scalable Response Prediction for Display Advertising. <i>ACM Transactions on Intelligent Systems and Technology</i> , 2015, 5, 1-34.	2.9	175
3529	Monitoring Metric First-Order Temporal Properties. <i>Journal of the ACM</i> , 2015, 62, 1-45.	1.8	116
3530	Anomaly Detection from Log Files Using Data Mining Techniques. <i>Lecture Notes in Electrical Engineering</i> , 2015, , 449-457.	0.3	36
3531	Managing Change for Big Data Driven Innovation. , 2015, , 125-153.		13
3532	CSTORE: A desktop-oriented distributed public cloud storage system. <i>Computers and Electrical Engineering</i> , 2015, 42, 60-73.	3.0	10
3533	Large Scale Frequent Pattern Mining Using MPI One-Sided Model. , 2015, , .		4
3534	Predicting Scheduling Failures in the Cloud: A Case Study with Google Clusters and Hadoop on Amazon EMR. , 2015, , .		19
3535	ATL model checking in the cloud. , 2015, , .		0
3536	A Cloud-Assisted Framework for Bag-of-Features Tagging in Social Networks. , 2015, , .		3

#	ARTICLE	IF	CITATIONS
3537	MR-SPS: Scalable parallel scheduler for YARN/MapReduce platform. , 2015, , .		1
3538	Complex Network Analysis on Distributed Systems. , 2015, , .		4
3539	Integrating Java 8 Streams with The Real-Time Specification for Java. , 2015, , .		6
3540	A System for Tender Price Evaluation of Construction Project Based on Big Data. Procedia Engineering, 2015, 123, 606-614.	1.2	24
3541	Map-Reduce based framework for instrument detection in large-scale surgical videos. , 2015, , .		6
3542	Advancing Integrated and Personalized Healthcare Services, the AEGLE Approach. , 2015, , .		2
3543	Optimal balance between energy and performance in hybrid computing applications. , 2015, , .		7
3544	Spark on entropy: A reliable & efficient scheduler for low-latency parallel jobs in heterogeneous cloud. , 2015, , .		13
3545	A Parallel Algorithm for Statistical Multiword Term Extraction from Very Large Corpora. , 2015, , .		4
3546	Scaling GMM Expectation Maximization algorithm using bulk synchronous Parallel approach. , 2015, , .		1
3547	Panopticon: A lock broker architecture for scalable transactions in the datacenter. , 2015, , .		0
3548	wFlatnet: Introducing Wireless in Flatnet Data Center Network. , 2015, , .		5
3549	Information Theory Based Genome-Scale Gene Networks Construction Using MapReduce. , 2015, , .		1
3550	Dual-centric Data Center Network Architectures. , 2015, , .		6
3551	Query execution evaluation in wireless network using MyHadoop. , 2015, , .		6
3552	Traffic and failure aware VM placement for multi-tenant cloud computing. , 2015, , .		18
3553	Stream my models: Reactive peer-to-peer distributed models@run.time. , 2015, , .		16
3554	A delay-based TCP for solving incast problem in data centers. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
3555	A boundary filtering based spatial join query processing optimization algorithm. , 2015, , .		0
3556	Cocktail: A hybrid system combining Hadoop and Storm. , 2015, , .		0
3557	The performance evaluations for Hadoop in diverse system architectures. , 2015, , .		0
3558	Drug discovery for breast cancer based on big data analytics techniques. , 2015, , .		3
3559	Distributed multistage alert correlation architecture based on Hadoop. , 2015, , .		2
3560	Clister: A framework for iterative MapReduce. , 2015, , .		1
3561	Optical network configuration methods considering end-to-end latency in data centers. , 2015, , .		1
3562	A customizable MapReduce framework for complex data-intensive workflows on GPUs. , 2015, , .		3
3563	Routing optimization in data center networks. , 2015, , .		1
3564	Goal-Based Analytic Composition for On-and Off-line Execution at Scale. , 2015, , .		0
3565	Integro: Leveraging Victim Prediction for Robust Fake Account Detection in OSNs. , 2015, , .		64
3566	Web service for detecting credit card fraud in near real-time. , 2015, , .		4
3567	A MapReduce Approach for Processing Student Data Activity in a Peer-to-Peer Networked Setting. , 2015, , .		2
3568	Small files storing and computing optimization in Hadoop parallel rendering. , 2015, , .		2
3569	Efficient Parallelization of the Google Trigram Method for Document Relatedness Computation. , 2015, , .		0
3570	Watt Watcher: Fine-Grained Power Estimation for Emerging Workloads. , 2015, , .		9
3571	An Overview of Monitoring Tools for Big Data and Cloud Applications. , 2015, , .		9
3572	Employing in-memory data grids for distributed graph processing. , 2015, , .		4

#	ARTICLE	IF	CITATIONS
3573	WDCloud: An end to end system for large-scale watershed delineation on cloud. , 2015, , .		1
3574	Regularized and sparse stochastic k-means for distributed large-scale clustering. , 2015, , .		7
3575	An Autonomous Data Structure for Brute Force Calculations in the Cloud. , 2015, , .		0
3576	SoSE Architecture of Data Intensive Computing for Healthcare Information System. , 2015, , .		0
3577	A Survey on MapReduce Scheduling in Cloud Computing. , 2015, , .		0
3578	Enriching megamodel management with collection-based operators. , 2015, , .		10
3579	GEM: A Framework for Developing Shared-Memory Parallel Genomic Applications on Memory Constrained Architectures. , 2015, , .		1
3580	Scalable Euclidean Embedding for Big Data. , 2015, , .		0
3581	Mining network traffic anomaly based on adjustable piecewise entropy. , 2015, , .		5
3582	Millipedes: Distributed and Set-Based Sub-Task Scheduler of Computing Engines Running on Yarn Cluster. , 2015, , .		4
3583	Modeling randomized data streams in caching, data processing, and crawling applications. , 2015, , .		1
3584	OMO: Optimize MapReduce overlap with a good start (reduce) and a good finish (map). , 2015, , .		9
3585	Memory-Based Data Storing Technologies on Hadoop Distribution File System. , 2015, , .		1
3586	Incast-Aware Switch-Assisted TCP Congestion Control for Data Centers. , 2015, , .		10
3587	TAPS: Software Defined Task-Level Deadline-Aware Preemptive Flow Scheduling in Data Centers. , 2015, , .		1
3588	On estimating the Spectral Radius of large graphs through subgraph sampling. , 2015, , .		5
3589	Timely Long Tail Identification through Agent Based Monitoring and Analytics. , 2015, , .		12
3590	Large-scale subspace clustering using random sketching and validation. , 2015, , .		4

#	ARTICLE	IF	CITATIONS
3591	A LogP Extension for Modeling Tree Aggregation Networks. , 2015, , .		1
3592	When Good-Enough is Enough: Complex Queries at Fixed Cost. , 2015, , .		3
3593	Providing Fairer Resource Allocation for Multi-tenant Cloud-Based Systems. , 2015, , .		3
3594	Evolutionary Neural Network Based Energy Consumption Forecast for Cloud Computing. , 2015, , .		16
3595	Scalable RDF triple store using summary of hashed information and Bit comparison. , 2015, , .		3
3596	A Hybrid Processing System for Large-Scale Traffic Sensor Data. IEEE Access, 2015, 3, 2341-2351.	2.6	18
3597	GossipMap. , 2015, , .		19
3598	PL2AP. , 2015, , .		4
3599	MapReduce-based k-prototypes clustering method for big data. , 2015, , .		16
3600	Load balancing in cloud computing environment based on an improved particle swarm optimization. , 2015, , .		20
3601	Distributed Discord Discovery: Spark Based Anomaly Detection in Time Series. , 2015, , .		7
3602	Encapsulation of partitioned parallelism into open-source database management systems. Programming and Computer Software, 2015, 41, 350-360.	0.5	3
3603	MRPrimer: a MapReduce-based method for the thorough design of valid and ranked primers for PCR. Nucleic Acids Research, 2015, 43, e130-e130.	6.5	14
3604	Evaluating the Combined Impact of Node Architecture and Cloud Workload Characteristics on Network Traffic and Performance/Cost. , 2015, , .		6
3605	Study of attribute word extraction in Chinese product reviews. , 2015, , .		0
3606	MIPr - a framework for distributed image processing using Hadoop. , 2015, , .		8
3607	Towards distributed region growing image segmentation based on MapReduce. , 2015, , .		10
3608	Secure image processing inside cloud file sharing environment using lightweight containers. , 2015, , .		8

#	ARTICLE	IF	CITATIONS
3609	Analysis of Data Preprocessing Increasing the Oversampling Ratio for Extremely Imbalanced Big Data Classification. , 2015, , .		19
3610	A Grid-Based k-Nearest Neighbor Join for Large Scale Datasets on MapReduce. , 2015, , .		4
3611	Distributed model-to-model transformation with ATL on MapReduce. , 2015, , .		15
3612	VDMR-DBSCAN: Varied Density MapReduce DBSCAN. Lecture Notes in Computer Science, 2015, , 134-150.	1.0	5
3613	REFBSS: Reference based similarity search in biological network databases. , 2015, , .		1
3614	Jupiter Rising. Computer Communication Review, 2015, 45, 183-197.	1.5	228
3615	Using Straggler Replication to Reduce Latency in Large-scale Parallel Computing. Performance Evaluation Review, 2015, 43, 7-11.	0.4	86
3616	MREv: An Automatic MapReduce Evaluation Tool for Big Data Workloads. Procedia Computer Science, 2015, 51, 80-89.	1.2	7
3617	Scaling and contextualizing personalized healthcare: A case study of disease prediction algorithm integration. Journal of Biomedical Informatics, 2015, 57, 377-385.	2.5	14
3618	A Productive Cloud Computing Platform Research for Big Data Analytics. , 2015, , .		4
3619	Gemini: An Adaptive Performance-Fairness Scheduler for Data-Intensive Cluster Computing. , 2015, , .		13
3620	Resource aware scheduler for heterogeneous workload based on estimated task processing time. , 2015, , .		1
3621	Big database stores a review on various big data datastores. , 2015, , .		3
3622	ATLAS: An Adaptive Failure-Aware Scheduler for Hadoop. , 2015, , .		9
3623	An Improved MapReduce Design of Kmeans with Iteration Reducing for Clustering Stock Exchange Very Large Datasets. , 2015, , .		2
3624	Efficiency Improvements in Social Network Communication via MapReduce. , 2015, , .		5
3625	Flex: Flexible and Energy Efficient Scheduling for Big Data Storage. , 2015, , .		0
3626	Scalable developments for big data analytics in remote sensing. , 2015, , .		6

#	ARTICLE	IF	CITATIONS
3627	SDN-based TCP congestion control in data center networks. , 2015, , .		8
3628	Tempe: Live scripting for live data. , 2015, , .		11
3629	Workload scheduling in distributed stream processors using graph partitioning. , 2015, , .		23
3630	Early experience with optimizing I/O performance using high-performance SSDs for in-memory cluster computing. , 2015, , .		17
3631	GraphSC: Parallel Secure Computation Made Easy. , 2015, , .		68
3632	Semantics preserving MapReduce process for RDB to RDF transformation. International Journal of Metadata, Semantics and Ontologies, 2015, 10, 229.	0.2	4
3633	The implementation and use of a generic dataflow behaviour in Erlang. , 2015, , .		1
3634	Identification of (near) Real-time Traffic Congestion in the Cities of Australia through Twitter. , 2015, , .		18
3635	RE-PAGE: Domain-Specific REplication and PARallel Processing of GENomic Data. , 2015, , .		0
3636	An optimized cloud based big data processing mechanism using Self-Organizing Map in Hadoop environments. , 2015, , .		0
3637	A Resource Supply-Demand based Approach for Automatic MapReduce Job Optimization. , 2015, , .		1
3638	Performance evaluation of fair and capacity scheduling in Hadoop YARN. , 2015, , .		6
3639	Modified Delay Scheduling: A Heuristic Approach for Hadoop Scheduling to Improve Fairness and Response Time. Parallel Processing Letters, 2015, 25, 1550009.	0.4	2
3640	On Whether OCS Maximizes Application Throughput in All-Optical Datacenter Networks. Journal of Optical Communications and Networking, 2015, 7, 1135.	3.3	7
3641	An Approach to Instantly Detecting Fake Plates Based on Large-Scale ANPR Data. , 2015, , .		2
3642	Distributed processing of biological interactions using Hadoop. , 2015, , .		0
3643	In-memory Query System for Scientific Databases. , 2015, , .		6
3644	Rapier: Integrating routing and scheduling for coflow-aware data center networks. , 2015, , .		141

#	ARTICLE	IF	CITATIONS
3645	Enhancing TCP Incast congestion control over large-scale datacenter networks. , 2015, , .		0
3646	Testing data transformations in MapReduce programs. , 2015, , .		9
3647	Performance Evaluation of a MapReduce Hadoop-Based Implementation for Processing Large Virtual Campus Log Files. , 2015, , .		5
3648	A Multi-layer Framework for Graph Processing via Overlay Composition. Lecture Notes in Computer Science, 2015, , 515-527.	1.0	4
3649	Performance Tuning of MapReduce Jobs Using Surrogate-based Modeling. Procedia Computer Science, 2015, 51, 49-59.	1.2	4
3650	Secure Proxy Service Using p-Fibonacci Transformation of Cosine Coefficients on Cloud File Sharing Environment. , 2015, , .		1
3651	High-level synthesizable dataflow MapReduce accelerator for FPGA-coupled data centers. , 2015, , .		6
3652	TADOOP: Mining Network Traffic Anomalies with Hadoop. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2015, , 175-192.	0.2	5
3653	Change data capture in NoSQL databases: A functional and performance comparison. , 2015, , .		4
3654	Realization of secondary indexing to NoSQL database with intelligent adaptation. , 2015, , .		1
3655	Boosting of Association Rules for Robust Emergency Detection. , 2015, , .		4
3656	Semantic HMC: A Predictive Model Using Multi-label Classification for Big Data. , 2015, , .		6
3657	A predictive scheduling framework for fast and distributed stream data processing. , 2015, , .		27
3658	Cluster Scheduler on Heterogeneous Cloud. , 2015, , .		0
3659	Key Nodes Discovery in Large-Scale Logistics Network Based on MapReduce. , 2015, , .		4
3660	LuBase: A Search-Efficient Hybrid Storage System for Massive Text Data. Lecture Notes in Computer Science, 2015, , 134-148.	1.0	1
3661	A cloud-based approach for Gene Regulatory Networks dynamics simulations. , 2015, , .		2
3662	An iterative methodology for big data management, analysis and visualization. , 2015, , .		8

#	ARTICLE	IF	CITATIONS
3663	SDFS: Secure distributed file system for data-at-rest security for Hadoop-as-a-service. , 2015, , .		9
3664	High quality clustering of big data and solving empty-clustering problem with an evolutionary hybrid algorithm. , 2015, , .		10
3665	A new approach to the massive processing of satellite images. , 2015, , .		3
3666	Toward an analytical framework for proteomics software. , 2015, , .		0
3667	To Optimize Graph Based Power Iteration for Big Data Based on MapReduce Paradigm. Lecture Notes in Computer Science, 2015, , 370-381.	1.0	0
3668	Retrieval of Distinctive Regions of Interest from Video Surveillance Footage: A Real Use Case Study. , 2015, , .		3
3669	Computational Fields Meet Augmented Reality: Perspectives and Challenges. , 2015, , .		0
3670	Gray Computing: An Analysis of Computing with Background JavaScript Tasks. , 2015, , .		5
3671	Large-scale human action recognition with spark. , 2015, , .		8
3672	On Streaming Consistency of Big Data Stream Processing in Heterogenous Clusters. , 2015, , .		8
3673	Time series discord detection in medical data using a parallel relational database. , 2015, , .		7
3674	Adaptive scheduling in the cloud - SLA for Hadoop job scheduling. , 2015, , .		5
3675	Parallel Training GBRT Based on KMeans Histogram Approximation for Big Data. Lecture Notes in Computer Science, 2015, , 52-65.	1.0	0
3676	CF4BDA: A Conceptual Framework for Big Data Analytics Applications in the Cloud. IEEE Access, 2015, 3, 1944-1952.	2.6	6
3677	Weighted Finite Automata Based Image Compression on Hadoop MapReduce Framework. , 2015, , .		1
3678	AzureRender: A Cloud-Based Parallel and Distributed Rendering System. , 2015, , .		2
3679	Research into Big data for smart grids. , 2015, , .		1
3680	A MapReduce-Based k-Nearest Neighbor Approach for Big Data Classification. , 2015, , .		53

#	ARTICLE	IF	CITATIONS
3681	The PigMix Benchmark on Pig, MapReduce, and HPC Systems. , 2015, , .		10
3682	A Scalable Spatio-temporal Data Storage for Intelligent Transportation Systems Based on HBase. , 2015, , .		3
3683	Parallel computation of information gain using Hadoop and MapReduce. , 0, , .		13
3684	A Smart Work Performance Measurement System for Police Officers. IEEE Access, 2015, 3, 1755-1764.	2.6	11
3685	Dynamic Token Based Improving MapReduce Performance in Cloud Computing. , 2015, , .		1
3686	SRIdent: A novel pipeline for real-time identification of species from high-throughput sequencing reads in Metagenomics and clinical diagnostic assays. , 2015, 2015, 6481-4.		0
3687	Distributed Entropy Minimization Discretizer for Big Data Analysis under Apache Spark. , 2015, , .		17
3688	Handwritten digit recognition of Indian scripts: A cascade of distances approach. , 2015, , .		3
3689	Biomarker discovery, high performance and cloud computing: A comprehensive review. , 2015, , .		3
3690	Query sense disambiguation leveraging large scale user behavioral data. , 2015, , .		3
3691	Parallel information fusion method for microarray data analysis. , 2015, , .		0
3692	Scalable preference queries for high-dimensional data using map-reduce. , 2015, , .		1
3693	Performance evaluation of enabling logistic regression for big data with R. , 2015, , .		14
3694	A Knowledge Management Process to Enhance Trustworthiness-based Security in On-line Learning Teams. , 2015, , .		3
3695	Distributed SAR image change detection based on Spark. , 2015, , .		6
3696	Vispark: GPU-accelerated distributed visual computing using spark. , 2015, , .		2
3697	The RAMCloud Storage System. ACM Transactions on Computer Systems, 2015, 33, 1-55.	0.6	162
3698	A Load-Balancing Algorithm for Hadoop Distributed File System. , 2015, , .		14

#	ARTICLE	IF	CITATIONS
3699	One trillion edges. Proceedings of the VLDB Endowment, 2015, 8, 1804-1815.	2.1	274
3700	Multitask Oriented GPU Resource Sharing and Virtualization in Cloud Environment. Lecture Notes in Computer Science, 2015, , 509-524.	1.0	1
3701	SEIP: System for Efficient Image Processing on Distributed Platform. Journal of Computer Science and Technology, 2015, 30, 1215-1232.	0.9	6
3702	Enabling scientific data storage and processing on big-data systems. , 2015, , .		5
3703	Towards a big data theory model. , 2015, , .		2
3704	A Top-k query algorithm for big data based on MapReduce. , 2015, , .		1
3705	Large scale thematic mapping by supervised machine learning on ‘big data’ distributed cluster computing frameworks. , 2015, , .		1
3706	A Framework for Privacy-Aware Computing on Hybrid Clouds with Mixed-Sensitivity Data. , 2015, , .		11
3707	Research and improve on K-means algorithm based on hadoop. , 2015, , .		4
3708	Machine Learning for Data Science: Mathematical or Computational. , 2015, , 63-74.		1
3709	Chronos: Failure-aware scheduling in shared Hadoop clusters. , 2015, , .		14
3710	An Ordinal Random Forest and Its Parallel Implementation with MapReduce. , 2015, , .		2
3711	Generic menu optimization for multi-profile customer systems. , 2015, , .		2
3712	SpeedStream: A real-time stream data processing platform in the cloud. , 2015, , .		1
3713	Automated, high throughput exploration of process“structure”property relationships using the MapReduce paradigm. Materials Discovery, 2015, 1, 21-28.	3.3	23
3714	A Scalable Algorithm for Bichromatic Reverse Nearest Neighbor with Grids. , 2015, , .		0
3715	Acceleration of CUDA programs for non-GPU users using cloud. , 2015, , .		2
3716	Toward Optimal Resource Provisioning for Cloud MapReduce and Hybrid Cloud Applications. , 2015, , .		6

#	ARTICLE	IF	CITATIONS
3717	Scalable Property Aggregation for Linked Data Recommender Systems. , 2015, , .		1
3718	Hadoop based real-time Big Data Architecture for remote sensing Earth Observatory System. , 2015, , .		13
3719	PIC: Enable Large-Scale Privacy Preserving Content-Based Image Search on Cloud. , 2015, , .		15
3720	DPX10: An Efficient X10 Framework for Dynamic Programming Applications. , 2015, , .		2
3721	Beyond the lambda architecture: Effective scheduling for large scale EO information mining and interactive thematic mapping. , 2015, , .		6
3722	COMET: Client-Oriented METadata Service for Highly Available Distributed File Systems. , 2015, , .		2
3723	Designing an efficient persistent in-memory file system. , 2015, , .		12
3724	Development of MapReduce and MPI Programs for Motif Search. , 2015, , .		0
3725	Reflex-Tree: A Biologically Inspired Parallel Architecture for Future Smart Cities. , 2015, , .		10
3726	On web-based domain-specific language for Internet of Things. , 2015, , .		10
3727	Large interactive visualization of density functions on big data infrastructure. , 2015, , .		40
3728	Smart Crawler. , 2015, , .		0
3729	Where to go and what to play: Towards summarizing popular information from massive tourism blogs. Journal of Information Science, 2015, 41, 830-854.	2.0	21
3730	Key Frame Extraction and Foreground Modelling Using K-Means Clustering. , 2015, , .		6
3731	MatrixMap: Programming Abstraction and Implementation of Matrix Computation for Big Data Applications. , 2015, , .		3
3732	Accelerating Support Vector Machine Learning with GPU-Based MapReduce. , 2015, , .		3
3733	Rewriting complex SPARQL analytical queries for efficient cloud-based processing. , 2015, , .		0
3734	A scalable parallel XQuery processor. , 2015, , .		2

#	ARTICLE	IF	CITATIONS
3735	Job-Optimized Map-Side Join Processing Using MapReduce and HBase with Abstract RDF Data. , 2015, , .		3
3736	Distributed, MapReduce-Based Nearest Neighbor and E-Ball Kernel k-Means. , 2015, , .		3
3737	Hierarchical Mahalanobis Distance Clustering Based Technique for Prognostics in Applications Generating Big Data. , 2015, , .		0
3738	A Succinct Distributive Big Data Clustering Algorithm Based on Local-Remote Coordination. , 2015, , .		2
3739	Log-based change data capture from schema-free document stores using MapReduce. , 2015, , .		8
3740	MapReduce-Based H-Mine Algorithm. , 2015, , .		2
3741	A novel framework for mitigating insider attacks in big data systems. , 2015, , .		5
3742	SMASH: A Cloud-Based Architecture for Big Data Processing and Visualization of Traffic Data. , 2015, , .		7
3743	Digital patient: Personalized and translational data management through the MyHealthAvatar EU project. , 2015, 2015, 1397-400.		27
3744	Realtime file processing based on Map-Reduce framework. , 2015, , .		0
3745	Dependency-aware distributed video transcoding in the cloud. , 2015, , .		10
3746	Scalable Security Event Aggregation for Situation Analysis. , 2015, , .		4
3747	Spotgres - parallel data analytics on Spot Instances. , 2015, , .		8
3748	A MapReduce Optimization Method on Hadoop Cluster. , 2015, , .		3
3749	Controlling TCP Incast congestion in data centre networks. , 2015, , .		1
3750	Seagull – A Real-Time Coflow Scheduling System. , 2015, , .		4
3751	Research on Parallel Data Stream Clustering Algorithm Based on Grid and Density. , 2015, , .		3
3752	Cost Minimization for Big Data Processing in Geo-Distributed Data Centers. Wireless Networks, 2015, , 59-78.	0.3	8

#	ARTICLE	IF	CITATIONS
3753	PADP: A parallel data possession audit model for cloud storage. , 2015, , .		0
3754	Genomic analysis with MapReduce. , 2015, , .		3
3755	A study on improvement of internet traffic measurement and analysis using Hadoop system. , 2015, , .		3
3756	Coresets for visual summarization with applications to loop closure. , 2015, , .		17
3757	Temporal Aspects of Big Data Management: State-of-the-Art Analysis and Future Research Directions. , 2015, , .		12
3758	Improving efficiency of wireless sensor networks through lightweight in-memory compression. , 2015, , .		1
3759	MR-COF: A Genetic MapReduce Configuration Optimization Framework. Lecture Notes in Computer Science, 2015, , 344-357.	1.0	11
3760	Data Quality Monitoring of Cloud Databases Based on Data Quality SLAs. , 2015, , 3-20.		5
3761	A survey on Big Data storage strategies. , 2015, , .		6
3762	Autonomous and Self Controlling Smart Objects for the Future Internet. , 2015, , .		6
3763	AEGLE: A big bio-data analytics framework for integrated health-care services. , 2015, , .		4
3764	A MapReduce Based Technique for Mining Behavioral Patterns from Sensor Data. Lecture Notes in Computer Science, 2015, , 145-153.	1.0	1
3765	Composable and efficient functional big data processing framework. , 2015, , .		3
3766	Distributed fuzzy rough prototype selection for Big Data regression. , 2015, , .		5
3767	A Fair-Rank Ant Colony Algorithm in Distributed Mass Storage System. Canadian Journal of Electrical and Computer Engineering, 2015, 38, 338-345.	1.5	2
3768	Constructing Scalable Data-Flows on Hadoop with Legacy Components. , 2015, , .		0
3769	Knowledge reduction method based on information entropy for port big data using MapReduce. , 2015, , .		2
3770	Toward Scheduling I/O Request of Mapreduce Tasks Based on Markov Model. Lecture Notes in Computer Science, 2015, , 78-89.	1.0	1

#	ARTICLE	IF	CITATIONS
3771	Fast Parallel Mining of Maximally Informative k-Itemsets in Big Data. , 2015, , .		1
3772	Node Capability Modeling for Reduce Phaseâ€™s Scheduling in MapReduce Environment. Lecture Notes in Computer Science, 2015, , 217-231.	1.0	0
3773	A novel approach for feature selection based on MapReduce for biomarker discovery. , 2015, , .		0
3774	Load balancing in MapReduce on homogeneous and heterogeneous clusters: an in-depth review. International Journal of Communication Networks and Distributed Systems, 2015, 15, 149.	0.3	5
3775	A MapReduce-based fuzzy associative classifier for big data. , 2015, , .		15
3776	QoS: A QoS-Aware Storage Cloud Based on HDFS. , 2015, , .		3
3777	Scalable Learning of k-dependence Bayesian Classifiers under MapReduce. , 2015, , .		4
3778	Bandlimited OLAP cubes for interactive big data visualization. , 2015, , .		5
3779	A Hybrid Cloud Infrastructure for Big Data Applications. , 2015, , .		5
3780	Programming Support for an Integrated Multi-Party Computation and MapReduce Infrastructure. , 2015, , .		3
3781	A Theoretical Model for Big Data Analytics using Machine Learning Algorithms. , 2015, , .		1
3782	Real-time digital forensic triaging for cloud data analysis using MapReduce on Hadoop framework. International Journal of Electronic Security and Digital Forensics, 2015, 7, 119.	0.1	1
3783	G-Storm: GPU-enabled high-throughput online data processing in Storm. , 2015, , .		20
3784	Body Parts Segmentation with Attached Props Using RGB-D Imaging. , 2015, , .		20
3785	Digital Library Recommender System on Hadoop. , 2015, , .		3
3786	Cloud Networking for Big Data. Wireless Networks, 2015, , .	0.3	2
3787	Semantic presentation and fusion framework of unstructured data in smart cities. , 2015, , .		3
3788	Bipartite Matching Based Speculative Execution to Improve Cloud MapReduce Performance. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
3789	Clash of the titans. Proceedings of the VLDB Endowment, 2015, 8, 2110-2121.	2.1	169
3790	The research for storage scheme based on Hadoop. , 2015, , .		3
3791	Parallel NoSQL Entity Resolution Approach with MapReduce. , 2015, , .		4
3792	Window-Based Feature Extraction Framework for Multi-Sensor Data: A Posture Recognition Case Study. , 0, , .		10
3793	A MapReduce-Based Algorithm for Nearest Neighbor Search with Keywords. Lecture Notes in Electrical Engineering, 2015, , 1377-1384.	0.3	1
3794	<sc>Big</sc> data and analytics in higher education: Opportunities and challenges. British Journal of Educational Technology, 2015, 46, 904-920.	3.9	359
3795	GPU-based MapReduce for large-scale near-duplicate video retrieval. Multimedia Tools and Applications, 2015, 74, 10515-10534.	2.6	10
3797	Real-Time Semiparametric Regression for Distributed Data Sets. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 545-557.	4.0	7
3798	Distributed learning for Random Vector Functional-Link networks. Information Sciences, 2015, 301, 271-284.	4.0	134
3799	Survey of Large-Scale Data Management Systems for Big Data Applications. Journal of Computer Science and Technology, 2015, 30, 163-183.	0.9	37
3800	The Who-To-Follow System at Twitter: Strategy, Algorithms, and Revenue Impact. Interfaces, 2015, 45, 98-107.	1.6	25
3801	 x^2 MapReduce: Incremental MapReduce for Mining Evolving Big Data. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 1906-1919.	4.0	40
3802	Big Data: What Is It and What Does It Mean for Cardiovascular Research and Prevention Policy. Current Cardiovascular Risk Reports, 2015, 9, 1.	0.8	6
3803	Big Data and Cloud: A Survey. Advances in Intelligent Systems and Computing, 2015, , 773-778.	0.5	9
3804	Efficient XML Data Processing Based on MapReduce Framework. Lecture Notes in Electrical Engineering, 2015, , 97-104.	0.3	0
3805	Towards a Big Data Analytics Framework for IoT and Smart City Applications. Modeling and Optimization in Science and Technologies, 2015, , 257-282.	0.7	93
3806	ScaDiPaSi: An Effective Scalable and Distributable MapReduce-Based Method to Find Patient Similarity on Huge Healthcare Networks. Big Data Research, 2015, 2, 19-27.	2.6	56
3807	CWC: A Distributed Computing Infrastructure Using Smartphones. IEEE Transactions on Mobile Computing, 2015, 14, 1587-1600.	3.9	25

#	ARTICLE	IF	CITATIONS
3808	Promises and Challenges of Big Data Computing in Health Sciences. Big Data Research, 2015, 2, 2-11.	2.6	185
3809	Designing efficient high performance server-centric data center network architecture. Computer Networks, 2015, 79, 283-296.	3.2	29
3810	Heterogeneous Cloud Framework for Big Data Genome Sequencing. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2015, 12, 166-178.	1.9	33
3811	Modeling and optimizing MapReduce programs. Concurrency Computation Practice and Experience, 2015, 27, 1734-1766.	1.4	12
3812	Sketch and Validate for Big Data Clustering. IEEE Journal on Selected Topics in Signal Processing, 2015, 9, 678-690.	7.3	23
3813	A Parallel and Incremental Approach for Data-Intensive Learning of Bayesian Networks. IEEE Transactions on Cybernetics, 2015, 45, 2890-2904.	6.2	45
3814	On the Network Power Effectiveness of Data Center Architectures. IEEE Transactions on Computers, 2015, 64, 3237-3248.	2.4	7
3815	Impact of MapReduce Policies on Job Completion Reliability and Job Energy Consumption. IEEE Transactions on Parallel and Distributed Systems, 2015, 26, 1364-1378.	4.0	10
3816	BulkAligner: A novel sequence alignment algorithm based on graph theory and Trinity. Information Sciences, 2015, 303, 120-133.	4.0	2
3817	Scalable Distributed Processing of K Nearest Neighbor Queries over Moving Objects. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 1383-1396.	4.0	59
3818	Degradation science: Mesoscopic evolution and temporal analytics of photovoltaic energy materials. Current Opinion in Solid State and Materials Science, 2015, 19, 212-226.	5.6	51
3819	PRISM: Fine-Grained Resource-Aware Scheduling for MapReduce. IEEE Transactions on Cloud Computing, 2015, 3, 182-194.	3.1	48
3820	Reactive Resource Provisioning Heuristics for Dynamic Dataflows on Cloud Infrastructure. IEEE Transactions on Cloud Computing, 2015, 3, 105-118.	3.1	46
3821	On Data Center Network Architectures for Interconnecting Dual-Port Servers. IEEE Transactions on Computers, 2015, 64, 3210-3222.	2.4	26
3822	An Iterative MapReduce Based Frequent Subgraph Mining Algorithm. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 608-620.	4.0	66
3823	Constructing compact and effective graphs for recommender systems via node and edge aggregations. Expert Systems With Applications, 2015, 42, 3396-3409.	4.4	9
3824	Hypergraph partitioning for multiple communication cost metrics: Model and methods. Journal of Parallel and Distributed Computing, 2015, 77, 69-83.	2.7	39
3825	Research on MapReduce Based Incremental Iterative Model and Framework. IETE Journal of Research, 2015, 61, 32-40.	1.8	1

#	ARTICLE	IF	CITATIONS
3826	A Survey of Approaches used in Parallel Architectures and Multi-core Processors, For Performance Improvement. <i>Advances in Intelligent Systems and Computing</i> , 2015, , 537-545.	0.5	6
3827	Using Memory in the Right Way to Accelerate Big Data Processing. <i>Journal of Computer Science and Technology</i> , 2015, 30, 30-41.	0.9	8
3828	Architectural support for task scheduling: hardware scheduling for dataflow on NUMA systems. <i>Journal of Supercomputing</i> , 2015, 71, 2309-2338.	2.4	2
3829	Progressive online aggregation in a distributed stream system. <i>Journal of Systems and Software</i> , 2015, 102, 146-157.	3.3	2
3830	Scaling up MapReduce-based Big Data Processing on Multi-GPU systems. <i>Cluster Computing</i> , 2015, 18, 369-383.	3.5	66
3831	Improving the performance of GIS polygon overlay computation with MapReduce for spatial big data processing. <i>Cluster Computing</i> , 2015, 18, 507-516.	3.5	46
3832	Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. <i>International Journal of Production Economics</i> , 2015, 165, 223-233.	5.1	318
3833	Automatic Taxonomy Construction from Keywords via Scalable Bayesian Rose Trees. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2015, 27, 1861-1874.	4.0	19
3834	Design and Development of a Medical Big Data Processing System Based on Hadoop. <i>Journal of Medical Systems</i> , 2015, 39, 23.	2.2	67
3835	A MapReduce Approach to Address Big Data Classification Problems Based on the Fusion of Linguistic Fuzzy Rules. <i>International Journal of Computational Intelligence Systems</i> , 2015, 8, 422.	1.6	86
3836	Distributed discovery of frequent subgraphs of a network using MapReduce. <i>Computing (Vienna/New)</i> Tj ETQq0 0 0 rgBT /Overlock 10 T	3.2	19
3837	SAMES: deadline-constraint scheduling in MapReduce. <i>Frontiers of Computer Science</i> , 2015, 9, 128-141.	1.6	9
3838	Identifying Recurring Faulty Functions in Field Traces of a Large Industrial Software System. <i>IEEE Transactions on Reliability</i> , 2015, 64, 269-283.	3.5	4
3839	Efficient community identification and maintenance at multiple resolutions on distributed datastores. <i>Data and Knowledge Engineering</i> , 2015, 100, 133-147.	2.1	0
3840	Task partitioning optimization algorithm for energy saving and load balance on NoC-based MPSoCs. , 2015, , .		2
3841	Efficient Motif Discovery for Large-Scale Time Series in Healthcare. <i>IEEE Transactions on Industrial Informatics</i> , 2015, 11, 583-590.	7.2	57
3842	An architecture and platform for developing distributed recommendation algorithms on large-scale social networks. <i>Journal of Information Science</i> , 2015, 41, 686-704.	2.0	24
3843	B.Hive: A Zero Configuration Forms Honeytrap for Productive Web Applications. <i>IFIP Advances in Information and Communication Technology</i> , 2015, , 267-280.	0.5	1

#	ARTICLE	IF	CITATIONS
3844	Concrete Models and Empirical Evaluations for the Categorical Compositional Distributional Model of Meaning. <i>Computational Linguistics</i> , 2015, 41, 71-118.	2.5	24
3845	Understandable Big Data: A survey. <i>Computer Science Review</i> , 2015, 17, 70-81.	10.2	237
3846	Privacy-Preserving Scanning of Big Content for Sensitive Data Exposure with MapReduce. , 2015, , .		40
3847	Elastic Stream Processing with Latency Guarantees. , 2015, , .		78
3848	Meta-MapReduce for scalable data mining. <i>Journal of Big Data</i> , 2015, 2, .	6.9	14
3849	Application of Machine Learning for Multicenter Learning. , 2015, , 71-97.		0
3850	Accounting and Accountability in Content Distribution Architectures. <i>ACM Computing Surveys</i> , 2015, 47, 1-35.	16.1	19
3851	Parallel Algorithms for Constrained Tensor Factorization via Alternating Direction Method of Multipliers. <i>IEEE Transactions on Signal Processing</i> , 2015, 63, 5450-5463.	3.2	83
3852	Design adaptive task allocation scheduler to improve MapReduce performance in heterogeneous clouds. <i>Journal of Network and Computer Applications</i> , 2015, 57, 61-70.	5.8	42
3853	ParaDIME: Parallel Distributed Infrastructure for Minimization of Energy for data centers. <i>Microprocessors and Microsystems</i> , 2015, 39, 1174-1189.	1.8	8
3854	Node anomaly detection for homogeneous distributed environments. <i>Expert Systems With Applications</i> , 2015, 42, 7012-7025.	4.4	4
3855	Optimistic Recovery for Iterative Dataflows in Action. , 2015, , .		13
3856	REEF. , 2015, 2015, 1343-1355.		11
3857	Making Real Time Data Analytics Available as a Service. , 2015, , .		8
3858	Quartz. <i>Computer Communication Review</i> , 2015, 44, 283-294.	1.5	19
3859	A Top-Down Parallel Semisort. , 2015, , .		19
3860	A time-â€‘energy performance analysis of MapReduce on heterogeneous systems with GPUs. <i>Performance Evaluation</i> , 2015, 91, 255-269.	0.9	5
3861	Dynamic Provisioning of Data Intensive Computing Middleware Frameworks. , 2015, , .		1

#	ARTICLE	IF	CITATIONS
3862	OpenDBDDAS Toolkit: Secure MapReduce and Hadoop-like Systems. <i>Procedia Computer Science</i> , 2015, 51, 1675-1684.	1.2	2
3863	Peacock. <i>ACM Transactions on Intelligent Systems and Technology</i> , 2015, 6, 1-23.	2.9	28
3864	Parallelizing fully homomorphic encryption for a cloud environment. <i>Journal of Applied Research and Technology</i> , 2015, 13, 245-252.	0.6	22
3865	On scalable parallel recursive backtracking. <i>Journal of Parallel and Distributed Computing</i> , 2015, 84, 65-75.	2.7	14
3866	Ensembles of incremental learners to detect anomalies in ad hoc sensor networks. <i>Ad Hoc Networks</i> , 2015, 35, 14-36.	3.4	60
3867	Locality and loading aware virtual machine mapping techniques for optimizing communications in MapReduce applications. <i>Future Generation Computer Systems</i> , 2015, 53, 43-54.	4.9	45
3868	A fast map-reduce algorithm for burst errors in big data cloud storage. , 2015, , .		4
3869	Evolutionary Replicative Data Reorganization with Prioritization for Efficient Workload Processing. <i>Procedia Computer Science</i> , 2015, 51, 2357-2366.	1.2	1
3870	Accelerating Big Data analytics with Collaborative Planning in Teradata Aster 6. , 2015, , .		2
3871	Large-scale Predictive Analytics in Vertica. , 2015, , .		11
3872	MzJava: An open source library for mass spectrometry data processing. <i>Journal of Proteomics</i> , 2015, 129, 63-70.	1.2	25
3873	Making mapreduce scheduling effective in erasure-coded storage clusters. , 2015, , .		0
3874	Enabling Concurrent Failure Recovery for Regenerating-Coding-Based Storage Systems: From Theory to Practice. <i>IEEE Transactions on Computers</i> , 2015, 64, 1898-1911.	2.4	20
3875	Cloud-Based Programmable Sensor Data Provision. , 2015, , .		0
3876	Application of description logic learning in abnormal behaviour detection in smart homes. , 2015, , .		1
3877	Efficient Similarity Join Based on Earth Mover's Distance Using MapReduce. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2015, 27, 2148-2162.	4.0	9
3878	Ranking User Tags in Micro-Blogging Website. , 2015, , .		11
3879	A New Middleware for Distributed Data Processing in CUBRID DBMS. <i>Procedia Computer Science</i> , 2015, 52, 654-658.	1.2	1

#	ARTICLE	IF	CITATIONS
3880	A novel big data architecture in support of ADS-B data analytic. , 2015, , .		15
3881	Aura: An IoT Based Cloud Infrastructure for Localized Mobile Computation Outsourcing. , 2015, , .		45
3882	Yadoop: An Elastic Resource Management Solution of YARN. , 2015, , .		1
3883	Service-oriented mobile malware detection system based on mining strategies. Pervasive and Mobile Computing, 2015, 24, 101-116.	2.1	23
3884	High-Performance Biomedical Association Mining with MapReduce. , 2015, , .		2
3885	FACADE. , 2015, , .		61
3886	BigDancing. , 2015, , .		94
3887	Cost-based Fault-tolerance for Parallel Data Processing. , 2015, , .		15
3888	Forecasting the 2013â€“2014 Influenza Season Using Wikipedia. PLoS Computational Biology, 2015, 11, e1004239.	1.5	122
3889	A Novel Approach for Improving Security and Storage Efficiency on HDFS. Procedia Computer Science, 2015, 52, 631-635.	1.2	8
3890	Classifying Uncertain and Evolving Data Streams with Distributed Extreme Learning Machine. Journal of Computer Science and Technology, 2015, 30, 874-887.	0.9	13
3891	Collaborative filtering recommendation algorithm based on Hadoop and Spark. , 2015, , .		15
3892	Large-Scale Schema-Free Data Deduplication Approach with Adaptive Sliding Window Using MapReduce. Computer Journal, 2015, 58, 3187-3201.	1.5	20
3893	Multi-aspect-streaming tensor analysis. Knowledge-Based Systems, 2015, 89, 332-345.	4.0	17
3895	Big Data Processing Systems: State-of-the-Art and Open Challenges. , 2015, , .		3
3896	Big data analytics for logistics and transportation. , 2015, , .		13
3897	FatTreeSim. , 2015, , .		12
3898	ROSEFW-RF: The winner algorithm for the ECBDLâ€™14 big data competition: An extremely imbalanced big data bioinformatics problem. Knowledge-Based Systems, 2015, 87, 69-79.	4.0	116

#	ARTICLE	IF	CITATIONS
3899	A Study of Cloud/IX Operating System for the ARM-based Data Center Server Platform. <i>Procedia Engineering</i> , 2015, 100, 1696-1705.	1.2	1
3900	Aggregation and multidimensional analysis of big data for large-scale scientific applications. , 2015, , .		14
3901	Supporting sense-making and decision-making through time evolution analysis of open sources. , 2015, , .		1
3902	Live datastore transformation for optimizing big data applications in cloud environments. , 2015, , .		6
3903	Fault Tolerant Computation with the Sparse Grid Combination Technique. <i>SIAM Journal of Scientific Computing</i> , 2015, 37, C331-C353.	1.3	14
3904	An Energy-Efficient Nonvolatile In-Memory Computing Architecture for Extreme Learning Machine by Domain-Wall Nanowire Devices. <i>IEEE Nanotechnology Magazine</i> , 2015, 14, 998-1012.	1.1	71
3905	Big other: Surveillance Capitalism and the Prospects of an Information Civilization. <i>Journal of Information Technology</i> , 2015, 30, 75-89.	2.5	1,707
3906	Research on Performance Comparison of Data Center Between PM and VM. <i>Lecture Notes in Computer Science</i> , 2015, , 226-238.	1.0	0
3907	Bayesian computation: a summary of the current state, and samples backwards and forwards. <i>Statistics and Computing</i> , 2015, 25, 835-862.	0.8	130
3908	Improved collaborative filtering with intensity-based contraction. <i>Journal of Ambient Intelligence and Humanized Computing</i> , 2015, 6, 661-674.	3.3	6
3909	Efficient Prototyping of Fault Tolerant Map-Reduce Applications with Docker-Hadoop. , 2015, , .		4
3910	Field-Programmable Wiring Systems. <i>Proceedings of the IEEE</i> , 2015, 103, 1159-1180.	16.4	4
3911	Willow: Saving Data Center Network Energy for Network-Limited Flows. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2015, 26, 2610-2620.	4.0	18
3912	Computable Bounds in Fork-Join Queueing Systems. , 2015, , .		15
3913	Context-aware Scheduling for Apache Hadoop over Pervasive Environments. <i>Procedia Computer Science</i> , 2015, 52, 202-209.	1.2	8
3914	A Parallel Approach for Decision Trees Learning from Big Data Streams. <i>Lecture Notes in Business Information Processing</i> , 2015, , 3-15.	0.8	1
3915	Deriving All Minimal Hitting Sets Based on Join Relation. <i>IEEE Transactions on Systems, Man, and Cybernetics: Systems</i> , 2015, 45, 1063-1076.	5.9	12
3916	MR-RBAT: Anonymizing Large Transaction Datasets Using MapReduce. <i>Lecture Notes in Computer Science</i> , 2015, , 3-18.	1.0	0

#	ARTICLE	IF	CITATIONS
3917	Community discovery by propagating local and global information based on the MapReduce model. Information Sciences, 2015, 323, 73-93.	4.0	68
3918	Effective Information Retrieval Approach based on Parallel Matrix Method and MapReduce Framework. , 2015, , .		1
3919	An Efficient Algorithm for Discovering Motifs in Large DNA Data Sets. IEEE Transactions on Nanobioscience, 2015, 14, 535-544.	2.2	18
3920	The performance evaluation of k-means by two MapReduce frameworks, Hadoop vs. Twister. , 2015, , .		4
3921	Product whole life-cycle and omni-channels data convergence oriented enterprise networks integration in a sensing environment. Computers in Industry, 2015, 70, 23-45.	5.7	48
3922	Accelerating wrapper-based feature selection with K-nearest-neighbor. Knowledge-Based Systems, 2015, 83, 81-91.	4.0	138
3923	A geoprocessing workflow system for environmental monitoring and integrated modelling. Environmental Modelling and Software, 2015, 69, 128-140.	1.9	70
3924	ExPregel: a new computational model for large-scale graph processing. Concurrency Computation Practice and Experience, 2015, 27, 4954-4969.	1.4	16
3925	A data parallel approach to modelling and simulation of large crowd. Cluster Computing, 2015, 18, 1307-1316.	3.5	8
3926	Heterogeneous Environment Aware Streaming Graph Partitioning. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 1560-1572.	4.0	31
3928	Big Data service engine (BISE): Integration of Big Data technologies for human centric wellness data. , 2015, , .		6
3929	Semantic client-side approach for web personalization of SaaS-based cloud services. Concurrency Computation Practice and Experience, 2015, 27, 2144-2169.	1.4	5
3930	A cloudification methodology for multidimensional analysis: Implementation and application to a railway power simulator. Simulation Modelling Practice and Theory, 2015, 55, 46-62.	2.2	8
3931	SSFile: A novel column-store for efficient data analysis in Hadoop-based distributed systems. Information Sciences, 2015, 316, 68-86.	4.0	8
3932	MRDataCube: Data cube computation using MapReduce. , 2015, , .		13
3933	DTN MapEx: Disaster area mapping through distributed computing over a Delay Tolerant Network. , 2015, , .		29
3934	Fair scheduling of bag-of-tasks applications on large-scale platforms. Future Generation Computer Systems, 2015, 49, 28-44.	4.9	13
3935	Compositional reasoning about active objects with shared futures. Formal Aspects of Computing, 2015, 27, 551-572.	1.4	29

#	ARTICLE	IF	CITATIONS
3936	Efficient distributed subgraph similarity matching. VLDB Journal, 2015, 24, 369-394.	2.7	28
3937	A generalized topic modeling approach for automatic document annotation. International Journal on Digital Libraries, 2015, 16, 111-128.	1.1	29
3938	Storage schema and ontology-independent SPARQL to HiveQL translation. Journal of Supercomputing, 2015, 71, 2694-2719.	2.4	7
3939	A token-based authentication security scheme for Hadoop distributed file system using elliptic curve cryptography. Journal of Computer Virology and Hacking Techniques, 2015, 11, 137-142.	1.6	15
3940	Classification Framework of MapReduce Scheduling Algorithms. ACM Computing Surveys, 2015, 47, 1-38.	16.1	49
3941	aTRAM - automated target restricted assembly method: a fast method for assembling loci across divergent taxa from next-generation sequencing data. BMC Bioinformatics, 2015, 16, 98.	1.2	78
3942	Data processing of high-rate low-voltage distribution grid recordings for smart grid monitoring and analysis. Eurasip Journal on Advances in Signal Processing, 2015, 2015, .	1.0	29
3943	Assessing the benefits of search-based approaches when designing self-adaptive systems: a controlled experiment. Journal of Software Engineering Research and Development, 2015, 3, .	1.0	2
3944	A survey on platforms for big data analytics. Journal of Big Data, 2015, 2, 8.	6.9	269
3945	Contextual anomaly detection framework for big sensor data. Journal of Big Data, 2015, 2, .	6.9	85
3946	Taxonomies of workflow scheduling problem and techniques in the cloud. Future Generation Computer Systems, 2015, 52, 1-12.	4.9	84
3947	CHive: Bandwidth Optimized Continuous Querying in Distributed Clouds. IEEE Transactions on Cloud Computing, 2015, 3, 219-232.	3.1	10
3948	A novel parallel scheme for fast similarity search in large time series. China Communications, 2015, 12, 129-140.	2.0	8
3949	A Survey of Big Data Analytics in Healthcare and Government. Procedia Computer Science, 2015, 50, 408-413.	1.2	218
3950	Workflow scheduling in cloud: a survey. Journal of Supercomputing, 2015, 71, 3373-3418.	2.4	188
3951	Big Data Parallelism: Issues in different X-Information Paradigms. Procedia Computer Science, 2015, 50, 395-400.	1.2	8
3952	Exploring Performance Models of Hadoop Applications on Cloud Architecture. , 2015, , .		13
3953	Big data parallelism: Challenges in different computational paradigms. , 2015, , .		3

#	ARTICLE	IF	CITATIONS
3954	Accelerating Iterative Big Data Computing Through MPI. Journal of Computer Science and Technology, 2015, 30, 283-294.	0.9	13
3955	A MapReduce-based approach for shortest path problem in large-scale networks. Engineering Applications of Artificial Intelligence, 2015, 41, 151-165.	4.3	40
3956	Metrics and Algorithms for Routing Questions to User Communities. ACM Transactions on Information Systems, 2015, 33, 1-29.	3.8	6
3957	Review on Cloud Computing Application in P2P Video Streaming. Procedia Computer Science, 2015, 50, 185-190.	1.2	5
3958	Optimizing OLAP Cubes Construction by Improving Data Placement on Multi-nodes Clusters. , 2015, , .		4
3959	Schemes for Fast Transmission of Flows in Data Center Networks. IEEE Communications Surveys and Tutorials, 2015, 17, 1391-1422.	24.8	39
3960	Graph Processing Platforms at Scale: Practices and Experiences. , 2015, , .		11
3961	Processing and Analytics of Big Data Streams with Yahoo!S4. , 2015, , .		22
3963	MapReduce implementation for minimum reduct using parallel genetic algorithm. , 2015, , .		4
3964	Accelerating Big Data Processing on Modern Clusters. , 2015, , .		2
3965	Smart Cache: An Optimized MapReduce Implementation of Frequent Itemset Mining. , 2015, , .		6
3966	Management and application of mobile big data. International Journal of Embedded Systems, 2015, 7, 63.	0.2	19
3967	Near Duplicate Image Discovery on One Billion Images. , 2015, , .		9
3968	A framework for processing large scale geospatial and remote sensing data in MapReduce environment. Computers and Graphics, 2015, 49, 37-46.	1.4	45
3969	Parallel k-modes algorithm based on MapReduce. , 2015, , .		1
3970	Handling Big Data Efficiently by Using Map Reduce Technique. , 2015, , .		23
3971	NNB: An efficient nearest neighbor search method for hierarchical clustering on large datasets. , 2015, , .		5
3972	Improving Energy Efficiency of IO-Intensive MapReduce Jobs. , 2015, , .		4

#	ARTICLE	IF	CITATIONS
3973	PAXQuery: Efficient Parallel Processing of Complex XQuery. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 1977-1991.	4.0	9
3974	Computational fluid dynamics simulation based on Hadoop Ecosystem and heterogeneous computing. Computers and Fluids, 2015, 115, 1-10.	1.3	6
3975	Iterative sampling based frequent itemset mining for big data. International Journal of Machine Learning and Cybernetics, 2015, 6, 875-882.	2.3	19
3976	Graph relationship discovery using Pregel computing model. , 2015, , .		1
3977	SparkXS: Efficient Access Control for Intelligent and Large-Scale Streaming Data Applications. , 2015, , .		5
3978	Workload characteristics and resource aware Hadoop scheduler. , 2015, , .		5
3979	Performance modelling and analysis of mapreduce/hadoop workloads. , 2015, , .		4
3980	Semantic Mining of Social Networks. Synthesis Lectures on the Semantic Web: Theory and Technology, 2015, 5, 1-205.	5.0	0
3981	OSDT: A scalable application-level scheduling scheme for TCP Incast problem. , 2015, , .		3
3982	Simulation Runner: A Cloud-Based Parallel and Distributed HPC Platform. , 2015, , .		1
3984	Gossip-based spectral clustering of distributed data streams. , 2015, , .		3
3985	Entity Resolution in the Web of Data. Synthesis Lectures on the Semantic Web: Theory and Technology, 2015, 5, 1-122.	5.0	44
3986	FINGER: A novel erasure coding scheme using fine granularity blocks to improve Hadoop write and update performance. , 2015, , .		5
3987	Discretionary Information Flow Control for Interaction-Oriented Specifications. Lecture Notes in Computer Science, 2015, , 427-450.	1.0	5
3988	Performance evaluation and tuning for MapReduce computing in Hadoop distributed file system. , 2015, , .		8
3989	EHadoop: Network I/O Aware Scheduler for Elastic MapReduce Cluster. , 2015, , .		9
3990	On Datacenter-Network-Aware Load Balancing in MapReduce. , 2015, , .		1
3991	SpatialHadoop: A MapReduce framework for spatial data. , 2015, , .		375

#	ARTICLE	IF	CITATIONS
3992	A Pattern Specification and Optimizations Framework for Accelerating Scientific Computations on Heterogeneous Clusters. , 2015, , .		3
3993	SMapReduce: Optimising Resource Allocation by Managing Working Slots at Runtime. , 2015, , .		1
3994	SPLATT: Efficient and Parallel Sparse Tensor-Matrix Multiplication. , 2015, , .		137
3995	Cashmere: Heterogeneous Many-Core Computing. , 2015, , .		10
3996	On fast timing closure: speeding up incremental path-based timing analysis with mapreduce. , 2015, , .		5
3997	Boosting Degraded Reads in Heterogeneous Erasure-Coded Storage Systems. IEEE Transactions on Computers, 2015, 64, 2145-2157.	2.4	19
3998	An Integrated Approach to Sequence-Independent Local Alignment of Protein Binding Sites. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2015, 12, 298-308.	1.9	0
3999	Digital forensics as a service: Game on. Digital Investigation, 2015, 15, 20-38.	3.2	46
4000	A MapReduce scratchpad memory for multi-core cloud computing applications. Microprocessors and Microsystems, 2015, 39, 599-608.	1.8	9
4001	The general model of secure computation system. , 2015, , .		0
4002	Leveraging Large Data with Weak Supervision for Joint Feature and Opinion Word Extraction. Journal of Computer Science and Technology, 2015, 30, 903-916.	0.9	3
4003	Research on performance optimization and visualization tool of Hadoop. , 2015, , .		3
4004	Aggregate Programming for the Internet of Things. Computer, 2015, 48, 22-30.	1.2	138
4005	Towards an Efficient and Distributed DBSCAN Algorithm Using MapReduce. Lecture Notes in Business Information Processing, 2015, , 75-90.	0.8	2
4006	Distributed Graph Layout with Spark. , 2015, , .		11
4007	Optical multicast system for data center networks. Optics Express, 2015, 23, 22162.	1.7	37
4008	Optimizing data placement in heterogeneous Hadoop clusters. Cluster Computing, 2015, 18, 1465-1480.	3.5	33
4009	MapReduce Based Text Detection in Big Data Natural Scene Videos. Procedia Computer Science, 2015, 53, 216-223.	1.2	25

#	ARTICLE	IF	CITATIONS
4010	Overcoming Hadoop Scaling Limitations through Distributed Task Execution. , 2015, , .		39
4011	Simulating Cheated Results Dissemination for Volunteer Computing. , 2015, , .		4
4012	Accelerating Large-scale Image Retrieval on Heterogeneous Architectures with Spark. , 2015, , .		6
4013	A distributed frequent itemset mining algorithm using Spark for Big Data analytics. Cluster Computing, 2015, 18, 1493-1501.	3.5	73
4014	Large Java arrays and their applications. , 2015, , .		2
4015	Exemplar-based large vocabulary speech recognition using k-nearest neighbors. , 2015, , .		2
4016	Refining Social Graph Connectivity via Shortcut Edge Addition. ACM Transactions on Knowledge Discovery From Data, 2015, 10, 1-35.	2.5	33
4017	Data de duplication using NOSQL Databases in Cloud. , 2015, , .		4
4018	ActCap: Accelerating MapReduce on heterogeneous clusters with capability-aware data placement. , 2015, , .		18
4019	Processing geo-dispersed big data in an advanced mapreduce framework. IEEE Network, 2015, 29, 24-30.	4.9	8
4020	Redundancy in Linked Data Partitioning for Efficient Query Evaluation. , 2015, , .		2
4021	SAASFEE. Proceedings of the VLDB Endowment, 2015, 8, 1892-1895.	2.1	17
4022	The Parallel Improved Apriori Algorithm Research Based on Spark. , 2015, , .		17
4023	Sherlock's Toolkit: A forensic DNA analysis system. , 2015, , .		11
4024	Distributed Maximal Clique Computation and Management. IEEE Transactions on Services Computing, 2015, , 1-1.	3.2	24
4025	CEPSim: A Simulator for Cloud-Based Complex Event Processing. , 2015, , .		10
4026	Distributed Real-Time Event Analysis. , 2015, , .		8
4027	A MapReduce-based algorithm for parallelizing collusion detection in Hadoop. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
4028	Privacy and Big Data. Computer Communications and Networks, 2015, , 37-55.	0.8	5
4030	MR-CLOPE: A MapReduce based transactional clustering algorithm for DNS query log analysis. Journal of Central South University, 2015, 22, 3485-3494.	1.2	2
4031	Phylogenetic Analysis Using MapReduce Programming Model. , 2015, , .		1
4032	Cross-Layer Scheduling in Cloud Systems. , 2015, , .		8
4033	Hybrid human-machine information systems: Challenges and opportunities. Computer Networks, 2015, 90, 5-13.	3.2	23
4034	Efficient k-Nearest Neighbors Search in High Dimensions Using MapReduce. , 2015, , .		13
4035	CHCF: A Cloud-Based Heterogeneous Computing Framework for Large-Scale Image Retrieval. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25, 1900-1913.	5.6	13
4036	Application of a Key-Value Paradigm to Logic Factoring. Proceedings of the IEEE, 2015, 103, 2076-2092.	16.4	2
4037	Towards intelligent GIServices. Earth Science Informatics, 2015, 8, 463-481.	1.6	53
4038	Extending Scientific Workflow Systems to Support MapReduce Based Applications in the Cloud. , 2015, , .		2
4039	A time-efficient connected densest subgraph discovery algorithm for big data. , 2015, , .		5
4040	User-Constraint and Self-Adaptive Fault Tolerance for Event Stream Processing Systems. , 2015, , .		13
4041	Parallel overlapping community discovery based on grey relational analysis. , 2015, , .		2
4042	Cloud Computing Research Analysis Using Bibliometric Method. International Journal of Software Engineering and Knowledge Engineering, 2015, 25, 551-571.	0.6	10
4043	Evolving Classifier TEDAClass for Big Data. Procedia Computer Science, 2015, 53, 9-18.	1.2	15
4044	Large-scale cluster management at Google with Borg. , 2015, , .		754
4045	MALT. , 2015, , .		37
4046	Constructing Bisimulation Summaries on a Multi-Core Graph Processing Framework. , 2015, , .		4

#	ARTICLE	IF	CITATIONS
4047	Real-time network anomaly detection system using machine learning. , 2015, , .		42
4048	A Distributed Ensemble Scheme for nonlinear Support Vector Machine. , 2015, , .		0
4049	Joint Structural Learning to Rank with Deep Linear Feature Learning. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 2756-2769.	4.0	6
4050	Bio-inspired clustering: Basic features and future trends in the era of Big Data. , 2015, , .		8
4051	A Cloud Computing Course. , 2015, , .		5
4052	Set Cover at Web Scale. , 2015, , .		9
4053	Analysis and prediction of natural disaster using spatial data mining technique. , 2015, , .		7
4054	A Parallel Method for Scalable Anonymization of Transaction Data. , 2015, , .		3
4055	A Cooperative Coevolution Framework for Parallel Learning to Rank. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 3152-3165.	4.0	13
4056	Large-Scale Image Retrieval Based on Compressed Camera Identification. IEEE Transactions on Multimedia, 2015, 17, 1439-1449.	5.2	29
4057	Mining Big Data for Detecting, Extracting and Recommending Architectural Design Concepts. , 2015, , .		5
4058	Big Data Collection and Analysis Framework Research for Public Digital Culture Sharing Service. , 2015, , .		4
4059	BigDataDIRAC: Deploying Distributed Big Data Applications. , 2015, , .		1
4060	Creating a Data Science Platform for Developing Complication Risk Models for Personalized Treatment Planning in Radiation Oncology. , 2015, , .		6
4061	Cognitive Devanagari (Marathi) Text-to-Speech System. , 2015, , .		3
4062	Admission control in YARN clusters based on dynamic resource reservation. , 2015, , .		6
4063	Hadoop eco system for big data security and privacy. , 2015, , .		18
4064	Load balancing in data center networks with folded-Clos architectures. , 2015, , .		10

#	ARTICLE	IF	CITATIONS
4065	A Scheduling Framework for Periodic Tasks in Geo-Distributed Data Centers. , 2015, , .		1
4066	Cost-Efficient High-Performance Internet-Scale Data Analytics over Multi-cloud Environments. , 2015, , .		4
4067	Honeypot based unauthorized data access detection in MapReduce systems. , 2015, , .		8
4068	Spectral clustering of large-scale communities via random sketching and validation. , 2015, , .		1
4069	Collective citizens' behavior modelling with support of the Internet of Things and Big Data. , 2015, , .		2
4070	Image filtering with MapReduce in pseudo-distribution mode. , 2015, , .		1
4071	Tengu: An Experimentation Platform for Big Data Applications. , 2015, , .		9
4072	SparkBench. , 2015, , .		117
4073	Building Discriminative User Profiles for Large-scale Content Recommendation. , 2015, , .		13
4074	User Modeling for a Personal Assistant. , 2015, , .		43
4075	CommGuard. , 2015, , .		6
4076	Efficient batch processing of proximity queries with MapReduce. , 2015, , .		2
4077	Accelerated Path-Based Timing Analysis with MapReduce. , 2015, , .		12
4078	A Demonstration of Rubato DB. , 2015, , .		6
4079	ShareInsights. , 2015, , .		3
4080	Automated Assessment of Secure Search Systems. Operating Systems Review (ACM), 2015, 49, 22-30.	1.5	11
4081	Mahout on heterogeneous clusters using HadoopCL. , 2015, , .		1
4082	An evaluation and analysis of graph processing frameworks on five key issues. , 2015, , .		12

#	ARTICLE	IF	CITATIONS
4083	DARE. , 2015, , .		70
4084	Space and Time Efficient Parallel Graph Decomposition, Clustering, and Diameter Approximation. , 2015, , .		11
4085	Lasp. , 2015, , .		30
4086	Approximating closed fork-join queueing networks using product-form stochastic Petri-nets. Journal of Systems and Software, 2015, 110, 264-278.	3.3	9
4088	Big data analytics: a literature review. Journal of Management Analytics, 2015, 2, 175-201.	1.6	67
4089	Distributed Personalization. , 2015, , .		2
4090	Parallel K-prototypes for Clustering Big Data. Lecture Notes in Computer Science, 2015, , 628-637.	1.0	3
4091	An Empirical Study on Quality Issues of Production Big Data Platform. , 2015, , .		22
4092	A Scalable Hierarchical Clustering Algorithm Using Spark. , 2015, , .		29
4093	Elastic provisioning of virtual Hadoop clusters in OpenStack-based Clouds. , 2015, , .		6
4094	CUDA-enabled Hadoop cluster for Sparse Matrix Vector Multiplication. , 2015, , .		3
4095	MapReduce: Simplified Data Analysis of Big Data. Procedia Computer Science, 2015, 57, 563-571.	1.2	52
4096	Hadoop scalable Video Transcoding technique in cloud environment. , 2015, , .		2
4097	Efficient path query and reasoning method based on rare axis. Transactions of Tianjin University, 2015, 21, 278-283.	3.3	0
4098	An Efficient Electricity Generation Forecasting System Using Artificial Neural Network Approach with Big Data. , 2015, , .		7
4099	Research of road disease big data process based on Hadoop system. , 2015, , .		1
4100	Location Analytics as a Service: Providing Insights for Heterogeneous Spatiotemporal Data. , 2015, , .		5
4101	An evaluation on big data generalization using k-Anonymity algorithm on cloud. , 2015, , .		10

#	ARTICLE	IF	CITATIONS
4102	Large-scale machine learning based on functional networks for biomedical big data with high performance computing platforms. <i>Journal of Computational Science</i> , 2015, 11, 69-81.	1.5	49
4103	Pivot tracing. , 2015, , .		126
4104	Datix. <i>Computer Communication Review</i> , 2015, 45, 21-28.	1.5	13
4105	Big Data Analytics of Social Networks for the Discovery of "Following" Patterns. <i>Lecture Notes in Computer Science</i> , 2015, , 123-135.	1.0	48
4106	Do You Feel the Lag of Your Hadoop?. , 2015, , .		1
4107	Toward an SDN-enabled big data platform for social TV analytics. <i>IEEE Network</i> , 2015, 29, 43-49.	4.9	41
4108	Another Look at Secure Big Data Processing: Formal Framework and a Potential Approach. , 2015, , .		1
4109	Improving Serviceability for Virtual Clusters in Bandwidth-Constrained Datacenters. , 2015, , .		4
4110	Research on the Performance of Mining Packets of Educational Network for Malware Detection between PM and VM. , 2015, , .		1
4111	Optimal communication structures for big data aggregation. , 2015, , .		14
4112	Autonomous Orchestration of Distributed Discrete Event Simulations in the Presence of Resource Uncertainty. <i>ACM Transactions on Autonomous and Adaptive Systems</i> , 2015, 10, 1-20.	0.4	4
4113	DPPACS: A Novel Data Partitioning and Placement Aware Computation Scheduling Scheme for Data-Intensive Cloud Applications. <i>Computer Journal</i> , 0, , bxv062.	1.5	5
4114	Scalable collective spatial keyword query. , 2015, , .		6
4115	A framework for big data as a service. , 2015, , .		5
4116	LFRTTrainer: Large-Scale Face Recognition Training System. , 2015, , .		0
4117	Aggregation on the fly: reducing traffic for big data in the cloud. <i>IEEE Network</i> , 2015, 29, 17-23.	4.9	15
4118	A MapReduce Based Distributed Framework for Similarity Search in Healthcare Big Data Environment. <i>Lecture Notes in Computer Science</i> , 2015, , 173-182.	1.0	1
4119	MPJ Express Meets YARN: Towards Java HPC on Hadoop Systems. <i>Procedia Computer Science</i> , 2015, 51, 2678-2682.	1.2	3

#	ARTICLE	IF	CITATIONS
4120	Dynamic Price Prediction for Amazon Spot Instances. , 2015, , .		33
4121	Metadata-as-a-Service. , 2015, , .		6
4122	Preserving Privacy in MapReduce Based Clouds: Insight into Frameworks and Approaches. , 2015, , .		2
4123	Hadoop Image Processing Framework. , 2015, , .		15
4124	HM: A Column-Oriented MapReduce System on Hybrid Storage. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 3304-3317.	4.0	7
4125	A Study of Data Locality in YARN. , 2015, , .		9
4126	Opass: Analysis and Optimization of Parallel Data Access on Distributed File Systems. , 2015, , .		15
4127	Towards building an analytics platform in the cloud. , 2015, , .		0
4128	Towards Provenance-Based Anomaly Detection in MapReduce. , 2015, , .		18
4129	Scalable Network Traffic Classification Using Distributed Support Vector Machines. , 2015, , .		12
4130	The emergence of commodity-scale genetic manipulation. Current Opinion in Chemical Biology, 2015, 28, 150-155.	2.8	4
4131	MapReduce-DBMS: An Integration Model for Big Data Management and Optimization. Lecture Notes in Computer Science, 2015, , 430-439.	1.0	4
4132	Active Data: A programming model to manage data life cycle across heterogeneous systems and infrastructures. Future Generation Computer Systems, 2015, 53, 25-42.	4.9	16
4133	Predicting the Performance of Parallel Computing Models Using Queuing System. , 2015, , .		6
4134	Novel Data-Distribution Technique for Hadoop in Heterogeneous Cloud Environments. , 2015, , .		20
4135	Performance comparison of parallel graph coloring algorithms on BSP model using hadoop. , 2015, , .		13
4136	Increasing performance of parallel and distributed systems in high performance computing using weight based approach. , 2015, , .		5
4137	Comparitive study on healthcare prediction systems using big data. , 2015, , .		5

#	ARTICLE	IF	CITATIONS
4138	High-Performance Classification of Phishing URLs Using a Multi-modal Approach with MapReduce. , 2015, , .		6
4139	Fields for efficient analysis of big data. , 2015, , .		0
4140	Phase-Configurable Shuffle Optimization for Hadoop MapReduce. IEEE Transactions on Cloud Computing, 2015, , 1-1.	3.1	28
4141	sPCA. , 2015, , .		27
4142	Protelis. , 2015, , .		76
4143	GuardMR. , 2015, , .		37
4144	A review of metaheuristic scheduling techniques in cloud computing. Egyptian Informatics Journal, 2015, 16, 275-295.	4.4	293
4145	Aurora: Adaptive Block Replication in Distributed File Systems. , 2015, , .		11
4146	A high performance VoLTE traffic classification method using HTCondor. , 2015, , .		3
4147	BigBWA: approaching the Burrows-Wheeler aligner to Big Data technologies. Bioinformatics, 2015, 31, 4003-4005.	1.8	107
4148	Optimising Bootstrapping Algorithms Using R and Hadoop. , 2015, , .		0
4149	A Platform Based Distributed Service Framework for Large-Scale Cloud Ecosystem Development. , 2015, , .		0
4150	MapReduce-based efficient algorithm for finding large patterns. , 2015, , .		0
4151	Dynamic resource management in a MapReduce-style platform for fast data processing. , 2015, , .		0
4152	Big Data Processing Algorithms. Studies in Big Data, 2015, , 61-91.	0.8	5
4153	Sequence-Growth: A Scalable and Effective Frequent Itemset Mining Algorithm for Big Data Based on MapReduce Framework. , 2015, , .		17
4154	Cost-Aware Cloud Metering with Scalable Service Management Infrastructure. , 2015, , .		11
4155	Big data: Emerging technological paradigm and challenges. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
4156	Scalable Multimodal Search with Distributed Indexing by Sparse Hashing. , 2015, , .		3
4157	Exploiting Matrix Dependency for Efficient Distributed Matrix Computation. , 2015, , .		24
4158	A review of research on MapReduce scheduling algorithms in Hadoop. , 2015, , .		9
4159	A Distributed Approach Toward Discriminative Distance Metric Learning. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26, 2111-2122.	7.2	20
4160	An Improved Concurrent Programming Architectural Model Based on Cloud Computing. Lecture Notes in Electrical Engineering, 2015, , 1079-1086.	0.3	0
4161	Parallel Eclat for Opportunistic Mining of Frequent Itemsets. Lecture Notes in Computer Science, 2015, , 401-415.	1.0	12
4162	Scaling Machine Learning for Target Prediction in Drug Discovery using Apache Spark. , 2015, , .		7
4163	Adaptive sorted neighborhood blocking for entity matching with MapReduce. , 2015, , .		12
4164	The Management of Application of Big Data in Internet of Thing in Environmental Protection in China. , 2015, , .		7
4165	FLOWPROPHET: Generic and Accurate Traffic Prediction for Data-Parallel Cluster Computing. , 2015, , .		17
4166	Blind men and an elephant coalescing open-source, academic, and industrial perspectives on BigData. , 2015, , .		4
4167	On Understanding Big Data Impacts in Remotely Sensed Image Classification Using Support Vector Machine Methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8, 4634-4646.	2.3	71
4168	A Parallel Distributed Weka Framework for Big Data Mining Using Spark. , 2015, , .		40
4169	Optimizing Query Prices for Data-as-a-Service. , 2015, , .		6
4170	k-Means Performance Improvements with Centroid Calculation Heuristics Both for Serial and Parallel Environments. , 2015, , .		7
4171	How to Choose a Recommender System: Insights and Experiences for Large-Scale User Personalization. , 2015, , .		0
4172	Classification of multi-genomic data using MapReduce paradigm. , 2015, , .		0
4173	Efficient data preprocessing for genetic-fuzzy mining with MapReduce. , 2015, , .		2

#	ARTICLE	IF	CITATIONS
4174	DREAMS: Dynamic resource allocation for MapReduce with data skew. , 2015, , .		30
4175	Knowledge Engineering with Big Data. IEEE Intelligent Systems, 2015, 30, 46-55.	4.0	61
4176	Real time data analysis of taxi rides using StreamMine3G. , 2015, , .		2
4177	A Clustered Approach for Fast Computation of Betweenness Centrality in Social Networks. , 2015, , .		11
4178	Apache Tez. , 2015, , .		138
4179	Optimizing the MapReduce Framework for CPU-MIC Heterogeneous Cluster. Lecture Notes in Computer Science, 2015, , 33-44.	1.0	0
4180	Toward the Association Rules of Meteorological Data Mining Based on Cloud Computing. Lecture Notes in Electrical Engineering, 2015, , 1051-1059.	0.3	1
4181	Modified Fuzzy K-mean clustering using MapReduce in Hadoop and cloud. , 2015, , .		5
4182	SYNC or ASYNC: time to fuse for distributed graph-parallel computation. , 2015, , .		61
4183	Scalable and parallel machine learning algorithms for statistical data mining - Practice & experience. , 2015, , .		1
4184	MapReduce Algorithms. , 2015, , .		3
4185	Feature Selection and Classification of Microarray Data using MapReduce based ANOVA and K-Nearest Neighbor. Procedia Computer Science, 2015, 54, 301-310.	1.2	52
4186	Modeling of the resource allocation in cloud computing centers. Computer Networks, 2015, 91, 453-470.	3.2	53
4187	A Survey on Distributed File System Technology. Journal of Physics: Conference Series, 2015, 608, 012039.	0.3	8
4188	Distributed Programming over Time-Series Graphs. , 2015, , .		10
4189	Cost Estimation of Parallel Constrained Producer-Consumer Algorithms. , 2015, , .		0
4190	Approximate block coordinate descent for large scale hierarchical classification. , 2015, , .		2
4191	Efficient subspace skyline query based on user preference using MapReduce. Ad Hoc Networks, 2015, 35, 105-115.	3.4	20

#	ARTICLE	IF	CITATIONS
4192	A survey on data center networking for cloud computing. Computer Networks, 2015, 91, 528-547.	3.2	102
4193	Classification of microarray using MapReduce based proximal support vector machine classifier. Knowledge-Based Systems, 2015, 89, 584-602.	4.0	49
4194	Bankruptcy Prediction of Construction Businesses: Towards a Big Data Analytics Approach. , 2015, , .		13
4195	A Hybrid Network Architecture for Data Centers. , 2015, , .		1
4196	Benchmark of Map-Reduce Framework in MongoDB. , 2015, , .		1
4197	Integrated execution framework for catastrophe modeling. , 2015, , .		3
4198	Solutions for Processing K Nearest Neighbor Joins for Massive Data on MapReduce. , 2015, , .		18
4199	Extending a Peer-Based Coordination Model with Composable Design Patterns. , 2015, , .		2
4200	The Packing Server for real-time scheduling of MapReduce workflows. , 2015, , .		8
4201	Expressing and Composing Actors for Deterministic and Scalable Programming in Cloud. , 2015, , .		2
4202	Tentacle: distributed quantification of genes in metagenomes. GigaScience, 2015, 4, 40.	3.3	10
4203	Hadoop, MapReduce and HDFS: A Developers Perspective. Procedia Computer Science, 2015, 48, 45-50.	1.2	109
4204	Big Graph Processing Systems: State-of-the-Art and Open Challenges. , 2015, , .		6
4205	Network-Constrained Packing of Brokered Workloads in Virtualized Environments. , 2015, , .		2
4206	YARNsim: Simulating Hadoop YARN. , 2015, , .		11
4207	Confuga: Scalable Data Intensive Computing for POSIX Workflows. , 2015, , .		4
4208	Table2Graph: A Scalable Graph Construction from Relational Tables Using Map-Reduce. , 2015, , .		8
4209	VC3: Trustworthy Data Analytics in the Cloud Using SGX. , 2015, , .		380

#	ARTICLE	IF	CITATIONS
4210	Shared Memory Optimization in Virtualized Cloud. , 2015, , .		7
4211	DRS: Dynamic Resource Scheduling for Real-Time Analytics over Fast Streams. , 2015, , .		58
4212	A Hybrid Cloud Framework for Scientific Computing. , 2015, , .		0
4213	Global sensitivity analysis for large-scale socio-hydrological models using Hadoop. Environmental Modelling and Software, 2015, 73, 231-243.	1.9	25
4214	Stream Processing on Demand for Lambda Architectures. Lecture Notes in Computer Science, 2015, , 243-257.	1.0	8
4215	A survey on knowledge representation in materials science and engineering: An ontological perspective. Computers in Industry, 2015, 73, 8-22.	5.7	32
4216	A Comparative Analysis of Scheduling Mechanisms for Virtual Screening Workflow in a Shared Resource Environment. , 2015, , .		5
4217	KTV-Tree: Interactive Top-K Aggregation on Dynamic Large Dataset in the Cloud. , 2015, , .		1
4218	High performance declarative memory systems through MapReduce. , 2015, , .		2
4219	Active Manifold Learning with Twitter Big Data. Procedia Computer Science, 2015, 53, 208-215.	1.2	3
4221	Framework and challenges for Wireless body area networks based on big data. , 2015, , .		13
4222	The Collaborative Filtering Algorithm with Time Weight Based on MapReduce. Lecture Notes in Computer Science, 2015, , 386-395.	1.0	9
4223	UniCrawl: A Practical Geographically Distributed Web Crawler. , 2015, , .		14
4224	Developing a secure remote patient monitoring system. , 2015, , .		1
4225	Cloudets: Cloud-based cognition for large streaming data. , 2015, , .		4
4226	Towards bandwidth guaranteed energy efficient data center networking. Journal of Cloud Computing: Advances, Systems and Applications, 2015, 4, .	2.1	20
4227	Adaptive virtual machine assignment for multi-tenant data center networks. , 2015, , .		1
4228	Taming Non-local Stragglers Using Efficient Prefetching in MapReduce. , 2015, , .		2

#	ARTICLE	IF	CITATIONS
4229	On the architecture of a big data classification tool based on a map reduce approach for hyperspectral image analysis. , 2015, , .		6
4230	A video cloud platform combing online and offline cloud computing technologies. Personal and Ubiquitous Computing, 2015, 19, 1099-1110.	1.9	37
4231	Performance Prediction for Apache Spark Platform. , 2015, , .		91
4232	Processes Meet Big Data: Connecting Data Science with Process Science. IEEE Transactions on Services Computing, 2015, 8, 810-819.	3.2	66
4233	A MapReduce Reinforced Distributed Sequential Pattern Mining Algorithm. Lecture Notes in Computer Science, 2015, , 183-197.	1.0	7
4234	Apache Spark a Big Data Analytics Platform for Smart Grid. Procedia Technology, 2015, 21, 171-178.	1.1	109
4235	Light combinators for finite fields arithmetic. Science of Computer Programming, 2015, 111, 365-394.	1.5	1
4236	Evaluation of a Sector-Hash Based Rapid File Detection Method for Monitoring Infrastructure-as-a-Service Cloud Platforms. , 2015, , .		5
4237	Genome Data Analysis Using MapReduce Paradigm. , 2015, , .		1
4238	Efficient and Scalable SPARQL Query Processing with Transformed Table. , 2015, , .		1
4239	DBSCAN on Resilient Distributed Datasets. , 2015, , .		33
4240	Interactive and Scalable Exploration of Big Spatial Data – A Data Management Perspective. , 2015, , .		13
4241	Lightning fast and space efficient inequality joins. Proceedings of the VLDB Endowment, 2015, 8, 2074-2085.	2.1	21
4242	Scalable Data Quality for Big Data: The Pythia Framework for Handling Missing Values. Big Data, 2015, 3, 159-172.	2.1	3
4243	Contemporary challenges for data-intensive scientific workflow management systems. , 2015, , .		8
4244	An efficient framework for online virtual network embedding in virtualized cloud data centers. , 2015, , .		4
4245	Thinking Like a Vertex. ACM Computing Surveys, 2015, 48, 1-39.	16.1	216
4246	Enabling the big data analysis in the smart grid. , 2015, , .		16

#	ARTICLE	IF	CITATIONS
4247	Knowledge discovery from social media using big data-provided sentiment analysis (SoMABiT). Journal of Information Science, 2015, 41, 779-798.	2.0	25
4248	A Spark-Based Big Data Platform for Massive Remote Sensing Data Processing. Lecture Notes in Computer Science, 2015, , 120-126.	1.0	18
4249	Improving MapReduce Performance with Partial Speculative Execution. Journal of Grid Computing, 2015, 13, 587-604.	2.5	26
4250	Hamiltonian Properties of DCell Networks. Computer Journal, 2015, 58, 2944-2955.	1.5	38
4251	A Framework for Cost-Effective Scheduling of MapReduce Applications. , 2015, , .		4
4252	A high performance framework for modeling and simulation of large-scale complex systems. Future Generation Computer Systems, 2015, 51, 132-141.	4.9	19
4253	A cloud computing based system for cyber security management. International Journal of Parallel, Emergent and Distributed Systems, 2015, 30, 29-45.	0.7	18
4255	A Parallel Matrix-Based Method for Computing Approximations in Incomplete Information Systems. IEEE Transactions on Knowledge and Data Engineering, 2015, 27, 326-339.	4.0	76
4256	Big Data Storage Techniques for Spatial Databases: Implications of Big Data Architecture on Spatial Query Processing. Studies in Big Data, 2015, , 297-323.	0.8	2
4257	Flexible architecture for cluster evolution in cloud computing. Computers and Electrical Engineering, 2015, 42, 90-106.	3.0	10
4258	Robust network supercomputing with unreliable workers. Journal of Parallel and Distributed Computing, 2015, 75, 81-92.	2.7	3
4259	The Solid Architecture for real-time management of big semantic data. Future Generation Computer Systems, 2015, 47, 62-79.	4.9	37
4260	IO Cost Evaluation of OLAP Query Processing with MapReduce. Lecture Notes in Electrical Engineering, 2015, , 997-1002.	0.3	0
4261	FESTAL: Fault-Tolerant Elastic Scheduling Algorithm for Real-Time Tasks in Virtualized Clouds. IEEE Transactions on Computers, 2015, 64, 2545-2558.	2.4	83
4263	Towards cost-sensitive adaptation: When is it worth updating your predictive model?. Neurocomputing, 2015, 150, 240-249.	3.5	28
4264	Energy-efficient algorithms for distributed storage system based on block storage structure reconfiguration. Journal of Network and Computer Applications, 2015, 48, 71-86.	5.8	9
4265	Remote sensing big data computing: Challenges and opportunities. Future Generation Computer Systems, 2015, 51, 47-60.	4.9	444
4266	Big Data Technologies and Cloud Computing. , 2015, , 17-49.		11

#	ARTICLE	IF	CITATIONS
4267	MapReduce based parallel gene selection method. Applied Intelligence, 2015, 42, 147-156.	3.3	21
4268	Emerging trends and technologies in big data processing. Concurrency Computation Practice and Experience, 2015, 27, 2078-2091.	1.4	93
4269	Incentivising resource sharing in social clouds. Concurrency Computation Practice and Experience, 2015, 27, 1483-1497.	1.4	12
4270	Density-based data partitioning strategy to approximate large-scale subgraph mining. Information Systems, 2015, 48, 213-223.	2.4	32
4271	Enrichment of the phenotypic and genotypic Data Warehouse analysis using Question Answering systems to facilitate the decision making process in cereal breeding programs. Ecological Informatics, 2015, 26, 203-216.	2.3	6
4272	Active learning for sentiment analysis on data streams: Methodology and workflow implementation in the ClowdFlows platform. Information Processing and Management, 2015, 51, 187-203.	5.4	58
4273	An Adaptive and Memory Efficient Sampling Mechanism for Partitioning in MapReduce. International Journal of Parallel Programming, 2015, 43, 489-507.	1.1	33
4274	Studying the relationship between logging characteristics and the code quality of platform software. Empirical Software Engineering, 2015, 20, 1-27.	3.0	53
4275	Resource Management in Clouds: Survey and Research Challenges. Journal of Network and Systems Management, 2015, 23, 567-619.	3.3	382
4276	Ontology-based library recommender system using MapReduce. Cluster Computing, 2015, 18, 113-121.	3.5	26
4277	OFScheduler: A Dynamic Network Optimizer for MapReduce in Heterogeneous Cluster. International Journal of Parallel Programming, 2015, 43, 472-488.	1.1	22
4278	Parallel sampling from big data with uncertainty distribution. Fuzzy Sets and Systems, 2015, 258, 117-133.	1.6	23
4279	BlinkFS: A distributed file system for intermittent power. Sustainable Computing: Informatics and Systems, 2015, 6, 69-80.	1.6	1
4280	Cost-sensitive linguistic fuzzy rule based classification systems under the MapReduce framework for imbalanced big data. Fuzzy Sets and Systems, 2015, 258, 5-38.	1.6	223
4281	Paving the road toward Smart Grids through large-scale advanced metering infrastructures. Electric Power Systems Research, 2015, 120, 194-205.	2.1	63
4282	A knowledge-based platform for Big Data analytics based on publish/subscribe services and stream processing. Knowledge-Based Systems, 2015, 79, 3-17.	4.0	83
4283	MRPR: A MapReduce solution for prototype reduction in big data classification. Neurocomputing, 2015, 150, 331-345.	3.5	204
4284	Measuring the veracity of web event via uncertainty. Journal of Systems and Software, 2015, 102, 226-236.	3.3	13

#	ARTICLE	IF	CITATIONS
4285	DynamicCloudSim: Simulating heterogeneity in computational clouds. <i>Future Generation Computer Systems</i> , 2015, 46, 85-99.	4.9	54
4286	Cloud-Based Grasp Analysis and Planning for Toleranced Parts Using Parallelized Monte Carlo Sampling. <i>IEEE Transactions on Automation Science and Engineering</i> , 2015, 12, 455-470.	3.4	25
4287	Web technologies for environmental Big Data. <i>Environmental Modelling and Software</i> , 2015, 63, 185-198.	1.9	184
4288	Multi-criteria scheduling of Bag-of-Tasks applications on heterogeneous interlinked clouds with simulated annealing. <i>Journal of Systems and Software</i> , 2015, 101, 1-14.	3.3	64
4289	Anomaly detection in traffic using L1-norm minimization extreme learning machine. <i>Neurocomputing</i> , 2015, 149, 415-425.	3.5	48
4290	Elastic extreme learning machine for big data classification. <i>Neurocomputing</i> , 2015, 149, 464-471.	3.5	77
4291	Cost-sensitive learning of hierarchical tree classifiers for large-scale image classification and novel category detection. <i>Pattern Recognition</i> , 2015, 48, 1673-1687.	5.1	39
4292	Resisting Skew-Accumulation for Time-Stepped Applications in the Cloud via Exploiting Parallelism. <i>IEEE Transactions on Cloud Computing</i> , 2015, 3, 54-65.	3.1	9
4293	Structural link prediction based on ant colony approach in social networks. <i>Physica A: Statistical Mechanics and Its Applications</i> , 2015, 419, 80-94.	1.2	65
4294	Community-based complex cloud data center. <i>Physica A: Statistical Mechanics and Its Applications</i> , 2015, 419, 356-372.	1.2	9
4295	Scalable real-time OLAP on cloud architectures. <i>Journal of Parallel and Distributed Computing</i> , 2015, 79-80, 31-41.	2.7	26
4296	LIBRA: Lightweight Data Skew Mitigation in MapReduce. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2015, 26, 2520-2533.	4.0	91
4297	Energy-Aware Scheduling of MapReduce Jobs for Big Data Applications. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2015, 26, 2720-2733.	4.0	106
4298	Heterogeneous programming with Single Operation Multiple Data. <i>Journal of Computer and System Sciences</i> , 2015, 81, 16-37.	0.9	4
4299	Parallel online sequential extreme learning machine based on MapReduce. <i>Neurocomputing</i> , 2015, 149, 224-232.	3.5	86
4300	LsPS: A Job Size-Based Scheduler for Efficient Task Assignments in Hadoop. <i>IEEE Transactions on Cloud Computing</i> , 2015, 3, 411-424.	3.1	25
4301	Parallel Detrended Fluctuation Analysis for Fast Event Detection on Massive PMU Data. <i>IEEE Transactions on Smart Grid</i> , 2015, 6, 360-368.	6.2	43
4302	Big Data computing and clouds: Trends and future directions. <i>Journal of Parallel and Distributed Computing</i> , 2015, 79-80, 3-15.	2.7	583

#	ARTICLE	IF	CITATIONS
4303	Estimating graph distance and centrality on shared nothing architectures. <i>Concurrency Computation Practice and Experience</i> , 2015, 27, 3587-3613.	1.4	1
4304	Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. <i>CAD Computer Aided Design</i> , 2015, 59, 1-14.	1.4	491
4305	An analysis of definition and placement of virtual machines for high performance applications on Clouds. <i>Concurrency Computation Practice and Experience</i> , 2015, 27, 1789-1814.	1.4	1
4306	An Incremental and Distributed Inference Method for Large-Scale Ontologies Based on MapReduce Paradigm. <i>IEEE Transactions on Cybernetics</i> , 2015, 45, 53-64.	6.2	37
4307	CoSMiC: A hierarchical cloudlet-based storage architecture for mobile clouds. <i>Simulation Modelling Practice and Theory</i> , 2015, 50, 3-19.	2.2	17
4308	A feasible MapReduce peer-to-peer framework for distributed computing applications. <i>Vietnam Journal of Computer Science</i> , 2015, 2, 57-66.	1.0	1
4309	An effective and economical architecture for semantic-based heterogeneous multimedia big data retrieval. <i>Journal of Systems and Software</i> , 2015, 102, 207-216.	3.3	34
4310	A Survey on Software-Defined Networking. <i>IEEE Communications Surveys and Tutorials</i> , 2015, 17, 27-51.	24.8	751
4311	Low-dimensional tracking of association structures in categorical data. <i>Statistics and Computing</i> , 2015, 25, 1009-1022.	0.8	4
4312	Cost-Effective Resource Provisioning for MapReduce in a Cloud. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2015, 26, 1265-1279.	4.0	82
4313	High frequency batch-oriented computations over large sliding time windows. <i>Future Generation Computer Systems</i> , 2015, 43-44, 1-11.	4.9	2
4314	Design and application of augmented reality query-answering system in mobile phone information navigation. <i>Expert Systems With Applications</i> , 2015, 42, 810-820.	4.4	32
4315	Hierarchical attribute reduction algorithms for big data using MapReduce. <i>Knowledge-Based Systems</i> , 2015, 73, 18-31.	4.0	94
4316	The rise of "big data" on cloud computing: Review and open research issues. <i>Information Systems</i> , 2015, 47, 98-115.	2.4	1,853
4317	Formal representation of the SS-DB benchmark and experimental evaluation in EXTASCID. <i>Distributed and Parallel Databases</i> , 2015, 33, 277-317.	1.0	15
4318	MR-ELM: a MapReduce-based framework for large-scale ELM training in big data era. <i>Neural Computing and Applications</i> , 2016, 27, 101-110.	3.2	29
4319	Resource Allocation in Cloud Computing Using the Uncertainty Principle of Game Theory. <i>IEEE Systems Journal</i> , 2016, 10, 637-648.	2.9	108
4320	Adaptive Task Scheduling Strategy Based on Dynamic Workload Adjustment for Heterogeneous Hadoop Clusters. <i>IEEE Systems Journal</i> , 2016, 10, 471-482.	2.9	32

#	ARTICLE	IF	CITATIONS
4321	An efficient query processing optimization based on ELM in the cloud. <i>Neural Computing and Applications</i> , 2016, 27, 35-44.	3.2	8
4322	Enabling collaborative MapReduce on the Cloud with a single-sign-on mechanism. <i>Computing (Vienna/New York)</i> , 2016, 98, 55-72.	3.2	13
4324	PESSCARA: An Example Infrastructure for Big Data Research. , 0, , .		0
4325	A New Multi-Dimensional Hyperbolic Structure for Cloud Service Indexing. <i>International Journal of Database Management Systems</i> , 2016, 8, 01-13.	0.2	0
4326	Climate Analytics as a Service. , 2016, , 187-219.		8
4327	An Intelligent Approval System for City Construction based on Cloud Computing and Big Data. <i>International Journal of Grid and High Performance Computing</i> , 2016, 8, 57-69.	0.7	16
4328	A common framework of partition-based clustering for large scale dataset using sampling and its MapReduce implementation. <i>Tehnicki Vjesnik</i> , 2016, 23, .	0.3	1
4329	Customer Satisfaction through Technological Integration. <i>International Journal of Technology and Educational Marketing</i> , 2016, 6, 49-78.	0.1	0
4330	Users in the urban sensing process. , 2016, , 45-95.		6
4331	Proper Global Shared Preference Detection Based on Golden Section and Genetic Algorithm for Affinity Propagation Clustering. <i>International Journal of Distributed Sensor Networks</i> , 2016, 12, 9807206.	1.3	1
4332	Models and Techniques for Cloud-Based Data Analysis. , 2016, , 45-76.		2
4333	Query and Analysis of Data on Electric Consumption Based on Hadoop. <i>International Journal of Database Theory and Application</i> , 2016, 9, 153-160.	0.2	0
4334	Query Optimization Techniques in Graph Databases. <i>International Journal of Database Management Systems</i> , 2016, 8, 01-14.	0.2	10
4335	Optimizing ETL by a Two-Level Data Staging Method. <i>International Journal of Data Warehousing and Mining</i> , 2016, 12, 32-50.	0.4	4
4336	Survey Big Data Analytics, Applications and Privacy Concerns. <i>Indian Journal of Science and Technology</i> , 2016, 9, .	0.5	4
4337	Hadoop in the Cloud to Analyze Climate Datasets. , 2016, , 245-276.		1
4338	Big Data Networking : Requirements, Architecture and Issues. <i>International Journal of Wireless and Mobile Networks</i> , 2016, 8, 35-49.	0.1	3
4339	A parallelized network traffic classification framework based on selective ensemble and diversity measures. <i>Journal of High Speed Networks</i> , 2016, 22, 35-42.	0.6	0

#	ARTICLE	IF	CITATIONS
4340	The ATLAS EventIndex: data flow and inclusion of other metadata. Journal of Physics: Conference Series, 2016, 762, 012028.	0.3	0
4341	A Fast Alignment-Free Approach for De Novo Detection of Protein Conserved Regions. PLoS ONE, 2016, 11, e0161338.	1.1	3
4342	Database Techniques for Big Data. , 2016, , 139-159.		0
4343	Real-Time Analytics. , 2016, , 39-61.		12
4344	Low-Frequency Investment with High-Frequency Measures: Is it Profitable?. SSRN Electronic Journal, 2016, , .	0.4	2
4345	Near real-time landslide monitoring with the smart soil particles. Japanese Geotechnical Society Special Publication, 2016, 2, 1031-1034.	0.2	6
4346	In-memory indexation of event streams. , 2016, , .		0
4347	Multi-threading the generation of Burrows-Wheeler Alignment. Genetics and Molecular Research, 2016, 15, .	0.3	3
4348	A Visualization Review of Cloud Computing Algorithms in the Last Decade. Sustainability, 2016, 8, 1008.	1.6	13
4349	A Survey of Big Data Analytics. International Journal of Information Sciences and Techniques, 2016, 6, 35-43.	0.3	14
4350	An Enhanced Frequent Pattern Growth Based on MapReduce for Mining Association Rules. International Journal of Data Mining & Knowledge Management Process, 2016, 6, 19-28.	0.1	4
4351	Stream processing in IoT: foundations, state-of-the-art, and future directions. , 2016, , 145-161.		8
4352	Large Scale Data Enabled Evolution of Spoken Language Research and Applications. Handbook of Statistics, 2016, , 301-340.	0.4	13
4353	Parallel Mining Small Patterns from Business Process Traces. International Journal of Software Science and Computational Intelligence, 2016, 8, 32-45.	1.8	3
4354	A Workload Assignment Strategy for Efficient ROLAP Data Cube Computation in Distributed Systems. International Journal of Data Warehousing and Mining, 2016, 12, 51-71.	0.4	1
4355	Recent Developments on Security and Reliability in Large-Scale Data Processing with MapReduce. International Journal of Data Warehousing and Mining, 2016, 12, 49-68.	0.4	8
4356	Discovering Patterns using Process Mining. International Journal of Rough Sets and Data Analysis, 2016, 3, 21-31.	1.0	5
4357	MapReduce: A Technical Review. Indian Journal of Science and Technology, 2016, 9, .	0.5	1

#	ARTICLE	IF	CITATIONS
4358	Multivariate Multiple Regression Models for a Big Data-Empowered SON Framework in Mobile Wireless Networks. <i>Mobile Information Systems</i> , 2016, 2016, 1-10.	0.4	0
4359	A Natural Language Interface Concordant with a Knowledge Base. <i>Computational Intelligence and Neuroscience</i> , 2016, 2016, 1-15.	1.1	5
4360	MHDFS: A Memory-Based Hadoop Framework for Large Data Storage. <i>Scientific Programming</i> , 2016, 2016, 1-12.	0.5	0
4361	Parallelizing Backpropagation Neural Network Using MapReduce and Cascading Model. <i>Computational Intelligence and Neuroscience</i> , 2016, 2016, 1-11.	1.1	20
4362	Optimized Virtual Machine Placement with Traffic-Aware Balancing in Data Center Networks. <i>Scientific Programming</i> , 2016, 2016, 1-10.	0.5	9
4363	Distributed Parallel Endmember Extraction of Hyperspectral Data Based on Spark. <i>Scientific Programming</i> , 2016, 2016, 1-9.	0.5	2
4364	Remote Blood Glucose Monitoring in mHealth Scenarios: A Review. <i>Sensors</i> , 2016, 16, 1983.	2.1	37
4365	GeoWeb Crawler: An Extensible and Scalable Web Crawling Framework for Discovering Geospatial Web Resources. <i>ISPRS International Journal of Geo-Information</i> , 2016, 5, 136.	1.4	13
4366	A Cloud Computing-Enabled Spatio-Temporal Cyber-Physical Information Infrastructure for Efficient Soil Moisture Monitoring. <i>ISPRS International Journal of Geo-Information</i> , 2016, 5, 81.	1.4	11
4367	Resource Management in Big Data Processing Systems. , 2016, , 161-188.		1
4368	Big data analytics and cloud computing for sustainable building energy efficiency. , 2016, , 397-412.		10
4369	Comparison of Acceleration Techniques for Selected Low-Level Bioinformatics Operations. <i>Frontiers in Genetics</i> , 2016, 7, 5.	1.1	4
4370	Preparing Laboratory and Real-World EEG Data for Large-Scale Analysis: A Containerized Approach. <i>Frontiers in Neuroinformatics</i> , 2016, 10, 7.	1.3	24
4371	An Effective and Efficient MapReduce Algorithm for Computing BFS-Based Traversals of Large-Scale RDF Graphs. <i>Algorithms</i> , 2016, 9, 7.	1.2	4
4372	Main Issues in Big Data Security. <i>Future Internet</i> , 2016, 8, 44.	2.4	41
4373	Automatic Scaling Hadoop in the Cloud for Efficient Process of Big Geospatial Data. <i>ISPRS International Journal of Geo-Information</i> , 2016, 5, 173.	1.4	37
4374	Exploring Multi-Scale Spatiotemporal Twitter User Mobility Patterns with a Visual-Analytics Approach. <i>ISPRS International Journal of Geo-Information</i> , 2016, 5, 187.	1.4	20
4375	An Effective NoSQL-Based Vector Map Tile Management Approach. <i>ISPRS International Journal of Geo-Information</i> , 2016, 5, 215.	1.4	14

#	ARTICLE	IF	CITATIONS
4376	Challenges, Solutions, and Quality Metrics of Personal Genome Assembly in Advancing Precision Medicine. <i>Pharmaceutics</i> , 2016, 8, 15.	2.0	13
4377	Estimation Accuracy on Execution Time of Run-Time Tasks in a Heterogeneous Distributed Environment. <i>Sensors</i> , 2016, 16, 1386.	2.1	12
4378	GMRCube: A GPGPU accelerated MapReduce DataCube construction model. , 2016, , .		0
4379	Incremental attribute reduction algorithm for big data using MapReduce. <i>Journal of Computational Methods in Sciences and Engineering</i> , 2016, 16, 641-652.	0.1	5
4380	Performance Evaluation for Distributed Join Based on MapReduce. , 2016, , .		2
4381	Transparently Resilient Task Parallelism for Chapel. , 2016, , .		1
4382	Distributed gene expression modelling for exploring variability in epigenetic function. <i>BMC Bioinformatics</i> , 2016, 17, 446.	1.2	0
4383	Context driven metadata representation for SaaS. , 2016, , .		0
4384	Value proposition discovery in big data enabled business model innovation. , 2016, , .		1
4385	Scalable Predictive Analysis in Critically Ill Patients Using a Visual Open Data Analysis Platform. <i>PLoS ONE</i> , 2016, 11, e0145791.	1.1	35
4386	PERFORMANCE ANALYSIS OF TASK REPLICATION IN LARGE-SCALE PARALLEL-DISTRIBUTED PROCESSING: AN EXTREME VALUE THEORY APPROACH. <i>Journal of the Operations Research Society of Japan</i> , 2016, 59, 174-194.	0.3	0
4387	Patient Similarity: Emerging Concepts in Systems and Precision Medicine. <i>Frontiers in Physiology</i> , 2016, 7, 561.	1.3	60
4388	Analyzing Big Data in Psychology: A Split/Analyze/Meta-Analyze Approach. <i>Frontiers in Psychology</i> , 2016, 7, 738.	1.1	41
4389	Big Data Analytics = Machine Learning + Cloud Computing. , 2016, , 3-38.		38
4390	Heavy-Traffic Limits for a Fork-Join Network in the Halfin-Whitt Regime. <i>Stochastic Systems</i> , 2016, 6, 519-600.	0.8	9
4393	Local, Domain-independent Heuristics for the FEIII Challenge. , 2016, , .		1
4394	Parallel evaluation of multi-semi-joins. <i>Proceedings of the VLDB Endowment</i> , 2016, 9, 732-743.	2.1	1
4395	Nswap2L. , 2016, , .		4

#	ARTICLE	IF	CITATIONS
4396	eScience and Big Data Workflows in Clouds. , 2016, , 431-455.		3
4397	ROSCC: An Efficient Remote Sensing Observation-Sharing Method Based on Cloud Computing for Soil Moisture Mapping in Precision Agriculture. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9, 5588-5598.	2.3	42
4398	Extending Science Gateway Frameworks to Support Big Data Applications in the Cloud. Journal of Grid Computing, 2016, 14, 589-601.	2.5	14
4399	Presto: Towards efficient online virtual network embedding in virtualized cloud data centers. Computer Networks, 2016, 106, 196-208.	3.2	91
4400	ENTRADA: A high-performance network traffic data streaming warehouse. , 2016, , .		38
4401	MEMoMR: Accelerate MapReduce via reuse of intermediate results. Concurrency Computation Practice and Experience, 2016, 28, 3814-3829.	1.4	3
4402	Architecture for Internet of Things Analytical Ecosystem. Advances in Intelligent Systems and Computing, 2016, , 385-393.	0.5	0
4403	Data science: Accelerating innovation and discovery in chemical engineering. AIChE Journal, 2016, 62, 1402-1416.	1.8	63
4404	Fast Playback Framework for Analysis of Ground-Based Doppler Radar Observations Using MapReduce Technology. Journal of Atmospheric and Oceanic Technology, 2016, 33, 621-634.	0.5	5
4405	A Naturalâ€Languageâ€Based Approach to Intelligent Data Retrieval and Representation for Cloud BIM. Computer-Aided Civil and Infrastructure Engineering, 2016, 31, 18-33.	6.3	113
4406	Factory: Master Node High-Availability for Big Data Applications and Beyond. Lecture Notes in Computer Science, 2016, , 379-389.	1.0	3
4407	How Can SMEs Benefit from Big Data? Challenges and a Path Forward. Quality and Reliability Engineering International, 2016, 32, 2151-2164.	1.4	134
4408	Big data and hydroinformatics. Journal of Hydroinformatics, 2016, 18, 599-614.	1.1	43
4409	Edge Computing: Vision and Challenges. IEEE Internet of Things Journal, 2016, 3, 637-646.	5.5	4,860
4410	OpticV: An energy-efficient datacenter network architecture by MEMS-based all-optical bypassing. , 2016, , .		0
4411	ROUTE: runâ€time robust reducer workload estimation for MapReduce. International Journal of Network Management, 2016, 26, 224-244.	1.4	4
4412	CLORIFI: software vulnerability discovery using code clone verification. Concurrency Computation Practice and Experience, 2016, 28, 1900-1917.	1.4	22
4413	A synthetic data generator for online social network graphs. Social Network Analysis and Mining, 2016, 6, 1.	1.9	13

#	ARTICLE	IF	CITATIONS
4414	Analytics, challenges and applications in big data environment: a survey. Journal of Management Analytics, 2016, 3, 206-239.	1.6	52
4415	The era of Big Spatial Data. , 2016, , .		9
4416	Towards a distributed multi-tier file system for cluster computing. , 2016, , .		0
4417	Flame-MR: An event-driven architecture for MapReduce applications. Future Generation Computer Systems, 2016, 65, 46-56.	4.9	14
4418	Towards felicitous decision making: An overview on challenges and trends of Big Data. Information Sciences, 2016, 367-368, 747-765.	4.0	190
4419	Voting-based instance selection from large data sets with MapReduce and random weight networks. Information Sciences, 2016, 367-368, 1066-1077.	4.0	40
4420	Large-scale ontology matching: A review of the literature. , 2016, , .		13
4421	Optimization for Statistical Machine Translation: A Survey. Computational Linguistics, 2016, 42, 1-54.	2.5	22
4422	An Introduction to Big Data. , 2016, , 537-550.		3
4423	Efficient similarity join based on Earth Mover's Distance using MapReduce. , 2016, , .		1
4424	Accelerated join evaluation in Semantic Web databases by using FPGAs. Concurrency Computation Practice and Experience, 2016, 28, 2031-2051.	1.4	6
4425	Performance evaluation of job schedulers on Hadoop YARN. Concurrency Computation Practice and Experience, 2016, 28, 2711-2728.	1.4	17
4427	The on-line viewing system of the distributed gigapixel image. , 2016, , .		1
4428	MR-GABiT: Map reduce based genetic algorithm for biclustering time series data. , 2016, , .		2
4429	Factory: Non-stop batch jobs without checkpointing. , 2016, , .		5
4430	Scaling applications on cloud using GPGPU- trends and techniques. , 2016, , .		1
4431	Providing efficient, scalable and privacy preserved verification mechanism in remote attestation. , 2016, , .		7
4432	Improved Data-Aware Task Dispatching for Batch Queuing Systems. , 2016, , .		1

#	ARTICLE	IF	CITATIONS
4433	Replichard. , 2016, , .		4
4434	AT-GIS. , 2016, , .		2
4435	Big-DAMA. , 2016, , .		18
4436	Big data emerging issues: Hadoop security and privacy. , 2016, , .		11
4437	MapReduce Based Framework for Searchable Encryption. , 2016, , .		2
4438	A Novel Performance Evaluation and Optimization Model for Big Data System. , 2016, , .		2
4439	A MapReduce based approach for classification. , 2016, , .		2
4440	Distributed data augmented support vector machine on Spark. , 2016, , .		8
4441	Adaptive Anomaly Detection in Cloud Using Robust and Scalable Principal Component Analysis. , 2016, , .		7
4442	Sentiment analysis on Twitter Data-set using Naive Bayes algorithm. , 2016, , .		67
4443	Improving the performance of heterogeneous Hadoop cluster. , 2016, , .		2
4444	Research on MapReduce-Based Fuzzy Associative Classifier for Big Probabilistic Numerical Data. , 2016, , .		1
4445	Development of a Scalable Method for Creating Food Groups Using the NHANES Dataset and MapReduce. , 2016, , .		0
4446	Parallel subgraph matching on massive graphs. , 2016, , .		0
4447	Locality premised reducer scheduling in Hadoop. , 2016, , .		2
4448	Adaptive Block and Batch Sizing for Batched Stream Processing System. , 2016, , .		19
4449	Image retrieval and classification on deep convolutional SparkNet. , 2016, , .		8
4450	Noise Removal and Structured Data Detection to improve search for personality features. , 2016, , .		2

#	ARTICLE	IF	CITATIONS
4451	Unifying HDFS and GPFS. , 2016, , .		2
4452	Limited random walk algorithm for big graph data clustering. Journal of Big Data, 2016, 3, .	6.9	17
4453	Towards Memory-Optimized Data Shuffling Patterns for Big Data Analytics. , 2016, , .		7
4454	Quantifying the difference in resource demand among classic and modern NoC workloads. , 2016, , .		4
4455	Current Flow Betweenness Centrality with Apache Spark. Lecture Notes in Computer Science, 2016, , 270-278.	1.0	1
4456	Scaling DBSCAN-like Algorithms for Event Detection Systems in Twitter. Lecture Notes in Computer Science, 2016, , 356-373.	1.0	4
4457	Equation-based Multicast Congestion control in Data Center networks. , 2016, , .		2
4458	Embedding Concurrent Generators. , 2016, , .		0
4459	Signal Processing Techniques Restructure The Big Data Era. , 2016, , .		2
4460	Wind farm flow modeling using an input-output reduced-order model. , 2016, , .		25
4461	A distributed approach to weighted frequent Subgraph mining. , 2016, , .		7
4462	URBAN-NET: A network-based infrastructure monitoring and analysis system for emergency management and public safety. , 2016, , .		9
4463	In-cache MapReduce: Leverage Tiling to Boost Temporal Locality-Sensitive MapReduce Computations. , 2016, , .		1
4464	Intelligent agents based improved map-reduce method for multiple databases. , 2016, , .		0
4465	Implicit effect of decoding time on fault tolerance in erasure coded Cloud Storage systems. , 2016, , .		2
4466	Identifying performance bottlenecks in Hive: Use of processor counters. , 2016, , .		3
4467	Testing Interestingness Measures in Practice: A Large-Scale Analysis of Buying Patterns. , 2016, , .		11
4468	Processing of big educational data in the cloud using Apache Hadoop. , 2016, , .		8

#	ARTICLE	IF	CITATIONS
4469	Optimization and realization of parallel frequent item set mining algorithm. , 2016, , .		0
4470	Private databases on the cloud: Models, issues and research perspectives. , 2016, , .		4
4471	Parallel Document Clustering using Iterative MapReduce. , 2016, , .		1
4472	A scalable model for efficient information diffusion in large real world networks. , 2016, , .		3
4473	AQUIR - a system to generate quantitative and customized vision measurement template. , 2016, , .		0
4474	Partition Aware Connected Component Computation in Distributed Systems. , 2016, , .		6
4475	An implementation of heterogeneous architecture based MapReduce in the clouds. , 2016, , .		0
4476	LRDM: Local Record-Driving Mechanism for Big Data Privacy Preservation in Social Networks. , 2016, , .		6
4477	Optimization of multi kernel parallel support vector machine based on Hadoop. , 2016, , .		2
4478	Distributed data deduplication. Proceedings of the VLDB Endowment, 2016, 9, 864-875.	2.1	55
4479	Remote health care cyber-physical system: quality of service (QoS) challenges and opportunities. IET Cyber-Physical Systems: Theory and Applications, 2016, 1, 40-48.	1.9	58
4480	MOHA: Many-task computing meets the big data platform. , 2016, , .		6
4481	Scalable attack propagation model and algorithms for honeypot systems. , 2016, , .		4
4482	An End-to-End Privacy Preserving Design of a Map-Reduce Framework. , 2016, , .		2
4483	Subgroup Discovery on Big Data: Exhaustive Methodologies Using Map-Reduce. , 2016, , .		5
4484	Adapting to data sparsity for efficient parallel PARAFAC tensor decomposition in Hadoop. , 2016, , .		2
4485	GillesPy: A Python Package for Stochastic Model Building and Simulation. IEEE Life Sciences Letters, 2016, 2, 35-38.	1.2	38
4486	Semantic body parts segmentation for quadrupedal animals. , 2016, , .		20

#	ARTICLE	IF	CITATIONS
4487	QryGraph: A graphical tool for Big Data analytics. , 2016, , .		3
4488	PrefRank: Quantification and Aggregation of Subjective User Preferences. , 2016, , .		0
4489	BOLAS+: Scalable Lightweight Locality-Aware Scheduling for Hadoop. , 2016, , .		3
4490	An optimized frequent pattern mining algorithm with multiple minimum supports. , 2016, , .		2
4491	Tapas: An Implicitly Parallel Programming Framework for Hierarchical N-Body Algorithms. , 2016, , .		3
4492	DCSPARK: Virtualizing spark using Docker containers. , 2016, , .		2
4493	Improvements to Online Distributed Monitoring Systems. , 2016, , .		0
4494	Towards Application-Aware In-Network Bandwidth Management in Data Centers. , 2016, , .		1
4495	Inverted Voronoi-Based kNN Query Processing with MapReduce. , 2016, , .		5
4496	S-Storm: A Slot-Aware Scheduling Strategy for Even Scheduler in Storm. , 2016, , .		8
4497	DScheduler: Dynamic Network Scheduling Method for MapReduce in Distributed Controllers. , 2016, , .		2
4498	An Efficient Framework for Incremental Graph Computation. , 2016, , .		0
4499	Distributed live data search architecture for resource discovery on Internet of Things. , 2016, , .		6
4500	A de novo genome assembler based on MapReduce and bi-directed de Bruijn graph. , 2016, , .		0
4501	SciSpark: Highly interactive in-memory science data analytics. , 2016, , .		5
4502	Optimizing VM provisioning of map reduce tasks on public clouds. , 2016, , .		0
4503	Reducing Access Latency in Virtual Machine Assignments. , 2016, , .		1
4504	Graph Mining for Complex Data Analytics. , 2016, , .		6

#	ARTICLE	IF	CITATIONS
4505	Enabling ECN over Generic Packet Scheduling. , 2016, , .		40
4506	Cloud-based machine learning for predictive analytics: Tool wear prediction in milling. , 2016, , .		29
4507	Visually programming dataflows for distributed data analytics. , 2016, , .		4
4508	Big data availability: Selective partial checkpointing for in-memory database queries. , 2016, , .		0
4509	A Parallel K-Medoids Algorithm for Clustering based on MapReduce. , 2016, , .		10
4510	PredicT-ML: a tool for automating machine learning model building with big clinical data. Health Information Science and Systems, 2016, 4, 5.	3.4	33
4511	An experimental study of big spatial data systems. , 2016, , .		10
4512	The technical hashtag in Twitter data: A hadoop experience. , 2016, , .		4
4513	Steal-A-GC: Framework to Trigger GC during Idle Periods in Distributed Systems. , 2016, , .		0
4514	Smart online vehicle tracking system for security applications. , 2016, , .		1
4515	Datalography: Scaling datalog graph analytics on graph processing systems. , 2016, , .		10
4516	Parallel computation of k-nearest neighbor joins using MapReduce. , 2016, , .		11
4517	Data quality: Experiences and lessons from operationalizing big data. , 2016, , .		6
4518	Event segmentation using MapReduce based big data clustering. , 2016, , .		4
4519	Real-time processing of proteomics data: The internet of things and the connected laboratory. , 2016, , .		2
4520	Hobbits: Hadoop and Hive based Internet traffic analysis. , 2016, , .		9
4521	Performance evaluation of big data frameworks for large-scale data analytics. , 2016, , .		37
4522	Parallel clustering method for non-disjoint partitioning of large-scale data based on spark framework. , 2016, , .		4

#	ARTICLE	IF	CITATIONS
4523	Online Credit Card Fraud Detection: A Hybrid Framework with Big Data Technologies. , 2016, , .		20
4524	Dynamic Contextual Multi Arm Bandits in Display Advertisement. , 2016, , .		6
4525	Idempotent Task Cache System for Handling Intermediate Data Skew in MapReduce on Cloud Computing. , 2016, , .		1
4526	Resource allocation for data-parallel computing in networks with data locality. , 2016, , .		1
4527	A hybrid cloud platform for diversified real-time demand in transformation equipment monitoring. , 2016, , .		0
4528	Subgroup discovery on big data: Pruning the search space on exhaustive search algorithms. , 2016, , .		2
4529	Thrill: High-performance algorithmic distributed batch data processing with C++. , 2016, , .		23
4530	Improved Canopy-FCM algorithm based on MapReduce. , 2016, , .		1
4531	Some pairs problems. , 2016, , .		0
4532	A preliminary review of influential works in data-driven discovery. SpringerPlus, 2016, 5, 1266.	1.2	6
4533	Scalable and Distributed Sea Port Operational Areas Estimation from AIS Data. , 2016, , .		8
4534	Research the Data Analysis and Processing between MapReduce and Spark. , 2016, , .		1
4535	Identification of critical parameters for MapReduce energy efficiency using statistical Design of Experiments. , 2016, , .		6
4536	GraphFlow: Workflow-based big graph processing. , 2016, , .		5
4537	Security evaluation of RC4 using big data analytics. , 2016, , .		0
4538	Recommendation system for financial analytics. , 2016, , .		3
4539	Feedback analysis using big data tools. , 2016, , .		9
4540	Genetic-fuzzy mining with MapReduce. , 2016, , .		1

#	ARTICLE	IF	CITATIONS
4541	SpongeNet+: A Two-Layer Bandwidth Allocation Framework for Cloud Datacenter. , 2016, , .		0
4542	Accuracy Improvement for Backup Tasks in Hadoop Speculative Algorithm. , 2016, , .		1
4543	Research on map-reduce distributed computing model based on .NET platform. , 2016, , .		0
4544	Block-parallel data analysis with DIY2. , 2016, , .		25
4545	Analysis of OpenCL Work-Group Reduce for Intel GPUs. , 2016, , .		1
4546	Energy Aware Resource Management for MapReduce Jobs with Service Level Agreements in Cloud Data Centers. , 2016, , .		6
4547	DFA minimization in map-reduce. , 2016, , .		5
4548	Credential and security issues of cloud service models. , 2016, , .		9
4549	Social manufacturing realizing personalization production: A state-of-the-art review. , 2016, , .		10
4550	Spreading fuzzy random forests with MapReduce. , 2016, , .		3
4551	Approximate Iterative Method for Fast Data Analysis in Internet of Things Environment. , 2016, , .		0
4552	Map reduce programming model: Construction of inverted index for automated document clustering. , 2016, , .		0
4553	Outer bounds on the storage-repair bandwidth trade-off of exact-repair regenerating codes. International Journal of Information and Coding Theory, 2016, 3, 255.	0.3	11
4554	Efficient analytical architecture for sensor networks using Hadoop. , 2016, , .		0
4555	Semantic inference method using ontologies. , 2016, , .		1
4556	A novel technique using slot configuration on Hadoop clusters. , 2016, , .		0
4557	SARVAVID. , 2016, , .		14
4558	Cultural Ant Colony Optimization on GPUs for Travelling Salesman Problem. Lecture Notes in Computer Science, 2016, , 317-329.	1.0	2

#	ARTICLE	IF	CITATIONS
4559	Multi-objective Hadoop Configuration Optimization Using Steady-State NSGA-II. , 2016, , .		1
4560	A Dual Perturbation Approach for Differential Private ADMM-Based Distributed Empirical Risk Minimization. , 2016, , .		9
4561	Programming Scalable Cloud Services with AEON. , 2016, , .		7
4562	Cost function based event triggered Model Predictive Controllers application to Big Data Cloud services. , 2016, , .		6
4563	Parallel Discrete-Event Simulation on Data Processing Engines. , 2016, , .		0
4564	Evaluation of Datacenter Network Topology Influence on Hadoop MapReduce Performance. , 2016, , .		5
4565	Performance optimization of applications based on non-relational databases. , 2016, , .		4
4566	A Comparative Survey of the HPC and Big Data Paradigms: Analysis and Experiments. , 2016, , .		19
4567	Smart-MLlib: A High-Performance Machine-Learning Library. , 2016, , .		5
4568	Improving Multipath TCP for Latency Sensitive Flows in the Cloud. , 2016, , .		5
4569	Performance Evaluation of Edge Cloud Computing System for Big Data Applications. , 2016, , .		15
4570	Dynamic Fault Tolerant Scheduling Mechanism for Real Time Tasks in cloud computing. , 2016, , .		11
4571	Custody: Towards Data-Aware Resource Sharing in Cloud-Based Big Data Processing. , 2016, , .		2
4572	Record Skipping in Parallel Data Processing Systems. , 2016, , .		0
4573	A High-Performance Retrieval Method of Mass Data Oriented to Cloud Computing. , 2016, , .		0
4574	Parallelization of Vertical Search Engine using Hadoop and MapReduce. , 2016, , .		3
4575	A MapReduce-Based Approach for Prefix-Based Labeling of Large XML Data. Lecture Notes in Computer Science, 2016, , 83-98.	1.0	1
4576	An Accelerated MapReduce-Based K-prototypes for Big Data. Lecture Notes in Computer Science, 2016, , 13-25.	1.0	2

#	ARTICLE	IF	CITATIONS
4577	Probabilistic Network-Aware Task Placement for MapReduce Scheduling. , 2016, , .		22
4578	A unified software approach to specify pipeline and spatial parallelism in FPGA hardware. , 2016, , .		5
4579	Scheduler technologies in support of high performance data analysis. , 2016, , .		14
4580	AEGEUS. , 2016, , .		3
4581	MRUniNovo: an efficient tool for <i>de novo</i> peptide sequencing utilizing the hadoop distributed computing framework. <i>Bioinformatics</i> , 2017, 33, 944-946.	1.8	15
4582	A FIRM approach for software-defined service composition. , 2016, , .		2
4583	Scalable Processing in 5G Cloud-RAN Networks Using MapReduce Framework. , 2016, , .		3
4584	KOHA: Building a Kafka-Based Distributed Queue System on the Fly in a Hadoop Cluster. , 2016, , .		12
4586	Improving Load Balance for Data-Intensive Computing on Cloud Platforms. , 2016, , .		9
4587	Efficient Graph Clustering Algorithm and its use in Prediction of Students Performance. , 2016, , .		3
4588	Deadline constrained Cost Effective Workflow scheduler for Hadoop clusters in cloud datacenter. , 2016, , .		4
4589	Composite-Path Switching. , 2016, , .		4
4590	Firework: Big Data Sharing and Processing in Collaborative Edge Environment. , 2016, , .		53
4591	Graph database partitioning: A study. , 2016, , .		1
4592	Building Blocks of Mayan: Componentizing the eScience Workflows through Software-Defined Service Composition. , 2016, , .		5
4593	On parallelizing analysis of power systems in cloud environment. , 2016, , .		0
4594	Maximizing container-based network isolation in parallel computing clusters. , 2016, , .		3
4595	A Data Driven Approach for the Science of Cyber Security: Challenges and Directions. , 2016, , .		23

#	ARTICLE	IF	CITATIONS
4596	Efficient Intermediate Data Placement in Federated Cloud Data Centers Storage. Lecture Notes in Computer Science, 2016, , 1-15.	1.0	1
4597	A unified memory network architecture for in-memory computing in commodity servers. , 2016, , .		12
4598	Locality based data partitioning in Map reduce. , 2016, , .		2
4599	Infrastructure-Aware Functional Testing of MapReduce Programs. , 2016, , .		4
4600	ViTeNA: An SDN-based virtual network embedding algorithm for multi-tenant data centers. , 2016, , .		3
4601	Empirical analysis of asymptotic ensemble learning for big data. , 2016, , .		6
4602	Automating Analytics: How to Learn Metadata Such That Our Buildings Can Learn from Us. , 2016, , .		4
4604	A parallel algorithm of optimal power flow on Hadoop platform. , 2016, , .		1
4605	A Unified View of Data-Intensive Flows in Business Intelligence Systems: A Survey. Lecture Notes in Computer Science, 2016, , 66-107.	1.0	4
4606	VOLAP: A Scalable Distributed System for Real-Time OLAP with High Velocity Data. , 2016, , .		6
4607	Bulk I/O Storage Management for Big Data Applications. , 2016, , .		4
4608	Optimization of Chi-FRBCS-BigDataCS Fuzzy Rule-Based Classification System. , 2016, , .		1
4609	Towards adaptable and tunable cloud-based map-matching strategy for GPS trajectories. Frontiers of Information Technology and Electronic Engineering, 2016, 17, 1305-1319.	1.5	3
4610	On Distributed Deep Network for Processing Large-Scale Sets of Complex Data. , 2016, , .		2
4611	Sampling Method in Traffic Logs Analyzing. , 2016, , .		3
4612	V-Hadoop: Virtualized Hadoop using containers. , 2016, , .		3
4613	GI bleeding detection in wireless capsule endoscopy images based on pattern recognition and a MapReduce framework. , 2016, , .		11
4614	Bridging the Semantic Gaps of GPU Acceleration for Scale-out CNN-based Big Data Processing. , 2016, , .		17

#	ARTICLE	IF	CITATIONS
4615	Scalable metagenomics alignment research tool (SMART): a scalable, rapid, and complete search heuristic for the classification of metagenomic sequences from complex sequence populations. BMC Bioinformatics, 2016, 17, 292.	1.2	25
4616	FPGA-Based Accelerator for Deep Convolutional Neural Networks for the SPARK Environment. , 2016, , .		7
4617	Middleware Support for Seamless Integration of Domain Specific Cores in Cloud Datacenters. , 2016, , .		0
4618	Considerations on Geospatial Big Data. IOP Conference Series: Earth and Environmental Science, 2016, 46, 012058.	0.2	8
4619	Analyzing Healthcare Big Data With Prediction for Future Health Condition. IEEE Access, 2016, 4, 9786-9799.	2.6	110
4620	Improving network traffic in MapReduce for big data applications. , 2016, , .		2
4621	A First Approach in Evolutionary Fuzzy Systems based on the lateral tuning of the linguistic labels for Big Data classification. , 2016, , .		10
4622	Optimizing VM Provisioning of MapReduce Tasks on Public Cloud. , 2016, , .		3
4623	Non-Volatile In-Memory Computing by Spintronics. Synthesis Lectures on Emerging Engineering Technologies, 2016, 2, 1-161.	0.2	3
4624	Multi-agent frameworks in the service of the online product reputation extraction. , 2016, , .		1
4625	Reducing repair-bandwidth using codes based on factor graphs. , 2016, , .		4
4626	Optimization of spark storage solutions. , 2016, , .		1
4627	An Asynchronous Distributed ADMM Algorithm and Efficient Communication Model. , 2016, , .		10
4628	Implementation of MapReduce over structured peer-to-peer overlay of underutilized resources. , 2016, , .		0
4629	Smart grid communication network traffic anomaly detection based on entropy analysis. , 2016, , .		0
4630	Goldfish: In-Memory Massive Parallel Processing SQL Engine Based on Columnar Store. , 2016, , .		1
4631	Hierarchical Consistent Hashing for Heterogeneous Object-Based Storage. , 2016, , .		11
4632	The impact of Moore's Law and loss of Dennard scaling: Are DSP SoCs an energy efficient alternative to x86 SoCs?. Journal of Physics: Conference Series, 2016, 762, 012022.	0.3	10

#	ARTICLE	IF	CITATIONS
4633	A pattern collection for knowledge processing system architecture. , 2016, , .		3
4634	Flow Driven Energy-Aware Routing Algorithm in Data Center Network. , 2016, , .		0
4635	Different Scheduling Options in YARN. , 2016, , .		0
4636	Performance evaluation of Twitter datasets onÂSQL and NoSQL DBMS. Web Intelligence, 2016, 14, 275-286.	0.1	0
4637	Tasklets: "Better than Best-Effort" Computing. , 2016, , .		32
4638	BAShuffler. , 2016, , .		7
4639	Parallel evaluation of a DSP algorithm using julia. , 2016, , .		1
4640	Optimizing coflow completion times with utility max-min fairness. , 2016, , .		42
4641	Parallel Algorithms for Summing Floating-Point Numbers. , 2016, , .		0
4642	The research of mapreduce load balancing based on multiple partition algorithm. , 2016, , .		5
4643	Parallelizing label propagation for overlapping community detection. , 2016, , .		1
4644	On exploiting data locality for iterative mapreduce applications in hybrid clouds. , 2016, , .		7
4645	Application of Bloom Filter for Duplicate URL Detection in a Web Crawler. , 2016, , .		3
4646	Hadoop-HBase for finding association rules using Apriori MapReduce algorithm. , 2016, , .		4
4647	Modeling and Verifying HDFS Using CSP. , 2016, , .		4
4648	Experimental evaluation of memory configurations of Hadoop in Docker environments. , 2016, , .		6
4649	APS: Dynamic platform selector for big data analysis. , 2016, , .		0
4650	Thinking about the relationship between IT and food safety. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
4651	Unified and lightweight tasks and conduits: A high level parallel programming framework. , 2016, , .		1
4652	Utilizing semantic big data for realizing a national-scale infrastructure vulnerability analysis system. , 2016, , .		5
4653	A data fusion technique for smart home energy management and analysis. , 2016, , .		5
4654	Dynamic Virtual Network Traffic Engineering with Energy Efficiency in Multi-location Data Center Networks. , 2016, , .		10
4655	Observations on factors affecting performance of MapReduce based Apriori on Hadoop cluster. , 2016, , .		3
4656	Little Big Data: Shaping Minds for the Cloud. , 2016, , .		1
4657	Visual Reasoning Indexing and Retrieval Using In-Memory Computing. International Journal of Semantic Computing, 2016, 10, 299-322.	0.4	4
4658	The role of high performance, grid and cloud computing in high-throughput sequencing. , 2016, , .		1
4659	Accelerating BLAST Computation on an FPGA-enhanced PC Cluster. , 2016, , .		2
4660	Data intensive, computing and network aware (DCN) cloud VMs scheduling algorithm. , 2016, , .		1
4661	Defining software architectures for big data enabled operator support systems. , 2016, , .		12
4662	Implementation of page rank algorithm in Hadoop MapReduce framework. , 2016, , .		3
4663	eSplash: Efficient speculation in large scale heterogeneous computing systems. , 2016, , .		14
4664	Network Scheduling Aware Task Placement in Datacenters. , 2016, , .		18
4665	Transport-Layer Control to Increase Throughput in Bufferless Optical Packet-Switching Networks. Journal of Optical Communications and Networking, 2016, 8, 947.	3.3	12
4666	Finding group structures in "Big Data" in healthcare research using mixture models. , 2016, , .		0
4667	Enhanced Gender Identification through Social Media Analytics on the Cloud. , 2016, , .		2
4668	Analyzing users in parallel computing: A user-oriented study. , 2016, , .		6

#	ARTICLE	IF	CITATIONS
4669	An Adaptively Disperse Centroids K-Means Algorithm Based on MapReduce Model. , 2016, , .		3
4670	Designing Virtualization-Aware and Automatic Topology Detection Schemes for Accelerating Hadoop on SR-IOV-Enabled Clouds. , 2016, , .		1
4671	SLA-Aware Scheduling of Map-Reduce Applications on Public Clouds. , 2016, , .		5
4672	A Combined Analytical Modeling Machine Learning Approach for Performance Prediction of MapReduce Jobs in Cloud Environment. , 2016, , .		12
4673	Exploring Controlled RDF Distribution. , 2016, , .		6
4674	A distributed framework for event log analysis using MapReduce. , 2016, , .		8
4675	A performance-satisfied and affection-aware MapReduce allocation scheme for intelligent information applications. , 2016, , .		0
4676	PNKASR " A Positive Negative Keyword-Aware Service Recommendation methodology for big data applications. , 2016, , .		0
4677	Density-Based Algorithm in MapReduce. , 2016, , .		0
4678	Real time streaming data storage and processing using storm and analytics with Hive. , 2016, , .		7
4679	RedLib: Nestable Reductions for Collections in Java. , 2016, , .		1
4680	DRASH: A Data Replication-Aware Scheduler in Geo-Distributed Data Centers. , 2016, , .		10
4681	A Configurable Energy Aware Resource Management Technique for Optimization of Performance and Energy Consumption on Clouds. , 2016, , .		3
4682	Agricultural Knowledge Extraction from Text Sources Using a Distributed MapReduce Cluster. , 2016, , .		2
4683	Designing Parallel Data Processing for Large-Scale Sensor Orchestration. , 2016, , .		0
4684	CRED: Cloud Right-Sizing to Meet Execution Deadlines and Data Locality. , 2016, , .		5
4685	Gnafuy: A Framework for Ubiquitous Mobile Computation. , 2016, , .		0
4686	A High Performance Computing Framework for Data Mining. , 2016, , .		5

#	ARTICLE	IF	CITATIONS
4687	High Frequency Trading with Complex Event Processing. , 2016, , .		1
4688	Using CouchDB to Compute Temporal Aggregates. , 2016, , .		2
4689	Workload interleaving with performance guarantees in data centers. , 2016, , .		4
4690	Big data analytics systems. , 2016, , 137-160.		2
4691	MultiStack: Multi-Cloud Big Data Research Framework/Platform. , 2016, , .		1
4692	Early-Stage Engagement: Applying Big Data Analytics on Collaborative Learning Environment for Measuring Learners' Engagement Rate. , 2016, , .		4
4693	Tracking Language Mobility in the Twitter Landscape. , 2016, , .		8
4694	BRPS: A Big Data Placement Strategy for Data Intensive Applications. , 2016, , .		7
4695	A Big Data Analytics Framework for Supporting Multidimensional Mining over Big Healthcare Data. , 2016, , .		15
4696	Towards Scalable Subgraph Pattern Matching over Big Graphs on MapReduce. , 2016, , .		1
4697	An Implement of Actor Model - uJavaActor's Improvement. , 2016, , .		0
4698	PY-PITS: A Scalable Python Runtime System for the Computation of Partially Idempotent Tasks. , 2016, , .		5
4699	Low Latency and Resource-Aware Program Composition for Large-Scale Data Analysis. , 2016, , .		3
4700	2PTS: A Two-Phase Task Scheduling Algorithm for MapReduce. IEICE Transactions on Information and Systems, 2016, E99.D, 2377-2380.	0.4	0
4701	Insight of big data analytics in healthcare industry. , 2016, , .		11
4702	The Improvement of a Data Cached Strategy Based on Eventual Consistency Theory. , 2016, , .		0
4703	A Platform for Overcrowding Detection in Indoor Events Using Scalable Technologies. , 2016, , .		0
4704	ETSCH: Partition-Centric Graph Processing. , 2016, , .		1

#	ARTICLE	IF	CITATIONS
4705	Progress Estimation in Parallel Data Processing Systems. , 2016, , .		0
4706	KVBTREE: A Key/Value Based Storage Structure for Large-Scale Electric Power Data. , 2016, , .		1
4707	Location-Aware Data Block Allocation Strategy for HDFS-Based Applications in the Cloud. , 2016, , .		2
4708	CloudMeasure: A Platform for Performance Analysis of Cloud Computing Systems. , 2016, , .		1
4709	CPU Frequency Tuning to Improve Energy Efficiency of MapReduce Systems. , 2016, , .		4
4710	A system of systems framework for sustainable fashion supply chain management in the big data era. , 2016, , .		11
4711	Big data analytics concepts and management techniques. , 2016, , .		4
4712	Efficient online coflow routing and scheduling. , 2016, , .		58
4713	Distributed video transcoding on a heterogeneous computing platform. , 2016, , .		8
4714	Spatial data analysis with ArcGIS and MapReduce. , 2016, , .		8
4715	Hadoop in Flight: Migrating Live MapReduce Jobs for Power-Shifting Data Centers. , 2016, , .		2
4716	Architecture of an Extendable and Cloud-Ready Knowledge Management and Processing Framework for the Agricultural Domain. , 2016, , .		3
4717	MapReduce model for k-medoid clustering. , 2016, , .		1
4718	An efficient approach for rule matching in production system based on multi-agent. , 2016, , .		0
4719	A Versatile Event-Driven Data Model in HBase Database for Multi-source Data of Power Grid. , 2016, , .		6
4720	Synergistic Analysis of Evolving Graphs. Transactions on Architecture and Code Optimization, 2016, 13, 1-27.	1.6	36
4721	Profit-oriented task scheduling algorithm in Hadoop cluster. Eurasip Journal on Embedded Systems, 2016, 2016, .	1.2	4
4722	A hybrid integrated architecture for energy consumption prediction. Future Generation Computer Systems, 2016, 63, 131-147.	4.9	20

#	ARTICLE	IF	CITATIONS
4723	TEMPORARY REMOVAL: Cloud-based emerging services systems. International Journal of Information Management, 2016, , .	10.5	12
4724	ECJ+HADOOP: An Easy Way to Deploy Massive Runs of Evolutionary Algorithms. Lecture Notes in Computer Science, 2016, , 91-106.	1.0	10
4725	HPCâ€œGAP: engineering a 21stâ€œcentury highâ€œperformance computer algebra system. Concurrency Computation Practice and Experience, 2016, 28, 3606-3636.	1.4	3
4726	J2M: a Java to MapReduce translator for cloud computing. Journal of Supercomputing, 2016, 72, 1928-1945.	2.4	7
4727	A cloud-based system framework for performing online viewing, storage, and analysis on big data of massive BIMs. Automation in Construction, 2016, 71, 34-48.	4.8	93
4728	Inferring Social Strength from Spatiotemporal Data. ACM Transactions on Database Systems, 2016, 41, 1-47.	1.5	17
4729	MapReduce-based fast fuzzy c-means algorithm for large-scale underwater image segmentation. Future Generation Computer Systems, 2016, 65, 90-101.	4.9	42
4730	Bayes and big data: the consensus Monte Carlo algorithm. International Journal of Management Science and Engineering Management, 2016, 11, 78-88.	2.6	148
4731	Compressed Nonnegative Matrix Factorization Is Fast and Accurate. IEEE Transactions on Signal Processing, 2016, 64, 2269-2283.	3.2	28
4732	Large-scale seismic waveform quality metric calculation using Hadoop. Computers and Geosciences, 2016, 94, 18-30.	2.0	16
4733	Setting Up a Big Data Project: Challenges, Opportunities, Technologies and Optimization. Studies in Big Data, 2016, , 17-47.	0.8	10
4734	Leveraging MapReduce to efficiently extract associations between biomedical concepts from large text data. Microprocessors and Microsystems, 2016, 46, 202-210.	1.8	9
4735	Spark-Based Political Event Coding. , 2016, , .		14
4736	Efficient Cloud-Based Framework for Big Data Classification. , 2016, , .		3
4737	Grasping Popular Applications in Cellular Networks With Big Data Analytics Platforms. IEEE Transactions on Network and Service Management, 2016, 13, 681-695.	3.2	8
4739	Benchmarking performance for migrating a relational application to a parallel implementation. Future Generation Computer Systems, 2016, 63, 148-156.	4.9	6
4740	Critical Study of Performance Parameters on Distributed File Systems Using MapReduce. Procedia Computer Science, 2016, 78, 224-232.	1.2	7
4741	Security and privacy aspects in MapReduce on clouds: A survey. Computer Science Review, 2016, 20, 1-28.	10.2	67

#	ARTICLE	IF	CITATIONS
4742	A Density-Grid Based Clustering Algorithm on Data Stream Using Resilient Distributed Datasets. Lecture Notes in Computer Science, 2016, , 316-322.	1.0	1
4743	Distributed Gaussian Mixture Model Summarization Using the MapReduce Framework. Lecture Notes in Computer Science, 2016, , 323-335.	1.0	2
4744	Twister Platform for MapReduce Applications on a Docker Container. , 2016, , .		0
4745	LCP Array Construction in External Memory. Journal of Experimental Algorithmics, 2016, 21, 1-22.	0.7	6
4746	Efficient MapReduce Kernel k-Means for Big Data Clustering. , 2016, , .		7
4747	Architecture for speeding up program execution with cloud technology. Journal of Supercomputing, 2016, 72, 3601-3618.	2.4	5
4748	Transparent fault tolerance for scalable functional computation. Journal of Functional Programming, 2016, 26, .	0.5	6
4749	DBStream: A holistic approach to large-scale network traffic monitoring and analysis. Computer Networks, 2016, 107, 5-19.	3.2	24
4750	micMR: An efficient MapReduce framework for CPU&MIC heterogeneous architecture. Journal of Parallel and Distributed Computing, 2016, 93-94, 120-131.	2.7	3
4751	MapReduce short jobs optimization based on resource reuse. Microprocessors and Microsystems, 2016, 47, 178-187.	1.8	10
4752	A Binary-based MapReduce Analysis for Cloud Logs. Procedia Computer Science, 2016, 83, 1213-1218.	1.2	9
4753	Increasing digital investigator availability through efficient workflow management and automation. , 2016, , .		4
4754	Big Data Meet Green Challenges: Greening Big Data. IEEE Systems Journal, 2016, 10, 873-887.	2.9	189
4755	A Parallel Patient Treatment Time Prediction Algorithm and Its Applications in Hospital Queuing-Recommendation in a Big Data Environment. IEEE Access, 2016, 4, 1767-1783.	2.6	44
4756	Spark-Based Large-Scale Matrix Inversion for Big Data Processing. IEEE Access, 2016, 4, 2166-2176.	2.6	43
4757	Capabilities of Raspberry Pi 2 for Big Data and Video Streaming Applications in Data Centres. Lecture Notes in Computer Science, 2016, , 183-198.	1.0	1
4758	Dimension independent cosine similarity for collaborative filtering using MapReduce. , 2016, , .		1
4759	Detecting Events in Streaming Multimedia with Big Data Techniques. , 2016, , .		2

#	ARTICLE	IF	CITATIONS
4760	A Big Data Clustering Algorithm for Mitigating the Risk of Customer Churn. IEEE Transactions on Industrial Informatics, 2016, 12, 1270-1281.	7.2	96
4761	Energy-Efficient Big Data Analytics in Datacenters. Advances in Computers, 2016, , 59-101.	1.2	22
4762	CAT: A Cost-Aware Translator for SQL-query workflow to MapReduce jobflow. Data and Knowledge Engineering, 2016, 102, 42-56.	2.1	1
4763	Dimensions for Evaluating Cloud Resource Orchestration Frameworks. Computer, 2016, 49, 24-33.	1.2	11
4764	Negative Correlation Discovery for Big Multimedia Data Semantic Concept Mining and Retrieval. , 2016, , .		9
4765	USTF: A Unified System of Team Formation. IEEE Transactions on Big Data, 2016, 2, 70-84.	4.4	24
4767	Evaluation of non-linear power estimation models in a computing cluster. Sustainable Computing: Informatics and Systems, 2016, 11, 26-37.	1.6	0
4768	Energy-Aware Programming Model for Distributed Infrastructures. , 2016, , .		5
4769	Assessing Big Data SQL Frameworks for Analyzing Event Logs. , 2016, , .		6
4770	Black-Box Optimization of Hadoop Parameters Using Derivative-Free Optimization. , 2016, , .		4
4771	A streaming graph partitioning approach on imbalance cluster. , 2016, , .		0
4772	Probabilistic Graphical Models on Multi-Core CPUs Using Java 8. IEEE Computational Intelligence Magazine, 2016, 11, 41-54.	3.4	9
4773	Performance Analysis of a Parallel, Multi-node Pipeline for DNA Sequencing. Lecture Notes in Computer Science, 2016, , 233-242.	1.0	2
4774	The Performance Evaluation of the Java Implementation of Graph500. Lecture Notes in Computer Science, 2016, , 221-230.	1.0	5
4775	Cloud Computing for the Distribution of Numerical Weather Prediction Outputs. , 2016, , 121-135.		3
4776	An efficient algorithm to construct disjoint path covers of DCell networks. Theoretical Computer Science, 2016, 609, 197-210.	0.5	38
4777	Bounded link prediction in very large networks. Physica A: Statistical Mechanics and Its Applications, 2016, 457, 202-214.	1.2	17
4778	Privacy Preserving Unstructured Big Data Analytics: Issues and Challenges. Procedia Computer Science, 2016, 78, 120-124.	1.2	20

#	ARTICLE	IF	CITATIONS
4779	Impacts of Task Re-Execution Policy on MapReduce Jobs. Computer Journal, 2016, 59, 701-714.	1.5	3
4780	MapReduce: Review and open challenges. Scientometrics, 2016, 109, 389-422.	1.6	74
4781	Join Query Processing in Data Quality Management. Lecture Notes in Computer Science, 2016, , 329-342.	1.0	3
4782	D-Ocean: an unstructured data management system for data ocean environment. Frontiers of Computer Science, 2016, 10, 353-369.	1.6	12
4783	Robots at the Edge of the Cloud. Lecture Notes in Computer Science, 2016, , 3-13.	1.0	3
4784	mtDNA-Server: next-generation sequencing data analysis of human mitochondrial DNA in the cloud. Nucleic Acids Research, 2016, 44, W64-W69.	6.5	144
4786	A Nearest Neighbours-Based Algorithm for Big Time Series Data Forecasting. Lecture Notes in Computer Science, 2016, , 174-185.	1.0	20
4787	Efficient top-k similarity document search utilizing distributed file systems and cosine similarity. Cluster Computing, 2016, 19, 109-126.	3.5	17
4788	Optimizing jobs' completion time in cloud systems during virtual machine placement. , 2016, , .		4
4789	Conflicts to Harmony: A Framework for Resolving Conflicts in Heterogeneous Data by Truth Discovery. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 1986-1999.	4.0	78
4790	Big Data Usage. , 2016, , 143-165.		13
4791	Nomadic Computing for Big Data Analytics. Computer, 2016, 49, 52-60.	1.2	3
4792	SOF: Zero Configuration Simulation Optimization Framework on the Cloud. , 2016, , .		4
4793	K Nearest Neighbour Joins for Big Data on MapReduce: A Theoretical and Experimental Analysis. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 2376-2392.	4.0	36
4794	GraphOps. , 2016, , .		65
4795	Warehousing and Protecting Big Data. , 2016, , .		6
4796	Big data projects: just jump right in!. International Journal of Pervasive Computing and Communications, 2016, 12, 260-288.	1.1	21
4797	Analysis of Network IO Performance in Hadoop Cluster Environments Based on Docker Containers. Advances in Intelligent Systems and Computing, 2016, , 227-237.	0.5	2

#	ARTICLE	IF	CITATIONS
4798	A Borda count for collective sentiment analysis. <i>Annals of Mathematics and Artificial Intelligence</i> , 2016, 77, 281-302.	0.9	12
4799	Wiggins. , 2016, , .		0
4800	Similarity Search on Massive Data Based on FPGA. <i>Lecture Notes in Computer Science</i> , 2016, , 343-352.	1.0	0
4801	A threat monitoring system in enterprise networks with smart mobiles. <i>International Journal of Security and Networks</i> , 2016, 11, 66.	0.1	0
4802	A View on Fuzzy Systems for Big Data: Progress and Opportunities. <i>International Journal of Computational Intelligence Systems</i> , 2016, 9, 69.	1.6	47
4803	Enhance parallel input/output with cross-bundle aggregation. <i>International Journal of High Performance Computing Applications</i> , 2016, 30, 241-256.	2.4	3
4804	Accounting for the social: Investigating commensuration and Big Data practices at Facebook. <i>Big Data and Society</i> , 2016, 3, 205395171663136.	2.6	15
4805	A secure and efficiently searchable health information architecture. <i>Journal of Biomedical Informatics</i> , 2016, 61, 237-246.	2.5	9
4806	Cyber security data sources for dynamic network research. <i>Security Science and Technology</i> , 2016, , 37-65.	0.5	72
4807	MIRACLE: A multiple independent random walks community parallel detection algorithm for big graphs. <i>Journal of Network and Computer Applications</i> , 2016, 70, 89-101.	5.8	15
4808	CloudFIT, a PaaS Platform for IoT Applications over Pervasive Networks. <i>Communications in Computer and Information Science</i> , 2016, , 20-32.	0.4	2
4810	Tools, Use Cases, and Discussions. , 2016, , 135-150.		0
4811	Workloads in the Clouds. <i>Springer Series in Reliability Engineering</i> , 2016, , 525-550.	0.3	33
4812	SHYAM: A System for Autonomic Management of Virtual Clusters in Hybrid Clouds. <i>Communications in Computer and Information Science</i> , 2016, , 363-373.	0.4	0
4813	Characterizing gene sets using discriminative random walks with restart on heterogeneous biological networks. <i>Bioinformatics</i> , 2016, 32, 2167-2175.	1.8	38
4814	Optimization for Speculative Execution in Big Data Processing Clusters. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2016, , 1-1.	4.0	21
4815	Cooperative Big Data Processing Engine for Fast Reaction in Internet of Things Environment: Greater Than the Sum of Its Parts. <i>Lecture Notes in Electrical Engineering</i> , 2016, , 145-149.	0.3	5
4816	cOSPREY: A Cloud-Based Distributed Algorithm for Large-Scale Computational Protein Design. <i>Journal of Computational Biology</i> , 2016, 23, 737-749.	0.8	3

#	ARTICLE	IF	CITATIONS
4817	Rail-RNA: scalable analysis of RNA-seq splicing and coverage. <i>Bioinformatics</i> , 2017, 33, 4033-4040.	1.8	57
4818	Efficient and Scalable Workflows for Genomic Analyses. , 2016, , .		7
4819	Robust Large-Scale Machine Learning in the Cloud. , 2016, , .		12
4821	Optimization Analysis of Hadoop. <i>Communications in Computer and Information Science</i> , 2016, , 520-532.	0.4	0
4822	Clustering Protein Structures with Hadoop. <i>Lecture Notes in Computer Science</i> , 2016, , 141-153.	1.0	2
4823	Matrix Computations and Optimization in Apache Spark. , 2016, , .		51
4824	Big data analytics in bioinformatics: architectures, techniques, tools and issues. <i>Network Modeling Analysis in Health Informatics and Bioinformatics</i> , 2016, 5, 1.	1.2	28
4825	NXgraph: An efficient graph processing system on a single machine. , 2016, , .		63
4826	Using smartphones for prototyping semantic sensor analysis systems. , 2016, , .		4
4827	Faster cloud Star Joins with Reduced Disk Spill and Network Communication. <i>Procedia Computer Science</i> , 2016, 80, 74-85.	1.2	10
4828	RPK-table based efficient algorithm for join-aggregate query on MapReduce. <i>CAAI Transactions on Intelligence Technology</i> , 2016, 1, 79-89.	3.4	1
4829	A Four-Layer Architecture for Online and Historical Big Data Analytics. , 2016, , .		3
4830	Effective implementation of data segregation & extraction using big data in E - health insurance as a service. , 2016, , .		10
4831	A Data Science Model for Big Data Analytics of Frequent Patterns. , 2016, , .		7
4832	A Two-Phase MapReduce Algorithm for Scalable Preference Queries over High-Dimensional Data. , 2016, , .		6
4833	MapReduce for Elliptic Curve Discrete Logarithm Problem. , 2016, , .		3
4834	CURT MapReduce: Caching and Utilizing Results of Tasks for MapReduce on Cloud Computing. , 2016, , .		1
4835	Efficient Parallel Algorithms for k-Center Clustering. , 2016, , .		2

#	ARTICLE	IF	CITATIONS
4836	SpongeNet: Towards bandwidth guarantees of cloud datacenter with two-phase VM placement. , 2016, , .		3
4837	Guided Walk. , 2016, , .		9
4838	Incremental, iterative data processing with timely dataflow. Communications of the ACM, 2016, 59, 75-83.	3.3	22
4839	Scalable and cooperative big data mining platform design for smart grid. , 2016, , .		1
4840	GfLink: An In-Memory Computing Architecture on Heterogeneous CPU-GPU Clusters for Big Data. , 2016, , .		16
4841	A Cloud Computing Strategy for Region-Growing Segmentation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9, 5294-5303.	2.3	11
4842	Fast Social Service Network Construction Using Map-Reduce for Efficient Service Discovery. , 2016, , .		1
4843	An Introduction to Big Data, High Performance Computing, High-Throughput Computing, and Hadoop. , 2016, , 1-12.		6
4844	Big data applications in operations/supply-chain management: A literature review. Computers and Industrial Engineering, 2016, 101, 528-543.	3.4	286
4845	Application domain and programming language readability yardsticks. , 2016, , .		4
4846	Challenges and Approaches in Spatial Big Data Management. , 2016, , 3-13.		0
4847	Hadoop Performance Self-Tuning Using a Fuzzy-Prediction Approach. , 2016, , .		10
4848	Cloud-based Data-intensive Framework towards fault diagnosis in large-scale petrochemical plants. , 2016, , .		9
4849	The modeling of big traffic data processing based on cloud computing. , 2016, , .		2
4850	On Efficient Hierarchical Storage for Big Data Processing. , 2016, , .		17
4851	Cloud computing driven efficient mapping on soil moisture under sensor web environment. , 2016, , .		5
4852	A survey on reconfigurable accelerators for cloud computing. , 2016, , .		64
4853	Performance analysis of container-based hadoop cluster: OpenVZ and LXC. , 2016, , .		11

#	ARTICLE	IF	CITATIONS
4854	An Approach for Modeling and Ranking Node-Level Stragglers in Cloud Datacenters. , 2016, , .		6
4855	Faster: A Low Overhead Framework for Massive Data Analysis. , 2016, , .		0
4856	Graph Analytics Through Fine-Grained Parallelism. , 2016, , .		11
4857	Toward large-scale crop production forecasts for global food security. IBM Journal of Research and Development, 2016, 60, 5:1-5:11.	3.2	5
4858	Phoenix: A MapReduce implementation with new enhancements. , 2016, , .		2
4859	Proxy-Guided Load Balancing of Graph Processing Workloads on Heterogeneous Clusters. , 2016, , .		12
4860	Parallel k-Means++ for Multiple Shared-Memory Architectures. , 2016, , .		5
4861	High-Responsive Scheduling with MapReduce Performance Prediction on Hadoop YARN. , 2016, , .		5
4862	Big Data-Driven Service Composition Using Parallel Clustered Particle Swarm Optimization in Mobile Environment. IEEE Transactions on Services Computing, 2016, 9, 806-817.	3.2	73
4863	Counting Clusters in Twitter Posts. , 2016, , .		5
4864	Enterprise Architecture Analytics and Decision Support. Intelligent Systems Reference Library, 2016, , 201-218.	1.0	4
4865	Searchable Encryption for Biometric Identification Revisited. Lecture Notes in Computer Science, 2016, , 113-129.	1.0	0
4866	Contextual abstraction in a type system for component-based high performance computing platforms. Science of Computer Programming, 2016, 132, 96-128.	1.5	5
4867	A Hierarchical Distributed Processing Framework for Big Image Data. IEEE Transactions on Big Data, 2016, 2, 297-309.	4.4	18
4868	Efran (O): "Efficient Scalar Homomorphic Scheme on MapReduce for Data Privacy Preserving". , 2016, , .		3
4869	Large-Scale Processing Systems of Structured Data. SpringerBriefs in Computer Science, 2016, , 41-52.	0.2	0
4870	Large-Scale Stream Processing Systems. SpringerBriefs in Computer Science, 2016, , 75-89.	0.2	0
4871	Integration of Big Data for Connected Cars Applications Based on Tethered Connectivity. Procedia Computer Science, 2016, 98, 554-559.	1.2	5

#	ARTICLE	IF	CITATIONS
4872	Quantifying the performance impact of large pages on in-memory big-data workloads. , 2016, , .		9
4873	Scheduling jobs with non-uniform demands on multiple servers without interruption. , 2016, , .		9
4874	CHIEF: Controller Farm for Clouds of Software-Defined Community Networks. , 2016, , .		7
4875	The Design and Implementation of Big Data Platform for Telecom Operators. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2016, , 3-11.	0.2	0
4876	Improving the efficiency of IRWLS SVMs using parallel Cholesky factorization. Pattern Recognition Letters, 2016, 84, 91-98.	2.6	7
4877	MapReduce-Based Distributed Video Encoding Using Content-Aware Video Segmentation and Scheduling. IEEE Access, 2016, 4, 6802-6815.	2.6	7
4878	Distributed Incremental Graph Analysis. , 2016, , .		0
4879	Macaca: A Scalable and Energy-Efficient Platform for Coupling Cloud Computing with Distributed Embedded Computing. , 2016, , .		1
4880	Collating time-series resource data for system-wide job profiling. , 2016, , .		2
4881	Feature Extraction and Malware Detection on Large HTTPS Data Using MapReduce. Lecture Notes in Computer Science, 2016, , 311-324.	1.0	2
4882	Towards Automatic Memory Tuning for In-Memory Big Data Analytics in Clusters. , 2016, , .		7
4883	BAYWATCH: Robust Beaconing Detection to Identify Infected Hosts in Large-Scale Enterprise Networks. , 2016, , .		17
4884	Providing bandwidth guarantees, work conservation and low latency simultaneously in the cloud. , 2016, , .		28
4885	Application of Big Data Analytics via Cloud Computing. , 2016, , .		5
4886	Distributed Top-k Keyword Search over Very Large Databases with MapReduce. , 2016, , .		3
4887	Optimizing MapReduce Framework through Joint Scheduling of Overlapping Phases. , 2016, , .		9
4888	Implementation of parallel multi-objective artificial bee colony algorithm based on spark platform. , 2016, , .		4
4889	Optimization of RDF reasoning based on MapReduce. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
4890	A general and fast distributed system for large-scale dynamic programming applications. Parallel Computing, 2016, 60, 1-21.	1.3	4
4891	SD-HDFS: Secure Deletion in Hadoop Distributed File System. , 2016, , .		4
4892	A Scalable Probabilistic Change Detection Algorithm for Very High Resolution (VHR) Satellite Imagery. , 2016, , .		0
4894	Big Data Technologies in Automotive Development. ATZelektronik Worldwide, 2016, 11, 42-47.	0.1	2
4895	Application of scaling genetic algorithms using mapreduce in optimizing the configuration of PMU. , 2016, , .		2
4896	Software defined Network-on-Chip for scalable CMPs. , 2016, , .		11
4897	Efficient Virtual Network Embedding for Variable Size Virtual Machines in Fat-Tree Data Centers. , 2016, , .		2
4898	NEPTUNE: Real Time Stream Processing for Internet of Things and Sensing Environments. , 2016, , .		22
4899	A Case for Data Commons: Toward Data Science as a Service. Computing in Science and Engineering, 2016, 18, 10-20.	1.2	58
4900	Cross-Lingual Topic Discovery From Multilingual Search Engine Query Log. ACM Transactions on Information Systems, 2016, 35, 1-28.	3.8	10
4901	Energy efficiency of large scale graph processing platforms. , 2016, , .		2
4902	Empowering R with High Performance Computing Resources for Big Data Analytics. , 2016, , 191-217.		9
4903	Efficient computation of the Weighted Clustering Coefficient. Internet Mathematics, 2016, 12, 381-401.	0.7	0
4904	Modified particle swarm optimization for solving traveling salesman problem based on a Hadoop MapReduce framework. , 2016, , .		4
4905	A Scheduling Algorithm for Hadoop MapReduce Workflows with Budget Constraints in the Heterogeneous Cloud. , 2016, , .		8
4906	Efficient large-scale trace checking using mapreduce. , 2016, , .		12
4907	Reduction Drawing. , 2016, , .		19
4908	Hybrid scheduler to overcome the negative impact of job preemption for heterogeneous Hadoop systems. , 2016, , .		5

#	ARTICLE	IF	CITATIONS
4909	A Scalability Comparison Study of Data Management Approaches for Smart Metering Systems. , 2016, , .		5
4910	IPARBC: An Improved Parallel Association Rule Based on MapReduce Framework. , 2016, , .		1
4911	Sarcastic sentiment detection in tweets streamed in real time: a big data approach. Digital Communications and Networks, 2016, 2, 108-121.	2.7	97
4912	Tolerating correlated failures in Massively Parallel Stream Processing Engines. , 2016, , .		25
4914	Incremental time series algorithms for IoT analytics. , 2016, , .		4
4915	Multithreaded and Spark parallelization of feature selection filters. Journal of Computational Science, 2016, 17, 609-619.	1.5	35
4916	Research and implementation of big data preprocessing system based on Hadoop. , 2016, , .		10
4917	Scalable analytics to detect DNS misuse for establishing stealthy communication channels. IBM Journal of Research and Development, 2016, 60, 3:1-3:14.	3.2	8
4918	An Interoperability Framework and Distributed Platform for Fast Data Applications. , 2016, , 3-39.		1
4919	Efficient Online Surveillance Video Processing Based on Spark Framework. Lecture Notes in Computer Science, 2016, , 309-318.	1.0	7
4920	High-performance predictor for critical unstable generators based on scalable parallelized neural networks. Journal of Modern Power Systems and Clean Energy, 2016, 4, 414-426.	3.3	15
4921	Strategies and Principles of Distributed Machine Learning on Big Data. Engineering, 2016, 2, 179-195.	3.2	107
4922	The Six Pillars for Building Big Data Analytics Ecosystems. ACM Computing Surveys, 2017, 49, 1-36.	16.1	32
4923	Molecular Dynamics Simulations of Hydration Effects on Solvation, Diffusivity, and Permeability in Chitosan/Chitin Films. Journal of Physical Chemistry B, 2016, 120, 8997-9010.	1.2	32
4925	Stochastic bounds in Forkâ€œJoin queueing systems under full and partial mapping. Queueing Systems, 2016, 83, 261-291.	0.6	23
4926	Advances in Large-Scale Metrology â€œ Review and future trends. CIRP Annals - Manufacturing Technology, 2016, 65, 643-665.	1.7	158
4927	A Unified Approach to Data Modeling and Management in Big Data Era. , 2016, , 95-116.		4
4928	Network-aware embedding of virtual machine clusters onto federated cloud infrastructure. Journal of Systems and Software, 2016, 120, 89-104.	3.3	28

#	ARTICLE	IF	CITATIONS
4929	The anatomy of big data computing. Software - Practice and Experience, 2016, 46, 79-105.	2.5	125
4930	Big Data in Ocean Observation: Opportunities and Challenges. Lecture Notes in Computer Science, 2016, , 212-222.	1.0	5
4931	Big Graph Mining: Frameworks and Techniques. Big Data Research, 2016, 6, 1-10.	2.6	44
4932	Jupiter rising. Communications of the ACM, 2016, 59, 88-97.	3.3	46
4933	General-Purpose Big Data Processing Systems. SpringerBriefs in Computer Science, 2016, , 15-39.	0.2	3
4934	Approximate Parallel High Utility Itemset Mining. Big Data Research, 2016, 6, 26-42.	2.6	47
4935	Next-generation genotype imputation service and methods. Nature Genetics, 2016, 48, 1284-1287.	9.4	2,828
4936	Twitter UserRank using Hadoop MapReduce. , 2016, , .		3
4937	Large-Scale Graph Processing Systems. SpringerBriefs in Computer Science, 2016, , 53-73.	0.2	2
4938	Efficient Large Outer Joins over MapReduce. Lecture Notes in Computer Science, 2016, , 334-346.	1.0	3
4939	H-WorD: Supporting Job Scheduling in Hadoop with Workload-Driven Data Redistribution. Lecture Notes in Computer Science, 2016, , 306-320.	1.0	0
4940	A genetic algorithm-based job scheduling model for big data analytics. Eurasip Journal on Wireless Communications and Networking, 2016, 2016, 152.	1.5	17
4941	Adding data analytics capabilities to scaled-out object store. Journal of Systems and Software, 2016, 121, 16-27.	3.3	1
4942	A fine-grained load balancing technique for improving partition-parallel-based ontology matching approaches. Knowledge-Based Systems, 2016, 111, 17-26.	4.0	7
4943	Parallel Processing Systems for Big Data: A Survey. Proceedings of the IEEE, 2016, 104, 2114-2136.	16.4	71
4944	A Novel J wave Detection Method Based on Massive ECG Data and MapReduce. Lecture Notes in Computer Science, 2016, , 399-408.	1.0	2
4945	Review based service recommendation for big data. , 2016, , .		8
4946	Similarity-based Change Detection for RDF in MapReduce. Procedia Computer Science, 2016, 91, 789-797.	1.2	4

#	ARTICLE	IF	CITATIONS
4947	A Collective Communication Layer for the Software Stack of Big Data Analytics. , 2016, , .		1
4948	Scheduling for response time in Hadoop MapReduce. , 2016, , .		9
4949	Classification of customer requirements on Map Reduce-based Naive Bayes. , 2016, , .		4
4950	Graph Topology Abstraction for Distributed Path Queries. , 2016, , .		0
4951	An unsupervised classification process for large datasets using web reasoning. , 2016, , .		0
4952	On the Use of Abstract Models for RDF/S Provenance. , 2016, , 419-440.		0
4953	Energy big data: A survey. IEEE Access, 2016, 4, 3844-3861.	2.6	275
4954	Decisions as a service for application centric real time analytics. , 2016, , .		0
4955	An Efficient Parallel Block Processing Approach for K -Means Algorithm for High Resolution Orthoimagery Satellite Images. Procedia Computer Science, 2016, 89, 623-631.	1.2	6
4956	Simple is Good. , 2016, , .		11
4957	A scalable and productive workflow-based cloud platform for big data analytics. , 2016, , .		2
4958	Study on the mobile cloud framework for sociology: An empirical implementation. , 2016, , .		0
4959	S-FRULER: Scalable fuzzy rule learning through evolution for regression. Knowledge-Based Systems, 2016, 110, 255-266.	4.0	20
4960	On the architecture of a clustering platform for the analysis of big volumes of data. , 2016, , .		4
4961	Increasing the Throughput of Pipe-and-Filter Architectures by Integrating the Task Farm Parallelization Pattern. , 2016, , .		7
4962	Management of Distributed Big Data for Social Networks. , 2016, , .		14
4964	Medusa: An Efficient Cloud Fault-Tolerant MapReduce. , 2016, , .		16
4965	KOALA-F: A Resource Manager for Scheduling Frameworks in Clusters. , 2016, , .		10

#	ARTICLE	IF	CITATIONS
4966	Towards a Resource Manager for Scheduling Frameworks. , 2016, , .		0
4967	Resilient distributed computing platforms for big data analysis using Spark and Hadoop. , 2016, , .		3
4968	Coupling GPU and MPTCP to improve Hadoop/MapReduce performance. , 2016, , .		6
4969	An Improved MapReduce Algorithm for Mining Closed Frequent Itemsets. , 2016, , .		3
4970	A high accuracy surface modeling method based on GPU accelerated multi-Grid method. Transactions in GIS, 2016, 20, 991-1003.	1.0	3
4971	On the usability of shortest remaining time first policy in shared Hadoop clusters. , 2016, , .		4
4972	Sketch-based data placement among geo-distributed datacenters for cloud storages. , 2016, , .		15
4973	Towards Managing Variability in the Cloud. , 2016, , .		1
4974	Incremental Query Processing on Big Data Streams. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 2998-3012.	4.0	24
4975	Towards MapReduce based classification approaches for Intrusion Detection. , 2016, , .		3
4976	Scalable Overlapping Community Detection. , 2016, , .		6
4977	Probabilistic provisioning and scheduling in uncertain Cloud environments. , 2016, , .		7
4978	Big Data Analytics: Security and privacy challenges. , 2016, , .		77
4979	PoN: Open source solution for real-time data analysis. , 2016, , .		3
4980	A Practical Parallel Algorithm for Diameter Approximation of Massive Weighted Graphs. , 2016, , .		12
4981	Network security analysis using Big Data technology. , 2016, , .		13
4982	Data integration and visualisation for demanding marine operations. , 2016, , .		7
4983	A MapReduce architecture for embedded multiprocessor system-on-chips. IEICE Electronics Express, 2016, 13, 20151025-20151025.	0.3	1

#	ARTICLE	IF	CITATIONS
4984	Scalable memory fabric for silicon interposer-based multi-core systems. , 2016, , .		12
4985	Efficient Embedding of Dynamic Languages in Big-Data Analytics. , 2016, , .		0
4986	A novel process of viral vector barcoding and library preparation enables high-diversity library generation and recombination-free paired-end sequencing. Scientific Reports, 2016, 6, 37563.	1.6	22
4987	Accelerated learning of discriminative spatio-temporal features for action recognition. , 2016, , .		0
4988	BigDebug: interactive debugger for big data analytics in Apache Spark. , 2016, , .		10
4989	Distributed Big Data Techniques for Health Sensor Information Processing. Lecture Notes in Computer Science, 2016, , 217-227.	1.0	1
4990	Evolutionary undersampling for extremely imbalanced big data classification under apache spark. , 2016, , .		49
4991	Running krill herd algorithm on Hadoop: A performance study. , 2016, , .		1
4992	Visual Reasoning Indexing and Retrieval Using In-memory Computing. , 2016, , .		1
4993	Building a Hybrid Warehouse. ACM Transactions on Database Systems, 2016, 41, 1-38.	1.5	12
4994	Towards a Redundancy-Aware Network Stack for Data Centers. , 2016, , .		8
4995	Managing the Synchronization in the Lambda Architecture for Optimized Big Data Analysis. IEICE Transactions on Communications, 2016, E99.B, 297-306.	0.4	5
4996	On Big Data and Hydroinformatics. Procedia Engineering, 2016, 154, 184-191.	1.2	5
4997	Classification of Medical MRI Brain Images based on Hadoop. , 2016, , .		2
4998	A forensic cloud environment to address the big data challenge in digital forensics. , 2016, , .		9
4999	Scalable Cloud-Based Analysis Framework for Medical Big-Data. , 2016, , .		5
5000	A Reference Architecture for Supporting Secure Big Data Analytics over Cloud-Enabled Relational Databases. , 2016, , .		2
5001	Performance evaluation of in-memory computing on scale-up and scale-out cluster. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
5002	Spark-based large-scale matrix inversion for big data processing. , 2016, , .		1
5003	Hadoop based clustering system for genome sequencing. , 2016, , .		0
5004	Introducing a New Arabic Endpoint for DBpedia Internationalization Project. , 2016, , .		2
5005	Mining the Twitter Stream: Unravel Events, Interactions, and Communities in Real-Time. , 2016, , .		1
5006	On the Feasibility of Distributed Kernel Regression for Big Data. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 3041-3052.	4.0	20
5007	Traffic at-a-glance: Time-bounded analytics on large visual traffic data. , 2016, , .		1
5008	SmartGrids: MapReduce framework using Hadoop. , 2016, , .		5
5009	The Memory Challenge in Reduce Phase of MapReduce Applications. IEEE Transactions on Big Data, 2016, 2, 380-386.	4.4	13
5010	MrHeter: improving MapReduce performance in heterogeneous environments. Cluster Computing, 2016, 19, 1691-1701.	3.5	29
5011	Hadoop multi node cluster resource analysis. , 2016, , .		2
5012	STAC: Simultaneous Transmitting and Air Computing in Wireless Data Center Networks. IEEE Journal on Selected Areas in Communications, 2016, 34, 4024-4034.	9.7	10
5013	Using Object-Oriented Big Data Analytics to Reveal Server Performance Dead Zone. , 2016, , .		0
5014	A single-phase Phasor Measurement Unit for smart distribution systems. , 2016, , .		0
5015	Using High Performance Computing for Conquering Big Data. , 2016, , 13-30.		0
5016	An empirical comparison of Big Graph frameworks in the context of network analysis. Social Network Analysis and Mining, 2016, 6, 1.	1.9	8
5017	Computational challenges for sentiment analysis in life sciences. , 2016, , .		7
5018	OpERA: Opportunistic and Efficient Resource Allocation in Hadoop YARN by Harnessing Idle Resources. , 2016, , .		6
5019	Border crossing data acquisition, analysis, and use. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
5020	Optimized Multiple Platforms for Big Data Analysis. , 2016, , .		4
5021	A scalable GPU-enabled framework for training deep neural networks. , 2016, , .		4
5022	A Parallel Encoding Method of XML User Query Patterns. , 2016, , .		0
5023	Middleware Support for Self-aware Computing Systems. Natural Computing Series, 2016, , 215-238.	2.2	1
5024	Reference Architecture for Self-aware and Self-expressive Computing Systems. Natural Computing Series, 2016, , 37-49.	2.2	6
5025	Big data management processing with Hadoop MapReduce and spark technology: A comparison. , 2016, , .		38
5026	Improved Productivity of Mosaic Image by K-medoids and Feature Selection Mechanism on a Hadoop-Based Framework. , 2016, , .		0
5027	A prediction approach for correlated failures in distributed computing systems. , 2016, , .		3
5028	SparkArray: An Array-Based Scientific Data Management System Built on Apache Spark. , 2016, , .		8
5029	Harmony Search for Data Mining with Big Data. Lecture Notes in Computer Science, 2016, , 553-565.	1.0	1
5030	Towards cost-effective and high-performance caching middleware for distributed systems. International Journal of Big Data Intelligence, 2016, 3, 92.	0.4	7
5031	Mining large-scale repetitive sequences in a MapReduce setting. International Journal of Data Mining and Bioinformatics, 2016, 14, 210.	0.1	4
5032	A novel near-parallel version of k-means algorithm for n-dimensional data objects using MPI. International Journal of Grid and Utility Computing, 2016, 7, 80.	0.1	20
5033	A framework for cloud-aware development of bag-of-tasks scientific applications. International Journal of Grid and Utility Computing, 2016, 7, 130.	0.1	5
5034	Improving Mapreduce for Incremental Processing Using Map Data Storage. Procedia Computer Science, 2016, 87, 288-293.	1.2	3
5035	Using Scalable Data Mining for Predicting Flight Delays. ACM Transactions on Intelligent Systems and Technology, 2017, 8, 1-20.	2.9	65
5036	Dynamic Process Migration Based on Block Access Patterns Occurring in Storage Servers. Transactions on Architecture and Code Optimization, 2016, 13, 1-20.	1.6	1
5037	Real-time social network graph analysis using StreamMine3G. , 2016, , .		5

#	ARTICLE	IF	CITATIONS
5038	Blocking for large-scale Entity Resolution: Challenges, algorithms, and practical examples. , 2016, , .		6
5039	Fast Covariant VLAD for Image Search. IEEE Transactions on Multimedia, 2016, 18, 1843-1854.	5.2	6
5040	Scatter matrix concordance as a diagnostic for regressions on subsets of data. Statistical Analysis and Data Mining, 2016, 9, 249-259.	1.4	0
5041	Big Data in the construction industry: A review of present status, opportunities, and future trends. Advanced Engineering Informatics, 2016, 30, 500-521.	4.0	428
5042	Storing shared documents that are customized by users in cloud computing. Computing (Vienna/New) Tj ETQq0 0 0 rgBT /Overlock 10 T	3.2	4
5043	TetriSched. , 2016, , .		107
5044	Parallel Priority-Flood depression filling for trillion cell digital elevation models on desktops or clusters. Computers and Geosciences, 2016, 96, 56-68.	2.0	31
5045	A new data-grouping-aware dynamic data placement method that take into account jobs execute frequency for Hadoop. Microprocessors and Microsystems, 2016, 47, 161-169.	1.8	12
5046	CloudPhylo: a fast and scalable tool for phylogeny reconstruction. Bioinformatics, 2017, 33, 438-440.	1.8	7
5047	Towards co-designed optimizations in parallel frameworks. , 2016, , .		2
5048	Sensing, processing and analytics. , 2016, , .		5
5049	Storage and Database Management for Big Data. , 2016, , 15-41.		3
5050	Challenges and Approaches in Spatial Big Data Management. , 2016, , 19-30.		2
5051	Tolhit " A Scheduling Algorithm for Hadoop Cluster. Procedia Computer Science, 2016, 89, 203-208.	1.2	18
5052	A One-Class NIDS for SDN-Based SCADA Systems. , 2016, , .		17
5053	Providing caches for reduce tasks in a MapReduce cloud. , 2016, , .		1
5054	Big Data capabilities and readiness of South African retail organisations. , 2016, , .		16
5055	A performance benchmark for NetFlow data analysis on distributed stream processing systems. , 2016, , .		12

#	ARTICLE	IF	CITATIONS
5056	Dynamic Scheduling Algorithm for Reducing Start Time in Hadoop. , 2016, , .		2
5057	End-to-end internet speed analysis of mobile networks with mapReduce. , 2016, , .		3
5058	Call center performance evaluation using big data analytics. , 2016, , .		12
5059	A Multi-objective evolutionary fuzzy system for big data. , 2016, , .		5
5060	Distributed Algorithms for Computing Very Large Thresholded Covariance Matrices. ACM Transactions on Knowledge Discovery From Data, 2016, 11, 1-25.	2.5	1
5061	GEMs: shared-memory parallel programming for Node.js. , 2016, , .		2
5062	Healthy behavior with social apps: Proposal for evolution study of the use of fitness social apps on Facebook. , 2016, , .		1
5063	Parallel Hardware Merge Sorter. , 2016, , .		39
5064	A cloud-based large-scale distributed video analysis system. , 2016, , .		9
5065	CloudArray: Easing huge image processing. , 2016, , .		3
5066	Data-aware optimization of bioinformatics workflows in hybrid clouds. Journal of Big Data, 2016, 3, .	6.9	2
5067	Research of Construction and Application of Cloud Storage in the Environment of Industry 4.0. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2016, , 104-113.	0.2	1
5068	A Study of Big Data Computing Platforms: Fairness and Energy Consumption. , 2016, , .		3
5069	On Information Value of Top N Statistics. , 2016, , .		1
5070	Parallel Prime Number Labeling of Large XML Data Using MapReduce. , 2016, , .		3
5071	Challenges of Programming a System with Heterogeneous Memories and Heterogeneous Processors. , 2016, , .		2
5072	Nubytics: Scalable cloud services for data analysis and prediction. , 2016, , .		1
5073	Simple, efficient location-based routing for data center network using IP address hierarchy. International Journal of Network Management, 2016, 26, 492-514.	1.4	3

#	ARTICLE	IF	CITATIONS
5074	DSS: A Scalable and Efficient Stratified Sampling Algorithm for Large-Scale Datasets. Lecture Notes in Computer Science, 2016, , 133-146.	1.0	2
5075	Iterative sparse matrix-vector multiplication on in-memory cluster computing accelerated by GPUs for big data. , 2016, , .		2
5076	New techniques to curtail the tail latency in stream processing systems. , 2016, , .		5
5077	An efficient strategy for the collection and storage of large volumes of data for computation. Journal of Big Data, 2016, 3, .	6.9	12
5078	Distributed XPath query processing over large XML data based on MapReduce framework. , 2016, , .		1
5079	Prophet: Scheduling Executors with Time-Varying Resource Demands on Data-Parallel Computation Frameworks. , 2016, , .		15
5080	Performance and Monetary Cost of Large-Scale Distributed Graph Processing on Amazon Cloud. , 2016, , .		1
5081	Parallel Algorithms for Constructing Range and Nearest-Neighbor Searching Data Structures. , 2016, , .		18
5082	elephant56. , 2016, , .		12
5083	On exploring efficient shuffle design for in-memory MapReduce. , 2016, , .		4
5084	Tight bounds on one- and two-pass MapReduce algorithms for matrix multiplication. , 2016, , .		2
5086	A novel K-means based clustering algorithm for big data. , 2016, , .		11
5087	SAR image change detection based on Spark-FLICM algorithm. , 2016, , .		1
5088	Exploring machine learning methods for the Star/Galaxy Separation Problem. , 2016, , .		6
5089	Locality-Aware and Energy-Aware Job Pre-Assignment for Mapreduce. , 2016, , .		2
5090	Imbalanced Data Classification: A Novel Re-sampling Approach Combining Versatile Improved SMOTE and Rough Sets. Lecture Notes in Computer Science, 2016, , 31-42.	1.0	11
5091	Modeling Bidirectional Reflectance Factor of Complex Scene Using Iterative MapReduce. Communications in Computer and Information Science, 2016, , 68-77.	0.4	0
5092	Classification model to predict the learners' academic performance using big data. , 2016, , .		6

#	ARTICLE	IF	CITATIONS
5093	A comparative evaluation of open-source graph processing platforms. , 2016, , .		1
5094	Distributed mean shift clustering with approximate nearest neighbours. , 2016, , .		4
5095	Supporting Semantic Concept Retrieval with Negative Correlations in a Multimedia Big Data Mining System. International Journal of Semantic Computing, 2016, 10, 247-267.	0.4	12
5096	Distributed Decision Tree. , 2016, , .		8
5097	G-Storm: A GPU-Aware Storm Scheduler. , 2016, , .		3
5098	Automated neurovascular tracing and analysis of the knife-edge scanning microscope Rat Nissl data set using a computing cluster. , 2016, 2016, 6445-6448.		2
5099	Towards a generic computation model for smart city platforms. , 2016, , .		9
5100	Deploying and monitoring hadoop MapReduce analytics on single-chip cloud computer. , 2016, , .		1
5101	Edge-Computing-Aware Deployment of Stream Processing Tasks Based on Topology-External Information: Model, Algorithms, and a Storm-Based Prototype. , 2016, , .		9
5102	Performance Optimization of the SSVD Collaborative Filtering Algorithm on MapReduce Architectures. , 2016, , .		2
5103	SPSRG: a prediction approach for correlated failures in distributed computing systems. Cluster Computing, 2016, 19, 1703-1721.	3.5	7
5104	UNICORE 7 " Middleware services for distributed and federated computing. , 2016, , .		20
5105	<i>WorkStream</i> -- A Design Pattern for Multicore-Enabled Finite Element Computations. ACM Transactions on Mathematical Software, 2017, 43, 1-29.	1.6	45
5106	TwigStack-MR: An Approach to Distributed XML Twig Query Using MapReduce. , 2016, , .		3
5107	Fuzzy Based Scalable Clustering Algorithms for Handling Big Data Using Apache Spark. IEEE Transactions on Big Data, 2016, 2, 339-352.	4.4	51
5108	Perspectives of Machine Learning Techniques in Big Data Mining of Cancer. , 2016, , 317-336.		0
5109	Parallel Algorithms for Multirelational Data Mining: Application to Life Science Problems. Computer Communications and Networks, 2016, , 339-363.	0.8	0
5110	DVMP: Incremental traffic-aware VM placement on heterogeneous servers in data centers. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
5111	Architecting Time-Critical Big-Data Systems. IEEE Transactions on Big Data, 2016, 2, 310-324.	4.4	58
5112	Finding database contention hotspots under large-scale workloads - A big data approach. , 2016, , .		0
5113	A novel energy efficient platform based model to enable mobile Cloud applications. , 2016, , .		4
5114	On MBR codes with replication. , 2016, , .		10
5115	Massive Data Analysis: Tasks, Tools, Applications, and Challenges. , 2016, , 11-40.		6
5117	Performance Enhanced Multiset Similarity Joins. , 2016, , .		0
5118	Parallelizing Active Memory Ants with MapReduce for Clustering Financial Time Series Data. , 2016, , .		6
5119	Effectively Handling New Relationship Formations in Closeness Centrality Analysis of Social Networks Using Anytime Anywhere Methodology. , 2016, , .		4
5120	A Parallel DBSCAN Algorithm Based on Spark. , 2016, , .		30
5121	An Analysis Platform of Road Traffic Management System Log Data Based on Distributed Storage and Parallel Computing Techniques. , 2016, , .		9
5122	Diagnosing Performance Bottlenecks in Massive Data Parallel Programs. , 2016, , .		5
5123	A Self-Correcting Connected Components Algorithm. , 2016, , .		6
5124	Credit Card Fraud Detection using Big Data Analytics. , 2016, , .		38
5125	ScaleKB: scalable learning and inference over large knowledge bases. VLDB Journal, 2016, 25, 893-918.	2.7	28
5126	The golden age for popularizing big data. Science China Information Sciences, 2016, 59, 1.	2.7	4
5127	A Configuration Management Study to Fast Massive Writing for Distributed NoSQL System. IEICE Transactions on Information and Systems, 2016, E99.D, 2269-2282.	0.4	3
5128	Toward a Learning Health-care System – Knowledge Delivery at the Point of Care Empowered by Big Data and NLP. Biomedical Informatics Insights, 2016, 8s1, BII.S37977.	4.6	56
5129	Comparative analysis of Gaussian mixture model, logistic regression and random forest for big data classification using map reduce. , 2016, , .		5

#	ARTICLE	IF	CITATIONS
5130	A statistical threat detection method based on dynamic time threshold. , 2016, , .		0
5132	Distributed RDFS Rules Reasoning for Large-Scaled RDF Graphs Using Spark. , 2016, , .		3
5133	Generic and massively concurrent computation of belief combination rules. , 2016, , .		0
5134	On the optimality of averaging in distributed statistical learning. Information and Inference, 2016, 5, 379-404.	0.9	37
5135	Analyzing and scripting indian election strategies using big data via Apache Hadoop framework. , 2016, , .		5
5136	A Protocol for Extending Analytics Capability of SQL Database. , 2016, , .		1
5137	Performance Comparison and Analysis of Yarn's Schedulers with Stress Cases. , 2016, , .		1
5138	IMFSSC: An In-Memory Distributed File System Framework for Super Computing. , 2016, , .		0
5139	Rack-Scale Storage Fabric: A Practical Way to Build Best-Fit Infrastructure for High-Performance Data Processing. MATEC Web of Conferences, 2016, 76, 04040.	0.1	1
5140	Log-Assisted Straggler-Aware I/O Scheduler for High-End Computing. , 2016, , .		9
5142	Replica parallelism to utilize the granularity of data. , 2016, , .		0
5143	Efficient model partitioning for distributed model transformations. , 2016, , .		6
5144	The Efficient Algorithms for Constructing Enhanced Quadtrees Using MapReduce. IEICE Transactions on Information and Systems, 2016, E99.D, 918-926.	0.4	1
5145	Function passing: a model for typed, distributed functional programming. , 2016, , .		8
5146	HTSFinder: Powerful Pipeline of DNA Signature Discovery by Parallel and Distributed Computing. Evolutionary Bioinformatics, 2016, 12, EBO.S35545.	0.6	3
5147	Security and privacy controls for streaming data in extended intelligent environments. Journal of Ambient Intelligence and Smart Environments, 2016, 8, 467-483.	0.8	5
5148	Robust and Probabilistic Failure-Aware Placement. , 2016, , .		9
5149	Minimum value sampling algorithm based on CDC. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
5150	Big Data Application in Biomedical Research and Health Care: A Literature Review. Biomedical Informatics Insights, 2016, 8, Bil.S31559.	4.6	348
5151	SAMURAI: A batch and streaming context architecture for large-scale intelligent applications and environments. Journal of Ambient Intelligence and Smart Environments, 2016, 8, 63-78.	0.8	18
5152	MRFP: Discovery Frequent Patterns Using MapReduce Frequent Pattern Growth. , 2016, , .		3
5153	Multi-agent framework for social CRM: Extracting and analyzing opinions. , 2016, , .		2
5154	Efficient semantic-aware coflow scheduling for data-parallel jobs. , 2016, , .		4
5155	Real-time big data analytics architecture for remote sensing application. , 2016, , .		4
5156	Integrating Big data paradigm in WSNs. , 2016, , .		2
5157	A cost-efficient approach for measuring Moran's index of spatial autocorrelation in geostationary satellite data. , 2016, , .		9
5158	Partitioning uncertain workloads. NETNOMICS: Economic Research and Electronic Networking, 2016, 17, 233-253.	0.9	0
5159	Index based quasi-identifier approach for privacy preservation data sets on cloud. , 2016, , .		0
5160	The Role of Big Data Analysis in New Product Development. , 2016, , .		11
5161	A subgraph mining algorithm on big data. , 2016, , .		0
5162	Scheduling Dependent Coflows with Guaranteed Job Completion Time. , 2016, , .		2
5163	Surveillance Methods of Running Condition of Chemical Equipments Based on Skyline. , 2016, , .		0
5164	Navigating the Local Modes of Big Data: The Case of Topic Models. , 2016, , 51-97.		50
5165	Predicting the risk of diabetes in big data electronic health Records by using scalable random forest classification algorithm. , 2016, , .		19
5166	A Virtualization-Based Hybrid Storage System for a Map-Reduce Framework. IEICE Transactions on Information and Systems, 2016, E99.D, 2248-2258.	0.4	0
5167	GISQAF: MapReduce guided spatial query processing and analytics system. Software - Practice and Experience, 2016, 46, 1329-1349.	2.5	2

#	ARTICLE	IF	CITATIONS
5169	An efficient graph data processing system for large-scale social network service applications. Concurrency Computation Practice and Experience, 2016, 28, 729-747.	1.4	3
5170	A MapReduce-based parallel <i>K</i> -means clustering for large-scale CIM data verification. Concurrency Computation Practice and Experience, 2016, 28, 3096-3114.	1.4	9
5171	WatershedEng: an extensible distributed stream processing framework. Concurrency Computation Practice and Experience, 2016, 28, 2487-2502.	1.4	2
5172	Multi-cache: Dynamic, Efficient Partitioning for Multi-tier Caches in Consolidated VM Environments. , 2016, , .		8
5173	Awan: Locality-Aware Resource Manager for Geo-Distributed Data-Intensive Applications. , 2016, , .		10
5174	Surviving Failures with Performance-Centric Bandwidth Allocation in Private Datacenters. , 2016, , .		5
5175	Barrier-Aware Max-Min Fair Bandwidth Sharing and Path Selection in Datacenter Networks. , 2016, , .		8
5176	Data volume reduction in covering approximation spaces with respect to twenty-two types of covering based rough sets. International Journal of Approximate Reasoning, 2016, 75, 13-38.	1.9	6
5177	The role of big data in smart city. International Journal of Information Management, 2016, 36, 748-758.	10.5	763
5178	The Invention of the Object: Object Orientation and the Philosophical Development of Programming Languages. Philosophy and Technology, 2016, 29, 335-356.	2.6	6
5179	A MapReduce approach to diminish imbalance parameters for big deoxyribonucleic acid dataset. Computer Methods and Programs in Biomedicine, 2016, 131, 191-206.	2.6	85
5180	A social community detection algorithm based on parallel grey label propagation. Computer Networks, 2016, 107, 133-143.	3.2	32
5181	Benchmarking modern distributed streaming platforms. , 2016, , .		21
5182	Towards Peer-to-Peer Solution for Utilization of Volunteer Resources to Provide Computation-as-a-Service. , 2016, , .		3
5183	Fuzzy Based Clustering Algorithms to Handle Big Data with Implementation on Apache Spark. , 2016, , .		16
5184	S4 Applications Simulator for Performance Evaluation. , 2016, , .		0
5185	Exact Vs. Approximated Diameter Calculation in Large Graphs. , 2016, , .		0
5186	An in-memory based framework for scientific data analytics. , 2016, , .		11

#	ARTICLE	IF	CITATIONS
5187	Enorm. , 2016, , .		5
5188	Performance Evaluation of TCP Congestion Control Algorithms in Data Center Networks. , 2016, , .		10
5189	Venice: Exploring server architectures for effective resource sharing. , 2016, , .		10
5190	A Practical and Sustainable Model for Learning and Teaching Data Science. , 2016, , .		29
5191	Distributed execution environment for data mining as service. , 2016, , .		0
5192	Expected similarity estimation for large-scale batch and streaming anomaly detection. Machine Learning, 2016, 105, 305-333.	3.4	36
5193	Integration and optimization of multiple big data processing platforms. Engineering Computations, 2016, 33, 1680-1704.	0.7	10
5194	Advance Reservation System for Datacenters. , 2016, , .		2
5195	SQL-SA for big data discovery polymorphic and parallelizable SQL user-defined scalar and aggregate infrastructure in Teradata Aster 6.20. , 2016, , .		3
5196	Model driven performance simulation of cloud provisioned Hadoop mapreduce applications. , 2016, , .		3
5197	Extending search-based software testing techniques to big data applications. , 2016, , .		2
5198	Java2SDG: Stateful big data processing for the masses. , 2016, , .		2
5199	Scalability in Pattern Mining. , 2016, , 177-190.		1
5200	Evaluate H2Hadoop and Amazon EMR performances by processing MR jobs in text data sets. , 2016, , .		1
5201	Ontological Pathfinding. , 2016, , .		35
5202	Smart Sampling and Optimal Dimensionality Reduction of Big Data Using Compressed Sensing. Studies in Big Data, 2016, , 251-280.	0.8	0
5203	HDCache: A Distributed Cache System for Real-Time Cloud Services. Journal of Grid Computing, 2016, 14, 407-428.	2.5	9
5204	Mining Entity Rankings. Datenbank-Spektrum, 2016, 16, 27-38.	1.2	1

#	ARTICLE	IF	CITATIONS
5205	Towards a Comprehensive Data Analytics Framework for Smart Healthcare Services. Big Data Research, 2016, 4, 44-58.	2.6	126
5206	Vertex-disjoint paths in DCell networks. Journal of Parallel and Distributed Computing, 2016, 96, 38-44.	2.7	32
5207	Global communication schemes for the numerical solution of high-dimensional PDEs. Parallel Computing, 2016, 52, 78-105.	1.3	6
5208	Demand-Addressable Sensor Network: Toward Large-Scale Active Information Acquisition. IEEE Sensors Journal, 2016, 16, 7421-7432.	2.4	3
5209	Event detection from traffic tensors: A hybrid model. Neurocomputing, 2016, 203, 22-33.	3.5	31
5210	Soft Real-Time Hadoop Scheduler for Big Data Processing in Smart Cities. , 2016, , .		9
5211	An Exploration of Designing a Hybrid Scale-Up/Out Hadoop Architecture Based on Performance Measurements. IEEE Transactions on Parallel and Distributed Systems, 2016, , 1-1.	4.0	23
5212	Biomedical Ontology Quality Assurance Using a Big Data Approach. ACM Transactions on Knowledge Discovery From Data, 2016, 10, 1-28.	2.5	16
5213	Large-Scale and Big Optimization Based on Hadoop. Studies in Big Data, 2016, , 375-389.	0.8	0
5214	Challenges in Managing Real-Time Data in Health Information System (HIS). Lecture Notes in Computer Science, 2016, , 305-313.	1.0	7
5215	Straggler Detection in Parallel Computing Systems through Dynamic Threshold Calculation. , 2016, , .		19
5216	Computing at Massive Scale: Scalability and Dependability Challenges. , 2016, , .		21
5217	Semi-automated initialization of simulations: an application to healthcare. Journal of Defense Modeling and Simulation, 2016, 13, 171-182.	1.2	4
5218	HPG pore: an efficient and scalable framework for nanopore sequencing data. BMC Bioinformatics, 2016, 17, 107.	1.2	9
5219	A survey of machine learning for big data processing. Eurasip Journal on Advances in Signal Processing, 2016, 2016, .	1.0	386
5220	An efficient Hadoop data replication method design for heterogeneous clusters. , 2016, , .		7
5222	Large-Scale Industrial Alarm Reduction and Critical Events Mining Using Graph Analytics on Spark. , 2016, , .		13
5223	Air Quality Simulations Using Big Data Programming Models. , 2016, , .		10

#	ARTICLE	IF	CITATIONS
5224	The Matsu Wheel: A Cloud-Based Framework for Efficient Analysis and Reanalysis of Earth Satellite Imagery. , 2016, , .		4
5225	Logical Aspects of Massively Parallel and Distributed Systems. , 2016, , .		0
5226	Experiences with Performing MapReduce Analysis of Scientific Data on HPC Platforms. , 2016, , .		1
5227	ENTRADA: enabling DNS big data applications. , 2016, , .		6
5228	Adaptive and Hierarchical Runtime Manager for Energy-Aware Thermal Management of Embedded Systems. Transactions on Embedded Computing Systems, 2016, 15, 1-25.	2.1	40
5229	Towards a Virtual Domain Based Authentication on MapReduce. IEEE Access, 2016, 4, 1658-1675.	2.6	5
5230	An Efficient Multidimensional Big Data Fusion Approach in Machine-to-Machine Communication. Transactions on Embedded Computing Systems, 2016, 15, 1-25.	2.1	15
5231	An integrated solutionâ€”KAGFM for mass customization in customer-oriented product design under cloud manufacturing environment. International Journal of Advanced Manufacturing Technology, 2016, 84, 85-101.	1.5	40
5232	Large-scale supervised similarity learning in networks. Knowledge and Information Systems, 2016, 48, 707-740.	2.1	2
5233	A fast and efficient Hamming LSH-based scheme for accurate linkage. Knowledge and Information Systems, 2016, 49, 861-884.	2.1	23
5234	Scalable offline monitoring of temporal specifications. Formal Methods in System Design, 2016, 49, 75-108.	0.9	18
5235	eHSim: An Efficient Hybrid Similarity Search with MapReduce. , 2016, , .		4
5236	Flexible replica placement for optimized P2P backup on heterogeneous, unreliable machines. Concurrency Computation Practice and Experience, 2016, 28, 2166-2186.	1.4	3
5237	BSP cost and scalability analysis for MapReduce operations. Concurrency Computation Practice and Experience, 2016, 28, 2503-2527.	1.4	13
5238	Online anomaly detection for multiâ€”source VMware using a distributed streaming framework. Software - Practice and Experience, 2016, 46, 1479-1497.	2.5	8
5239	The improved (2 D) 2 PCA algorithm and its parallel implementation based on image block. Microprocessors and Microsystems, 2016, 47, 170-177.	1.8	6
5240	CPL: a core language for cloud computing. , 2016, , .		7
5241	SparkGIS: Efficient Comparison and Evaluation of Algorithm Results in Tissue Image Analysis Studies. Lecture Notes in Computer Science, 2016, 9579, 134-146.	1.0	7

#	ARTICLE	IF	CITATIONS
5242	A Map Reduce-Based Nearest Neighbor Approach for Big-Data-Driven Traffic Flow Prediction. IEEE Access, 2016, 4, 2920-2934.	2.6	50
5243	Mesa. Communications of the ACM, 2016, 59, 117-125.	3.3	11
5244	Scaling Spark on HPC Systems. , 2016, , .		41
5245	Estimating rates of rare events through a multidimensional dynamic hierarchical Bayesian framework. Applied Stochastic Models in Business and Industry, 2016, 32, 340-353.	0.9	1
5246	The Research and Application of Purchase Amount Based User-KNN Algorithm on Cloud Platform in Coal Sales System. , 2016, , .		1
5247	A Load-Aware Pluggable Cloud Framework for Real-Time Video Processing. IEEE Transactions on Industrial Informatics, 2016, 12, 2166-2176.	7.2	13
5248	A framework for enriching Data Warehouse analysis with Question Answering systems. Journal of Intelligent Information Systems, 2016, 46, 61-82.	2.8	7
5249	An energy-aware bi-level optimization model for multi-job scheduling problems under cloud computing. Soft Computing, 2016, 20, 303-317.	2.1	32
5250	A distributed transcoding and content protection system. Telecommunication Systems, 2016, 61, 59-76.	1.6	5
5251	Single-scan: a fast star-join query processing algorithm. Software - Practice and Experience, 2016, 46, 319-339.	2.5	4
5252	Exploiting Analytics Shipping with Virtualized MapReduce on HPC Backend Storage Servers. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 185-196.	4.0	8
5253	Scalable graph-based OLAP analytics over process execution data. Distributed and Parallel Databases, 2016, 34, 379-423.	1.0	48
5254	Scalable Process Discovery Using Map-Reduce. IEEE Transactions on Services Computing, 2016, 9, 469-481.	3.2	27
5255	Authenticated Key Exchange Protocols for Parallel Network File Systems. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 92-105.	4.0	4
5256	OverFlow: Multi-Site Aware Big Data Management for Scientific Workflows on Clouds. IEEE Transactions on Cloud Computing, 2016, 4, 76-89.	3.1	32
5257	On Traffic-Aware Partition and Aggregation in MapReduce for Big Data Applications. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 818-828.	4.0	35
5258	Parallel implementation of multilayered neural networks based on Map-Reduce on cloud computing clusters. Soft Computing, 2016, 20, 1471-1483.	2.1	12
5259	Scalable Algorithms for Nearest-Neighbor Joins on Big Trajectory Data. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 785-800.	4.0	34

#	ARTICLE	IF	CITATIONS
5260	Distributed In-Memory Processing of All k Nearest Neighbor Queries. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 925-938.	4.0	44
5261	Wireless Communication in Data Centers: A Survey. IEEE Communications Surveys and Tutorials, 2016, 18, 1572-1595.	24.8	89
5262	The Beckman report on database research. Communications of the ACM, 2016, 59, 92-99.	3.3	82
5263	OPTIMA: On-Line Partitioning Skew Mitigation for MapReduce with Resource Adjustment. Journal of Network and Systems Management, 2016, 24, 859-883.	3.3	14
5264	Expanded cloud plumes hiding Big Data ecosystem. Future Generation Computer Systems, 2016, 59, 63-92.	4.9	59
5265	Self-healing in mobile networks with big data. IEEE Communications Magazine, 2016, 54, 114-120.	4.9	35
5266	Placement and Performance Analysis of Virtual Multicast Networks in Fat-Tree Data Center Networks. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 3013-3028.	4.0	14
5267	Cloud-based automatic test data generation framework. Journal of Computer and System Sciences, 2016, 82, 712-738.	0.9	31
5268	A New Design of In-Memory File System Based on File Virtual Address Framework. IEEE Transactions on Computers, 2016, 65, 2959-2972.	2.4	43
5269	A hierarchical-coevolutionary-MapReduce-based knowledge reduction algorithm with robust ensemble Pareto equilibrium. Information Sciences, 2016, 342, 153-175.	4.0	7
5270	Data Centric Text Processing Using MapReduce. Advances in Intelligent Systems and Computing, 2016, , 129-137.	0.5	0
5272	epiC: an extensible and scalable system for processing Big Data. VLDB Journal, 2016, 25, 3-26.	2.7	9
5274	Achieving Optimal Inter-Node Communication in Graph Partitioning Using Random Selection and Breadth-First Search. International Journal of Parallel Programming, 2016, 44, 772-800.	1.1	1
5275	Mining Large Scale Tandem Mass Spectrometry Data for Protein Modifications Using Spectral Libraries. Journal of Proteome Research, 2016, 15, 721-731.	1.8	26
5276	Benefit of Compression in Hadoop: A Case Study of Improving IO Performance on Hadoop. , 2016, , 879-890.		4
5277	ISODAC: A high performance solution for indexing and searching heterogeneous data. Journal of Systems and Software, 2016, 118, 115-133.	3.3	7
5278	A Novel Pipeline Approach for Efficient Big Data Broadcasting. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 17-28.	4.0	16
5279	A Resource Aware MapReduce Based Parallel SVM for Large Scale Image Classifications. Neural Processing Letters, 2016, 44, 161-184.	2.0	5

#	ARTICLE	IF	CITATIONS
5280	Analytic Queries over Geospatial Time-Series Data Using Distributed Hash Tables. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 1408-1422.	4.0	17
5281	An efficient key partitioning scheme for heterogeneous MapReduce clusters. , 2016, , .		4
5282	DistR: A Distributed Method for the Reachability Query over Large Uncertain Graphs. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 3172-3185.	4.0	24
5283	Scalable Linear Visual Feature Learning via Online Parallel Nonnegative Matrix Factorization. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27, 2628-2642.	7.2	11
5284	ASC: Improving spark driver performance with automatic spark checkpoint. , 2016, , .		9
5285	VENUS: A System for Streamlined Graph Computation on a Single PC. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 2230-2245.	4.0	2
5286	Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data. GigaScience, 2016, 5, 12.	3.3	114
5287	AutoFDO: automatic feedback-directed optimization for warehouse-scale applications. , 2016, , .		61
5288	Data adapter for querying and transformation between SQL and NoSQL database. Future Generation Computer Systems, 2016, 65, 111-121.	4.9	55
5289	HDInsight4PSi: Boosting performance of 3D protein structure similarity searching with HDInsight clusters in Microsoft Azure cloud. Information Sciences, 2016, 349-350, 77-101.	4.0	29
5290	Gaussian Limits for a Fork-Join Network with Nonexchangeable Synchronization in Heavy Traffic. Mathematics of Operations Research, 2016, 41, 560-595.	0.8	15
5291	Security Theories and Practices for Big Data. , 2016, , 157-192.		7
5292	An Evaluation of Distributed Processing Models for Random Walk-Based Link Prediction Algorithms Over Social Big Data. Advances in Intelligent Systems and Computing, 2016, , 919-928.	0.5	0
5293	Workload Adaptive Shared Memory Management for High Performance Network I/O in Virtualized Cloud. IEEE Transactions on Computers, 2016, 65, 3480-3494.	2.4	14
5294	Machine Learning with Big Data An Efficient Electricity Generation Forecasting System. Big Data Research, 2016, 5, 9-15.	2.6	64
5295	Efficiency Experiments on Hadoop and Giraph with PageRank. , 2016, , .		3
5296	GBC3: A Versatile Cube-Based Server-Centric Network for Data Centers. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 2895-2910.	4.0	25
5297	A Service-Based Approach to Traffic Sensor Data Integration and Analysis to Support Community-Wide Green Commute in China. IEEE Transactions on Intelligent Transportation Systems, 2016, 17, 2648-2657.	4.7	29

#	ARTICLE	IF	CITATIONS
5298	The real cost of sequencing: scaling computation to keep pace with data generation. <i>Genome Biology</i> , 2016, 17, 53.	3.8	264
5299	Learning the Multilingual Translation Representations for Question Retrieval in Community Question Answering via Non-Negative Matrix Factorization. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2016, 24, 1305-1314.	4.0	28
5300	Fault-Tolerant Scheduling for Real-Time Scientific Workflows with Elastic Resource Provisioning in Virtualized Clouds. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2016, 27, 3501-3517.	4.0	106
5301	Iterative big data clustering algorithms: a review. <i>Software - Practice and Experience</i> , 2016, 46, 107-129.	2.5	45
5302	Big Data – A Theory Model. , 2016, , .		2
5304	MapReduce based distributed learning algorithm for Restricted Boltzmann Machine. <i>Neurocomputing</i> , 2016, 198, 4-11.	3.5	26
5305	HHS: an efficient network topology for large-scale data centers. <i>Journal of Supercomputing</i> , 2016, 72, 874-899.	2.4	9
5306	An Index Scheme for Similarity Search on Cloud Computing using MapReduce over Docker Container. , 2016, , .		10
5307	A novel metadata extraction method for surveillance video. , 2016, , .		3
5308	Efficient parallelised search engine based on virtual cluster. <i>International Journal of Computational Science and Engineering</i> , 2016, 12, 53.	0.4	5
5309	Single-cell Transcriptome Study as Big Data. <i>Genomics, Proteomics and Bioinformatics</i> , 2016, 14, 21-30.	3.0	35
5310	P-Scheduler. , 2016, , .		27
5311	A formal definition of Big Data based on its essential features. <i>Library Review</i> , 2016, 65, 122-135.	1.5	623
5312	Implementing Fusion to Improve the Efficiency of Information Retrieval Using Clustering and Map Reduction. <i>Advances in Intelligent Systems and Computing</i> , 2016, , 879-884.	0.5	0
5313	A Bootstrap Metropolisâ€Hastings Algorithm for Bayesian Analysis of Big Data. <i>Technometrics</i> , 2016, 58, 304-318.	1.3	9
5314	Datacenter Design and Management: A Computer Architect's Perspective. <i>Synthesis Lectures on Computer Architecture</i> , 2016, 11, 1-121.	1.3	3
5315	Attribute Equilibrium Dominance Reduction Accelerator (DCCAEDR) Based on Distributed Coevolutionary Cloud and Its Application in Medical Records. <i>IEEE Transactions on Systems, Man, and Cybernetics: Systems</i> , 2016, 46, 384-400.	5.9	13
5316	Towards scalable rough set based attribute subset selection for intrusion detection using parallel genetic algorithm in MapReduce. <i>Simulation Modelling Practice and Theory</i> , 2016, 64, 18-29.	2.2	39

#	ARTICLE	IF	CITATIONS
5317	Handling data skew in join algorithms using MapReduce. Expert Systems With Applications, 2016, 51, 286-299.	4.4	27
5319	A Survey on Resiliency Techniques in Cloud Computing Infrastructures and Applications. IEEE Communications Surveys and Tutorials, 2016, 18, 2244-2281.	24.8	110
5320	A Survey of Role Mining. ACM Computing Surveys, 2016, 48, 1-37.	16.1	70
5322	Design Issues of Big Data Parallelisms. Advances in Intelligent Systems and Computing, 2016, , 209-217.	0.5	6
5323	An efficient MapReduce algorithm for similarity join in metric spaces. Journal of Supercomputing, 2016, 72, 1179-1200.	2.4	6
5324	Slot Utilization and Performance Improvement in Hadoop Cluster. Advances in Intelligent Systems and Computing, 2016, , 49-62.	0.5	2
5325	Constrained-meta-path-based ranking in heterogeneous information network. Knowledge and Information Systems, 2016, 49, 719-747.	2.1	21
5327	Cost-aware DAG scheduling algorithms for minimizing execution cost on cloud resources. Journal of Supercomputing, 2016, 72, 985-1012.	2.4	48
5328	Highway traffic accident prediction using VDS big data analysis. Journal of Supercomputing, 2016, 72, 2815-2831.	2.4	56
5329	An expandable web-based platform for visually analyzing basin-scale hydro-climate time series data. Environmental Modelling and Software, 2016, 78, 97-105.	1.9	17
5330	An Analysis of Public Clouds Elasticity in the Execution of Scientific Applications: a Survey. Journal of Grid Computing, 2016, 14, 193-216.	2.5	51
5331	Scalable clustering algorithms for continuous environmental flow cytometry. Bioinformatics, 2016, 32, 417-423.	1.8	15
5332	Simulation Algorithms With Exponential Integration for Time-Domain Analysis of Large-Scale Power Delivery Networks. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2016, 35, 1681-1694.	1.9	19
5333	Geospatial Big Data for Environmental and Agricultural Applications. , 2016, , 353-390.		16
5334	Analysis of microarray leukemia data using an efficient MapReduce-based K-nearest-neighbor classifier. Journal of Biomedical Informatics, 2016, 60, 395-409.	2.5	23
5335	A Hybrid Data Center Architecture for Big Data. Big Data Research, 2016, 3, 29-40.	2.6	21
5336	Space Shuffle: A Scalable, Flexible, and High-Performance Data Center Network. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 3351-3365.	4.0	21
5337	Scalable P2P bot detection system based on network data stream. Peer-to-Peer Networking and Applications, 2016, 9, 1209-1225.	2.6	11

#	ARTICLE	IF	CITATIONS
5338	A distributed spatial-temporal weighted model on MapReduce for short-term traffic flow forecasting. <i>Neurocomputing</i> , 2016, 179, 246-263.	3.5	148
5339	Towards Practical and Near-Optimal Coflow Scheduling for Data Center Networks. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2016, 27, 3366-3380.	4.0	54
5340	Dynamic Resource Allocation for MapReduce with Partitioning Skew. <i>IEEE Transactions on Computers</i> , 2016, 65, 3304-3317.	2.4	19
5341	A MapReduce framework to improve template matching uncertainty. , 2016, , .		5
5342	Handling big data: research challenges and future directions. <i>Journal of Supercomputing</i> , 2016, 72, 1494-1516.	2.4	103
5343	Personalized Recommendation System on Hadoop and HBase. <i>Communications in Computer and Information Science</i> , 2016, , 34-45.	0.4	1
5344	Declarative coordination of graph-based parallel programs. , 2016, , .		4
5345	External sorting on flash storage: reducing cell wearing and increasing efficiency by avoiding intermediate writes. <i>VLDB Journal</i> , 2016, 25, 495-518.	2.7	7
5346	Rapid Screening of Big Data Against Inadvertent Leaks. , 2016, , 193-235.		1
5347	An Efficient High-Dimensional Big Data Storage Structure Based on US-ELM. <i>Proceedings in Adaptation, Learning and Optimization</i> , 2016, , 489-500.	1.5	0
5348	A smart speculative execution strategy based on node classification for heterogeneous Hadoop systems. , 2016, , .		0
5349	Improving the performance of Apache Hadoop on pervasive environments through context-aware scheduling. <i>Journal of Ambient Intelligence and Humanized Computing</i> , 2016, 7, 333-345.	3.3	12
5350	Big data architecture for construction waste analytics (CWA): A conceptual framework. <i>Journal of Building Engineering</i> , 2016, 6, 144-156.	1.6	130
5351	Performance evaluation of MRDataCube for data cube computation algorithm using MapReduce. , 2016, , .		4
5352	Scheduling in Big Data Heterogeneous Distributed System Using Hadoop. <i>Advances in Intelligent Systems and Computing</i> , 2016, , 119-131.	0.5	1
5353	Pharmacy robotic dispensing and planogram analysis using association rule mining with prescription data. <i>Expert Systems With Applications</i> , 2016, 57, 296-310.	4.4	32
5354	Single-pass and linear-time k-means clustering based on MapReduce. <i>Information Systems</i> , 2016, 60, 1-12.	2.4	52
5355	Big data analytic architecture for intruder detection in heterogeneous wireless sensor networks. <i>Journal of Network and Computer Applications</i> , 2016, 66, 236-249.	5.8	23

#	ARTICLE	IF	CITATIONS
5357	Truth Discovery in Crowdsourced Detection of Spatial Events. IEEE Transactions on Knowledge and Data Engineering, 2016, 28, 1047-1060.	4.0	44
5358	The Edge Weight Computation with MapReduce for Extracting Weighted Graphs. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 3659-3672.	4.0	3
5359	Pattern mining: current status and emerging topics. Progress in Artificial Intelligence, 2016, 5, 165-170.	1.5	8
5360	Context-Aware Recommendation Model based on Mobile Application Analysis Platform. Multimedia Tools and Applications, 2016, 75, 14783-14794.	2.6	8
5361	Modelling with stakeholders â€œ Next generation. Environmental Modelling and Software, 2016, 77, 196-220.	1.9	405
5362	Analysis and evaluation of MapReduce solutions on an HPC cluster. Computers and Electrical Engineering, 2016, 50, 200-216.	3.0	15
5363	Implementing Randomized Matrix Algorithms in Parallel and Distributed Environments. Proceedings of the IEEE, 2016, 104, 58-92.	16.4	20
5365	Towards a data processing plane: An automata-based distributed dynamic data processing model. Future Generation Computer Systems, 2016, 59, 21-32.	4.9	3
5366	Parallel and Streaming Truth Discovery in Large-Scale Quantitative Crowdsourcing. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 2984-2997.	4.0	24
5367	Measure-based diffusion grid construction and high-dimensional data discretization. Applied and Computational Harmonic Analysis, 2016, 40, 207-228.	1.1	3
5368	Data locality in MapReduce: A network perspective. Performance Evaluation, 2016, 96, 1-11.	0.9	15
5369	An Unlicensed Taxi Identification Model Based on Big Data Analysis. IEEE Transactions on Intelligent Transportation Systems, 2016, 17, 1703-1713.	4.7	80
5370	Big Data Analysis for Media Production. Proceedings of the IEEE, 2016, 104, 2085-2113.	16.4	6
5371	A-ELMâž: Adaptive Distributed Extreme Learning Machine with MapReduce. Neurocomputing, 2016, 174, 368-374.	3.5	22
5372	A scalable approach for detecting plagiarized mobile applications. Knowledge and Information Systems, 2016, 49, 143-169.	2.1	3
5373	CEPSim: Modelling and simulation of Complex Event Processing systems in cloud environments. Future Generation Computer Systems, 2016, 65, 122-139.	4.9	37
5374	High-dimentional data authentication protocol based on hash chain for Hadoop systems. Cluster Computing, 2016, 19, 475-484.	3.5	4
5375	Toward integrity assurance of outsourced computing â€œ a game theoretic perspective. Future Generation Computer Systems, 2016, 55, 87-100.	4.9	8

#	ARTICLE	IF	CITATIONS
5376	Personalized Prediction and Sparsity Pursuit in Latent Factor Models. Journal of the American Statistical Association, 2016, 111, 241-252.	1.8	21
5377	A Coarse-Grained Reconfigurable Architecture for Compute-Intensive MapReduce Acceleration. IEEE Computer Architecture Letters, 2016, 15, 69-72.	1.0	6
5378	HadoopCL2: Motivating the Design of a Distributed, Heterogeneous Programming System With Machine-Learning Applications. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 762-775.	4.0	11
5379	<sc>Heads-Join</sc>: Efficient Earth Mover's Distance Similarity Joins on Hadoop. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1660-1673.	4.0	15
5380	Efficient Parallel Skyline Evaluation Using MapReduce. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1996-2009.	4.0	36
5381	Towards Exploring Data-Intensive Scientific Applications at Extreme Scales through Systems and Simulations. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1824-1837.	4.0	21
5382	Organisational sustainability modellingâ€”An emerging service and analytics model for evaluating Cloud Computing adoption with two case studies. International Journal of Information Management, 2016, 36, 167-179.	10.5	63
5383	How computer science can help in understanding the 3D genome architecture. Briefings in Bioinformatics, 2016, 17, 733-744.	3.2	10
5384	Enhancing the Performance of MapReduce Default Scheduler by Detecting Prolonged TaskTrackers in Heterogeneous Environments. Advances in Intelligent Systems and Computing, 2016, , 225-233.	0.5	0
5385	Energy efficiency aware load distribution and electricity cost volatility control for cloud service providers. Journal of Network and Computer Applications, 2016, 59, 185-197.	5.8	28
5386	DCloud: Deadline-Aware Resource Allocation for Cloud Computing Jobs. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 2248-2260.	4.0	44
5387	MapReduce Parallel Programming Model: A State-of-the-Art Survey. International Journal of Parallel Programming, 2016, 44, 832-866.	1.1	50
5388	A MapReduce solution for associative classification of big data. Information Sciences, 2016, 332, 33-55.	4.0	90
5389	Fast Compression of Large Semantic Web Data Using X10. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 2603-2617.	4.0	9
5390	Data Center Energy Consumption Modeling: A Survey. IEEE Communications Surveys and Tutorials, 2016, 18, 732-794.	24.8	695
5391	A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys and Tutorials, 2016, 18, 1153-1176.	24.8	1,775
5392	Cloud-Based Information Infrastructure for Next-Generation Power Grid: Conception, Architecture, and Applications. IEEE Transactions on Smart Grid, 2016, 7, 1896-1912.	6.2	77
5394	Common Pitfalls of Benchmarking Big Data Systems. IEEE Transactions on Services Computing, 2016, 9, 152-160.	3.2	7

#	ARTICLE	IF	CITATIONS
5395	PRS: Parallel Relaxation Simulation for Massive Graphs. Computer Journal, 2016, 59, 848-860.	1.5	13
5396	Big Data Essentials. Integrated Series on Information Systems, 2016, , 17-29.	0.1	1
5397	A Survey of Different Technologies and Recent Challenges of Big Data. Smart Innovation, Systems and Technologies, 2016, , 537-548.	0.5	9
5398	Unsupervised Topic Modeling in a Large Free Text Radiology Report Repository. Journal of Digital Imaging, 2016, 29, 59-62.	1.6	32
5399	Hybrid Job-Driven Scheduling for Virtual MapReduce Clusters. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1687-1699.	4.0	25
5402	Pull-based recommendations in mobile environments. Computer Standards and Interfaces, 2016, 44, 185-204.	3.8	23
5403	Fast and scalable vector similarity joins with MapReduce. Journal of Intelligent Information Systems, 2016, 46, 473-497.	2.8	3
5404	Social big data: Recent achievements and new challenges. Information Fusion, 2016, 28, 45-59.	11.7	629
5405	PR-ELM: Parallel regularized extreme learning machine based on cluster. Neurocomputing, 2016, 173, 1073-1081.	3.5	39
5406	Parallel ensemble of online sequential extreme learning machine based on MapReduce. Neurocomputing, 2016, 174, 352-367.	3.5	39
5407	Big-Data Approach for Three-Dimensional Building Extraction from Aerial Laser Scanning. Journal of Computing in Civil Engineering, 2016, 30, .	2.5	21
5408	Big Data Privacy: Challenges to Privacy Principles and Models. Data Science and Engineering, 2016, 1, 21-28.	4.6	76
5409	A Crowdsourcing Worker Quality Evaluation Algorithm on MapReduce for Big Data Applications. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1879-1888.	4.0	23
5410	MBFS: a parallel metadata search method based on Bloomfilters using MapReduce for large-scale file systems. Journal of Supercomputing, 2016, 72, 3006-3032.	2.4	6
5411	Systematic Data Placement Optimization in Multi-Cloud Storage for Complex Requirements. IEEE Transactions on Computers, 2016, 65, 1964-1977.	2.4	36
5412	MapTask Scheduling in MapReduce With Data Locality: Throughput and Heavy-Traffic Optimality. IEEE/ACM Transactions on Networking, 2016, 24, 190-203.	2.6	130
5413	A grid density based framework for classifying streaming data in the presence of concept drift. Journal of Intelligent Information Systems, 2016, 46, 179-211.	2.8	34
5414	A general perspective of Big Data: applications, tools, challenges and trends. Journal of Supercomputing, 2016, 72, 3073-3113.	2.4	104

#	ARTICLE	IF	CITATIONS
5415	Will solid-state drives accelerate your bioinformatics? In-depth profiling, performance analysis and beyond. Briefings in Bioinformatics, 2016, 17, 713-727.	3.2	12
5416	Achieving Optimal Block Pipelining in Organized Network Coded Gossip. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 627-639.	4.0	0
5417	Dynamic Job Ordering and Slot Configurations for MapReduce Workloads. IEEE Transactions on Services Computing, 2016, 9, 4-17.	3.2	26
5418	Revisiting TCP Congestion Control in a Virtual Cluster Environment. IEEE/ACM Transactions on Networking, 2016, 24, 2154-2167.	2.6	2
5419	P-DOT: a model of computation for big data. International Journal of Parallel, Emergent and Distributed Systems, 2016, 31, 233-253.	0.7	0
5420	Enabling Mobile Cloud Wide Spread Through an Evolutionary Market-Based Approach. IEEE Systems Journal, 2016, 10, 839-846.	2.9	2
5421	Predictive analytics using statistical, learning, and ensemble methods to support real-time exploration of discrete event simulations. Future Generation Computer Systems, 2016, 56, 360-374.	4.9	28
5422	Heuristics for periodical batch job scheduling in a MapReduce computing framework. Information Sciences, 2016, 326, 119-133.	4.0	30
5423	Novel heuristic speculative execution strategies in heterogeneous distributed environments. Computers and Electrical Engineering, 2016, 50, 166-179.	3.0	16
5424	Power-efficient assignment of virtual machines to physical machines. Future Generation Computer Systems, 2016, 54, 82-94.	4.9	27
5425	Skyline Discovery and Composition of Multi-Cloud Mashup Services. IEEE Transactions on Services Computing, 2016, 9, 72-83.	3.2	45
5426	FiDoop: Parallel Mining of Frequent Itemsets Using MapReduce. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016, 46, 313-325.	5.9	90
5427	Processing Cassandra Datasets with Hadoop-Streaming Based Approaches. IEEE Transactions on Services Computing, 2016, 9, 46-58.	3.2	29
5428	Parallel community detection on large graphs with MapReduce and GraphChi. Data and Knowledge Engineering, 2016, 104, 17-31.	2.1	26
5429	End-to-End Optimization for Geo-Distributed MapReduce. IEEE Transactions on Cloud Computing, 2016, 4, 293-306.	3.1	44
5430	Governing energy consumption in Hadoop through CPU frequency scaling: An analysis. Future Generation Computer Systems, 2016, 54, 219-232.	4.9	56
5431	An efficient MapReduce-based rule matching method for production system. Future Generation Computer Systems, 2016, 54, 478-489.	4.9	7
5432	Hadoop Performance Modeling for Job Estimation and Resource Provisioning. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 441-454.	4.0	85

#	ARTICLE	IF	CITATIONS
5433	Analyzing Hadoop power consumption and impact on application QoS. <i>Future Generation Computer Systems</i> , 2016, 55, 213-223.	4.9	30
5434	Scheduling algorithm based on prefetching in MapReduce clusters. <i>Applied Soft Computing Journal</i> , 2016, 38, 1109-1118.	4.1	22
5435	Spatiotemporal Segmentation of Metro Trips Using Smart Card Data. <i>IEEE Transactions on Vehicular Technology</i> , 2016, 65, 1137-1149.	3.9	36
5436	Fast, Ad Hoc Query Evaluations over Multidimensional Geospatial Datasets. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 28-42.	3.1	16
5437	Self-Adjusting Slot Configurations for Homogeneous and Heterogeneous Hadoop Clusters. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 344-357.	3.1	26
5438	Designing Self-Tuning Split-Map-Merge Applications for High Cost-Efficiency in the Cloud. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 303-316.	3.1	10
5439	Developing a decision support system for improving sustainability performance of manufacturing processes. <i>Journal of Intelligent Manufacturing</i> , 2017, 28, 1421-1440.	4.4	24
5440	DyScale: A MapReduce Job Scheduler for Heterogeneous Multicore Processors. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 317-330.	3.1	9
5441	Orchestrating Bulk Data Transfers across Geo-Distributed Datacenters. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 112-125.	3.1	67
5442	HFSP: Bringing Size-Based Scheduling To Hadoop. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 43-56.	3.1	22
5443	BDCaM: Big Data for Context-Aware Monitoring – A Personalized Knowledge Discovery Framework for Assisted Healthcare. <i>IEEE Transactions on Cloud Computing</i> , 2017, 5, 628-641.	3.1	74
5444	The classification of imbalanced large data sets based on MapReduce and ensemble of ELM classifiers. <i>International Journal of Machine Learning and Cybernetics</i> , 2017, 8, 1009-1017.	2.3	87
5445	Priority Scheduling for Heterogeneous Workloads: Tradeoff Between Evictions and Response Time. <i>IEEE Systems Journal</i> , 2017, 11, 684-695.	2.9	1
5446	BPLG – BMCS: GPU-sorting algorithm using a tuning skeleton library. <i>Journal of Supercomputing</i> , 2017, 73, 4-16.	2.4	2
5447	Enumerating Maximal Bicliques from a Large Graph Using MapReduce. <i>IEEE Transactions on Services Computing</i> , 2017, 10, 771-784.	3.2	31
5448	MedGraph: a graph-based representation and computation to handle large sets of images. <i>Multimedia Tools and Applications</i> , 2017, 76, 2769-2785.	2.6	4
5449	Data-Aware Task Dispatching for Batch Queuing System. <i>IEEE Systems Journal</i> , 2017, 11, 889-897.	2.9	5
5450	Recent Development in Big Data Analytics for Business Operations and Risk Management. <i>IEEE Transactions on Cybernetics</i> , 2017, 47, 81-92.	6.2	177

#	ARTICLE	IF	CITATIONS
5451	An Infrastructure Service Recommendation System for Cloud Applications with Real-time QoS Requirement Constraints. <i>IEEE Systems Journal</i> , 2017, 11, 2960-2970.	2.9	22
5452	A Dynamic Rule Creation Based Anomaly Detection Method for Identifying Security Breaches in Log Records. <i>Wireless Personal Communications</i> , 2017, 94, 497-511.	1.8	32
5453	Speeding up the multimedia feature extraction: a comparative study on the big data approach. <i>Multimedia Tools and Applications</i> , 2017, 76, 7497-7517.	2.6	7
5454	Service Pattern: An Integrated Business Process Model for Modern Service Industry. <i>IEEE Transactions on Services Computing</i> , 2017, 10, 841-853.	3.2	14
5455	Enabling fast failure recovery in shared Hadoop clusters: Towards failure-aware scheduling. <i>Future Generation Computer Systems</i> , 2017, 74, 208-219.	4.9	30
5456	SLA-aware energy-efficient scheduling scheme for Hadoop YARN. <i>Journal of Supercomputing</i> , 2017, 73, 3526-3546.	2.4	29
5457	An FPGA-based Integrated MapReduce Accelerator Platform. <i>Journal of Signal Processing Systems</i> , 2017, 87, 357-369.	1.4	8
5458	The Parallelization of Back Propagation Neural Network in MapReduce and Spark. <i>International Journal of Parallel Programming</i> , 2017, 45, 760-779.	1.1	31
5459	SurvSurf: human retrieval on large surveillance video data. <i>Multimedia Tools and Applications</i> , 2017, 76, 6521-6549.	2.6	13
5460	A Two-Way Street: Green Big Data Processing for a Greener Smart Grid. <i>IEEE Systems Journal</i> , 2017, 11, 784-795.	2.9	46
5461	Cloud Computing-Based Time Series Analysis for Structural Damage Detection. <i>Journal of Engineering Mechanics - ASCE</i> , 2017, 143, .	1.6	20
5462	A parallel algorithm for mining constrained frequent patterns using MapReduce. <i>Soft Computing</i> , 2017, 21, 2237-2249.	2.1	12
5463	Efficient Role Mining for Context-Aware Service Recommendation Using a High-Performance Cluster. <i>IEEE Transactions on Services Computing</i> , 2017, 10, 914-926.	3.2	17
5464	Short Read Mapping: An Algorithmic Tour. <i>Proceedings of the IEEE</i> , 2017, 105, 436-458.	16.4	63
5465	Online Learning-Based Server Provisioning for Electricity Cost Reduction in Data Center. <i>IEEE Transactions on Control Systems Technology</i> , 2017, 25, 1044-1051.	3.2	15
5466	kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data. <i>Knowledge-Based Systems</i> , 2017, 117, 3-15.	4.0	216
5467	Dependable large scale behavioral patterns mining from sensor data using Hadoop platform. <i>Information Sciences</i> , 2017, 379, 128-145.	4.0	28
5468	Using adaptive resource allocation to implement an elastic MapReduce framework. <i>Software - Practice and Experience</i> , 2017, 47, 349-360.	2.5	1

#	ARTICLE	IF	CITATIONS
5469	Cache-conscious run-time decomposition of data parallel computations. <i>Journal of Supercomputing</i> , 2017, 73, 3366-3390.	2.4	1
5470	Large-scale biological meta-database management. <i>Future Generation Computer Systems</i> , 2017, 67, 481-489.	4.9	7
5471	Fuzzy rule based classification systems for big data with MapReduce: granularity analysis. <i>Advances in Data Analysis and Classification</i> , 2017, 11, 711-730.	0.9	33
5472	Scaling machine learning for target prediction in drug discovery using Apache Spark. <i>Future Generation Computer Systems</i> , 2017, 67, 409-417.	4.9	39
5473	Nebo: An efficient, parallel, and portable domain-specific language for numerically solving partial differential equations. <i>Journal of Systems and Software</i> , 2017, 125, 389-400.	3.3	9
5474	Continuous validation of performance test workloads. <i>Automated Software Engineering</i> , 2017, 24, 189-231.	2.2	17
5475	Towards Systematic Parallelization of Graph Transformations Over Pregel. <i>International Journal of Parallel Programming</i> , 2017, 45, 320-339.	1.1	4
5476	Proposal reviewer recommendation system based on big data for a national research management institute. <i>Journal of Information Science</i> , 2017, 43, 147-158.	2.0	13
5477	Analysis of power consumption in heterogeneous virtual machine environments. <i>Soft Computing</i> , 2017, 21, 4531-4542.	2.1	13
5478	Adaptive flow scheduling for modular datacenter networks. <i>Peer-to-Peer Networking and Applications</i> , 2017, 10, 1142-1151.	2.6	0
5479	FiDooP-DP: Data Partitioning in Frequent Itemset Mining on Hadoop Clusters. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2017, 28, 101-114.	4.0	52
5480	A Mobility Analytical Framework for Big Mobile Data in Densely Populated Area. <i>IEEE Transactions on Vehicular Technology</i> , 2017, 66, 1443-1455.	3.9	78
5481	Collective Sensing Platforms. <i>Understanding Complex Systems</i> , 2017, , 115-133.	0.3	1
5482	An efficient parallel method for batched OS-ELM training using MapReduce. <i>Memetic Computing</i> , 2017, 9, 183-197.	2.7	6
5483	UFalloc: Towards Utility Max-min Fairness of Bandwidth Allocation for Applications in Datacenter Networks. <i>Mobile Networks and Applications</i> , 2017, 22, 161-173.	2.2	6
5484	Applications for Environmental Sensing in EveryAware. <i>Understanding Complex Systems</i> , 2017, , 135-155.	0.3	0
5485	A systematic review and comparative analysis of cross-document coreference resolution methods and tools. <i>Computing (Vienna/New York)</i> , 2017, 99, 313-349.	3.2	32
5486	A highly scalable parallel algorithm for maximally informative k-itemset mining. <i>Knowledge and Information Systems</i> , 2017, 50, 1-26.	2.1	12

#	ARTICLE	IF	CITATIONS
5487	Classification with boosting of extreme learning machine over arbitrarily partitioned data. <i>Soft Computing</i> , 2017, 21, 2269-2281.	2.1	16
5488	SuperMIC: Analyzing Large Biological Datasets in Bioinformatics with Maximal Information Coefficient. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2017, 14, 783-795.	1.9	7
5489	Leaky Buffer: A Novel Abstraction for Relieving Memory Pressure from Cluster Data Processing Frameworks. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2017, 28, 128-140.	4.0	9
5490	ALDCNN: a distributed crossing semantic gap learning for person re-identification. <i>Concurrency Computation Practice and Experience</i> , 2017, 29, e3766.	1.4	8
5491	Balancing push and pull in Confuga, an active storage cluster file system for scientific workflows. <i>Concurrency Computation Practice and Experience</i> , 2017, 29, e3834.	1.4	0
5492	iShuffle: Improving Hadoop Performance with Shuffle-on-Write. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2017, 28, 1649-1662.	4.0	60
5493	Small files storing and computing optimization in Hadoop parallel rendering. <i>Concurrency Computation Practice and Experience</i> , 2017, 29, e3847.	1.4	6
5494	A dynamic and parallel approach for repetitive prime labeling of XML with MapReduce. <i>Journal of Supercomputing</i> , 2017, 73, 810-836.	2.4	4
5495	An efficient approximation to the K-means clustering for massive data. <i>Knowledge-Based Systems</i> , 2017, 117, 56-69.	4.0	163
5496	A topology-based scaling mechanism for Apache Storm. <i>International Journal of Network Management</i> , 2017, 27, e1933.	1.4	10
5497	DASC: data aware algorithm for scalable clustering. <i>Knowledge and Information Systems</i> , 2017, 50, 851-881.	2.1	2
5498	gSched: a resource aware Hadoop scheduler for heterogeneous cloud computing environments. <i>Concurrency Computation Practice and Experience</i> , 2017, 29, e3841.	1.4	7
5499	Efficient Processing of Skyline Queries Using MapReduce. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2017, 29, 1031-1044.	4.0	23
5500	Evaluating Scalable Distributed Erlang for Scalability and Reliability. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2017, 28, 2244-2257.	4.0	11
5501	Disk-based shortest path discovery using distance index over large dynamic graphs. <i>Information Sciences</i> , 2017, 382-383, 201-215.	4.0	9
5502	Mobile web and cloud services enabling Internet of Things. <i>CSI Transactions on ICT</i> , 2017, 5, 109-117.	0.7	19
5503	Copernicus, a hybrid dataflow and peer-to-peer scientific computing platform for efficient large-scale ensemble sampling. <i>Future Generation Computer Systems</i> , 2017, 71, 18-31.	4.9	15
5504	ViSiBiD: A learning model for early discovery and real-time prediction of severe clinical events using vital signs as big data. <i>Computer Networks</i> , 2017, 113, 244-257.	3.2	63

#	ARTICLE	IF	CITATIONS
5505	Probabilistic modeling to achieve load balancing in Expert Clouds. <i>Ad Hoc Networks</i> , 2017, 59, 12-23.	3.4	27
5506	High-performance parallel approaches for three-dimensional light detection and ranging point clouds gridding. <i>Journal of Applied Remote Sensing</i> , 2017, 11, 016011.	0.6	3
5507	Next-generation sequencing: big data meets high performance computing. <i>Drug Discovery Today</i> , 2017, 22, 712-717.	3.2	108
5508	A similarity query system for road traffic data based on a NoSQL document store. <i>Journal of Systems and Software</i> , 2017, 127, 28-51.	3.3	5
5509	Coordinated Multi-Point Clustering Schemes: A Survey. <i>IEEE Communications Surveys and Tutorials</i> , 2017, 19, 743-764.	24.8	128
5510	SparkBench: a spark benchmarking suite characterizing large-scale in-memory data analytics. <i>Cluster Computing</i> , 2017, 20, 2575-2589.	3.5	26
5511	Parallel programming with pictures is a Snap!. <i>Journal of Parallel and Distributed Computing</i> , 2017, 105, 150-162.	2.7	19
5512	Performance Improvement of MapReduce for Heterogeneous Clusters Based on Efficient Locality and Replica Aware Scheduling (ELRAS) Strategy. <i>Wireless Personal Communications</i> , 2017, 95, 2709-2733.	1.8	10
5513	A Multi-agent System Approach to Load-Balancing and Resource Allocation for Distributed Computing. <i>Springer Proceedings in Complexity</i> , 2017, , 41-54.	0.2	25
5514	A Set of Methods to Support Object-Based Distributed Analysis of Large Volumes of Earth Observation Data. <i>IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing</i> , 2017, 10, 681-690.	2.3	8
5515	OpenFlow-Based Scalable Routing With Hybrid Addressing in Data Center Networks. <i>IEEE Communications Letters</i> , 2017, 21, 1047-1050.	2.5	8
5516	Distributed Parallel PCA for Modeling and Monitoring of Large-Scale Plant-Wide Processes With Big Data. <i>IEEE Transactions on Industrial Informatics</i> , 2017, 13, 1877-1885.	7.2	227
5517	Fast Connected Components Computation in Large Graphs by Vertex Pruning. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2017, 28, 760-773.	4.0	22
5518	Partitioning dynamic graph asynchronously with distributed FENNEL. <i>Future Generation Computer Systems</i> , 2017, 71, 32-42.	4.9	9
5520	The Parvicursor infrastructure to facilitate the design of Grid and Cloud computing systems. <i>Computing (Vienna/New York)</i> , 2017, 99, 979-1006.	3.2	3
5521	Analyzing and Enhancing Dynamic Threshold Policy of Data Center Switches. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2017, 28, 2454-2470.	4.0	15
5522	Designing parallel data processing for enabling large-scale sensor applications. <i>Personal and Ubiquitous Computing</i> , 2017, 21, 457-473.	1.9	6
5523	Handling multi-dimensional complex queries in key-value data stores. <i>Information Systems</i> , 2017, 66, 82-96.	2.4	6

#	ARTICLE	IF	CITATIONS
5524	Social Crowdsourcing to Friends: An Incentive Mechanism for Multi-Resource Sharing. IEEE Journal on Selected Areas in Communications, 2017, 35, 795-808.	9.7	17
5525	High-Resolution Multi-decadal Simulation of Tropical Cyclones. , 2017, , 187-211.		9
5526	Flow Stealer: lightweight load balancing by stealing flows in distributed SDN controllers. Science China Information Sciences, 2017, 60, 1.	2.7	30
5527	A characterization of workflow management systems for extreme-scale applications. Future Generation Computer Systems, 2017, 75, 228-238.	4.9	100
5528	CBMR: An optimized MapReduce for item-based collaborative filtering recommendation algorithm with empirical analysis. Concurrency Computation Practice and Experience, 2017, 29, e4092.	1.4	21
5529	Modeling and Verifying HDFS Using Process Algebra. Mobile Networks and Applications, 2017, 22, 318-331.	2.2	4
5530	HB-File: An efficient and effective high-dimensional big data storage structure based on US-ELM. Neurocomputing, 2017, 261, 184-192.	3.5	5
5531	Assessing sensitivities in algorithmic detection of tropical cyclones in climate data. Geophysical Research Letters, 2017, 44, 1141-1149.	1.5	67
5532	FSQCN: Fast and simple quantized congestion notification in data center ethernet. Journal of Network and Computer Applications, 2017, 83, 53-62.	5.8	12
5533	An evolving Takagi-Sugeno model based on aggregated trapezium clouds for anomaly detection in large datasets. Journal of Intelligent and Fuzzy Systems, 2017, 32, 2295-2308.	0.8	14
5534	Design and evaluation of small-large outer joins in cloud computing environments. Journal of Parallel and Distributed Computing, 2017, 110, 2-15.	2.7	25
5535	Exploiting NoSQL Graph Databases and in Memory Architectures for Extracting Graph Structural Data Summaries. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2017, 25, 81-109.	0.9	9
5536	An efficient spark-based adaptive windowing for entity matching. Journal of Systems and Software, 2017, 128, 1-10.	3.3	18
5537	A parallel C4.5 decision tree algorithm based on MapReduce. Concurrency Computation Practice and Experience, 2017, 29, e4015.	1.4	23
5538	One step at a time: Parallelism in an introductory programming course. Journal of Parallel and Distributed Computing, 2017, 105, 4-17.	2.7	18
5539	A parallel approximate SS-ELM algorithm based on MapReduce for large-scale datasets. Journal of Parallel and Distributed Computing, 2017, 108, 85-94.	2.7	12
5540	Exploiting Efficient Densest Subgraph Discovering Methods for Big Data. IEEE Transactions on Big Data, 2017, 3, 334-348.	4.4	4
5541	A parallel fuzzy clustering algorithm for large graphs using Pregel. Expert Systems With Applications, 2017, 78, 135-144.	4.4	24

#	ARTICLE	IF	CITATIONS
5543	On a cyberinfrastructure platform for multidisciplinary, data-intensive scientific research. , 2017, , .		0
5544	A Multicore Path to Connectomics-on-Demand. , 2017, , .		8
5545	Big Data Programming Models. , 2017, , 31-63.		10
5546	Programming Platforms for Big Data Analysis. , 2017, , 65-99.		2
5547	Large-Scale Data Stream Processing Systems. , 2017, , 219-260.		6
5548	BigDimETL: ETL for Multidimensional Big Data. Advances in Intelligent Systems and Computing, 2017, , 935-944.	0.5	4
5549	Emerging Cost Effective Big Data Architectures. , 2017, , 755-776.		1
5550	Toward intrusion detection using belief decision trees for big data. Knowledge and Information Systems, 2017, 53, 671-698.	2.1	15
5551	Wearable Medical Sensor-Based System Design: A Survey. IEEE Transactions on Multi-Scale Computing Systems, 2017, 3, 124-138.	2.5	104
5552	An insight into imbalanced Big Data classification: outcomes and challenges. Complex & Intelligent Systems, 2017, 3, 105-120.	4.0	153
5553	Enabling data science in the Gaia mission archive: The present-day mass function and age distribution. Astronomy and Computing, 2017, 19, 1-15.	0.8	3
5554	Reasoning on data partitioning for single-round multi-join evaluation in massively parallel systems. Communications of the ACM, 2017, 60, 93-100.	3.3	5
5555	A Distributed Stream Library for Java 8. IEEE Transactions on Big Data, 2017, 3, 262-275.	4.4	7
5556	Polymorphic malware detection using sequence classification methods and ensembles. Eurasip Journal on Information Security, 2017, 2017, .	2.2	39
5557	Competitive algorithms for multistage online scheduling. European Journal of Operational Research, 2017, 260, 468-481.	3.5	6
5558	Stream-based live data replication approach of in-memory cache. Concurrency Computation Practice and Experience, 2017, 29, e4052.	1.4	10
5559	Exhaustive search algorithms to mine subgroups on Big Data using Apache Spark. Progress in Artificial Intelligence, 2017, 6, 145-158.	1.5	11
5560	HFIM: a Spark-based hybrid frequent itemset mining algorithm for big data processing. Journal of Supercomputing, 2017, 73, 3652-3668.	2.4	50

#	ARTICLE	IF	CITATIONS
5561	Distributed extreme learning machine with alternating direction method of multiplier. <i>Neurocomputing</i> , 2017, 261, 164-170.	3.5	25
5562	A Faster Algorithm for Enumerating Connected Convex Subgraphs in Acyclic Digraphs. <i>IEEE Embedded Systems Letters</i> , 2017, 9, 9-12.	1.3	3
5563	A framework for distributed nearest neighbor classification using Hadoop. <i>Journal of Computational Methods in Sciences and Engineering</i> , 2017, 17, S11-S19.	0.1	2
5564	A parallel k -means clustering algorithm based on redundance elimination and extreme points optimization employing MapReduce. <i>Concurrency Computation Practice and Experience</i> , 2017, 29, e4109.	1.4	17
5565	Overcrowding detection in indoor events using scalable technologies. <i>Personal and Ubiquitous Computing</i> , 2017, 21, 507-519.	1.9	6
5566	Bioinformatics from a Big Data Perspective: Meeting the Challenge. <i>Lecture Notes in Computer Science</i> , 2017, , 349-359.	1.0	0
5567	Efficient MapReduce Matrix Multiplication with Optimized Mapper Set. <i>Advances in Intelligent Systems and Computing</i> , 2017, , 186-196.	0.5	5
5568	An Apache Spark Implementation for Sentiment Analysis on Twitter Data. <i>Lecture Notes in Computer Science</i> , 2017, , 15-25.	1.0	20
5569	Dynamic algorithm modeling application for comparing distributed computing frameworks. <i>CSI Transactions on ICT</i> , 2017, 5, 341-359.	0.7	0
5570	GB-CENT. , 2017, , .		18
5571	Energy-Efficient dynamic virtual network traffic engineering for north-south traffic in multi-location data center networks. <i>Computer Networks</i> , 2017, 125, 90-102.	3.2	15
5572	Prediction of Population Health Indices from Social Media using Kernel-based Textual and Temporal Features. , 2017, , .		14
5573	Atrak: a MapReduce-based data warehouse for big data. <i>Journal of Supercomputing</i> , 2017, 73, 4596-4610.	2.4	8
5574	What Can the Big Data Eco-System and Data Analytics Do for E-Health? A Smooth Review Study. <i>Lecture Notes in Computer Science</i> , 2017, , 629-641.	1.0	0
5575	Business Intelligence and Analytics: Big Systems for Big Data. , 2017, , 7-49.		1
5576	Statistical Analysis of Latency Through Semantic Profiling. , 2017, , .		15
5577	A Survey on Network Methodologies for Real-Time Analytics of Massive IoT Data and Open Research Issues. <i>IEEE Communications Surveys and Tutorials</i> , 2017, 19, 1457-1477.	24.8	300
5578	Coding for Distributed Fog Computing. , 2017, 55, 34-40.		106

#	ARTICLE	IF	CITATIONS
5579	Using the Launcher for Executing High Throughput Workloads. Big Data Research, 2017, 8, 57-64.	2.6	0
5580	Parallel computation of the reachability graph of petri net models with semantic information. Software - Practice and Experience, 2017, 47, 647-668.	2.5	2
5581	Scalable MapReduce-based Fuzzy Min-Max Neural Network for Pattern Classification. , 2017, , .		4
5582	A Survey on Big Data and Collective Intelligence. Lecture Notes in Computer Science, 2017, , 169-181.	1.0	1
5583	Performance modeling of cloud apps using message queueing as a service (MaaS). , 2017, , .		2
5584	Data center traffic engineering using Markov approximation. , 2017, , .		0
5585	Browser-based Harnessing of Voluntary Computational Power. Foundations of Computing and Decision Sciences, 2017, 42, 3-42.	0.5	12
5586	Query Processing Techniques for Big Spatial-Keyword Data. , 2017, , .		14
5587	Cloud Engineering Principles and Technology Enablers for Medical Image Processing-as-a-Service. , 2017, 2017, 127-137.		7
5588	Performance evaluation of bloom filter size in map-side and reduce-side bloom joins. , 2017, , .		2
5589	MapReduce in Computational Biology - A Synopsis. Communications in Computer and Information Science, 2017, , 53-64.	0.4	3
5590	MR-DIS: democratic instance selection for big data by MapReduce. Progress in Artificial Intelligence, 2017, 6, 211-219.	1.5	24
5591	In Defense of MapReduce. IEEE Internet Computing, 2017, 21, 94-98.	3.2	2
5592	MapReduce for accurate error correction of next-generation sequencing data. Bioinformatics, 2017, 33, 3844-3851.	1.8	15
5593	Data transformation as a means towards dynamic data storage and polyglot persistence. International Journal of Network Management, 2017, 27, e1976.	1.4	5
5594	Starvation Avoidance with Deadline Constraints in Hadoop Environments. , 2017, , .		0
5595	Defining the execution semantics of stream processing engines. Journal of Big Data, 2017, 4, .	6.9	14
5596	Turing's Theory of Morphogenesis: Where We Started, Where We Are and Where We Want to Go. Theory and Applications of Computability, 2017, , 219-235.	0.8	11

#	ARTICLE	IF	CITATIONS
5597	Recent advances in feature selection and its applications. Knowledge and Information Systems, 2017, 53, 551-577.	2.1	238
5598	Distributed snapshot maintenance in wide-column NoSQL databases using partitioned incremental ETL pipelines. Information Systems, 2017, 70, 48-58.	2.4	4
5599	Avoiding Slow Running Nodes in Distributed Systems. Lecture Notes in Networks and Systems, 2017, , 411-420.	0.5	3
5600	DECO3RUM: A Differential Evolution learning approach for generating compact Mamdani fuzzy rule-based models. Expert Systems With Applications, 2017, 83, 257-272.	4.4	24
5601	Large-Scale Encryption in the Hadoop Environment: Challenges and Solutions. IEEE Access, 2017, 5, 7156-7163.	2.6	33
5602	Deep net architectures for visual-based clothing image recognition on large database. Soft Computing, 2017, 21, 2923-2939.	2.1	12
5603	P3-Mobile. , 2017, , .		5
5604	Large-scale virtual screening on public cloud resources with Apache Spark. Journal of Cheminformatics, 2017, 9, 15.	2.8	16
5605	A General-Purpose Architecture for Replicated Metadata Services in Distributed File Systems. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 2747-2759.	4.0	3
5606	Memory-Centric Reconfigurable Accelerator for Classification and Machine Learning Applications. ACM Journal on Emerging Technologies in Computing Systems, 2017, 13, 1-24.	1.8	4
5607	DALiuGE: A graph execution framework for harnessing the astronomical data deluge. Astronomy and Computing, 2017, 20, 1-15.	0.8	18
5608	Efficiently predicting large-scale protein-protein interactions using MapReduce. Computational Biology and Chemistry, 2017, 69, 202-206.	1.1	13
5609	Dynamic Delay Predictions for Large-Scale Railway Networks: Deep and Shallow Extreme Learning Machines Tuned via Thresholdout. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47, 2754-2767.	5.9	72
5610	Concurrency bugs in open source software: a case study. Journal of Internet Services and Applications, 2017, 8, .	1.6	11
5611	Meta-synthesis of big data impacts on information systems development. Journal of Management Analytics, 2017, 4, 182-201.	1.6	7
5612	PPHOPCM: Privacy-Preserving High-Order Possibilistic c-Means Algorithm for Big Data Clustering with Cloud Computing. IEEE Transactions on Big Data, 2022, 8, 25-34.	4.4	85
5613	Efficient vCore Based Container Deployment Algorithm for Improving Heterogeneous Hadoop YARN Performance. Lecture Notes in Electrical Engineering, 2017, , 191-201.	0.3	2
5614	GPU-Accelerated Parallel Hierarchical Extreme Learning Machine on Flink for Big Data. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47, 2740-2753.	5.9	92

#	ARTICLE	IF	CITATIONS
5615	Parallel non-divergent flow accumulation for trillion cell digital elevation models on desktops or clusters. Environmental Modelling and Software, 2017, 92, 202-212.	1.9	20
5616	Distributed Compliance Monitoring of Business Processes over MapReduce Architectures. , 2017, , .		5
5617	Efficient Redundancy Techniques for Latency Reduction in Cloud Systems. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2017, 2, 1-30.	0.8	80
5618	Public Interest Analysis Based on Implicit Feedback of IPTV Users. IEEE Transactions on Industrial Informatics, 2017, 13, 2077-2086.	7.2	7
5619	<scp>Internet of People</scp>. Concurrency Computation Practice and Experience, 2017, 29, e4050.	1.4	5
5620	Simulation of MapReduce Across Geographically Distributed Datacentres Using CloudSim. Lecture Notes in Computer Science, 2017, , 70-81.	1.0	1
5621	Guest Editorial: The Parallel Storage, Processing and Analysis for Big Data. International Journal of Parallel Programming, 2017, 45, 731-733.	1.1	0
5622	The stellar transformation: From interconnection networks to datacenter networks. Computer Networks, 2017, 113, 29-45.	3.2	12
5623	UniAD. ACM Transactions on Database Systems, 2017, 42, 1-42.	1.5	2
5624	Large-scale parallel genome assembler over cloud computing environment. Journal of Bioinformatics and Computational Biology, 2017, 15, 1740003.	0.3	5
5625	Efficient GPU utilization in heterogeneous big data cluster using token-based scheduler. , 2017, , .		0
5626	Throughput optimization of TCP incast congestion control in large-scale datacenter networks. Computer Networks, 2017, 124, 46-60.	3.2	14
5627	OctopusFS. , 2017, , .		39
5628	Maxâ€flow minâ€cut algorithm with application to road networks. Concurrency Computation Practice and Experience, 2017, 29, e4099.	1.4	4
5629	Adding network bandwidth resource management to Hadoop YARN. , 2017, , .		4
5630	MoDNN: Local distributed mobile computing system for Deep Neural Network. , 2017, , .		168
5631	Integrating Map-Reduce and Stream-Processing for Efficiency (MRSP). Communications in Computer and Information Science, 2017, , 3-15.	0.4	1
5632	Smart Partitioning Mechanism for Dealing with Intermediate Data Skew in Reduce Task on Cloud Computing. , 2017, , .		5

#	ARTICLE	IF	CITATIONS
5633	Query Centric Partitioning and Allocation for Partially Replicated Database Systems. , 2017, , .		22
5634	aHDFS: An Erasure-Coded Data Archival System for Hadoop Clusters. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 3060-3073.	4.0	14
5635	Detecting Strong Ties Using Network Motifs. , 2017, , .		34
5637	A general and efficient divide-and-conquer algorithm framework for multi-core clusters. Cluster Computing, 2017, 20, 2605-2626.	3.5	6
5638	SMOTE-GPU: Big Data preprocessing on commodity hardware for imbalanced classification. Progress in Artificial Intelligence, 2017, 6, 347-354.	1.5	16
5639	A Hadoop-based Principle Component Analysis on embedded heterogeneous platform. , 2017, , .		1
5640	Topological analysis of data. EPJ Data Science, 2017, 6, .	1.5	59
5641	Fair workload distribution for multi-server systems with pulling strategies. Performance Evaluation, 2017, 113, 26-41.	0.9	5
5642	Exploiting Different Types of Parallelism in Distributed Analysis of Remote Sensing Data. IEEE Geoscience and Remote Sensing Letters, 2017, 14, 1298-1302.	1.4	9
5643	Nebula: Distributed Edge Cloud for Data Intensive Computing. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 3229-3242.	4.0	53
5644	Modeling and exploiting bibliographic Network for publication ranking. , 2017, , .		0
5645	A framework for scalable real-time anomaly detection over voluminous, geospatial data streams. Concurrency Computation Practice and Experience, 2017, 29, e4106.	1.4	8
5646	Top-k temporal keyword search over social media data. World Wide Web, 2017, 20, 1049-1069.	2.7	3
5647	Embracing social big data in wireless system design. Journal of Communications and Information Networks, 2017, 2, 81-96.	3.5	3
5648	MeshFS: A distributed file system for cloud-based wireless mesh network. Journal of Systems and Software, 2017, 131, 201-217.	3.3	1
5649	Efficient Implicit Parallel Patterns for Geographic Information System. Procedia Computer Science, 2017, 108, 545-554.	1.2	1
5650	Space-based parallel program design process with high-level communication channels. Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers, Series A/Chung-kuo Kung Ch'eng Hsuch K'an, 2017, 40, 347-354.	0.6	1
5651	Towards a scalable and energy-efficient resource manager for coupling cluster computing with distributed embedded computing. Cluster Computing, 2017, 20, 3707-3720.	3.5	1

#	ARTICLE	IF	CITATIONS
5652	Exploring correlation for fast skyline computation. Journal of Supercomputing, 2017, 73, 5071-5102.	2.4	2
5653	An online classification algorithm for large scale data streams: iNGSVM. Neurocomputing, 2017, 262, 67-76.	3.5	5
5654	An innovative GPS trajectory data based model for geographic recommendation service. Transactions in GIS, 2017, 21, 880-896.	1.0	13
5655	Leveraging Hardware Isolation for Process Level Access Control & Authentication. , 2017, , .		3
5656	Inferring software behavioral models with MapReduce. Science of Computer Programming, 2017, 145, 13-36.	1.5	8
5657	A scalable architecture for data-intensive natural language processing. Natural Language Engineering, 2017, 23, 709-731.	2.1	2
5658	MRSIM: Mitigating Reducer Skew In MapReduce. , 2017, , .		6
5659	Optimizing Spatial Queries in MapReduce. , 2017, , .		0
5660	Game theoretic decision making in IoT-assisted activity monitoring of defence personnel. Multimedia Tools and Applications, 2017, 76, 21911-21935.	2.6	16
5661	Parallel gene selection and dynamic ensemble pruning based on Affinity Propagation. Computers in Biology and Medicine, 2017, 87, 8-21.	3.9	7
5662	Improved routing algorithms in the dual-port datacenter networks HCN and BCN. Future Generation Computer Systems, 2017, 75, 58-71.	4.9	5
5663	Time Series Data Storage System Based on Big Data Technologies: Architecture and Implementation. International Journal of Emerging Electric Power Systems, 2017, 18, .	0.6	1
5664	Manufacturing Analytics and Industrial Internet of Things. IEEE Intelligent Systems, 2017, 32, 74-79.	4.0	185
5665	Cost-Aware Big Data Processing Across Geo-Distributed Datacenters. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 3114-3127.	4.0	43
5666	Feisu: Fast Query Execution over Heterogeneous Data Sources on Large-Scale Clusters. , 2017, , .		2
5667	From Clarity to Efficiency for Distributed Algorithms. ACM Transactions on Programming Languages and Systems, 2017, 39, 1-41.	1.7	22
5668	Dataclay: A distributed data store for effective inter-player data sharing. Journal of Systems and Software, 2017, 131, 129-145.	3.3	19
5669	A Parallel Algorithm for Large-Scale Nonconvex Penalized Quantile Regression. Journal of Computational and Graphical Statistics, 2017, 26, 935-939.	0.9	30

#	ARTICLE	IF	CITATIONS
5670	Blocking cost-driven multicast scheduling in fat-tree data center networks. Concurrency Computation Practice and Experience, 2017, 29, e4166.	1.4	2
5671	Improving the Quality of an R-Tree Using the Map-Reduce Framework. Lecture Notes in Electrical Engineering, 2017, , 164-170.	0.3	4
5672	IoE-MPP: A mobile portal platform for internet of everything. Journal of Intelligent and Fuzzy Systems, 2017, 32, 3069-3080.	0.8	3
5673	Mobile Big Data: The Fuel for Data-Driven Wireless. IEEE Internet of Things Journal, 2017, 4, 1489-1516.	5.5	72
5674	Modeling Multiclass Task-Based Applications on Heterogeneous Distributed Environments. Lecture Notes in Computer Science, 2017, , 166-180.	1.0	0
5675	CATS: cache-aware task scheduling for Hadoop-based systems. Cluster Computing, 2017, 20, 3691-3705.	3.5	4
5676	SD-CPS: Taming the challenges of Cyber-Physical Systems with a Software-Defined approach. , 2017, , .		16
5677	MapReduce Implementation of Strassen's Algorithm for Matrix Multiplication. , 2017, , .		4
5678	Improving the performance and reproducibility of experiments on large-scale testbeds with k -cores. Computer Communications, 2017, 110, 35-47.	3.1	1
5679	A scalable method for link prediction in large real world networks. Journal of Parallel and Distributed Computing, 2017, 109, 89-101.	2.7	28
5680	On the Maximum Rate of Networked Computation in a Capacitated Network. IEEE/ACM Transactions on Networking, 2017, 25, 2444-2458.	2.6	0
5681	An architecture for mitigating near earth object's impact to the earth. , 2017, , .		0
5682	Analysis of massive industrial data using MapReduce framework for parallel processing. , 2017, , .		1
5683	Self-organized dynamic provisioning for big data. Cluster Computing, 2017, 20, 2749-2762.	3.5	1
5684	Efficient Nonlinear Regression-Based Compression of Big Sensing Data on Cloud. , 2017, , 83-98.		3
5685	Extreme Learning Machine and Its Applications in Big Data Processing. , 2017, , 117-150.		20
5686	Smart Sensor Networks for Building Safety. , 2017, , 241-255.		1
5687	The Internet of Things and Its Applications. , 2017, , 256-279.		3

#	ARTICLE	IF	CITATIONS
5688	A MapReduce-based scalable discovery and indexing of structured big data. Future Generation Computer Systems, 2017, 73, 32-43.	4.9	28
5689	Data placement in massively distributed environments for fast parallel mining of frequent itemsets. Knowledge and Information Systems, 2017, 53, 207-237.	2.1	2
5690	A Spark-based genetic algorithm for sensor placement in large scale drinking water distribution systems. Cluster Computing, 2017, 20, 1089-1099.	3.5	29
5691	Scalable two-phase co-occurring sensitive pattern hiding using MapReduce. Journal of Big Data, 2017, 4, .	6.9	9
5693	Performance modeling of big data applications in the cloud centers. Journal of Supercomputing, 2017, 73, 2258-2283.	2.4	15
5694	Non-negative matrix factorisation of large mass spectrometry datasets. Chemometrics and Intelligent Laboratory Systems, 2017, 163, 76-85.	1.8	38
5695	Distributed asteroid discovery system for large astronomical data. Journal of Network and Computer Applications, 2017, 93, 27-37.	5.8	6
5696	Improving Execution Concurrency of Large-Scale Matrix Multiplication on Distributed Data-Parallel Platforms. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 2539-2552.	4.0	19
5697	Large-Scale Clustering Algorithms. Studies in Big Data, 2017, , 3-28.	0.8	0
5698	Improving Fair Scheduling Performance on Hadoop. , 2017, , .		3
5699	Big Data for Effective Management of Smart Grids. Studies in Big Data, 2017, , 209-229.	0.8	1
5700	Mitigating Service Variability in MapReduce Clusters via Task Cloning: A Competitive Analysis. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 2866-2880.	4.0	3
5701	Modeling temporal aspects of sensor data for MongoDB NoSQL database. Journal of Big Data, 2017, 4, .	6.9	32
5702	Distributed and weighted extreme learning machine for imbalanced big data learning. Tsinghua Science and Technology, 2017, 22, 160-173.	4.1	20
5703	Knapp: A Packet Processing Framework for Manycore Accelerators. , 2017, , .		1
5704	Distributed and scalable sequential pattern mining through stream processing. Knowledge and Information Systems, 2017, 53, 365-390.	2.1	21
5705	Estimation of the prevalence of adverse drug reactions from social media. International Journal of Medical Informatics, 2017, 102, 130-137.	1.6	33
5706	Current Challenges with BIGDATA Analytics in Structural Health Monitoring. Conference Proceedings of the Society for Experimental Mechanics, 2017, , 79-84.	0.3	16

#	ARTICLE	IF	CITATIONS
5707	M-Grid: a distributed framework for multidimensional indexing and querying of location based data. Distributed and Parallel Databases, 2017, 35, 55-81.	1.0	12
5708	Performance Comparison of Distributed Processing of Large Volume of Data on Top of Xen and Docker-Based Virtual Clusters. Lecture Notes in Computer Science, 2017, , 103-113.	1.0	5
5709	A speculative execution strategy based on node classification and hierarchy index mechanism for heterogeneous Hadoop systems. , 2017, , .		2
5710	A scalable approach for tree segmentation within small-footprint airborne LiDAR data. Computers and Geosciences, 2017, 102, 139-147.	2.0	14
5711	Dynamic control of data center network and computation resources. , 2017, , .		0
5712	SEINA: A stealthy and effective internal attack in Hadoop systems. , 2017, , .		9
5713	D-Cube. , 2017, , .		39
5715	Efficient Ranking and Selection in Parallel Computing Environments. Operations Research, 2017, 65, 821-836.	1.2	58
5716	Measuring Moran's I in a Cost-Efficient Manner to Describe a Land-Cover Change Pattern in Large-Scale Remote Sensing Imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10, 2631-2639.	2.3	27
5717	Flow distribution-aware load balancing for the datacenter. Computer Communications, 2017, 106, 136-146.	3.1	16
5718	Putting the Pieces Back Together Again. , 2017, , .		16
5719	Top-k spatial-keyword publish/subscribe over sliding window. VLDB Journal, 2017, 26, 301-326.	2.7	28
5720	De-Bruijn graph with MapReduce framework towards metagenomic data classification. International Journal of Information Technology (Singapore), 2017, 9, 59-75.	1.8	40
5721	Traffic At-a-Glance: Time-Bounded Analytics on Large Visual Traffic Data. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 2703-2717.	4.0	5
5722	Design and implementation of HDFS data encryption scheme using ARIA algorithm on Hadoop. , 2017, , .		6
5723	Building environment analysis based on clustering methods from sensor data on top of the Hadoop platform. , 2017, , .		2
5724	An efficient cache management scheme for accessing small files in Distributed File Systems. , 2017, , .		2
5725	A Novel Distributed Recommendation Framework Using Big Data in Social Context. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31, 1759015.	0.7	1

#	ARTICLE	IF	CITATIONS
5726	Knowledge-Based Decision Support Systems for Personalized u-lifecare Big Data Services. Intelligent Systems Reference Library, 2017, , 187-203.	1.0	2
5727	Parallel Incremental Frequent Itemset Mining for Large Data. Journal of Computer Science and Technology, 2017, 32, 368-385.	0.9	4
5728	Big data for secure healthcare system: a conceptual design. Complex & Intelligent Systems, 2017, 3, 133-151.	4.0	56
5729	Anomaly detection for big log data using a Hadoop ecosystem. , 2017, , .		5
5730	The design and benchmarking of a Cloud-based platform for processing and visualization of traffic data. , 2017, , .		13
5731	On the mining and usage of Movement Patterns in large traffic networks. , 2017, , .		1
5732	Improved Apriori Algorithm Using Power Set on Hadoop. Advances in Intelligent Systems and Computing, 2017, , 245-254.	0.5	2
5733	Intelligent monitor system based on cloud and convolutional neural networks. Journal of Supercomputing, 2017, 73, 3260-3276.	2.4	13
5734	A high performance query analytical framework for supporting data-intensive climate studies. Computers, Environment and Urban Systems, 2017, 62, 210-221.	3.3	19
5735	Resilient application placement for geo-distributed cloud networks. Journal of Network and Computer Applications, 2017, 85, 14-31.	5.8	19
5736	Enhancing network I/o performance for a virtualized Hadoop cluster. Concurrency Computation Practice and Experience, 2017, 29, e3974.	1.4	1
5737	Leukocytes Classification and Segmentation in Microscopic Blood Smear: A Resource-Aware Healthcare Service in Smart Cities. IEEE Access, 2017, 5, 3475-3489.	2.6	81
5738	Joint Compression of Near-Duplicate Videos. IEEE Transactions on Multimedia, 2017, 19, 908-920.	5.2	13
5739	Chronological Big data Curation: A Study on the Enhanced Information Retrieval System. IEEE Access, 2017, 5, 11269-11277.	2.6	10
5740	SmallClient for big data: an indexing framework towards fast data retrieval. Cluster Computing, 2017, 20, 1193-1208.	3.5	14
5741	Multi-instance dictionary learning via multivariate performance measure optimization. Pattern Recognition, 2017, 66, 448-459.	5.1	1
5742	Using Data Variety for Efficient Progressive Big Data Processing in Warehouse-Scale Computers. IEEE Computer Architecture Letters, 2017, 16, 166-169.	1.0	15
5743	Goal-based composition of scalable hybrid analytics for heterogeneous architectures. Journal of Parallel and Distributed Computing, 2017, 108, 59-73.	2.7	3

#	ARTICLE	IF	CITATIONS
5744	Optimizing VM allocation and data placement for data-intensive applications in cloud using ACO metaheuristic algorithm. Engineering Science and Technology, an International Journal, 2017, 20, 616-628.	2.0	76
5745	A Nonrelational Data Warehouse for the Analysis of Field and Laboratory Data From Multiple Heterogeneous Photovoltaic Test Sites. IEEE Journal of Photovoltaics, 2017, 7, 230-236.	1.5	40
5746	Towards the Tradeoffs in Designing Data Center Network Architectures. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 260-273.	4.0	22
5747	Hive-Based Anomaly Detection in Hadoop Log Data Management. Lecture Notes in Electrical Engineering, 2017, , 837-842.	0.3	0
5748	Logical query optimization for Cloudera Impala system. Journal of Systems and Software, 2017, 125, 35-46.	3.3	5
5749	The Digital Music Lab. Journal on Computing and Cultural Heritage, 2017, 10, 1-21.	1.2	14
5750	Multi-query processing of XML data streams on multicore. Journal of Supercomputing, 2017, 73, 2339-2368.	2.4	2
5751	Practical-oriented protocols for privacy-preserving outsourced big data analysis: Challenges and future research directions. Computers and Security, 2017, 69, 97-113.	4.0	32
5752	Cost-Effective Online Trending Topic Detection and Popularity Prediction in Microblogging. ACM Transactions on Information Systems, 2017, 35, 1-36.	3.8	12
5753	Document Classification Using Word2Vec and Chi-square on Apache Spark. Lecture Notes in Electrical Engineering, 2017, , 867-872.	0.3	1
5754	Parallel meta-blocking for scaling entity resolution over big heterogeneous data. Information Systems, 2017, 65, 137-157.	2.4	44
5755	Adaptive fault detection and diagnosis using parsimonious Gaussian mixture models trained with distributed computing techniques. Journal of the Franklin Institute, 2017, 354, 2543-2572.	1.9	11
5756	HBase storage schemas for massive spatial vector data. Cluster Computing, 2017, 20, 3657-3666.	3.5	8
5757	Towards a Foundational API for Resilient Distributed Systems Design. , 2017, , .		8
5758	Bringing Parallel Patterns Out of the Corner. Transactions on Architecture and Code Optimization, 2017, 14, 1-26.	1.6	18
5759	Computation semantics of the functional scientific workflow language Cuneiform. Journal of Functional Programming, 2017, 27, .	0.5	9
5760	The IPTV video evaluation model based on big data. , 2017, , .		1
5761	A parallel fuzzy rule discovery algorithm and future goods automated trading system. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
5762	Significance of restful web services in bigdata analytics. , 2017, , .		0
5763	On the Design of Resilient Multicloud MapReduce. IEEE Cloud Computing, 2017, 4, 74-82.	5.3	4
5764	Scalable Hadoop-Based Pooled Time Series of Big Video Data from the Deep Web. , 2017, , .		1
5765	GATE Monte Carlo simulation of dose distribution using MapReduce in a cloud computing environment. Australasian Physical and Engineering Sciences in Medicine, 2017, 40, 777-783.	1.4	0
5766	Big data in government organizations: A review. , 2017, , .		0
5767	Privacy-preserving pattern matching over encrypted genetic data in cloud computing. , 2017, , .		29
5768	Optimization Scheme Based on Parallel Computing Technology. Communications in Computer and Information Science, 2017, , 504-513.	0.4	0
5769	freeCycles - Efficient Multi-Cloud Computing Platform. Journal of Grid Computing, 2017, 15, 501-526.	2.5	2
5770	Modelling and Control of Big Data Frameworks * *This work was partially supported by the Swedish Foundation for Strategic Research under the project "Future factories in the cloud (FiC)" with grant number GMT14-0032.. IFAC-PapersOnLine, 2017, 50, 6110-6115.	0.5	2
5771	Optimizing Data-Intensive Applications Automatically By Leveraging Parallel Data Processing Frameworks. , 2017, , .		5
5772	Towards effective scheduling policies for many-task applications: Practice and experience based on HTCaaS. Concurrency Computation Practice and Experience, 2017, 29, e4242.	1.4	1
5773	A cooperative scheduling method based on the device load feedback for multiple tasks scheduling. Journal of Network and Computer Applications, 2017, 99, 110-119.	5.8	4
5774	From conceptual design to performance optimization of ETL workflows: current state of research and open problems. VLDB Journal, 2017, 26, 777-801.	2.7	44
5775	Samza. Proceedings of the VLDB Endowment, 2017, 10, 1634-1645.	2.1	194
5776	Streaming Algorithm for Euler Characteristic Curves of Multidimensional Images. Lecture Notes in Computer Science, 2017, , 397-409.	1.0	2
5777	Design of Distributed Calculation Scheme Using Network Address Translation for Ad-hoc Wireless Positioning Network. Communications in Computer and Information Science, 2017, , 33-48.	0.4	0
5778	SCENIC: single-cell regulatory network inference and clustering. Nature Methods, 2017, 14, 1083-1086.	9.0	3,086
5779	A Survey of Active Object Languages. ACM Computing Surveys, 2018, 50, 1-39.	16.1	65

#	ARTICLE	IF	CITATIONS
5780	Incremental Frequent Itemsets Mining with MapReduce. Lecture Notes in Computer Science, 2017, , 247-261.	1.0	0
5781	A Comparison of Distributed Machine Learning Platforms. , 2017, , .		19
5782	Research of Hadoop-Based Digital Library Data Service System. , 2017, , .		4
5783	Towards Clustering Validation in Big Data Context. , 2017, , .		4
5784	SparkGA. , 2017, , .		21
5785	A containerized analytics framework for data and compute-intensive pipeline applications. , 2017, , .		3
5786	MapReduce Based Parallel Bayesian Network for Manufacturing Quality Control. Chinese Journal of Mechanical Engineering (English Edition), 2017, 30, 1216-1226.	1.9	5
5787	Big Data for Context Aware Computing “ Perspectives and Challenges. Big Data Research, 2017, 10, 33-43.	2.6	38
5788	Efficient Parallelization of Regular Expression Matching for Deep Inspection. , 2017, , .		4
5789	Method of predicting bus arrival time based on MapReduce combining clustering with neural network. , 2017, , .		11
5790	Trends and challenges in smart healthcare research: A journey from data to wisdom. , 2017, , .		27
5791	iiHadoop: an asynchronous distributed framework for incremental iterative computations. Journal of Big Data, 2017, 4, .	6.9	5
5792	Multi-class Resource Sharing with Batch Arrivals and Complete Blocking. Lecture Notes in Computer Science, 2017, , 157-169.	1.0	10
5793	MapReduce and Its Applications, Challenges, and Architecture: a Comprehensive Review and Directions for Future Research. Journal of Grid Computing, 2017, 15, 295-321.	2.5	50
5794	Sphinx: Empowering Impala for Efficient Execution of SQL Queries on Big Spatial Data. Lecture Notes in Computer Science, 2017, , 65-83.	1.0	23
5795	A Review of Resource Scheduling in Large-Scale Server Cluster. Communications in Computer and Information Science, 2017, , 494-505.	0.4	1
5796	Distributed Extreme Learning Machine with Kernels Based on Mapreduce for Spectral-Spatial Classification of Hyperspectral Image. , 2017, , .		1
5797	Big data analytics architecture for real-time traffic control. , 2017, , .		83

#	ARTICLE	IF	CITATIONS
5798	FlowDB. , 2017, , .		12
5799	An Overview of the MapReduce Model. Lecture Notes in Computer Science, 2017, , 312-317.	1.0	1
5800	Caching for Pattern Matching Queries in Time Evolving Graphs: Challenges and Approaches. , 2017, , .		5
5801	ADMM for Penalized Quantile Regression in Big Data. International Statistical Review, 2017, 85, 494-518.	1.1	27
5802	Managing Provenance of Implicit Data Flows in Scientific Experiments. ACM Transactions on Internet Technology, 2017, 17, 1-22.	3.0	3
5803	Parallelization of Massive Textstream Compression Based on Compressed Sensing. ACM Transactions on Information Systems, 2018, 36, 1-18.	3.8	7
5804	The SusCity Big Data Warehousing Approach for Smart Cities. , 2017, , .		14
5805	A Case for Term Weighting Using a Dictionary on GPUs. Lecture Notes in Computer Science, 2017, , 103-117.	1.0	0
5806	A Fine-Grained Distribution Approach for ETL Processes in Big Data Environments. Data and Knowledge Engineering, 2017, 111, 114-136.	2.1	20
5807	A distributed graph approach for pre-processing linked RDF data using supercomputers. , 2017, , .		0
5808	Minimal Coflow Routing and Scheduling in OpenFlow-Based Cloud Storage Area Networks. , 2017, , .		5
5809	DeepDeath: Learning to predict the underlying cause of death with Big Data. , 2017, 2017, 3373-3376.		2
5810	Hierarchical Self-Tuning of Concurrency and Resource Units in Data-Analytics Frameworks. , 2017, , .		5
5811	An Auto-Scaling Framework for Heterogeneous Hadoop Systems. International Journal of Cooperative Information Systems, 2017, 26, 1750004.	0.6	5
5812	Distributed Local Outlier Detection in Big Data. , 2017, , .		22
5813	Credit-Scheduled Delay-Bounded Congestion Control for Datacenters. , 2017, , .		144
5814	SAMPGA task scheduling algorithm in cloud computing. , 2017, , .		6
5815	Execution of Recursive Queries in Apache Spark. Lecture Notes in Computer Science, 2017, , 289-302.	1.0	5

#	ARTICLE	IF	CITATIONS
5816	Graph Computation and Its Applications in Smart Grid. , 2017, , .		21
5817	On the energy efficiency of MapReduce shuffling operations in data centers. , 2017, , .		6
5818	Scalable Performance Tuning of Hadoop MapReduce: A Noisy Gradient Approach. , 2017, , .		11
5819	Detection and Classification of Bleeding Region in WCE Images using Color Feature. , 2017, , .		20
5820	Encoded distributed optimization. , 2017, , .		33
5821	Megalloc: Fast Distributed Memory Allocator for NVM-Based Cluster. , 2017, , .		1
5822	An Efficient Map-Reduce Framework to Mine Periodic Frequent Patterns. Lecture Notes in Computer Science, 2017, , 120-129.	1.0	3
5823	An experimental analysis of limitations of MapReduce for iterative algorithms on Spark. Cluster Computing, 2017, 20, 3593-3604.	3.5	4
5824	DRS: Auto-Scaling for Real-Time Stream Analytics. IEEE/ACM Transactions on Networking, 2017, 25, 3338-3352.	2.6	40
5825	A local rational model approach for H^{∞} norm estimation: With application to an active vibration isolation system. Control Engineering Practice, 2017, 69, 63-70.	1.2	1
5826	Hybrid storage architecture and efficient MapReduce processing for unstructured data. Parallel Computing, 2017, 69, 63-77.	1.3	10
5827	Simultaneous Data Management in Sensor-Based Systems using Metadata, Disaggregation and Processing. Procedia Computer Science, 2017, 112, 2117-2126.	1.2	4
5828	A Real-Time Scheduling Strategy Based on Processing Framework of Hadoop. , 2017, , .		2
5829	Coflourish: An SDN-Assisted Coflow Scheduling Framework for Clouds. , 2017, , .		0
5830	Smart urban planning using big data analytics based internet of things. , 2017, , .		12
5831	Security Issues and Challenges in Big Data Analytics in Distributed Environment. Scalable Computing and Communications, 2017, , 83-94.	0.5	2
5832	Utilizing Big Data Analytics for Automatic Building of Language-agnostic Semantic Knowledge Bases. Scalable Computing and Communications, 2017, , 137-160.	0.5	0
5833	StreamLearner. , 2017, , .		3

#	ARTICLE	IF	CITATIONS
5834	Archimedes. SIGMOD Record, 2017, 46, 30-35.	0.7	3
5835	Big data storage technologies: a survey. Frontiers of Information Technology and Electronic Engineering, 2017, 18, 1040-1070.	1.5	76
5837	RMHC-MR: Instance selection by random mutation hill climbing algorithm with MapReduce in big data. Procedia Computer Science, 2017, 111, 252-259.	1.2	4
5838	Task Scheduling in Big Data Platforms: A Systematic Literature Review. Journal of Systems and Software, 2017, 134, 170-189.	3.3	37
5839	Internet of Things: From Small- to Large-Scale Orchestration. , 2017, , .		6
5840	Multi-representation Based Data Processing Architecture for IoT Applications. , 2017, , .		8
5841	Proposing stochastic probability-based math model and algorithms utilizing social networking and academic data for good fit students prediction. Social Network Analysis and Mining, 2017, 7, 1.	1.9	11
5842	Smart urban planning using Big Data analytics to contend with the interoperability in Internet of Things. Future Generation Computer Systems, 2017, 77, 65-76.	4.9	89
5843	MapReduce-Based Algorithms for Managing Big RDF Graphs: State-of-the-Art Analysis, Paradigms, and Future Directions. , 2017, , .		6
5844	Docode 5: Building a real-world plagiarism detection system. Engineering Applications of Artificial Intelligence, 2017, 64, 261-271.	4.3	9
5845	Chrysaor: Fine-Grained, Fault-Tolerant Cloud-of-Clouds MapReduce. , 2017, , .		4
5846	Efficient Event Correlation over Distributed Systems. , 2017, , .		18
5847	Cloud Resource Scaling for Big Data Streaming Applications Using a Layered Multi-dimensional Hidden Markov Model. , 2017, , .		15
5848	Task-aware TCP in Data Center Networks. , 2017, , .		0
5849	Analysis and Evaluation of the GAS Model for Distributed Graph Computation. , 2017, , .		1
5850	An efficient approach for outlier detection in big sensor data of health care. International Journal of Communication Systems, 2017, 30, e3352.	1.6	26
5851	Quantum Operating Systems. , 2017, , .		11
5852	Real-Time Machine Learning. , 2017, , .		31

#	ARTICLE	IF	CITATIONS
5853	Industrial Big Data for Fault Diagnosis: Taxonomy, Review, and Applications. IEEE Access, 2017, 5, 17368-17380.	2.6	125
5854	An Improved Straggler Identification Scheme for Data-Intensive Computing on Cloud Platforms. , 2017, , .		3
5855	Performance evaluation of cluster algorithms for Big Data analysis on cloud. , 2017, , .		1
5856	Sparse representations based distributed attribute learning for person re-identification. Multimedia Tools and Applications, 2017, 76, 25015-25037.	2.6	6
5857	Keddah: Capturing Hadoop Network Behaviour. , 2017, , .		2
5858	Toward Stream-Based IP Flow Analysis. , 2017, 55, 70-76.		18
5859	Predicting seismic events in coal mines based on underground sensor measurements. Engineering Applications of Artificial Intelligence, 2017, 64, 83-94.	4.3	26
5861	Parallel resolution of large-scale GNSS network un-difference ambiguity. Advances in Space Research, 2017, 60, 2637-2647.	1.2	5
5862	Enabling Wide Area Data Analytics with Collaborative Distributed Processing Pipelines (CDPPs). , 2017, , .		1
5863	Finding the Best Box-Cox Transformation in Big Data with Meta-Model Learning: A Case Study on QCT Developer Cloud. , 2017, , .		2
5864	Big-data-driven innovation for enterprises. , 2017, , .		5
5865	Toward a FraMewOrk for EmbeDded & CollAborative data AnaLyls with HeTerogeneous Devlces. , 2017, , .		2
5866	Designing ETL processes to integrate multi-field digital information resources. , 2017, , .		1
5867	Data mining in distributed environment: a survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2017, 7, e1216.	4.6	91
5868	Research on formalization of efficient query application problems with compound condition in software development. Journal of China Universities of Posts and Telecommunications, 2017, 24, 18-30.	0.8	0
5869	Process state synchronization for mobility support in mobile cloud computing. , 2017, , .		15
5870	Modeling and Simulation of Extreme-Scale Fat-Tree Networks for HPC Systems and Data Centers. ACM Transactions on Modeling and Computer Simulation, 2017, 27, 1-23.	0.6	3
5871	Random access in nondelimited variable-length record collections for parallel reading with Hadoop. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
5872	A Lightweight MapReduce Framework for Secure Processing with SGX. , 2017, , .		19
5873	A platform for MBDAaaS based on patterns and skeletons: The python based algorithms compiler. , 2017, , .		5
5874	Integration of big data for querying CAN bus data from connected car. , 2017, , .		9
5875	Enterprise data breach: causes, challenges, prevention, and future directions. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2017, 7, e1211.	4.6	114
5876	Parallelizing Sequential Graph Computations. , 2017, , .		50
5877	ABCpy. , 2017, , .		20
5878	PADP: A parallel data possession audit model for cloud storage. Concurrency Computation Practice and Experience, 2017, 29, e4154.	1.4	1
5879	Making a case for the on-demand multiple distributed message queue system in a Hadoop cluster. Cluster Computing, 2017, 20, 2095-2106.	3.5	4
5880	Leveraging multiple coflow attributes for information-agnostic coflow scheduling. , 2017, , .		5
5881	KunPeng. , 2017, , .		31
5882	State of the art in MapReduce. , 2017, , .		0
5883	NFP. , 2017, , .		169
5884	Energy-aware coflow and antenna scheduling for hybrid server-centric data center networks. , 2017, , .		3
5885	A cloud-enabled automatic disaster analysis system of multi-sourced data streams: An example synthesizing social media, remote sensing and Wikipedia data. Computers, Environment and Urban Systems, 2017, 66, 23-37.	3.3	51
5886	A parallel computation of skyline using multiple regression analysis-based filtering on MapReduce. Distributed and Parallel Databases, 2017, 35, 383-409.	1.0	4
5887	Cellular matrix model for parallel combinatorial optimization algorithms in Euclidean plane. Applied Soft Computing Journal, 2017, 61, 642-660.	4.1	4
5888	A Longitudinal Study of Google Play. IEEE Transactions on Computational Social Systems, 2017, 4, 135-149.	3.2	7
5889	Activity Recognition Based on Pattern Recognition of Myoelectric Signals for Rehabilitation. Scalable Computing and Communications, 2017, , 427-442.	0.5	7

#	ARTICLE	IF	CITATIONS
5890	Using Spark and GraphX to Parallelize Large-Scale Simulations of Bacterial Populations over Host Contact Networks. Lecture Notes in Computer Science, 2017, , 591-600.	1.0	2
5891	Hayabusa. , 2017, , .		2
5892	Small Batch or Large Batch?. , 2017, , .		6
5893	SCARF: A container-based approach to cloud-scale digital forensic processing. Digital Investigation, 2017, 22, S39-S47.	3.2	21
5894	Scaling Reliably. ACM Transactions on Programming Languages and Systems, 2017, 39, 1-46.	1.7	6
5895	HPAT. , 2017, , .		9
5896	Distributed eigenfaces for massive face image data. Multimedia Tools and Applications, 2017, 76, 25983-26000.	2.6	4
5897	Pattern generation from event oriented sensor data using distributed sensor transaction model. , 2017, , .		1
5898	Large Scale Medical Data Mining for Accurate Diagnosis: A Blueprint. Scalable Computing and Communications, 2017, , 157-176.	0.5	16
5899	Association rules mining on MapReduce. , 2017, , .		1
5900	Science and data science. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114, 8689-8692.	3.3	119
5901	Performance Study of Ceph Storage with Intel Cache Acceleration Software: Decoupling Hadoop MapReduce and HDFS over Ceph Storage. , 2017, , .		5
5902	Improvising block placement policy in HADOOP framework. , 2017, , .		0
5903	DataFlow Systems: From Their Origins to Future Applications in Data Analytics, Deep Learning, and the Internet of Things. Computer Communications and Networks, 2017, , 127-148.	0.8	1
5904	When Hashes Met Wedges. , 2017, , .		4
5905	Automatic Arabic Text Summarization for Large Scale Multiple Documents Using Genetic Algorithm and MapReduce. , 2017, , .		2
5906	Model-Based Performance Evaluation of Batch and Stream Applications for Big Data. , 2017, , .		6
5907	Partitioning-Insensitive Watermarking Approach for Distributed Relational Databases. Lecture Notes in Computer Science, 2017, , 172-192.	1.0	4

#	ARTICLE	IF	CITATIONS
5908	<i>AliMe Assist</i> . , 2017, , .		65
5909	Query-able Kafka. <i>Proceedings of the VLDB Endowment</i> , 2017, 10, 1646-1657.	2.1	6
5911	Scalable and Parallel Processing of Influence Maximization for Large-Scale Social Networks. , 2017, , .		0
5912	Parallel Processing over Spatial-Temporal Datasets from Geo, Bio, Climate and Social Science Communities: A Research Roadmap. , 2017, , .		18
5913	Hybrid data warehouse model for climate big data analysis. , 2017, , .		7
5914	Adaptive distributed energy-saving data gathering technique for wireless sensor networks. , 2017, , .		21
5915	PSOM”partitioning-based scalable ontology matching using MapReduce. <i>Sadhana - Academy Proceedings in Engineering Sciences</i> , 2017, 42, 2009-2024.	0.8	3
5916	Optimized on-demand data streaming from sensor nodes. , 2017, , .		18
5917	Adaptive real-time anomaly detection in cloud infrastructures. <i>Concurrency Computation Practice and Experience</i> , 2017, 29, e4193.	1.4	15
5918	Is Big Data a Big Deal for Applied Microeconomics?. , 2017, , 35-52.		1
5920	MapReduce-based skyline query processing scheme using adaptive two-level grids. <i>Cluster Computing</i> , 2017, 20, 3605-3616.	3.5	7
5921	A bibliometric approach to tracking big data research trends. <i>Journal of Big Data</i> , 2017, 4, .	6.9	70
5922	Big Media Data Analysis. <i>Signal Processing: Image Communication</i> , 2017, 59, 105-108.	1.8	5
5923	OpenCluster: A Flexible Distributed Computing Framework for Astronomical Data Processing. <i>Publications of the Astronomical Society of the Pacific</i> , 2017, 129, 024001.	1.0	8
5924	ICT for renewable energy integration into smart buildings: IoT and big data approach. , 2017, , .		25
5925	QoS in NaaS (Network-as-a-Service) Using Software Defined Networking. <i>Service Science: Research and Innovations in the Service Economy</i> , 2017, , 115-139.	1.1	3
5926	An Improved FP-Growth Algorithm Based on SOM Partition. <i>Communications in Computer and Information Science</i> , 2017, , 166-178.	0.4	2
5927	Fraud detection in big data using supervised and semi-supervised learning techniques. , 2017, , .		32

#	ARTICLE	IF	CITATIONS
5928	An enhanced pre-processing model for big data processing: A quality framework. , 2017, , .		5
5929	A new lexicon learning algorithm for sentiment analysis of big data. , 2017, , .		3
5930	Properties, Principles, and Metrics in Transportation CPS. SpringerBriefs in Computer Science, 2017, , 51-63.	0.2	0
5931	Classification of scientific papers with big data technologies. , 2017, , .		3
5932	Magnet: Real-Time Trace Stream Analytics Framework for 5G Operations Support Systems. IEEE Network, 2017, 31, 6-13.	4.9	0
5933	Scheduling Coflows in Datacenter Networks. Performance Evaluation Review, 2017, 45, 29-30.	0.4	5
5934	ReForeSt: Random Forests in Apache Spark. Lecture Notes in Computer Science, 2017, , 331-339.	1.0	3
5935	IncBricks. Operating Systems Review (ACM), 2017, 51, 795-809.	1.5	6
5936	Big data analysis to Features Opinions Extraction of customer. Procedia Computer Science, 2017, 112, 906-916.	1.2	29
5937	Introducing Design Patterns to Knowledge Processing Systems in the Context of Big Data and Cloud Platforms. , 2017, , .		0
5938	OFLoad: An OpenFlow-Based Dynamic Load Balancing Strategy for Datacenter Networks. IEEE Transactions on Network and Service Management, 2017, 14, 792-803.	3.2	31
5939	Parallel Multi-Label Propagation Based on Influence Model and Its Application to Overlapping Community Discovery. International Journal on Artificial Intelligence Tools, 2017, 26, 1760013.	0.7	4
5940	Scalable implementation of dependence clustering in Apache Spark. , 2017, , .		2
5941	Seismic data attribute extraction based on Hadoop platform. , 2017, , .		8
5942	A Survey on Fault Management in Software-Defined Networks. IEEE Communications Surveys and Tutorials, 2017, 19, 2284-2321.	24.8	92
5943	Efficient Anytime Anywhere Algorithms for Vertex Additions in Large and Dynamic Graphs. , 2017, , .		3
5944	Data Science: A New Paradigm in the Age of Big-Data Science and Analytics. New Mathematics and Natural Computation, 2017, 13, 119-143.	0.4	11
5945	Towards Engineering a Web-Scale Multimedia Service. , 2017, , .		9

#	ARTICLE	IF	CITATIONS
5946	Skipping congestion-links for coflow scheduling. , 2017, , .		3
5947	Cost-efficient negotiation over multiple resources with reinforcement learning. , 2017, , .		3
5948	A Parallel Conditional Random Fields Model Based on Spark Computing Environment. Journal of Grid Computing, 2017, 15, 323-342.	2.5	7
5949	StreamQRE: modular specification and efficient evaluation of quantitative queries over streaming data. , 2017, , .		26
5950	CRED: Cloud Right-Sizing with Execution Deadlines and Data Locality. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 3389-3400.	4.0	7
5951	Chronograph. , 2017, , .		10
5952	A Review of Big Graph Mining Research. IOP Conference Series: Materials Science and Engineering, 2017, 180, 012065.	0.3	7
5953	Storage-Tag-Aware Scheduler for Hadoop Cluster. IEEE Access, 2017, 5, 13742-13755.	2.6	22
5954	A Big Data Architecture for Automotive Applications: PSA Group Deployment Experience. , 2017, , .		7
5955	Parallel Progressive Approach to Entity Resolution Using MapReduce. , 2017, , .		17
5956	A theoretical framework for cloud resource provisioning using MapReduce technique. , 2017, , .		0
5957	A load balancing scheme for sensing and analytics on a mobile edge computing network. , 2017, , .		15
5958	BBQ: Elastic MapReduce over Cloud Platforms. , 2017, , .		3
5959	Cost Model and Analysis of Iterative MapReduce Applications for Hybrid Cloud Bursting. , 2017, , .		6
5960	Flexible Scheduling of Distributed Analytic Applications. , 2017, , .		4
5961	AnalyzeThat: A Programmable Shared-Memory System for an Array of Processing-In-Memory Devices. , 2017, , .		3
5962	Optimized MapFile Based Storage of Small Files in Hadoop. , 2017, , .		11
5963	Evaluation of Data Locality Strategies for Hybrid Cloud Bursting of Iterative MapReduce. , 2017, , .		9

#	ARTICLE	IF	CITATIONS
5964	Terabyte Sort on FPGA-Accelerated Flash Storage. , 2017, , .		27
5965	FitRank " Social app to combat physical inactivity study of the use of fitness social apps on Facebook's users profiles. , 2017, , .		0
5966	Addressing Performance Heterogeneity in MapReduce Clusters with Elastic Tasks. , 2017, , .		8
5967	Predicting Large Scale Fine Grain Energy Consumption. Energy Procedia, 2017, 111, 1079-1088.	1.8	9
5968	The MapReduce-based approach to improve vehicle controls on big traffic events. , 2017, , .		3
5969	WevQuery. Proceedings of the ACM on Human-Computer Interaction, 2017, 1, 1-17.	2.5	15
5970	Scalability and Realtime on Big Data, MapReduce, NoSQL and Spark. Lecture Notes in Business Information Processing, 2017, , 79-104.	0.8	0
5971	An evolutionary algorithm for mining rare association rules: A Big Data approach. , 2017, , .		7
5972	A first attempt on global evolutionary undersampling for imbalanced big data. , 2017, , .		13
5973	An adaptive scheduling algorithm for heterogeneous Hadoop systems. , 2017, , .		8
5974	Towards Scalable Non-Monotonic Stream Reasoning via Input Dependency Analysis. , 2017, , .		3
5975	FFTree: A flexible architecture for data center networks towards configurability and cost efficiency. , 2017, , .		3
5976	Kernel-based features for predicting population health indices from geocoded social media data. Decision Support Systems, 2017, 102, 22-31.	3.5	25
5977	CCFinder: using Spark to find clustering coefficient in big graphs. Journal of Supercomputing, 2017, 73, 4683-4710.	2.4	7
5978	A Taxonomy and Survey of Fault-Tolerant Workflow Management Systems in Cloud and Distributed Computing Environments. , 2017, , 285-320.		20
5979	Host managed contention avoidance storage solutions for Big Data. Journal of Big Data, 2017, 4, .	6.9	7
5980	DLRankSVM: an efficient distributed algorithm for linear RankSVM. Journal of Supercomputing, 2017, 73, 2157-2186.	2.4	2
5981	Structured I/O streams in Clive: a toolbox approach for wide area network computing. Journal of Internet Services and Applications, 2017, 8, .	1.6	1

#	ARTICLE	IF	CITATIONS
5982	Pattern matching of signature-based IDS using Myers algorithm under MapReduce framework. Eurasip Journal on Information Security, 2017, 2017, .	2.4	30
5983	Scaling up Bayesian variational inference using distributed computing clusters. International Journal of Approximate Reasoning, 2017, 88, 435-451.	1.9	7
5984	On the combinatorial design of data centre network topologies. Journal of Computer and System Sciences, 2017, 89, 328-348.	0.9	1
5985	A Distributed Multi-exemplar Affinity Propagation Clustering Algorithm Based on MapReduce. , 2017, , .		1
5986	PIC: Enable Large-Scale Privacy Preserving Content-Based Image Search on Cloud. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 3258-3271.	4.0	70
5987	Emerging Role of Data Analytics in Combating Financial Frauds. , 2017, , .		1
5988	Query-Driven Learning for Predictive Analytics of Data Subspace Cardinality. ACM Transactions on Knowledge Discovery From Data, 2017, 11, 1-46.	2.5	17
5989	Combining Vertex-Centric Graph Processing with SPARQL for Large-Scale RDF Data Analytics. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 3374-3388.	4.0	19
5990	Multi-Capacity Bin Packing with Dependent Items and its Application to the Packing of Brokered Workloads in Virtualized Environments. Future Generation Computer Systems, 2017, 72, 129-144.	4.9	5
5991	An analysis of fault detection strategies in wireless sensor networks. Journal of Network and Computer Applications, 2017, 78, 267-287.	5.8	134
5992	Leveraging Adaptive I/O to Optimize Collective Data Shuffling Patterns for Big Data Analytics. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 1663-1674.	4.0	19
5993	NoSQL database systems: a survey and decision guidance. Computer Science - Research and Development, 2017, 32, 353-365.	2.7	67
5994	vLocality: Revisiting Data Locality for MapReduce in Virtualized Clouds. IEEE Network, 2017, 31, 28-35.	4.9	19
5995	Big data challenges in ocean observation: a survey. Personal and Ubiquitous Computing, 2017, 21, 55-65.	1.9	31
5996	Model-Based Multiscale Gigapixel Image Formation Pipeline on GPU. IEEE Transactions on Computational Imaging, 2017, 3, 493-502.	2.6	3
5997	Optimizing Cloud-Service Performance: Efficient Resource Provisioning via Optimal Workload Allocation. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 1689-1702.	4.0	44
5998	Big Data and Cloud Computing: A Survey of the State-of-the-Art and Research Challenges. Studies in Big Data, 2017, , 23-41.	0.8	43
5999	MRCP-RM: A Technique for Resource Allocation and Scheduling of MapReduce Jobs with Deadlines. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 1375-1389.	4.0	29

#	ARTICLE	IF	CITATIONS
6001	The Unicorn Runtime: Efficient Distributed Shared Memory Programming for Hybrid CPU-GPU Clusters. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 1518-1534.	4.0	9
6002	An efficient MapReduce scheduling scheme for processing large multimedia data. Multimedia Tools and Applications, 2017, 76, 17273-17296.	2.6	13
6003	Curtailing job completion time in MapReduce clouds through improved Virtual Machine allocation. Computers and Electrical Engineering, 2017, 58, 190-202.	3.0	6
6004	RPig., 2017, , 193-216.		1
6006	Optimal construction of virtual networks for Cloud-based MapReduce workflows. Computer Networks, 2017, 112, 194-207.	3.2	7
6007	Generating automatic linguistic descriptions with big data. Information Sciences, 2017, 380, 12-30.	4.0	14
6008	Different aspects of workflow scheduling in large-scale distributed systems. Simulation Modelling Practice and Theory, 2017, 70, 120-134.	2.2	28
6009	Big Data and cloud computing: innovation opportunities and challenges. International Journal of Digital Earth, 2017, 10, 13-53.	1.6	537
6010	An automatic image-text alignment method for large-scale web image retrieval. Multimedia Tools and Applications, 2017, 76, 21401-21421.	2.6	2
6011	Big Data Behavioral Analytics Meet Graph Theory: On Effective Botnet Takedowns. IEEE Network, 2017, 31, 18-26.	4.9	25
6012	Estimation of Passenger Route Choice Pattern Using Smart Card Data for Complex Metro Systems. IEEE Transactions on Intelligent Transportation Systems, 2017, 18, 790-801.	4.7	87
6013	Sparse Functional Dynamical Models – A Big Data Approach. Journal of Computational and Graphical Statistics, 2017, 26, 319-329.	0.9	3
6014	Hashedcubes: Simple, Low Memory, Real-Time Visual Exploration of Big Data. IEEE Transactions on Visualization and Computer Graphics, 2017, 23, 671-680.	2.9	66
6015	A MapReduce-based artificial bee colony for large-scale data clustering. Pattern Recognition Letters, 2017, 93, 78-84.	2.6	44
6016	Persisting big-data: The NoSQL landscape. Information Systems, 2017, 63, 1-23.	2.4	108
6017	Accelerating big data analytics on HPC clusters using two-level storage. Parallel Computing, 2017, 61, 18-34.	1.3	19
6018	Failure detector abstractions for MapReduce-based systems. Information Sciences, 2017, 379, 112-127.	4.0	10
6019	MapReduce-based entity matching with multiple blocking functions. Frontiers of Computer Science, 2017, 11, 895-911.	1.6	5

#	ARTICLE	IF	CITATIONS
6020	Intelligent and independent processes for overcoming big graphs. Journal of Supercomputing, 2017, 73, 1438-1466.	2.4	1
6021	Visual Analysis of MOOC Forums with iForum. IEEE Transactions on Visualization and Computer Graphics, 2017, 23, 201-210.	2.9	55
6022	An Adaptive MapReduce Scheduler for Scalable Heterogeneous Systems. Advances in Intelligent Systems and Computing, 2017, , 603-611.	0.5	12
6023	MEFASD-BD: Multi-objective evolutionary fuzzy algorithm for subgroup discovery in big data environments - A MapReduce solution. Knowledge-Based Systems, 2017, 117, 70-78.	4.0	33
6024	A high-throughput energy-efficient passive optical datacenter network. Photonic Network Communications, 2017, 33, 258-274.	1.4	0
6025	Online MapReduce scheduling problem of minimizing the makespan. Journal of Combinatorial Optimization, 2017, 33, 590-608.	0.8	12
6026	Optimizing Crawler4j using MapReduce Programming Model. Journal of the Institution of Engineers (India): Series B, 2017, 98, 329-336.	1.3	0
6027	Hierarchical k-Nearest Neighbor with GPUs and a High Performance Cluster: Application to Handwritten Character Recognition. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31, 1750005.	0.7	9
6028	Makespan minimization for MapReduce systems with different servers. Future Generation Computer Systems, 2017, 67, 13-21.	4.9	12
6029	Finding efficiencies in frequent pattern mining from big uncertain data. World Wide Web, 2017, 20, 571-594.	2.7	12
6030	Enable cloud DevOps approach for industry and higher education. Concurrency Computation Practice and Experience, 2017, 29, e3937.	1.4	15
6031	Parallel pattern classification utilizing GPU-based kernelized Slackmin algorithm. Journal of Parallel and Distributed Computing, 2017, 99, 90-99.	2.7	5
6032	Dynamic Differential Privacy for ADMM-Based Distributed Classification Learning. IEEE Transactions on Information Forensics and Security, 2017, 12, 172-187.	4.5	130
6033	Availability/Network-aware MapReduce over the Internet. Information Sciences, 2017, 379, 94-111.	4.0	12
6034	Educational Big Data Mining: How to Enhance Virtual Learning Environments. Advances in Intelligent Systems and Computing, 2017, , 681-690.	0.5	13
6035	Fast and scalable inequality joins. VLDB Journal, 2017, 26, 125-150.	2.7	18
6036	Automating the license compatibility process in open source software with SPDX. Journal of Systems and Software, 2017, 131, 386-401.	3.3	18
6037	FARMS: Efficient mapreduce speculation for failure recovery in short jobs. Parallel Computing, 2017, 61, 68-82.	1.3	11

#	ARTICLE	IF	CITATIONS
6038	Detecting subgraph isomorphism with MapReduce. Journal of Supercomputing, 2017, 73, 1810-1851.	2.4	6
6039	A Comprehensive Study of MapReduce Over Lustre for Intermediate Data Placement and Shuffle Strategies on HPC Clusters. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 633-646.	4.0	20
6040	Distributed Computing Architecture for Optimal Control of Distribution Feeders With Smart Loads. IEEE Transactions on Smart Grid, 2017, 8, 1469-1478.	6.2	11
6041	Technological Evolution of the ICT Sector. Internet of Things, 2017, , 53-87.	1.3	1
6042	Hadoop Based Parallel Binary Bat Algorithm for Network Intrusion Detection. International Journal of Parallel Programming, 2017, 45, 1194-1213.	1.1	27
6044	A flexible and high-performance data center network topology. Journal of Supercomputing, 2017, 73, 1484-1503.	2.4	9
6045	Making the case for reforming the I/O software stack of extreme-scale systems. Advances in Engineering Software, 2017, 111, 26-31.	1.8	6
6046	Effective hybrid load scheduling of online and offline clusters for e-health service. Neurocomputing, 2017, 220, 60-66.	3.5	3
6047	A case study of tuning MapReduce for efficient Bioinformatics in the cloud. Parallel Computing, 2017, 61, 83-95.	1.3	14
6048	Cooperative particle swarm optimization using MapReduce. Soft Computing, 2017, 21, 6593-6603.	2.1	8
6049	XHAMI - extended HDFS and MapReduce interface for Big Data image processing applications in cloud computing environments. Software - Practice and Experience, 2017, 47, 455-472.	2.5	15
6050	Uploading multiply deferrable big data to the cloud platform using cost-effective online algorithms. Future Generation Computer Systems, 2017, 67, 276-285.	4.9	5
6051	A method to construct reduced-order parameter-varying models. International Journal of Robust and Nonlinear Control, 2017, 27, 582-597.	2.1	39
6052	Towards the efficient parallelization of multi-pass adaptive blocking for entity matching. Journal of Parallel and Distributed Computing, 2017, 101, 27-40.	2.7	5
6053	Metric Similarity Joins Using MapReduce. IEEE Transactions on Knowledge and Data Engineering, 2017, 29, 656-669.	4.0	22
6054	Using Large-Scale Social Media Networks as a Scalable Sensing System for Modeling Real-Time Energy Utilization Patterns. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47, 2627-2640.	5.9	11
6055	An Empirical Comparison of Algorithms to Find Communities in Directed Graphs and Their Application in Web Data Analytics. IEEE Transactions on Big Data, 2017, 3, 289-306.	4.4	20
6056	A Novel MapReduce Based k-Means Clustering. Advances in Intelligent Systems and Computing, 2017, , 247-255.	0.5	0

#	ARTICLE	IF	CITATIONS
6058	MapReduce-based clustering for near-duplicate image identification. Multimedia Tools and Applications, 2017, 76, 23291-23307.	2.6	0
6059	CloudFlows: Online workflows for distributed big data mining. Future Generation Computer Systems, 2017, 68, 38-58.	4.9	57
6060	A New Cloud Computing Architecture for the Classification of Remote Sensing Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10, 409-416.	2.3	39
6061	ExCCC-DCN: A Highly Scalable, Cost-Effective and Energy-Efficient Data Center Structure. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 1046-1060.	4.0	28
6062	Fully Scalable Methods for Distributed Tensor Factorization. IEEE Transactions on Knowledge and Data Engineering, 2017, 29, 100-113.	4.0	51
6063	Optimizing Graph Processing on GPUs. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 1149-1162.	4.0	17
6065	Information system for image classification based on frequency curve proximity. Information Systems, 2017, 64, 12-21.	2.4	0
6066	Networking for Big Data: A Survey. IEEE Communications Surveys and Tutorials, 2017, 19, 531-549.	24.8	174
6067	Privacy preserving big data publishing: a scalable <i>k</i>-anonymization approach using MapReduce. IET Software, 2017, 11, 271-276.	1.5	27
6068	Future trend of deep learning frameworks " From the perspective of big data analytics and HPC. , 2017, , .		2
6069	Streaming Distributed DNA Sequence Alignment Using Apache Spark. , 2017, , .		8
6070	Role of parallel support vector machine and map-reduce in risk analysis. , 2017, , .		1
6071	MDMP: A new algorithm to create inverted index files in BigData, using MapReduce. , 2017, , .		3
6072	Image processing on multinode hadoop cluster. , 2017, , .		6
6073	A new hybrid approach for document clustering. , 2017, , .		2
6074	Literature survey on existing analytical schemes to optimize the mining results using incremental map reduce. , 2017, , .		2
6075	On power consumption profiles for data intensive workloads in virtualized hadoop clusters. , 2017, , .		4
6076	Audition for multimedia computing. , 2017, , 31-50.		0

#	ARTICLE	IF	CITATIONS
6077	Hawkes processes for events in social media. , 2017, , 191-218.		42
6078	Cloud gaming. , 2017, , 287-314.		3
6079	External memory pipelining made easy with TPIE. , 2017, , .		2
6080	MPI-LiFE: Designing High-Performance Linear Fascicle Evaluation of Brain Connectome with MPI. , 2017, , .		1
6081	A learning-based mapreduce scheduler in heterogeneous environments. , 2017, , .		1
6082	Particle swarm optimization for large-scale clustering on apache spark. , 2017, , .		8
6083	A Survey of Data Scientists in South Africa. Communications in Computer and Information Science, 2017, , 175-191.	0.4	1
6084	On Spatial Joins in MapReduce. , 2017, , .		26
6085	An Efficient Task-based All-Reduce for Machine Learning Applications. , 2017, , .		3
6086	A domain specific language for business transaction processing. , 2017, , .		4
6087	Comparing consensus Monte Carlo strategies for distributed Bayesian computation. Brazilian Journal of Probability and Statistics, 2017, 31, .	0.1	12
6088	Integrating Big Data technologies in a dynamic environment EIAH dedicated to e-learning systems based on cloud infrastructure. , 2017, , .		4
6089	Big spatial data processing with Apache Spark. , 2017, , .		6
6090	Distributed decision tree v.2.0. , 2017, , .		5
6091	Towards Ex Vivo Testing of MapReduce Applications. , 2017, , .		5
6092	Victream. , 2017, , .		2
6093	Rainfall interception calculation modeling using parallel computation systems on NVIDIA CUDA. , 2017, , .		0
6094	Matrix multiplication of big data using MapReduce: A review. , 2017, , .		5

#	ARTICLE	IF	CITATIONS
6095	From VFX project management to predictive forecasting. , 2017, , .		2
6096	Exact fuzzy k-nearest neighbor classification for big datasets. , 2017, , .		13
6097	A Big Data Analysis Method for Online Education. , 2017, , .		4
6098	Distributed algorithm for sequential pattern mining on a large sequence dataset. , 2017, , .		3
6099	Short-term load forecasting based automatic distribution network reconfiguration. , 2017, , .		13
6100	Assessing and improving the performance and scalability of an iterative algorithm for Hadoop. , 2017, , .		1
6101	Network security and anomaly detection with Big-DAMA, a big data analytics framework. , 2017, , .		20
6102	PeerHunter: Detecting peer-to-peer botnets through community behavior analysis. , 2017, , .		20
6103	LASER: A Deep Learning Approach for Speculative Execution and Replication of Deadline-Critical Jobs in Cloud. , 2017, , .		9
6104	Exploratory Research on Developing Hadoop-Based Data Analytics Tools. , 2017, , .		0
6105	Name node performance enlarging by aggregator based HADOOP framework. , 2017, , .		3
6106	State management in Apache Flink®. Proceedings of the VLDB Endowment, 2017, 10, 1718-1729.	2.1	152
6107	GML learning, a generic machine learning model for network measurements analysis. , 2017, , .		16
6108	A Software Solution for Preventing Alzheimer's Disease Based on MapReduce Framework. , 2017, , .		2
6109	Distributed Statistical Machine Learning in Adversarial Settings. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2017, 1, 1-25.	1.4	189
6110	SparkGIS. , 2017, , .		35
6111	eTRIKS analytical environment: A modular high performance framework for medical data analysis. , 2017, , .		6
6112	Bigdata analysis and comparison of bigdata analytic approaches. , 2017, , .		4

#	ARTICLE	IF	CITATIONS
6114	Analysis of TPC-DS. , 2017, , .		30
6115	Discovery of Action Rules at Lowest Cost in Spark. , 2017, , .		3
6116	Holistic and Algebraic Data Cube Computation Using MapReduce. , 2017, , .		0
6117	Grid information management innovation of government in the era of big data. , 2017, , .		0
6118	A comparison of big data application programming approaches: A travel companion case study. , 2017, , .		1
6119	Cloud detection on the Google Earth engine platform. , 2017, , .		3
6120	The impact of data locality on the performance of a SaaS cloud with real-time data-intensive applications. , 2017, , .		11
6121	Quantifying volume, velocity, and variety to support (Big) data-intensive application development. , 2017, , .		14
6122	Scheduling Recurring Distributed Dataflow Jobs Based on Resource Utilization and Interference. , 2017, , .		8
6123	Integrating flight-related information into a (Big) data lake. , 2017, , .		13
6124	Cloud-POA: A cloud-based map only implementation of PO-MSA on Amazon multi-node EC2 Hadoop Cluster. , 2017, , .		2
6125	Deadline-aware MapReduce scheduling with selective speculative execution. , 2017, , .		3
6126	VSOC - A Virtual Security Operating Center. , 2017, , .		5
6127	A novel implementation for FastSLAM 2.0 algorithm based on cloud robotics. , 2017, , .		4
6128	Efficient spark-based framework for big geospatial data query processing and analysis. , 2017, , .		14
6129	Software systems for processing and analysis of big data and event streams. , 2017, , .		2
6130	Developing a localized web server for survival, generic and protein data analysis with high performance computing technology. , 2017, , .		0
6131	An Efficient Algorithm for Parallel Computation of Rough Entropy Using CUDA. , 2017, , .		3

#	ARTICLE	IF	CITATIONS
6132	Indexing for Large Scale Data Querying Based on Spark SQL. , 2017, , .		2
6133	Distributed single pass clustering algorithm based on MapReduce. , 2017, , .		1
6134	Machine learning techniques in Hadoop environment: A survey. , 2017, , .		1
6135	Survey on Big Data and Cloud Computing. , 2017, , .		4
6136	Measuring data locality ratio in virtual MapReduce cluster using WorkflowSim. , 2017, , .		2
6137	Deep learning for video classification and captioning. , 2017, , 3-29.		50
6138	Drizzle. , 2017, , .		97
6139	Quality of Information in Mobile Crowdsensing. ACM Transactions on Sensor Networks, 2017, 13, 1-43.	2.3	129
6140	ArrayUDF. , 2017, , .		28
6141	LAVEA. , 2017, , .		161
6142	Adaptive Feedforward and Feedback Control for Cloud Services. IFAC-PapersOnLine, 2017, 50, 5504-5509.	0.5	4
6143	Modeling energy and performance of Phoenix++ based parallel programs. , 2017, , .		1
6144	User clustering based on Canopy+K-means algorithm in cloud computing. Journal of Interdisciplinary Mathematics, 2017, 20, 1489-1492.	0.4	7
6145	Research and implementation of improved random forest algorithm based on Spark. , 2017, , .		8
6146	Scalable OLAP-Based Big Data Analytics over Cloud Infrastructures. , 2017, , .		8
6147	Real time financial analysis using big data technologies. , 2017, , .		4
6148	Sentiment classification on big data using Naïve bayes and logistic regression. , 2017, , .		39
6149	Spectral-spatial classification of hyperspectral image using Distributed Extreme Learning Machine with MapReduce. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
6150	HeatPipe: High Throughput, Low Latency Big Data Heatmap with Spark Streaming. , 2017, , .		8
6151	Software Engineering Education: Converging with the Startup Industry. , 2017, , .		18
6152	Accelerating big data applications using lightweight virtualization framework on enterprise cloud. , 2017, , .		45
6153	Big Data Management Performance Evaluation in Hadoop Ecosystem. , 2017, , .		2
6154	Metadata Service for Random Frame Reading of Video Stored in Distributed File System. , 2017, , .		0
6155	Fast Matrix Multiplication with Big Sparse Data. Cybernetics and Information Technologies, 2017, 17, 16-30.	0.4	1
6156	Connectivity-Aware Task Outsourcing and Scheduling in D2D Networks. , 2017, , .		5
6157	Scheduling jobs across geo-distributed datacenters with max-min fairness. , 2017, , .		16
6158	Mining on Big Data Using Hadoop MapReduce Model. IOP Conference Series: Materials Science and Engineering, 2017, 263, 042007.	0.3	2
6159	Eley: On the Effectiveness of Burst Buffers for Big Data Processing in HPC Systems. , 2017, , .		11
6160	Viyojit. , 2017, , .		14
6161	Parallelization of K-Means Clustering Algorithm for Data Mining. ITM Web of Conferences, 2017, 12, 03048.	0.4	0
6162	MTM: A Reliable Multiple Trees Multicast for Data Center Network. , 2017, , .		3
6163	Using Deep Learning to Predict and Optimize Hadoop Data Analytic Service in a Cloud Platform. , 2017, , .		8
6164	An Online-Offline Combined Big Data Mining Platform. , 2017, , .		0
6165	HCudaBLAST: an implementation of BLAST on Hadoop and Cuda. Journal of Big Data, 2017, 4, .	6.9	6
6166	Optimizing MapReduce Partitioner Using Naive Bayes Classifier. , 2017, , .		2
6167	Inferring Workflows with Job Dependencies from Distributed Processing Systems Logs. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
6168	Automated debugging in data-intensive scalable computing. , 2017, 2017, 520-534.		18
6169	Research on query analysis and optimization based on spark. , 2017, , .		4
6170	Ensuring reliability in geo-distributed edge cloud. , 2017, , .		11
6171	Efficient entity resolution using multiple blocking keys for bibliographic dataset. , 2017, , .		3
6172	Range queries in spatial index research based on the Spark. , 2017, , .		1
6173	A Multicore Path to Connectomics-on-Demand. ACM SIGPLAN Notices, 2017, 52, 267-281.	0.2	3
6174	Multimedia fog computing: minions in the cloud and crowd. , 2017, , 255-286.		2
6175	On the Optimization of Improved Apriori Algorithm via Linked-list Trie. , 2017, , .		3
6176	ParallelCharMax: An Effective Maximal Frequent Itemset Mining Algorithm Based on MapReduce Framework. , 2017, , .		3
6177	Filtering the big data based on volume, variety and velocity by using Kalman filter recursive approach. , 2017, , .		3
6178	Exploiting Data Dependency to Mitigate Stragglers in Distributed Spatial Simulation. , 2017, , .		8
6179	SwiftTuna: Responsive and incremental visual exploration of large-scale multidimensional data. , 2017, , .		1
6180	A Comparative Study of Data Skew in Hadoop. , 2017, , .		5
6181	Scheduling Coflows in Datacenter Networks. , 2017, , .		5
6182	KP-S: A Spark-Based Design of the K-Prototypes Clustering for Big Data. , 2017, , .		4
6183	A novel key partitioning schema for efficient execution of MapReduce applications. , 2017, , .		0
6184	Is Intel high performance analytics toolkit a good alternative to Apache Spark?. , 2017, , .		1
6185	Hadoop based mining of distributed association rules from big data. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
6186	Coflow-aware dynamic routing for SDN-based data center networks. , 2017, , .		0
6187	A Fault-Tolerant ONVIF Standard Extension for Cloud-Based Surveillance Systems. , 2017, , .		0
6188	Big Data Processing: Scalability with Extreme Single-Node Performance. , 2017, , .		1
6189	Shuffle phase optimization in spark. , 2017, , .		1
6190	Cooperation between data modeling and simulation modeling for performance analysis of Hadoop. , 2017, , .		2
6191	Modeling and Verifying Identity Authentication Security of HDFS Using CSP. , 2017, , .		5
6192	Parallel support vector machine used in map-reduce for risk analysis. , 2017, , .		5
6193	A data-driven approach for multivariate contextualized anomaly detection: Industry use case. , 2017, , .		5
6194	A fast large-size production data transformation scheme for supporting smart manufacturing in semiconductor industry. , 2017, , .		2
6195	An Evaluation Study of a Proposed Hadoop for Hybrid Networks (HHN). , 2017, , .		0
6196	A Cost Efficient Multicast Nonblocking Interconnection Network. , 2017, , .		0
6197	Unrolled Memory Inner-Products: An Abstract GPU Operator for Efficient Vision-Related Computations. , 2017, , .		4
6198	Encrypted domain multimedia content analysis. , 2017, , 75-104.		0
6199	Ashman: A Bandwidth Fragmentation-Based Dynamic Flow Scheduling for Data Center Networks. Computer Journal, 2017, 60, 1498-1509.	1.5	4
6200	Reserved, on demand or serverless: Model-based simulations for cloud budget planning. , 2017, , .		11
6201	Distributed randomized singular value decomposition using count sketch. , 2017, , .		3
6202	A review on big data protection of Hadoop. , 2017, , .		3
6203	An Overview of Practical Impacts of Functional Programming. , 2017, , .		4

#	ARTICLE	IF	CITATIONS
6204	A Computational Pipeline for the IUCN Risk Assessment for Meso-American Reef Ecosystem. , 2017, , .		1
6205	Understanding recurring quality problems and their impact on code sharing in block-based software. , 2017, , .		26
6206	Solving mean-shift clustering using MapReduce Hadoop. , 2017, , .		3
6207	Computation Capability Deduction Architecture for MapReduce on Cloud Computing. , 2017, , .		0
6208	Reducing Imbalance Ratio in MapReduce. , 2017, , .		2
6209	An efficient technique to improve resources utilization for hadoop MapReduce in heterogeneous system. , 2017, , .		2
6210	TASMR: Towards advanced secure mapreduc framework across untrusted hybrid clouds. , 2017, , .		3
6211	Perldoop2: A Big Data-Oriented Source-to-Source Perl-Java Compiler. , 2017, , .		3
6212	A Big Aurora Data Management Framework Toward Aurora Classification. , 2017, , .		0
6213	Network-Adaptive Scheduling of Data-Intensive Parallel Jobs with Dependencies in Clusters. , 2017, , .		5
6214	Discovering Interactions in Big Data Research: A Learning-Enhanced Bibliometric Study. , 2017, , .		1
6215	Visualization Analysis of NoSQL Research Field Based on SCI by CiteSpace V. , 2017, , .		0
6216	DeepSpotCloud: Leveraging Cross-Region GPU Spot Instances for Deep Learning. , 2017, , .		24
6217	Parallel graph Laplacian for large datasets. , 2017, , .		0
6218	Strike the Balance between System Utilization and Data Locality under Deadline Constraint for MapReduce Clusters. , 2017, , .		0
6219	Harvesting Underutilized Resources to Improve Responsiveness and Tolerance to Crash and Silent Faults for Data-Intensive Applications. , 2017, , .		2
6220	Tasks management of manned space ground data processing based on OSWorkflow engine and Web Services. , 2017, , .		0
6221	Modeling and Analyzing Latency in the Memcached system. , 2017, , .		6

#	ARTICLE	IF	CITATIONS
6222	Kaleido: Enabling Efficient Scientific Data Processing on Big-Data Systems. , 2017, , .		3
6223	A Survey of Semantics-Aware Performance Optimization for Data-Intensive Computing. , 2017, , .		4
6224	A Fast Distributed Principal Component Analysis with Variance Reduction. , 2017, , .		1
6225	MR-Droid: A Scalable and Prioritized Analysis of Inter-App Communication Risks. , 2017, , .		17
6226	Detection of I/O overloads in local HPC cluster rocks with LUSTRE file system. , 2017, , .		0
6227	The MapReduce application of matrix multiplication implemented on field programmable gate arrays. , 2017, , .		0
6228	Super learning for anomaly detection in cellular networks. , 2017, , .		7
6229	A first approach to handle fuzzy emerging patterns mining on big data problems: The EvAEFP-spark algorithm. , 2017, , .		4
6230	An Efficient MapReduce-Based Adaptive K-Means Clustering for Large Dataset. , 2017, , .		2
6231	Distributed PathGraph: A Cluster Centric Framework for Distributed Processing Graph. , 2017, , .		1
6232	Jointly optimizing task granularity and concurrency for in-memory mapreduce frameworks. , 2017, , .		6
6233	Visual analytics with unparalleled variety scaling for big earth data. , 2017, , .		4
6234	Crack random forest for arbitrary large datasets. , 2017, , .		1
6235	Trilogy: Data placement to improve performance and robustness of cloud computing. , 2017, , .		2
6236	RadiusSketch: Massively Distributed Indexing of Time Series. , 2017, , .		3
6237	Locality-aware hybrid coded MapReduce for server-rack architecture. , 2017, , .		7
6238	Communication vs distributed computation: An alternative trade-off curve. , 2017, , .		40
6239	The benefit of being flexible in distributed computation. , 2017, , .		10

#	ARTICLE	IF	CITATIONS
6240	Interconnect Energy Savings and Lower Latency Networks in Hadoop Clusters: The Missing Link. , 2017, , .		2
6241	Performance Modelling and Cost Effective Execution for Distributed Graph Processing on Configurable VMs. , 2017, , .		4
6242	Coded computation over heterogeneous clusters. , 2017, , .		55
6243	SOMR: Towards a Security-Oriented MapReduce Infrastructure. , 2017, , .		6
6244	Scalable Data Stream Clustering with k Estimation. , 2017, , .		3
6245	Shade: A differentially-private wrapper for enterprise big data. , 2017, , .		5
6246	CloudEC: A MapReduce-based algorithm for correcting errors in next-generation sequencing big data. , 2017, , .		7
6247	A scalable hybrid architecture for high performance data-parallel applications. , 2017, , .		1
6248	On Heterogeneous Coded Distributed Computing. , 2017, , .		45
6249	Distributing Frank-Wolfe via Map-Reduce. , 2017, , .		3
6250	Adapting MapReduce for Efficient Watermarking of Large Relational Dataset. , 2017, , .		4
6251	Online and Distributed Robust Regressions Under Adversarial Data Corruption. , 2017, , .		6
6252	An adaptive SLA-based data flow mechanism for stream processing engines. , 2017, , .		8
6253	A-MapCG: An Adaptive MapReduce Framework for GPUs. , 2017, , .		3
6254	A Transactional Model for Parallel Programming of Graph Applications on Computing Clusters. , 2017, , .		3
6255	GreenCloudTax: A Flexible IaaS Tax Approach as Stimulus for Green Cloud Computing. , 2017, , .		1
6256	P-Spar(k)ql: SPARQL Evaluation Method on Spark GraphX with Parallel Query Plan. , 2017, , .		5
6257	Design and implementation of a scalable distributed web crawler based on Hadoop. , 2017, , .		4

#	ARTICLE	IF	CITATIONS
6258	Enhanced protection for multimedia content in cloud. , 2017, , .		1
6259	Map-reduce based parallel support vector machine for risk analysis. , 2017, , .		1
6260	MeDNN: A distributed mobile system with enhanced partition and deployment for large-scale DNNs. , 2017, , .		58
6261	Research on Preprocessing Technology of Aircraft Flight Test Data Based on Spark. , 2017, , .		0
6262	Tula: A disk latency aware balancing and block placement strategy for Hadoop. , 2017, , .		13
6263	A novel compression algorithm decision method for spark shuffle process. , 2017, , .		3
6264	ECL-watch: A big data application performance tuning tool in the HPCC systems platform. , 2017, , .		4
6265	Data skew alleviation by balanced range assignment. , 2017, , .		0
6266	Planning in compute-aggregate problems as optimization problems on graphs. , 2017, , .		2
6267	Classification of microarray data using SVM mapreduce. , 2017, , .		4
6269	Hadoop and MapReduce technology as a solution for Wireless Body Area Networks in e-Health. , 2017, , .		2
6270	A global distributed approach to the Chi et al. fuzzy rule-based classification system for big data classification problems. , 2017, , .		2
6271	CP scheduler in GPU enabled Hadoop cluster. , 2017, , .		1
6272	Preparing HPC Applications for the Exascale Era: A Decoupling Strategy. , 2017, , .		2
6273	A big data framework for spacecraft prognostics and health monitoring. , 2017, , .		2
6274	Service Integrity Assurance for Distributed Computation Outsourcing. IEEE Transactions on Services Computing, 2017, , 1-1.	3.2	2
6275	dSpark: Deadline-Based Resource Allocation for Big Data Applications in Apache Spark. , 2017, , .		24
6276	Research on information security evaluation model of public institution. , 2017, , .		0

#	ARTICLE	IF	CITATIONS
6277	A RST-based stateful data analytics within spark. , 2017, , .		0
6278	Resource optimised workflow scheduling in Hadoop using stochastic hill climbing technique. IET Software, 2017, 11, 239-244.	1.5	6
6279	Parallel computing based distribution network reliability evaluation technology research. , 2017, , .		0
6280	Enabling stream processing for people-centric IoT based on the fog computing paradigm. , 2017, , .		9
6281	A novel approach for mining patterns from large uncertain data using MapReduce model. , 2017, , .		1
6282	GSC: Greedy shard caching algorithm for improved I/O efficiency in GraphChi. , 2017, , .		0
6283	Minimizing Data Access Latencies via Virtual Machine Placement Method in Datacenter. , 2017, , .		0
6284	Towards location-aware joint job and data assignment in cloud data centers with NVM. , 2017, , .		7
6285	Joint source selection and transfer optimization for erasure coding storage system. , 2017, , .		0
6286	Mitigate data skew caused stragglers through ImKP partition in MapReduce. , 2017, , .		3
6287	Addressing Memory Pressure in Data-intensive Parallel Programs via Container Based Virtualization. , 2017, , .		1
6288	Constructing gazetteers from volunteered Big Geo-Data based on Hadoop. Computers, Environment and Urban Systems, 2017, 61, 172-186.	3.3	110
6289	Parallel implementation of Kaufmanâ€™s initialization for clustering large remote sensing images on clouds. Computers, Environment and Urban Systems, 2017, 61, 153-162.	3.3	7
6290	MERRA Analytic Services: Meeting the Big Data challenges of climate science through cloud-enabled Climate Analytics-as-a-Service. Computers, Environment and Urban Systems, 2017, 61, 198-211.	3.3	81
6292	A Multi-Source Data Aggregation and Multidimensional Analysis Model for Big Data. ITM Web of Conferences, 2017, 12, 05009.	0.4	0
6293	HVTS: Hadoop-Based Video Transcoding System for Media Services. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2017, E100.A, 1248-1253.	0.2	2
6294	Apache REEF. ACM Transactions on Computer Systems, 2017, 35, 1-31.	0.6	12
6295	IncBricks. , 2017, , .		66

#	ARTICLE	IF	CITATIONS
6296	IncBricks. ACM SIGPLAN Notices, 2017, 52, 795-809.	0.2	2
6297	IncBricks. Computer Architecture News, 2017, 45, 795-809.	2.5	2
6298	Design and Realization of Intelligent Public-Transport Flow Monitoring System Based on Big Data. , 2017, , .		0
6299	Multi-node Big Data VM Platform and Job Submission Portlet. , 2017, , .		0
6300	TempestExtremes: a framework for scale-insensitive pointwise feature tracking on unstructured grids. Geoscientific Model Development, 2017, 10, 1069-1090.	1.3	144
6301	HC-Middleware. , 2017, , .		0
6302	Improving Triangle-Graph Based Text Summarization using Hybrid Similarity Function. Indian Journal of Science and Technology, 2017, 10, 1-15.	0.5	6
6303	Task replication and control for highly parallel in-memory stores. , 2017, , .		0
6304	Efficient Imbalanced Multimedia Concept Retrieval by Deep Learning on Spark Clusters. International Journal of Multimedia Data Engineering and Management, 2017, 8, 1-20.	0.3	22
6305	System architecture of coastal remote sensing data mining and services based on cloud computing. International Journal of Computational Science and Engineering, 2017, 15, 267.	0.4	0
6306	Advanced cloud fault tolerance system. IOP Conference Series: Materials Science and Engineering, 2017, 263, 042060.	0.3	0
6307	A strategy to load balancing for non-connectivity MapReduce job. IOP Conference Series: Materials Science and Engineering, 2017, 231, 012038.	0.3	0
6308	Optimization of Multiple Correlated Queries by Detecting Similar Data Source with Hadoop/Hive. Indian Journal of Science and Technology, 2017, 10, 1-8.	0.5	2
6309	Web Evolution - The Shift from Information Publishing to Reasoning. International Journal of Artificial Intelligence & Applications, 2017, 8, 11-28.	0.3	4
6310	Role of Non-Axiomatic Logic in a Distributed Reasoning Environment. Lecture Notes in Computer Science, 2017, , 381-388.	1.0	2
6311	Cyber Infrastructure as a Service to Empower Multidisciplinary, Data-Driven Scientific Research. International Journal of Computer Science and Information Technology, 2017, 9, 31-41.	0.3	1
6312	Automated long-term recording and analysis of neural activity in behaving animals. ELife, 2017, 6, .	2.8	122
6313	Data Driven Calibration Approach. SAE International Journal of Commercial Vehicles, 2017, 10, 353-359.	0.4	3

#	ARTICLE	IF	CITATIONS
6314	Distributed mixture-of-experts for Big Data using PETUUM framework. , 2017, , .		0
6315	Research on the Parallelization of the DBSCAN Clustering Algorithm for Spatial Data Mining Based on the Spark Platform. Remote Sensing, 2017, 9, 1301.	1.8	28
6316	System Framework for Cardiovascular Disease Prediction Based on Big Data Technology. Symmetry, 2017, 9, 293.	1.1	17
6317	Incremental algorithm based on wedge sampling for estimating clustering coefficient with MapReduce. , 2017, , .		2
6318	A scalable frequent user association mining algorithm using MapReduce. , 2017, , .		0
6319	DIMSpan. , 2017, , .		9
6320	Shadow: Exploiting the Power of Choice for Efficient Shuffling in MapReduce. , 2017, , .		0
6321	Unicorn: Unified resource orchestration for multi-domain, geo-distributed data analytics. , 2017, , .		3
6322	HALA: Hive advanced lexical analyzer. , 2017, , .		0
6323	MCSA: A multi-criteria shuffling algorithm for the MapReduce framework. , 2017, , .		0
6324	Large Scale Implementations for Twitter Sentiment Classification. Algorithms, 2017, 10, 33.	1.2	44
6325	Algorithm and Implementation of Distributed ESN Using Spark Framework and Parallel PSO. Applied Sciences (Switzerland), 2017, 7, 353.	1.3	7
6326	Blockchain-Empowered Fair Computational Resource Sharing System in the D2D Network. Future Internet, 2017, 9, 85.	2.4	24
6327	AHS Model: Efficient Topological Operators for a Sensor Web Publish/Subscribe System. ISPRS International Journal of Geo-Information, 2017, 6, 54.	1.4	5
6328	Big Data of the Past. Frontiers in Digital Humanities, 2017, 4, .	1.2	25
6329	Privacy-Preserving k -Means Clustering under Multiowner Setting in Distributed Cloud Environments. Security and Communication Networks, 2017, 1-19.	1.0	9
6330	Load rebalancing for Hadoop Distributed File System using distributed hash table. , 2017, , .		6
6331	Big data in mobile and pervasive computing. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
6332	Hadoop based short term traffic flow prediction on D ² using correlation model and KNN HSine. , 2017, , .		0
6333	An effective anonymization technique of big data using suppression slicing method. , 2017, , .		2
6334	A constraint programming-based resource allocation and scheduling of map reduce jobs with service level agreement. , 2017, , .		1
6335	ECLAT: Frequent itemset using MapReduce. , 2017, , .		2
6336	A Distributed K-Means Segmentation Algorithm Applied to Lobesia botrana Recognition. Complexity, 2017, 2017, 1-14.	0.9	51
6337	Scalable recommender system based on MapReduce framework. , 2017, , .		0
6338	A Novel Distributed Quantum-Behaved Particle Swarm Optimization. Journal of Optimization, 2017, 2017, 1-9.	6.0	9
6339	Towards Large-Scale, Heterogeneous Anomaly Detection Systems in Industrial Networks: A Survey of Current Trends. Security and Communication Networks, 2017, 2017, 1-17.	1.0	13
6340	A heuristic based algorithm for distance-aware virtual machine allocation in cloud. , 2017, , .		0
6341	Wireless Sensor Development for Malaysian Forest Monitoring and Tracking System. , 2017, , .		0
6342	Work-in-Progress: Towards Efficient and Scalable Big Data Analytics: Mapreduce vs. RDD™s. , 2017, , .		3
6343	Seamless service migration for human-centered computing in future Internet architecture. , 2017, , .		1
6344	A fourier spectral pattern analysis to design credit scoring models. , 2017, , .		5
6345	Data Science Tools and Techniques to Support Data Analytics in Transportation Applications. , 2017, , 69-90.		0
6346	Real- Time Mining Techniques: A Big Data Perspective for a Smart Future. Indian Journal of Science and Technology, 2017, 10, 1-7.	0.5	5
6347	Towards a distributed computing framework for Fog. , 2017, , .		14
6348	Social-sensed multimedia computing. , 2017, , 137-157.		1
6349	Bloomfish: A Highly Scalable Distributed K-mer Counting Framework. , 2017, , .		12

#	ARTICLE	IF	CITATIONS
6350	Costs of task allocation with local feedback: Effects of colony size and extra workers in social insects and other multi-agent systems. <i>PLoS Computational Biology</i> , 2017, 13, e1005904.	1.5	3
6351	ML-NA: A Machine Learning Based Node Performance Analyzer Utilizing Straggler Statistics. , 2017, , .		4
6352	Estimating Clustering Coefficient via Random Walk on MapReduce. , 2017, , .		1
6353	iSpot: Achieving Predictable Performance for Big Data Analytics with Cloud Transient Servers. , 2017, , .		1
6354	TDAG: A Tunable Distributed Data Processing Model for Data Stream. , 2017, , .		0
6355	Dynamic Load Balancing and Channel Strategy for Apache Flume Collecting Real-Time Data Stream. , 2017, , .		4
6356	Detecting Anomalies from End-to-End Internet Performance Measurements (PingER) Using Cluster Based Local Outlier Factor. , 2017, , .		5
6357	Scalable non-negative matrix tri-factorization. <i>BioData Mining</i> , 2017, 10, 41.	2.2	7
6358	MapReduce particle filtering with exact resampling and deterministic runtime. <i>Eurasip Journal on Advances in Signal Processing</i> , 2017, 2017, 71.	1.0	5
6359	Efficient resource allocation and dimensioning of media edge clouds infrastructure. <i>Journal of Cloud Computing: Advances, Systems and Applications</i> , 2017, 6, .	2.1	4
6360	Large-scale distributed L-BFGS. <i>Journal of Big Data</i> , 2017, 4, .	6.9	22
6361	Meeting deadlines for approximation processing in MapReduce environments. <i>Frontiers of Information Technology and Electronic Engineering</i> , 2017, 18, 1754-1772.	1.5	0
6362	An Improvement of Distributed Indexing Method Based on the Map/Reduce. , 2017, , .		0
6363	Mapreduce model for finding closely knit communities in large scale networks. , 2017, , .		5
6364	DSC: Dynamic stripe construction for asynchronous encoding in clustered file system. , 2017, , .		11
6365	Optimizing stochastic scheduling in fork-join queueing models: Bounds and applications. , 2017, , .		7
6366	SMiPE: Estimating the Progress of Recurring Iterative Distributed Dataflows. , 2017, , .		4
6367	Time-Saving First: Coflow Scheduling for Datacenter Networks. , 2017, , .		2

#	ARTICLE	IF	CITATIONS
6368	A Framework for Efficient Energy Scheduling of Spark Workloads. , 2017, , .		10
6369	Transparent heterogeneous backing store for file systems. , 2017, , .		0
6370	Research on mixed tasks scheduling in YARN. , 2017, , .		0
6371	A Scalable Spark-Based Fault Diagnosis Platform for Gearbox Fault Diagnosis in Wind Farms. , 2017, , .		10
6372	Characterizing and Improving the Performance of Many-Core Task-Based Parallel Programming Runtimes. , 2017, , .		1
6373	MURS: Mitigating Memory Pressure in Service-Oriented Data Processing System. , 2017, , .		2
6374	Standards-Based Function Shipping - How to Use TOSCA for Shipping and Executing Data Analytics Software in Remote Manufacturing Environments. , 2017, , .		5
6375	A scalable evolutionary linguistic fuzzy system with adaptive defuzzification in big data. , 2017, , .		11
6376	Seflow: Efficient Flow Scheduling for Data-Parallel Jobs. , 2017, , .		0
6377	The high dynamic reconstruction of high-resolution image based on MapReduce. , 2017, , .		1
6378	Mimir: Memory-Efficient and Scalable MapReduce for Large Supercomputing Systems. , 2017, , .		19
6379	ExtraV. Proceedings of the VLDB Endowment, 2017, 10, 1706-1717.	2.1	43
6380	Addressing job processing variability through redundant execution and opportunistic checkpointing: A competitive analysis. , 2017, , .		3
6381	Assessing the quality of a Consensus determined using a multi-level approach. , 2017, , .		4
6382	Towards Designing Effective Data Persistence through Tradeoff Space Analysis. , 2017, , .		3
6383	Big Data Storage and Management: Challenges and Opportunities. IFIP Advances in Information and Communication Technology, 2017, , 28-38.	0.5	4
6385	Software-defined storage for fast trajectory queries using a deltaFS indexed massive directory. , 2017, , .		6
6386	Putting the Pieces Back Together Again: Contest Webs for Large-Scale Problem Solving. SSRN Electronic Journal, 0, , .	0.4	5

#	ARTICLE	IF	CITATIONS
6387	Resources Management and Performance Analysis in Datacenter Networks. , 2017, , .		0
6388	Cross-Checking Multiple Data Sources Using Multiway Join in MapReduce. Scientific Programming, 2017, 2017, 1-9.	0.5	0
6389	Adaptive Online Hyper-Parameters Tuning for Ad Event-Prediction Models. , 2017, , .		11
6390	Walmart's Sales Data Analysis - A Big Data Analytics Perspective. , 2017, , .		13
6391	Applying of Decision Tree Analysis to Risk Factors Associated with Pressure Ulcers in Long-Term Care Facilities. Healthcare Informatics Research, 2017, 23, 43.	1.0	27
6392	SDTCP: Towards Datacenter TCP Congestion Control with SDN for IoT Applications. Sensors, 2017, 17, 109.	2.1	33
6393	Implication of Restful Web Services in Bigdata Analytics. , 2017, , .		1
6394	Accelerating Relevance Vector Machine for Large-Scale Data on Spark. ITM Web of Conferences, 2017, 12, 01024.	0.4	1
6395	Designing Low-Complexity Heavy-Traffic Delay-Optimal Load Balancing Schemes. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2017, 1, 1-30.	1.4	29
6396	Multimodal analysis of free-standing conversational groups. , 2017, , 51-74.		0
6397	Efficient similarity search. , 2017, , 105-134.		3
6399	Frequent Itemset Mining Using Improved Apriori Algorithm with MapReduce. , 2017, , .		2
6401	A Fast Synthetic Aperture Radar Raw Data Simulation Using Cloud Computing. Sensors, 2017, 17, 113.	2.1	9
6402	From Citation Network to Study Map. , 2017, , .		6
6403	Invulnerability study of transit network based on complex network. AIP Conference Proceedings, 2017, , .	0.3	0
6405	NS3 Based HDFS Data Placement Algorithm Evaluation Framework. , 2017, , .		1
6406	A Review on Job Scheduling for Hadoop Mapreduce. , 2017, , .		2
6407	Asymptotically Optimal Approximation Algorithms for Coflow Scheduling. , 2017, , .		10

#	ARTICLE	IF	CITATIONS
6408	Utilizing implicit user cues for multimedia analytics. , 2017, , 219-251.		0
6409	Parallel Distributed Patterns Mining Using Hadoop MapReduce Framework. International Journal of Grid and High Performance Computing, 2017, 9, 70-85.	0.7	10
6410	An Insight into State-of-the-Art Techniques for Big Data Classification. International Journal of Information System Modeling and Design, 2017, 8, 24-42.	0.9	0
6411	Massive data processing for effective trustworthiness modeling. , 2017, , 83-104.		0
6412	Frequent Itemset Mining in Large Datasets a Survey. International Journal of Information Retrieval Research, 2017, 7, 37-49.	0.6	0
6413	Personalized Online Analytical Processing in Big Data Context Using User Profile and Search Context. International Journal of Strategic Information Technology and Applications, 2017, 8, 67-80.	0.6	1
6414	Aras. International Journal of Distributed Systems and Technologies, 2017, 8, 47-60.	0.6	5
6415	Towards a New Model of Storage and Access to Data in Big Data and Cloud Computing. International Journal of Ambient Computing and Intelligence, 2017, 8, 31-44.	0.8	38
6416	Situation recognition using multimodal data. , 2017, , 159-189.		0
6417	GPUMap. , 2017, , .		1
6418	QFrag. , 2017, , .		21
6419	Efficient Implementation of Hadoop MapReduce based Business Process Dataflow. International Journal of Decision Support System Technology, 2017, 9, 49-60.	0.4	4
6420	Distributed Top-K Join Queries Optimizing for RDF Datasets. International Journal of Web Services Research, 2017, 14, 67-83.	0.5	4
6421	Efficient Skew Handling for Outer Joins in a Cloud Computing Environment. IEEE Transactions on Cloud Computing, 2018, 6, 558-571.	3.1	15
6422	Data Transfer Scheduling for Maximizing Throughput of Big-Data Computing in Cloud Systems. IEEE Transactions on Cloud Computing, 2018, 6, 87-98.	3.1	11
6423	MapReduce Scheduling for Deadline-Constrained Jobs in Heterogeneous Cloud Computing Systems. IEEE Transactions on Cloud Computing, 2018, 6, 127-140.	3.1	44
6424	Towards Privacy Preserving Publishing of Set-Valued Data on Hybrid Cloud. IEEE Transactions on Cloud Computing, 2018, 6, 316-329.	3.1	27
6425	On Optimal and Fair Service Allocation in Mobile Cloud Computing. IEEE Transactions on Cloud Computing, 2018, 6, 815-828.	3.1	28

#	ARTICLE	IF	CITATIONS
6426	Dependency-Aware Data Locality for MapReduce. IEEE Transactions on Cloud Computing, 2018, 6, 667-679.	3.1	9
6427	Parallel Semi-Supervised Multi-Ant Colonies Clustering Ensemble Based on MapReduce Methodology. IEEE Transactions on Cloud Computing, 2018, 6, 857-867.	3.1	12
6428	Achieving Energy Efficiency in Data Centers Using an Artificial Intelligence Abstraction Model. IEEE Transactions on Cloud Computing, 2018, 6, 612-624.	3.1	12
6429	Collective Energy-Efficiency Approach to Data Center Networks Planning. IEEE Transactions on Cloud Computing, 2018, 6, 656-666.	3.1	15
6430	EPAS: A Sampling Based Similarity Identification Algorithm for the Cloud. IEEE Transactions on Cloud Computing, 2018, 6, 720-733.	3.1	4
6431	H2Hadoop: Improving Hadoop Performance Using the Metadata of Related Jobs. IEEE Transactions on Cloud Computing, 2018, 6, 1031-1040.	3.1	15
6432	Datacenter Traffic Control: Understanding Techniques and Tradeoffs. IEEE Communications Surveys and Tutorials, 2018, 20, 1492-1525.	24.8	106
6433	Energy Efficient Big Data Networks: Impact of Volume and Variety. IEEE Transactions on Network and Service Management, 2018, 15, 458-474.	3.2	60
6434	Anomaly Detection Techniques Based on Kappa-Pruned Ensembles. IEEE Transactions on Reliability, 2018, 67, 212-229.	3.5	26
6435	Non-Asymptotic Delay Bounds for Multi-Server Systems with Synchronization Constraints. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1545-1559.	4.0	15
6436	TPTVer: A trusted third party based trusted verifier for multi-layered outsourced big data system in cloud environment. China Communications, 2018, 15, 122-137.	2.0	12
6437	GNCâ€‘app: A new Cytoscape app to rate gene networks biological coherence using geneâ€‘gene indirect relationships. BioSystems, 2018, 166, 61-65.	0.9	14
6438	Design and Implementation of Virtual Hadoop Cluster on Private Cloud. Communications in Computer and Information Science, 2018, , 61-71.	0.4	0
6439	iHOME: Index-Based JOIN Query Optimization for Limited Big Data Storage. Journal of Grid Computing, 2018, 16, 345-380.	2.5	13
6440	Machine learning and systems for building the next generation of EDA tools. , 2018, , .		12
6441	Are We Ready to Drive Software-Defined Networks? A Comprehensive Survey on Management Tools and Techniques. ACM Computing Surveys, 2019, 51, 1-35.	16.1	27
6442	SelfSplit parallelization for mixed-integer linear programming. Computers and Operations Research, 2018, 93, 101-112.	2.4	1
6443	A New Processing Approach for Reducing Computational Complexity in Cloud-RAN Mobile Networks. IEEE Access, 2018, 6, 6927-6946.	2.6	5

#	ARTICLE	IF	CITATIONS
6444	Data lake: a new ideology in big data era. ITM Web of Conferences, 2018, 17, 03025.	0.4	64
6445	Occopus: a Multi-Cloud Orchestrator to Deploy and Manage Complex Scientific Infrastructures. Journal of Grid Computing, 2018, 16, 19-37.	2.5	37
6446	Real-Time Probabilistic Data Fusion for Large-Scale IoT Applications. IEEE Access, 2018, 6, 10015-10027.	2.6	61
6447	Entropy4Cloud: Using Entropy-Based Complexity to Optimize Cloud Service Resource Management. IEEE Transactions on Emerging Topics in Computational Intelligence, 2018, 2, 13-24.	3.4	7
6448	FML-kNN: scalable machine learning on Big Data using k-nearest neighbor joins. Journal of Big Data, 2018, 5, .	6.9	22
6449	A New Data Layout Scheme for Energy-Efficient MapReduce Processing Tasks. Journal of Grid Computing, 2018, 16, 285-298.	2.5	7
6450	Design of material management system of mining group based on Hadoop. IOP Conference Series: Earth and Environmental Science, 2018, 108, 042052.	0.2	1
6451	Efficient data distribution and results merging for parallel data clustering in mapreduce environment. Applied Intelligence, 2018, 48, 2408-2428.	3.3	4
6452	GPUhd. , 2018, , .		4
6453	VM Deployment Methods for DaaS Model in Clouds. Lecture Notes on Data Engineering and Communications Technologies, 2018, , 371-382.	0.5	0
6454	BIGMiner: a fast and scalable distributed frequent pattern miner for big data. Cluster Computing, 2018, 21, 1507-1520.	3.5	30
6455	A power load forecast approach based on spatial-temporal clustering of load data. Concurrency Computation Practice and Experience, 2018, 30, e4386.	1.4	9
6456	Managing big RDF data in clouds: Challenges, opportunities, and solutions. Sustainable Cities and Society, 2018, 39, 375-386.	5.1	29
6457	A Big Data smart library recommender system for an educational institution. Library Hi Tech, 2018, 36, 498-523.	3.7	47
6458	SIMPO. Transactions on Architecture and Code Optimization, 2018, 15, 1-28.	1.6	9
6459	Data-Based Fault Diagnosis Model Using a Bayesian Causal Analysis Framework. International Journal of Information Technology and Decision Making, 2018, 17, 583-620.	2.3	8
6460	A CAM-Free Exascalable HPC Router for Low-Energy Communications. Lecture Notes in Computer Science, 2018, , 99-111.	1.0	6
6461	A novel adaptive e-learning model based on Big Data by using competence-based knowledge and social learner activities. Applied Soft Computing Journal, 2018, 69, 14-32.	4.1	53

#	ARTICLE	IF	CITATIONS
6462	BDEv 3.0: Energy efficiency and microarchitectural characterization of Big Data processing frameworks. <i>Future Generation Computer Systems</i> , 2018, 86, 565-581.	4.9	22
6463	An experimental survey on big data frameworks. <i>Future Generation Computer Systems</i> , 2018, 86, 546-564.	4.9	91
6464	An Approach to Improve Load Balancing in Distributed Storage Systems for NoSQL Databases: MongoDB. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 529-538.	0.5	1
6465	A Survey on NoSQL Stores. <i>ACM Computing Surveys</i> , 2019, 51, 1-43.	16.1	129
6466	Investigation of Replication Factor for Performance Enhancement in the Hadoop Distributed File System. , 2018, , .		7
6467	Protecting Personal Data with Blockchain Technology. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 119-125.	0.5	1
6468	Opportunities and obstacles for deep learning in biology and medicine. <i>Journal of the Royal Society Interface</i> , 2018, 15, 20170387.	1.5	1,282
6469	Revisiting Scalable Coherent Shared Memory. <i>Computer</i> , 2018, 51, 40-49.	1.2	5
6470	MapReduce Approach to Build Network User Profiles with Top-k Rankings for Network Security. , 2018, , .		1
6471	High-Performance Architecture Using Fast Dynamic Reconfigurable Accelerators. <i>IEEE Transactions on Very Large Scale Integration (VLSI) Systems</i> , 2018, 26, 1209-1222.	2.1	1
6472	Detecting malicious domain names using deep learning approaches at scale. <i>Journal of Intelligent and Fuzzy Systems</i> , 2018, 34, 1355-1367.	0.8	73
6473	MapReduce functions to remote sensing distributed data processingâ€”Global vegetation drought monitoring as example. <i>Software - Practice and Experience</i> , 2018, 48, 1352-1367.	2.5	21
6474	A multi-objective decision support framework for virtual machine placement in cloud data centers: a real case study. <i>Journal of Supercomputing</i> , 2018, 74, 2984-3015.	2.4	28
6475	Effects of standardization on the evolution of information and communications technology. <i>Technological Forecasting and Social Change</i> , 2018, 132, 308-317.	6.2	16
6476	Towards a compiler analysis for parallel algorithmic skeletons. , 2018, , .		3
6477	The Research of Big Data Analysis Platform of Oil & Gas Production. , 2018, , .		2
6478	Unifying computing resources and access interface to support parallel and distributed computing education. <i>Journal of Parallel and Distributed Computing</i> , 2018, 118, 201-212.	2.7	5
6479	Review of social media analytics process and Big Data pipeline. <i>Social Network Analysis and Mining</i> , 2018, 8, 1.	1.9	36

#	ARTICLE	IF	CITATIONS
6480	Analysis of Mahout Big Data Clustering Algorithms. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 999-1008.	0.5	4
6481	Enhancing in-memory efficiency for MapReduce-based data processing. <i>Journal of Parallel and Distributed Computing</i> , 2018, 120, 323-338.	2.7	6
6482	Distributing relational model transformation on MapReduce. <i>Journal of Systems and Software</i> , 2018, 142, 1-20.	3.3	10
6483	Engineering Resilient Collective Adaptive Systems by Self-Stabilisation. <i>ACM Transactions on Modeling and Computer Simulation</i> , 2018, 28, 1-28.	0.6	68
6484	Classification of Customer Tweets Using Big Data Analytics. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 169-180.	0.5	2
6485	Evaluation of MapReduce-Based Distributed Parallel Machine Learning Algorithms. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 101-111.	0.5	5
6486	Big data handling mechanisms in the healthcare applications: A comprehensive and systematic literature review. <i>Journal of Biomedical Informatics</i> , 2018, 82, 47-62.	2.5	76
6487	Modular design of a factor-graph-based inference engine on a System-On-Chip (SoC). <i>Microprocessors and Microsystems</i> , 2018, 60, 53-64.	1.8	1
6488	Handling distributed XML queries over large XML data based on MapReduce framework. <i>Information Sciences</i> , 2018, 453, 1-20.	4.0	12
6489	Components of Artificial Intelligence and Data Analytics. , 2018, , 101-148.		4
6490	A Novel Network Structure with Power Efficiency and High Availability for Data Centers. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2018, 29, 254-268.	4.0	16
6491	Learning communication patterns for malware discovery in HTTPs data. <i>Expert Systems With Applications</i> , 2018, 101, 129-142.	4.4	20
6492	Fuzzy integral-based ELM ensemble for imbalanced big data classification. <i>Soft Computing</i> , 2018, 22, 3519-3531.	2.1	43
6493	Efficient parallel algorithm for computing rough set approximation on GPU. <i>Soft Computing</i> , 2018, 22, 7553-7569.	2.1	7
6494	CHI-PG: A fast prototype generation algorithm for Big Data classification problems. <i>Neurocomputing</i> , 2018, 287, 22-33.	3.5	16
6496	Merging weighted SVMs for parallel incremental learning. <i>Neural Networks</i> , 2018, 100, 25-38.	3.3	13
6497	Competitive Partial Computation Offloading for Maximizing Energy Efficiency in Mobile Cloud Computing. <i>IEEE Access</i> , 2018, 6, 899-912.	2.6	25
6498	Ensembles of Lasso Screening Rules. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2018, 40, 2841-2852.	9.7	12

#	ARTICLE	IF	CITATIONS
6499	Detection and Mitigation of Congestion in SDN Enabled Data Center Networks: A Survey. IEEE Access, 2018, 6, 1730-1740.	2.6	38
6500	Harnessing the Power of HPC in Simulation and Optimization of Large Transportation Networks: Spatio-Temporal Traffic Management in the Greater Toronto Area. IEEE Intelligent Transportation Systems Magazine, 2018, 10, 95-106.	2.6	11
6501	Achieving Load Balance for Parallel Data Access on Distributed File Systems. IEEE Transactions on Computers, 2018, 67, 388-402.	2.4	24
6502	Scheduling Data-Intensive Workloads in Large-Scale Distributed Systems: Trends and Challenges. Studies in Big Data, 2018, , 19-43.	0.8	15
6503	Practical Verifiable Computation—A MapReduce Case Study. IEEE Transactions on Information Forensics and Security, 2018, 13, 1376-1391.	4.5	6
6504	BDWatchdog: Real-time monitoring and profiling of Big Data applications and frameworks. Future Generation Computer Systems, 2018, 87, 420-437.	4.9	10
6505	A survey towards an integration of big data analytics to big insights for value-creation. Information Processing and Management, 2018, 54, 758-790.	5.4	267
6506	GfLink: An In-Memory Computing Architecture on Heterogeneous CPU-GPU Clusters for Big Data. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1275-1288.	4.0	80
6507	Learning-Based Memory Allocation Optimization for Delay-Sensitive Big Data Processing. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1332-1341.	4.0	13
6508	Framework for Fast and Efficient Cloud Video Transcoding System Using Intelligent Splitter and Hadoop MapReduce. Wireless Personal Communications, 2018, 102, 2117-2132.	1.8	6
6509	The role of Information and Communication Technologies in healthcare: taxonomies, perspectives, and challenges. Journal of Network and Computer Applications, 2018, 107, 125-154.	5.8	256
6510	An Enhanced and Efficient Approach for Improving the Performance of HPC Environment Using Map-Reduce With MARIANE. Materials Today: Proceedings, 2018, 5, 1980-1988.	0.9	1
6511	LDPC Code Design for Distributed Storage: Balancing Repair Bandwidth, Reliability, and Storage Overhead. IEEE Transactions on Communications, 2018, 66, 507-520.	4.9	14
6512	A Scalable Multi-Granular Data Model for Data Parallel Workflows. , 2018, , .		1
6513	Cloud computing for genomic data analysis and collaboration. Nature Reviews Genetics, 2018, 19, 208-219.	7.7	205
6514	VOLAP: A Scalable Distributed Real-Time OLAP System for High-Velocity Data. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 226-239.	4.0	3
6515	Distributed ReliefF-based feature selection in Spark. Knowledge and Information Systems, 2018, 57, 1-20.	2.1	59
6516	Big Data Analysis-Based Secure Cluster Management for Optimized Control Plane in Software-Defined Networks. IEEE Transactions on Network and Service Management, 2018, 15, 27-38.	3.2	168

#	ARTICLE	IF	CITATIONS
6517	Molecular barcoding of viral vectors enables mapping and optimization of mRNA<i>trans</i>-splicing. Rna, 2018, 24, 673-687.	1.6	5
6518	Scalable Methods for Measuring the Connectivity and Quality of Large Numbers of Linked Datasets. Journal of Data and Information Quality, 2017, 9, 1-49.	1.5	14
6519	A fast rank mutual information based decision tree and its implementation via Map–Reduce. Concurrency Computation Practice and Experience, 2018, 30, e4387.	1.4	3
6520	Task Scheduling for Processing Big Graphs in Heterogeneous Commodity Clusters. Communications in Computer and Information Science, 2018, , 235-249.	0.4	0
6521	Efficient hybrid multicast approach in wireless data center network. Future Generation Computer Systems, 2018, 83, 27-36.	4.9	11
6522	Big Data Analytics in Uncertainty Quantification: Application to Structural Diagnosis and Prognosis. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 2018, 4, .	1.1	10
6523	A Study on Garbage Collection Algorithms for Big Data Environments. ACM Computing Surveys, 2019, 51, 1-35.	16.1	22
6524	EC-DNN: A new method for parallel training of deep neural networks. Neurocomputing, 2018, 287, 118-127.	3.5	3
6525	Fog Assisted-IoT Enabled Patient Health Monitoring in Smart Homes. IEEE Internet of Things Journal, 2018, 5, 1789-1796.	5.5	332
6526	Fast and Scalable Distributed Loopy Belief Propagation on Real-World Graphs. , 2018, , .		2
6527	Scalable processing and autocovariance computation of big functional data. Software - Practice and Experience, 2018, 48, 123-140.	2.5	2
6528	Collaborative matrix factorization mechanism for group recommendation in big data-based library systems. Library Hi Tech, 2018, 36, 458-481.	3.7	16
6529	sSketch. , 2018, , .		2
6530	Development and Application of Big Data Platform for “Bohai granary“. Wireless Personal Communications, 2018, 103, 275-293.	1.8	5
6531	Manufacturing Systems at Scale with Big Data Streaming and Online Machine Learning. Studies in Computational Intelligence, 2018, , 253-264.	0.7	12
6532	DFuzzy: a deep learning-based fuzzy clustering model for large graphs. Knowledge and Information Systems, 2018, 57, 159-181.	2.1	37
6533	Adaptive Speculation for Efficient Internetwork Application Execution in Clouds. ACM Transactions on Internet Technology, 2018, 18, 1-22.	3.0	11
6534	Systematic Review of Clustering High-Dimensional and Large Datasets. ACM Transactions on Knowledge Discovery From Data, 2018, 12, 1-68.	2.5	50

#	ARTICLE	IF	CITATIONS
6535	Augmentation Techniques for Mobile Cloud Computing. ACM Computing Surveys, 2019, 51, 1-38.	16.1	67
6536	Big Data Science. KI - Kunstliche Intelligenz, 2018, 32, 27-36.	2.2	0
6537	An efficient TCTL model checking algorithm and a reduction technique for verification of timed actor models. Science of Computer Programming, 2018, 153, 1-29.	1.5	13
6538	Parallel Computing Algorithms for Bigdata Frequent Pattern Mining. Lecture Notes on Data Engineering and Communications Technologies, 2018, , 113-123.	0.5	0
6539	Iteration acceleration for distributed learning systems. Parallel Computing, 2018, 72, 29-41.	1.3	4
6540	A scalable parallel framework for microstructure analysis of large-scale molecular dynamics simulations data. Computational Materials Science, 2018, 144, 322-330.	1.4	3
6541	A survey and taxonomy of resource optimisation for executing bag-of-task applications on public clouds. Future Generation Computer Systems, 2018, 82, 1-11.	4.9	42
6542	The Virtual Savant: Automatic generation of parallel solvers. Information Sciences, 2018, 432, 411-430.	4.0	8
6543	Distributed simulation optimization and parameter exploration framework for the cloud. Simulation Modelling Practice and Theory, 2018, 83, 108-123.	2.2	10
6544	Analyzing large scale genomic data on the cloud with Sparkhit. Bioinformatics, 2018, 34, 1457-1465.	1.8	14
6545	Effect of Live Migration on Virtual Hadoop Cluster. Lecture Notes in Computer Science, 2018, , 152-157.	1.0	0
6546	Efficient Algorithms for Local Density Based Anomaly Detection. Lecture Notes in Computer Science, 2018, , 336-342.	1.0	1
6547	HPC Cloud for Scientific and Business Applications. ACM Computing Surveys, 2019, 51, 1-29.	16.1	93
6548	A distributed approach to compliance monitoring of business process event streams. Future Generation Computer Systems, 2018, 82, 104-118.	4.9	15
6549	IncPregel: an incremental graph parallel computation model. Frontiers of Computer Science, 2018, 12, 1076-1089.	1.6	4
6550	A Pearson's correlation coefficient based decision tree and its parallel implementation. Information Sciences, 2018, 435, 40-58.	4.0	191
6551	A novel parallel community detection scheme based on label propagation. World Wide Web, 2018, 21, 1377-1398.	2.7	10
6552	Understanding and Combating Memory Bloat in Managed Data-Intensive Systems. ACM Transactions on Software Engineering and Methodology, 2017, 26, 1-41.	4.8	20

#	ARTICLE	IF	CITATIONS
6553	Big Data Challenges and Data Aggregation Strategies in Wireless Sensor Networks. IEEE Access, 2018, 6, 20558-20571.	2.6	109
6554	Alignment-free clustering of large data sets of unannotated protein conserved regions using minhashing. BMC Bioinformatics, 2018, 19, 83.	1.2	10
6555	Efficient iterative virtual screening with Apache Spark and conformal prediction. Journal of Cheminformatics, 2018, 10, 8.	2.8	24
6556	Toward Smarter Hadoop's Slaves Nodes by Deploying Game Theory Strategies. Lecture Notes in Networks and Systems, 2018, , 263-268.	0.5	0
6557	VM instance selection for deadline constraint job on agent-based interconnected cloud. Future Generation Computer Systems, 2018, 87, 470-487.	4.9	4
6558	High-Throughput Infrastructure for Advanced ITS Services: A Case Study on Air Pollution Monitoring. IEEE Transactions on Intelligent Transportation Systems, 2018, 19, 2246-2257.	4.7	8
6559	GB-PANDAS: Performance Evaluation Review, 2018, 45, 2-14.	0.4	19
6560	A query execution scheduling scheme for Impala system. Concurrency Computation Practice and Experience, 2018, 30, e4392.	1.4	1
6561	Fair bandwidth allocating and strip-aware prefetching for concurrent read streams and striped RAIDs in distributed file systems. Journal of Supercomputing, 2018, 74, 3904-3932.	2.4	2
6562	Distributed nearest neighbor classification for large-scale multi-label data on spark. Future Generation Computer Systems, 2018, 87, 66-82.	4.9	35
6563	An Improved Approximation for Scheduling Malleable Tasks with Precedence Constraints via Iterative Method. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1937-1946.	4.0	8
6564	Automatic translation from Java to Spark. Concurrency Computation Practice and Experience, 2018, 30, e4459.	1.4	1
6565	A framework for big data analytics approach to failure prediction of construction firms. Applied Computing and Informatics, 2018, 16, 207-222.	3.7	11
6566	Parallel swarm intelligence strategies for large-scale clustering based on MapReduce with application to epigenetics of aging. Applied Soft Computing Journal, 2018, 69, 771-783.	4.1	3
6567	FSCOAL-Parallel simultaneous fuzzy co-clustering and learning. International Journal of Intelligent Systems, 2018, 33, 1364-1380.	3.3	0
6568	Design and implementation of DeepDSL: A DSL for deep learning. Computer Languages, Systems and Structures, 2018, 54, 39-70.	1.4	6
6569	MRQAR: A generic MapReduce framework to discover quantitative association rules in big data problems. Knowledge-Based Systems, 2018, 153, 176-192.	4.0	35
6570	A Collaborative Scheduling-Based Parallel Solution for HEVC Encoding on Multicore Platforms. IEEE Transactions on Multimedia, 2018, 20, 2935-2948.	5.2	8

#	ARTICLE	IF	CITATIONS
6571	T-Brain: A Collaboration Platform for Data Scientists. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 295-309.	0.5	0
6572	A fast and low idle time method for mining frequent patterns in distributed and many-task computing environments. <i>Distributed and Parallel Databases</i> , 2018, 36, 613-641.	1.0	0
6573	A soft computing approach to big data summarization. <i>Fuzzy Sets and Systems</i> , 2018, 348, 4-20.	1.6	8
6574	An Adaptive Synchronous Parallel Strategy for Distributed Machine Learning. <i>IEEE Access</i> , 2018, 6, 19222-19230.	2.6	27
6575	Amdahl's Law in Big Data Analytics: Alive and Kicking in TPCx-BB (BigBench). , 2018, , .		6
6576	Open and scalable analytics of large Earth observation datasets: From scenes to multidimensional arrays using SciDB and GDAL. <i>ISPRS Journal of Photogrammetry and Remote Sensing</i> , 2018, 138, 47-56.	4.9	35
6577	Distributed Intrusion Detection System for Cloud Environments based on Data Mining techniques. <i>Procedia Computer Science</i> , 2018, 127, 35-41.	1.2	87
6578	Load Balancing in Data Center Networks: A Survey. <i>IEEE Communications Surveys and Tutorials</i> , 2018, 20, 2324-2352.	24.8	115
6579	FBSGraph: Accelerating Asynchronous Graph Processing via Forward and Backward Sweeping. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2018, 30, 895-907.	4.0	16
6580	Efficient Performance-Centric Bandwidth Allocation with Fairness Tradeoff. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2018, 29, 1693-1706.	4.0	6
6581	A productivity improvement of distributed software testing using checkpoint. , 2018, , .		0
6582	A productivity improvement of distributed software testing using checkpoint. , 2018, , .		1
6583	Capacity-aware key partitioning scheme for heterogeneous big data analytic engines. , 2018, , .		0
6584	BCDC: A High-Performance, Server-Centric Data Center Network. <i>Journal of Computer Science and Technology</i> , 2018, 33, 400-416.	0.9	32
6585	Improving the Effectiveness of Burst Buffers for Big Data Processing in HPC Systems with Eley. <i>Future Generation Computer Systems</i> , 2018, 86, 308-318.	4.9	7
6586	Principal Components Analysis Random Discretization Ensemble for Big Data. <i>Knowledge-Based Systems</i> , 2018, 150, 166-174.	4.0	32
6587	Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies, and Solutions. <i>IEEE Communications Surveys and Tutorials</i> , 2018, 20, 2429-2453.	24.8	726
6588	Industrial Cyberphysical Systems: Realizing Cloud-Based Big Data Infrastructures. <i>IEEE Industrial Electronics Magazine</i> , 2018, 12, 25-35.	2.3	73

#	ARTICLE	IF	CITATIONS
6589	MCL: A Cost-Efficient Nonblocking Multicast Interconnection Network. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 2046-2058.	4.0	5
6590	Data driven composite shape descriptor design for shape retrieval with a VoR-Tree. Applied Mathematics, 2018, 33, 88-106.	0.6	1
6591	Reasoning about attribute value equivalence in relational data. Information Systems, 2018, 75, 1-12.	2.4	2
6592	Performance Evaluation of Big Data Analysis. , 2018, , 1-6.		2
6593	Automatic Testing of Design Faults in MapReduce Applications. IEEE Transactions on Reliability, 2018, 67, 717-732.	3.5	13
6594	Building edge intelligence for online activity recognition in service-oriented IoT systems. Future Generation Computer Systems, 2018, 87, 557-567.	4.9	31
6595	MR-IoT: An information centric MapReduce framework for IoT. , 2018, , .		9
6596	Multiple Mode SAR Raw Data Simulation and Parallel Acceleration for Gaofen-3 Mission. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11, 2115-2126.	2.3	54
6597	Firework: Data Processing and Sharing for Hybrid Cloud-Edge Analytics. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 2004-2017.	4.0	64
6598	Finding and Counting Tree-Like Subgraphs Using MapReduce. IEEE Transactions on Multi-Scale Computing Systems, 2018, 4, 217-230.	2.5	2
6599	CoMan: Managing Bandwidth Across Computing Frameworks in Multiplexed Datacenters. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1013-1029.	4.0	8
6600	Double-layered big data analytics architecture for solar cells series welding machine. Computers in Industry, 2018, 97, 17-23.	5.7	8
6601	A programming model and foundation for lineage-based distributed computation. Journal of Functional Programming, 2018, 28, .	0.5	3
6602	Machine learning based generic violation waiver system with application on electromigration sign-off. , 2018, , .		6
6603	Modeling Service Applications for Optimal Parallel Embedding. IEEE Transactions on Cloud Computing, 2018, 6, 1067-1079.	3.1	10
6604	Modulo Based Data Placement Algorithm for Energy Consumption Optimization of MapReduce System. Journal of Grid Computing, 2018, 16, 409-424.	2.5	13
6605	A Workflow Management System for Scalable Data Mining on Clouds. IEEE Transactions on Services Computing, 2018, 11, 480-492.	3.2	78
6606	SVDC: A Highly Scalable Isolation Architecture for Virtualized Layer-2 Data Center Networks. IEEE Transactions on Cloud Computing, 2018, 6, 1178-1190.	3.1	1

#	ARTICLE	IF	CITATIONS
6607	Profit Maximization for Geographically Dispersed Green Data Centers. IEEE Transactions on Smart Grid, 2018, 9, 703-711.	6.2	47
6608	A Big Data Scale Algorithm for Optimal Scheduling of Integrated Microgrids. IEEE Transactions on Smart Grid, 2018, 9, 274-282.	6.2	36
6609	Feedback Autonomic Provisioning for Guaranteeing Performance in MapReduce Systems. IEEE Transactions on Cloud Computing, 2018, 6, 1004-1016.	3.1	16
6610	Optimizing Live Migration of Multiple Virtual Machines. IEEE Transactions on Cloud Computing, 2018, 6, 1096-1109.	3.1	24
6611	Accelerating MapReduce on Commodity Clusters: An SSD-Empowered Approach. IEEE Transactions on Big Data, 2018, 4, 396-407.	4.4	2
6612	An intermediate data placement algorithm for load balancing in Spark computing environment. Future Generation Computer Systems, 2018, 78, 287-301.	4.9	59
6613	MtMR: Ensuring MapReduce Computation Integrity with Merkle Tree-Based Verifications. IEEE Transactions on Big Data, 2018, 4, 418-431.	4.4	33
6614	pipsCloud: High performance cloud computing for remote sensing big data management and processing. Future Generation Computer Systems, 2018, 78, 353-368.	4.9	140
6615	On Distributed Fuzzy Decision Trees for Big Data. IEEE Transactions on Fuzzy Systems, 2018, 26, 174-192.	6.5	101
6616	Cloud computing-based parallel genetic algorithm for gene selection in cancer classification. Neural Computing and Applications, 2018, 30, 1601-1610.	3.2	23
6617	Mobile cloud computing for computation offloading: Issues and challenges. Applied Computing and Informatics, 2018, 14, 1-16.	3.7	148
6618	Prediction-Based and Locality-Aware Task Scheduling for Parallelizing Video Transcoding Over Heterogeneous MapReduce Cluster. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28, 1009-1020.	5.6	18
6619	Modeling SmallClient indexing framework for big data analytics. Journal of Supercomputing, 2018, 74, 5241-5262.	2.4	3
6620	Storage tier-aware replicative data reorganization with prioritization for efficient workload processing. Future Generation Computer Systems, 2018, 79, 618-629.	4.9	5
6621	Scalable Architecture for SDN Traffic Classification. IEEE Systems Journal, 2018, 12, 3203-3214.	2.9	31
6622	A Multicommodity Flow-Based Detailed Router With Efficient Acceleration Techniques. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2018, 37, 217-230.	1.9	17
6623	Using machine learning and big data approaches to predict travel time based on historical and real-time data from Taiwan electronic toll collection. Soft Computing, 2018, 22, 5707-5718.	2.1	56
6624	AllMDs: An Integrated Framework of Automatic Idiopathic Inflammatory Myopathy Diagnosis for Muscle. IEEE Journal of Biomedical and Health Informatics, 2018, 22, 942-954.	3.9	4

#	ARTICLE	IF	CITATIONS
6625	Replication-Based Fault-Tolerance for Large-Scale Graph Processing. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1621-1635.	4.0	10
6626	A fault-tolerant ONVIF protocol extension for seamless surveillance video stream recording. Computer Standards and Interfaces, 2018, 55, 55-72.	3.8	1
6627	JouleMR: Towards Cost-Effective and Green-Aware Data Processing Frameworks. IEEE Transactions on Big Data, 2018, 4, 258-272.	4.4	15
6628	Large-Scale Multimodality Attribute Reduction With Multi-Kernel Fuzzy Rough Sets. IEEE Transactions on Fuzzy Systems, 2018, 26, 226-238.	6.5	110
6629	Parallel knowledge acquisition algorithms for big data using MapReduce. International Journal of Machine Learning and Cybernetics, 2018, 9, 1007-1021.	2.3	15
6630	DPM: A novel distributed large-scale social graph processing framework for link prediction algorithms. Future Generation Computer Systems, 2018, 78, 474-480.	4.9	15
6631	SOR: An optimized semantic ontology retrieval algorithm for heterogeneous multimedia big data. Journal of Computational Science, 2018, 28, 455-465.	1.5	30
6632	A general-purpose framework for parallel processing of large-scale LiDAR data. International Journal of Digital Earth, 2018, 11, 26-47.	1.6	30
6633	Characterizing and diagnosing out of memory errors in MapReduce applications. Journal of Systems and Software, 2018, 137, 399-414.	3.3	6
6634	Multilevel Data Processing Using Parallel Algorithms for Analyzing Big Data in High-Performance Computing. International Journal of Parallel Programming, 2018, 46, 508-527.	1.1	32
6635	MCRS: A course recommendation system for MOOCs. Multimedia Tools and Applications, 2018, 77, 7051-7069.	2.6	69
6636	A glimpse on big data analytics in the framework of marketing strategies. Soft Computing, 2018, 22, 325-342.	2.1	82
6637	HDM: A Composable Framework for Big Data Processing. IEEE Transactions on Big Data, 2018, 4, 150-163.	4.4	12
6638	Towards cloud-based parallel metaheuristics. International Journal of High Performance Computing Applications, 2018, 32, 693-705.	2.4	10
6639	Improved performance optimization for massive small files in cloud computing environment. Annals of Operations Research, 2018, 265, 305-317.	2.6	5
6640	Towards Quality Aware Information Integration in Distributed Sensing Systems. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 198-211.	4.0	12
6641	New scheduling approach using reinforcement learning for heterogeneous distributed systems. Journal of Parallel and Distributed Computing, 2018, 117, 292-302.	2.7	108
6642	Internet of Things based activity surveillance of defence personnel. Journal of Ambient Intelligence and Humanized Computing, 2018, 9, 2061-2076.	3.3	20

#	ARTICLE	IF	CITATIONS
6643	RankMap: A Framework for Distributed Learning From Dense Data Sets. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29, 2717-2730.	7.2	1
6644	Distributed Kernel-Based Gradient Descent Algorithms. Constructive Approximation, 2018, 47, 249-276.	1.8	47
6645	Pornographic video detection with MapReduce. International Journal of Machine Learning and Cybernetics, 2018, 9, 2105-2115.	2.3	3
6646	Exploiting coarse-grained reused-based opportunities in Big Data multi-query optimization. Journal of Computational Science, 2018, 26, 432-452.	1.5	17
6647	Research and implementation of user clustering based on MapReduce in multimedia big data. Multimedia Tools and Applications, 2018, 77, 10017-10031.	2.6	13
6649	A Survey of Recent Trends in Testing Concurrent Software Systems. IEEE Transactions on Software Engineering, 2018, 44, 747-783.	4.3	38
6650	Train Delay Prediction Systems: A Big Data Analytics Perspective. Big Data Research, 2018, 11, 54-64.	2.6	85
6651	An Ingestion and Analytics Architecture for IoT Applied to Smart City Use Cases. IEEE Internet of Things Journal, 2018, 5, 765-774.	5.5	70
6652	Content-Aware Partial Compression for Textual Big Data Analysis in Hadoop. IEEE Transactions on Big Data, 2018, 4, 459-472.	4.4	3
6653	LPiTrack: Eye movement pattern recognition algorithm and application to biometric identification. Machine Learning, 2018, 107, 313-331.	3.4	10
6654	Angel: a new large-scale machine learning system. National Science Review, 2018, 5, 216-236.	4.6	53
6655	Implementation of scalable fuzzy relational operations in MapReduce. Soft Computing, 2018, 22, 3061-3075.	2.1	7
6656	Real-Time Big Data Stream Processing Using GPU with Spark Over Hadoop Ecosystem. International Journal of Parallel Programming, 2018, 46, 630-646.	1.1	67
6657	Using Hadoop MapReduce for Parallel Genetic Algorithms: A Comparison of the Global, Grid and Island Models. Evolutionary Computation, 2018, 26, 535-567.	2.3	30
6658	An extended Intelligent Water Drops algorithm for workflow scheduling in cloud computing environment. Egyptian Informatics Journal, 2018, 19, 33-55.	4.4	57
6659	A Fuzzy-Based Intelligent Cloud Broker with MapReduce Framework to Evaluate the Trust Level of Cloud Services Using Customer Feedback. International Journal of Fuzzy Systems, 2018, 20, 339-347.	2.3	19
6660	GEODIS: towards the optimization of data locality-aware job scheduling in geo-distributed data centers. Computing (Vienna/New York), 2018, 100, 21-46.	3.2	23
6661	Multi-criteria matrix localization and integration for personalized collaborative filtering in IoT environments. Multimedia Tools and Applications, 2018, 77, 4697-4730.	2.6	8

#	ARTICLE	IF	CITATIONS
6662	QoS-aware Big service composition using MapReduce based evolutionary algorithm with guided mutation. <i>Future Generation Computer Systems</i> , 2018, 86, 1008-1018.	4.9	52
6663	Benchmarking Big Data Systems: A Review. <i>IEEE Transactions on Services Computing</i> , 2018, 11, 580-597.	3.2	50
6664	AEGEUS++: an energy-aware online partition skew mitigation algorithm for mapreduce in cloud. <i>Cluster Computing</i> , 2018, 21, 1243-1260.	3.5	2
6665	A new decision-making approach for multiple criteria sorting with an imbalanced set of assignment examples. <i>European Journal of Operational Research</i> , 2018, 265, 598-620.	3.5	13
6666	Big data for open innovation in SMEs and large corporations: Trends, opportunities, and challenges. <i>Creativity and Innovation Management</i> , 2018, 27, 6-22.	1.9	145
6667	CHI-BD: A fuzzy rule-based classification system for Big Data classification problems. <i>Fuzzy Sets and Systems</i> , 2018, 348, 75-101.	1.6	60
6668	Using machine learning to optimize parallelism in big data applications. <i>Future Generation Computer Systems</i> , 2018, 86, 1076-1092.	4.9	57
6669	Deploying massive runs of evolutionary algorithms with ECJ and Hadoop: Reducing interest points required for face recognition. <i>International Journal of High Performance Computing Applications</i> , 2018, 32, 706-720.	2.4	3
6670	Performance optimization of MapReduce-based Apriori algorithm on Hadoop cluster. <i>Computers and Electrical Engineering</i> , 2018, 67, 348-364.	3.0	55
6671	Levering Mobile Cloud Computing for Mobile Big Data Analytics. <i>Lecture Notes on Data Engineering and Communications Technologies</i> , 2018, , 21-39.	0.5	4
6674	Effectively and Efficiently Supporting Encrypted OLAP Queries over Big Data: Models, Issues, Challenges. <i>Lecture Notes in Electrical Engineering</i> , 2018, , 329-336.	0.3	0
6675	Energy-harvesting based on internet of things and big data analytics for smart health monitoring. <i>Sustainable Computing: Informatics and Systems</i> , 2018, 20, 155-164.	1.6	31
6676	Recommender Systems for Large-Scale Social Networks: A review of challenges and solutions. <i>Future Generation Computer Systems</i> , 2018, 78, 413-418.	4.9	134
6677	Toward modeling and optimization of features selection in Big Data based social Internet of Things. <i>Future Generation Computer Systems</i> , 2018, 82, 715-726.	4.9	68
6678	A Big Data-as-a-Service Framework: State-of-the-Art and Perspectives. <i>IEEE Transactions on Big Data</i> , 2018, 4, 325-340.	4.4	136
6679	Apriori Versions Based on MapReduce for Mining Frequent Patterns on Big Data. <i>IEEE Transactions on Cybernetics</i> , 2018, 48, 2851-2865.	6.2	54
6680	PPS: Parallel Pincer Search for Mining Frequent Itemsets Based on Spark. <i>Advances in Intelligent Systems and Computing</i> , 2018, , 351-363.	0.5	0
6681	An approach to validity indices for clustering techniques in Big Data. <i>Progress in Artificial Intelligence</i> , 2018, 7, 81-94.	1.5	19

#	ARTICLE	IF	CITATIONS
6682	Using Map-Reduce for Image Analysis in Cloud Environment. Lecture Notes in Networks and Systems, 2018, , 109-118.	0.5	1
6683	Intelligent inversion method for pre-stack seismic big data based on MapReduce. Computers and Geosciences, 2018, 110, 81-89.	2.0	24
6684	Big Data: Tutorial and guidelines on information and process fusion for analytics algorithms with MapReduce. Information Fusion, 2018, 42, 51-61.	11.7	122
6685	High-Level Programming Abstractions for Distributed Graph Processing. IEEE Transactions on Knowledge and Data Engineering, 2018, 30, 305-324.	4.0	39
6686	A novel Bit Vector Product algorithm for mining frequent itemsets from large datasets using MapReduce framework. Cluster Computing, 2018, 21, 1365-1380.	3.5	2
6687	Dynamic frequency based parallel k-bat algorithm for massive data clustering (DFBPKBA). International Journal of Systems Assurance Engineering and Management, 2018, 9, 866-874.	1.5	25
6688	Mining massive hierarchical data using a scalable probabilistic graphical model. Information Sciences, 2018, 425, 62-75.	4.0	5
6689	Confluence: Speeding Up Iterative Distributed Operations by Key-Dependency-Aware Partitioning. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 351-364.	4.0	2
6690	Mobile Agent Based MapReduce Framework for Big Data Processing. Advances in Intelligent Systems and Computing, 2018, , 391-402.	0.5	1
6691	Scalable Aspect-Based Summarization in the Hadoop Environment. Advances in Intelligent Systems and Computing, 2018, , 439-449.	0.5	6
6692	Robust Fuzzy Neuro system for Big Data Analytics. Advances in Intelligent Systems and Computing, 2018, , 543-552.	0.5	4
6693	Big Data: Issues, Challenges, and Techniques in Business Intelligence. Advances in Intelligent Systems and Computing, 2018, , 613-628.	0.5	20
6694	Generalized Query Processing Mechanism in Cloud Database Management System. Advances in Intelligent Systems and Computing, 2018, , 641-648.	0.5	3
6695	Cloud-Based Parallel Machine Learning for Tool Wear Prediction. Journal of Manufacturing Science and Engineering, Transactions of the ASME, 2018, 140, .	1.3	46
6696	Review of Parallel Apriori Algorithm on MapReduce Framework for Performance Enhancement. Advances in Intelligent Systems and Computing, 2018, , 403-411.	0.5	6
6697	Learning with Big Data Technology: The Future of Education. Advances in Intelligent Systems and Computing, 2018, , 209-217.	0.5	10
6698	An efficient distributed protein disorder prediction with pasted samples. Computers and Electrical Engineering, 2018, 65, 342-356.	3.0	2
6699	Online MapReduce processing on two identical parallel machines. Journal of Combinatorial Optimization, 2018, 35, 216-223.	0.8	5

#	ARTICLE	IF	CITATIONS
6700	Optimization of stream-based live data migration strategy in the cloud. <i>Concurrency Computation Practice and Experience</i> , 2018, 30, e4293.	1.4	7
6701	Coping with demand volatility in retail pharmacies with the aid of big data exploration. <i>Computers and Operations Research</i> , 2018, 98, 343-354.	2.4	27
6702	Tractable queries on big data via preprocessing with logarithmic-size output. <i>Knowledge and Information Systems</i> , 2018, 56, 141-163.	2.1	2
6703	QoS-guaranteed resource provisioning for cloud-based MapReduce in dynamical environments. <i>Future Generation Computer Systems</i> , 2018, 78, 18-30.	4.9	9
6704	A general framework for real-time analysis of massive multimedia streams. <i>Multimedia Systems</i> , 2018, 24, 391-406.	3.0	14
6705	Routing and Traffic Engineering in Data Center Networks. , 2018, , 396-422.		1
6707	Privacy-preserving tabular data publishing: A comprehensive evaluation from web to cloud. <i>Computers and Security</i> , 2018, 72, 74-95.	4.0	24
6708	Academia versus social media: A psycho-linguistic analysis. <i>Journal of Computational Science</i> , 2018, 25, 228-237.	1.5	3
6709	Computing Crowd Consensus with Partial Agreement. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2018, 30, 1-14.	4.0	38
6710	Supervised Distributed Hashing for Large-Scale Multimedia Retrieval. <i>IEEE Transactions on Multimedia</i> , 2018, 20, 675-686.	5.2	33
6711	A Computationally Efficient Implementation of Tabulated Combustion Chemistry based on Polynomials and Automatic Source Code Generation. <i>Flow, Turbulence and Combustion</i> , 2018, 100, 119-146.	1.4	6
6712	The industrial management of SMEs in the era of Industry 4.0. <i>International Journal of Production Research</i> , 2018, 56, 1118-1136.	4.9	702
6713	A unifying analysis for the supervised descriptive rule discovery via the weighted relative accuracy. <i>Knowledge-Based Systems</i> , 2018, 139, 89-100.	4.0	20
6714	A Distributed Fuzzy Associative Classifier for Big Data. <i>IEEE Transactions on Cybernetics</i> , 2018, 48, 2656-2669.	6.2	37
6715	Game Theoretic Approaches in Mobile Cloud Computing Systems for Big Data Applications: A Systematic Literature Review. <i>Lecture Notes on Data Engineering and Communications Technologies</i> , 2018, , 41-62.	0.5	6
6716	Evaluating Decision Analytics from Mobile Big Data using Rough Set Based Ant Colony. <i>Lecture Notes on Data Engineering and Communications Technologies</i> , 2018, , 217-231.	0.5	2
6717	Balancing search space partitions by sparse coding for distributed redundant media indexing and retrieval. <i>International Journal of Multimedia Information Retrieval</i> , 2018, 7, 57-70.	3.6	3
6718	IoTDeM: An IoT Big Data-oriented MapReduce performance prediction extended model in multiple edge clouds. <i>Journal of Parallel and Distributed Computing</i> , 2018, 118, 316-327.	2.7	88

#	ARTICLE	IF	CITATIONS
6719	Performance prediction of parallel computing models to analyze cloud-based big data applications. Cluster Computing, 2018, 21, 1439-1454.	3.5	8
6720	FastSLAM 2.0 tracking and mapping as a Cloud Robotics service. Computers and Electrical Engineering, 2018, 69, 412-421.	3.0	9
6721	Informed Prefetching for Distributed Multi-Level Storage Systems. Journal of Signal Processing Systems, 2018, 90, 619-640.	1.4	7
6722	MetReS, an Efficient Database for Genomic Applications. Journal of Computational Biology, 2018, 25, 200-213.	0.8	1
6723	The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems With Applications, 2018, 97, 205-227.	4.4	420
6724	Energy-efficient hadoop for big data analytics and computing: A systematic review and research insights. Future Generation Computer Systems, 2018, 86, 1351-1367.	4.9	42
6725	Distributed data stream processing and edge computing: A survey on resource elasticity and future directions. Journal of Network and Computer Applications, 2018, 103, 1-17.	5.8	232
6726	Data-driven cost estimation for additive manufacturing in cybermanufacturing. Journal of Manufacturing Systems, 2018, 46, 115-126.	7.6	102
6727	From P2P to NoSQL: A continuous metric for classifying large-scale storage systems. Journal of Parallel and Distributed Computing, 2018, 113, 227-249.	2.7	6
6728	Aura: An incentive-driven ad-hoc IoT cloud framework for proximal mobile computation offloading. Future Generation Computer Systems, 2018, 86, 821-835.	4.9	38
6729	Data-Driven Voltage Unbalance Analysis in Power Distribution Networks. , 2018, , 321-342.		0
6730	Mobile 3D assembly process information construction and transfer to the assembly station of complex products. International Journal of Computer Integrated Manufacturing, 2018, 31, 11-26.	2.9	19
6731	Health Recommender System and Its Applicability with MapReduce Framework. Advances in Intelligent Systems and Computing, 2018, , 255-266.	0.5	0
6732	Context-Aware Computing, Learning, and Big Data in Internet of Things: A Survey. IEEE Internet of Things Journal, 2018, 5, 1-27.	5.5	299
6733	Replicable parallel branch and bound search. Journal of Parallel and Distributed Computing, 2018, 113, 92-114.	2.7	15
6734	EDAWS: A distributed framework with efficient data analytics workspace towards discriminative services for critical infrastructures. Future Generation Computer Systems, 2018, 81, 78-93.	4.9	3
6735	Data exfiltration: A review of external attack vectors and countermeasures. Journal of Network and Computer Applications, 2018, 101, 18-54.	5.8	72
6736	\$k\$ NN-DP: Handling Data Skewness in \$k\$NN\$ Joins Using MapReduce. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 600-613.	4.0	25

#	ARTICLE	IF	CITATIONS
6737	Effective Allocation of Resources and Task Scheduling in Cloud Environment using Social Group Optimization. <i>Arabian Journal for Science and Engineering</i> , 2018, 43, 4265-4272.	1.7	44
6738	Arc4nix: A cross-platform geospatial analytical library for cluster and cloud computing. <i>Computers and Geosciences</i> , 2018, 111, 159-166.	2.0	5
6739	Distributing extreme learning machines with Apache Spark for NetFlow-based malware activity detection. <i>Pattern Recognition Letters</i> , 2018, 101, 14-20.	2.6	19
6740	Migration from Relational Databases to HBase: A Feasibility Assessment. <i>Lecture Notes in Networks and Systems</i> , 2018, , 383-395.	0.5	1
6741	A hybrid MapReduce-based k-means clustering using genetic algorithm for distributed datasets. <i>Journal of Supercomputing</i> , 2018, 74, 1562-1579.	2.4	36
6743	Parallel Bat Algorithm-Based Clustering Using MapReduce. <i>Lecture Notes on Data Engineering and Communications Technologies</i> , 2018, , 73-82.	0.5	32
6744	Combining Hadoop with MPI to Solve Metagenomics Problems that are both Data- and Compute-intensive. <i>International Journal of Parallel Programming</i> , 2018, 46, 762-775.	1.1	3
6745	An efficient parallel processing method for skyline queries in MapReduce. <i>Journal of Supercomputing</i> , 2018, 74, 886-935.	2.4	8
6746	Power consumption minimization by distributive particle swarm optimization for luminance control and its parallel implementations. <i>Expert Systems With Applications</i> , 2018, 96, 479-491.	4.4	18
6747	Fair multi-agent task allocation for large datasets analysis. <i>Knowledge and Information Systems</i> , 2018, 54, 591-615.	2.1	9
6748	Formalised Composition and Interaction for Heterogeneous Structured Parallelism. <i>International Journal of Parallel Programming</i> , 2018, 46, 120-151.	1.1	1
6749	FTLLS: A fault tolerant, low latency, distributed scheduling approach based on sparrow. <i>Peer-to-Peer Networking and Applications</i> , 2018, 11, 1129-1140.	2.6	0
6751	Cost efficient scheduling of MapReduce applications on public clouds. <i>Journal of Computational Science</i> , 2018, 26, 375-388.	1.5	16
6752	Computing all minimal hitting sets by subset recombination. <i>Applied Intelligence</i> , 2018, 48, 257-270.	3.3	6
6753	Scalable system scheduling for HPC and big data. <i>Journal of Parallel and Distributed Computing</i> , 2018, 111, 76-92.	2.7	67
6754	Dynamic Capabilities of Decision-oriented Service Systems. <i>International Journal of Information Systems in the Service Sector</i> , 2018, 10, 41-63.	0.2	2
6755	Proportion Scheduler to Improve the Mismatched Locality in YARN. , 2018, , .		0
6756	Asynchronous Group-Based ADMM Algorithm under Efficient Communication Structure. , 2018, , .		2

#	ARTICLE	IF	CITATIONS
6757	ShenTu: Processing Multi-Trillion Edge Graphs on Millions of Cores in Seconds. , 2018, , .		42
6758	Bibliometrical Analysis of International Big Data Research: Based on Citespace and VOSviewer. , 2018, , .		18
6759	An On-the-Fly Scheduling Strategy for Distributed Stream Processing Platform. , 2018, , .		2
6760	A Proposed Technique for Conversion of Unstructured Agro-Data to Semi-Structured or Structured Data. , 2018, , .		5
6761	Leveraging Coding Techniques for Speeding up Distributed Computing. , 2018, , .		22
6762	Scheduling Coflows with Incomplete Information. , 2018, , .		4
6763	HPC4AI. , 2018, , .		24
6764	From Think Parallel to Think Sequential. SIGMOD Record, 2018, 47, 15-22.	0.7	0
6765	Query Processing “kNN. , 2018, , 1-5.		0
6766	Hadoop. , 2018, , 1-5.		0
6767	Comparison of the HPC and Big Data Java Libraries Spark, PCJ and APGAS. , 2018, , .		3
6768	Classification of Textual Data in Distributed Environment. , 2018, , .		0
6769	Defining a Sandbox for Responsible AI. SSRN Electronic Journal, 0, , .	0.4	0
6770	From Java to FPGA: An Experience with the Intel HARP System. , 2018, , .		6
6771	SDPaxos. , 2018, , .		15
6772	Optimizing Machine Learning on Apache Spark in HPC Environments. , 2018, , .		1
6773	Orpheus. , 2018, , .		5
6774	MatGap: A Systematic Approach to Perform Match and Gap Analysis among SBVR-Based Domain Specific Business Rules. , 2018, , .		2

#	ARTICLE	IF	CITATIONS
6775	An Efficient Distributed Approach on High Dimensional Data Similarity Searchable Encryption. , 2018, , .		0
6776	A Novel Solution for Scaling Video Shot Boundary Detection Based on Hadoop. International Journal of Distributed Systems and Technologies, 2018, 9, 39-52.	0.6	1
6777	Robust and Probabilistic Failure-Aware Placement. ACM Transactions on Parallel Computing, 2018, 5, 1-30.	1.2	3
6778	Aggressive Synchronization with Partial Processing for Iterative ML Jobs on Clusters. , 2018, , .		13
6779	Parallelization and vectorization of ROOT fitting classes. Journal of Physics: Conference Series, 2018, 1085, 032024.	0.3	0
6780	Speeding DBLP querying using hadoop and spark. IOP Conference Series: Materials Science and Engineering, 0, 459, 012003.	0.3	0
6781	Efficient Scheduling for Scalable Bioinformatics Analysis Platform with Microservices. , 2018, , .		0
6782	Improving I/O Efficiency in Hadoop-Based Massive Data Analysis Programs. Scientific Programming, 2018, 2018, 1-9.	0.5	1
6783	SparkOT: Diagnosing Operation Level Inefficiency in Spark. , 2018, , .		1
6784	Doppio: I/O-Aware Performance Analysis, Modeling and Optimization for In-memory Computing Framework. , 2018, , .		12
6785	Hadoop vs. Spark: Impact on Performance of the Hammer Query Engine for Open Data Corpora. Algorithms, 2018, 11, 209.	1.2	5
6786	Integrated High-Performance Platform for Fast Query Response in Big Data with Hive, Impala, and SparkSQL: A Performance Evaluation. Applied Sciences (Switzerland), 2018, 8, 1514.	1.3	2
6787	Sparse-Sensor Placement for Wind Farm Control. Journal of Physics: Conference Series, 2018, 1037, 032019.	0.3	14
6788	<i>Gaia</i> Data Release 2. Astronomy and Astrophysics, 2018, 616, A3.	2.1	124
6789	Improved Performance of Hive Using Index-Based Operation on Big Data. , 2018, , .		2
6790	A Reinforcement Learning Based Resource Management Approach for Time-critical Workloads in Distributed Computing Environment. , 2018, , .		12
6791	Big Data Indexing. , 2018, , 1-11.		0
6792	Learning-based Automatic Parameter Tuning for Big Data Analytics Frameworks. , 2018, , .		16

#	ARTICLE	IF	CITATIONS
6793	A Generic and Scalable Pipeline for Large-Scale Analytics of Continuous Aircraft Engine Data. , 2018, , .		7
6794	A Multi-Criteria Experimental Ranking of Distributed SPARQL Evaluators. , 2018, , .		5
6795	Parallel K-Medoids Improved Algorithm Based on MapReduce. , 2018, , .		0
6796	ENEDI: Energy Saving in Datacenters. , 2018, , .		2
6797	Interactive Launch of 16,000 Microsoft Windows Instances on a Supercomputer. , 2018, , .		3
6798	Novel Scalable Deep Learning Approaches for Big Data Analytics Applied to ECG Processing. International Journal of Applied Metaheuristic Computing, 2018, 9, 33-51.	0.5	1
6799	COBRA: Toward Provably Efficient Semi-Clairvoyant Scheduling in Data Analytics Systems. , 2018, , .		2
6800	On a Dynamic Data Placement Strategy for Heterogeneous Hadoop Clusters. , 2018, , .		4
6801	Technical Analysis on Financial Time Series Data Based on Map-Reduce Programming Model: A Case Study. , 2018, , .		3
6802	Exploiting a Synergy between Greedy Approach and NSGA for Scheduling in Computing Clusters. , 2018, , .		0
6803	Scheduling in the Hybrid Cloud Constrained by Process Mining. , 2018, , .		5
6804	Forecasting Smart Meter Energy Usage Using Distributed Systems and Machine Learning. , 2018, , .		12
6805	SparkFSM: A Highly Scalable Frequent Subgraph Mining Approach using Apache Spark. , 2018, , .		1
6806	Review of Graph Processing Frameworks. , 2018, , .		2
6807	Workflows Orchestrating Workflows: Thousands of Queries and Their Fault Tolerance Using APIs of Omics Web Resources. , 2018, , .		0
6808	De-duplicating the OpenAIRE Scholarly Communication Big Graph. , 2018, , .		1
6809	FDQ: Advance Analytics Over Real Scientific Array Datasets. , 2018, , .		1
6810	Research Challenges in Big Data Solutions in Different Applications. SSRN Electronic Journal, 2018, , .	0.4	0

#	ARTICLE	IF	CITATIONS
6811	Impact of Link Failures on the Performance of MapReduce in Data Center Networks. , 2018, , .		2
6812	Data distribution method for scalable actionable pattern mining. , 2018, , .		2
6813	Graph Database and Graph Computing for Cyber-Physical Power Systems. , 2018, , .		1
6814	Code and Data Transformations to Address Garbage Collector Performance in Big Data Processing. , 2018, , .		0
6815	Parallelization of Multi-label classification for large data sets. , 2018, , .		0
6816	Evaluation of Performance Saturation Using the Hadoop Framework. , 2018, , .		0
6817	Automated Parallel Data Processing Engine with Application to Large-Scale Feature Extraction. , 2018, , .		3
6818	Reactive caching for composed services: polling at the speed of push. , 2018, 2, 1-28.		1
6819	Streaming Algorithm for Big Data Logistic Regression. , 2018, , .		3
6820	Automated Table Partitioner (ATAP) in Apache Hive. , 2018, , .		0
6821	A Dynamic Load Balancing Scheme for Distributed Formal Concept Analysis. , 2018, , .		2
6822	Evaluating Checkpoint Interval for Fault-Tolerance in MapReduce. , 2018, , .		0
6823	Research on Key Technologies of Cloud Energy Management for Wide Area Integrated Energy Internet. , 2018, , .		3
6824	Pattern-Based Automatic Parallelization of Representative-Based Clustering Algorithms. , 2018, , .		3
6825	Land Cover Classification of Fine-Resolution Remote Sensing Data. , 2018, , 17-28.		1
6826	Practical Cross Program Memoization with KeyChain. , 2018, , .		0
6827	File Placing Location Optimization on Hadoop SWIM. , 2018, , .		2
6828	Map Reduce by K-Nearest Neighbor Joins. , 2018, , .		1

#	ARTICLE	IF	CITATIONS
6829	A Private Cloud for Data Mining Education. , 2018, , .		4
6830	IncastGuard: An Efficient TCP-Incast Mitigation Mechanism for Cloud Networks. , 2018, , .		3
6831	Technology Enablers for Big Data, Multi-Stage Analysis in Medical Image Processing. , 2018, , .		1
6832	Scheduling Stream Processing Tasks on Geo-Distributed Heterogeneous Resources. , 2018, , .		10
6833	On the Usage of the Probability Integral Transform to Reduce the Complexity of Multi-Way Fuzzy Decision Trees in Big Data Classification Problems. , 2018, , .		3
6834	A Flexible Desktop Tool for the Deployment of Periodic Downstream Services. , 2018, , .		0
6835	Joint Scheduling of Tasks and Network Flows in Big Data Clusters. IEEE Access, 2018, 6, 66600-66611.	2.6	6
6836	Sitting Posture Correction Device Based on Infrared Distance Measurement. , 2018, , .		3
6837	DISTIL. , 2018, , .		8
6838	A Performance Study of Big Spatial Data Systems. , 2018, , .		7
6839	PETS: Bottleneck-Aware Spark Tuning with Parameter Ensembles. , 2018, , .		4
6840	Spark Sensing: A Cloud Computing Framework to Unfold Processing Efficiencies for Large and Multiscale Remotely Sensed Data, with Examples on Landsat 8 and MODIS Data. Journal of Sensors, 2018, 2018, 1-12.	0.6	4
6841	Foundry: a message-oriented, horizontally scalable ETL system for scientific data integration and enhancement. Database: the Journal of Biological Databases and Curation, 2018, 2018, .	1.4	8
6842	FEMRL: A Framework for Large-Scale Privacy-Preserving Linkage of Patientsâ€™ Electronic Health Records. , 2018, , .		2
6843	A qualitative survey on frequent subgraph mining. Open Computer Science, 2018, 8, 194-209.	1.3	8
6844	Event Analysis on Power Communication Networks With Big Data for Maintenance Forms. IEEE Access, 2018, 6, 72263-72274.	2.6	5
6845	On the Performance Analysis of Map-Reduce Programming Model on In-Memory NoSQL Storage Platforms: A Case Study. , 2018, , .		4
6846	An Efficient Dynamic Load-Balancing Large Scale Graph-Processing System. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6847	Security and Cryptographic Challenges for Authentication Based on Biometrics Data. <i>Cryptography</i> , 2018, 2, 39.	1.4	10
6848	POSUM: A Portfolio Scheduler for MapReduce Workloads. , 2018, , .		1
6849	Concurrent agent-based evolutionary computations as adaptive dataflows. <i>Concurrency Computation Practice and Experience</i> , 2018, 30, e4702.	1.4	1
6851	Big Data Processing in Fog - Smart Parking Case Study. , 2018, , .		9
6852	Skew-Aware Collective Communication for MapReduce Shuffling. , 2018, , .		0
6853	A vision of post-exascale programming. <i>Frontiers of Information Technology and Electronic Engineering</i> , 2018, 19, 1261-1266.	1.5	6
6854	From Application to Disk: Tracing I/O Through the Big Data Stack. <i>Lecture Notes in Computer Science</i> , 2018, , 89-102.	1.0	0
6855	Training bioinformaticians in High Performance Computing. <i>Heliyon</i> , 2018, 4, e01057.	1.4	5
6856	Big Data Processing Platform for Smart City. , 2018, , .		0
6857	Evidence Fusion for Malicious Bot Detection in IoT. , 2018, , .		11
6858	A Complete Data Science Work-flow For Insurance Field. , 2018, , .		0
6859	Actionable Pattern Mining - A Scalable Data Distribution Method Based on Information Granules. , 2018, , .		4
6860	Validation and Correction of Large Security Policies: A Clustering and Access Log Based Approach. , 2018, , .		6
6861	Online Job Dispatching and Scheduling to Minimize Job Completion Time and to Meet Deadlines. <i>Journal of Interconnection Networks</i> , 2018, 18, 1850012.	0.6	0
6862	Hengam a MapReduce-Based Distributed Data Warehouse for Big Data. <i>International Journal of Artificial Life Research</i> , 2018, 8, 16-35.	0.1	2
6863	A MapReduce-Based User Identification Algorithm in Web Usage Mining. <i>International Journal of Information Technology and Web Engineering</i> , 2018, 13, 11-23.	1.2	6
6864	K-mer Counting: memory-efficient strategy, parallel computing and field of application for Bioinformatics. , 2018, , .		2
6865	Massive Online Shopping Data Mining Based on Hadoop. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6866	DisTenC: A Distributed Algorithm for Scalable Tensor Completion on Spark. , 2018, , .		9
6867	Unleashing the Power of Hashtags in Tweet Analytics with Distributed Framework on Apache Storm. , 2018, , .		3
6868	A Proposal of Scalable and Performing Implementation of Algorithms for Anomaly and Community Detection. , 2018, , .		0
6869	A Decomposition-Based Partition Method For Matrix-Multiplication In The Mapreduce Model. , 2018, , .		1
6870	Big Data Analyzing for Control Valve Using the Spark Cluster Computing Engine. , 2018, , .		0
6871	Towards Multi-Objective Optimisation of Hadoop 2.x Application Deployment on Public Clouds. , 2018, , .		0
6872	Distributed Data Analytics Framework for Smart Transportation. , 2018, , .		11
6873	Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR. , 2018, , .		56
6874	BigDataNetSim: A Simulator for Data and Process Placement in Large Big Data Platforms. , 2018, , .		3
6875	MLB+-tree: A Multi-level B+-tree Index for Multidimensional Range Query on Seismic Data. , 2018, , .		2
6876	BiSpark: a Spark-based highly scalable aligner for bisulfite sequencing data. BMC Bioinformatics, 2018, 19, 472.	1.2	5
6877	Resource Management: Performance Assuredness in Distributed Cloud Computing via Online Reconfigurations. , 2018, , 160-236.		0
6878	Short-Term Traffic Flow Forecasting Using a Distributed Spatial-Temporal k Nearest Neighbors Model. , 2018, , .		9
6879	Challenges and Proposals for Enabling Dynamic Heterogeneous Execution of Big Data Frameworks. , 2018, , .		1
6880	The impact of checkpointing interval selection on the scheduling performance of Hadoop framework. , 2018, , .		2
6881	A survey of scheduling frameworks in big data systems. International Journal of Cloud Computing, 2018, 7, 103.	0.3	11
6882	Enabling lock-free concurrent workers over temporal graphs composed of multiple time-series. , 2018, , .		2
6883	A Scalable and Efficient Multi-Label CNN-Based License Plate Recognition on Spark. , 2018, , .		2

#	ARTICLE	IF	CITATIONS
6884	Research on Ship Berthing Monitoring Technology Based on 3D Laser Point Cloud Data. , 2018, , .		1
6885	Wireless MapReduce Distributed Computing. , 2018, , .		13
6886	Dynamic Programming-Based Vessel Speed Adjustment for Energy Saving and Emission Reduction. Energies, 2018, 11, 1273.	1.6	15
6887	A Computer Cluster for Big Data and Data Analytics Management. , 2018, , .		0
6888	Grid-Based Parallel Algorithms of Join Queries for Analyzing Multi-Dimensional Data on MapReduce. IEICE Transactions on Information and Systems, 2018, E101.D, 964-976.	0.4	2
6889	Quality of Cloud Services Determined by the Dynamic Management of Scheduling Models for Complex Heterogeneous Workloads. , 2018, , .		4
6890	Using Weaker Consistency Models with Monitoring and Recovery for Improving Performance of Key-Value Stores. , 2018, , .		2
6891	One Self-Adaptive Memory Scheduling Algorithm for the Shuffle Process in Spark Platform. , 2018, , .		0
6892	On the Power of Combiner Optimizations in MapReduce Over MPI Workflows. , 2018, , .		2
6893	Analyzing fault tolerance mechanism of Hadoop Mapreduce under different type of failures. , 2018, , .		0
6894	HMSPKmerCounter: Hadoop based Parallel, Scalable, Distributed Kmer Counter for Large Datasets. , 2018, , .		0
6895	A Review on Data locality in Hadoop MapReduce. , 2018, , .		3
6896	Coflow Deadline Scheduling via Network-Aware Optimization. , 2018, , .		3
6897	Parallelizing the Construction of a k-Dimensional Tree. , 2018, , .		2
6898	Parallel Acceleration of Subgraph Enumeration in the Process of Network Motif Detection. , 2018, , .		1
6899	PAST-HOAS: Fast Setup of Per-Address Sink Routing Trees by Reverse Flooding. , 2018, , .		0
6900	On Distributed Computing with Heterogeneous Communication Constraints. , 2018, , .		6
6901	Efficient Coflow Scheduling of Multi-Stage Jobs with Isolation Guarantee. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6902	Promoting and Balancing Resource Utilization via Integration by Exploiting the Time-varying Property of Resource Requirements. , 2018, , .		1
6903	Data Encoding for Byzantine-Resilient Distributed Gradient Descent. , 2018, , .		5
6904	Research on Data Sharing Analysis and Key Technology of Smart City. , 2018, , .		0
6905	A Lightweight Model for Right-Sizing Master-Worker Applications. , 2018, , .		7
6906	Evaluation of high-level query languages based on MapReduce in Big Data. Journal of Big Data, 2018, 5, .	6.9	12
6907	A Loading Balancing Equi-join Algorithm Based on Key Cost. , 2018, , .		0
6908	Aten: A Dispatcher for Big Data Applications in Heterogeneous Systems. , 2018, , .		3
6909	Performance Factor Analysis and Scope of Optimization for Big Data Processing on Cluster. , 2018, , .		0
6910	Finding simple sequence repeats (SSRs) within human genome using MapReduce based K-mer algorithm. , 2018, , .		1
6911	High Performance Energy Prediction using Hadoop with Spark. ITM Web of Conferences, 2018, 20, 02012.	0.4	0
6912	An Improved Speculative Strategy for Heterogeneous Spark Cluster. MATEC Web of Conferences, 2018, 173, 01018.	0.1	2
6913	Large-Scale Clustering Using MPI-Based Canopy. , 2018, , .		1
6914	Use of Big Data Analytics in WASH Sector. , 2018, , .		3
6915	MapReduce with Components for Processing Big Graphs. , 2018, , .		1
6916	Research of MapReduce architecture on busbar protection. , 2018, , .		0
6917	Towards Energy-Proportional Anomaly Detection in the Smart Grid. , 2018, , .		4
6918	CLIBE: Precise Cluster-Level I/O Bandwidth Enforcement in Distributed File System. , 2018, , .		0
6919	In-Memory Hadoop on Supercomputers Using Memcached-Like Nodes for Data Storage Only. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6920	Big Data Ensure Homologous Patient Enduring Therapy Time Forecast Algorithm by Healing Facility Echelon Recommendation. , 2018, , .		0
6921	DcCNC: Dynamic and Concurrent Offloading Strategy Considering Network Condition and Server Load. , 2018, , .		0
6922	Decoupled Strategy for Imbalanced Workloads in MapReduce Frameworks. , 2018, , .		0
6923	High-Performance Computational and Information Technologies for Numerical Models and Data Processing. , 2018, , .		0
6924	Using compiler snippets to exploit parallelism on heterogeneous hardware: a Java reduction case study. , 2018, , .		2
6925	NOVEL APPROACH FOR DENOISING USING HADOOP IMAGE PROCESSING INTERFACE. Procedia Computer Science, 2018, 132, 1327-1350.	1.2	3
6926	P-FHM+: Parallel high utility itemset mining algorithm for big data processing. Procedia Computer Science, 2018, 132, 918-927.	1.2	26
6927	MRF: MapReduce based Forecasting Algorithm for Time Series Data. Procedia Computer Science, 2018, 132, 92-102.	1.2	9
6928	An Ontology-driven MapReduce Framework for Association Rules Mining in Massive Data. Procedia Computer Science, 2018, 126, 224-233.	1.2	2
6929	Spark-based data analytics of sequence motifs in large omics data. Procedia Computer Science, 2018, 126, 596-605.	1.2	25
6930	BigDimETL with NoSQL Database. Procedia Computer Science, 2018, 126, 798-807.	1.2	14
6931	A Performance Comparison of Big Data Processing Platform Based on Parallel Clustering Algorithms. Procedia Computer Science, 2018, 139, 127-135.	1.2	6
6932	Big Data Analytics based on PANFIS MapReduce. Procedia Computer Science, 2018, 144, 140-152.	1.2	3
6933	CSTF. , 2018, , .		12
6934	Towards Probabilistic Networks of Polarized Evolutionary Processors. , 2018, , .		3
6935	Segmentation and Classification of Brain MR Images Using Big Data Analytics. , 2018, , .		3
6936	Seal: Efficient Training Large Scale Statistical Machine Translation Models on Spark. , 2018, , .		0
6937	Hierarchy Value Density - A Priority Algorithm for Task Scheduling on Cloud. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6938	Realization of Chaotic Sequence Encryption Algorithm in MapReduce Distributed Parallel Model. , 2018, , .		1
6939	Far-Sighted Multi-Stage Aware Coflow Scheduling. , 2018, , .		2
6940	Radar: Reducing Tail Latencies for Batched Stream Processing with Blank Scheduling. , 2018, , .		0
6941	Big Data and Internet of Things: A Survey. , 2018, , .		5
6942	Entropy Based Pruning of Backoff Maxent Language Models with Contextual Features. , 2018, , .		3
6943	Chicago Crime Data Analysis Using PIG in Hadoop. , 2018, , .		0
6944	Building efficient fuzzy regression trees for large scale and high dimensional problems. Journal of Big Data, 2018, 5, .	6.9	6
6945	Actualizing the Job Scheduling Scheme with Heterogeneous Virtual Map reduce Clusters. , 2018, , .		0
6946	Low-Rank Optimization for Data Shuffling in Wireless Distributed Computing. , 2018, , .		7
6947	Vertex-Cut Partitioning Performance Analysis for FASTCD Algorithm in Large-Scale Graph. , 2018, , .		0
6948	Research on Tracking and Tracing Bitcoin Fund Flows. , 2018, , .		5
6949	Shaping Deadline Coflows to Accelerate Non-Deadline Coflows. , 2018, , .		6
6950	CLF: An Online Coflow-Aware Packet Scheduling Algorithm. , 2018, , .		2
6951	LSTM-based Traffic Prediction in Support of Periodically Light Path Reconfiguration in Hybrid Data Center Network. , 2018, , .		7
6952	An Improved Index Based on MapReduce for Path Queries in Public Transportation Networks. , 2018, , .		1
6953	Adaptive Probe Size Estimation for Cluster Management in Datacenters. , 2018, , .		0
6954	Achieving Performance and Programmability for MapReduce(-Like) Frameworks. , 2018, , .		2
6955	Performance Analysis of Big Data Frameworks on Virtualized Clusters. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6956	SciAP: A Programmable, High-Performance Platform for Large-Scale Scientific Data. , 2018, , .		0
6957	Complex Workflows Implementation in Stream Data Processing. , 2018, , .		0
6958	Legal Files Management System Using Big Data. , 2018, , .		0
6959	Coded Computing for Distributed Graph Analytics. , 2018, , .		30
6960	Big Earth Data: a comprehensive analysis of visualization analytics issues. Big Earth Data, 2018, 2, 321-350.	2.0	10
6961	A New MapReduce Approach with Dynamic Fuzzy Inference for Big Data Classification Problems. International Journal of Cognitive Informatics and Natural Intelligence, 2018, 12, 40-54.	0.4	7
6962	Substituting Missing Values in End-to-End Internet Performance Measurements Using k-Nearest Neighbors. , 2018, , .		2
6963	Supervised Papers Classification on Large-Scale High-Dimensional Data with Apache Spark. , 2018, , .		0
6964	How Improve Set Similarity Join Based on Prefix Approach in Distributed Environment. , 2018, , .		0
6965	Towards a Better Replica Management for Hadoop Distributed File System. , 2018, , .		13
6966	Secure and Scalable Key Value Storage for Managing Big Data in Smart Cities Using Intel SGX. , 2018, , .		1
6967	Reusable Big Data System for Industrial Data Mining - A Case Study on Anomaly Detection in Chemical Plants. Lecture Notes in Computer Science, 2018, , 611-622.	1.0	6
6968	Hybrid Circuit/Packet Network Scheduling with Multiple Composite Paths. , 2018, , .		1
6969	Towards Making Distributed RDF Processing FLINKer. , 2018, , .		4
6970	Frequent Itemset Mining in Big Data With Effective Single Scan Algorithms. IEEE Access, 2018, 6, 68013-68026.	2.6	33
6971	Dynamic Heterogeneity-Aware Coded Cooperative Computation at the Edge. , 2018, , .		27
6972	EMPOWERING, a Smart Big Data Framework for Sustainable Electricity Suppliers. IEEE Access, 2018, 6, 71132-71142.	2.6	7
6973	Access Point Decoding Coded MapReduce for Tree Fog Network. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
6974	Efficiently finding minimal failing input in MapReduce programs. , 2018, , .		0
6975	Scalable Big Data Privacy with MapReduce. , 2018, , 1-8.		0
6976	MapFIM+: Memory Aware Parallelized Frequent Itemset Mining In Very Large Datasets. Lecture Notes in Computer Science, 2018, , 200-225.	1.0	2
6977	Online Aggregation: A Review. Lecture Notes in Computer Science, 2018, , 103-114.	1.0	2
6978	MapReduce a Comprehensive Review. , 2018, , .		5
6979	A Distributed Rule Engine for Streaming Big Data. Lecture Notes in Computer Science, 2018, , 123-130.	1.0	4
6980	HEAP: An Efficient and Fault-Tolerant Authentication and Key Exchange Protocol for Hadoop-Assisted Big Data Platform. IEEE Access, 2018, 6, 75342-75382.	2.6	17
6981	The Inverse Tree-OLAP Problem. , 2018, , .		0
6982	Spatial-Temporal K Nearest Neighbors Model on MapReduce for Traffic Flow Prediction. Lecture Notes in Computer Science, 2018, , 253-260.	1.0	3
6983	Fitted Fat-Tree for Localized Traffic in Data Center Networks. , 2018, , .		0
6984	Spam Detection Framework for Online Reviews Using Hadoop™ s Computational Capability. , 2018, , .		1
6985	Workload Alleviation Scheduling Framework to Alleviate Negative Performance Impact of Intermediate Data Skew in Small-Scale MapReduce Cloud. , 2018, , .		1
6986	Distributed gene clinical decision support system based on cloud computing. BMC Medical Genomics, 2018, 11, 100.	0.7	2
6987	Improving Fuzzy Rule Based Classification Systems in Big Data via Support-based Filtering. , 2018, , .		1
6988	Micro-Burst in Data Centers: Observations, Analysis, and Mitigations. , 2018, , .		30
6989	RPC: Joint Online Reducer Placement and Coflow Bandwidth Scheduling for Clusters. , 2018, , .		6
6990	An Automatic Deployment Support for Processing Remote Sensing Data in the Cloud. , 2018, , .		1
6991	Criso: An Incremental Scalable and Cost-Effective Data Center Interconnection by Using 2-Port Servers and low-end Switches. , 2018, , .		2

#	ARTICLE	IF	CITATIONS
6992	Workload Partitioning and Task Migration to Reduce Response Times in Heterogeneous Computing Environments. , 2018, , .		6
6993	Energy Efficiency of Server-Centric PON Data Center Architecture for Fog Computing. , 2018, , .		21
6994	Microscope: Pinpoint Performance Issues with Causal Graphs in Micro-service Environments. Lecture Notes in Computer Science, 2018, , 3-20.	1.0	78
6995	A Case Study of Accelerating Apache Spark with FPGA. , 2018, , .		7
6996	Toward Scalable Anonymization for Privacy-Preserving Big Data Publishing. Advances in Intelligent Systems and Computing, 2018, , 297-304.	0.5	3
6997	Fundamental Limits of Wireless Distributed Computing Networks. , 2018, , .		7
6998	Blind, Adaptive and Robust Flow Segmentation in Datacenters. , 2018, , .		4
6999	VeriTable: Fast Equivalence Verification of Multiple Large Forwarding Tables. , 2018, , .		3
7000	Online Partial Throughput Maximization for Multidimensional Coflow. , 2018, , .		6
7001	Extracting Knowledge from Open Source Projects to Improve Program Security. , 2018, , .		4
7002	A datalog-based computational model for coordination-free, data-parallel systems. Theory and Practice of Logic Programming, 2018, 18, 874-927.	1.1	1
7003	Lunar Geomorphology Three-Dimensional Visualization within WorldWide Telescope Based on GPU. , 2018, , .		0
7004	Hierarchical Clustering in Scalable Distributed Two-Layer Datastore for Big Data as a Service. , 2018, , .		3
7005	PADS: Performance-Aware Dynamic Scheduling for Effective MapReduce Computation in Heterogeneous Clusters. , 2018, , .		1
7006	A preliminary study on Hybrid Spill-Tree Fuzzy k-Nearest Neighbors for big data classification. , 2018, , .		4
7007	An Novel Parameter Tuning Alogrithm for MapReduce Model. , 2018, , .		1
7008	Deep Learning Convolutional Neural Networks with Dropout - A Parallel Approach. , 2018, , .		11
7009	Optimized Common Parameter Set Extraction by Benchmarking Applications on a Big Data Platform. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
7010	From E-911 to NG-911: Overview and Challenges in Ecuador. IEEE Access, 2018, 6, 42578-42591.	2.6	10
7011	Investigating the Efficiency of Machine Learning Algorithms on MapReduce Clusters with SSDs. , 2018, , .		0
7012	Optimising Cloud-Based Hadoop 2.x Applications. , 2018, , .		0
7013	Executing Analytics and Fusion Workloads on Transient Computing Resources in Tactical Environments. , 2018, , .		0
7014	Optimization of Real-World MapReduce Applications With Flame-MR: Practical Use Cases. IEEE Access, 2018, 6, 69750-69762.	2.6	1
7015	GDup: De-Duplication of Scholarly Communication Big Graphs. , 2018, , .		3
7016	Scalable Batch Stream Clustering with k Estimation. , 2018, , .		1
7017	Prophet: Fast Accurate Model-Based Throughput Prediction for Reactive Flow in DC Networks. , 2018, , .		5
7018	Research on Green Development of Shaanxi Industry Under The Background of Big Data. , 2018, , .		0
7019	A Balanced Partitioning Mechanism Using Collapsed-Condensed Trie in MapReduce. , 2018, , .		0
7020	SPARK Processing of Computing-Intensive Classification of Remote Sensing Images: The Case on K-Means Clustering Algorithm. , 2018, , .		2
7021	Parallelizing Sequential Graph Computations. ACM Transactions on Database Systems, 2018, 43, 1-39.	1.5	28
7023	Parallelizing Shortest Average-Distance Query Processing. , 2018, , .		0
7024	Concurrency Control Protocol for Parallel B-Tree Structures That Improves the Efficiency of Request Transfers and SMOs within a Node. IEICE Transactions on Information and Systems, 2018, E101.D, 152-170.	0.4	0
7025	Future Cloud Systems Design: Challenges and Research Directions. IEEE Access, 2018, 6, 74120-74150.	2.6	29
7026	Subsampling MCMC - an Introduction for the Survey Statistician. Sankhya A, 2018, 80, 33-69.	0.4	11
7027	Sparker: Optimizing Spark for Heterogeneous Clusters. , 2018, , .		3
7028	Protocol Programming: A Layered Programming Paradigm. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
7029	A Cloud-based Control System Architecture for Multi-UAV. , 2018, , .		9
7030	SmarTor. , 2018, , .		6
7031	Parallel MapReduce: Maximizing Cloud Resource Utilization and Performance Improvement Using Parallel Execution Strategies. BioMed Research International, 2018, 2018, 1-17.	0.9	2
7032	A oneM2M-Based Query Engine for Internet of Things (IoT) Data Streams. Sensors, 2018, 18, 3253.	2.1	8
7033	Multilinear EigenECGs and FisherECGs for Individual Identification from Information Obtained by an Electrocardiogram Sensor. Symmetry, 2018, 10, 487.	1.1	5
7034	Understanding and Minimizing Disk Contention Effects for Data-Intensive Processing in Virtualized Systems. , 2018, , .		7
7035	An Effective High Utility Itemset Mining Algorithm with Big Data Based on MapReduce Framework. , 2018, , .		2
7036	Parallel Graph Connectivity in Log Diameter Rounds. , 2018, , .		41
7037	AIRPORTS Metrics: A Big Data application for computing flights' performance indexes based on flown trajectories. , 2018, , .		2
7038	Improving Hamming distance-based fuzzy join in MapReduce using Bloom Filters. , 2018, , .		3
7039	A Parallel Fast Fourier Transform Algorithm for Large-Scale Signal Data Using Apache Spark in Cloud. Lecture Notes in Computer Science, 2018, , 293-310.	1.0	3
7040	Cooperative-Competitive Task Allocation in Edge Computing for Delay-Sensitive Social Sensing. , 2018, , .		50
7041	Exact and Approximate Map-Reduce Algorithms for Convex Hull. Lecture Notes in Computer Science, 2018, , 480-494.	1.0	0
7042	Cloud-based improved Monte Carlo localization algorithm with robust orientation estimation for mobile robots. Engineering Computations, 2018, 36, 178-203.	0.7	5
7043	Five years at the edge. , 2018, , .		31
7044	Challenges in High Performance Big Data Frameworks. , 2018, , .		0
7045	A Method to Accelerate K-Means and GMM Computation with GPU and Multi-Core CPU. , 2018, , .		1
7046	Privacy-Preserving Naive Bayes Classification Using Fully Homomorphic Encryption. Lecture Notes in Computer Science, 2018, , 349-358.	1.0	7

#	ARTICLE	IF	CITATIONS
7047	Multicasting in a 256-Port Sub- μ sec Latency λ Switch Architecture for Disaggregated DataCenters. , 2018, , .		0
7048	Non-local Data Fetch Scheme Based on Delay Distribution for Hadoop Clusters in Public Cloud. , 2018, , .		1
7049	Intermediate Results Materialization Selection and Format for Data-Intensive Flows*. Fundamenta Informaticae, 2018, 163, 111-138.	0.3	2
7050	In Search of Actionable Patterns of Lowest Cost – A Scalable Action Graph Method. , 2018, , .		1
7051	MapReduce Tuning to Improve Distributed Machine Learning Performance. , 2018, , .		6
7052	In-memory Spatial-Aware Framework for Processing Proximity-Alike Queries in Big Spatial Data. , 2018, , .		7
7053	Vote Parallel SVM: An Extension of Parallel Support Vector Machine. , 2018, , .		1
7054	[Engineering Paper] A Tool for Optimizing Java 8 Stream Software via Automated Refactoring. , 2018, , .		6
7055	Recognition of bacteria named entity using conditional random fields in Spark. BMC Systems Biology, 2018, 12, 106.	3.0	9
7056	Formalizing Compute-Aggregate Problems in Cloud Computing. Lecture Notes in Computer Science, 2018, , 377-391.	1.0	0
7057	Load Balancing through Block Rearrangement Policy for Hadoop Heterogeneous Cluster. , 2018, , .		10
7058	The Case of Big Data Platform Services for Semiconductor Wafer Fabrication Foundries. , 2018, , .		1
7059	Graphical Spark Programming in IoT Mashup Tools. , 2018, , .		7
7060	CloudRaid: hunting concurrency bugs in the cloud via log-mining. , 2018, , .		20
7061	Result Pattern Hiding Searchable Encryption for Conjunctive Queries. , 2018, , .		113
7062	An efficient load balancing multicast scheduling for solving congestion problem in social data center networks. , 2018, , .		0
7063	Sensitive Data Detection Using NN and KNN from Big Data. Lecture Notes in Computer Science, 2018, , 628-642.	1.0	2
7064	Accelerating Spark Shuffle with RDMA. , 2018, , .		1

#	ARTICLE	IF	CITATIONS
7065	Social Media Data Analysis Using MapReduce Programming Model and Training a Tweet Classifier Using Apache Mahout. , 2018, , .		5
7066	Online Machine Learning in Big Data Streams: Overview. , 2018, , 1-11.		20
7067	MapRDD. , 2018, , .		1
7068	HSASore: A Hierarchical Storage Architecture for Computing Systems Containing Large-Scale Intermediate Data. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2018, , 591-601.	0.2	1
7069	Distributed Hybrid Rice Algorithm Based on Hadoop. , 2018, , .		0
7070	Sprintz. , 2018, 2, 1-23.		66
7071	Three-dimensional image registration using distributed parallel computing. IET Image Processing, 2018, 12, 1713-1720.	1.4	2
7072	Cross-language optimizations in big data systems. , 2018, , .		0
7073	The Internet of Things Is Digitizing and Transforming Science. SLAS Technology, 2018, 23, 407-411.	1.0	8
7074	Towards secure dynamic product lines in the cloud. , 2018, , .		6
7075	Libra. , 2018, , .		0
7076	Atomic Information Disclosure of Off-Chained Computations Using Threshold Encryption. Lecture Notes in Computer Science, 2018, , 85-93.	1.0	0
7077	Energy-efficient shipping: An application of big data analysis for optimizing engine speed of inland ships considering multiple environmental factors. Ocean Engineering, 2018, 169, 457-468.	1.9	71
7078	HSRA: Hadoop-based spliced read aligner for RNA sequencing data. PLoS ONE, 2018, 13, e0201483.	1.1	12
7079	Mining Rules with Constants from Large Scale Knowledge Bases. Lecture Notes in Computer Science, 2018, , 521-535.	1.0	1
7080	RIDE: real-time massive image processing platform on distributed environment. Eurasip Journal on Image and Video Processing, 2018, 2018, .	1.7	9
7081	Multimedia big data computing and Internet of Things applications: A taxonomy and process model. Journal of Network and Computer Applications, 2018, 124, 169-195.	5.8	156
7082	Six Pass MapReduce Implementation of Strassen's Algorithm for Matrix Multiplication. , 2018, , .		3

#	ARTICLE	IF	CITATIONS
7083	HaDaap: A hotness-aware data placement strategy for improving storage efficiency in heterogeneous Hadoop clusters. <i>Concurrency Computation Practice and Experience</i> , 2018, 30, e4830.	1.4	9
7084	Supporting Continuous Skyline Queries in Dynamically Weighted Road Networks. <i>Mathematical Problems in Engineering</i> , 2018, 2018, 1-14.	0.6	4
7086	Case Study: Adaptive Test Automation for Testing an Adaptive Hadoop Resource Manager. , 2018, , .		1
7087	The Internet of Things, Fog and Cloud continuum: Integration and challenges. <i>Internet of Things (Netherlands)</i> , 2018, 3-4, 134-155.	4.9	195
7088	Adaptive ensemble simulations of biomolecules. <i>Current Opinion in Structural Biology</i> , 2018, 52, 87-94.	2.6	24
7089	Who Tests the Testers?. , 2018, , .		15
7090	A Sliding-Window Algorithm Implementation in MapReduce. <i>Lecture Notes in Social Networks</i> , 2018, , 69-83.	0.8	0
7091	Think Sequential, Run Parallel. <i>Lecture Notes in Computer Science</i> , 2018, , 1-25.	1.0	2
7092	Introduction to Supervised Descriptive Pattern Mining. , 2018, , 1-31.		2
7093	Foundations of the Hadoop Ecosystem. <i>Computational Biology</i> , 2018, , 137-150.	0.1	3
7094	LAS: Logical-Block Affinity Scheduling in Big Data Analytics Systems. , 2018, , .		1
7095	Spark-parSketch. , 2018, , .		4
7096	From Neural Re-Ranking to Neural Ranking. , 2018, , .		72
7097	CASS: A distributed network clustering algorithm based on structure similarity for large-scale network. <i>PLoS ONE</i> , 2018, 13, e0203670.	1.1	5
7098	Dynamic Model Evaluation to Accelerate Distributed Machine Learning. , 2018, , .		3
7099	Improving Utilization and Parallelism of Hadoop Cluster by Elastic Containers. , 2018, , .		7
7100	Adhesive Model for Collection and Auto Storage of Colossal Health Data for Epidemiological Studies. , 2018, , .		0
7101	Imbalanced Classification for Big Data. , 2018, , 327-349.		7

#	ARTICLE	IF	CITATIONS
7102	Foundations on Imbalanced Classification. , 2018, , 19-46.		9
7103	Health Big Data Analytics: A Technology Survey. IEEE Access, 2018, 6, 65661-65678.	2.6	72
7104	Dynamic Performance Aware Reduce Task Scheduling in MapReduce on Virtualized Environment. , 2018, , .		1
7105	Chabok: a Map-Reduce based method to solve data warehouse problems. Journal of Big Data, 2018, 5, .	6.9	7
7106	Error-bounded and Number-bounded Approximate Spatial Query for Interactive Visualization. Open Geosciences, 2018, 10, 491-503.	0.6	1
7107	Architectures. , 2018, , 1-11.		0
7108	Apache Flink. , 2018, , 1-8.		13
7109	Sensor Data Based System-Level Anomaly Prediction for Smart Manufacturing. , 2018, , .		18
7110	TSim: a system for discovering similar users on Twitter. Journal of Big Data, 2018, 5, .	6.9	9
7111	A Local Approximation Approach for Processing Time-Evolving Graphs. Symmetry, 2018, 10, 247.	1.1	3
7112	Big Data and Its Visualization With Fog Computing. International Journal of Fog Computing, 2018, 1, 51-82.	1.8	9
7113	A Preliminary Study of the Feasibility of Global Evolutionary Feature Selection for Big Datasets under Apache Spark. , 2018, , .		2
7114	Joint Optimization of MapReduce Scheduling and Network Policy in Hierarchical Clouds. , 2018, , .		6
7115	RIOS. , 2018, , .		11
7116	Uncertainty Propagation in Data Processing Systems. , 2018, , .		7
7117	Efficient Memory Caching for Erasure Coding Based Key-Value Storage Systems. Communications in Computer and Information Science, 2018, , 512-539.	0.4	2
7118	Framework Design of Storage and Visualization System for Massive Unmanned Aerial Vehicle (UAV) Remote Sensing Data. , 2018, , .		1
7119	High-Performance Unsupervised Anomaly Detection for Cyber-Physical System Networks. , 2018, , .		51

#	ARTICLE	IF	CITATIONS
7120	Design of memory efficient FIFO-based merge sorter. IEICE Electronics Express, 2018, 15, 20171272-20171272.	0.3	1
7121	The Competence of Volunteer Computing for MapReduce Big Data Applications. Communications in Computer and Information Science, 2018, , 8-23.	0.4	2
7122	Distributed Search Architecture for Object Tracking in the Internet of Things. IEEE Access, 2018, 6, 60152-60159.	2.6	11
7123	Multicasting in a High-Port Sub- μ s Latency Hipo λ Optical Packet Switch. IEEE Photonics Technology Letters, 2018, 30, 1535-1538.	1.3	7
7124	BEANS – a software package for distributed Big Data analysis. Monthly Notices of the Royal Astronomical Society, 2018, 477, 3076-3090.	1.6	3
7125	Training Big Random Forests with Little Resources. , 2018, , .		8
7126	LogDrive: a proactive data collection and analysis framework for time-traveling forensic investigation in IaaS cloud environments. Journal of Cloud Computing: Advances, Systems and Applications, 2018, 7, .	2.1	8
7127	The Transition from A Priori to A Posteriori Information: Bayesian Procedures in Distributed Large-Scale Data Processing Systems. Automatic Documentation and Mathematical Linguistics, 2018, 52, 203-213.	0.2	0
7128	Recommendations for All: Solving Thousands of Recommendation Problems Daily. , 2018, , .		2
7129	Scheduling Coflows of Multi-stage Jobs to Minimize the Total Weighted Job Completion Time. , 2018, , .		25
7130	Utilizing MapReduce to Improve Probe-Car Track Data Mining. ISPRS International Journal of Geo-Information, 2018, 7, 287.	1.4	4
7131	A Distributed Parallel Algorithm Based on Low-Rank and Sparse Representation for Anomaly Detection in Hyperspectral Images. Sensors, 2018, 18, 3627.	2.1	6
7132	Exploration of Apache Hadoop Techniques: Mapreduce and Hive for Big Data. Communications in Computer and Information Science, 2018, , 543-552.	0.4	0
7133	The Datacenter as a Computer: Designing Warehouse-Scale Machines, Third Edition. Synthesis Lectures on Computer Architecture, 2018, 13, i-189.	1.3	54
7134	FITS Data Source for Apache Spark. Computing and Software for Big Science, 2018, 2, 1.	1.3	6
7135	Dynamic Prediction Research of Silicon Content in Hot Metal Driven by Big Data in Blast Furnace Smelting Process under Hadoop Cloud Platform. Complexity, 2018, 2018, 1-16.	0.9	28
7136	Complex Power System Status Monitoring and Evaluation Using Big Data Platform and Machine Learning Algorithms: A Review and a Case Study. Complexity, 2018, 2018, 1-21.	0.9	31
7137	High performance deep learning techniques for big data analytics. Concurrency Computation Practice and Experience, 2018, 30, e5032.	1.4	0

#	ARTICLE	IF	CITATIONS
7138	Mapping the study of learning analytics in higher education. Behaviour and Information Technology, 2018, 37, 1142-1155.	2.5	20
7139	Performance Analysis of Large-Scale Distributed Stream Processing Systems on the Cloud. , 2018, , .		5
7140	Parallel Deep Neural Network for Detecting Computer Attacks in Information Telecommunication Systems. , 2018, , .		4
7141	FlowHacker: Detecting Unknown Network Attacks in Big Traffic Data Using Network Flows. , 2018, , .		10
7142	Towards Update-Efficient and Parallel-Friendly Content-Based Indexing Scheme in Cloud Computing. International Journal of Semantic Computing, 2018, 12, 191-213.	0.4	1
7143	A Comparative Study of Join Algorithms in Mapreduce. SSRN Electronic Journal, 2018, , .	0.4	1
7144	Dual-Scheme Block Management to Trade Off Storage Overhead, Performance and Reliability. Communications in Computer and Information Science, 2018, , 462-476.	0.4	0
7145	Monitoring Data Integrity in Big Data Analytics Services. , 2018, , .		2
7146	An Innovative Framework for Supporting Cognitive-Based Big Data Analytics for Frequent Pattern Mining. , 2018, , .		12
7147	An Approach of RDD Optimization in Big Data Analytics. IOP Conference Series: Materials Science and Engineering, 2018, 396, 012022.	0.3	2
7148	A Robust Distributed Big Data Clustering-based on Adaptive Density Partitioning using Apache Spark. Symmetry, 2018, 10, 342.	1.1	11
7149	A Fog Computing-Oriented, Highly Scalable IoT Framework for Monitoring Public Educational Buildings. , 2018, , .		13
7150	Mobile Edge Computing Empowers Internet of Things. IEICE Transactions on Communications, 2018, E101.B, 604-619.	0.4	80
7151	Declarative debugging of concurrent Erlang programs. Journal of Logical and Algebraic Methods in Programming, 2018, 101, 22-41.	0.4	8
7152	Intermediate Data Caching Optimization for Multi-Stage and Parallel Big Data Frameworks. , 2018, , .		27
7153	Applications and Services. Urban Computing, 2018, , 147-166.	0.9	1
7154	Spark Clustering Computing Platform Based Parallel Particle Swarm Optimizers for Computationally Expensive Global Optimization. Lecture Notes in Computer Science, 2018, , 424-435.	1.0	2
7155	Consistency Maintenance in Distributed Analytical Stream Processing. Communications in Computer and Information Science, 2018, , 413-422.	0.4	0

#	ARTICLE	IF	CITATIONS
7156	A novel generated method for high quality images. , 2018, , .		1
7157	RDF Data Storage and Query Processing Schemes. ACM Computing Surveys, 2019, 51, 1-36.	16.1	69
7158	Analysis of Cybersecurity Threats in Industry 4.0: The Case of Intrusion Detection. Lecture Notes in Computer Science, 2018, , 119-130.	1.0	14
7159	MapReduce Based Analysis of Sample Applications Using Hadoop. Communications in Computer and Information Science, 2018, , 34-44.	0.4	2
7160	MapReduce service provisioning for frequent big data jobs on clouds considering data transfers. Computers and Electrical Engineering, 2018, 71, 594-610.	3.0	2
7161	DistGear: A Lightweight Event-Driven Framework for Developing Distributed Applications. , 2018, , .		0
7162	2DPR-Tree: Two-Dimensional Priority R-Tree Algorithm for Spatial Partitioning in SpatialHadoop. ISPRS International Journal of Geo-Information, 2018, 7, 179.	1.4	10
7163	Sampling Based Histogram PCA and Its Mapreduce Parallel Implementation on Multicore. Symmetry, 2018, 10, 162.	1.1	1
7164	Scalable Random Sampling K-Prototypes Using Spark. Lecture Notes in Computer Science, 2018, , 317-326.	1.0	2
7165	A New Combinatorial Design of Coded Distributed Computing. , 2018, , .		19
7166	Coded Distributed Computing with Node Cooperation Substantially Increases Speedup Factors. , 2018, , .		13
7167	Distributed Computing Trade-offs with Random Connectivity. , 2018, , .		19
7168	Hierarchical Coding for Distributed Computing. , 2018, , .		55
7169	Novel Data Segmentation Techniques for Efficient Discovery of Correlated Patterns Using Parallel Algorithms. Lecture Notes in Computer Science, 2018, , 355-370.	1.0	0
7170	Experiences on Teaching Parallel and Distributed Computing for Undergraduates. , 2018, , .		14
7171	The Performance Analysis of Cache Architecture Based on Alluxio over Virtualized Infrastructure. , 2018, , .		4
7172	Handling Vertex Deletions in Memory Scalable Anytime Anywhere Algorithms for Large and Dynamic Social Networks. , 2018, , .		3
7173	Hadoop based Demography Big Data Management System. , 2018, , .		5

#	ARTICLE	IF	CITATIONS
7174	Redundancy-Guaranteed and Receiving-Constrained Disaster Backup in Cloud Data Center Network. IEEE Access, 2018, 6, 47666-47681.	2.6	13
7175	Energy-Efficient Speculative Execution using Advanced Reservation for Heterogeneous Clusters. , 2018, , .		8
7176	Shed: Optimal Dynamic Cloning to Meet Application Deadlines in Cloud. , 2018, , .		1
7177	Sentiment Analysis using Novel Distributed Word Embedding for Movie Reviews. , 2018, , .		3
7178	Experience-Availability Analysis of Online Cloud Services using Stochastic Models. , 2018, , .		3
7179	Survey on Map Reduce Scheduling Algorithms in Hadoop Heterogeneous Environments. , 2018, , .		2
7180	High-performance Database Integrating Transaction and Analysis. , 2018, , .		0
7181	Integrating Hands-on Components into Database Curriculum at Undergraduate and Graduate Levels. , 2018, , .		0
7182	A Text Preprocessing Framework for Text Mining on Big Data Infrastructure. , 2018, , .		5
7183	A novel parallel distance metric-based approach for diversified ranking on large graphs. Future Generation Computer Systems, 2018, 88, 79-91.	4.9	2
7184	Efficient scheduling for synchronized demands in stochastic networks. , 2018, , .		1
7185	MR-Mafia: Parallel Subspace Clustering Algorithm Based on MapReduce for Large Multi-dimensional Datasets. , 2018, , .		1
7186	FUSO: Fast Multi-Path Loss Recovery for Data Center Networks. IEEE/ACM Transactions on Networking, 2018, 26, 1376-1389.	2.6	24
7187	C-Stream. ACM Transactions on Parallel Computing, 2017, 4, 1-27.	1.2	4
7188	HDUMP: A Data Recovery Tool for Hadoop. Lecture Notes in Computer Science, 2018, , 821-824.	1.0	0
7189	APRA: An approximate parallel recommendation algorithm for Big Data. Knowledge-Based Systems, 2018, 157, 10-19.	4.0	25
7190	Rock you like a hurricane. , 2018, , .		16
7191	Parallelizing Hartley transform with Hadoop for fast detection of glass defects. Concurrency Computation Practice and Experience, 2018, 30, e4499.	1.4	2

#	ARTICLE	IF	CITATIONS
7192	A Novel Clustering Method Using Enhanced Grey Wolf Optimizer and MapReduce. Big Data Research, 2018, 14, 93-100.	2.6	94
7193	Vs-Driven Big Data Process Development. Communications in Computer and Information Science, 2018, , 57-71.	0.4	0
7194	Scalable Architectures for Big Data Analysis. , 2018, , 1-9.		0
7195	Incremental Learning for Large Scale Classification Systems. IFIP Advances in Information and Communication Technology, 2018, , 112-122.	0.5	3
7196	Big-data-driven safety decision-making: A conceptual framework and its influencing factors. Safety Science, 2018, 109, 46-56.	2.6	57
7197	Big Data Systems Meet Machine Learning Challenges: Towards Big Data Science as a Service. Big Data Research, 2018, 14, 1-11.	2.6	96
7198	Tracking Causal Order in AWS Lambda Applications. , 2018, , .		19
7199	On the Semantics of Distributed Reactive Programming: The Cost of Consistency. IEEE Transactions on Software Engineering, 2018, 44, 689-711.	4.3	19
7200	Temporal features and relations discovery of activities from sensor data. Journal of Ambient Intelligence and Humanized Computing, 2024, 15, 1911-1926.	3.3	6
7201	OPTiC: Opportunistic Graph Processing in Multi-Tenant Clusters. , 2018, , .		0
7202	An energy-aware resource provisioning scheme for real-time applications in a cloud data center. Software - Practice and Experience, 2018, 48, 1734-1757.	2.5	9
7203	Big data for online learning systems. Education and Information Technologies, 2018, 23, 2783-2800.	3.5	30
7204	A cloud services recommendation system based on Fuzzy Formal Concept Analysis. Data and Knowledge Engineering, 2018, 116, 100-123.	2.1	36
7205	The 10 Vs, Issues and Challenges of Big Data. , 2018, , .		45
7206	Using Google Earth Engine to detect land cover change: Singapore as a use case. European Journal of Remote Sensing, 2018, 51, 486-500.	1.7	129
7207	Big Data and Data Analytics in Aviation. , 2018, , 55-65.		4
7208	A Big Data architecture for spectrum monitoring in cognitive radio applications. Annales Des Telecommunications/Annals of Telecommunications, 2018, 73, 451-461.	1.6	7
7209	Optimized distributed systems achieve significant performance improvement on sorted merging of massive VCF files. GigaScience, 2018, 7, .	3.3	4

#	ARTICLE	IF	CITATIONS
7210	Empya: Saving Energy in the Face of Varying Workloads. , 2018, , .		1
7211	Toward Transparent Data Management in Multi-Layer Storage Hierarchy of HPC Systems. , 2018, , .		7
7212	A Vector-Scheduling Approach for Running Many-Task Applications in the Cloud. Lecture Notes in Computer Science, 2018, , 3-19.	1.0	2
7213	Partition based clustering of large datasets using MapReduce framework: An analysis of recent themes and directions. Future Computing and Informatics Journal, 2018, 3, 247-261.	0.4	22
7214	Mutomvo: Mutation testing framework for simulated cloud and HPC environments. Journal of Systems and Software, 2018, 143, 187-207.	3.3	23
7215	Cycle-accelerated simulation for three-dimensional near-data processing system with power, temperature, and latency analysis. , 2018, , .		0
7216	Big Data Management for Healthcare Systems: Architecture, Requirements, and Implementation. Advances in Bioinformatics, 2018, 2018, 1-10.	5.7	65
7217	Power and performance optimization in FPGAâ€accelerated clouds. Concurrency Computation Practice and Experience, 2018, 30, e4526.	1.4	4
7218	A new MapReduce associative classifier based on a new storage format for large-scale imbalanced data. Cluster Computing, 2018, 21, 1821-1847.	3.5	5
7219	A distributed Frankâ€Wolfe framework for learning low-rank matrices with the trace norm. Machine Learning, 2018, 107, 1457-1475.	3.4	5
7220	Distributed storage system for electric power data based on Hbase. Big Data Mining and Analytics, 2018, 1, 324-334.	7.5	14
7221	Multiple query optimization approach based on hive+1. Journal of Intelligent and Fuzzy Systems, 2018, 35, 47-53.	0.8	0
7222	S2FA. , 2018, , .		15
7224	On the Use of Random Discretization and Dimensionality Reduction in Ensembles for Big Data. Lecture Notes in Computer Science, 2018, , 15-26.	1.0	1
7225	Cloud Based Face Recognition for Google Glass. , 2018, , .		10
7226	Computational Techniques in Data Integration and Big Data Handling in Omics. , 2018, , 209-222.		0
7227	Intelligent Data Placement in Heterogeneous Hadoop Cluster. Communications in Computer and Information Science, 2018, , 568-579.	0.4	9
7228	Application and Analysis of Multicast Blocking Modelling in Fat-Tree Data Center Networks. Complexity, 2018, 2018, 1-12.	0.9	2

#	ARTICLE	IF	CITATIONS
7229	Lessons Learned with Laser Scanning Point Cloud Management in Hadoop HBase. Lecture Notes in Computer Science, 2018, , 231-253.	1.0	5
7230	Congestion control in high-speed lossless data center networks: A survey. Future Generation Computer Systems, 2018, 89, 360-374.	4.9	34
7231	Partitioning and offloading in smart mobile devices for mobile cloud computing: State of the art and future directions. Journal of Network and Computer Applications, 2018, 119, 83-96.	5.8	30
7232	Distributed X-ray photon correlation spectroscopy data reduction using Hadoop <i>MapReduce</i> . Journal of Synchrotron Radiation, 2018, 25, 1135-1143.	1.0	17
7233	The Design, Evolution, and Use of KernelF. Lecture Notes in Computer Science, 2018, , 3-55.	1.0	7
7234	The adaptive complexity of maximizing a submodular function. , 2018, , .		35
7235	ex-FTCD: A novel mapreduce model for distributed multi source shortest path problem. Journal of Intelligent and Fuzzy Systems, 2018, 34, 1643-1652.	0.8	5
7236	A MapReduce-Based Parallel Frequent Pattern Growth Algorithm for Spatiotemporal Association Analysis of Mobile Trajectory Big Data. Complexity, 2018, 2018, 1-16.	0.9	13
7237	Decoupled MapReduce for Shared-Memory Multi-Core Architectures. IEEE Computer Architecture Letters, 2018, 17, 143-146.	1.0	0
7238	An Improved Bound for Minimizing the Total Weighted Completion Time of Coflows in Datacenters. IEEE/ACM Transactions on Networking, 2018, 26, 1674-1687.	2.6	32
7239	Adaptive Asynchronous Parallelization of Graph Algorithms. , 2018, , .		15
7240	A Two-Stage Data Processing Algorithm to Generate Random Sample Partitions for Big Data Analysis. Lecture Notes in Computer Science, 2018, , 347-364.	1.0	13
7241	ATUN-HL: Auto Tuning of Hybrid Layouts Using Workload and Data Characteristics. Lecture Notes in Computer Science, 2018, , 200-215.	1.0	1
7242	Stochastic non-preemptive co-flow scheduling with time-indexed relaxation. , 2018, , .		2
7243	Defining, enforcing and checking privacy policies in data-intensive applications. , 2018, , .		13
7244	Kavosh: an effective Map-Reduce-based association rule mining method. Journal of Big Data, 2018, 5, .	6.9	11
7245	Geometric Spanners in the MapReduce Model. Lecture Notes in Computer Science, 2018, , 675-687.	1.0	2
7246	Relieving Both Storage and Recovery Burdens in Big Data Clusters with R-STAIR Codes. IEEE Internet Computing, 2018, 22, 15-26.	3.2	6

#	ARTICLE	IF	CITATIONS
7247	A MapReduce-based Adjoint method for preventing brain disease. Journal of Big Data, 2018, 5, .	6.9	2
7249	A collection of Aquamaps native layers in NetCDF format. Data in Brief, 2018, 17, 292-296.	0.5	13
7250	HDM-MC in-Action: A Framework for Big Data Analytics across Multiple Clusters. , 2018, , .		0
7251	MapReduce-based pattern classification for design space analysis. , 2018, , .		2
7252	Parallel Primitives for Vendor-Agnostic Implementation of Big Data Mining Algorithms. , 2018, , .		0
7253	PFCA: An influence-based parallel fuzzy clustering algorithm for large complex networks. Expert Systems, 2018, 35, e12295.	2.9	1
7254	Big Data Clustering with Kernel <i>k</i> -Means: Resources, Time and Performance. International Journal on Artificial Intelligence Tools, 2018, 27, 1860006.	0.7	2
7255	Analytics for the Internet of Things. ACM Computing Surveys, 2019, 51, 1-36.	16.1	137
7256	Distributed Collaborative Execution on the Edges and Its Application to AMBER Alerts. IEEE Internet of Things Journal, 2018, 5, 3580-3593.	5.5	29
7257	BigRoots: An Effective Approach for Root-Cause Analysis of Stragglers in Big Data System. IEEE Access, 2018, 6, 41966-41977.	2.6	14
7258	Developing a big data analytics platform for manufacturing systems: architecture, method, and implementation. International Journal of Advanced Manufacturing Technology, 2018, 99, 2193-2217.	1.5	51
7259	Real-Time Graph Partition and Embedding of Large Network. , 2018, , .		3
7260	BAMSI: a multi-cloud service for scalable distributed filtering of massive genome data. BMC Bioinformatics, 2018, 19, 240.	1.2	4
7261	Towards Intelligent Distributed Data Systems for Scalable Efficient and Accurate Analytics. , 2018, , .		2
7262	Synthesizable Higher-Order Functions for C++. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2018, 37, 2835-2844.	1.9	8
7263	Ignem: Upward Migration of Cold Data in Big Data File Systems. , 2018, , .		2
7264	SDAC: A model for analysis of the execution semantics of data processing framework in cloud. Computer Languages, Systems and Structures, 2018, 54, 406-426.	1.4	1
7265	Towards a Scalable Software Defined Network-on-Chip for Next Generation Cloud. Sensors, 2018, 18, 2330.	2.1	13

#	ARTICLE	IF	CITATIONS
7266	A Binary Cuckoo Search Big Data Algorithm Applied to Large-Scale Crew Scheduling Problems. Complexity, 2018, 2018, 1-15.	0.9	46
7267	DIMPL: a distributed in-memory drone flight path builder system. Journal of Big Data, 2018, 5, .	6.9	4
7268	Infrastructure for Transportation Cyber-Physical Systems. , 2018, , 153-171.		1
7269	Conceptual Integrity Without Concepts. International Journal of Software Engineering and Knowledge Engineering, 2018, 28, 955-981.	0.6	1
7270	Naive Bayes Classifier Based Partitioner for MapReduce. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2018, E101.A, 778-786.	0.2	2
7271	G-HBase: A High Performance Geographical Database Based on HBase. IEICE Transactions on Information and Systems, 2018, E101.D, 1053-1065.	0.4	9
7272	A Distributed Indexing Method for Timeline Similarity Query. Algorithms, 2018, 11, 41.	1.2	5
7273	Large Scale Product Recommendation of Supermarket Ware Based on Customer Behaviour Analysis. Big Data and Cognitive Computing, 2018, 2, 11.	2.9	16
7274	Development of Easily Accessible Electricity Consumption Model Using Open Data and GA-SVR. Energies, 2018, 11, 373.	1.6	9
7275	Big Data Analytics for Discovering Electricity Consumption Patterns in Smart Cities. Energies, 2018, 11, 683.	1.6	79
7276	A gray-box performance model for Apache Spark. Future Generation Computer Systems, 2018, 89, 58-67.	4.9	22
7277	RDF Query Answering Using Apache Spark: Review and Assessment. , 2018, , .		13
7278	Automatically Leveraging MapReduce Frameworks for Data-Intensive Applications. , 2018, , .		22
7279	Round compression for parallel matching algorithms. , 2018, , .		36
7280	Cloud agnostic Big Data platform focusing on scalability and cost-efficiency. Advances in Engineering Software, 2018, 125, 167-177.	1.8	10
7282	Distributed control framework for mapreduce cloud on cloud computing. , 2018, , .		2
7283	A comparison of Hadoop, Spark and Storm for the task of large scale image classification. , 2018, , .		9
7284	Big Data Bayesian Linear Regression and Variable Selection by Normal-Inverse-Gamma Summation. Bayesian Analysis, 2018, 13, .	1.6	3

#	ARTICLE	IF	CITATIONS
7285	High Performance Methods for Linked Open Data Connectivity Analytics. Information (Switzerland), 2018, 9, 134.	1.7	9
7286	Distributed Fast Self-Organized Maps for Massive Spectrophotometric Data Analysis. Sensors, 2018, 18, 1419.	2.1	3
7287	An Empirical Study on Visualizing the Intellectual Structure and Hotspots of Big Data Research from a Sustainable Perspective. Sustainability, 2018, 10, 667.	1.6	19
7288	Programmatic ETL. Lecture Notes in Business Information Processing, 2018, , 21-50.	0.8	1
7289	Janus: Diagnostics and reconfiguration of data parallel programs. Journal of Parallel and Distributed Computing, 2018, 120, 196-210.	2.7	2
7290	Big data time series forecasting based on nearest neighbours distributed computing with Spark. Knowledge-Based Systems, 2018, 161, 12-25.	4.0	40
7291	Divide and Conquer. , 2018, , .		0
7292	Towards safe refactoring for intelligent parallelization of Java 8 streams. , 2018, , .		8
7293	Large-Scale Simulations of Bacterial Populations Over Complex Networks. Journal of Computational Biology, 2018, 25, 850-861.	0.8	1
7294	Big Data and Its Key Technology in the Future. Computing in Science and Engineering, 2018, 20, 75-88.	1.2	8
7295	Efficient Stream Processing of Scientific Data. , 2018, , .		1
7296	From raw data packets to ingress egress traffic matrix: The distributed MapReduce-based solution. , 2018, , .		3
7297	Batch and online anomaly detection for scientific applications in a Kubernetes environment. , 2018, , .		11
7298	A Locality-aware Task Scheduling Algorithm for Video Transcoding over Heterogenous MapReduce Cluster. , 2018, , .		0
7299	Application deployment using Microservice and Docker containers: Framework and optimization. Journal of Network and Computer Applications, 2018, 119, 97-109.	5.8	61
7300	Lazy collaborative filtering with dynamic neighborhoods. Information Discovery and Delivery, 2018, 46, 95-109.	1.6	3
7301	Efficient Messaging for Java Applications Running in Data Centers. , 2018, , .		3
7302	A systematic fault-tolerant computational model for both crash failures and silent data corruption. , 2018, , .		0

#	ARTICLE	IF	CITATIONS
7303	An integer linear-programming based resource allocation method for SQL-like queries in the cloud. , 2018, , .		5
7304	An efficient theta-join query processing in distributed environment. Journal of Parallel and Distributed Computing, 2018, 121, 42-52.	2.7	2
7305	Hadoop Configuration Tuning With Ensemble Modeling and Metaheuristic Optimization. IEEE Access, 2018, 6, 44161-44174.	2.6	20
7306	Automated Extractions for Machine Generated Mail. , 2018, , .		7
7307	Robust Fingerprinting Method for Webtoon Identification in Large-Scale Databases. IEEE Access, 2018, 6, 37932-37946.	2.6	4
7308	On Conceptually Simple Algorithms for Variants of Online Bipartite Matching. Lecture Notes in Computer Science, 2018, , 253-268.	1.0	2
7309	Advances in Automation Technologies for Lower Extremity Neurorehabilitation: A Review and Future Challenges. IEEE Reviews in Biomedical Engineering, 2018, 11, 289-305.	13.1	43
7310	A Prototype Model for Resource Provisioning in Cloud Computing Using MapReduce Technique. Advances in Intelligent Systems and Computing, 2018, , 971-977.	0.5	0
7311	The MapReduce-based approach to improve the shortest path computation in large-scale road networks: the case of A* algorithm. Journal of Big Data, 2018, 5, .	6.9	11
7312	DCF: A Dataflow-Based Collaborative Filtering Training Algorithm. International Journal of Parallel Programming, 2018, 46, 686-698.	1.1	0
7313	A Modeling Language for MapReduce Programing in a Storage System Perspective. Journal of Signal Processing Systems, 2018, 90, 1133-1150.	1.4	0
7314	A survey of parallel particle tracing algorithms in flow visualization. Journal of Visualization, 2018, 21, 351-368.	1.1	6
7315	Big Data Analytics in Medicine and Healthcare. Journal of Integrative Bioinformatics, 2018, 15, .	1.0	188
7316	OEHadoop: Accelerate Hadoop Applications by Co-Designing Hadoop With Data Center Network. IEEE Access, 2018, 6, 25849-25860.	2.6	10
7317	BiGNoC: Accelerating Big Data Computing with Application-Specific Photonic Network-on-Chip Architectures. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 2402-2415.	4.0	17
7318	TSSA: A two step scheduling algorithm for the event-driven clusters. , 2018, , .		1
7319	Dataset Popularity Prediction for Caching of CMS Big Data. Journal of Grid Computing, 2018, 16, 211-228.	2.5	13
7320	Towards attack-resistant Aggregate Computing using trust mechanisms. Science of Computer Programming, 2018, 167, 114-137.	1.5	21

#	ARTICLE	IF	CITATIONS
7321	Cost-Aware Streaming Data Analysis. , 2018, , .		1
7322	An efficient placement of sinks and SDN controller nodes for optimizing the design cost of industrial IoT systems. Software - Practice and Experience, 2018, 48, 1893-1919.	2.5	14
7323	Dataflow management, dynamic load balancing, and concurrent processing for real-time embedded vision applications using Quasar. International Journal of Circuit Theory and Applications, 2018, 46, 1733-1755.	1.3	6
7324	New Techniques for Coding Political Events across Languages. , 2018, , .		6
7325	Chronos: A Unifying Optimization Framework for Speculative Execution of Deadline-Critical MapReduce Jobs. , 2018, , .		9
7326	Multi-Attributes-Based Coflow Scheduling Without Prior Knowledge. IEEE/ACM Transactions on Networking, 2018, 26, 1962-1975.	2.6	12
7327	An Efficient Approach to Improve Security for MapReduce Computation in Cloud System. , 2018, , .		1
7328	Data-Driven Parallel Video Transcoding for Content Delivery Network in the Cloud. , 2018, , .		4
7329	RDF Data Storage Techniques for Efficient SPARQL Query Processing Using Distributed Computation Engines. , 2018, , .		6
7330	Parallel and distributed association rule mining in life science: A novel parallel algorithm to mine genomics data. Information Sciences, 2021, 575, 747-761.	4.0	17
7331	GMC: Greening MapReduce Clusters Considering Both Computational Energy and Cooling Energy. , 2018, , .		1
7332	On Implementing the Push-Relabel Algorithm on Top of Pregel. New Generation Computing, 2018, 36, 419-449.	2.5	2
7333	A Task-Based Greedy Scheduling Algorithm for Minimizing Energy of MapReduce Jobs. Journal of Grid Computing, 2018, 16, 535-551.	2.5	9
7334	Hadoop-based analytic framework for cyber forensics. International Journal of Communication Systems, 2018, 31, e3772.	1.6	13
7335	Informal Learning in Twitter: Architecture of Data Analysis Workflow and Extraction of Top Group of Connected Hashtags. Communications in Computer and Information Science, 2018, , 3-15.	0.4	0
7336	Transforming Healthcare Via Big Data Analytics. , 2018, , 319-334.		0
7337	Planning of distributed data production for High Energy and Nuclear Physics. Cluster Computing, 2018, 21, 1949-1965.	3.5	0
7338	Exploring Textures in Traffic Matrices to Classify Data Center Communications. , 2018, , .		2

#	ARTICLE	IF	CITATIONS
7339	Big Data Analytics Using Multi-Classifer Approach with Rhadoop. , 2018, , .		1
7340	Can Adaptive Feedforward Control Improve Operation of Cloud Services?. , 2018, , .		0
7341	Big Data-Oriented PaaS Architecture with Disk-as-a-Resource Capability and Container-Based Virtualization. Journal of Grid Computing, 2018, 16, 587-605.	2.5	3
7342	The multi-scale high-order statistical moments of financial time series. Physica A: Statistical Mechanics and Its Applications, 2018, 512, 474-488.	1.2	2
7343	Efficient linear fusion of partial estimators. , 2018, 78, 265-283.		22
7344	Big data quality prediction in the process industry: A distributed parallel modeling framework. Journal of Process Control, 2018, 68, 1-13.	1.7	78
7345	MASKED: A MapReduce Solution for the Kappa-Pruned Ensemble-Based Anomaly Detection System. , 2018, , .		0
7346	A survey of state management in big data processing systems. VLDB Journal, 2018, 27, 847-872.	2.7	46
7347	Progressive Image Retrieval With Quality Guarantee Under MapReduce Framework. IEEE Access, 2018, 6, 44685-44697.	2.6	5
7348	Real-time data analysis using Spark and Hadoop. , 2018, , .		31
7349	Software-defined data services: Interoperable and network-aware big data executions. , 2018, , .		1
7350	Toward a MapReduce-Based K-Means Method for Multi-dimensional Time Serial Data Clustering. Advances in Intelligent Systems and Computing, 2018, , 816-825.	0.5	1
7351	Spline: Spark Lineage, not only for the Banking Industry. , 2018, , .		4
7352	An Efficient Triangle Enumeration on Parallel and Distributed Frameworks. , 2018, , .		0
7353	Towards a Cloud Computing Paradigm for Big Data Analysis in Smart Cities. Programming and Computer Software, 2018, 44, 181-189.	0.5	43
7354	Profiling distributed systems in lightweight virtualized environments with logs and resource metrics. , 2018, , .		7
7356	RT-DBSCAN: Real-Time Parallel Clustering of Spatio-Temporal Data Using Spark-Streaming. Lecture Notes in Computer Science, 2018, , 524-539.	1.0	10
7357	Mining of productive periodic-frequent patterns for IoT data analytics. Future Generation Computer Systems, 2018, 88, 512-523.	4.9	25

#	ARTICLE	IF	CITATIONS
7358	Fast and peer-to-peer vital signal learning system for cloud-based healthcare. Future Generation Computer Systems, 2018, 88, 220-233.	4.9	14
7359	Monitoring big process data of industrial plants with multiple operating modes based on Hadoop. Journal of the Taiwan Institute of Chemical Engineers, 2018, 91, 10-21.	2.7	25
7360	A survey of RDF management technologies and benchmark datasets. Journal of Ambient Intelligence and Humanized Computing, 2018, 9, 1693-1704.	3.3	17
7361	Hound. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2018, 2, 1-36.	1.4	13
7362	Skyway. , 2018, , .		23
7363	Discovering Progression Stages in Trillion-Scale Behavior Logs. , 2018, , .		1
7364	Classifying Big Data Analytic Approaches: A Generic Architecture. Communications in Computer and Information Science, 2018, , 268-295.	0.4	3
7365	DimBoost. , 2018, , .		27
7366	Prototyping a Web-Scale Multimedia Retrieval Service Using Spark. ACM Transactions on Multimedia Computing, Communications and Applications, 2018, 14, 1-24.	3.0	8
7367	Spark Based ANFIS Approach for Anomaly Detection Using Big Data. Communications in Computer and Information Science, 2018, , 450-458.	0.4	0
7368	Study on information recommendation of scientific and technological achievements based on user behavior modeling and big data mining. , 2018, , .		2
7369	Scalable Implementations of Rough Set Algorithms: A Survey. Lecture Notes in Computer Science, 2018, , 648-660.	1.0	2
7370	Elastic CPU Cap Mechanism for Timely Dataflow Applications. Lecture Notes in Computer Science, 2018, , 554-568.	1.0	1
7371	A novel spark-based multi-step forecasting algorithm for big data time series. Information Sciences, 2018, 467, 800-818.	4.0	28
7372	A literature review on the key technologies of processing big data. , 2018, , .		3
7373	Spatial hotspot detection using polygon propagation. International Journal of Digital Earth, 2019, 12, 825-842.	1.6	2
7374	The optimization of parallel convolutional RBM based on Spark. International Journal of Wavelets, Multiresolution and Information Processing, 2019, 17, 1940011.	0.9	5
7375	Scalable machine learning computing a data summarization matrix with a parallel array DBMS. Distributed and Parallel Databases, 2019, 37, 329-350.	1.0	9

#	ARTICLE	IF	CITATIONS
7376	Earth Science [Big] Data Analytics. , 2019, , 99-128.		3
7377	Technology in the 21st century: New challenges and opportunities. Technological Forecasting and Social Change, 2019, 143, 321-335.	6.2	36
7378	A data locality based scheduler to enhance MapReduce performance in heterogeneous environments. Future Generation Computer Systems, 2019, 90, 423-434.	4.9	43
7379	Locality-aware and load-balanced static task scheduling for MapReduce. Future Generation Computer Systems, 2019, 90, 49-61.	4.9	18
7380	A study of big data evolution and research challenges. Journal of Information Science, 2019, 45, 322-340.	2.0	67
7382	C2Net: A Network-Efficient Approach to Collision Counting LSH Similarity Join. IEEE Transactions on Knowledge and Data Engineering, 2019, 31, 423-436.	4.0	9
7383	FPGA-Based Hardware Accelerator for an Embedded Factor Graph with Configurable Optimization. Journal of Circuits, Systems and Computers, 2019, 28, 1950031.	1.0	7
7384	Nxt-Max: for supporting VDC-based maximum redundant bandwidth allocation in cloud datacenter. Cluster Computing, 2019, 22, 12567-12580.	3.5	2
7385	Scalable and parallel sequential pattern mining using spark. World Wide Web, 2019, 22, 295-324.	2.7	17
7386	Real-time data processing scheme using big data analytics in internet of things based smart transportation environment. Journal of Ambient Intelligence and Humanized Computing, 2019, 10, 4167-4177.	3.3	75
7387	Dandelion: A Unified Code Offloading System for Wearable Computing. IEEE Transactions on Mobile Computing, 2019, 18, 546-559.	3.9	18
7388	MapReduce in Computational Biology Via Hadoop and Spark. , 2019, , 221-229.		2
7389	DtCraft: A High-Performance Distributed Execution Engine at Scale. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2019, 38, 1070-1083.	1.9	10
7390	Discovering and forecasting interactions in big data research: A learning-enhanced bibliometric study. Technological Forecasting and Social Change, 2019, 146, 795-807.	6.2	50
7391	Computing over encrypted spatial data generated by IoT. Telecommunication Systems, 2019, 70, 193-229.	1.6	4
7392	Changing Paradigm in Map Production and Geoinformation Management—An Introduction. Lecture Notes in Geoinformation and Cartography, 2019, , 3-42.	0.5	0
7393	Research on Data Storage and Processing Optimization Based on Federation HDFS and Spark. Advances in Intelligent Systems and Computing, 2019, , 1047-1057.	0.5	0
7394	Probabilistic movement models and zones of control. Machine Learning, 2019, 108, 127-147.	3.4	24

#	ARTICLE	IF	CITATIONS
7395	RTSBL: Reduce Task Scheduling Based on the Load Balancing and the Data Locality in Hadoop. Lecture Notes in Networks and Systems, 2019, , 271-280.	0.5	1
7396	Study on the country risk rating with distributed crawling system. Journal of Supercomputing, 2019, 75, 6159-6177.	2.4	3
7397	Container-based bioinformatics with Pachyderm. Bioinformatics, 2019, 35, 839-846.	1.8	35
7398	Analysis of student behavior in learning management systems through a Big Data framework. Future Generation Computer Systems, 2019, 90, 262-272.	4.9	99
7399	A novel mapreduce algorithm for distributed mining of sequential patterns using co-occurrence information. Applied Intelligence, 2019, 49, 150-171.	3.3	12
7400	A Survey: Classification of Big Data. Advances in Intelligent Systems and Computing, 2019, , 299-306.	0.5	4
7401	A formally based parallelization of data mining algorithms for multi-core systems. Journal of Supercomputing, 2019, 75, 7909-7920.	2.4	3
7402	An Efficient Framework for Smart City Using Big Data Technologies and Internet of Things. Advances in Intelligent Systems and Computing, 2019, , 319-328.	0.5	8
7403	VisHive: Supporting web-based visualization through ad hoc computational clusters of mobile devices. Information Visualization, 2019, 18, 195-210.	1.2	6
7404	Microtask Programming. IEEE Transactions on Software Engineering, 2019, 45, 1106-1124.	4.3	9
7405	A Hybrid MapReduce Implementation of PCA on Tianhe-2. Journal of Physics: Conference Series, 2019, 1168, 052013.	0.3	0
7406	Subgroup discovery in MOOCs: a big data application for describing different types of learners. Interactive Learning Environments, 2019, , 1-19.	4.4	9
7407	Building A Scalable Forward Flux Sampling Framework using Big Data and HPC. , 2019, , .		3
7408	Design Exploration of Multi-tier Interconnection Networks for Exascale Systems. , 2019, , .		2
7409	Occam's Razor for Big Data? On Detecting Quality in Large Unstructured Datasets. Applied Sciences (Switzerland), 2019, 9, 3065.	1.3	15
7410	PARSUC: A Parallel Subsampling-Based Method for Clustering Remote Sensing Big Data. Sensors, 2019, 19, 3438.	2.1	8
7411	Examining characteristics of predictive models with imbalanced big data. Journal of Big Data, 2019, 6, .	6.9	23
7412	Aggregation rules based on stochastic gradient descent in Byzantine consensus. , 2019, , .		0

#	ARTICLE	IF	CITATIONS
7413	An Automated Framework for Explaining Facts Extracted From Mobility Datasets. , 2019, , .		0
7414	Toward Efficient Processing of Spatio-Temporal Workloads in a Distributed In-Memory System. , 2019, , .		6
7415	An Executable Specification of Map-Join-Reduce Using Haskell. IEEE Access, 2019, 7, 10892-10904.	2.6	3
7416	A Deep Dissertation of Data Science: Related Issues and its Applications. , 2019, , .		6
7417	On Conceptually Simple Algorithms for Variants of Online Bipartite Matching. Theory of Computing Systems, 2019, 63, 1781-1818.	0.7	1
7418	Network representation learning: models, methods and applications. SN Applied Sciences, 2019, 1, 1.	1.5	9
7419	Pillar K-Means Clustering Algorithm Using MapReduce Framework. IOP Conference Series: Earth and Environmental Science, 2019, 258, 012031.	0.2	0
7420	Hybrid Resource Management for HPC and Data Intensive Workloads. , 2019, , .		5
7421	A data structure perspective to the RDD-based Apriori algorithm on Spark. International Journal of Information Technology (Singapore), 2022, 14, 1585-1594.	1.8	2
7422	Big Data Issues in Smart Grids: A Survey. IEEE Systems Journal, 2019, 13, 4158-4168.	2.9	79
7423	A Flexible Heuristic to Schedule Distributed Analytic Applications in Compute Clusters. IEEE Transactions on Cloud Computing, 2019, , 1-1.	3.1	0
7424	Near Optimal Coded Data Shuffling for Distributed Learning. IEEE Transactions on Information Theory, 2019, 65, 7325-7349.	1.5	20
7425	Performance Evaluation of Big Data Processing Strategies for Neuroimaging. , 2019, , .		3
7426	A Scalable Automated Diagnostic Feature Extraction System for EEGs. , 2019, , .		0
7427	Overhead Aware Task Scheduling for Cloud Container Services. , 2019, , .		6
7428	MPR " A Partitioning-Replication Framework for Multi-Processing kNN Search on Road Networks. , 2019, , .		4
7429	Stage Delay Scheduling. , 2019, , .		13
7430	Search Engine for the Internet of Things: Lessons From Web Search, Vision, and Opportunities. IEEE Access, 2019, 7, 104673-104691.	2.6	28

#	ARTICLE	IF	CITATIONS
7431	Load-balancing distributed outer joins through operator decomposition. <i>Journal of Parallel and Distributed Computing</i> , 2019, 132, 21-35.	2.7	4
7432	DREAM. , 2019, , .		7
7433	A Probabilistic Range Query of Moving Objects in Road Network. <i>IEEE Access</i> , 2019, 7, 40165-40174.	2.6	1
7434	Efficient Distributed Range Query Processing in Apache Spark. , 2019, , .		5
7435	Transforming Query Sequences for High-Throughput B+ Tree Processing on Many-Core Processors. , 2019, , .		1
7436	Fast and Efficient Distributed Matrix-vector Multiplication Using Rateless Fountain Codes. , 2019, , .		24
7437	It Can Understand the Logs, Literally. , 2019, , .		4
7438	â€œShort-Dotâ€ Computing Large Linear Transforms Distributedly Using Coded Short Dot Products. <i>IEEE Transactions on Information Theory</i> , 2019, 65, 6171-6193.	1.5	75
7439	Gapprox: using Gallup approach for approximation in Big Data processing. <i>Journal of Big Data</i> , 2019, 6, .	6.9	27
7440	Big data in healthcare: management, analysis and future prospects. <i>Journal of Big Data</i> , 2019, 6, .	6.9	722
7441	Slow Replica and Shared Protection: Energy-Efficient and Reliable Task Assignment in Cloud Data Centers. <i>IEEE Transactions on Reliability</i> , 2021, 70, 931-943.	3.5	2
7442	RaSQL. , 2019, , .		20
7443	AssistantGraph: An Approach for Reusable and Composable Data-Driven Assistant Components. , 2019, , .		3
7444	Fast. Efficient Performance Predictions for Big Data Applications. , 2019, , .		1
7445	A Heuristic Virtual Machine Scheduling Algorithm in Cloud Data Center. , 2019, , .		3
7446	Frequent itemset mining: A 25 years review. <i>Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery</i> , 2019, 9, e1329.	4.6	138
7447	ISEE: An Intelligent Scene Exploration and Evaluation Platform for Large-Scale Visual Surveillance. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2019, 30, 2743-2758.	4.0	11
7448	Large scale microbiome profiling in the cloud. <i>Bioinformatics</i> , 2019, 35, i13-i22.	1.8	9

#	ARTICLE	IF	CITATIONS
7449	In Search of Big Medical Data Integration Solutions - A Comprehensive Survey. IEEE Access, 2019, 7, 91265-91290.	2.6	40
7450	Exact-Differential Simulation. ACM Transactions on Modeling and Computer Simulation, 2019, 29, 1-25.	0.6	1
7451	Polymorphic edge detection (PED): two efficient methods of polymorphism detection from next-generation sequencing data. BMC Bioinformatics, 2019, 20, 362.	1.2	3
7452	ObliDC. , 2019, , .		7
7453	A system for real-time interactive analysis of deep learning training. , 2019, , .		5
7454	An Improvement of a Checkpoint-based Distributed Testing Technique on a Big Data Environment. , 2019, , .		2
7455	Mitigating stragglers to avoid QoS violation for time-critical applications through dynamic server blacklisting. Future Generation Computer Systems, 2019, 101, 831-842.	4.9	7
7456	Per-point processing for detailed urban solar estimation with aerial laser scanning and distributed computing. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 155, 119-135.	4.9	16
7457	Evaluation of distributed stream processing frameworks for IoT applications in Smart Cities. Journal of Big Data, 2019, 6, .	6.9	66
7458	A Solution to Scientific Analysis Service in Data Cloud. , 2019, , .		0
7459	SLOPE: Structural Locality-Aware Programming Model for Composing Array Data Analysis. Lecture Notes in Computer Science, 2019, , 61-80.	1.0	2
7460	Active learning to detect DDoS attack using ranked features. Computer Communications, 2019, 145, 203-222.	3.1	30
7461	Performance Scaling of Cassandra on High-Thread Count Servers. , 2019, , .		2
7462	Job-Aware Optimization of File Placement in Hadoop. , 2019, , .		4
7463	Calculi for symmetric queries. Journal of Computer and System Sciences, 2019, 105, 54-86.	0.9	1
7464	Parallel Discovery of Trajectory Companions from Heterogeneous Streaming Data. , 2019, , .		1
7465	Distributed Deep Learning Strategies for Automatic Speech Recognition. , 2019, , .		17
7466	Wireless MapReduce Distributed Computing. IEEE Transactions on Information Theory, 2019, 65, 6101-6114.	1.5	37

#	ARTICLE	IF	CITATIONS
7467	Map-Matching on Big Data: a Distributed and Efficient Algorithm with a Hidden Markov Model. , 2019, , .		3
7468	DFix: automatically fixing timing bugs in distributed systems. , 2019, , .		10
7469	Capacity based deadline aware dynamic load balancing (CPDALB) model in cloud computing environment. International Journal of Computers and Applications, 2021, 43, 987-1001.	0.8	8
7470	Trade-off Between Energy Consumption and Makespan in the Mapreduce Resource Allocation Problem. Lecture Notes in Computer Science, 2019, , 239-250.	1.0	6
7471	Big Data technologies to process spatial and attribute data when designing and operating mine-engineering systems. International Journal of Image and Data Fusion, 2019, 10, 217-231.	0.8	1
7472	Real-time intelligent big data processing: technology, platform, and applications. Science China Information Sciences, 2019, 62, 1.	2.7	32
7473	Ignis: scaling distribution-oblivious systems with light-touch distribution. , 2019, , .		4
7474	c 2 AIDER : cognitive cloud exoskeleton system and its applications. Cognitive Computation and Systems, 2019, 1, 33-39.	0.8	9
7475	When Relational-Based Applications Go to NoSQL Databases: A Survey. Information (Switzerland), 2019, 10, 241.	1.7	4
7476	State of the art in big data applications in microgrid: A review. Advanced Engineering Informatics, 2019, 42, 100945.	4.0	52
7477	A survey on quality-assurance approximate stream processing and applications. Future Generation Computer Systems, 2019, 101, 1062-1080.	4.9	4
7478	When Crowdsourcing Meets Social IoT: An Efficient Privacy-Preserving Incentive Mechanism. IEEE Internet of Things Journal, 2019, 6, 9707-9721.	5.5	14
7479	Finding Users Who Act Alike. , 2019, , .		13
7480	iQCAR. , 2019, 2019, 918-935.		4
7481	On the Optimal Number of Computational Resources in MapReduce. Lecture Notes in Computer Science, 2019, , 240-252.	1.0	0
7482	Cluster-Scheduling Big Graph Traversal Task for Parallel Processing in Heterogeneous Cloud Based on DAG Transformation. IEEE Access, 2019, 7, 77070-77082.	2.6	8
7483	Two-Level Fault Diagnosis of SF6 Electrical Equipment Based on Big Data Analysis. Big Data and Cognitive Computing, 2019, 3, 4.	2.9	6
7484	Science for everyone (ScifE): A proposed framework for science as a service using interactive web technologies. Computers and Geosciences, 2019, 131, 70-79.	2.0	5

#	ARTICLE	IF	CITATIONS
7485	EPSim-C: A Parallel Epoch-Based Cycle-Accurate Microarchitecture Simulator Using Cloud Computing. Electronics (Switzerland), 2019, 8, 716.	1.8	2
7486	A cognitive technology based healthcare monitoring system and medical data transmission. Measurement: Journal of the International Measurement Confederation, 2019, 146, 322-332.	2.5	27
7487	Semantic-aware Workflow Construction and Analysis for Distributed Data Analytics Systems. , 2019, , .		5
7488	K-means Bayes algorithm for imbalanced fault classification and big data application. Journal of Process Control, 2019, 81, 54-64.	1.7	46
7489	Research of Flexible Load Analysis of Distribution Network Based on Big Data. , 2019, , .		2
7490	MIFO. , 2019, , .		3
7491	A Study on the Performance and Scalability of Apache Flink Over Hadoop MapReduce. International Journal of Fog Computing, 2019, 2, 61-73.	1.8	0
7492	DGST: Efficient and scalable suffix tree construction on distributed data-parallel platforms. Parallel Computing, 2019, 87, 87-102.	1.3	7
7493	Distributed Hayabusa. , 2019, , .		2
7494	Leveraging resource management for efficient performance of Apache Spark. Journal of Big Data, 2019, 6, .	6.9	20
7495	Revisiting Resource Management for Deep Learning Framework. Electronics (Switzerland), 2019, 8, 327.	1.8	0
7496	Maximizing MapReduce job speed and reliability in the mobile cloud by optimizing task allocation. Pervasive and Mobile Computing, 2019, 60, 101082.	2.1	12
7497	A parallel clustering algorithm for logs data based on Hadoop platform. , 2019, , .		2
7498	Investigating Random Undersampling and Feature Selection on Bioinformatics Big Data. , 2019, , .		18
7499	A Survey on Data Placement Strategy in Big Data Heterogeneous Environments. , 2019, , .		2
7500	Trust-Based Scheduling Framework for Big Data Processing with MapReduce. IEEE Transactions on Services Computing, 2022, 15, 279-293.	3.2	11
7501	Better Late Than Never. , 2019, , .		0
7502	UserDEV. , 2019, , .		1

#	ARTICLE	IF	CITATIONS
7503	A fault-tolerant programming model for distributed interactive applications. , 2019, 3, 1-29.		13
7504	Multivariate Big Data Analysis for intrusion detection: 5 steps from the haystack to the needle. Computers and Security, 2019, 87, 101603.	4.0	22
7505	Parallel Heuristics for Balanced Graph Partitioning Based on Richness of Implicit Knowledge. IEEE Access, 2019, 7, 96444-96454.	2.6	6
7506	A Resource Aware Parallelized Back Propagation Neural Network in Enabling Efficient Large-Scale Digital Health Data Processing. IEEE Access, 2019, 7, 114700-114713.	2.6	5
7507	Efficient scheduling for multi-stage coflows. CCF Transactions on Networking, 2019, 2, 83-97.	1.0	1
7508	Reco: Efficient Regularization-Based Coflow Scheduling in Optical Circuit Switches. , 2019, , .		5
7509	A QoS-oriented Scheduling and Autoscaling Framework for Deep Learning. , 2019, , .		0
7510	AutoGRD. , 2019, , .		21
7511	A Comprehensive Review of Access Control Mechanism Based on Attribute Based Encryption Scheme for Cloud Computing. International Journal of Advanced Pervasive and Ubiquitous Computing, 2019, 11, 33-52.	0.4	1
7512	The Zeitgeist Juncture of "Big Data" and its Future Trends. , 2019, , .		4
7513	A Comprehensive Survey on Secure Outsourced Computation and Its Applications. IEEE Access, 2019, 7, 159426-159465.	2.6	30
7514	Mapping Heterogeneity Does Not Affect Wireless Coded MapReduce. , 2019, , .		3
7515	Aegean. , 2019, , .		8
7516	Zone of control. , 2019, , .		1
7517	From Big to Smart Data: Iterative ensemble filter for noise filtering in Big Data classification. International Journal of Intelligent Systems, 2019, 34, 3260-3274.	3.3	12
7518	A cybermanufacturing and AI framework for laser powder bed fusion (LPBF) additive manufacturing process. Manufacturing Letters, 2019, 21, 41-44.	1.1	16
7519	Accuracy Prediction for Distributed Decision Tree using Machine Learning approach. , 2019, , .		24
7520	Decoding Parallel Program Execution by using Java Interactive Visualization Environment (JIVE): Behavioral and Performance Analysis. , 2019, , .		7

#	ARTICLE	IF	CITATIONS
7521	A distributed approximate nearest neighbors algorithm for efficient large scale mean shift clustering. Journal of Parallel and Distributed Computing, 2019, 134, 128-139.	2.7	25
7522	Distributed Selection of Continuous Features in Multilabel Classification Using Mutual Information. IEEE Transactions on Neural Networks and Learning Systems, 2019, 31, 1-14.	7.2	15
7523	Towards an Energy Efficient Computing With Coordinated Performance-Aware Scheduling in Large Scale Data Clusters. IEEE Access, 2019, 7, 140261-140277.	2.6	4
7524	Understanding Edge Computing: Engineering Evolution With Artificial Intelligence. IEEE Access, 2019, 7, 164229-164245.	2.6	85
7525	A Survey of Distributed Data Stream Processing Frameworks. IEEE Access, 2019, 7, 154300-154316.	2.6	84
7526	ML4IoT: A Framework to Orchestrate Machine Learning Workflows on Internet of Things Data. IEEE Access, 2019, 7, 152953-152967.	2.6	12
7527	Shadow: Exploiting the Power of Choice for Efficient Shuffling in MapReduce. IEEE Transactions on Big Data, 2022, 8, 253-267.	4.4	3
7528	d-Simplex: Adaptive Delaunay Triangulation for Performance Modeling and Prediction on Big Data Analytics. IEEE Transactions on Big Data, 2019, , 1-1.	4.4	14
7529	Pream: Enhancing HPC Storage System Performance with Pre-Allocated Metadata Management Mechanism. , 2019, , .		1
7530	Evolving large-scale data stream analytics based on scalable PANFIS. Knowledge-Based Systems, 2019, 166, 186-197.	4.0	7
7531	An Optimal Solution of Storing and Processing Small Image Files on Hadoop. Procedia Computer Science, 2019, 154, 581-587.	1.2	0
7532	EmuEdge: A Hybrid Emulator for Reproducible and Realistic Edge Computing Experiments. , 2019, , .		20
7533	Boa Meets Python: A Boa Dataset of Data Science Software in Python Language. , 2019, , .		14
7534	DeepStore. , 2019, , .		30
7535	Multi-layer intrusion detection system with ExtraTrees feature selection, extreme learning machine ensemble, and softmax aggregation. Eurasip Journal on Information Security, 2019, 2019, .	2.4	41
7536	The Application of Big Data in Cyberspace: A Survey. , 2019, , .		1
7537	Data Reliability Enhancement Method through Data Validation in Crowdsensing System. , 2019, , .		1
7538	Cost analysis of nondeterministic probabilistic programs. , 2019, , .		24

#	ARTICLE	IF	CITATIONS
7539	Positive and negative association rule mining in Hadoop's MapReduce environment. Journal of Big Data, 2019, 6, .	6.9	17
7540	An Integrated Big and Fast Data Analytics Platform for Smart Urban Transportation Management. IEEE Access, 2019, 7, 117652-117677.	2.6	42
7541	Energy- and locality-efficient multi-job scheduling based on MapReduce for heterogeneous datacenter. Service Oriented Computing and Applications, 2019, 13, 297-308.	1.3	8
7543	An Outlook to Declarative Languages for Big Streaming Data. , 2019, , .		4
7544	Finding Frequent Routes From Taxi Trips With Time Windows. , 2019, , .		0
7545	Parallel Framework for Data-Intensive Computing with XSEDE. , 2019, , .		1
7546	PerTract: Model Extraction and Specification of Big Data Systems for Performance Prediction by the Example of Apache Spark and Hadoop. Big Data and Cognitive Computing, 2019, 3, 47.	2.9	7
7547	Impact of class distribution on the detection of slow HTTP DoS attacks using Big Data. Journal of Big Data, 2019, 6, .	6.9	23
7548	QAECN: Dynamically Tuning ECN Threshold with Micro-burst in Multi-queue Data Centers. , 2019, , .		2
7549	Influence of Tasks Duration Variability on Task-Based Runtime Schedulers. , 2019, , .		4
7550	Minimizing Transmission and Processing Delay in a NFV-based Network. , 2019, , .		0
7551	Efficient Distributed k-clique Mining for Large Networks Using MapReduce. IEEE Transactions on Knowledge and Data Engineering, 2019, , 1-1.	4.0	0
7552	Research on Web Service Selection Based on Parallel Skyline Algorithm. , 2019, , .		0
7553	A Flexible Distributed Optimization Framework for Service of Concurrent Tasks in Processing Networks. , 2019, , .		1
7554	Large-Scale Training Framework for Video Annotation. , 2019, , .		1
7555	Granular Computing. , 2019, , .		6
7556	Massively Parallel Computation via Remote Memory Access. , 2019, , .		11
7557	White-box testing of big data analytics with complex user-defined functions. , 2019, , .		16

#	ARTICLE	IF	CITATIONS
7558	An advanced comparison on big data world computing frameworks. Journal of Physics: Conference Series, 2019, 1228, 012003.	0.3	0
7559	TTLcC: Taming Tail Latency for Erasure-Coded Cloud Storage Systems. IEEE Transactions on Network and Service Management, 2019, 16, 1609-1623.	3.2	12
7560	Two Simple Tricks for Fast Cache-Aware Parallel Particle Swarm Optimization. , 2019, , .		2
7561	MiniForest: Distributed and Dynamic Multicasting in Datacenter Networks. IEEE Transactions on Network and Service Management, 2019, 16, 1268-1281.	3.2	5
7562	A Novel Large-scale Chinese Encyclopedia Knowledge Parallel Refining Method Based on MapReduce. IEEE Access, 2019, 7, 111840-111857.	2.6	1
7563	An Integrated Fast Data Transmission Scheme Based on Network Coding. IEEE Access, 2019, 7, 112216-112228.	2.6	5
7564	Leveraging InfiniBand for Highly Concurrent Messaging in Java Applications. , 2019, , .		2
7565	Scalable Motor Movement Recognition from Electroencephalography using Machine Learning. , 2019, , .		1
7566	When Network Matters: Data Center Scheduling with Network Tasks. , 2019, , .		10
7567	GCache: Neighborhood-Guided Graph Caching in a Distributed Environment. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 2463-2477.	4.0	5
7568	Software Deployment on Heterogeneous Platforms: A Systematic Mapping Study. IEEE Transactions on Software Engineering, 2021, 47, 1683-1707.	4.3	5
7569	SmartCube: An Adaptive Data Management Architecture for the Real-Time Visualization of Spatiotemporal Datasets. IEEE Transactions on Visualization and Computer Graphics, 2019, 26, 1-1.	2.9	9
7570	A Performance Driven Micro Services-Based Architecture/System for Analyzing Noisy IoT Data. , 2019, , .		3
7571	Emergent Scheduling of Distributed Execution Frameworks. , 2019, , .		0
7572	Distributed High-Dimension Matrix Operation Optimization on Spark. , 2019, , .		0
7573	Joint Online Coflow Routing and Scheduling in Data Center Networks. IEEE/ACM Transactions on Networking, 2019, 27, 1771-1786.	2.6	31
7574	Hierarchical Request-Size-Aware Flash Translation Layer Based on Page-Level Mapping. Journal of Circuits, Systems and Computers, 2019, 28, 1950117.	1.0	3
7575	Optimized Speculative Execution Strategy for Different Workload Levels in Heterogeneous Spark Cluster. , 2019, , .		3

#	ARTICLE	IF	CITATIONS
7576	Terabyte-scale Particle Data Analysis. , 2019, , .		4
7577	Speculative Scheduling for Stochastic HPC Applications. , 2019, , .		6
7578	Shared data science infrastructure for genomics data. BMC Bioinformatics, 2019, 20, 436.	1.2	3
7579	Towards Cloud Distributed Image Indexing by Sparse Hashing. , 2019, , .		1
7580	Holoentropy based Correlative Naive Bayes classifier and MapReduce model for classifying the big data. Evolutionary Intelligence, 2022, 15, 1037-1050.	2.3	4
7581	Overcome Heterogeneity Impact in Modeled Fork-Join Queuing Networks for Tail Prediction. , 2019, , .		1
7582	Hadoop MapReduce and Dynamic Intelligent Splitter for Efficient and Speed transmission of Cloud-based video transforming. , 2019, , .		7
7583	Lamda-Flow: Automatic Pushdown of Dataflow Operators Close to the Data. , 2019, , .		4
7584	Towards a Geospatial Big Data Platform for Geospatial Information Services. , 2019, , .		0
7585	Fog Computing for Distributed Family Learning in Cyber-Manufacturing Modeling. , 2019, , .		9
7586	A Topology-aware Coding Framework for Distributed Graph Processing. , 2019, , .		2
7587	Safe Automated Refactoring for Intelligent Parallelization of Java 8 Streams. , 2019, , .		16
7588	Type Migration in Ultra-Large-Scale Codebases. , 2019, , .		13
7589	Towards Scalable Blockchain Analysis. , 2019, , .		11
7590	Demystifying Parallel and Distributed Deep Learning. ACM Computing Surveys, 2020, 52, 1-43.	16.1	292
7591	GPU-Accelerated Similarity Self-Join for Multi-Dimensional Data. , 2019, , .		4
7592	A Free Stale Synchronous Parallel Strategy for Distributed Machine Learning. , 2019, , .		3
7593	CalmAn an open source tool for scalable calcium imaging data analysis. ELife, 2019, 8, .	2.8	551

#	ARTICLE	IF	CITATIONS
7594	A Dynamic Scheduling Framework for Multi-Tenancy Clouds. , 2019, , .		2
7595	Cuckoo: Opportunistic MapReduce on Ephemeral and Heterogeneous Cloud Resources. , 2019, , .		7
7596	Efficient similarity computation for collaborative filtering in dynamic environments. , 2019, , .		2
7597	Big Data Analytics for Large-scale Wireless Networks. ACM Computing Surveys, 2020, 52, 1-36.	16.1	87
7598	Distributed Centrality Analysis of Social Network Data Using MapReduce. Algorithms, 2019, 12, 161.	1.2	29
7599	Big data processing framework for manufacturing. Procedia CIRP, 2019, 83, 661-664.	1.0	7
7600	IoT Data Processing in the Fog: Functions, Streams, or Batch Processing?. , 2019, , .		34
7601	Hybrid CPU/GPU clustering in shared memory on the billion point scale. , 2019, , .		6
7602	Pyne. , 2019, , .		0
7603	A New Framework for Evaluating Straggler Detection Mechanisms in MapReduce. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2019, 4, 1-23.	0.8	10
7604	A Learning-Based Data Placement Framework for Low Latency in Data Center Networks. IEEE Transactions on Cloud Computing, 2022, 10, 146-157.	3.1	23
7605	Orchestrating Big Data Analysis Workflows in the Cloud. ACM Computing Surveys, 2020, 52, 1-41.	16.1	35
7606	A Comprehensive Survey on Cloud Data Mining (CDM) Frameworks and Algorithms. ACM Computing Surveys, 2020, 52, 1-62.	16.1	20
7607	A review on big data based parallel and distributed approaches of pattern mining. Journal of King Saud University - Computer and Information Sciences, 2022, 34, 1639-1662.	2.7	26
7608	Crystal: A scalable and fault-tolerant Archimedean-based server-centric cloud data center network architecture. Computer Communications, 2019, 147, 159-179.	3.1	14
7609	Scalable Action Mining for Recommendations to Reduce Hospital Readmission. , 2019, , .		0
7610	Efficient distributed reachability querying of massive temporal graphs. VLDB Journal, 2019, 28, 871-896.	2.7	15
7611	Parallel computing-based online geometry triangulation for building information modeling utilizing big data. Automation in Construction, 2019, 107, 102942.	4.8	15

#	ARTICLE	IF	CITATIONS
7612	On the performance and convergence of distributed stream processing via approximate fault tolerance. VLDB Journal, 2019, 28, 821-846.	2.7	6
7613	Building secure protocols for extensible distributed coordination through secure extensions. Computers and Security, 2019, 87, 101583.	4.0	2
7614	Analysing billion-objects catalogue interactively: Apache Spark for physicists. Astronomy and Computing, 2019, 28, 100305.	0.8	7
7615	Models of everywhere revisited: A technological perspective. Environmental Modelling and Software, 2019, 122, 104521.	1.9	31
7616	Research paper classification systems based on TF-IDF and LDA schemes. Human-centric Computing and Information Sciences, 2019, 9, .	6.1	137
7617	An Efficient Block Assignment Policy in Hadoop Distributed File System for Multimedia Data Processing. IEICE Transactions on Information and Systems, 2019, E102.D, 1569-1571.	0.4	0
7618	On Managing Spatial Hypermedia with Document Stores. , 2019, , .		1
7619	Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges. Internet of Things (Netherlands), 2019, 8, 100118.	4.9	242
7620	Understanding Big Data Through a Systematic Literature Review: The ITMI Model. International Journal of Information Technology and Decision Making, 2019, 18, 1433-1461.	2.3	20
7621	Parallelism and partitioning in large-scale GAs using spark. , 2019, , .		3
7622	Task-Aware TCP in Data Center Networks. IEEE/ACM Transactions on Networking, 2019, 27, 389-404.	2.6	20
7623	Distributionally Robust Radio Frequency Localization. IEEE Transactions on Signal and Information Processing Over Networks, 2019, 5, 390-403.	1.6	0
7624	A hybrid multi-objective firefly and simulated annealing based algorithm for big data classification. Concurrency Computation Practice and Experience, 2019, 31, e4985.	1.4	14
7625	Predictive performance modeling for distributed batch processing using black box monitoring and machine learning. Information Systems, 2019, 82, 33-52.	2.4	43
7626	Predictive Analytics of Hyper-Connected Collaborative Network. International Journal of Business Data Communications and Networking, 2019, 15, 17-33.	1.2	1
7627	Receiver-driven fair congestion control for TCP outcast in data center networks. Journal of Network and Computer Applications, 2019, 131, 75-88.	5.8	20
7628	Architectural Tactics for Big Data Cybersecurity Analytics Systems: A Review. Journal of Systems and Software, 2019, 151, 81-118.	3.3	53
7629	A Roadmap for the Development of the "SP Machine"™ for Artificial Intelligence. Computer Journal, 2019, 62, 1584-1604.	1.5	5

#	ARTICLE	IF	CITATIONS
7630	Machine Learning and Integrative Analysis of Biomedical Big Data. <i>Genes</i> , 2019, 10, 87.	1.0	208
7631	Parallel Computing of Support Vector Machines. <i>ACM Computing Surveys</i> , 2019, 51, 1-38.	16.1	25
7632	Trajectory big data processing based on frequent activity. <i>Tsinghua Science and Technology</i> , 2019, 24, 317-332.	4.1	12
7633	MapReduce Data Skewness Handling: A Systematic Literature Review. <i>International Journal of Parallel Programming</i> , 2019, 47, 907-950.	1.1	10
7634	Extraction of MapReduce-based features from spectrograms for audio-based surveillance. , 2019, 87, 1-9.		12
7635	Scalable computational geometry in MapReduce. <i>VLDB Journal</i> , 2019, 28, 523-548.	2.7	23
7636	A Big Data platform for smart meter data analytics. <i>Computers in Industry</i> , 2019, 105, 250-259.	5.7	64
7637	Towards Integrations of Big Data Technology Components. <i>Lecture Notes in Business Information Processing</i> , 2019, , 114-120.	0.8	0
7638	PPR-partitioning: a distributed graph partitioning algorithm based on the personalized PageRank vectors in vertex-centric systems. <i>Knowledge and Information Systems</i> , 2019, 61, 847-871.	2.1	0
7639	OMUX: Optical Multicast and Unicast-capable interconnection network for data centers. <i>Optical Switching and Networking</i> , 2019, 33, 1-12.	1.2	5
7640	Energy-Aware Coflow and Antenna Scheduling for Hybrid Server-Centric Data Center Networks. <i>IEEE Transactions on Green Communications and Networking</i> , 2019, 3, 356-365.	3.5	3
7641	A Grammar-Guided Genetic Programming Algorithm for Associative Classification in Big Data. <i>Cognitive Computation</i> , 2019, 11, 331-346.	3.6	18
7642	Applications of Artificial Intelligence in Transport: An Overview. <i>Sustainability</i> , 2019, 11, 189.	1.6	299
7643	A Distributed Spatiotemporal Contingency Analysis for the Lebanese Power Grid. <i>IEEE Transactions on Computational Social Systems</i> , 2019, 6, 162-175.	3.2	1
7644	MISFP-Growth: Hadoop-Based Frequent Pattern Mining with Multiple Item Support. <i>Applied Sciences (Switzerland)</i> , 2019, 9, 2075.	1.3	11
7645	A Distributed Algorithm for Scalable Fuzzy Time Series. <i>Lecture Notes in Computer Science</i> , 2019, , 42-56.	1.0	4
7646	Framework for parallelisation on big data. <i>PLoS ONE</i> , 2019, 14, e0214044.	1.1	5
7647	Similarity joins for high-dimensional data using Spark. <i>Concurrency Computation Practice and Experience</i> , 2019, 31, e5339.	1.4	7

#	ARTICLE	IF	CITATIONS
7649	Application of a Smart City Model to a Traditional University Campus with a Big Data Architecture: A Sustainable Smart Campus. Sustainability, 2019, 11, 2857.	1.6	98
7650	Implementation of a Large-Scale Platform for Cyber-Physical System Real-Time Monitoring. IEEE Access, 2019, 7, 52455-52466.	2.6	39
7651	Genotype Imputation in Genome-Wide Association Studies. Current Protocols in Human Genetics, 2019, 102, e84.	3.5	22
7653	Quality Assessment of Peer-Produced Content in Knowledge Repositories using Development and Coordination Activities. Journal of Management Information Systems, 2019, 36, 478-512.	2.1	20
7654	Towards the application of recommender systems to secure coding. Eurasip Journal on Information Security, 2019, 2019, .	2.4	11
7655	Efficient Parallel Associative Classification Based on Rules Memoization. Lecture Notes in Computer Science, 2019, , 31-44.	1.0	2
7656	Towards Organization Management Using Exploratory Screening and Big Data Tests: A Case Study of the Spanish Red Cross. IEEE Access, 2019, 7, 80661-80674.	2.6	6
7657	An Incremental Approach to Address Big Data Classification Problems Using Cognitive Models. Cognitive Computation, 2019, 11, 347-366.	3.6	9
7658	Coded Computation Over Heterogeneous Clusters. IEEE Transactions on Information Theory, 2019, 65, 4227-4242.	1.5	127
7659	Accelerating Actor-Based Applications with Parallel Patterns. , 2019, , .		4
7660	An effective distributed predictive model with Matrix factorization and random forest for Big Data recommendation systems. Expert Systems With Applications, 2019, 137, 253-265.	4.4	23
7661	Assisted pattern mining for discovering interactive behaviours on the web. International Journal of Human Computer Studies, 2019, 130, 196-208.	3.7	8
7662	Addressing Skewness in Iterative ML Jobs with Parameter Partition. , 2019, , .		6
7663	Spark-Based Parallel Genetic Algorithm for Simulating a Solution of Optimal Deployment of an Underwater Sensor Network. Sensors, 2019, 19, 2717.	2.1	8
7664	Heterogeneity-Aware Data Placement in Hybrid Clouds. Lecture Notes in Computer Science, 2019, , 177-191.	1.0	0
7665	A MapReduce-Based Association Rule Mining Using Hadoop Cluster—An Application of Disease Analysis. Lecture Notes in Networks and Systems, 2019, , 533-541.	0.5	5
7666	DBEst. , 2019, , .		32
7667	Data-driven parallelizable traffic incident detection using spatio-temporally denoised robust thresholds. Transportation Research Part C: Emerging Technologies, 2019, 105, 81-99.	3.9	22

#	ARTICLE	IF	CITATIONS
7668	GeoBeam: A distributed computing framework for spatial data. Computers and Geosciences, 2019, 131, 15-22.	2.0	13
7669	Semantic preservation of standardized healthcare documents in big data. International Journal of Medical Informatics, 2019, 129, 133-145.	1.6	7
7670	Privacy preserving service selection using fully homomorphic encryption scheme on untrusted cloud service platform. Knowledge-Based Systems, 2019, 180, 104-115.	4.0	25
7671	MapReduce: an infrastructure review and research insights. Journal of Supercomputing, 2019, 75, 6934-7002.	2.4	18
7672	Toward Privacy-Aware Healthcare Data Fusion Systems. Communications in Computer and Information Science, 2019, , 26-37.	0.4	1
7673	ARFF Data Source Library for Distributed Single/Multiple Instance, Single/Multiple Output Learning on Apache Spark. Lecture Notes in Computer Science, 2019, , 173-179.	1.0	0
7674	A Distributed PCM Clustering Algorithm Based on Spark. , 2019, , .		1
7675	Map-Side Join Processing of SPARQL Queries Based on Abstract RDF Data Filtering. Journal of Database Management, 2019, 30, 22-40.	1.0	5
7676	Challenges and Future Trends for Microarray Analysis. Methods in Molecular Biology, 2019, 1986, 283-293.	0.4	14
7677	Efficient Approaches for Solving the Large-Scale k-Medoids Problem: Towards Structured Data. Studies in Computational Intelligence, 2019, , 199-219.	0.7	13
7678	A text classification framework for simple and effective early depression detection over social media streams. Expert Systems With Applications, 2019, 133, 182-197.	4.4	106
7679	Reconstruction of Grains in Polycrystalline Materials From Incomplete Data Using Laguerre Tessellations. Microscopy and Microanalysis, 2019, 25, 743-752.	0.2	10
7680	Auto-scaling for real-time stream analytics on HPC cloud. Service Oriented Computing and Applications, 2019, 13, 169-183.	1.3	4
7681	A Cyberinfrastructure for Big Data Transportation Engineering. Journal of Big Data Analytics in Transportation, 2019, 1, 83-94.	1.4	8
7682	Join Algorithms under Apache Spark. , 2019, , .		2
7683	Big data analytics for video surveillance. Multimedia Tools and Applications, 2019, 78, 26129-26162.	2.6	37
7684	Applying big data and stream processing to the real estate domain. Behaviour and Information Technology, 2019, 38, 950-958.	2.5	6
7685	Data Prefetching for Heterogeneous Hadoop Cluster. , 2019, , .		1

#	ARTICLE	IF	CITATIONS
7686	Research and Practice of Big Data Analysis Process Based on Hadoop Framework. , 2019, , .		2
7687	Community Detection. , 2019, , 275-314.		0
7688	Distributed Framework for Automating Opinion Discretization From Text Corpora on Facebook. IEEE Access, 2019, 7, 78675-78684.	2.6	7
7689	A distributed hypervisor for resource aggregation. , 2019, , .		0
7690	Intrusion Detection Systems: A Cross-Domain Overview. IEEE Communications Surveys and Tutorials, 2019, 21, 3639-3681.	24.8	61
7691	Towards large-scale face-based race classification on spark framework. Multimedia Tools and Applications, 2019, 78, 26729-26746.	2.6	6
7692	Analytical composite performance models for Big Data applications. Journal of Network and Computer Applications, 2019, 142, 63-75.	5.8	13
7693	Scalable Frequent Sequence Mining with Flexible Subsequence Constraints. , 2019, , .		2
7694	Distributed Bottleneck-Aware Coflow Scheduling in Data Centers. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1565-1579.	4.0	8
7695	A review of scientific research in defensive cyberspace operation tools and technologies. Journal of Cyber Security Technology, 2019, 3, 1-46.	1.8	11
7696	A Survey on Parallel Join Algorithms Using MapReduce on Hadoop. , 2019, , .		1
7697	Enumerating permutations sortable by k passes through a pop-stack. Advances in Applied Mathematics, 2019, 108, 79-96.	0.4	14
7698	A Review of Botnet Detection Approaches Based on DNS Traffic Analysis. Lecture Notes in Networks and Systems, 2019, , 305-321.	0.5	3
7699	Scalable I-Diversity. International Journal of Information Technology and Web Engineering, 2019, 14, 27-40.	1.2	3
7700	Userland CO-PAGER. , 2019, , .		2
7701	Efficient Straggler Replication in Large-Scale Parallel Computing. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2019, 4, 1-23.	0.8	26
7702	Multistructure-Based Collaborative Online Distillation. Entropy, 2019, 21, 357.	1.1	9
7703	Entropic Statistical Description of Big Data Quality in Hotel Customer Relationship Management. Entropy, 2019, 21, 419.	1.1	10

#	ARTICLE	IF	CITATIONS
7704	Efficient Dynamic Double Threshold Energy Detection of Cooperative Spectrum Sensing in Cognitive Radio. <i>Advances in Intelligent Systems and Computing</i> , 2019, , 479-492.	0.5	2
7705	A Bibliometrics analysis on big data research (2009â€“2018). <i>Journal of Data Information and Management</i> , 2019, 1, 3-15.	1.6	25
7706	SPARK-MSNA: Efficient algorithm on Apache Spark for aligning multiple similar DNA/RNA sequences with supervised learning. <i>Scientific Reports</i> , 2019, 9, 6631.	1.6	7
7707	Performance Evaluation of Tree Ensemble Classification Models Towards Challenges of Big Data Analytics. <i>Communications in Computer and Information Science</i> , 2019, , 141-154.	0.4	0
7708	Multi-agent Healthcare Information System on Hadoop. <i>IFMBE Proceedings</i> , 2019, , 185-193.	0.2	0
7709	k-DDD Measure and MapReduce Based Anonymity Model for Secured Privacy-Preserving Big Data Publishing. <i>International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems</i> , 2019, 27, 177-199.	0.9	14
7710	Data Categorization Using Hadoop MapReduce-Based Parallel K-Means Clustering. <i>Journal of the Institution of Engineers (India): Series B</i> , 2019, 100, 95-103.	1.3	13
7711	Scheduling dependent coflows to minimize the total weighted job completion time in datacenters. <i>Computer Networks</i> , 2019, 158, 193-205.	3.2	5
7712	Random Sample Partition: A Distributed Data Model for Big Data Analysis. <i>IEEE Transactions on Industrial Informatics</i> , 2019, 15, 5846-5854.	7.2	66
7713	IoT based pollution monitoring and health correlation: a case study on smart city. <i>International Journal of Systems Assurance Engineering and Management</i> , 2019, 10, 731-738.	1.5	5
7714	A Framework for Preprocessing Multivariate, Topology-Aware Time Series and Event Data in a Multi-System Environment. , 2019, , .		1
7715	HBSD. <i>International Journal of Information Technology and Web Engineering</i> , 2019, 14, 37-63.	1.2	2
7716	Skyline and reverse skyline query processing in SpatialHadoop. <i>Data and Knowledge Engineering</i> , 2019, 122, 55-80.	2.1	9
7717	Fuzzy String Matching Algorithm for Spam Detection in Twitter. <i>Communications in Computer and Information Science</i> , 2019, , 289-301.	0.4	13
7718	Data Shuffling in Wireless Distributed Computing via Low-Rank Optimization. <i>IEEE Transactions on Signal Processing</i> , 2019, 67, 3087-3099.	3.2	26
7719	Parametric Gaussian process regression for big data. <i>Computational Mechanics</i> , 2019, 64, 409-416.	2.2	22
7720	A write-friendly approach to manage namespace of Hadoop distributed file system by utilizing nonvolatile memory. <i>Journal of Supercomputing</i> , 2019, 75, 6632-6662.	2.4	6
7721	Dealing With Data Heterogeneity in a Data Fusion Perspective. <i>Data Handling in Science and Technology</i> , 2019, 31, 235-270.	3.1	12

#	ARTICLE	IF	CITATIONS
7722	A MapReduce solution for incremental mining of sequential patterns from big data. Expert Systems With Applications, 2019, 133, 109-125.	4.4	24
7723	Data-Intensive Workflow Management: For Clouds and Data-Intensive and Scalable Computing Environments. Synthesis Lectures on Data Management, 2019, 14, 1-179.	0.6	8
7724	Chemometrics. International Journal of Fog Computing, 2019, 2, 1-42.	1.8	1
7725	Cloud computing for storing and analyzing petabytes of genomic data. Journal of Industrial Information Integration, 2019, 15, 50-57.	4.3	18
7726	EPPADS: An Enhanced Phase-Based Performance-Aware Dynamic Scheduler for High Job Execution Performance in Large Scale Clusters. Lecture Notes in Computer Science, 2019, , 140-156.	1.0	2
7727	Performance Analysis of Scheduling Algorithms in Apache Hadoop. , 2019, , 45-57.		1
7728	Cost-efficient disaster backup for multiple data centers using capacity-constrained multicast. Concurrency Computation Practice and Experience, 2019, 31, e5266.	1.4	5
7729	Partially collapsed Gibbs sampling for latent Dirichlet allocation. Expert Systems With Applications, 2019, 131, 208-218.	4.4	15
7730	Towards a more efficient and cost-sensitive extreme learning machine: A state-of-the-art review of recent trend. Neurocomputing, 2019, 350, 70-90.	3.5	44
7731	Analyzing big datasets of genomic sequences: fast and scalable collection of k-mer statistics. BMC Bioinformatics, 2019, 20, 138.	1.2	17
7732	A survey of issues and solutions of health data management systems. Innovations in Systems and Software Engineering, 2019, 15, 155-166.	1.6	2
7733	Privacy preserving similarity joins using MapReduce. Information Sciences, 2019, 493, 20-33.	4.0	19
7734	Extending smart containers for data locality-aware skeleton programming. Concurrency Computation Practice and Experience, 2019, 31, e5003.	1.4	7
7735	Learning automata-based algorithms for MapReduce data skewness handling. Journal of Supercomputing, 2019, 75, 6488-6516.	2.4	5
7736	A Novel Approach of Fair Scheduling to Enhance Performance of Hadoop Distributed File System. , 2019, , .		1
7737	Cloud-based map alignment strategies for multi-robot FastSLAM 2.0. International Journal of Distributed Sensor Networks, 2019, 15, 155014771982932.	1.3	1
7738	SLA-Based Adaptation Schemes in Distributed Stream Processing Engines. Applied Sciences (Switzerland), 2019, 9, 1045.	1.3	3
7739	Combustion Optimization for Coal Fired Power Plant Boilers Based on Improved Distributed ELM and Distributed PSO. Energies, 2019, 12, 1036.	1.6	17

#	ARTICLE	IF	CITATIONS
7740	Survey of Breast Cancer Detection Using Machine Learning Techniques in Big Data. Journal of Cases on Information Technology, 2019, 21, 80-92.	0.7	4
7741	Literature review on Big Data and Its Application Fields. Journal of Physics: Conference Series, 2019, 1176, 042043.	0.3	2
7742	A hierarchic task-based programming model for distributed heterogeneous computing. International Journal of High Performance Computing Applications, 2019, 33, 987-997.	2.4	3
7743	Ensemble RBM-based classifier using fuzzy integral for big data classification. International Journal of Machine Learning and Cybernetics, 2019, 10, 3327-3337.	2.3	9
7744	A self-managed Mesos cluster for data analytics with QoS guarantees. Future Generation Computer Systems, 2019, 96, 449-461.	4.9	10
7745	Survey of agent-based cloud computing applications. Future Generation Computer Systems, 2019, 100, 223-236.	4.9	45
7746	Scalable RDMA RPC on Reliable Connection with Efficient Resource Sharing. , 2019, , .		49
7747	GANA-VDC: Application-Aware with Bandwidth Guarantee in Cloud Datacenters. Electronics (Switzerland), 2019, 8, 258.	1.8	1
7748	A parallel tree node splitting criterion for fuzzy decision trees. Concurrency Computation Practice and Experience, 2019, 31, e5268.	1.4	4
7749	MapReduce based integration of health hubs: a healthcare design approach. Health and Technology, 2019, 9, 737-750.	2.1	0
7750	A Compiler for Agnostic Programming and Deployment of Big Data Analytics on Multiple Platforms. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1920-1931.	4.0	6
7751	Achieving efficient source camera identification on Hadoop. Multimedia Tools and Applications, 2019, 78, 32999-33021.	2.6	4
7752	Efficient processing of all-k-nearest-neighbor queries in the MapReduce programming framework. Data and Knowledge Engineering, 2019, 121, 42-70.	2.1	14
7753	The role of big data analytics in industrial Internet of Things. Future Generation Computer Systems, 2019, 99, 247-259.	4.9	234
7754	LightStore. , 2019, , .		12
7755	Shared Nearest-Neighbor Quantum Game-Based Attribute Reduction With Hierarchical Coevolutionary Spark and Its Application in Consistent Segmentation of Neonatal Cerebral Cortical Surfaces. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30, 2013-2027.	7.2	39
7756	Substream-Centric Maximum Matchings on FPGA. , 2019, , .		10
7757	Distributing Frankâ€“Wolfe via map-reduce. Knowledge and Information Systems, 2019, 60, 665-690.	2.1	0

#	ARTICLE	IF	CITATIONS
7758	Efficient Scheduling of Weighted Coflows in Data Centers. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 2003-2017.	4.0	23
7759	Semantic-Aware Data Cube for Cloud Networks. , 2019, , 179-204.		1
7761	Data Partitioning Scheme for Efficient Distributed RDF Querying Using Apache Spark. , 2019, , .		9
7762	Semi-Clairvoyant Scheduling in Data Analytics Systems. IEEE Transactions on Computers, 2019, 68, 1376-1389.	2.4	1
7763	Efficient computation of key performance indicators in a distance learning university. Information Discovery and Delivery, 2019, 47, 96-105.	1.6	6
7764	MR-Edge: a MapReduce-based Protocol for IoT Edge Computing with Resource Constraints. , 2019, , .		6
7765	A Novel Parallel Biclustering Approach and Its Application to Identify and Segment Highly Profitable Telecom Customers. IEEE Access, 2019, 7, 28696-28711.	2.6	8
7766	A framework and a performance assessment for serverless MapReduce on AWS Lambda. Future Generation Computer Systems, 2019, 97, 259-274.	4.9	49
7767	Data Management in Machine Learning Systems. Synthesis Lectures on Data Management, 2019, 14, 1-173.	0.6	16
7768	An Important and Timely Field. , 2019, , 1-8.		6
7769	The History of Computing Education Research. , 2019, , 11-39.		26
7770	Computing Education Research Today. , 2019, , 40-55.		5
7771	Computing Education Literature Review and Voices from the Field. , 2019, , 56-78.		10
7772	A Study Design Process. , 2019, , 81-101.		1
7774	Inferential Statistics. , 2019, , 133-172.		2
7775	Qualitative Methods for Computing Education. , 2019, , 173-207.		9
7776	Learning Sciences for Computing Education. , 2019, , 208-230.		17
7777	Higher Education Pedagogy. , 2019, , 276-291.		4

#	ARTICLE	IF	CITATIONS
7778	Engineering Education Research. , 2019, , 292-322.		4
7779	Novice Programmers and Introductory Programming. , 2019, , 327-376.		60
7780	Programming Paradigms and Beyond. , 2019, , 377-413.		31
7781	Assessment and Plagiarism. , 2019, , 414-444.		6
7782	Pedagogic Approaches. , 2019, , 445-480.		13
7783	Equity and Diversity. , 2019, , 481-510.		10
7784	Computational Thinking. , 2019, , 513-546.		24
7785	Schools (Kâ€“12). , 2019, , 547-583.		5
7786	Computing for Other Disciplines. , 2019, , 584-605.		4
7787	New Programming Paradigms. , 2019, , 606-636.		1
7788	Tools and Environments. , 2019, , 639-662.		11
7789	Tangible Computing. , 2019, , 663-678.		35
7790	Leveraging the Integrated Development Environment for Learning Analytics. , 2019, , 679-706.		7
7791	Teacher Learning and Professional Development. , 2019, , 727-748.		1
7792	Learning Outside the Classroom. , 2019, , 749-772.		6
7793	Student Knowledge and Misconceptions. , 2019, , 773-800.		1
7794	Students As Teachers and Communicators. , 2019, , 827-858.		5
7795	A Case Study of Peer Instruction. , 2019, , 861-874.		3

#	ARTICLE	IF	CITATIONS
7796	A Case Study of Qualitative Methods. , 2019, , 875-894.		0
7798	A 1024-Port Optical Uni- and Multicast Packet Switch Fabric. Journal of Lightwave Technology, 2019, 37, 1415-1423.	2.7	24
7799	Evaluating Distributed Computing Infrastructures: An Empirical Study Comparing Hadoop Deployments on Cloud and Local Systems. IEEE Transactions on Cloud Computing, 2021, 9, 1075-1088.	3.1	2
7800	A Distributed Support Vector Machine Using Apache Spark for Semi-supervised Classification with Data Augmentation. Advances in Intelligent Systems and Computing, 2019, , 395-405.	0.5	2
7801	Coflow scheduling algorithm based density peaks clustering. Future Generation Computer Systems, 2019, 97, 805-813.	4.9	3
7802	Scaling entity resolution: A loosely schema-aware approach. Information Systems, 2019, 83, 145-165.	2.4	14
7803	Improving Data-Analytics Performance Via Autonomic Control of Concurrency and Resource Units. ACM Transactions on Autonomous and Adaptive Systems, 2018, 13, 1-25.	0.4	8
7804	ToothPic: Camera-Based Image Retrieval on Large Scales. IEEE MultiMedia, 2019, 26, 33-43.	1.5	1
7805	URSA. , 2019, , .		27
7806	SpiNNTools: The Execution Engine for the SpiNNaker Platform. Frontiers in Neuroscience, 2019, 13, 231.	1.4	25
7807	Provisioning and Performance Evaluation of Parallel Systems with Output Synchronization. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2019, 4, 1-31.	0.8	3
7809	Benchmarking Distributed Data Processing Systems for Machine Learning Workloads. Lecture Notes in Computer Science, 2019, , 42-57.	1.0	1
7810	Data Science for Entrepreneurship Research: Studying Demand Dynamics for Entrepreneurial Skills in the Netherlands. SSRN Electronic Journal, 0, , .	0.4	0
7811	MV-kWNN: A novel multivariate and multi-output weighted nearest neighbours algorithm for big data time series forecasting. Neurocomputing, 2019, 353, 56-73.	3.5	34
7812	Venice. ACM Transactions on Computer Systems, 2019, 36, 1-26.	0.6	11
7813	MAPSkew: Metaheuristic Approaches for Partitioning Skew in MapReduce. Algorithms, 2019, 12, 5.	1.2	3
7814	Stochastic Online Learning for Mobile Edge Computing: Learning from Changes. IEEE Communications Magazine, 2019, 57, 63-69.	4.9	82
7815	Coflow Scheduling in the Multi-Resource Environment. IEEE Transactions on Network and Service Management, 2019, 16, 783-796.	3.2	2

#	ARTICLE	IF	CITATIONS
7818	Scalable and optimal planning based on Pregel. Concurrency Computation Practice and Experience, 2019, 31, e4966.	1.4	0
7819	An Overview of Big Data and Machine Learning Paradigms. Advances in Intelligent Systems and Computing, 2019, , 237-251.	0.5	6
7820	Attention-Based Hierarchical Recurrent Neural Network for Phenotype Classification. Lecture Notes in Computer Science, 2019, , 465-476.	1.0	2
7821	Metaheuristics for data mining. 4or, 2019, 17, 115-139.	1.0	10
7822	On-the-fly scheduling versus reservation-based scheduling for unpredictable workflows. International Journal of High Performance Computing Applications, 2019, 33, 1140-1158.	2.4	5
7823	Study on the use of different quality measures within a multi-objective evolutionary algorithm approach for emerging pattern mining in big data environments. Big Data Analytics, 2019, 4, .	2.2	4
7824	Residual Capacity-Aware Virtual Machine Assignment for Reducing Network Loads in Multi-tenant Data Center Networks. Journal of Network and Systems Management, 2019, 27, 949-971.	3.3	2
7825	Efficient Subgraph Matching on Large RDF Graphs Using MapReduce. Data Science and Engineering, 2019, 4, 24-43.	4.6	34
7826	Data-driven monitoring of multimode continuous processes: A review. Chemometrics and Intelligent Laboratory Systems, 2019, 189, 56-71.	1.8	112
7827	KTMiner: Distributed k-truss detection in big graphs. Information Systems, 2019, 83, 195-216.	2.4	3
7828	Processing Unstructured Databases Using a Quantum Approach. Lecture Notes in Intelligent Transportation and Infrastructure, 2019, , 275-285.	0.3	2
7829	Survey of State-of-the-Art Mixed Data Clustering Algorithms. IEEE Access, 2019, 7, 31883-31902.	2.6	138
7830	A Higher-Order Calculus of Computational Fields. ACM Transactions on Computational Logic, 2019, 20, 1-55.	0.7	55
7831	A comprehensive survey of anomaly detection in banking, wireless sensor networks, social networks, and healthcare. Intelligent Decision Technologies, 2019, 13, 229-270.	0.6	16
7832	Tuberculosis detection analysis using texture features on CXRs images. AIP Conference Proceedings, 2019, , .	0.3	3
7833	Towards Efficient and Scalable Data-Intensive Content Delivery: State-of-the-Art, Issues and Challenges. Lecture Notes in Computer Science, 2019, , 88-137.	1.0	5
7834	Distributed parallel deep learning of Hierarchical Extreme Learning Machine for multimode quality prediction with big process data. Engineering Applications of Artificial Intelligence, 2019, 81, 450-465.	4.3	43
7835	An Energy Efficiency Optimization and Control Model for Hadoop Clusters. IEEE Access, 2019, 7, 40534-40549.	2.6	12

#	ARTICLE	IF	CITATIONS
7836	A systematic literature review on MapReduce scheduling methods. <i>Intelligent Decision Technologies</i> , 2019, 13, 1-21.	0.6	1
7837	Why High-Performance Modelling and Simulation for Big Data Applications Matters. <i>Lecture Notes in Computer Science</i> , 2019, , 1-35.	1.0	6
7838	Performance Data Analysis for Parallel Processing Using Bigdata Distribution. <i>Advances in Intelligent Systems and Computing</i> , 2019, , 602-611.	0.5	2
7839	SAIR: significance-aware approach to improve QoR of big data processing in case of budget constraint. <i>Journal of Supercomputing</i> , 2019, 75, 5760-5781.	2.4	12
7840	HPCC based framework for COPD readmission risk analysis. <i>Journal of Big Data</i> , 2019, 6, .	6.9	1
7841	Hybrid Recommendation Approach for Behavior Modification in the Chinese Construction Industry. <i>Journal of Construction Engineering and Management - ASCE</i> , 2019, 145, .	2.0	14
7842	Multi-Level Elasticity for Data Stream Processing. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2019, 30, 2326-2337.	4.0	14
7843	A rewrite-based optimizer for Spark. <i>Future Generation Computer Systems</i> , 2019, 98, 586-599.	4.9	6
7844	Parallel Markov chain Monte Carlo for Bayesian hierarchical models with big data, in two stages. <i>Journal of Applied Statistics</i> , 2019, 46, 1917-1936.	0.6	2
7845	A Hybrid MPI/OpenMP Parallelization of \$K\$-Means Algorithms Accelerated Using the Triangle Inequality. <i>IEEE Access</i> , 2019, 7, 42280-42297.	2.6	22
7846	Radio: Accelerating Disk-Based Graph Processing by Reducing Disk I/Os. <i>IEEE Transactions on Computers</i> , 2019, 68, 414-425.	2.4	2
7847	Data mining tools. <i>Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery</i> , 2019, 9, e1309.	4.6	16
7848	Big Data Analysis in UAV Surveillance for Wildfire Prevention and Management. <i>Lecture Notes in Business Information Processing</i> , 2019, , 47-58.	0.8	7
7849	Neural Article Pair Modeling for Wikipedia Sub-article Matching. <i>Lecture Notes in Computer Science</i> , 2019, , 3-19.	1.0	8
7850	iRODS metadata management for a cancer genome analysis workflow. <i>BMC Bioinformatics</i> , 2019, 20, 29.	1.2	3
7851	Decentralized iterative approaches for community clustering in the networks. <i>Journal of Supercomputing</i> , 2019, 75, 4894-4917.	2.4	3
7852	Evaluating associative classification algorithms for Big Data. <i>Big Data Analytics</i> , 2019, 4, .	2.2	11
7853	Scalable Processing of Spatial-Keyword Queries. <i>Synthesis Lectures on Data Management</i> , 2019, 11, 1-116.	0.6	9

#	ARTICLE	IF	CITATIONS
7854	VFTree: A Versatile Network Architecture for Data Centers. IEEE Transactions on Cloud Computing, 2019, , 1-1.	3.1	0
7855	Cognitive Sciences for Computing Education. , 2019, , 231-275.		22
7856	Teacher Knowledge for Inclusive Computing Learning. , 2019, , 709-726.		6
7857	Motivation, Attitudes, and Dispositions. , 2019, , 801-826.		15
7858	An improved integrated Grid and MapReduceâ€Hadoop architecture for spatial data: Hilbert TGS Râ€Treeâ€-based IGSIM. Concurrency Computation Practice and Experience, 2019, 31, e5202.	1.4	4
7859	HOLMeS: eHealth in the Big Data and Deep Learning Era. Information (Switzerland), 2019, 10, 34.	1.7	21
7860	GPU-accelerated Large-Scale Non-negative Matrix Factorization Using Spark. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2019, , 189-201.	0.2	1
7861	Parallel naive Bayes algorithm for large-scale Chinese text classification based on spark. Journal of Central South University, 2019, 26, 1-12.	1.2	23
7862	Urban Computing Leveraging Location-Based Social Network Data. ACM Computing Surveys, 2020, 52, 1-39.	16.1	60
7863	Predicting Unethical Physician Behavior At Scale: A Distributed Computing Framework. , 2019, , .		1
7864	Research on Massive Data Analysis of Power Utilization Based on Improved K-Means Algorithm and Cloud Computing. , 2019, , .		5
7865	Distributed Data Analytics Framework for Cluster Analysis of Parking Violation. , 2019, , .		0
7866	Energy-Efficient Edge-Facilitated Wireless Collaborative Computing using Map-Reduce. , 2019, , .		3
7867	Sparsifying Distributed Algorithms with Ramifications in Massively Parallel Computation and Centralized Local Computation. , 2019, , 1636-1653.		35
7868	Progressive Forwarding Disaster Backup among Cloud Datacenters. IEICE Transactions on Information and Systems, 2019, E102.D, 2135-2147.	0.4	6
7869	Analysis of KNN Algorithm with Mapreduce Technique on Big Data. IOP Conference Series: Materials Science and Engineering, 2019, 590, 012028.	0.3	1
7870	High Performance and Scalable Simulations of a Bio-inspired Computational Model. , 2019, , .		0
7871	BEANS â€ distributed data analysis for numerical simulations. Proceedings of the International Astronomical Union, 2019, 14, 460-463.	0.0	0

#	ARTICLE	IF	CITATIONS
7872	Optimize Neutral Framework With Fair Share Resource Allocator For Big Data Processing On Cloud Infrastructure. , 2019, , .		0
7873	Solar-Wind Energy Assessment by Big Data Analysis. , 0, , .		0
7874	Classification in a Distributed System - A Study of Random Forest in the Hadoop MapReduce Framework. , 2019, , .		0
7875	Big-Data Applications as Self-Adaptive Systems of Systems. , 2019, , .		2
7876	Evolutionary Approach based Scheduler for Speculative Task Execution. , 2019, , .		1
7877	Effective Straggler Mitigation with Cross-Layer Interference-Aware Optimization. , 2019, , .		1
7878	Cascaded Coded Distributed Computing on Heterogeneous Networks. , 2019, , .		22
7879	A Hive and SQL Case Study in Cloud Data Analytics. , 2019, , .		3
7880	An optimization approach to load balancing, scheduling and right sizing of cloud computing systems with data locality. , 2019, , .		5
7881	Cloud Computing Task Scheduling Based on Improved Differential Evolution Algorithm. , 2019, , .		2
7882	HaRD: a heterogeneity-aware replica deletion for HDFS. Journal of Big Data, 2019, 6, 94.	6.9	10
7883	Spatial-Aware Approximate Big Data Stream Processing. , 2019, , .		10
7884	Near-Optimal Massively Parallel Graph Connectivity. , 2019, , .		15
7885	Parallelized Topological Relaxation Algorithm. , 2019, , .		0
7886	Elastic Executor Provisioning for Iterative Workloads on Apache Spark. , 2019, , .		4
7887	Benchmarking and Performance Modelling of MapReduce Communication Pattern. , 2019, , .		2
7888	Nap: Network-Aware Data Partitions for Efficient Distributed Processing. , 2019, , .		1
7889	CAMR: Coded Aggregated MapReduce. , 2019, , .		8

#	ARTICLE	IF	CITATIONS
7890	A Performance Comparison of Dask and Apache Spark for Data-Intensive Neuroimaging Pipelines. , 2019, , .		9
7891	Goldilocks: Adaptive Resource Provisioning in Containerized Data Centers. , 2019, , .		9
7892	Understanding the Performance of In-Network Computing: A Case Study. , 2019, , .		3
7893	SwitchAgg: A Further Step Towards In-Network Computing. , 2019, , .		4
7894	Exponentially Faster Massively Parallel Maximal Matching. , 2019, , .		17
7895	A Cloud Resource Scheduling Framework for big data stream and analytics in Cloud Environment. , 2019, , .		1
7896	A Data-Driven Decision Making Framework for Value-Based Engineering Design of Complex Network Systems. , 2019, , .		0
7897	Distributed empirical risk minimization over directed graphs. , 2019, , .		2
7898	Processing Large Datasets of Fined Grained Source Code Changes. , 2019, , .		1
7899	Efficient Implementation of Dynamic Array SystemUsing MapReduce Framework. , 2019, , .		0
7900	Heterogeneity-aware Gradient Coding for Straggler Tolerance. , 2019, , .		17
7901	Data Encoding Methods for Byzantine-Resilient Distributed Optimization. , 2019, , .		12
7902	CNR:A Technique for Data Replication Organization in BigData. , 2019, , .		2
7903	High Dimensional Data Clustering by means of Distributed Dirichlet Process Mixture Models. , 2019, , .		3
7904	$2(1 - \frac{1}{n})$ -Factor Steiner Tree Approximation in $\tilde{O}(n^{1/3})$ Rounds in the CONGESTED CLIQUE. , 2019, , .		0
7905	OSTB: Optimizing Fairness and Efficiency for Coflow Scheduling without Prior Knowledge. , 2019, , .		1
7906	Parallel Processing of Frequent Itemset Based on MapReduce Programming Model. , 2019, , .		3
7907	Parallel Hybrid BBO Search Method for Twitter Sentiment Analysis of Large Scale Datasets Using MapReduce. International Journal of Information Security and Privacy, 2019, 13, 106-122.	0.6	11

#	ARTICLE	IF	CITATIONS
7908	Resource Scheduling of Workflow Tasks in Cloud Environment. , 2019, , .		4
7909	Comparison of Data Mining Algorithms for Predicting the Cancer Disease Using Python. , 2019, , .		1
7910	Multi-way Divide and Conquer Parallel Programming based on PLists. , 2019, , .		2
7911	Weather Data Analysis Using Hadoop: Applications and Challenges. IOP Conference Series: Materials Science and Engineering, 2019, 551, 012044.	0.3	4
7912	Dynamic Load Balancing Algorithm for Heterogeneous Hadoop Cluster. , 2019, , .		0
7913	Facing the full model selection problem in high volume datasets employing intelligent proxy models. Intelligent Data Analysis, 2019, 23, 1109-1129.	0.4	2
7914	MBECN: Enabling ECN with Micro-burst Traffic in Multi-queue Data Center. , 2019, , .		2
7915	Usable Disk Space Control Based on Hadoop Job Features. , 2019, , .		2
7916	Cloud Task Scheduling Algorithm Based on Three Queues and Dynamic Priority. , 2019, , .		8
7917	Scalable Video Classification using Bag of Visual Words on Spark. , 2019, , .		0
7918	MEER: Online Estimation of Optimal Memory Reservations for Long Lived Containers in In-Memory Cluster Computing. , 2019, , .		5
7919	Towards Adaptive Replication for Hot/Cold Blocks in HDFS using MemCached. , 2019, , .		2
7920	Conditional Hardness Results for Massively Parallel Computation from Distributed Lower Bounds. , 2019, , .		22
7921	Improving Plagiarism Checker Throughput with Apache Storm. , 2019, , .		0
7922	KerTSDroid: Detecting Android Malware at Scale through Kernel Task Structures. , 2019, , .		4
7923	Dynamic Control of CPU Cap Allocations in Stream Processing and Data-Flow Platforms. , 2019, , .		0
7924	Efficient task completion for parallel offloading in vehicular fog computing. China Communications, 2019, 16, 42-55.	2.0	27
7925	dislib: Large Scale High Performance Machine Learning in Python. , 2019, , .		8

#	ARTICLE	IF	CITATIONS
7926	Evaluation of Pilot Jobs for Apache Spark Applications on HPC Clusters. , 2019, , .		1
7927	Active Provenance for Data-Intensive Workflows: Engaging Users and Developers. , 2019, , .		3
7928	Smart Street Lights and Mobile Citizen Apps for Resilient Communication in a Digital City. , 2019, , .		4
7929	Secure Coded Cooperative Computation at the Heterogeneous Edge against Byzantine Attacks. , 2019, , .		7
7930	Co-scheduler: Accelerating Data-Parallel Jobs in Datacenter Networks with Optical Circuit Switching. , 2019, , .		6
7931	DYRS: Bandwidth-Aware Disk-to-Memory Migration of Cold Data in Big-Data File Systems. , 2019, , .		1
7932	i2Graph: An Incremental Iterative Computation Model for Large Scale Dynamic Graphs. , 2019, , .		0
7933	The Argument for a "Data Cube" for Large-Scale Psychometric Data. <i>Frontiers in Education</i> , 2019, 4, .	1.2	7
7934	Parallel Cellular Automata Markov Model for Land Use Change Prediction over MapReduce Framework. <i>ISPRS International Journal of Geo-Information</i> , 2019, 8, 454.	1.4	28
7935	Optimization of Shortest-Path Search on RDBMS-Based Graphs. <i>ISPRS International Journal of Geo-Information</i> , 2019, 8, 550.	1.4	2
7936	An Efficient MapReduce-Based Parallel Processing Framework for User-Based Collaborative Filtering. <i>Symmetry</i> , 2019, 11, 748.	1.1	9
7937	Big Data and Cloud Computing: An Emerging Perspective and Future Trends. , 2019, , .		2
7938	Annealing for Distributed Global Optimization. , 2019, , .		3
7939	Heterogeneous Coded Distributed Computing: Joint Design of File Allocation and Function Assignment. , 2019, , .		18
7940	New Parallel Algorithms for All Pairwise Computation on Large HPC Clusters. , 2019, , .		0
7941	On Realizing Distributed Deep Neural Networks: An Astrophysics Case Study. , 2019, , .		0
7942	The Impact of Bike-Sharing Ridership on Air Quality: A Scalable Data Science Framework. , 2019, , .		0
7943	Data Science Model Curriculum Implementation for Various Types of Big Data Infrastructure Courses. , 2019, , .		2

#	ARTICLE	IF	CITATIONS
7944	Dynamic Sizing of Continuously Divisible Jobs for Heterogeneous Resources. , 2019, , .		1
7945	On Distributed Information Composition in Big Data Systems. , 2019, , .		0
7946	DOD-ETL: distributed on-demand ETL for near real-time business intelligence. Journal of Internet Services and Applications, 2019, 10, .	1.6	20
7947	Empirical Analysis of MapReduce Job Scheduling with respect to Energy Consumption of Clusters. , 2019, , .		1
7948	Workload-Aware Scheduling for Data Analytics upon Heterogeneous Storage. , 2019, , .		1
7949	Mobile Application for Storage and Retrieval of e-learning videos Using Hadoop. , 2019, , .		5
7950	A Scalable Approach for Improving Implementation of a Frequent Pattern Mining Algorithm using MapReduce Programming. , 2019, , .		1
7951	A robust, efficient, and balanced parallel algorithm for finding connected components. , 2019, , .		1
7952	Hyper: Distributed Cloud Processing for Large-Scale Deep Learning Tasks. , 2019, , .		1
7953	HMM Optimized Modeling of SSD Storage for I/O MapReduce Workloads. , 2019, , .		0
7954	Online makespan minimization in MapReduce-like systems with unsplit map tasks. , 2019, , .		0
7955	Node Performance Load Balancing Algorithm for Hadoop Cluster. , 2019, , .		0
7956	Collaborative Learning with Limited Interaction: Tight Bounds for Distributed Exploration in Multi-armed Bandits. , 2019, , .		9
7957	Spark-Based Machine Learning Pipeline Construction Method. , 2019, , .		0
7958	High-Utility Itemset Mining in Big Dataset. , 2019, , .		0
7959	Practical Data Mid-Platform Design and Implementation for Medical Big Data. , 2019, , .		5
7960	Massively Distributed Datacube Processing. , 2019, , .		3
7961	Image Recognition with MapReduce Based Convolutional Neural Networks. , 2019, , .		2

#	ARTICLE	IF	CITATIONS
7962	Music Streaming Service Prediction with MapReduce-based Artificial Neural Network. , 2019, , .		1
7963	RAMP: Real-Time Anomaly Detection in Scientific Workflows. , 2019, , .		2
7964	Two-Stage Framework for Big Spatial Data Analytics to Support Disaster Response. , 2019, , .		4
7965	Mining High-Utility Sequential Patterns from Big Datasets. , 2019, , .		6
7966	An Approach for Scale Suspicious Network Events Detection. , 2019, , .		2
7967	OPAL: High performance platform for large-scale privacy-preserving location data analytics. , 2019, , .		3
7968	Real-Time Vehicular Traffic Analysis using Big Data Processing and IoT based Devices for Future Policy Predictions in Smart Transportation. , 2019, , .		4
7969	Adaptation of Classical Machine Learning Algorithms to Big Data Context: Problems and Challenges : Case Study: Hidden Markov Models Under Spark. , 2019, , .		6
7970	Finding Moving Objects in Astronomical Image Datasets. , 2019, , .		0
7971	Skew-Aware Collective Communication for MapReduce Shuffling. IEICE Transactions on Information and Systems, 2019, E102.D, 2389-2399.	0.4	0
7972	Suffix Tree Construction based Mapreduce. , 2019, , .		0
7973	Improving big-data automotive applications performance through adaptive resource allocation. , 2019, , .		0
7974	NOESIS: A Framework for Complex Network Data Analysis. Complexity, 2019, 2019, 1-14.	0.9	4
7975	In Search of a Fast and Efficient Serverless DAG Engine. , 2019, , .		21
7976	Cost-efficient Task Scheduling for Parallel Applications on Heterogeneous Cloud Environment. , 2019, , .		0
7977	Graph Computing based Parallel Power Flow Algorithm and Graph Visualization for Power Distribution Networks. , 2019, , .		2
7978	Survey: Utilizing excess cloud resources to reduce the cost and time of distributed processing. , 2019, , .		4
7979	Distributed Mining of Frequent Patterns in Big Data by Hybrid Strategies. , 2019, , .		0

#	ARTICLE	IF	CITATIONS
7980	MovCloud: A Cloud-Enabled Framework to Analyse Movement Behaviors. , 2019, , .		6
7981	YouTube Video Classification based on Title and Description Text. , 2019, , .		7
7982	Automated Configuration Parameter Classification Model for Hive Query Plan on the Apache Yarn. , 2019, , .		0
7983	Coded Distributed Tracking. , 2019, , .		1
7984	Coding for Heterogeneous UAV-Based Networked Airborne Computing. , 2019, , .		10
7985	Information Spaces for Big Data Processing: Unification and Parallelization of Sequential Information Accumulation Procedures. , 2019, , .		1
7986	Harmony: An Approach for Geo-distributed Processing of Big-Data Applications. , 2019, , .		4
7987	Taming Latency in Data Centers Via Active Congestion-Probing. , 2019, , .		4
7988	Robust Gradient Descent via Moment Encoding and LDPC Codes. , 2019, , .		28
7989	AZ-Code: An Efficient Availability Zone Level Erasure Code to Provide High Fault Tolerance in Cloud Storage Systems. , 2019, , .		23
7990	PDF-DS: Privacy-Preserving Data Filtering for Distributed Data Streams in Cloud. , 2019, , .		2
7991	Trend-Smooth: Accelerate Asynchronous SGD by Smoothing Parameters Using Parameter Trends. IEEE Access, 2019, 7, 156848-156859.	2.6	6
7992	Distributed, Numerically Stable Distance and Covariance Computation with MPI for Extremely Large Datasets. , 2019, , .		4
7993	Coded Matrix Multiplication on a Group-Based Model. , 2019, , .		19
7994	Field Programmable Gate Array Applications—A Scientometric Review. Computation, 2019, 7, 63.	1.0	46
7995	Toward High-Performance Computing and Big Data Analytics Convergence: The Case of Spark-DIY. IEEE Access, 2019, 7, 156929-156955.	2.6	8
7996	Cost-Minimizing Online Algorithms for Geo-Distributed Data Analytics. IEEE Access, 2019, 7, 163515-163525.	2.6	0
7997	Subdiffusive Load Balancing in Time-Varying Queueing Systems. Operations Research, 2019, 67, 1678-1698.	1.2	4

#	ARTICLE	IF	CITATIONS
7998	Adaptively Accelerating Map-Reduce/Spark with GPUs: A Case Study. , 2019, , .		3
7999	A Scalable (α, k)-Anonymization Approach using MapReduce for Privacy Preserving Big Data Publishing. , 2019, , .		2
8000	Computational Strategies for Scalable Genomics Analysis. Genes, 2019, 10, 1017.	1.0	9
8001	Big Data in Cloud Computing and Internet of Things. , 2019, , .		1
8002	An Asynchronous ADMM Algorithm for Distributed Optimization with Dynamic Scheduling Strategy. , 2019, , .		3
8003	Online Feature Selection (OFS) with Accelerated Bat Algorithm (ABA) and Ensemble Incremental Deep Multiple Layer Perceptron (EIDMLP) for big data streams. Journal of Big Data, 2019, 6, .	6.9	10
8004	A new Internet of Things architecture for real-time prediction of various diseases using machine learning on big data environment. Journal of Big Data, 2019, 6, .	6.9	49
8005	Rateless Codes for Distributed Computations with Sparse Compressed Matrices. , 2019, , .		6
8006	SLO-aware colocation: Harvesting transient resources from latency-critical services. Journal of Systems Architecture, 2019, 101, 101663.	2.5	5
8007	Coded MapReduce Schemes Based on Placement Delivery Array. , 2019, , .		12
8008	Workflow Environments for Advanced Cyberinfrastructure Platforms. , 2019, , .		3
8009	Matrix Profile XIV. , 2019, , .		35
8010	swCPD: Optimizing Canonical Polyadic Decomposition on Sunway Manycore Architecture. , 2019, , .		1
8011	Optimizing Speculative Execution in Spark Heterogeneous Environments. IEEE Transactions on Cloud Computing, 2022, 10, 568-582.	3.1	11
8012	Procella. Proceedings of the VLDB Endowment, 2019, 12, 2022-2034.	2.1	16
8013	Exploring Benefits of NVMe SSDs for BigData Processing in Enterprise Data Centers. , 2019, , .		6
8014	Fuzzy Clustering Based on Cloud and Fog Computing. , 2019, , .		0
8015	On using MapReduce to scale algorithms for Big Data analytics: a case study. Journal of Big Data, 2019, 6, .	6.9	7

#	ARTICLE	IF	CITATIONS
8016	CAP-based Examination of Popular NoSQL Database Technologies in Streaming Data Processing. , 2019, , .		2
8017	Privacy in Multi-party Computation in MapReduce. , 2019, , .		0
8018	Straggler Mitigation at Scale. IEEE/ACM Transactions on Networking, 2019, 27, 2266-2279.	2.6	30
8019	A Holistic Auto-Configurable Ensemble Machine Learning Strategy for Financial Trading. Computation, 2019, 7, 67.	1.0	11
8020	HybSMRP: a hybrid scheduling algorithm in Hadoop MapReduce framework. Journal of Big Data, 2019, 6, .	6.9	22
8021	swTensor: accelerating tensor decomposition on Sunway architecture. CCF Transactions on High Performance Computing, 2019, 1, 161-176.	1.1	2
8022	Cheetah: A Dynamic Performance Optimization Approach on Heterogeneous Big Data Analytics Cluster. , 2019, , .		2
8023	A Big Data Semantic Driven Context Aware Recommendation Method for Question-Answer Items. IEEE Access, 2019, 7, 182664-182678.	2.6	4
8024	BÃ¼yÃ¼k veri Ã¼zerine perspektif bir bakÃ±ÅŸ. , 2019, , .		2
8025	IPSO: A Scaling Model for Data-Intensive Applications. , 2019, , .		1
8026	Coronary Heart Disease Predictive Decision Scheme Using Big Data and RNN. , 2019, , .		6
8027	Supporting Data Stream Analytical Processing in Vehicular Sensor Networks. , 2019, , .		0
8028	SlimML: Removing Non-critical Input Data in Large-scale Iterative Machine Learning. IEEE Transactions on Knowledge and Data Engineering, 2019, , 1-1.	4.0	6
8029	A concept of a hardware design environment with the functional language elixir. , 2019, , .		0
8030	Financial Quantitative Big Data Platform Based on High Performance Computing. , 2019, , .		1
8031	Parallel Processing Strategies for Big Geospatial Data. Frontiers in Big Data, 2019, 2, 44.	1.8	6
8032	Performance Improvement of Parallel Programming Model Based on Parameterized Pipelined Map Reduce Approach. , 2019, , .		3
8033	A Chaos Engineering System for Live Analysis and Falsification of Exception-Handling in the JVM. IEEE Transactions on Software Engineering, 2021, 47, 2534-2548.	4.3	13

#	ARTICLE	IF	CITATIONS
8034	MapReduce Scheduler to Minimize the Size of Intermediate Data in Shuffle Phase. , 2019, , .		2
8035	Distributed Parameter Optimization Scheduling Strategy and System Design Based on Mesos. , 2019, , .		1
8036	KHyperLogLog: Estimating Reidentifiability and Joinability of Large Data at Scale. , 2019, , .		7
8037	Performance evaluation of K-means clustering on Hadoop infrastructure. Journal of Discrete Mathematical Sciences and Cryptography, 2019, 22, 1349-1363.	0.5	24
8038	MaRelA: a cloud MapReduce based high performance whole slide image analysis framework. Distributed and Parallel Databases, 2019, 37, 251-272.	1.0	6
8039	DataSynapse: A Social Data Curation Foundry. Distributed and Parallel Databases, 2019, 37, 351-384.	1.0	33
8040	A programmable shared-memory system for an array of processing-in-memory devices. Cluster Computing, 2019, 22, 385-398.	3.5	0
8041	Load balancing in join algorithms for skewed data in MapReduce systems. Journal of Supercomputing, 2019, 75, 228-254.	2.4	18
8042	The big data system, components, tools, and technologies: a survey. Knowledge and Information Systems, 2019, 60, 1165-1245.	2.1	78
8043	Smart Focused Web Crawler for Hidden Web. Lecture Notes in Networks and Systems, 2019, , 419-427.	0.5	1
8044	A scalable spatial skyline evaluation system utilizing parallel independent region groups. VLDB Journal, 2019, 28, 73-98.	2.7	7
8045	A scalable and distributed dendritic cell algorithm for big data classification. Swarm and Evolutionary Computation, 2019, 50, 100432.	4.5	10
8046	Hardware Accelerated Semantic Declarative Memory Systems through CUDA and MapReduce. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 601-614.	4.0	5
8047	Some reflections on combining meta-analysis and structural equation modeling. Research Synthesis Methods, 2019, 10, 15-22.	4.2	26
8048	Meta-Heuristic Techniques Study for Fault Tolerance in Cloud Computing Environment: A Survey Work. Advances in Intelligent Systems and Computing, 2019, , 1-11.	0.5	1
8049	Dependency-Aware Network Adaptive Scheduling of Data-Intensive Parallel Jobs. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 515-529.	4.0	11
8050	An Identity-Based Authentication Framework for Big Data Security. Lecture Notes in Networks and Systems, 2019, , 63-71.	0.5	12
8051	GAME-SCORE: Game-based energy-aware cloud scheduler and simulator for computational clouds. Simulation Modelling Practice and Theory, 2019, 93, 3-20.	2.2	29

#	ARTICLE	IF	CITATIONS
8052	Integrating Coflow and Circuit Scheduling for Optical Networks. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1346-1358.	4.0	15
8053	An Asymptotic Ensemble Learning Framework for Big Data Analysis. IEEE Access, 2019, 7, 3675-3693.	2.6	17
8054	Digger. ACM Transactions on Knowledge Discovery From Data, 2019, 13, 1-27.	2.5	4
8055	Aspect-Based Text Summarization Using MapReduce Optimization. EAI/Springer Innovations in Communication and Computing, 2019, , 131-139.	0.9	0
8056	The Internet of Things: A Review of Enabled Technologies and Future Challenges. IEEE Access, 2019, 7, 7606-7640.	2.6	152
8057	Cassandra—A Distributed Database System: An Overview. Advances in Intelligent Systems and Computing, 2019, , 519-526.	0.5	6
8058	Handling Big Data Using MapReduce Over Hybrid Cloud. Lecture Notes in Networks and Systems, 2019, , 135-144.	0.5	6
8059	A low-latency computing framework for time-evolving graphs. Journal of Supercomputing, 2019, 75, 3673-3692.	2.4	1
8060	Mining Big Data with Random Forests. Cognitive Computation, 2019, 11, 294-316.	3.6	15
8061	Libra: scalable k-mer-based tool for massive all-vs-all metagenome comparisons. GigaScience, 2019, 8, .	3.3	40
8062	Coded Computation Against Processing Delays for Virtualized Cloud-Based Channel Decoding. IEEE Transactions on Communications, 2019, 67, 28-38.	4.9	12
8063	Big Data Preprocessing for Modern World: Opportunities and Challenges. Lecture Notes on Data Engineering and Communications Technologies, 2019, , 335-343.	0.5	6
8064	Can deep learning techniques improve classification performance of vandalism detection in Wikipedia?. Engineering Applications of Artificial Intelligence, 2019, 78, 248-259.	4.3	13
8067	Ensembles for feature selection: A review and future trends. Information Fusion, 2019, 52, 1-12.	11.7	327
8068	ParStream-seq: An improved method of handling next generation sequence data. Genomics, 2019, 111, 1641-1650.	1.3	2
8069	Probabilistic modeling of renewable energy source based on Spark platform with large-scale sample data. International Transactions on Electrical Energy Systems, 2019, 29, e2759.	1.2	4
8070	A novel algorithm for handling reducer side data skew in MapReduce based on a learning automata game. Information Sciences, 2019, 501, 662-679.	4.0	14
8071	Cost-Effective Cloud Server Provisioning for Predictable Performance of Big Data Analytics. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1036-1051.	4.0	35

#	ARTICLE	IF	CITATIONS
8072	A taxonomy of software-based and hardware-based approaches for energy efficiency management in the Hadoop. Journal of Network and Computer Applications, 2019, 126, 162-177.	5.8	34
8073	A distributed data management system to support large-scale data analysis. Journal of Systems and Software, 2019, 148, 105-115.	3.3	18
8074	A framework for parallel map-matching at scale using Spark. Distributed and Parallel Databases, 2019, 37, 697-720.	1.0	5
8075	HSDC: A Highly Scalable Data Center Network Architecture for Greater Incremental Scalability. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1105-1119.	4.0	33
8076	Sentiment Analysis of Restaurant Reviews. Smart Innovation, Systems and Technologies, 2019, , 383-390.	0.5	4
8077	The Impact of Distributed Data in Big Data Platforms on Organizations. Advances in Intelligent Systems and Computing, 2019, , 1024-1036.	0.5	3
8078	A survey on cost-effective context-aware distribution of social data streams over energy-efficient data centres. Simulation Modelling Practice and Theory, 2019, 93, 42-64.	2.2	0
8079	OpenFlow Compatible Key-Based Routing Protocol: Adapting SDN Networks to Content/Service-Centric Paradigm. Journal of Network and Systems Management, 2019, 27, 730-755.	3.3	3
8080	A distributed overlapping community detection model for large graphs using autoencoder. Future Generation Computer Systems, 2019, 94, 16-26.	4.9	22
8081	Vehdoop: A Scalable Analytical Processing Framework for Vehicular Sensor Networks. IEEE Transactions on Intelligent Transportation Systems, 2019, 20, 3104-3114.	4.7	6
8082	Enhanced PeerHunter: Detecting Peer-to-Peer Botnets Through Network-Flow Level Community Behavior Analysis. IEEE Transactions on Information Forensics and Security, 2019, 14, 1485-1500.	4.5	28
8083	Design And Analysis of the Sparse Array for DoA Estimation of Noncircular Signals. IEEE Transactions on Signal Processing, 2019, 67, 460-473.	3.2	56
8084	Introducing a Vector Space Model to Perform a Proactive Credit Scoring. Communications in Computer and Information Science, 2019, , 125-148.	0.4	2
8085	QuickSquad: A new single-machine graph computing framework for detecting fake accounts in large-scale social networks. Peer-to-Peer Networking and Applications, 2019, 12, 1385-1402.	2.6	4
8086	Data Prefetching and Eviction Mechanisms of In-Memory Storage Systems Based on Scheduling for Big Data Processing. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1738-1752.	4.0	7
8087	Large-scale e-learning recommender system based on Spark and Hadoop. Journal of Big Data, 2019, 6, .	6.9	48
8088	LADRA: Log-based abnormal task detection and root-cause analysis in big data processing with Spark. Future Generation Computer Systems, 2019, 95, 392-403.	4.9	24
8089	STiMR k-Means: An Efficient Clustering Method for Big Data. International Journal of Pattern Recognition and Artificial Intelligence, 2019, 33, 1950013.	0.7	8

#	ARTICLE	IF	CITATIONS
8090	Optimizing Partition Granularity, Membership Function Parameters, and Rule Bases of Fuzzy Classifiers for Big Data by a Multi-objective Evolutionary Approach. <i>Cognitive Computation</i> , 2019, 11, 367-387.	3.6	14
8091	Review: machine learning techniques applied to cybersecurity. <i>International Journal of Machine Learning and Cybernetics</i> , 2019, 10, 2823-2836.	2.3	85
8092	Towards predictable performance via two-layer bandwidth allocation in cloud datacenter. <i>Journal of Parallel and Distributed Computing</i> , 2019, 126, 34-47.	2.7	7
8093	A Big Data Approach for the Extraction of Fuzzy Emerging Patterns. <i>Cognitive Computation</i> , 2019, 11, 400-417.	3.6	8
8094	Analysis of Twitter messages using big data tools to evaluate and locate the activity in the city of Valencia (Spain). <i>Cities</i> , 2019, 86, 37-50.	2.7	23
8095	Towards evaluation of cloud ontologies. <i>Journal of Parallel and Distributed Computing</i> , 2019, 126, 82-106.	2.7	13
8096	Clustering algorithms: A comparative approach. <i>PLoS ONE</i> , 2019, 14, e0210236.	1.1	303
8097	NoT: a high-level no-threading parallel programming method for heterogeneous systems. <i>Journal of Supercomputing</i> , 2019, 75, 3810-3841.	2.4	1
8098	Reducing partition skew on MapReduce: an incremental allocation approach. <i>Frontiers of Computer Science</i> , 2019, 13, 960-975.	1.6	7
8099	An effect of big data technology with ant colony optimization based routing in vehicular ad hoc networks: Towards smart cities. <i>Journal of Cleaner Production</i> , 2019, 217, 584-593.	4.6	71
8100	Regional-scale Detection of Fault Scarps and Other Tectonic Landforms: Examples From Northern California. <i>Journal of Geophysical Research: Solid Earth</i> , 2019, 124, 1016-1035.	1.4	8
8101	Clustering-based and QoS-aware services composition algorithm for ambient intelligence. <i>Information Sciences</i> , 2019, 482, 419-439.	4.0	42
8102	Privacy-Preserving Truth Discovery in Crowd Sensing Systems. <i>ACM Transactions on Sensor Networks</i> , 2019, 15, 1-32.	2.3	50
8103	Distributed FastShapelet Transform: a Big Data time series classification algorithm. <i>Information Sciences</i> , 2019, 496, 451-463.	4.0	27
8104	FastMFDs: a fast, efficient algorithm for mining minimal functional dependencies from large-scale distributed data with Spark. <i>Journal of Supercomputing</i> , 2019, 75, 2497-2517.	2.4	6
8105	Internet of Things (IoT), mobile cloud, cloudlet, mobile IoT, IoT cloud, fog, mobile edge, and edge emerging computing paradigms: Disambiguation and research directions. <i>Journal of Network and Computer Applications</i> , 2019, 128, 105-140.	5.8	230
8106	Client-side straggler-aware I/O scheduler for object-based parallel file systems. <i>Parallel Computing</i> , 2019, 82, 3-18.	1.3	8
8107	Evaluating the benefits of using proactive transformed-domain-based techniques in fraud detection tasks. <i>Future Generation Computer Systems</i> , 2019, 93, 18-32.	4.9	26

#	ARTICLE	IF	CITATIONS
8108	Distributed correlation-based feature selection in spark. Information Sciences, 2019, 496, 287-299.	4.0	33
8109	Proof of concept of a novel cloud computing approach for object-based remote sensing data analysis and classification. GIScience and Remote Sensing, 2019, 56, 536-553.	2.4	11
8110	Big Data in ITS: Concept, Case Studies, Opportunities, and Challenges. IEEE Transactions on Intelligent Transportation Systems, 2019, 20, 3189-3194.	4.7	14
8111	Big data analytics for healthcare industry: impact, applications, and tools. Big Data Mining and Analytics, 2019, 2, 48-57.	7.5	111
8112	The Influence of Big Data and IoT on Smart Cities. , 2019, , 151-172.		1
8113	Understanding the language of entrepreneurship. Journal of Economic and Administrative Sciences, 2019, 35, 113-127.	0.7	2
8114	Effectively Reconfigure the Optical Circuit Switching Layer Topology in Data Center Network by OCBridge. Journal of Lightwave Technology, 2019, 37, 897-908.	2.7	8
8115	Enhancing the scalability of expressive stream reasoning via input-driven parallelization. Semantic Web, 2019, 10, 457-474.	1.1	4
8116	Data Analytic for Improving Operations and Maintenance in Smart-Grid Environment. Power Systems, 2019, , 147-161.	0.3	1
8117	Overview of Scalable Partitional Methods for Big Data Clustering. Unsupervised and Semi-supervised Learning, 2019, , 1-23.	0.4	15
8118	Enabling Smart Data: Noise filtering in Big Data classification. Information Sciences, 2019, 479, 135-152.	4.0	103
8119	MULTI-CLASS RESOURCE SHARING WITH BATCH ARRIVALS. Probability in the Engineering and Informational Sciences, 2019, 33, 348-366.	0.6	1
8120	Design and implementation of reconfigurable acceleration for in-memory distributed big data computing. Future Generation Computer Systems, 2019, 92, 68-75.	4.9	6
8121	Finding tendencies in streaming data using Big Data frequent itemset mining. Knowledge-Based Systems, 2019, 163, 666-674.	4.0	36
8122	Cost-effective deadline-aware stochastic scheduling strategy for workflow applications on virtual machines in cloud computing. Concurrency Computation Practice and Experience, 2019, 31, e5006.	1.4	17
8123	A Multi-Layered Scheme for Distributed Simulations on the Cloud Environment. IEEE Transactions on Cloud Computing, 2019, 7, 5-18.	3.1	14
8124	Video Stream Analysis in Clouds: An Object Detection and Classification Framework for High Performance Video Analytics. IEEE Transactions on Cloud Computing, 2019, 7, 1152-1167.	3.1	52
8125	Formal specification and implementation of an automated pattern-based parallel-code generation framework. International Journal on Software Tools for Technology Transfer, 2019, 21, 183-202.	1.7	0

#	ARTICLE	IF	CITATIONS
8126	Heterogeneous Architectures for Big Data Batch Processing in MapReduce Paradigm. IEEE Transactions on Big Data, 2019, 5, 18-33.	4.4	17
8127	AR-RRNS: Configurable reliable distributed data storage systems for Internet of Things to ensure security. Future Generation Computer Systems, 2019, 92, 1080-1092.	4.9	76
8128	Parallel extraction of association rules from genomics data. Applied Mathematics and Computation, 2019, 350, 434-446.	1.4	28
8129	Big data and rule-based recommendation system in Internet of Things. Cluster Computing, 2019, 22, 1837-1846.	3.5	13
8130	One-pass MapReduce-based clustering method for mixed large scale data. Journal of Intelligent Information Systems, 2019, 52, 619-636.	2.8	14
8131	Learning the way to the cloud: Big Data Park. Concurrency Computation Practice and Experience, 2019, 31, e4234.	1.4	5
8132	Deep learning in big data Analytics: A comparative study. Computers and Electrical Engineering, 2019, 75, 275-287.	3.0	147
8133	Parallel and distributed core label propagation with graph coloring. Concurrency Computation Practice and Experience, 2019, 31, e4355.	1.4	6
8134	Parallel cosine nearest neighbor graph construction. Journal of Parallel and Distributed Computing, 2019, 129, 61-82.	2.7	1
8135	Resource stealing: a resource multiplexing method for mix workloads in cloud system. Journal of Supercomputing, 2019, 75, 33-49.	2.4	14
8136	Comparative Study of Optical Interconnection Architectures in Data Center Networks. Journal of Optical Communications, 2019, 40, 225-238.	4.0	1
8137	A Distributed Computing Platform for fMRI Big Data Analytics. IEEE Transactions on Big Data, 2019, 5, 109-119.	4.4	10
8138	WukaStore: Scalable, Configurable and Reliable Data Storage on Hybrid Volunteered Cloud and Desktop Systems. IEEE Transactions on Big Data, 2022, 8, 85-98.	4.4	17
8139	Speeding Up MCMC by Efficient Data Subsampling. Journal of the American Statistical Association, 2019, 114, 831-843.	1.8	66
8140	Programming models and systems for Big Data analysis. International Journal of Parallel, Emergent and Distributed Systems, 2019, 34, 632-652.	0.7	39
8141	Automatic parallelization of recursive functions with rewriting rules. Science of Computer Programming, 2019, 173, 128-152.	1.5	1
8142	An efficient parallel similarity matrix construction on MapReduce for collaborative filtering. Journal of Supercomputing, 2019, 75, 123-141.	2.4	15
8144	Gray Computing: A Framework for Computing with Background JavaScript Tasks. IEEE Transactions on Software Engineering, 2019, 45, 171-193.	4.3	3

#	ARTICLE	IF	CITATIONS
8145	Improved algorithm for parallel mining collaborative frequent itemsets in multiple data streams. Cluster Computing, 2019, 22, 6133-6141.	3.5	7
8146	Scheduling Jobs across Geo-Distributed Datacenters with Max-Min Fairness. IEEE Transactions on Network Science and Engineering, 2019, 6, 488-500.	4.1	25
8147	Very large-scale data classification based on K-means clustering and multi-kernel SVM. Soft Computing, 2019, 23, 3793-3801.	2.1	34
8148	The RePhrase Extended Pattern Set for Data Intensive Parallel Computing. International Journal of Parallel Programming, 2019, 47, 74-93.	1.1	2
8149	The Dinamica EGO virtual machine. Science of Computer Programming, 2019, 173, 3-20.	1.5	13
8150	Joint deadline-constrained and influence-aware design for allocating MapReduce jobs in cloud computing systems. Cluster Computing, 2019, 22, 6963-6976.	3.5	4
8151	On the role of message broker middleware for many-task computing on a big-data platform. Cluster Computing, 2019, 22, 2527-2540.	3.5	8
8152	GraphH: A Processing-in-Memory Architecture for Large-Scale Graph Processing. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2019, 38, 640-653.	1.9	75
8153	Spark-based intelligent parameter inversion method for prestack seismic data. Neural Computing and Applications, 2019, 31, 4577-4593.	3.2	12
8154	An improved method of locality-sensitive hashing for scalable instance matching. Knowledge and Information Systems, 2019, 58, 275-294.	2.1	8
8155	Socio-cyber network: The potential of cyber-physical system to define human behaviors using big data analytics. Future Generation Computer Systems, 2019, 92, 868-878.	4.9	21
8156	A Scientific Workflow Management System for orchestration of parallel components in a cloud of large-scale parallel processing services. Science of Computer Programming, 2019, 173, 95-127.	1.5	8
8157	Resource Provisioning for Task-Batch Based Workflows with Deadlines in Public Clouds. IEEE Transactions on Cloud Computing, 2019, 7, 814-826.	3.1	21
8158	Practical Privacy-Preserving MapReduce Based K-Means Clustering Over Large-Scale Dataset. IEEE Transactions on Cloud Computing, 2019, 7, 568-579.	3.1	34
8159	An Adaptive Efficiency-Fairness Meta-Scheduler for Data-Intensive Computing. IEEE Transactions on Services Computing, 2019, 12, 865-879.	3.2	8
8160	Adaptive Avatar Handoff in the Cloudlet Network. IEEE Transactions on Cloud Computing, 2019, 7, 664-676.	3.1	27
8161	Leveraging MapReduce and Synchrophasors for Real-Time Anomaly Detection in the Smart Grid. IEEE Transactions on Emerging Topics in Computing, 2019, 7, 392-403.	3.2	21
8162	Speedy Cloud: Cloud Computing with Support for Hardware Acceleration Services. IEEE Transactions on Cloud Computing, 2019, 7, 850-865.	3.1	6

#	ARTICLE	IF	CITATIONS
8163	Empirical Discovery of Power-Law Distribution in MapReduce Scalability. IEEE Transactions on Cloud Computing, 2019, 7, 744-755.	3.1	2
8164	Faster MapReduce Computation on Clouds Through Better Performance Estimation. IEEE Transactions on Cloud Computing, 2019, 7, 770-783.	3.1	7
8165	<i>Adia</i>: Achieving High Link Utilization with Coflow-Aware Scheduling in Data Center Networks. IEEE Transactions on Cloud Computing, 2019, 7, 431-441.	3.1	6
8166	Straggler Root-Cause and Impact Analysis for Massive-scale Virtualized Cloud Datacenters. IEEE Transactions on Services Computing, 2019, 12, 91-104.	3.2	50
8167	PISCES: Optimizing Multi-Job Application Execution in MapReduce. IEEE Transactions on Cloud Computing, 2019, 7, 273-286.	3.1	7
8168	Hadoop MapReduce for Mobile Clouds. IEEE Transactions on Cloud Computing, 2019, 7, 224-236.	3.1	12
8169	<i>Dithen</i>: A Computation-as-a-Service Cloud Platform for Large-Scale Multimedia Processing. IEEE Transactions on Cloud Computing, 2019, 7, 509-523.	3.1	8
8170	MSP: Multiple Sub-graph Query Processing using Structure-based Graph Partitioning Strategy and Map-Reduce. Journal of King Saud University - Computer and Information Sciences, 2019, 31, 22-34.	2.7	5
8171	Big Data Privacy in Biomedical Research. IEEE Transactions on Big Data, 2020, 6, 296-308.	4.4	22
8172	Using spatiotemporal distribution of geocoded Twitter data to predict US county-level health indices. Future Generation Computer Systems, 2020, 110, 620-628.	4.9	7
8173	A conceptual distributed framework for improved and secured healthcare system. International Journal of Healthcare Management, 2020, 13, 74-87.	1.2	15
8174	ODDS: Optimizing Data-Locality Access for Scientific Data Analysis. IEEE Transactions on Cloud Computing, 2020, 8, 220-231.	3.1	6
8175	Green Cloudlet Network: A Sustainable Platform for Mobile Cloud Computing. IEEE Transactions on Cloud Computing, 2020, 8, 180-192.	3.1	34
8176	Map-Balance-Reduce: An improved parallel programming model for load balancing of MapReduce. Future Generation Computer Systems, 2020, 105, 993-1001.	4.9	26
8177	Efficiently Translating Complex SQL Query to MapReduce Jobflow on Cloud. IEEE Transactions on Cloud Computing, 2020, 8, 508-517.	3.1	38
8178	Massively Distributed Time Series Indexing and Querying. IEEE Transactions on Knowledge and Data Engineering, 2020, 32, 108-120.	4.0	25
8179	HGeoHashBase: an optimized storage model of spatial objects for location-based services. Frontiers of Computer Science, 2020, 14, 208-218.	1.6	6
8180	Cornac: Tackling Huge Graph Visualization with Big Data Infrastructure. IEEE Transactions on Big Data, 2020, 6, 80-92.	4.4	8

#	ARTICLE	IF	CITATIONS
8181	A rewrite/merge approach for supporting real-time data warehousing via lightweight data integration. <i>Journal of Supercomputing</i> , 2020, 76, 3898-3922.	2.4	7
8182	Fast Recovery MapReduce (FAR-MR) to accelerate failure recovery in big data applications. <i>Journal of Supercomputing</i> , 2020, 76, 3572-3588.	2.4	8
8183	MapReduce scheduling algorithms: a review. <i>Journal of Supercomputing</i> , 2020, 76, 4915-4945.	2.4	24
8184	Big Data Processing Workflows Oriented Real-Time Scheduling Algorithm using Task-Duplication in Geo-Distributed Clouds. <i>IEEE Transactions on Big Data</i> , 2020, 6, 131-144.	4.4	38
8185	Hierarchical Edge-Cloud SDN Controller System With Optimal Adaptive Resource Allocation for Load-Balancing. <i>IEEE Systems Journal</i> , 2020, 14, 265-276.	2.9	17
8186	CFM-BD: A Distributed Rule Induction Algorithm for Building Compact Fuzzy Models in Big Data Classification Problems. <i>IEEE Transactions on Fuzzy Systems</i> , 2020, 28, 163-177.	6.5	30
8187	Aspect-based summarisation using distributed clustering and single-objective optimisation. <i>Journal of Information Science</i> , 2020, 46, 176-190.	2.0	9
8188	Profitable Taxi Travel Route Recommendation Based on Big Taxi Trajectory Data. <i>IEEE Transactions on Intelligent Transportation Systems</i> , 2020, 21, 653-668.	4.7	46
8189	Analysis of building maintenance requests using a text mining approach: building services evaluation. <i>Building Research and Information</i> , 2020, 48, 207-217.	2.0	51
8190	Indexes to Find the Optimal Number of Clusters in a Hierarchical Clustering. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 3-13.	0.5	1
8191	MPCM: Scalable and Accurate Multiple Network Alignment. <i>IEEE/ACM Transactions on Computational Biology and Bioinformatics</i> , 2020, 17, 2040-2052.	1.9	5
8192	An Efficient Framework for a Third Party Auditor in Cloud Computing Environments. <i>Computer Journal</i> , 2020, 63, 1285-1297.	1.5	6
8193	Big Data Tools, Technologies, and Applications: A Survey. <i>EAI/Springer Innovations in Communication and Computing</i> , 2020, , 453-490.	0.9	16
8194	Reliable stream data processing for elastic distributed stream processing systems. <i>Cluster Computing</i> , 2020, 23, 555-574.	3.5	6
8195	Scalable Bayesian Nonparametric Clustering and Classification. <i>Journal of Computational and Graphical Statistics</i> , 2020, 29, 53-65.	0.9	17
8196	Cloud resource management using 3Vs of Internet of Big data streams. <i>Computing (Vienna/New York)</i> , 2020, 102, 1463-1485.	3.2	5
8197	Learning of Gaussian Processes in Distributed and Communication Limited Systems. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2020, 42, 1928-1941.	9.7	7
8198	Distributed SmSVM Ensemble Learning. <i>Proceedings of the International Neural Networks Society</i> , 2020, , 7-16.	0.6	1

#	ARTICLE	IF	CITATIONS
8199	Canny edge detection and Hough transform for high resolution video streams using Hadoop and Spark. Cluster Computing, 2020, 23, 397-408.	3.5	31
8200	Fuzzy knowledge based performance analysis on big data. Neurocomputing, 2020, 389, 218-228.	3.5	7
8201	DNA short read alignment on apache spark. Applied Computing and Informatics, 2023, 19, 64-81.	3.7	2
8202	GraphMP: I/O-Efficient Big Graph Analytics on a Single Commodity Machine. IEEE Transactions on Big Data, 2020, 6, 816-829.	4.4	4
8203	Scalable Discovery of Hybrid Process Models in a Cloud Computing Environment. IEEE Transactions on Services Computing, 2020, 13, 368-380.	3.2	25
8204	Distributed time-respecting flow graph pattern matching on temporal graphs. World Wide Web, 2020, 23, 609-630.	2.7	12
8205	Asynchronous Approximation of a Single Component of the Solution to a Linear System. IEEE Transactions on Network Science and Engineering, 2020, 7, 975-986.	4.1	3
8206	RecFlow: SDN-based receiver-driven flow scheduling in datacenters. Cluster Computing, 2020, 23, 289-306.	3.5	6
8207	MR-I MaxMin-scalable two-phase border based knowledge hiding technique using MapReduce. Future Generation Computer Systems, 2020, 109, 538-550.	4.9	3
8208	Fine-grained scheduling in multi-resource clusters. Journal of Supercomputing, 2020, 76, 1931-1958.	2.4	0
8209	FPO tree and DP3 algorithm for distributed parallel Frequent Itemsets Mining. Expert Systems With Applications, 2020, 140, 112874.	4.4	12
8210	Measuring data-centre workflows complexity through process mining: the Google cluster case. Journal of Supercomputing, 2020, 76, 2449-2478.	2.4	15
8211	PTCP: A priority-based transport control protocol for timeout mitigation in commodity data center. Future Generation Computer Systems, 2020, 102, 619-632.	4.9	5
8212	Real-time big data processing for instantaneous marketing decisions: A problematization approach. Industrial Marketing Management, 2020, 90, 558-569.	3.7	74
8213	SP-BRAIN: scalable and reliable implementations of a supervised relevance-based machine learning algorithm. Soft Computing, 2020, 24, 7417-7434.	2.1	1
8214	Visualization analysis of big data research based on Citespace. Soft Computing, 2020, 24, 8173-8186.	2.1	84
8215	Context agnostic trajectory prediction based on $\langle \text{mml:math xmlns:mml="http://www.w3.org/1998/Math/MathML" display="inline" id="d1e652" altimg="si22.svg" \rangle \langle \text{mml:mi} \rangle \langle \text{mml:mi} \rangle \langle \text{mml:math} \rangle$ -architecture. Future Generation Computer Systems, 2020, 110, 531-539.	4.9	10
8216	Scaling modeling and simulation on high-performance computing clusters. Simulation, 2020, 96, 221-232.	1.1	4

#	ARTICLE	IF	CITATIONS
8217	Statistical strategies for the analysis of massive data sets. <i>Biometrical Journal</i> , 2020, 62, 270-281.	0.6	6
8218	Distributed multi-label feature selection using individual mutual information measures. <i>Knowledge-Based Systems</i> , 2020, 188, 105052.	4.0	85
8219	H2Pregel : A partition-based hybrid hierarchical graph computation approach. <i>Future Generation Computer Systems</i> , 2020, 104, 15-31.	4.9	3
8220	Manufacturing big data ecosystem: A systematic literature review. <i>Robotics and Computer-Integrated Manufacturing</i> , 2020, 62, 101861.	6.1	182
8221	DistriPlan: an optimized join execution framework for geo-distributed scientific data. <i>Distributed and Parallel Databases</i> , 2020, 38, 127-152.	1.0	1
8222	Approaches of enhancing interoperations among high performance computing and big data analytics via augmentation. <i>Cluster Computing</i> , 2020, 23, 953-988.	3.5	8
8223	Distributed outlier detection in hierarchically structured datasets with mixed attributes. <i>Quality Technology and Quantitative Management</i> , 2020, 17, 337-353.	1.1	2
8224	Implementation and evaluation distributed mixed pixels analysis algorithm for hyperspectral image based on constraint non-negative matrix factorization. <i>International Journal of Parallel, Emergent and Distributed Systems</i> , 2020, 35, 365-375.	0.7	0
8225	Parallel co-location mining with MapReduce and NoSQL systems. <i>Knowledge and Information Systems</i> , 2020, 62, 1433-1463.	2.1	21
8226	PowerHash: a hybrid grouping scheme by leveraging power-law properties of data. <i>International Journal of Data Science and Analytics</i> , 2020, 9, 273-284.	2.4	0
8227	A review of the literature on big data analytics in healthcare. <i>Journal of the Operational Research Society</i> , 2020, 71, 1511-1529.	2.1	57
8228	Pivot-based approximate k-NN similarity joins for big high-dimensional data. <i>Information Systems</i> , 2020, 87, 101410.	2.4	10
8229	Heterogeneous megamodel management using collection operators. <i>Software and Systems Modeling</i> , 2020, 19, 231-260.	2.2	4
8230	KPN-based parallelization of Wu's Manber algorithm on multi-core machines. <i>Cluster Computing</i> , 2020, 23, 725-734.	3.5	0
8231	Harnessing Remote Sensing for Civil Engineering: Then, Now, and Tomorrow. <i>Lecture Notes in Civil Engineering</i> , 2020, , 3-30.	0.3	10
8232	A Lightweight Indexing Approach for Efficient Batch Similarity Processing with MapReduce. <i>SN Computer Science</i> , 2020, 1, 1.	2.3	34
8233	Generating synthetic positive and negative business process traces through abduction. <i>Knowledge and Information Systems</i> , 2020, 62, 813-839.	2.1	2
8234	Online Decentralized and Cooperative Dispatch for Multi-Microgrids. <i>IEEE Transactions on Automation Science and Engineering</i> , 2020, 17, 450-462.	3.4	19

#	ARTICLE	IF	CITATIONS
8235	Optimal Scheduling of Tasks in Cloud Computing Using Hybrid Firefly-Genetic Algorithm. Learning and Analytics in Intelligent Systems, 2020, , 678-687.	0.5	22
8236	HeteroYARN: A Heterogeneous FPGA-Accelerated Architecture Based on YARN. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 2968-2980.	4.0	2
8237	Dual-PISA: An index for aggregation operations on time series data. Information Systems, 2020, 87, 101427.	2.4	3
8239	On the Optimal Recovery Threshold of Coded Matrix Multiplication. IEEE Transactions on Information Theory, 2020, 66, 278-301.	1.5	146
8240	UPS: Unified PMU-Data Storage System to Enhance T+D PMU Data Usability. IEEE Transactions on Smart Grid, 2020, 11, 739-748.	6.2	18
8241	Optimizing Geo-Distributed Data Analytics with Coordinated Task Scheduling and Routing. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 279-293.	4.0	20
8242	Threats on the horizon: understanding security threats in the era of cyber-physical systems. Journal of Supercomputing, 2020, 76, 2643-2664.	2.4	45
8243	MapReduce based improved quick reduct algorithm with granular refinement using vertical partitioning scheme. Knowledge-Based Systems, 2020, 189, 105104.	4.0	10
8244	A survey of current challenges in partitioning and processing of graph-structured data in parallel and distributed systems. Distributed and Parallel Databases, 2020, 38, 495-530.	1.0	13
8245	Photonic Multiply-Accumulate Operations for Neural Networks. IEEE Journal of Selected Topics in Quantum Electronics, 2020, 26, 1-18.	1.9	166
8246	Survey of external memory large-scale graph processing on a multi-core system. Journal of Supercomputing, 2020, 76, 549-579.	2.4	4
8247	A combined entropy-based approach for a proactive credit scoring. Engineering Applications of Artificial Intelligence, 2020, 87, 103292.	4.3	22
8248	Efficient development of high performance data analytics in Python. Future Generation Computer Systems, 2020, 111, 570-581.	4.9	6
8249	Towards a distributed local-search approach for partitioning large-scale social networks. Information Sciences, 2020, 508, 200-213.	4.0	6
8250	Autonomic Intrusion Detection and Response Using Big Data. IEEE Systems Journal, 2020, 14, 1984-1991.	2.9	18
8251	Challenges and Opportunities for Coping with the Smart Divide in Rural America. Annals of the American Association of Geographers, 2020, 110, 559-570.	1.5	9
8252	A parallel classification framework for protein fold recognition. Evolutionary Intelligence, 2020, 13, 525-535.	2.3	1
8253	Smart Intra-query Fault Tolerance for Massive Parallel Processing Databases. Data Science and Engineering, 2020, 5, 65-79.	4.6	12

#	ARTICLE	IF	CITATIONS
8254	VeriTable: Fast equivalence verification of multiple large forwarding tables. Computer Networks, 2020, 168, 106981.	3.2	0
8255	Improved Teaching Learning-Based QoS-Aware Services Composition for Internet of Things. IEEE Systems Journal, 2020, 14, 4155-4164.	2.9	13
8256	The growing role of integrated and insightful big and real-time data analytics platforms. Advances in Computers, 2020, , 165-186.	1.2	5
8257	Handling data skew at reduce stage in Spark by ReducePartition. Concurrency Computation Practice and Experience, 2020, 32, e5637.	1.4	4
8258	Detection outliers on internet of things using big data technology. Egyptian Informatics Journal, 2020, 21, 131-138.	4.4	30
8259	DQPFS: Distributed quadratic programming based feature selection for big data. Journal of Parallel and Distributed Computing, 2020, 138, 1-14.	2.7	10
8260	Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. Journal of Manufacturing Systems, 2020, 54, 138-151.	7.6	213
8261	Distributed mining of high utility time interval sequential patterns using mapreduce approach. Expert Systems With Applications, 2020, 141, 112967.	4.4	21
8262	Cost-aware load balancing for multilingual record linkage using MapReduce. Ain Shams Engineering Journal, 2020, 11, 419-433.	3.5	7
8263	A component-based framework for certification of components in a cloud of HPC services. Science of Computer Programming, 2020, 191, 102379.	1.5	3
8264	Joint Optimization of MapReduce Scheduling and Network Policy in Hierarchical Data Centers. IEEE Transactions on Cloud Computing, 2022, 10, 461-473.	3.1	2
8265	When Gaussian Process Meets Big Data: A Review of Scalable GPs. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31, 4405-4423.	7.2	342
8266	Discovering multi-dimensional motifs from multi-dimensional time series for air pollution control. Concurrency Computation Practice and Experience, 2020, 32, e5645.	1.4	2
8267	SMusket: Spark-based DNA error correction on distributed-memory systems. Future Generation Computer Systems, 2020, 111, 698-713.	4.9	8
8268	Hidden Markov Model-based Load Balancing in Data Center Networks. Computer Journal, 2020, 63, 1449-1462.	1.5	3
8269	Graph Computing-Based WLS Fast Decoupled State Estimation. IEEE Transactions on Smart Grid, 2020, 11, 2440-2451.	6.2	17
8270	Parallelizing Machine Learning as a service for the end-user. Future Generation Computer Systems, 2020, 105, 275-286.	4.9	6
8271	A multi-agent architecture for mobile sensing systems. Journal of Ambient Intelligence and Humanized Computing, 2020, 11, 4439-4451.	3.3	7

#	ARTICLE	IF	CITATIONS
8272	A distributed computing framework for wind speed big data forecasting on Apache Spark. Sustainable Energy Technologies and Assessments, 2020, 37, 100582.	1.7	31
8273	Distributed Pregel-based provenance-aware regular path query processing on RDF knowledge graphs. World Wide Web, 2020, 23, 1465-1496.	2.7	20
8274	Distributed simultaneous inference in generalized linear models via confidence distribution. Journal of Multivariate Analysis, 2020, 176, 104567.	0.5	15
8275	Comparison of MongoDB and Cassandra Databases for Spectrum Monitoring As-a-Service. IEEE Transactions on Network and Service Management, 2020, 17, 346-360.	3.2	29
8276	A power optimised and reprogrammable system for smart wireless vibration monitoring. Structural Control and Health Monitoring, 2020, 27, e2468.	1.9	6
8277	Towards product-level parallel computing of large-scale building information modeling data using graph theory. Building and Environment, 2020, 169, 106558.	3.0	10
8278	Scaling density-based community detection to large-scale social networks via MapReduce framework. Journal of Intelligent and Fuzzy Systems, 2020, 38, 1663-1674.	0.8	2
8279	Observing and Mitigating Micro-Burst Traffic in Data Center Networks. IEEE/ACM Transactions on Networking, 2020, 28, 98-111.	2.6	14
8280	Network Function Virtualization Resource Allocation Based on Joint Benders Decomposition and ADMM. IEEE Transactions on Vehicular Technology, 2020, 69, 1706-1718.	3.9	20
8281	Big data analytics in health: an overview and bibliometric study of research activity. Health Information and Libraries Journal, 2020, 37, 5-25.	1.3	23
8282	FPGAs in the Datacenters: the Case of Parallel Hybrid Super Scalar String Sample Sort. , 2020, , .		2
8283	Temporal Fusion Pointer network-based Reinforcement Learning algorithm for Multi-Objective Workflow Scheduling in the cloud. , 2020, , .		4
8284	Distributed Graph Diameter Approximation. Algorithms, 2020, 13, 216.	1.2	4
8285	Robust multi-view k-means clustering with outlier removal. Knowledge-Based Systems, 2020, 210, 106518.	4.0	18
8286	SplitCast: Optimizing Multicast Flows in Reconfigurable Datacenter Networks. , 2020, , .		10
8287	Coded Distributed Computing with Heterogeneous Function Assignments. , 2020, , .		15
8288	TUNERCAR: A Superoptimization Toolchain for Autonomous Racing. , 2020, , .		15
8289	Research on Intelligent Concentrator Based on Edge Computing Technology. , 2020, , .		1

#	ARTICLE	IF	CITATIONS
8290	Fairness-Efficiency Scheduling for Cloud Computing With Soft Fairness Guarantees. IEEE Transactions on Cloud Computing, 2022, 10, 1806-1818.	3.1	13
8291	Prophet: Toward Fast, Error-Tolerant Model-Based Throughput Prediction for Reactive Flows in DC Networks. IEEE/ACM Transactions on Networking, 2020, 28, 2475-2488.	2.6	5
8292	Partition Selection for Large-Scale Data Management Using KNN Join Processing. Mathematical Problems in Engineering, 2020, 2020, 1-14.	0.6	2
8293	Round Compression for Parallel Matching Algorithms. SIAM Journal on Computing, 2020, 49, STOC18-1-STOC18-44.	0.8	5
8294	WP-SGD: Weighted parallel SGD for distributed unbalanced-workload training system. Journal of Parallel and Distributed Computing, 2020, 145, 202-216.	2.7	2
8295	ImRP: A Predictive Partition Method for Data Skew Alleviation in Spark Streaming Environment. Parallel Computing, 2020, 100, 102699.	1.3	6
8296	Efficient Compilation and Execution of JVM-Based Data Processing Frameworks on Heterogeneous Co-Processors. , 2020, , .		0
8297	Improved feature decay algorithms for statistical machine translation. Natural Language Engineering, 2022, 28, 71-91.	2.1	3
8298	Cloud curriculum resource management platform based on Hadoop. Measurement and Control, 2020, 53, 1782-1790.	0.9	6
8299	Data-Driven Molecular Dynamics: A Multifaceted Challenge. Pharmaceuticals, 2020, 13, 253.	1.7	33
8300	A Review of Supervised Classification based on Contrast Patterns: Applications, Trends, and Challenges. Journal of Grid Computing, 2020, 18, 797-845.	2.5	30
8301	Big data in biology: The hope and present-day challenges in it. Gene Reports, 2020, 21, 100869.	0.4	22
8302	Performance optimization of non-equilibrium ionization simulations from MapReduce and GPU acceleration. Parallel Computing, 2020, 98, 102682.	1.3	0
8303	Virtual Machine Placement with Disk Anti-colocation Constraints Using Variable Neighborhood Search Heuristic. Information Systems Frontiers, 2021, 23, 1245-1271.	4.1	3
8304	Entity deduplication in big data graphs for scholarly communication. Data Technologies and Applications, 2020, 54, 409-435.	0.9	5
8305	An independent time optimized hybrid infrastructure for big data analytics. Modern Physics Letters B, 2020, 34, 2050311.	1.0	20
8306	A comprehensive survey of anomaly detection techniques for high dimensional big data. Journal of Big Data, 2020, 7, .	6.9	160
8307	BloTA: A Buildout IoT Application Language. IEEE Access, 2020, 8, 126443-126459.	2.6	9

#	ARTICLE	IF	CITATIONS
8308	Nonlinear Control Analysis of Quantized Congestion Notification in Data Center Networks. IEEE Access, 2020, 8, 125401-125411.	2.6	2
8309	Review of the Basic Technology of Big Data on Analysis and Storage Applied in Power Grids. IOP Conference Series: Earth and Environmental Science, 2020, 440, 032144.	0.2	0
8310	Scaling Blockchains: A Comprehensive Survey. IEEE Access, 2020, 8, 125244-125262.	2.6	138
8311	PFPMine: A parallel approach for discovering interacting data entities in data-intensive cloud workflows. Future Generation Computer Systems, 2020, 113, 474-487.	4.9	3
8312	Cloud computing simulators: A comprehensive review. Simulation Modelling Practice and Theory, 2020, 104, 102144.	2.2	55
8313	Vision Paper: Grand Challenges in Resilience: Autonomous System Resilience through Design and Runtime Measures. IEEE Open Journal of the Computer Society, 2020, 1, 155-172.	5.2	14
8314	DASSA: Parallel DAS Data Storage and Analysis for Subsurface Event Detection. , 2020, , .		5
8315	HydraMini: An FPGA-based Affordable Research and Education Platform for Autonomous Driving. , 2020, , .		5
8316	An effective multi-level synchronization clustering method based on a linear weighted Vicsek model. Applied Intelligence, 2020, 50, 4063-4080.	3.3	4
8317	Remote Method Delegation: a Platform for Grid Computing. Journal of Grid Computing, 2020, 18, 711-725.	2.5	2
8318	GigaSOM.jl: High-performance clustering and visualization of huge cytometry datasets. GigaScience, 2020, 9, .	3.3	8
8319	Optimizing performance of GATK workflows using Apache Arrow In-Memory data framework. BMC Genomics, 2020, 21, 683.	1.2	4
8320	An effective architecture of digital twin system to support human decision making and AI-driven autonomy. Concurrency Computation Practice and Experience, 2021, 33, e6111.	1.4	12
8322	FPGA-Based On-Board Hyperspectral Imaging Compression: Benchmarking Performance and Energy Efficiency against GPU Implementations. Remote Sensing, 2020, 12, 3741.	1.8	14
8323	ViSeR: Visual Self-Regularization. , 2020, , .		0
8324	A Zeroth-Order ADMM Algorithm for Stochastic Optimization over Distributed Processing Networks. , 2020, , .		1
8325	Application of Keyword Dynamic Query Software in Relational Database based on Big Data. , 2020, , .		2
8326	StragglerHelper: Alleviating Straggling in Computing Clusters via Sharing Memory Access Patterns. , 2020, , .		1

#	ARTICLE	IF	CITATIONS
8327	E2PAMEA: A fast evolutionary algorithm for extracting fuzzy emerging patterns in big data environments. <i>Neurocomputing</i> , 2020, 415, 60-73.	3.5	5
8328	Optimizing data center network throughput by solving TCP Incast problem using k-means algorithm. <i>International Journal of Communication Systems</i> , 0, , e4535.	1.6	2
8329	Evaluating the Effects of Modern Storage Devices on the Efficiency of Parallel Machine Learning Algorithms. <i>International Journal on Artificial Intelligence Tools</i> , 2020, 29, 2060008.	0.7	0
8330	SwiftIDS: Real-time intrusion detection system based on LightGBM and parallel intrusion detection mechanism. <i>Computers and Security</i> , 2020, 97, 101984.	4.0	101
8331	Big Data-Driven Abnormal Behavior Detection in Healthcare Based on Association Rules. <i>IEEE Access</i> , 2020, 8, 129002-129011.	2.6	20
8332	Q-Flink: A QoS-Aware Controller for Apache Flink. , 2020, , .		9
8333	Can We Build a Search Engine over Spark?. , 2020, , .		0
8334	Memory-Efficient and Skew-Tolerant MapReduce Over MPI for Supercomputing Systems. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2020, 31, 2734-2748.	4.0	4
8335	Multidimensional Parallel Dynamic Programming Algorithm Based on Spark for Large-Scale Hydropower Systems. <i>Water Resources Management</i> , 2020, 34, 3427-3444.	1.9	9
8336	Multi-level feature filtering algorithm based on MapReduce. <i>Journal of Physics: Conference Series</i> , 2020, 1486, 032027.	0.3	0
8337	Enforcing Deadlines for Skeleton-based Parallel Programming. , 2020, , .		3
8338	Communication Pattern Based Data Authentication (CPDA) Designed for Big Data Processing in a Multiple Public Cloud Environment. <i>IEEE Access</i> , 2020, 8, 107716-107748.	2.6	2
8339	Reservation and Checkpointing Strategies for Stochastic Jobs. , 2020, , .		1
8340	Traffic Behavior in Cloud Data Centers: A Survey. , 2020, , .		6
8341	Bert: Scalable Source Routed Multicast for Cloud Data Centers. , 2020, , .		1
8342	Spark implementation of the enhanced Scatter Search metaheuristic: Methodology and assessment. <i>Swarm and Evolutionary Computation</i> , 2020, 59, 100748.	4.5	9
8343	Big Spatial Data Management for the Internet of Things: A Survey. <i>Journal of Network and Systems Management</i> , 2020, 28, 990-1035.	3.3	13
8344	NumbaSummarizer: A Python Library for Simplified Vectorization Reports. , 2020, , .		2

#	ARTICLE	IF	CITATIONS
8345	MapReduce Analytics Based Fast Search Algorithms. Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers, Series A/Chung-kuo Kung Ch'eng Hsueh K'uan, 2020, 43, 831-837.	0.6	1
8346	Kulla, a container-centric construction model for building infrastructure-agnostic distributed and parallel applications. Journal of Systems and Software, 2020, 168, 110665.	3.3	6
8347	Review of Apriori based Frequent Itemset Mining Solutions on Big Data. , 2020, , .		6
8348	Think Cities: the accelerator for sustainable planning. , 2020, , .		0
8349	TOSS: A Topology-based Scheduler for Storm Clusters. , 2020, , .		2
8350	MTHAEL: Cross-Architecture IoT Malware Detection Based on Neural Network Advanced Ensemble Learning. IEEE Transactions on Computers, 2020, 69, 1654-1667.	2.4	50
8351	Communication-Efficient Jaccard similarity for High-Performance Distributed Genome Comparisons. , 2020, , .		34
8352	Fog Computing-inspired Smart Home Framework for Predictive Veterinary Healthcare. Microprocessors and Microsystems, 2020, 78, 103227.	1.8	21
8353	Distributed Local Bayesian Network for Gene Regulatory Network Reconstruction. , 2020, , .		2
8354	Salamander: a Holistic Scheduling of MapReduce Jobs on Ephemeral Cloud Resources. , 2020, , .		3
8355	Performance benefits of Intel® Optane™ DC persistent memory for the parallel processing of large neuroimaging data. , 2020, , .		3
8356	On Scheduling Coflows. Algorithmica, 2020, 82, 3604-3629.	1.0	3
8357	Marabunta: Continuous Distributed Processing of Skewed Streams. , 2020, , .		3
8358	Towards Performance Optimization for Hadoop MapReduce Applications. , 2020, , .		0
8359	Scheduling Stochastic Real-Time Jobs In Unreliable Workers. , 2020, , .		0
8360	A Survey on Large-Scale Machine Learning. IEEE Transactions on Knowledge and Data Engineering, 2020, , 1-1.	4.0	43
8361	The Rise of NoSQL Systems. Journal of Database Management, 2020, 31, 67-82.	1.0	5
8362	GeoClone: Online Task Replication and Scheduling for Geo-Distributed Analytics under Uncertainties. , 2020, , .		3

#	ARTICLE	IF	CITATIONS
8363	Data Analysis as a Tool for the Application of Adaptive Learning in a University Environment. Applied Sciences (Switzerland), 2020, 10, 7016.	1.3	14
8364	Scalable Recommendation Using Large Scale Graph Partitioning With Pregel and Giraph. International Journal of Cognitive Informatics and Natural Intelligence, 2020, 14, 42-61.	0.4	6
8365	A Distributed Quantum-Behaved Particle Swarm Optimization Using Opposition-Based Learning on Spark for Large-Scale Optimization Problem. Mathematics, 2020, 8, 1860.	1.1	6
8366	TMaR: a two-stage MapReduce scheduler for heterogeneous environments. Human-centric Computing and Information Sciences, 2020, 10, .	6.1	4
8367	MRS-DP: Improving Performance and Resource Utilization of Big Data Applications with Deadlines and Priorities. Big Data, 2020, 8, 323-331.	2.1	1
8368	2Lbp-RRNS: Two-Levels RRNS With Backpropagation for Increased Reliability and Privacy-Preserving of Secure Multi-Clouds Data Storage. IEEE Access, 2020, 8, 199424-199439.	2.6	6
8369	Improving Collection Dynamics by Monotonic Filtering. , 2020, , .		3
8370	A Low Complexity Decentralized Neural Net with Centralized Equivalence using Layer-wise Learning. , 2020, , .		3
8371	Facing Big Data by an Agent-Based Multimodal Evolutionary Approach to Classification. , 2020, , .		1
8372	DiPLIP: Distributed Parallel Processing Platform for Stream Image Processing Based on Deep Learning Model Inference. Electronics (Switzerland), 2020, 9, 1664.	1.8	5
8373	Comparison of MPI and Spark for Data Science Applications. , 2020, , .		4
8374	Analyzing spatial analytics systems based on Hadoop and Spark: A user perspective. Software - Practice and Experience, 2020, 50, 2121-2144.	2.5	6
8375	New Approaches for Performance Optimization and Analysis of Large-Scale Dynamic Social Network Analysis using Anytime Anywhere Algorithms. , 2020, , .		1
8376	SeQual: Big Data Tool to Perform Quality Control and Data Preprocessing of Large NGS Datasets. IEEE Access, 2020, 8, 146075-146084.	2.6	5
8377	Optimization for Large-Scale Fuzzy Joins Using Fuzzy Filters in MapReduce. , 2020, , .		1
8378	Low Complexity Distributed Computing via Binary Matrices with Extension to Stragglers. , 2020, , .		5
8379	LPA: An Automated, Accurate and Efficient Differentially Private Big-Data Mining System. , 2020, , .		1
8380	DeepAalo: Auto-adjusting Demotion Thresholds for Information-agnostic Coflow Scheduling. , 2020, , .		6

#	ARTICLE	IF	CITATIONS
8381	Methods and Experiences for Developing Abstractions for Data-intensive, Scientific Applications. , 2020, , .		0
8383	Is Stability Achievable?. , 2020, , 107-123.		0
8384	Fluid Limits, Fluid Equations, and Positive Recurrence. , 2020, , 124-147.		0
8385	Proving Fluid Model Stability Using Lyapunov Functions. , 2020, , 160-193.		0
8386	Max-Weight and Back-Pressure Control. , 2020, , 194-218.		0
8387	Proportionally Fair Resource Allocation. , 2020, , 219-251.		0
8388	Task Allocation in Server Farms. , 2020, , 252-267.		0
8391	Pairs (Re)Loaded: System Design & Benchmarking For Scalable Geospatial Applications. , 2020, , .		2
8392	Distributed Training and Inference of Deep Learning Models for Multi-Modal Land Cover Classification. Remote Sensing, 2020, 12, 2670.	1.8	13
8393	Comparison of Different Computing Platforms for Implementing Parallel Genetic Programming. , 2020, , .		2
8394	On Byzantine-Resilient High-Dimensional Stochastic Gradient Descent. , 2020, , .		3
8395	Spark-based parallel calculation of 3D fourier shell correlation for macromolecule structure local resolution estimation. BMC Bioinformatics, 2020, 21, 391.	1.2	1
8396	A Security Monitoring Framework for Mobile Devices. Electronics (Switzerland), 2020, 9, 1197.	1.8	7
8397	Performance Evaluation of an Independent Time Optimized Infrastructure for Big Data Analytics that Maintains Symmetry. Symmetry, 2020, 12, 1274.	1.1	29
8398	ACQuA: A Parallel Accelerator Architecture for Pure Functional Programs. , 2020, , .		1
8399	Pricing and personal data collection strategies of online platforms in the face of privacy concerns. Electronic Commerce Research, 2022, 22, 539-559.	3.0	16
8400	Coded Computation Against Straggling Channel Decoders in the Cloud for Gaussian Channels. , 2020, , .		1
8401	Sequential addition of coded sub-tasks for straggler mitigation. , 2020, , .		8

#	ARTICLE	IF	CITATIONS
8403	Stochastic Processing Networks. , 2020, , 29-53.		0
8404	Markov Representations. , 2020, , 54-78.		0
8405	Extensions and Complements. , 2020, , 79-106.		0
8406	Fluid Equations That Characterize Specific Policies. , 2020, , 148-159.		0
8407	Multihop Packet Networks. , 2020, , 268-310.		0
8413	Typhoon Quantitative Rainfall Prediction from Big Data Analytics by Using the Apache Hadoop Spark Parallel Computing Framework. Atmosphere, 2020, 11, 870.	1.0	12
8414	The Conditional Reliability Evaluation of Data Center Network BCDC. Computer Journal, 2021, 64, 1451-1464.	1.5	10
8415	5G-enabled Edge Computing for MapReduce-based Data Pre-processing. , 2020, , .		2
8416	Coflow Scheduling with Performance Guarantees for Data Center Applications. , 2020, , .		2
8417	Resolvable Designs for Speeding Up Distributed Computing. IEEE/ACM Transactions on Networking, 2020, 28, 1657-1670.	2.6	20
8418	Video Big Data Analytics in the Cloud: A Reference Architecture, Survey, Opportunities, and Open Research Issues. IEEE Access, 2020, 8, 152377-152422.	2.6	17
8419	Network Scheduling and Compute Resource Aware Task Placement in Datacenters. IEEE/ACM Transactions on Networking, 2020, 28, 2435-2448.	2.6	4
8420	A Distributed Low-Complexity Coding Solution for Large-Scale Distributed FFT. IEEE Transactions on Communications, 2020, 68, 6617-6628.	4.9	6
8421	Big Data Processing: Batch-based processing and stream-based processing. , 2020, , .		6
8422	Coordinative Scheduling of Computation and Communication in Data-parallel Systems. IEEE Transactions on Computers, 2020, , 1-1.	2.4	1
8423	Scheduling-Guided Automatic Processing of Massive Hyperspectral Image Classification on Cloud Computing Architectures. IEEE Transactions on Cybernetics, 2021, 51, 3588-3601.	6.2	54
8424	Machine Learning-Based Method for Remaining Range Prediction of Electric Vehicles. IEEE Access, 2020, 8, 212423-212441.	2.6	22
8425	Cascaded Coded Distributed Computing Schemes Based on Placement Delivery Arrays. IEEE Access, 2020, 8, 221385-221395.	2.6	7

#	ARTICLE	IF	CITATIONS
8426	The role of randomness in the broadcast congested clique model. Information and Computation, 2020, , 104669.	0.5	0
8427	A Three-Level Optimization Model for Nonlinearly Separable Clustering. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34, 3211-3218.	3.6	5
8428	A Scalable Virtualized Server Cluster Providing Sensor Data Storage and Web Services. Symmetry, 2020, 12, 1942.	1.1	3
8429	Exploiting Sharing Join Opportunities in Big Data Multiquery Optimization with Flink. Complexity, 2020, 2020, 1-25.	0.9	2
8430	A New Big Data Benchmark for OLAP Cube Design Using Data Pre-Aggregation Techniques. Applied Sciences (Switzerland), 2020, 10, 8674.	1.3	11
8431	Research on High Efficiency IO Control Technology of Seismic Data Based on Data Segmentation. IOP Conference Series: Materials Science and Engineering, 2020, 729, 012016.	0.3	0
8432	QoSComm: A Data Flow Allocation Strategy among SDN-Based Data Centers for IoT Big Data Analytics. Applied Sciences (Switzerland), 2020, 10, 7586.	1.3	2
8433	gMSR: A Multi-GPU Algorithm to Accelerate a Massive Validation of Biclusters. Electronics (Switzerland), 2020, 9, 1782.	1.8	1
8434	Summarizer: Fuzzy Rule-Based Classification Systems for Vertical and Horizontal Big Data. , 2020, , .		1
8435	In-Memory Data Anonymization Using Scalable and High Performance RDD Design. Electronics (Switzerland), 2020, 9, 1732.	1.8	9
8436	A Local Feature Engineering Strategy to Improve Network Anomaly Detection. Future Internet, 2020, 12, 177.	2.4	23
8437	Bottleneck-Aware Coflow Scheduling Without Prior Knowledge. , 2020, , .		7
8438	FCPP: an efficient and extensible Field Calculus framework. , 2020, , .		17
8439	A deep learning approach for forecasting non-stationary big remote sensing time series. Arabian Journal of Geosciences, 2020, 13, 1.	0.6	15
8440	Adaptive Density-Based Spatial Clustering for Massive Data Analysis. IEEE Access, 2020, 8, 23346-23358.	2.6	19
8441	Importance of Data Distribution on Hive-Based Systems for Query Performance: An Experimental Study. , 2020, , .		2
8442	A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security. IEEE Communications Surveys and Tutorials, 2020, 22, 1646-1685.	24.8	576
8443	Performance Enhancement in Big Data handling. , 2020, , .		2

#	ARTICLE	IF	CITATIONS
8444	Minimizing Latency for Secure Coded Computing Using Secret Sharing via Staircase Codes. IEEE Transactions on Communications, 2020, 68, 4609-4619.	4.9	34
8445	Dominoes: An Interactive Exploratory Data Analysis Tool for Software Relationships. IEEE Transactions on Software Engineering, 2022, 48, 377-396.	4.3	1
8446	A modelling platform for optimizing time-dependent transit fares in large-scale multimodal networks. Transport Policy, 2020, 92, 38-54.	3.4	10
8447	POF-SVLM: pareto optimized framework for seamless VM live migration. Computing (Vienna/New York), 2020, 102, 2159-2183.	3.2	3
8448	Parallel performance of molecular dynamics trajectory analysis. Concurrency Computation Practice and Experience, 2020, 32, e5789.	1.4	1
8449	Towards Extract-Transform-Load Operations in a Big Data context. International Journal of Sociotechnology and Knowledge Development, 2020, 12, 77-95.	0.4	13
8450	Automated Deployment of a Spark Cluster with Machine Learning Algorithm Integration. Big Data Research, 2020, 19-20, 100135.	2.6	5
8451	An Optimal Locality-Aware Task Scheduling Algorithm Based on Bipartite Graph Modelling for Spark Applications. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 2406-2420.	4.0	21
8452	Endpoint-Flexible Coflow Scheduling Across Geo-Distributed Datacenters. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 2466-2481.	4.0	14
8453	TACC: Topology-Aware Coded Computing for Distributed Graph Processing. IEEE Transactions on Signal and Information Processing Over Networks, 2020, 6, 508-525.	1.6	3
8454	Models for Internal Clustering Validation Indexes Based on Hadoop-MapReduce. International Journal of Distributed Systems and Technologies, 2020, 11, 42-67.	0.6	1
8457	The four dimensions of social network analysis: An overview of research methods, applications, and software tools. Information Fusion, 2020, 63, 88-120.	11.7	143
8458	Persistence of RDF Data into NoSQL: A Survey and a Reference Architecture. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 1370-1389.	4.0	2
8459	Two-Step Classification with SVD Preprocessing of Distributed Massive Datasets in Apache Spark. Algorithms, 2020, 13, 71.	1.2	7
8460	Efficient scientific workflow scheduling for deadline-constrained parallel tasks in cloud computing environments. Information Sciences, 2020, 531, 31-46.	4.0	55
8461	Unbiased Markov Chain Monte Carlo Methods with Couplings. Journal of the Royal Statistical Society Series B: Statistical Methodology, 2020, 82, 543-600.	1.1	54
8462	Bigflow: A General Optimization Layer for Distributed Computing Frameworks. Journal of Computer Science and Technology, 2020, 35, 453-467.	0.9	1
8463	Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. Computers in Industry, 2020, 120, 103244.	5.7	103

#	ARTICLE	IF	CITATIONS
8464	Cost effective stream workflow scheduling to handle application structural changes. Future Generation Computer Systems, 2020, 112, 348-361.	4.9	8
8465	Sensing the Web for Induction of Association Rules and their Composition through Ensemble Techniques. Procedia Computer Science, 2020, 169, 851-859.	1.2	2
8466	Efficient orchestration of Node-RED IoT workflows using a Vector Symbolic Architecture. Future Generation Computer Systems, 2020, 111, 117-131.	4.9	7
8467	High-Performance Machine Learning for Large-Scale Data Classification considering Class Imbalance. Scientific Programming, 2020, 2020, 1-16.	0.5	11
8468	Modeling and Research of Processing Big Data Sets in Distributed Information Systems. , 2020, , .		0
8469	A Fuzzy Mining Approach for Energy Efficiency in a Big Data Framework. IEEE Transactions on Fuzzy Systems, 2020, 28, 2747-2758.	6.5	13
8470	MaRe: Processing Big Data with application containers on Apache Spark. GigaScience, 2020, 9, .	3.3	6
8471	Coordinate Cache Management for Performance Improvement in Spark. , 2020, , .		0
8472	Trading Cost and Throughput in Geo-Distributed Analytics With A Two Time Scale Approach. IEEE Transactions on Cloud Computing, 2022, 10, 2163-2177.	3.1	3
8473	Five Years at the Edge: Watching Internet From the ISP Network. IEEE/ACM Transactions on Networking, 2020, 28, 561-574.	2.6	51
8474	JUST: JD Urban Spatio-Temporal Data Engine. , 2020, , .		38
8475	Asynchronous runtime with distributed manager for task-based programming models. Parallel Computing, 2020, 97, 102664.	1.3	2
8476	Simple Fault-tolerant Computing for Field Solvers. International Journal of Computational Fluid Dynamics, 2020, 34, 583-596.	0.5	1
8477	AcoustiCloud: A cloud-based system for managing large-scale bioacoustics processing. Environmental Modelling and Software, 2020, 131, 104778.	1.9	5
8478	DIGDUG: Scalable Separable Dense Graph Pruning and Join Operations in MapReduce. IEEE Transactions on Big Data, 2020, , 1-1.	4.4	2
8479	The Effects of Data Sampling with Deep Learning and Highly Imbalanced Big Data. Information Systems Frontiers, 2020, 22, 1113-1131.	4.1	50
8480	Online Multi-Resource Social Welfare Maximization for Non-Preemptive Jobs. IEEE Access, 2020, 8, 97920-97934.	2.6	2
8481	Shuffle Scheduling for MapReduce Jobs Based on Periodic Network Status. IEEE/ACM Transactions on Networking, 2020, 28, 1832-1844.	2.6	2

#	ARTICLE	IF	CITATIONS
8482	Latency Optimization for Coded Computation Straggled by Wireless Transmission. IEEE Wireless Communications Letters, 2020, , 1-1.	3.2	16
8483	Secure and Efficient <i>k</i> -NN Classification for Industrial Internet of Things. IEEE Internet of Things Journal, 2020, 7, 10945-10954.	5.5	29
8484	DCE-miner: an association rule mining algorithm for multimedia based on the MapReduce framework. Multimedia Tools and Applications, 2020, 79, 16771-16793.	2.6	8
8485	Data Flow Dependent Component Placement of Data Processing Cloud Applications. , 2020, , .		2
8486	Conditional MaxRS Query for Evolving Spatial Data. Frontiers in Big Data, 2020, 3, 20.	1.8	2
8487	TSpool: Transactions on a stream processor. Journal of Parallel and Distributed Computing, 2020, 140, 65-79.	2.7	10
8488	Shed+: Optimal Dynamic Speculation to Meet Application Deadlines in Cloud. IEEE Transactions on Network and Service Management, 2020, 17, 1515-1526.	3.2	6
8489	Improving Online Education Using Big Data Technologies. , 0, , .		7
8490	A holistic cross-layer optimization approach for mitigating stragglers in in-memory data processing. Journal of Systems Architecture, 2020, 111, 101801.	2.5	1
8491	PACC: Large scale connected component computation on Hadoop and Spark. PLoS ONE, 2020, 15, e0229936.	1.1	5
8492	Tails in the cloud: a survey and taxonomy of straggler management within large-scale cloud data centres. Journal of Supercomputing, 2020, 76, 10050-10089.	2.4	13
8493	Locator: A Cloud-Fog-Enabled Framework for Facilitating Efficient Location based Services. , 2020, , .		11
8494	An Extended Nonstrict Partially Ordered Set-Based Configurable Linear Sorter on FPGAs. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2020, 39, 1031-1044.	1.9	3
8495	Automatic parallelization of representative-based clustering algorithms for multicore cluster systems. International Journal of Data Science and Analytics, 2020, 10, 135-159.	2.4	3
8496	GasChecker: Scalable Analysis for Discovering Gas-Inefficient Smart Contracts. IEEE Transactions on Emerging Topics in Computing, 2021, 9, 1433-1448.	3.2	39
8497	Parallel Computational Structure and Semantics for Soil Quality Analysis Based on LoRa and Apache Spark. , 2020, , .		2
8498	Improving the performance of Hadoop MapReduce Applications via Optimization of concurrent containers per Node. , 2020, , .		7
8499	Hawkeye: Adaptive Straggler Identification on Heterogeneous Spark Cluster With Reinforcement Learning. IEEE Access, 2020, 8, 57822-57832.	2.6	7

#	ARTICLE	IF	CITATIONS
8500	An Event-Driven Approach to Serverless Seismic Imaging in the Cloud. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 2032-2049.	4.0	15
8501	A Blockchain-Powered Decentralized and Secure Computing Paradigm. IEEE Transactions on Emerging Topics in Computing, 2021, 9, 2201-2222.	3.2	21
8502	Integrating Sentiment Analysis on Hybrid Collaborative Filtering Method in a Big Data Environment. International Journal of Information Technology and Decision Making, 2020, 19, 385-412.	2.3	8
8503	Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence. Information (Switzerland), 2020, 11, 193.	1.7	205
8504	Distributed Evolutionary Hyperparameter Optimization for Fuzzy Time Series. IEEE Transactions on Network and Service Management, 2020, 17, 1309-1321.	3.2	17
8505	Automating the Configuration of MapReduce: A Reinforcement Learning Scheme. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50, 4183-4196.	5.9	7
8506	Cloud-Based RDF Data Management. Synthesis Lectures on Data Management, 2020, 15, 1-103.	0.6	1
8507	A Simple Semantic-Based Data Storage Layout for Querying Point Clouds. ISPRS International Journal of Geo-Information, 2020, 9, 72.	1.4	7
8508	Job failure prediction in Hadoop based on log file analysis. International Journal of Computers and Applications, 2022, 44, 260-269.	0.8	3
8509	The impact factors on the competence of big data processing. International Journal of Computers and Applications, 2020, , 1-18.	0.8	1
8510	Unleashing Constraint Optimisation Problem solving in Big Data environments. Journal of Computational Science, 2020, 45, 101180.	1.5	0
8511	Graph reconstruction in the congested clique. Journal of Computer and System Sciences, 2020, 113, 1-17.	0.9	4
8512	A Unified Framework for User Identification Across Online and Offline Data. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 1562-1575.	4.0	4
8513	CloudRaid: Detecting Distributed Concurrency Bugs via Log Mining and Enhancement. IEEE Transactions on Software Engineering, 2022, 48, 662-677.	4.3	5
8514	Verifying Properties of MapReduce-Based Big Data Processing. IEEE Transactions on Reliability, 2022, 71, 321-338.	3.5	4
8515	Mapreduce-Based Distributed Clustering Method Using CF+ Tree. IEEE Access, 2020, 8, 104232-104246.	2.6	2
8516	Alpaca: AR Graphics Extensions for Web Applications. , 2020, , .		0
8517	Distributed synthesized association mining for big transactional data. Sadhana - Academy Proceedings in Engineering Sciences, 2020, 45, 1.	0.8	2

#	ARTICLE	IF	CITATIONS
8518	Parallel and distributed architecture of genetic algorithm on Apache Hadoop and Spark. Applied Soft Computing Journal, 2020, 95, 106497.	4.1	24
8519	Communication-Efficient Edge AI: Algorithms and Systems. IEEE Communications Surveys and Tutorials, 2020, 22, 2167-2191.	24.8	200
8520	RHEEMix in the data jungle: a cost-based optimizer for cross-platform systems. VLDB Journal, 2020, 29, 1287-1310.	2.7	8
8521	Towards distributed node similarity search on graphs. World Wide Web, 2020, 23, 3025-3053.	2.7	4
8522	Fog-inspired smart home environment for domestic animal healthcare. Computer Communications, 2020, 160, 521-533.	3.1	13
8523	Extending BigTim Distributed Clustering Platform to Support Mobile Devices. , 2020, , .		0
8524	A smart data approach for Spatial Big Data analytics. , 2020, , .		1
8525	An analysis of two-way equi-join algorithms under MapReduce. Journal of King Saud University - Computer and Information Sciences, 2020, , .	2.7	3
8526	Revisiting the Feldâ€™s Friendship Paradox in Online Social Networks. IEEE Access, 2020, 8, 24062-24071.	2.6	0
8527	SparkRA: Enabling Big Data Scalability for the GATK RNA-seq Pipeline with Apache Spark. Genes, 2020, 11, 53.	1.0	7
8528	Adaptive cache pre-forwarding policy for distributed deep learning. Computers and Electrical Engineering, 2020, 82, 106558.	3.0	2
8529	A Pliable Index Coding Approach to Data Shuffling. IEEE Transactions on Information Theory, 2020, 66, 1333-1353.	1.5	16
8530	A Fast and Scalable Workflow for SNPs Detection in Genome Sequences Using Hadoop Map-Reduce. Genes, 2020, 11, 166.	1.0	7
8531	From digital to sustainable: A scientometric review of smart city literature between 1990 and 2019. Journal of Cleaner Production, 2020, 258, 120689.	4.6	133
8532	A survey of data partitioning and sampling methods to support big data analysis. Big Data Mining and Analytics, 2020, 3, 85-101.	7.5	152
8533	Modeling viral diffusion using quantum computational network simulation. Quantum Engineering, 2020, 2, e29.	1.2	25
8534	Random access with a distributed Bitmap Join Index for Star Joins. Heliyon, 2020, 6, e03342.	1.4	0
8535	Application and Performance Optimization of MapReduce Model in Image Segmentation. IEEE Access, 2020, 8, 31835-31844.	2.6	3

#	ARTICLE	IF	CITATIONS
8536	Low-Overhead Compressibility Prediction for High-Performance Lossless Data Compression. IEEE Access, 2020, 8, 37105-37123.	2.6	1
8537	Real-Time Optimization of Dynamic Speed Scaling for Distributed Data Centers. IEEE Transactions on Network Science and Engineering, 2020, 7, 2090-2103.	4.1	5
8538	A GRID-Based Spatial Interpolation Method as a Tool Supporting Real Estate Market Analyses. ISPRS International Journal of Geo-Information, 2020, 9, 39.	1.4	6
8539	ScienceEarth: A Big Data Platform for Remote Sensing Data Processing. Remote Sensing, 2020, 12, 607.	1.8	36
8540	Distributed Feature Extraction on Apache Spark for Human Action Recognition. , 2020, , .		2
8541	Knowledge and Data-Driven Mapping of Environmental Status Indicators from Remote Sensing and VGI. Remote Sensing, 2020, 12, 495.	1.8	9
8542	RDF 1.1: Knowledge Representation and Data Integration Language for the Web. Symmetry, 2020, 12, 84.	1.1	18
8543	Survey on memory management techniques in heterogeneous computing systems. IET Computers and Digital Techniques, 2020, 14, 47-60.	0.9	11
8544	Target Classification and Recognition for High-Resolution Remote Sensing Images: Using the Parallel Cross-Model Neural Cognitive Computing Algorithm. IEEE Geoscience and Remote Sensing Magazine, 2020, 8, 50-62.	4.9	8
8545	A Scalable Smartwatch-Based Medication Intake Detection System Using Distributed Machine Learning. Journal of Medical Systems, 2020, 44, 76.	2.2	28
8546	High-performance attribute reduction on graphics processing unit. Journal of Experimental and Theoretical Artificial Intelligence, 2020, 32, 977-996.	1.8	2
8547	Why big data, and what it is: basics to Advanced big data journey for the medical industry. , 2020, , 221-249.		0
8548	A parallel computing framework for solving user equilibrium problem on computer clusters. Transportmetrica A: Transport Science, 2020, 16, 550-573.	1.3	7
8549	Distributed parallel algorithms for online virtual network embedding applications. International Journal of Communication Systems, 2023, 36, e4325.	1.6	9
8550	A low-power task scheduling algorithm for heterogeneous cloud computing. Journal of Supercomputing, 2020, 76, 7290-7314.	2.4	11
8551	Resource-Constrained Replication Strategies for Hierarchical and Heterogeneous Tasks. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 793-804.	4.0	10
8552	Financial Big Data Analysis and Early Warning Platform: A Case Study. IEEE Access, 2020, 8, 36515-36526.	2.6	12
8553	Predicting Hadoop misconfigurations using machine learning. Software - Practice and Experience, 2020, 50, 1168-1183.	2.5	1

#	ARTICLE	IF	CITATIONS
8554	MERIT: Tensor Transform for Memory-Efficient Vision Processing on Parallel Architectures. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2020, 28, 791-804.	2.1	4
8556	Designing a MapReduce performance model in distributed heterogeneous platforms based on benchmarking approach. Journal of Supercomputing, 2020, 76, 7177-7203.	2.4	7
8557	Deep learning techniques for optimizing medical big data. , 2020, , 187-211.		8
8558	Requirements for Big Data Adoption for Railway Asset Management. IEEE Access, 2020, 8, 15543-15564.	2.6	28
8559	Generalized communication cost efficient multi-way spatial join: revisiting the curse of the last reducer. Geoinformatica, 2020, 24, 557-589.	2.0	3
8560	Fast scalable construction of ([compressed] static minimal perfect hash) functions. Information and Computation, 2020, 273, 104517.	0.5	4
8561	SOOM: Sort-Based Optimizer for Big Data Multi-Query. Big Data, 2020, 8, 38-61.	2.1	4
8562	Recovery of original individual person data (IPD) inferences from empirical IPD summaries only: Applications to distributed computing under disclosure constraints. Statistics in Medicine, 2020, 39, 1183-1198.	0.8	12
8563	Data linking over RDF knowledge graphs: A survey. Concurrency Computation Practice and Experience, 2020, 32, e5746.	1.4	6
8564	Large-scale graph processing systems: a survey. Frontiers of Information Technology and Electronic Engineering, 2020, 21, 384-404.	1.5	12
8565	An Analysis of Distributed Programming Models and Frameworks for Large-scale Graph Processing. IETE Journal of Research, 2020, , 1-9.	1.8	3
8566	Recognizing Indonesian Acronym and Expansion Pairs with Supervised Learning and MapReduce. Information (Switzerland), 2020, 11, 210.	1.7	3
8567	Toward a domain-specific language for scientific workflow-based applications on multicloud system. Concurrency Computation Practice and Experience, 2020, 33, e5802.	1.4	5
8568	Novel data-placement scheme for improving the data locality of Hadoop in heterogeneous environments. Concurrency Computation Practice and Experience, 2021, 33, e5752.	1.4	6
8569	A Bayesian perspective of statistical machine learning for big data. Computational Statistics, 2020, 35, 893-930.	0.8	15
8570	In-Memory Cache and Intra-Node Combiner Approaches for Optimizing Execution Time in High-Performance Computing. SN Computer Science, 2020, 1, 1.	2.3	2
8571	Machine learning based solutions for security of Internet of Things (IoT): A survey. Journal of Network and Computer Applications, 2020, 161, 102630.	5.8	266
8572	Streaming Smart Meter Data Integration to Enable Dynamic Demand Assignment for Real-Time Hydraulic Simulation. Journal of Water Resources Planning and Management - ASCE, 2020, 146, .	1.3	18

#	ARTICLE	IF	CITATIONS
8573	Offloading Online MapReduce tasks with Stateful Programmable Data Planes. , 2020, , .		3
8574	A-ECN Minimizing Queue Length for Datacenter Networks. IEEE Access, 2020, 8, 49100-49111.	2.6	4
8575	Improving Flow Scheduling Scheme With Mix-Traffic in Multi-Tenant Data Centers. IEEE Access, 2020, 8, 64666-64677.	2.6	2
8576	Scalable Data Storage Design for Nonstationary IoT Environment With Adaptive Security and Reliability. IEEE Internet of Things Journal, 2020, 7, 10171-10188.	5.5	42
8577	Distributed Soft Clustering Algorithm for IoT Based on Finite Time Average Consensus. IEEE Internet of Things Journal, 2021, 8, 16096-16107.	5.5	4
8578	Optimal Server Selection for Straggler Mitigation. IEEE/ACM Transactions on Networking, 2020, 28, 709-721.	2.6	14
8579	Context Deep Neural Network Model for Predicting Depression Risk Using Multiple Regression. IEEE Access, 2020, 8, 18171-18181.	2.6	58
8580	Scedar: A scalable Python package for single-cell RNA-seq exploratory data analysis. PLoS Computational Biology, 2020, 16, e1007794.	1.5	9
8581	A parallel fuzzy rule-base based decision tree in the framework of map-reduce. Pattern Recognition, 2020, 103, 107326.	5.1	12
8582	A debugging approach for live Big Data applications. Science of Computer Programming, 2020, 194, 102460.	1.5	5
8583	Multiswarm Artificial Bee Colony algorithm based on spark cloud computing platform for medical image registration. Computer Methods and Programs in Biomedicine, 2020, 192, 105432.	2.6	14
8584	LSQ: Load Balancing in Large-Scale Heterogeneous Systems With Multiple Dispatchers. IEEE/ACM Transactions on Networking, 2020, 28, 1186-1198.	2.6	18
8585	Cross-Cloud MapReduce for Big Data. IEEE Transactions on Cloud Computing, 2020, 8, 375-386.	3.1	25
8586	Splitting Large Medical Data Sets Based on Normal Distribution in Cloud Environment. IEEE Transactions on Cloud Computing, 2020, 8, 518-531.	3.1	26
8587	Enabling Cloud Applications to Negotiate Multiple Resources in a Cost-Efficient Manner. IEEE Transactions on Services Computing, 2021, 14, 413-425.	3.2	1
8588	Parallel computation of probabilistic skyline queries using MapReduce. Journal of Supercomputing, 2021, 77, 418-444.	2.4	4
8589	Cache-Based Multi-Query Optimization for Data-Intensive Scalable Computing Frameworks. Information Systems Frontiers, 2021, 23, 35-51.	4.1	8
8590	DiffPageRank: an efficient differential PageRank approach in MapReduce. Journal of Supercomputing, 2021, 77, 188-211.	2.4	3

#	ARTICLE	IF	CITATIONS
8591	TextBenDS: a Generic Textual Data Benchmark for Distributed Systems. Information Systems Frontiers, 2021, 23, 81-100.	4.1	9
8592	Investigating the performance of Hadoop and Spark platforms on machine learning algorithms. Journal of Supercomputing, 2021, 77, 1273-1300.	2.4	50
8593	Review on Oscillatory Stability in Power Grids With Renewable Energy Sources: Monitoring, Analysis, and Control Using Synchrophasor Technology. IEEE Transactions on Industrial Electronics, 2021, 68, 519-531.	5.2	54
8594	An efficient algorithm for identifying (â, “, d) motif from huge DNA datasets. Journal of Ambient Intelligence and Humanized Computing, 2021, 12, 485-495.	3.3	2
8595	A survey on novel classification of deduplication storage systems. Distributed and Parallel Databases, 2021, 39, 201-230.	1.0	7
8596	A Spark-based differential evolution with grouping topology model for large-scale global optimization. Cluster Computing, 2021, 24, 515-535.	3.5	16
8597	Decision support based on optimized data mining techniques: Application to mobile telecommunication companies. Concurrency Computation Practice and Experience, 2021, 33, .	1.4	1
8598	A Distributed Rough Evidential <i>K</i> -NN Classifier: Integrating Feature Reduction and Classification. IEEE Transactions on Fuzzy Systems, 2021, 29, 2322-2335.	6.5	17
8599	Pruning techniques for parallel processing of reverse top-k queries. Distributed and Parallel Databases, 2021, 39, 169-199.	1.0	0
8600	Location Planning of PEV Fast Charging Station: An Integrated Approach Under Traffic and Power Grid Requirements. IEEE Transactions on Intelligent Transportation Systems, 2021, 22, 483-492.	4.7	37
8601	A heuristic method towards deadline-aware energy-efficient mapreduce scheduling problem in Hadoop YARN. Cluster Computing, 2021, 24, 683-699.	3.5	16
8602	Trigger-Based Incremental Data Processing with Unified Sync and Async Model. IEEE Transactions on Cloud Computing, 2021, 9, 372-385.	3.1	5
8603	A Micro-Extended Belief Rule-Based System for Big Data Multiclass Classification Problems. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51, 420-440.	5.9	29
8604	Cloud Resource Scaling for Time-Bounded and Unbounded Big Data Streaming Applications. IEEE Transactions on Cloud Computing, 2021, 9, 504-517.	3.1	5
8605	New YARN Non-Exclusive Resource Management Scheme through Opportunistic Idle Resource Assignment. IEEE Transactions on Cloud Computing, 2021, 9, 696-709.	3.1	1
8606	Towards complex dynamic fog network orchestration using embedded neural switch. International Journal of Computers and Applications, 2021, 43, 91-108.	0.8	9
8607	D-SRTF: Distributed Shortest Remaining Time First Scheduling for Data Center Networks. IEEE Transactions on Cloud Computing, 2021, 9, 562-575.	3.1	13
8608	Providing Bandwidth Guarantees, Work Conservation and Low Latency Simultaneously in the Cloud. IEEE Transactions on Cloud Computing, 2021, 9, 763-776.	3.1	10

#	ARTICLE	IF	CITATIONS
8609	Hierarchical Density-Based Clustering Using MapReduce. IEEE Transactions on Big Data, 2021, 7, 102-114.	4.4	25
8610	Summarizing and linking electronic health records. Distributed and Parallel Databases, 2021, 39, 321-360.	1.0	3
8611	A parallel content-based image retrieval system using spark and tachyon frameworks. Journal of King Saud University - Computer and Information Sciences, 2021, 33, 141-149.	2.7	15
8612	Fast Communication-Efficient Spectral Clustering over Distributed Data. IEEE Transactions on Big Data, 2021, 7, 158-168.	4.4	6
8613	Scheduling Mix-Coflows in Datacenter Networks. IEEE Transactions on Network and Service Management, 2021, 18, 2002-2015.	3.2	6
8614	Joint coflow routing and scheduling in leaf-spine data centers. Journal of Parallel and Distributed Computing, 2021, 148, 83-95.	2.7	2
8615	Large-Scale High-Utility Sequential Pattern Analytics in Internet of Things. IEEE Internet of Things Journal, 2021, 8, 12669-12678.	5.5	31
8616	Optimal Load Allocation for Coded Distributed Computation in Heterogeneous Clusters. IEEE Transactions on Communications, 2021, 69, 44-58.	4.9	14
8617	Air quality monitoring and analysis with dynamic training using deep learning. Journal of Supercomputing, 2021, 77, 5586-5605.	2.4	3
8618	Big Data Driven Marine Environment Information Forecasting: A Time Series Prediction Network. IEEE Transactions on Fuzzy Systems, 2021, 29, 4-18.	6.5	117
8619	Chicken swarm foraging algorithm for big data classification using the deep belief network classifier. Data Technologies and Applications, 2021, 55, 332-352.	0.9	2
8620	A hybrid distributed batch-stream processing approach for anomaly detection. Information Sciences, 2021, 543, 309-327.	4.0	16
8621	Ignite-GPU: a GPU-enabled in-memory computing architecture on clusters. Journal of Supercomputing, 2021, 77, 3165-3192.	2.4	7
8622	Sparse evolutionary deep learning with over one million artificial neurons on commodity hardware. Neural Computing and Applications, 2021, 33, 2589-2604.	3.2	21
8623	Federated Self-Supervised Learning of Multisensor Representations for Embedded Intelligence. IEEE Internet of Things Journal, 2021, 8, 1030-1040.	5.5	55
8624	Sampling business process event logs using graph-based ranking model. Concurrency Computation Practice and Experience, 2021, 33, e5974.	1.4	15
8625	Process mining-constrained scheduling in the hybrid cloud. Concurrency Computation Practice and Experience, 2021, 33, e6025.	1.4	5
8626	SeisNoise.jl: Ambient Seismic Noise Cross Correlation on the CPU and GPU in Julia. Seismological Research Letters, 2021, 92, 517-527.	0.8	5

#	ARTICLE	IF	CITATIONS
8627	The acoustic inverse problem in the framework of alternating direction method of multipliers. <i>Mechanical Systems and Signal Processing</i> , 2021, 149, 107220.	4.4	20
8628	An efficient deadline constrained and data locality aware dynamic scheduling framework for multitenancy clouds. <i>Concurrency Computation Practice and Experience</i> , 2021, 33, e6037.	1.4	3
8629	Interval graph multi-coloring-based resource reservation for energy-efficient containerized cloud data centers. <i>Journal of Supercomputing</i> , 2021, 77, 4484-4532.	2.4	2
8630	E-Commerce data classification in the cloud environment based on bayesian algorithm. <i>Journal of Intelligent and Fuzzy Systems</i> , 2021, 40, 5819-5826.	0.8	1
8631	Mining userâ€“user communities for a weighted bipartite network using spark GraphFrames and Flink Gelly. <i>Journal of Supercomputing</i> , 2021, 77, 5984-6035.	2.4	5
8632	Multi-class imbalanced big data classification on Spark. <i>Knowledge-Based Systems</i> , 2021, 212, 106598.	4.0	51
8633	Multi-source information fusion based on rough set theory: A review. <i>Information Fusion</i> , 2021, 68, 85-117.	11.7	131
8634	Efficient algorithms for discovering high-utility patterns with strong frequency affinities. <i>Expert Systems With Applications</i> , 2021, 169, 114464.	4.4	8
8635	A real-time explainable traffic collision inference framework based on probabilistic graph theory. <i>Knowledge-Based Systems</i> , 2021, 212, 106442.	4.0	4
8636	Link maintenance for integrity in linked open data evolution: Literature survey and open challenges. <i>Semantic Web</i> , 2021, 12, 517-541.	1.1	1
8637	Engineering collective intelligence at the edge with aggregate processes. <i>Engineering Applications of Artificial Intelligence</i> , 2021, 97, 104081.	4.3	36
8638	Noninvasive MapReduce Performance Tuning Using Multiple Tuning Methods on Hadoop. <i>IEEE Systems Journal</i> , 2021, 15, 2906-2917.	2.9	4
8639	Fuzzy-Based Kernelized Clustering Algorithms for Handling Big Data Using Apache Spark. <i>Advances in Intelligent Systems and Computing</i> , 2021, , 423-435.	0.5	3
8640	Parallelization of network motif discovery using star contraction. <i>Parallel Computing</i> , 2021, 101, 102734.	1.3	3
8641	Hybrid KNN-join: Parallel nearest neighbor searches exploiting CPU and GPU architectural features. <i>Journal of Parallel and Distributed Computing</i> , 2021, 149, 119-137.	2.7	8
8642	A fast parallel attribute reduction algorithm using Apache Spark. <i>Knowledge-Based Systems</i> , 2021, 212, 106582.	4.0	9
8643	SPO: A Secure and Performance-aware Optimization for MapReduce Scheduling. <i>Journal of Network and Computer Applications</i> , 2021, 176, 102944.	5.8	11
8644	Cyber intrusion detection through association rule mining on multi-source logs. <i>Applied Intelligence</i> , 2021, 51, 4043-4057.	3.3	14

#	ARTICLE	IF	CITATIONS
8645	Embedded Index Coding. IEEE Transactions on Information Theory, 2021, 67, 1461-1477.	1.5	6
8646	Profiles of Upcoming HPC Applications and Their Impact on Reservation Strategies. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 1178-1190.	4.0	4
8647	Flow Scheduling of Service Chain Processing in a NFV-Based Network. IEEE Transactions on Network Science and Engineering, 2021, 8, 389-399.	4.1	7
8648	Scheduling coflows of multi-stage jobs under network resource constraints. Computer Networks, 2021, 184, 107686.	3.2	5
8649	Breaking the linear-memory barrier in MPC: Fast MIS on trees with strongly sublinear memory. Theoretical Computer Science, 2021, 849, 22-34.	0.5	0
8650	Unified programming concepts for unobtrusive integration of cloud-based and local parallel computing. Future Generation Computer Systems, 2021, 115, 700-719.	4.9	4
8651	Rendezvous algorithms for large-scale modeling and simulation. Journal of Parallel and Distributed Computing, 2021, 147, 184-195.	2.7	3
8652	Global Consensus Monte Carlo. Journal of Computational and Graphical Statistics, 2021, 30, 249-259.	0.9	10
8653	Notions of architecture in fog computing. Computing (Vienna/New York), 2021, 103, 51-73.	3.2	22
8654	Evaluating the effects of access control policies within NoSQL systems. Future Generation Computer Systems, 2021, 114, 491-505.	4.9	4
8655	Towards Greening MapReduce Clusters Considering Both Computation Energy and Cooling Energy. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 931-942.	4.0	2
8656	CLONE: Collaborative Learning on the Edges. IEEE Internet of Things Journal, 2021, 8, 10222-10236.	5.5	7
8657	Optimizing the Response Time of Memcached Systems via Model and Quantitative Analysis. IEEE Transactions on Computers, 2021, 70, 1458-1471.	2.4	2
8658	SMOaaS: a Scalable Matrix Operation as a Service model in Cloud. Journal of Supercomputing, 2021, 77, 3381-3401.	2.4	0
8659	Platonica: an efficient and high-performance dual-centric data center network architecture. Cluster Computing, 2021, 24, 997-1032.	3.5	3
8660	TVD-MRDL: traffic violation detection system using MapReduce-based deep learning for large-scale data. Multimedia Tools and Applications, 2021, 80, 2489-2516.	2.6	10
8661	Accelerating ELM training over data streams. International Journal of Machine Learning and Cybernetics, 2021, 12, 87-102.	2.3	0
8662	PSpec-SQL: Enabling Fine-Grained Control for Distributed Data Analytics. IEEE Transactions on Dependable and Secure Computing, 2021, 18, 810-824.	3.7	0

#	ARTICLE	IF	CITATIONS
8663	On communication efficient dataflow computing in software defined networking enabled cloud. Concurrency Computation Practice and Experience, 2021, 33, 1-1.	1.4	13
8664	Integrated process planning and scheduling in networked manufacturing systems for I4.0: a review and framework proposal. Wireless Networks, 2021, 27, 1587-1599.	2.0	34
8665	SegGuard: Segmentation-based Anonymization of Network Data in Clouds for Privacy-Preserving Security Auditing. IEEE Transactions on Dependable and Secure Computing, 2021, , 1-1.	3.7	3
8666	On Parallel Real-Time Security Improvement Using Mixed-Integer Programming. IEEE Access, 2021, 9, 58824-58837.	2.6	1
8667	Online Learning-Based Co-task Dispatching with Function Configuration in Edge Computing. Lecture Notes in Computer Science, 2021, , 198-209.	1.0	1
8668	Change Detection in Smart Grids Using Dynamic Mixtures of t-Distributions. Lecture Notes in Mechanical Engineering, 2021, , 53-68.	0.3	0
8669	Proposed big data architecture for facial recognition using machine learning. AIMS Electronics and Electrical Engineering, 2021, 5, 68-92.	0.8	8
8670	Run-Time Dynamic Resource Adjustment for Mitigating Skew in MapReduce. CMES - Computer Modeling in Engineering and Sciences, 2021, 126, 771-790.	0.8	3
8671	Novel Approach for Mining Patterns. International Journal of Applied Evolutionary Computation, 2021, 12, 27-42.	0.7	1
8672	Collaborative Filtering Recommendation Using Nonnegative Matrix Factorization in GPU-Accelerated Spark Platform. Scientific Programming, 2021, 2021, 1-15.	0.5	6
8673	A Review of Anomalies Detection Based on Association Rules Techniques. Lecture Notes in Networks and Systems, 2021, , 1155-1166.	0.5	0
8674	Big Data and Healthcare Data. , 2021, , 1914-1944.		0
8675	Scalable Privacy-Preserving Distributed Learning. Proceedings on Privacy Enhancing Technologies, 2021, 2021, 323-347.	2.3	23
8676	An Efficient Rolling Bearing Fault Diagnosis Method Based on Spark and Improved Random Forest Algorithm. IEEE Access, 2021, 9, 37866-37882.	2.6	49
8678	FLCD: A Flexible Low Complexity Design of Coded Distributed Computing. IEEE Transactions on Cloud Computing, 2023, 11, 470-483.	3.1	2
8679	Trends and Technologies in Big Data Processing. , 2021, , 42-67.		0
8680	Security and Privacy Challenges in Big Data Environment. , 2020, , 89-99.		1
8682	Big Data-Based Structural Health Monitoring of Concrete structures”A Perspective Review. Lecture Notes in Mechanical Engineering, 2021, , 529-540.	0.3	3

#	ARTICLE	IF	CITATIONS
8683	Private and rateless adaptive coded matrix-vector multiplication. <i>Eurasip Journal on Wireless Communications and Networking</i> , 2021, 2021, .	1.5	10
8684	Fast SQL/Row Pattern Recognition Query Processing Using Parallel Primitives on GPUs. <i>Lecture Notes in Computer Science</i> , 2021, , 22-34.	1.0	0
8685	mbkmeans: Fast clustering for single cell data using mini-batch k-means. <i>PLoS Computational Biology</i> , 2021, 17, e1008625.	1.5	36
8686	Federated Learning With Heterogeneity-Aware Probabilistic Synchronous Parallel on Edge. <i>IEEE Transactions on Services Computing</i> , 2022, 15, 614-626.	3.2	5
8687	Applying Patterns to Support Deployment in Cloud-Edge Environments: A Case Study. <i>Lecture Notes in Networks and Systems</i> , 2021, , 139-148.	0.5	1
8688	Differentially Oblivious Data Analysis With Intel SGX: Design, Optimization, and Evaluation. <i>IEEE Transactions on Dependable and Secure Computing</i> , 2022, 19, 3741-3758.	3.7	4
8689	An application layer technique to overcome TCP incast in data center network using delayed server response. <i>International Journal of Information Technology (Singapore)</i> , 2021, 13, 703-711.	1.8	1
8690	Variable Expanding Structure for Data Center Interconnection Networks. <i>Journal of Advanced Computational Intelligence and Intelligent Informatics</i> , 2021, 25, 13-22.	0.5	1
8691	An Application of Automatic Text Revision for Power Defect Log. <i>Journal of Physics: Conference Series</i> , 2021, 1757, 012027.	0.3	0
8692	Frequent Subgraph Mining Algorithms in Static and Temporal Graph-Transaction Settings: A Survey. <i>IEEE Transactions on Big Data</i> , 2021, , 1-1.	4.4	4
8693	Epidemic Monitoring and Forecasting System Based on Big Data. <i>Computer Science and Application</i> , 2021, 11, 867-875.	0.0	0
8694	Performance Comparison for Data Retrieval from NoSQL and SQL Databases: A Case Study for COVID-19 Genome Sequence Dataset. , 2021, , .		6
8695	Big Data Resource Management & Networks: Taxonomy, Survey, and Future Directions. <i>IEEE Communications Surveys and Tutorials</i> , 2021, 23, 2098-2130.	24.8	33
8696	A Dynamic Ridesplitting Method With Potential Pick-Up Probability Based on GPS Trajectories. <i>IEEE Transactions on Intelligent Transportation Systems</i> , 2022, 23, 10786-10802.	4.7	2
8697	ScaleG: A Distributed Disk-based System for Vertex-centric Graph Processing. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2021, , 1-1.	4.0	6
8698	Distributed Mining of High Utility Time Interval Sequential Patterns with Multiple Minimum Utility Thresholds. <i>Lecture Notes in Computer Science</i> , 2021, , 86-97.	1.0	2
8699	Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities. <i>Multimedia Systems</i> , 2022, 28, 1339-1371.	3.0	53
8701	From WHOIS to WHOWAS: A Large-Scale Measurement Study of Domain Registration Privacy under the GDPR. , 2021, , .		10

#	ARTICLE	IF	CITATIONS
8702	Stochastic Gradient-Based Distributed Bayesian Estimation in Cooperative Sensor Networks. IEEE Transactions on Signal Processing, 2021, 69, 1713-1724.	3.2	4
8703	Big data platforms: in the lens of selection and evaluation approach. Journal of Decision Systems, 2022, 32, 19-48.	2.2	4
8704	Loupe: A Visualization Tool for High-Level Execution Plans in SystemDS. Lecture Notes in Computer Science, 2021, , 627-630.	1.0	0
8705	Distributed Mining of High Utility Sequential Patterns with Negative Item Values. International Journal of Advanced Computer Science and Applications, 2021, 12, .	0.5	1
8706	Big Data Clustering Techniques: Recent Advances and Survey. , 2021, , 57-79.		3
8707	An overview of healthcare data analytics with applications to the COVID-19 pandemic. IEEE Transactions on Big Data, 2021, , 1-1.	4.4	10
8708	Clustered Multicast Source Routing for Large-Scale Cloud Data Centers. IEEE Access, 2021, 9, 12693-12705.	2.6	4
8709	Big-Parallel-ETL: New ETL for Multidimensional NoSQL Graph Oriented Data. Journal of Physics: Conference Series, 2021, 1743, 012037.	0.3	3
8710	High-Performance Computation in Residue Number System Using Floating-Point Arithmetic. Computation, 2021, 9, 9.	1.0	10
8711	An Easy-to-Use Execution Environment for the Parallelisation of Computationally Intensive Data Science Applications. , 2021, , 16-23.		0
8712	Research Trends in Social Media/Big Data with the Emphasis on Data Collection and Data Management: A Bibliometric Analysis. Human Dynamics in Smart Cities, 2021, , 47-63.	0.2	4
8713	MespaConfig: Memory-Sparing Configuration Auto-Tuning for Co-Located In-Memory Cluster Computing Jobs. IEEE Transactions on Services Computing, 2022, 15, 2883-2896.	3.2	5
8714	Big Data Management in Smart Grids: Technologies and Challenges. IEEE Access, 2021, 9, 73046-73059.	2.6	20
8715	<i>LogM</i>: Log Analysis for Multiple Components of Hadoop Platform. IEEE Access, 2021, 9, 73522-73532.	2.6	10
8716	On Benchmarking for Concurrent Runtime Verification. Lecture Notes in Computer Science, 2021, , 3-23.	1.0	8
8717	Distributed Machine Learning for Wireless Communication Networks: Techniques, Architectures, and Applications. IEEE Communications Surveys and Tutorials, 2021, 23, 1458-1493.	24.8	53
8718	A Flexible Distributed Stochastic Optimization Framework for Concurrent Tasks in Processing Networks. IEEE/ACM Transactions on Networking, 2021, , 1-14.	2.6	0
8719	Cooperative Co-Evolution and MapReduce. International Journal of Information Technology Project Management, 2021, 12, 29-62.	0.3	5

#	ARTICLE	IF	CITATIONS
8720	Maximizing Last-Minute Backup in Endangered Time-Varying Inter-Datacenter Networks. IEEE/ACM Transactions on Networking, 2021, 29, 2646-2663.	2.6	2
8721	A scalable medication intake monitoring system. , 2021, , 217-240.		0
8722	Advances in MapReduce Big Data Processing: Platform, Tools, and Algorithms. Studies in Big Data, 2021, , 105-128.	0.8	9
8723	Spark Based Text Clustering Method Using Hashing. Lecture Notes in Computer Science, 2021, , 137-142.	1.0	1
8724	Cloud Computing Education Strategies. , 2021, , 2519-2530.		0
8725	A Comprehensive Study on Internet of Things Based on Key Artificial Intelligence Technologies and Industry 4.0. , 2021, , 171-191.		1
8726	Dynamic Capabilities of Decision-oriented Service Systems. , 2021, , 240-266.		0
8727	Dynamic Resource Provisioning for Iterative Workloads on Apache Spark. IEEE Transactions on Cloud Computing, 2023, 11, 639-652.	3.1	5
8728	Big Data analytics in medical imaging. , 2021, , 113-136.		0
8729	An Experimental Model of the Collect/Report Paradigm. IOP Conference Series: Materials Science and Engineering, 2021, 1031, 012073.	0.3	0
8730	A Data-Aware Scheduling Strategy for Executing Large-Scale Distributed Workflows. IEEE Access, 2021, 9, 47354-47364.	2.6	3
8731	Performance Enhancement of Hadoop MapReduce by Combining Data Inside the Mapper. , 2021, , .		1
8732	Factors Accountable for Diabetes Using Artificial Intelligence in Medico-Care. Advances in Intelligent Systems and Computing, 2021, , 43-53.	0.5	0
8733	A study on Medical Device Security Status in Medical Convergence Industry. , 2021, , .		0
8734	Simple Fault-Tolerant Computing for CFD Codes. , 2021, , .		0
8735	Deployment Service for Scalable Distributed Deep Learning Training on Multiple Clouds. , 2021, , .		0
8736	The Design and Implementation of KMeans Based on Unified Batch and Streaming Processing. Communications in Computer and Information Science, 2021, , 639-649.	0.4	0
8737	Speed-Robust Scheduling. Lecture Notes in Computer Science, 2021, , 283-296.	1.0	1

#	ARTICLE	IF	CITATIONS
8738	A Service-based Joint Model Used for Distributed Learning: Application for Smart Agriculture. IEEE Transactions on Emerging Topics in Computing, 2021, , 1-1.	3.2	16
8739	NileOS: A Distributed Asymmetric Core-Based Micro-Kernel for Big Data Processing. IEEE Access, 2021, 9, 3696-3711.	2.6	0
8740	Large Scale Graph Based Network Forensics Analysis. Lecture Notes in Computer Science, 2021, , 457-469.	1.0	0
8741	Towards HPC and Big Data Analytics Convergence: Design and Experimental Evaluation of a HPDA Framework for eScience at Scale. IEEE Access, 2021, 9, 73307-73326.	2.6	10
8742	Rolling Bearing Fault Diagnosis Method Based on Parallel QPSO-BPNN Under Spark-GPU Platform. IEEE Access, 2021, 9, 56786-56801.	2.6	6
8743	Modeling and Analyzing Resource-Sensitive Actors: A Tutorial Introduction. Lecture Notes in Computer Science, 2021, , 3-19.	1.0	5
8745	Leveraging User-Diversity in Energy-Efficient Edge-Facilitated Collaborative Fog Computing. IEEE Access, 2021, 9, 95636-95650.	2.6	1
8746	Spark Accelerated Implementation of Parallel Attribute Reduction from Incomplete Data. Lecture Notes in Computer Science, 2021, , 203-217.	1.0	2
8747	A Modular Remote Sensing Big Data Framework. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60, 1-11.	2.7	4
8748	Denial-of-Service and Botnet Analysis, Detection, and Mitigation. , 2021, , 20-48.		1
8749	Scalable I-Diversity. , 2021, , 1051-1065.		0
8750	Distributed Top-K Join Queries Optimizing for RDF Datasets. , 2021, , 1895-1913.		0
8751	Data Science and Big Data Practice Using Apache Spark and Python. Advances in Data Mining and Database Management Book Series, 2021, , 67-95.	0.4	0
8752	Dynamic Capabilities of Decision-oriented Service Systems. , 2021, , 957-984.		0
8753	A Comprehensive Review of Access Control Mechanism Based on Attribute Based Encryption Scheme for Cloud Computing. , 2021, , 899-921.		0
8755	FasTensor Programming Model. SpringerBriefs in Computer Science, 2021, , 9-22.	0.2	0
8756	Crime Forecasting Using Time Series Analysis. Asset Analytics, 2021, , 253-262.	0.4	0
8758	IBM PAIRS: Scalable Big Geospatial-Temporal Data and Analytics As-a-Service. , 2021, , 3-34.		4

#	ARTICLE	IF	CITATIONS
8759	AEML: An Acceleration Engine for Multi-GPU Load-balancing in Distributed Heterogeneous Environment. IEEE Transactions on Computers, 2021, , 1-1.	2.4	6
8760	Big Data Analysis in Bioinformatics. , 2021, , 405-429.		0
8761	Trade-Off Analysis Between Parallelism and Accuracy of SLIC on Apache Spark. , 2021, , .		2
8762	Security by Design for Big Data Frameworks Over Cloud Computing. IEEE Transactions on Engineering Management, 2022, 69, 3676-3693.	2.4	44
8763	SDPSO: Spark Distributed PSO-based approach for feature selection and cancer disease prognosis. Journal of Big Data, 2021, 8, .	6.9	10
8764	End-to-End Data Pipeline in Games for Real-Time Data Analytics. Lecture Notes on Data Engineering and Communications Technologies, 2021, , 269-275.	0.5	1
8765	Data Encoding for Byzantine-Resilient Distributed Optimization. IEEE Transactions on Information Theory, 2021, 67, 1117-1140.	1.5	11
8766	Content analytics based on random forest classification technique: An empirical evaluation using online news dataset. International Journal of Advanced and Applied Sciences, 2021, 8, 77-84.	0.2	3
8767	An Overview of End-to-End Entity Resolution for Big Data. ACM Computing Surveys, 2021, 53, 1-42.	16.1	99
8768	A Bottom-Up Tree Based Storage Approach for Efficient IoT Data Analytics in Cloud Systems. Journal of Grid Computing, 2021, 19, 1.	2.5	8
8769	MapReduce based parallel attribute reduction in Incomplete Decision Systems. Knowledge-Based Systems, 2021, 213, 106677.	4.0	7
8770	Dragoon: a hybrid and efficient big trajectory management system for offline and online analytics. VLDB Journal, 2021, 30, 287-310.	2.7	14
8771	Building and Operating a Large-Scale Enterprise Data Analytics Platform. Big Data Research, 2021, 23, 100181.	2.6	11
8772	A survey on bandwidth-aware geo-distributed frameworks for big-data analytics. Journal of Big Data, 2021, 8, .	6.9	7
8774	Internet of Everything. , 2021, , 21-56.		2
8775	Historical data based approach for straggler avoidance in a heterogeneous Hadoop cluster. Journal of Ambient Intelligence and Humanized Computing, 2021, 12, 9573-9589.	3.3	5
8776	Joint Reducer Placement and Coflow Bandwidth Scheduling for Computing Clusters. IEEE/ACM Transactions on Networking, 2021, 29, 438-451.	2.6	7
8777	Providing Consistent State to Distributed Storage System. Computers, 2021, 10, 23.	2.1	3

#	ARTICLE	IF	CITATIONS
8778	Distributed mining of convoys in large scale datasets. <i>Geoinformatica</i> , 2021, 25, 353-396.	2.0	4
8780	The vertex-pancyclicity of data center networks. <i>Theoretical Computer Science</i> , 2021, 855, 74-89.	0.5	5
8781	Distributed anomaly detection using concept drift detection based hybrid ensemble techniques in streamed network data. <i>Cluster Computing</i> , 2021, 24, 2099-2114.	3.5	19
8782	Unsupervised detection of botnet activities using frequent pattern tree mining. <i>Complex & Intelligent Systems</i> , 2022, 8, 761-769.	4.0	4
8783	An efficient apriori algorithm for frequent pattern mining using mapreduce in healthcare data. <i>Bulletin of Electrical Engineering and Informatics</i> , 2021, 10, 390-403.	0.6	19
8784	Modernizing parallel code with pattern analysis. , 2021, , .		1
8785	ShadowVM. , 2021, , .		0
8786	kNN-join Query Processing Algorithm on Mapreduce for Large Amounts of Data. , 2021, , .		1
8787	Hadoop Deep Neural Network for offending drivers. <i>Journal of Ambient Intelligence and Humanized Computing</i> , 2022, 13, 659-671.	3.3	6
8788	A heterogeneous key performance indicator metadata model for air quality monitoring in sustainable cities. <i>Environmental Modelling and Software</i> , 2021, 136, 104955.	1.9	2
8789	Research on Hybrid Index Method of Double-Layer B+ Tree for Power Big Data Considering Knowledge Graph. <i>Journal of Physics: Conference Series</i> , 2021, 1771, 012004.	0.3	1
8791	Investigating a Streaming Analytics Framework for Data Analytics Applications in the Aircraft Cabin. , 0, , .		0
8792	Stargate. , 2021, , .		0
8793	Movie Recommendation through Multiple Bias Analysis. <i>Applied Sciences (Switzerland)</i> , 2021, 11, 2817.	1.3	1
8794	FASTA/Q data compressors for MapReduce-Hadoop genomics: space and time savings made easy. <i>BMC Bioinformatics</i> , 2021, 22, 144.	1.2	6
8795	Intelligent Integrated Community Model for Ethnic Minority Areas Using Data Analysis and Machine Learning. <i>Mobile Information Systems</i> , 2021, 2021, 1-8.	0.4	3
8796	Coded Wireless Distributed Matrix Multiplication with Flexible Assignment of Output Vectors. , 2021, , .		0
8797	Accelerated execution via eager-release of dependencies in task-based workflows. <i>International Journal of High Performance Computing Applications</i> , 2021, 35, 325-343.	2.4	3

#	ARTICLE	IF	CITATIONS
8798	Multi-objective automated guided vehicle scheduling based on MapReduce framework. <i>Advances in Production Engineering and Management</i> , 2021, 16, 37-46.	0.8	7
8799	An Optimization of Bitmap Index Compression Technique in Bulk Data Movement Infrastructure. <i>IOP Conference Series: Materials Science and Engineering</i> , 2021, 1099, 012074.	0.3	0
8800	Conception and Software Implementation of a Nuclear Data Evaluation Pipeline. <i>Nuclear Data Sheets</i> , 2021, 173, 239-284.	0.7	11
8801	BSF: A parallel computation model for scalability estimation of iterative numerical algorithms on cluster computing systems. <i>Journal of Parallel and Distributed Computing</i> , 2021, 149, 193-206.	2.7	12
8802	Scalable, High-Performance, and Generalized Subtree Data Anonymization Approach for Apache Spark. <i>Electronics (Switzerland)</i> , 2021, 10, 589.	1.8	12
8803	The characters of teaching physical Qigong in coastal area based on geological management information system. <i>Arabian Journal of Geosciences</i> , 2021, 14, 1.	0.6	1
8805	Towards efficient canonical polyadic decomposition on sunway many-core processor. <i>Information Sciences</i> , 2021, 549, 221-248.	4.0	2
8806	Feature selection using cloud-based parallel genetic algorithm for intrusion detection data classification. <i>Neural Computing and Applications</i> , 2021, 33, 11861-11873.	3.2	15
8807	Review on ranking and selection: A new perspective. <i>Frontiers of Engineering Management</i> , 2021, 8, 321-343.	3.3	35
8808	Fuzzy based decision making approach for big data research on health information management system. <i>Journal of Ambient Intelligence and Humanized Computing</i> , 2021, 12, 3363-3371.	3.3	0
8809	Linking place records using multi-view encoders. <i>Neural Computing and Applications</i> , 2021, 33, 12103-12119.	3.2	8
8810	Aggregate centrality measures for IoT-based coordination. <i>Science of Computer Programming</i> , 2021, 203, 102584.	1.5	3
8811	Relationship between land use carbon emission and economic growth based on GIS. <i>Arabian Journal of Geosciences</i> , 2021, 14, 1.	0.6	1
8812	In Big Data Era: Analysis of Hadoop Cluster Performance. , 2021, , .		0
8813	Implicitly distributed fuzzy random forests. , 2021, , .		1
8814	Dynamic evaluation of logistics enterprise competitiveness based on machine learning and improved neural network. <i>Journal of Ambient Intelligence and Humanized Computing</i> , 0, , 1.	3.3	4
8815	The Four Generations of Entity Resolution. <i>Synthesis Lectures on Data Management</i> , 2021, 16, 1-170.	0.6	13
8816	Application Research of Big Data Real-time Processing Technology in Smart Grid. <i>IOP Conference Series: Earth and Environmental Science</i> , 2021, 696, 012044.	0.2	3

#	ARTICLE	IF	CITATIONS
8817	Processes meet Big Data: Scaling process discovery algorithms in Big Data environment. Journal of King Saud University - Computer and Information Sciences, 2022, 34, 8478-8489.	2.7	2
8818	Have We Reached Consensus? An Analysis of Distributed Systems Syllabi. , 2021, , .		2
8819	Fly-path: Traffic-based multi-hop routing approach for hybrid wireless data centers. Computer Communications, 2021, 170, 95-108.	3.1	0
8820	Design and Evaluation of a New Machine Learning Framework for IoT and Embedded Devices. Electronics (Switzerland), 2021, 10, 600.	1.8	11
8821	Anomaly Detection for Time Series Data Stream. , 2021, , .		4
8822	Machine-Learning Based Memory Prediction Model for Data Parallel Workloads in Apache Spark. Symmetry, 2021, 13, 697.	1.1	0
8823	Energy and SLA-driven MapReduce Job Scheduling Framework for Cloud-based Cyber-Physical Systems. ACM Transactions on Internet Technology, 2021, 21, 1-24.	3.0	1
8824	Open-source multi-purpose remote laboratory for IoT education. , 2021, , .		9
8825	SEIZE: Runtime Inspection for Parallel Dataflow Systems. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 842-854.	4.0	1
8826	Potential Deep Learning Solutions to Persistent and Emerging Big Data Challengesâ€™A Practitionersâ€™ Cookbook. ACM Computing Surveys, 2021, 54, 1-39.	16.1	1
8827	MODELING OF SYSTEMS UNDER CLOUD ENVIRONMENT. ASEAN Engineering Journal, 2021, 11, 190-198.	0.2	0
8828	Understanding join strategies in distributed systems. , 2021, , .		0
8829	New Results on the Computation-Communication Tradeoff for Heterogeneous Coded Distributed Computing. IEEE Transactions on Communications, 2021, 69, 2254-2270.	4.9	11
8830	Multi-Layer Data Model of a Complex Transportation Network. International Journal of Embedded and Real-Time Communication Systems, 2021, 12, 21-36.	0.3	0
8831	MapReduce based parallel fuzzy-rough attribute reduction using discernibility matrix. Applied Intelligence, 2022, 52, 154-173.	3.3	11
8832	Single-Cell Transcriptomics: Current Methods and Challenges in Data Acquisition and Analysis. Frontiers in Neuroscience, 2021, 15, 591122.	1.4	53
8833	A Survey on Subgraph Counting. ACM Computing Surveys, 2022, 54, 1-36.	16.1	43
8835	A Survey on Distributed Graph Pattern Matching in Massive Graphs. ACM Computing Surveys, 2022, 54, 1-35.	16.1	17

#	ARTICLE	IF	CITATIONS
8836	A parallel text clustering method using Spark and hashing. Computing (Vienna/New York), 2021, 103, 2007-2031.	3.2	3
8837	Investigating Approaches of Integrating BIM, IoT, and Facility Management for Renovating Existing Buildings: A Review. Sustainability, 2021, 13, 3930.	1.6	41
8838	Widening: using parallel resources to improve model quality. Data Mining and Knowledge Discovery, 2021, 35, 1258.	2.4	0
8840	Optimizations for filter-based join algorithms in MapReduce. Journal of Intelligent and Fuzzy Systems, 2021, 40, 8963-8980.	0.8	5
8841	GPrimer: a fast GPU-based pipeline for primer design for qPCR experiments. BMC Bioinformatics, 2021, 22, 220.	1.2	5
8842	Performance Models of Data Parallel DAG Workflows for Large Scale Data Analytics. , 2021, , .		4
8843	Formalization and taxonomy of compute-aggregate problems for cloud computing applications. Computer Networks, 2021, 189, 107915.	3.2	0
8844	Self-organizing fuzzy inference ensemble system for big streaming data classification. Knowledge-Based Systems, 2021, 218, 106870.	4.0	22
8845	PaSh. , 2021, , .		8
8846	Synergy via Redundancy: Adaptive Replication Strategies and Fundamental Limits. IEEE/ACM Transactions on Networking, 2021, 29, 737-749.	2.6	5
8847	Distributed linear regression by averaging. Annals of Statistics, 2021, 49, .	1.4	17
8848	Modeling population dynamics for information dissemination through Facebook. Concurrency Computation Practice and Experience, 2023, 35, e6333.	1.4	1
8849	User-centric genomics infrastructure: trends and technologies. Genome, 2021, 64, 467-475.	0.9	1
8850	Automatically exploiting the memory hierarchy of GPUs through just-in-time compilation. , 2021, , .		1
8851	EigenCloud: A Cooperation and Trust-Aware Dependable Cloud File-Sharing Network. IEEE Transactions on Computational Social Systems, 2021, 8, 522-536.	3.2	3
8852	Executing cyclic scientific workflows in the cloud. Journal of Cloud Computing: Advances, Systems and Applications, 2021, 10, .	2.1	5
8853	Research on the application of crawler technology in machine learning. Journal of Physics: Conference Series, 2021, 1865, 042040.	0.3	3
8854	Machine Learning Frameworks. , 2021, , 173-219.		0

#	ARTICLE	IF	CITATIONS
8855	HBFPF-DC: A parallel frequent itemset mining using Spark. <i>Parallel Computing</i> , 2021, 101, 102738.	1.3	8
8856	Coding for Distributed Fog Computing in Internet of Mobile Things. <i>IEEE Transactions on Mobile Computing</i> , 2021, 20, 1337-1350.	3.9	12
8857	CIAO: An Optimization Framework for Client-Assisted Data Loading. , 2021, , .		2
8858	Trillion-scale Graph Processing Simulation based on Top-Down Graph Upscaling. , 2021, , .		0
8859	A Comparative Survey of Big Data Computing and HPC: From a Parallel Programming Model to a Cluster Architecture. <i>International Journal of Parallel Programming</i> , 2022, 50, 27-64.	1.1	8
8860	Guided container selection for data streaming through neural learning in cloud. <i>International Journal of Systems Assurance Engineering and Management</i> , 0, , 1.	1.5	20
8861	Energy-Efficient Scheduling of MapReduce Tasks Based on Load Balancing and Deadline Constraint in Heterogeneous Hadoop YARN Cluster. , 2021, , .		4
8862	Query Rewriting for Incremental Continuous Query Evaluation in HIFUN. <i>Algorithms</i> , 2021, 14, 149.	1.2	1
8863	Compressed Coded Distributed Computing. <i>IEEE Transactions on Communications</i> , 2021, 69, 2773-2783.	4.9	5
8864	Incentive-Based Coded Distributed Computing Management for Latency Reduction in IoT Services—A Game Theoretic Approach. <i>IEEE Internet of Things Journal</i> , 2021, 8, 8259-8278.	5.5	11
8865	Rack-Scaling: An efficient rack-based redistribution method to accelerate the scaling of cloud disk arrays. , 2021, , .		2
8866	Rider Chicken Optimization Algorithm-Based Recurrent Neural Network for Big Data Classification in Spark Architecture. <i>Computer Journal</i> , 2022, 65, 2183-2196.	1.5	1
8867	FTRLIM: Distributed Instance Matching Framework for Large-Scale Knowledge Graph Fusion. <i>Entropy</i> , 2021, 23, 602.	1.1	4
8868	Big Data in Smart City: Management Challenges. <i>Applied Sciences (Switzerland)</i> , 2021, 11, 4557.	1.3	11
8869	Adaptive IWD-based algorithm for deployment of business processes into cloud federations. <i>International Journal of Pervasive Computing and Communications</i> , 2021, ahead-of-print, .	1.1	0
8870	DeepMal: maliciousness-Preserving adversarial instruction learning against static malware detection. <i>Cybersecurity</i> , 2021, 4, .	3.1	12
8871	Revolution of Smart Healthcare Materials in Big Data Analytics. <i>Materials Today: Proceedings</i> , 2023, 81, 834-841.	0.9	9
8872	What serverless computing is and should become. <i>Communications of the ACM</i> , 2021, 64, 76-84.	3.3	102

#	ARTICLE	IF	CITATIONS
8873	Practical Analysis of Replication-Based Systems. , 2021, , .		0
8874	Big Data as a Tool for Building a Predictive Model of Mill Roll Wear. Symmetry, 2021, 13, 859.	1.1	38
8875	CASH: A Credit Aware Scheduling for Public Cloud Platforms. , 2021, , .		3
8876	Differentially Private Decision Tree Based on Pearson's Correlation Coefficient. , 2021, , .		2
8877	Rim. , 2021, , .		8
8878	Collaborative Learning Based Straggler Prevention in Large-Scale Distributed Computing Framework. Security and Communication Networks, 2021, 2021, 1-9.	1.0	47
8879	Distributed temporal graph analytics with GRADOOP. VLDB Journal, 2022, 31, 375-401.	2.7	13
8880	Scheduling for heterogeneous systems in accelerator-rich environments. Journal of Supercomputing, 2022, 78, 200-221.	2.4	4
8881	Robust Service Mapping in Multi-Tenant Clouds. , 2021, , .		3
8882	Sherlock: an open-source data platform to store, analyze and integrate Big Data for biology. F1000Research, 0, 10, 409.	0.8	1
8883	Collaborative Fog Computing Architecture for Privacy-Preserving Data Aggregation. , 2021, , .		3
8884	Data science for oceanography: from small data to big data. Big Earth Data, 2022, 6, 236-250.	2.0	10
8885	Distributed Large-scale Time-series Data Processing and Analysis System Based on Spark Platform. , 2021, , .		2
8886	Doughnutie: An efficient and low-latency cloud data center network architecture. Concurrency Computation Practice and Experience, 2021, 33, e6337.	1.4	2
8887	Providing In-network Support to Coflow Scheduling. , 2021, , .		2
8888	A Theoretical and Experimental Comparison of Large-Scale Join Algorithms in Spark. SN Computer Science, 2021, 2, 1.	2.3	2
8889	Autonomous Load Balancing in Distributed Hash Tables Using Churn and the Sybil Attack. , 2021, , .		0
8890	Memory Efficient Edge Addition Designs for Large and Dynamic Social Networks. , 2021, , .		0

#	ARTICLE	IF	CITATIONS
8891	Phased synthesis of divide and conquer programs. , 2021, , .		4
8892	NPSCS: Non-Preemptive Stochastic Coflow Scheduling With Time-Indexed LP Relaxation. IEEE Transactions on Network and Service Management, 2021, 18, 2377-2387.	3.2	2
8893	Criso: An Incremental Scalable and Cost-Effective Network Architecture for Data Centers. IEEE Transactions on Network and Service Management, 2021, 18, 2016-2029.	3.2	7
8894	PMAB: A Public Mutual Audit Blockchain for Outsourced Data in Cloud Storage. Security and Communication Networks, 2021, 2021, 1-11.	1.0	3
8895	NodeRank: Finding influential nodes in social networks based on interests. Journal of Supercomputing, 2022, 78, 2098-2124.	2.4	9
8896	Algorithms for a Topology-aware Massively Parallel Computation Model. , 2021, , .		1
8897	Artificial intelligence and database technology work together. Journal of Physics: Conference Series, 2021, 1955, 012004.	0.3	1
8898	Graph Sparsification for Derandomizing Massively Parallel Computation with Low Space. ACM Transactions on Algorithms, 2021, 17, 1-27.	0.9	6
8899	CodedSketch: A Coding Scheme for Distributed Computation of Approximated Matrix Multiplication. IEEE Transactions on Information Theory, 2021, 67, 4185-4196.	1.5	15
8900	Scalable feature selection using ReliefF aided by locality-sensitive hashing. International Journal of Intelligent Systems, 2021, 36, 6161-6179.	3.3	11
8901	Stargazer: Toward efficient data analytics scheduling via task completion time inference. Computers and Electrical Engineering, 2021, 92, 107092.	3.0	0
8902	Django: Bilateral coflow scheduling with predictive concurrent connections. Journal of Parallel and Distributed Computing, 2021, 152, 45-56.	2.7	3
8903	Towards Large Scale Image Retrieval System Using Parallel Frameworks. , 0, , .		0
8904	An optimized tensor completion library for multiple GPUs. , 2021, , .		1
8905	Characterizing Job-Task Dependency in Cloud Workloads Using Graph Learning. , 2021, , .		3
8906	ABC-DIMM: Alleviating the Bottleneck of Communication in DIMM-based Near-Memory Processing with Inter-DIMM Broadcast. , 2021, , .		9
8907	Deploying a Strategy to Unlock Big Data Research and Teaching Activities in the West Balkan Region. , 2021, , .		1
8908	Real-time Data Infrastructure at Uber. , 2021, , .		18

#	ARTICLE	IF	CITATIONS
8909	Big data analytics for intelligent manufacturing systems: A review. Journal of Manufacturing Systems, 2022, 62, 738-752.	7.6	155
8910	s2p: Provenance Research for Stream Processing System. Applied Sciences (Switzerland), 2021, 11, 5523.	1.3	3
8911	Satellite Image Time Series Analysis for Big Earth Observation Data. Remote Sensing, 2021, 13, 2428.	1.8	36
8912	From warm to hot starts. , 2021, , .		12
8913	Doing more with less. , 2021, , .		3
8914	Treeplication: An Erasure Code for Distributed Full Recovery Under the Random Multiset Channel. IEEE Transactions on Information Theory, 2021, 67, 3542-3556.	1.5	1
8915	HTPC: heterogeneous traffic-aware partition coding for random packet spraying in data center networks. Journal of Cloud Computing: Advances, Systems and Applications, 2021, 10, .	2.1	3
8916	Cover or Pack. , 2021, , .		3
8917	Screening hardware and volume factors in distributed machine learning algorithms on spark. Computing (Vienna/New York), 2021, 103, 2203-2225.	3.2	4
8918	CiM3D: Comparator-in-Memory Designs Using Monolithic 3-D Technology for Accelerating Data-Intensive Applications. IEEE Journal on Exploratory Solid-State Computational Devices and Circuits, 2021, 7, 79-87.	1.1	2
8919	Analysis of Bayesian optimization algorithms for big data classification based on Map Reduce framework. Journal of Big Data, 2021, 8, .	6.9	8
8920	Scale-down experiments on TPCx-HS. , 2021, , .		0
8921	Intelligent Data Collaboration in Heterogeneous-device IoT Platforms. ACM Transactions on Sensor Networks, 2021, 17, 1-17.	2.3	11
8922	A Scalable Platform for Distributed Object Tracking Across a Many-Camera Network. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 1479-1493.	4.0	10
8923	Regularization-Based Coflow Scheduling in Optical Circuit Switches. IEEE/ACM Transactions on Networking, 2021, 29, 1280-1293.	2.6	6
8924	Cheating Like The Neighbors: Logarithmic Complexity For Fitness Evaluation In Genetic Algorithms. , 2021, , .		1
8925	Predictive, reactive and replication-based load balancing of tasks in Chameleon and sam(oa) ². , 2021, , .		3
8926	X-composer. , 2021, , .		0

#	ARTICLE	IF	CITATIONS
8927	Security-aware task scheduling with deadline constraints on heterogeneous hybrid clouds. Journal of Parallel and Distributed Computing, 2021, 153, 15-28.	2.7	24
8928	Preprocessing of Datasets Using Sequential and Parallel Approach: A Comparison. Lecture Notes in Networks and Systems, 2022, , 311-320.	0.5	0
8929	Massively Parallel Algorithms for Distance Approximation and Spanners. , 2021, , .		4
8930	Robustness Comparison of Scheduling Algorithms in MapReduce Framework. Lecture Notes in Networks and Systems, 2022, , 494-508.	0.5	0
8931	Performance and scaling behavior of bioinformatic applications in virtualization environments to create awareness for the efficient use of compute resources. PLoS Computational Biology, 2021, 17, e1009244.	1.5	7
8932	Modelling Place Visit Probability Sequences during Trajectory Data Gaps Based on Movement History. ISPRS International Journal of Geo-Information, 2021, 10, 456.	1.4	0
8933	A comparative experimental study of distributed storage engines for big spatial data processing using GeoSpark. Journal of Supercomputing, 2022, 78, 2556-2579.	2.4	4
8934	Basketball Data Analysis Using Spark Framework and K-Means Algorithm. Journal of Healthcare Engineering, 2021, 2021, 1-7.	1.1	3
8935	Byzantine-Resilient SGD in High Dimensions on Heterogeneous Data. , 2021, , .		5
8936	Bandwidth-Aware Rescheduling Mechanism in SDN-Based Data Center Networks. Electronics (Switzerland), 2021, 10, 1774.	1.8	3
8937	Self-organizing Data Processing for Time Series Using SPARK. Lecture Notes on Data Engineering and Communications Technologies, 2022, , 239-248.	0.5	0
8939	Artificial intelligence-based Kubernetes container for scheduling nodes of energy composition. International Journal of Systems Assurance Engineering and Management, 0, , 1.	1.5	15
8940	Accelerating Gossip-Based Deep Learning in Heterogeneous Edge Computing Platforms. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 1591-1602.	4.0	17
8941	MP-CREDIT: Multi-path credit for high-speed data center transports. Computer Networks, 2021, 193, 108061.	3.2	3
8942	A Theoretical Model for Global Optimization of Parallel Algorithms. Mathematics, 2021, 9, 1685.	1.1	4
8943	On Batch-Processing Based Coded Computing for Heterogeneous Distributed Computing Systems. IEEE Transactions on Network Science and Engineering, 2021, 8, 2438-2454.	4.1	11
8945	FDR2-BD: A Fast Data Reduction Recommendation Tool for Tabular Big Data Classification Problems. Electronics (Switzerland), 2021, 10, 1757.	1.8	3
8946	Performance Analysis of Hadoop YARN Job Schedulers in a Multi-Tenant Environment on HiBench Benchmark Suite. International Journal of Distributed Systems and Technologies, 2021, 12, 64-82.	0.6	3

#	ARTICLE	IF	CITATIONS
8947	Big data classification with optimization driven MapReduce framework. International Journal of Knowledge-Based and Intelligent Engineering Systems, 2021, 25, 173-183.	0.7	4
8948	Data Lake Versus Data Warehouse Architecture: A Comparative Study. Lecture Notes in Electrical Engineering, 2022, , 201-210.	0.3	4
8949	Employing Vertical Elasticity for Efficient Big Data Processing in Container-Based Cloud Environments. Applied Sciences (Switzerland), 2021, 11, 6200.	1.3	4
8950	SentiLog. , 2021, , .		10
8951	A Blockchain-Based Decentralized Framework for Fair Data Processing. IEEE Transactions on Network Science and Engineering, 2021, 8, 2301-2315.	4.1	2
8952	Fault tolerance in big data storage and processing systems: A review on challenges and solutions. Ain Shams Engineering Journal, 2022, 13, 101538.	3.5	22
8953	SPARQL2Flink: Evaluation of SPARQL Queries on Apache Flink. Applied Sciences (Switzerland), 2021, 11, 7033.	1.3	2
8954	An Enhanced Security Measure for Multimedia Images Using Hadoop Cluster. International Journal of Operations Research and Information Systems, 2021, 12, 1-7.	1.0	8
8955	Content-aware data distribution over cluster nodes. Intelligent Data Analysis, 2021, 25, 907-927.	0.4	3
8956	An order-aware dataflow model for parallel Unix pipelines. , 2021, 5, 1-28.		1
8957	Non-Communicable Diseases and Social Media: A Heart Disease Symptoms Application. Journal of Information and Knowledge Management, 2021, 20, .	0.8	0
8958	Towards big services: a synergy between service computing and parallel programming. Computing (Vienna/New York), 2021, 103, 2479-2519.	3.2	7
8959	<i>Hone:</i> Mitigating Stragglers in Distributed Stream Processing With Tuple Scheduling. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 2021-2034.	4.0	10
8960	A parallelization model for performance characterization of Spark Big Data jobs on Hadoop clusters. Journal of Big Data, 2021, 8, .	6.9	10
8961	TempestExtremes v2.1: a community framework for feature detection, tracking, and analysis in large datasets. Geoscientific Model Development, 2021, 14, 5023-5048.	1.3	53
8962	Distributed Deep Learning for Remote Sensing Data Interpretation. Proceedings of the IEEE, 2021, 109, 1320-1349.	16.4	16
8963	Map Reduce based on Cloak DHT Data Replication Evaluation. International Journal of Database Management Systems, 2021, 13, 9-25.	0.2	0
8964	A Comprehensive Survey for Hadoop Distributed File System. Asian Journal of Research in Computer Science, 0, , 46-57.	0.0	9

#	ARTICLE	IF	CITATIONS
8965	Balanced Multipath Transport Protocol for Mitigating MPTCP Incast in Data Center Networks. International Journal of Next-generation Computing, 2021, 12, .	1.1	0
8966	MapEmbed. , 2021, , .		6
8967	Hoplite. , 2021, , .		7
8968	Parallel and Distributed Computing for Anomaly Detection From Hyperspectral Remote Sensing Imagery. Proceedings of the IEEE, 2021, 109, 1306-1319.	16.4	12
8969	Recent Developments in Parallel and Distributed Computing for Remotely Sensed Big Data Processing. Proceedings of the IEEE, 2021, 109, 1282-1305.	16.4	45
8970	A parallel and accurate method for large-scale image segmentation on a cloud environment. Journal of Supercomputing, 2022, 78, 4330-4357.	2.4	2
8971	Towards enabling I/O awareness in task-based programming models. Future Generation Computer Systems, 2021, 121, 74-89.	4.9	1
8972	The Optimization Strategies on Clarification of the Misconceptions of Big Data Processing in Dynamic and Opportunistic Environments. Big Data and Cognitive Computing, 2021, 5, 38.	2.9	0
8973	A Fault-Tolerant Distributed Framework for Asynchronous Iterative Computations. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 2062-2073.	4.0	0
8974	Increasing the Prediction Accuracy for Thyroid Disease: A Step Towards Better Health for Society. Wireless Personal Communications, 2022, 122, 1921-1938.	1.8	17
8975	Confidential machine learning on untrusted platforms: a survey. Cybersecurity, 2021, 4, 30.	3.1	8
8978	How Many CPU Cores is an FPGA Worth? Lessons Learned from Accelerating String Sorting on a CPU-FPGA System. Journal of Signal Processing Systems, 0, , 1.	1.4	2
8979	Fiber Clustering Acceleration With a Modified Kmeans++ Algorithm Using Data Parallelism. Frontiers in Neuroinformatics, 2021, 15, 727859.	1.3	7
8980	An extensive review of data security infrastructure and legislature. Security and Privacy, 2022, 5, e189.	1.9	0
8981	A Combinatorial Design for Cascaded Coded Distributed Computing on General Networks. IEEE Transactions on Communications, 2021, 69, 5686-5700.	4.9	4
8982	Advancing Logistics 4.0 with the Implementation of a Big Data Warehouse: A Demonstration Case for the Automotive Industry. Electronics (Switzerland), 2021, 10, 2221.	1.8	22
8983	ProSPECT: Proactive Storage Using Provenance for Efficient Compute and Tiering. , 2022, 7, 219-234.		0
8984	Review of urban computing in air quality management as smart city service: An integrated IoT, AI, and cloud technology perspective. Urban Climate, 2021, 39, 100972.	2.4	70

#	ARTICLE	IF	CITATIONS
8985	MR-MVPP: A map-reduce-based approach for creating MVPP in data warehouses for big data applications. <i>Information Sciences</i> , 2021, 570, 200-224.	4.0	9
8986	Spark-based parallel processing whale optimization algorithm. <i>Concurrency Computation Practice and Experience</i> , 2022, 34, e6607.	1.4	3
8987	Optimization Techniques for a Distributed In-Memory Computing Platform by Leveraging SSD. <i>Applied Sciences (Switzerland)</i> , 2021, 11, 8476.	1.3	1
8988	Small files™ problem in Hadoop: A systematic literature review. <i>Journal of King Saud University - Computer and Information Sciences</i> , 2022, 34, 8658-8674.	2.7	10
8989	An integrated multi-node Hadoop framework to predict high-risk factors of Diabetes Mellitus using a Multilevel MapReduce based Fuzzy Classifier (MMR-FC) and Modified DBSCAN algorithm. <i>Applied Soft Computing Journal</i> , 2021, 108, 107423.	4.1	5
8990	Efficient and Flexible Aggregation and Distribution of MODIS Atmospheric Products Based on Climate Analytics as a Service Framework. <i>Remote Sensing</i> , 2021, 13, 3541.	1.8	2
8991	Straggler identification approach in large data processing frameworks using ensembled gradient boosting in smart-cities cloud services. <i>International Journal of Systems Assurance Engineering and Management</i> , 2022, 13, 146-155.	1.5	1
8992	A New Combinatorial Coded Design for Heterogeneous Distributed Computing. <i>IEEE Transactions on Communications</i> , 2021, 69, 5672-5685.	4.9	11
8993	High-volume point cloud data simplification based on decomposed graph filtering. <i>Automation in Construction</i> , 2021, 129, 103815.	4.8	13
8994	A MapReduce-based K-means clustering algorithm. <i>Journal of Supercomputing</i> , 0, , 1.	2.4	3
8995	Massively Parallel Computation via Remote Memory Access. <i>ACM Transactions on Parallel Computing</i> , 2021, 8, 1-25.	1.2	4
8997	Mining Profitable and Concise Patterns in Large-Scale Internet of Things Environments. <i>Wireless Communications and Mobile Computing</i> , 2021, 2021, 1-12.	0.8	1
8998	Analysis and improvement of the latency-based congestion control algorithm DX. <i>Future Generation Computer Systems</i> , 2021, 123, 206-218.	4.9	2
8999	Effective *-flow schedule for optical circuit switching based data center networks: A comprehensive survey. <i>Computer Networks</i> , 2021, 197, 108321.	3.2	6
9000	Candidate region aware nested named entity recognition. <i>Neural Networks</i> , 2021, 142, 340-350.	3.3	12
9001	Hadoop Perfect File: A fast and memory-efficient metadata access archive file to face small files problem in HDFS. <i>Journal of Parallel and Distributed Computing</i> , 2021, 156, 119-130.	2.7	14
9002	Unsupervised skill transfer learning for autonomous robots using distributed Growing Self Organizing Maps. <i>Robotics and Autonomous Systems</i> , 2021, 144, 103835.	3.0	9
9003	Rotation Forest for Big Data. <i>Information Fusion</i> , 2021, 74, 39-49.	11.7	8

#	ARTICLE	IF	CITATIONS
9004	Parallel algorithm for improving the performance of spatial queries in SQL: The use cases of SQLite/SpatialLite and PostgreSQL/PostGIS databases. Computers and Geosciences, 2021, 155, 104840.	2.0	7
9005	Instance Matching in Knowledge Graphs through random walks and semantics. Future Generation Computer Systems, 2021, 123, 73-84.	4.9	3
9006	MapReduce-Based Crow Search-Adopted Partitional Clustering Algorithms for Handling Large-Scale Data. International Journal of Cognitive Informatics and Natural Intelligence, 2021, 15, 1-23.	0.4	0
9007	Evidential instance selection for K-nearest neighbor classification of big data. International Journal of Approximate Reasoning, 2021, 138, 123-144.	1.9	20
9008	Coflow Scheduling in Data Centers: Routing and Bandwidth Allocation. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 2661-2675.	4.0	9
9009	Data Analytic Models That Redress the Limitations of MapReduce. International Journal of Web-Based Learning and Teaching Technologies, 2021, 16, 1-15.	0.6	1
9010	Scalable and Flexible Two-Phase Ensemble Algorithms for Causality Discovery. Big Data Research, 2021, 26, 100252.	2.6	4
9011	Optimal task replication considering reliability, performance, and energy consumption for parallel computing in cloud systems. Reliability Engineering and System Safety, 2021, 215, 107834.	5.1	8
9012	Scaling the Growing Neural Gas for Visual Cluster Analysis. Big Data Research, 2021, 26, 100254.	2.6	6
9013	Sentiment Weighted Word Embedding for Big Text Data. International Journal of Web-Based Learning and Teaching Technologies, 2021, 16, 1-17.	0.6	0
9014	An Incremental Iterative Acceleration Architecture in Distributed Heterogeneous Environments With GPUs for Deep Learning. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 2823-2837.	4.0	3
9015	Preemptive scheduling on unrelated machines with fractional precedence constraints. Journal of Parallel and Distributed Computing, 2021, 157, 280-286.	2.7	1
9016	Big data analytics and machine learning: A retrospective overview and bibliometric analysis. Expert Systems With Applications, 2021, 184, 115561.	4.4	50
9017	ARENA: Asynchronous Reconfigurable Accelerator Ring to Enable Data-Centric Parallel Computing. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 2880-2892.	4.0	11
9018	RENDa: Resource and Network Aware Data Placement Algorithm for Periodic Workloads in Cloud. IEEE Transactions on Parallel and Distributed Systems, 2021, 32, 2906-2920.	4.0	16
9019	Efficient closed high-utility pattern fusion model in large-scale databases. Information Fusion, 2021, 76, 122-132.	11.7	34
9020	Design-time performance modeling of compositional parallel programs. Parallel Computing, 2021, 108, 102839.	1.3	1
9021	Device Hopping. Transactions on Architecture and Code Optimization, 2021, 18, 1-25.	1.6	0

#	ARTICLE	IF	CITATIONS
9022	Optimizing Network Transfers for Data Analytic Jobs Across Geo-Distributed Datacenters. IEEE Transactions on Parallel and Distributed Systems, 2022, 33, 403-414.	4.0	8
9023	Effective and efficient skyline query processing over attribute-order-preserving-free encrypted data in cloud-enabled databases. Future Generation Computer Systems, 2022, 126, 237-251.	4.9	12
9024	Towards a data science platform for improving SME collaboration through Industry 4.0 technologies. Technological Forecasting and Social Change, 2022, 174, 121242.	6.2	55
9025	Topic Modeling Using Latent Dirichlet allocation. ACM Computing Surveys, 2022, 54, 1-35.	16.1	64
9026	Coded computing for on-device cooperative inference. , 2022, , 83-101.		1
9027	A Dual-Index Based Representation for Processing XPath Queries on Very Large XML Documents. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2021, , 18-30.	0.2	0
9028	Assessment of Complexity in Cloud Computing Adoption: a Case Study of Local Governments in Australia. Information Systems Frontiers, 2022, 24, 595-617.	4.1	10
9029	iAnalysis V1.0: An Interactive Analysis Service System. , 2021, , .		0
9030	Decision Support in the Era of Social Media and User-Generated Content. Integrated Series on Information Systems, 2021, , 79-95.	0.1	0
9031	Distributed Tree-Based Machine Learning for Short-Term Load Forecasting With Apache Spark. IEEE Access, 2021, 9, 57372-57384.	2.6	7
9033	Adaptive and Heterogeneity-Aware Coded Cooperative Computation at the Edge. IEEE Transactions on Mobile Computing, 2021, , 1-1.	3.9	3
9034	Optimizing Internal Overlaps by Self-Adjusting Resource Allocation in Multi-Stage Computing Systems. IEEE Access, 2021, 9, 88805-88819.	2.6	1
9035	Crop Management with the IoT: An Interdisciplinary Survey. Agronomy, 2021, 11, 181.	1.3	31
9036	Open problems in queueing theory inspired by datacenter computing. Queueing Systems, 2021, 97, 3-37.	0.6	27
9037	Bringing AI to the edge: a formal M&S specification to deploy effective IoT architectures. Journal of Simulation, 2022, 16, 494-511.	1.0	6
9038	An improved approach for mining association rules in parallel using Spark Streaming. International Journal of Circuit Theory and Applications, 2021, 49, 1028-1039.	1.3	14
9039	â€œSaksham Modelâ€•Performance Improvisation Using Node Capability Evaluation in Apache Hadoop. , 2021, , 1282-1302.		0
9040	Open Source Software (OSS) for Big Data. , 2021, , 858-875.		0

#	ARTICLE	IF	CITATIONS
9041	Toward Efficient Online Scheduling for Distributed Machine Learning Systems. IEEE Transactions on Network Science and Engineering, 2022, 9, 1951-1969.	4.1	2
9042	Multi-job Merging Framework and Scheduling Optimization for Apache Flink. Lecture Notes in Computer Science, 2021, , 20-36.	1.0	0
9043	A MOM-based ensemble method for robustness, subsampling and hyperparameter tuning. Electronic Journal of Statistics, 2021, 15, .	0.4	0
9044	Spark Based Intrusion Detection System Using Practical Swarm Optimization Clustering. Studies in Big Data, 2021, , 197-216.	0.8	0
9045	Multi-Client Functional Encryption for Separable Functions. Lecture Notes in Computer Science, 2021, , 724-753.	1.0	3
9046	Optimization of Systems-Development Life Cycle Through Automation Using Ansible. Lecture Notes in Electrical Engineering, 2021, , 229-240.	0.3	0
9048	MBTree: Detecting Encryption RATs Communication Using Malicious Behavior Tree. IEEE Transactions on Information Forensics and Security, 2021, 16, 3589-3603.	4.5	16
9049	A Comprehensive Survey on Coded Distributed Computing: Fundamentals, Challenges, and Networking Applications. IEEE Communications Surveys and Tutorials, 2021, 23, 1800-1837.	24.8	38
9050	A Ray-Based Remote Distributed Computing System for Autonomous Driving. Computer Science and Application, 2021, 11, 1334-1340.	0.0	0
9051	Programmable Implementation and Blockchain Security Scheme Based on Edge Computing Firework Model. International Journal of Information Technology and Web Engineering, 2021, 16, 1-22.	1.2	2
9052	Single-Forking of Coded Subtasks for Straggler Mitigation. IEEE/ACM Transactions on Networking, 2021, 29, 2413-2424.	2.6	4
9053	Federated Machine Learning: Survey, Multi-Level Classification, Desirable Criteria and Future Directions in Communication and Networking Systems. IEEE Communications Surveys and Tutorials, 2021, 23, 1342-1397.	24.8	243
9054	Big Data Processing on Cloud Computing Using Hadoop Mapreduce and Apache Spark. , 2021, , 824-845.		0
9055	Big Data Applications and Architectures for Emerging Countries. Advances in Higher Education and Professional Development Book Series, 2021, , 120-142.	0.1	1
9056	Fair and near-optimal coflow scheduling without prior knowledge of coflow size. Journal of Supercomputing, 2021, 77, 7690-7717.	2.4	1
9057	Analysis and implementation of reactive fault tolerance techniques in Hadoop: a comparative study. Journal of Supercomputing, 2021, 77, 7184-7210.	2.4	6
9058	Software as a Service, Semantic Web, and Big Data. , 2021, , 1179-1201.		3
9059	Software Development Methodology for Cloud Computing and Its Impact. , 2021, , 151-172.		1

#	ARTICLE	IF	CITATIONS
9060	Integrated Electronic HEI Performance Management. Advances in Logistics, Operations, and Management Science Book Series, 2021, , 1108-1131.	0.3	0
9061	Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines. IEEE Access, 2021, 9, 120043-120065.	2.6	137
9062	OODIDA: On-Board/Off-Board Distributed Real-Time Data Analytics for Connected Vehicles. Data Science and Engineering, 2021, 6, 102-117.	4.6	10
9064	Big data and machine learning framework for clouds and its usage for text classification. Concurrency Computation Practice and Experience, 2021, 33, e6164.	1.4	10
9068	Convergence in Language Design: A Case of Lightning Striking Four Times in the Same Place. Lecture Notes in Computer Science, 2006, , 2-12.	1.0	3
9069	Grids, Clouds, and Virtualization. , 2011, , 1-21.		21
9070	Scientific Data Management in the Cloud: A Survey of Technologies, Approaches and Challenges. , 2010, , 517-533.		3
9071	Grids, P2P and Services Computing. , 2010, , .		1
9073	Random Prism: An Alternative to Random Forests. , 2011, , 5-18.		11
9074	Information Retrieval and Digital Libraries. , 2014, , 613-641.		5
9075	Development of Cloud Applications in Hybrid Clouds with Support for Multi-scheduling. Computer Communications and Networks, 2013, , 283-301.	0.8	2
9076	Multi-sensor Fusion. , 2014, , 301-354.		22
9077	Big Data, Cloud Computing, and Internet of Things. , 2019, , 905-932.		1
9079	Geospatial Data Management with Terraflly. , 2011, , 637-665.		2
9080	GPGPU Computing for Cloud Auditing. , 2014, , 259-282.		1
9081	A Petabit Bufferless Optical Switch for Data Center Networks. Optical Networks Series, 2013, , 135-154.	1.1	46
9082	The Learning Registry: Applying Social Metadata for Learning Resource Recommendations. , 2014, , 77-95.		3
9084	GPU-Accelerated Cloud Computing for Data-Intensive Applications. , 2014, , 105-129.		4

#	ARTICLE	IF	CITATIONS
9085	Relational Databases and Biomedical Big Data. <i>Methods in Molecular Biology</i> , 2017, 1617, 69-81.	0.4	2
9086	Developing Genetic Algorithms Using Different MapReduce Frameworks: MPI vs. Hadoop. <i>Lecture Notes in Computer Science</i> , 2018, , 262-272.	1.0	2
9088	lIlastic: Linked Data Generation Workflow and User Interface for iMinds Scholarly Data. <i>Lecture Notes in Computer Science</i> , 2018, , 15-32.	1.0	1
9089	A Distributed Approach for Machine Learning in Large Scale Manufacturing Systems. <i>Studies in Computational Intelligence</i> , 2019, , 41-52.	0.7	5
9090	An Efficient Batch Similarity Processing with MapReduce. <i>Lecture Notes in Computer Science</i> , 2018, , 158-171.	1.0	1
9092	Relational Equivalence Proofs Between Imperative and MapReduce Algorithms. <i>Lecture Notes in Computer Science</i> , 2018, , 248-266.	1.0	7
9093	Exploiting Complex Protein Domain Networks for Protein Function Annotation. <i>Studies in Computational Intelligence</i> , 2019, , 598-610.	0.7	7
9094	Analysis of Streaming Data Using Big Data and Hybrid Machine Learning Approach. , 2019, , 629-643.		9
9095	Facial Recognition on Cloud for Android Based Wearable Devices. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 98-110.	0.5	3
9096	Parallelizing Convergent Cross Mapping Using Apache Spark. <i>Lecture Notes in Computer Science</i> , 2019, , 133-142.	1.0	3
9097	Aggregate Processes in Field Calculus. <i>Lecture Notes in Computer Science</i> , 2019, , 200-217.	1.0	11
9098	Hadoop: A Standard Framework for Computer Cluster. , 2019, , 341-356.		5
9099	Breaking the Linear-Memory Barrier in MPC : Fast MIS on Trees with Strongly Sublinear Memory. <i>Lecture Notes in Computer Science</i> , 2019, , 124-138.	1.0	6
9100	Implementation of NN Algorithm Based on the MapReduce Framework. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 709-717.	0.5	1
9101	Towards a Cost Model to Optimize User-Defined Functions in an ETL Workflow Based on User-Defined Performance Metrics. <i>Lecture Notes in Computer Science</i> , 2019, , 441-456.	1.0	8
9102	Hadoop Scalability and Performance Testing in Homogeneous Clusters. <i>Lecture Notes in Electrical Engineering</i> , 2020, , 907-917.	0.3	1
9103	A Review of Big Data and Its Applications in Healthcare and Public Sector. <i>Studies in Big Data</i> , 2020, , 55-66.	0.8	10
9104	Cyber-Physical Cloud Computing Systems and Internet of Everything. <i>Intelligent Systems Reference Library</i> , 2020, , 201-227.	1.0	7

#	ARTICLE	IF	CITATIONS
9105	Data Capsule: A New Paradigm for Automatic Compliance with Data Privacy Regulations. Lecture Notes in Computer Science, 2019, , 3-23.	1.0	18
9106	Augmented Reality Technique Assists Target Language Learning. Lecture Notes in Computer Science, 2019, , 558-567.	1.0	5
9107	RDD-Eclat: Approaches to Parallelize Eclat Algorithm on Spark RDD Framework. Lecture Notes on Data Engineering and Communications Technologies, 2020, , 755-768.	0.5	6
9109	Reduction of Traffic on Roads Using Big Data Applications. Lecture Notes on Data Engineering and Communications Technologies, 2020, , 333-339.	0.5	2
9110	An Empirical Study on the Use and Misuse of Java 8 Streams. Lecture Notes in Computer Science, 2020, , 97-118.	1.0	12
9111	Healthcare 4.0: A Voyage of Fog Computing with IOT, Cloud Computing, Big Data, and Machine Learning. Signals and Communication Technology, 2021, , 177-210.	0.4	10
9112	Geospatial Big Data Handling with High Performance Computing: Current Approaches and Future Directions. Geotechnologies and the Environment, 2020, , 53-76.	0.3	16
9113	FLY: A Domain-Specific Language for Scientific Computing on FaaS. Lecture Notes in Computer Science, 2020, , 531-544.	1.0	1
9114	Artificial Intelligence in Health Care: Predictive Analysis on Diabetes Using Machine Learning Algorithms. Lecture Notes in Computer Science, 2020, , 354-366.	1.0	4
9115	Scalable Reference Genome Assembly from Compressed Pan-Genome Index with Spark. Lecture Notes in Computer Science, 2020, , 68-84.	1.0	2
9116	Distributed and Parallel Ensemble Classification for Big Data Based on Kullback-Leibler Random Sample Partition. Lecture Notes in Computer Science, 2020, , 448-464.	1.0	8
9117	A Comparative Study of Join Algorithms in Spark. Lecture Notes in Computer Science, 2020, , 185-198.	1.0	3
9118	Load Balancing Approach for a MapReduce Job Running on a Heterogeneous Hadoop Cluster. Lecture Notes in Computer Science, 2021, , 289-298.	1.0	5
9119	A Two-Step Feature Space Transforming Method to Improve Credit Scoring Performance. Communications in Computer and Information Science, 2020, , 134-157.	0.4	1
9121	Analysis and Improvement of SPRINT Algorithm Based on Hadoop. Lecture Notes in Electrical Engineering, 2014, , 209-217.	0.3	2
9122	A Parallel and Convergent Support Vector Machine Based on MapReduce. Lecture Notes in Electrical Engineering, 2014, , 585-592.	0.3	2
9123	GraphDB â€œ Storing Large Graphs on Secondary Memory. Advances in Intelligent Systems and Computing, 2014, , 177-186.	0.5	4
9124	Distributed Processing of XPath Queries Using MapReduce. Advances in Intelligent Systems and Computing, 2014, , 69-77.	0.5	8

#	ARTICLE	IF	CITATIONS
9125	Pocket Data Mining. Studies in Big Data, 2014, , .	0.8	22
9126	Big Data for Enhanced Learning Analytics: A Case for Large-Scale Comparative Assessments. Communications in Computer and Information Science, 2013, , 225-233.	0.4	2
9127	On Scheduling Algorithms for MapReduce Jobs in Heterogeneous Clouds with Budget Constraints. Lecture Notes in Computer Science, 2013, , 251-265.	1.0	6
9128	A Practice of TPC-DS Multidimensional Implementation on NoSQL Database Systems. Lecture Notes in Computer Science, 2014, , 93-108.	1.0	10
9129	A Research Agenda for the Study of Entropic Social Structural Evolution, Functional Roles, Adhocratic Leadership Styles, and Credibility in Online Organizations and Knowledge Markets. , 2015, , 3-33.		7
9130	Large-Scale Similarity Join with Edit-Distance Constraints. Lecture Notes in Computer Science, 2014, , 328-342.	1.0	7
9131	BigOP: Generating Comprehensive Big Data Workloads as a Benchmarking Framework. Lecture Notes in Computer Science, 2014, , 483-492.	1.0	9
9133	Big Data Storage. SpringerBriefs in Computer Science, 2014, , 33-49.	0.2	6
9134	Are Public Clouds Elastic Enough for Scientific Computing?. Lecture Notes in Computer Science, 2014, , 294-307.	1.0	2
9135	Self-splitting of Workload in Parallel Computation. Lecture Notes in Computer Science, 2014, , 394-404.	1.0	12
9136	Finding Vacant Taxis Using Large Scale GPS Traces. Lecture Notes in Computer Science, 2014, , 793-804.	1.0	3
9137	Constructing Virtual Private Supercomputer Using Virtualization and Cloud Technologies. Lecture Notes in Computer Science, 2014, , 341-354.	1.0	18
9138	Research and Application of DBSCAN Algorithm Based on Hadoop Platform. Lecture Notes in Computer Science, 2014, , 73-87.	1.0	3
9139	Fast Computation of Recurrences in Long Time Series. Springer Proceedings in Mathematics and Statistics, 2014, , 17-29.	0.1	8
9140	Lessons from the Congested Clique Applied to MapReduce. Lecture Notes in Computer Science, 2014, , 149-164.	1.0	17
9141	Trace Checking of Metric Temporal Logic with Aggregating Modalities Using MapReduce. Lecture Notes in Computer Science, 2014, , 144-158.	1.0	17
9142	A BigBench Implementation in the Hadoop Ecosystem. Lecture Notes in Computer Science, 2014, , 3-18.	1.0	11
9143	A Multidimensional OLAP Engine Implementation in Key-Value Database Systems. Lecture Notes in Computer Science, 2014, , 155-170.	1.0	2

#	ARTICLE	IF	CITATIONS
9144	On Interlinking Linked Data Sources by Using Ontology Matching Techniques and the Map-Reduce Framework. Lecture Notes in Computer Science, 2014, , 53-60.	1.0	5
9145	SPKV: A Multi-dimensional Index System for Large Scale Key-Value Stores. Lecture Notes in Computer Science, 2014, , 367-378.	1.0	2
9146	An Adaptive Skew Insensitive Join Algorithm for Large Scale Data Analytics. Lecture Notes in Computer Science, 2014, , 494-502.	1.0	4
9147	HPSO: Prefetching Based Scheduling to Improve Data Locality for MapReduce Clusters. Lecture Notes in Computer Science, 2014, , 82-95.	1.0	9
9148	Count Me In: Viable Distributed Summary Statistics for Securing High-Speed Networks. Lecture Notes in Computer Science, 2014, , 320-340.	1.0	4
9149	Big Data Mining or Turning Data Mining into Predictive Analytics from Large-Scale 3Vs Data: The Future Challenge for Knowledge Discovery. Lecture Notes in Computer Science, 2014, , 4-8.	1.0	10
9150	Fields as a Generic Data Type for Big Spatial Data. Lecture Notes in Computer Science, 2014, , 159-172.	1.0	10
9151	μ-Controlled-Replicate: An Improved Controlled-Replicate Algorithm for Multi-way Spatial Join Processing on Map-Reduce. Lecture Notes in Computer Science, 2014, , 278-293.	1.0	3
9152	Stateless Stabilization Bootstrap (Extended Abstract). Lecture Notes in Computer Science, 2014, , 180-194.	1.0	2
9153	Populating Entity Name Systems for Big Data Integration. Lecture Notes in Computer Science, 2014, , 521-528.	1.0	5
9155	Exploratory Professional Search through Semantic Post-Analysis of Search Results. Lecture Notes in Computer Science, 2014, , 166-192.	1.0	6
9158	Towards Efficient Power Management in MapReduce: Investigation of CPU-Frequencies Scaling on Power Efficiency in Hadoop. Lecture Notes in Computer Science, 2014, , 147-164.	1.0	7
9159	Power-Efficient Assignment of Virtual Machines to Physical Machines. Lecture Notes in Computer Science, 2014, , 71-88.	1.0	3
9160	Processing Spatio-temporal Data On Map-Reduce. Lecture Notes in Computer Science, 2014, , 57-59.	1.0	3
9161	Two Stages Query Processing Optimization Based on ELM in the Cloud. Proceedings in Adaptation, Learning and Optimization, 2015, , 91-102.	1.5	1
9162	Map Matching for Taxi GPS Data with Extreme Learning Machine. Lecture Notes in Computer Science, 2014, , 447-460.	1.0	3
9163	Disruptive Innovation: Large Scale Multimedia Data Mining. , 2015, , 3-28.		1
9164	Can Road Traffic Volume Information Improve Partitioning for Distributed SUMO?. Lecture Notes in Mobility, 2015, , 61-74.	0.2	2

#	ARTICLE	IF	CITATIONS
9165	Discussion of BigBench: A Proposed Industry Standard Performance Benchmark for Big Data. Lecture Notes in Computer Science, 2015, , 44-63.	1.0	17
9167	Spatial Join with Hadoop. , 2017, , 2032-2036.		3
9169	Code Mobility Meets Self-organisation: A Higher-Order Calculus of Computational Fields. Lecture Notes in Computer Science, 2015, , 113-128.	1.0	27
9170	Ontology-Based Integrated Monitoring of Hadoop Clusters in Industrial Environments with OPC UA and RESTful Web Services. Communications in Computer and Information Science, 2015, , 162-171.	0.4	3
9171	Big Data Processing by Volunteer Computing Supported by Intelligent Agents. Lecture Notes in Computer Science, 2015, , 268-278.	1.0	2
9172	High-Performance Peer-to-Peer Systems. Computer Communications and Networks, 2015, , 317-337.	0.8	1
9174	Optimizing the Data-Process Relationship for Fast Mining of Frequent Itemsets in MapReduce. Lecture Notes in Computer Science, 2015, , 217-231.	1.0	3
9175	Rule-Based Exploration of Structured Data in the Browser. Lecture Notes in Computer Science, 2015, , 161-175.	1.0	2
9176	Human Migration and Spatial Synchrony: Spatial Patterns in Temporal Trends. , 2016, , 325-351.		1
9178	An Efficient Distributed Index for Geospatial Databases. Lecture Notes in Computer Science, 2015, , 28-42.	1.0	9
9179	Cross-Checking Data Sources in MapReduce. Communications in Computer and Information Science, 2015, , 165-174.	0.4	1
9180	CLUS: Parallel Subspace Clustering Algorithm on Spark. Communications in Computer and Information Science, 2015, , 175-185.	0.4	12
9181	Massively Parallel Unsupervised Feature Selection on Spark. Communications in Computer and Information Science, 2015, , 186-196.	0.4	2
9182	NPEPE: Massive Natural Computing Engine for Optimally Solving NP-complete Problems in Big Data Scenarios. Communications in Computer and Information Science, 2015, , 207-217.	0.4	9
9183	A New Link Prediction Algorithm Based on Local Links. Lecture Notes in Computer Science, 2015, , 16-28.	1.0	7
9184	Mining Personal Interests of Microbloggers Based on Free Tags in SINA Weibo. Lecture Notes in Computer Science, 2015, , 79-87.	1.0	2
9185	Research on Multiple Files Input Programming Method Based on MapReduce. Lecture Notes in Computer Science, 2015, , 336-342.	1.0	2
9186	A Balanced Vertex Cut Partition Method in Distributed Graph Computing. Lecture Notes in Computer Science, 2015, , 43-54.	1.0	2

#	ARTICLE	IF	CITATIONS
9187	The Dinamica Virtual Machine for Geosciences. Lecture Notes in Computer Science, 2015, , 44-58.	1.0	1
9188	Cloud-Based Intelligent Transportation Systems Using Model Predictive Control. Lecture Notes in Computer Science, 2015, , 464-477.	1.0	14
9189	What If Mixing Technologies for Big Data Mining and Queries Optimization. Lecture Notes in Computer Science, 2015, , 619-627.	1.0	2
9190	Governance of Societies. , 2016, , 75-115.		2
9191	CSRA: An Efficient Resource Allocation Algorithm in MapReduce Considering Data Skewness. Lecture Notes in Computer Science, 2015, , 651-662.	1.0	3
9192	An Ensemble Matchers Based Rank Aggregation Method for Taxonomy Matching. Lecture Notes in Computer Science, 2015, , 190-202.	1.0	2
9193	Multi-agent Systems Meet Aggregate Programming: Towards a Notion of Aggregate Plan. Lecture Notes in Computer Science, 2015, , 49-64.	1.0	3
9194	Inferring Software Behavioral Models with MapReduce. Lecture Notes in Computer Science, 2015, , 135-149.	1.0	3
9195	TrustedMR: A Trusted MapReduce System Based on Tamper Resistance Hardware. Lecture Notes in Computer Science, 2015, , 38-56.	1.0	3
9197	DFIS: A Scalable Distributed Fingerprint Identification System. Lecture Notes in Computer Science, 2015, , 162-175.	1.0	2
9198	Discovery of Frequent Episodes in Event Logs. Lecture Notes in Business Information Processing, 2015, , 1-31.	0.8	22
9199	Detecting and Dismantling Composite Visualizations in the Scientific Literature. Lecture Notes in Computer Science, 2015, , 247-266.	1.0	5
9201	You Are How You Query: Deriving Behavioral Fingerprints from DNS Traffic. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2015, , 348-366.	0.2	13
9202	Achieving Application-Level Utility Max-Min Fairness of Bandwidth Allocation in Datacenter Networks. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2016, , 36-46.	0.2	2
9203	Temporal Event Tracing on Big Healthcare Data Analytics. International Series on Computer Entertainment and Media Technology, 2016, , 95-108.	0.7	3
9205	Aggregate Programming: From Foundations to Applications. Lecture Notes in Computer Science, 2016, , 233-260.	1.0	5
9206	The Design of the Efficient Theta-Join in Map-Reduce Environment. Communications in Computer and Information Science, 2016, , 204-215.	0.4	1
9207	Data Dictionary Extraction for Robust Emergency Detection. Smart Innovation, Systems and Technologies, 2016, , 25-37.	0.5	1

#	ARTICLE	IF	CITATIONS
9210	Finding Electric Energy Consumption Patterns in Big Time Series Data. <i>Advances in Intelligent Systems and Computing</i> , 2016, , 231-238.	0.5	17
9212	Profiting from Several Recommendation Algorithms Using a Scalable Approach. <i>Studies in Computational Intelligence</i> , 2017, , 357-375.	0.7	1
9213	Automated Spark Clusters Deployment for Big Data with Standalone Applications Integration. <i>Lecture Notes in Computer Science</i> , 2016, , 150-159.	1.0	1
9214	Local Exceptionality Detection on Social Interaction Networks. <i>Lecture Notes in Computer Science</i> , 2016, , 298-302.	1.0	6
9215	Scalable and Fast Top-k Most Similar Trajectories Search Using MapReduce In-Memory. <i>Lecture Notes in Computer Science</i> , 2016, , 228-241.	1.0	5
9216	Walmart Online Grocery Personalization: Behavioral Insights and Basket Recommendations. <i>Lecture Notes in Computer Science</i> , 2016, , 49-64.	1.0	14
9217	Dynamic Topic-Based Sentiment Analysis of Large-Scale Online News. <i>Lecture Notes in Computer Science</i> , 2016, , 3-18.	1.0	9
9218	Comparison and Improvement of Hadoop MapReduce Performance Prediction Models in the Private Cloud. <i>Lecture Notes in Computer Science</i> , 2016, , 77-91.	1.0	4
9219	A Synthesized Urban Science in the Context of Big Data and Cyberinfrastructure. <i>Advances in Geographic Information Science</i> , 2018, , 435-448.	0.3	3
9220	How Can Quality Awareness Support Rapid Software Development? – A Research Preview. <i>Lecture Notes in Computer Science</i> , 2017, , 167-173.	1.0	25
9221	Scalability of a Genomic Data Analysis in the BioTest Platform. <i>Lecture Notes in Computer Science</i> , 2017, , 741-752.	1.0	4
9222	The Rise of Cloud Computing in the Era of Emerging Networked Society. <i>Computer Communications and Networks</i> , 2017, , 3-25.	0.8	7
9223	Rough Sets in Machine Learning: A Review. <i>Studies in Computational Intelligence</i> , 2017, , 87-118.	0.7	17
9225	An Efficient Bulk Loading Approach of Secondary Index in Distributed Log-Structured Data Stores. <i>Lecture Notes in Computer Science</i> , 2017, , 87-102.	1.0	5
9226	An Efficient Implementation of the Algorithm by LukÅŕÅŕi et al. on Hadoop. <i>Lecture Notes in Computer Science</i> , 2017, , 475-489.	1.0	6
9227	Cloud-Based Whole Slide Image Analysis Using MapReduce. <i>Lecture Notes in Computer Science</i> , 2017, , 62-77.	1.0	3
9228	Scalable Forecasting Techniques Applied to Big Electricity Time Series. <i>Lecture Notes in Computer Science</i> , 2017, , 165-175.	1.0	9
9230	On Scheduling Coflows. <i>Lecture Notes in Computer Science</i> , 2017, , 13-24.	1.0	10

#	ARTICLE	IF	CITATIONS
9231	Massively Distributed Environments and Closed Itemset Mining: The DCIM Approach. Lecture Notes in Computer Science, 2017, , 231-246.	1.0	1
9232	Adaptive Cheat Detection in Decentralized Volunteer Computing with Untrusted Nodes. Lecture Notes in Computer Science, 2017, , 192-205.	1.0	2
9233	A Bayes Classifier-Based OVFD Algorithm for Massive Stream Data Mining on Big Data Platform. Advances in Intelligent Systems and Computing, 2018, , 537-546.	0.5	1
9234	TeslaML: Steering Machine Learning Automatically in Tencent. Lecture Notes in Computer Science, 2017, , 313-318.	1.0	1
9235	MapReduce-Based Complex Big Data Analytics over Uncertain and Imprecise Social Networks. Lecture Notes in Computer Science, 2017, , 130-145.	1.0	21
9236	TARDIS: Optimal Execution of Scientific Workflows in Apache Spark. Lecture Notes in Computer Science, 2017, , 74-87.	1.0	2
9237	MiSeRe-Hadoop: A Large-Scale Robust Sequential Classification Rules Mining Framework. Lecture Notes in Computer Science, 2017, , 105-119.	1.0	3
9238	MapFIM: Memory Aware Parallelized Frequent Itemset Mining in Very Large Datasets. Lecture Notes in Computer Science, 2017, , 478-495.	1.0	4
9239	Migration from Web Services to Cloud Services. Lecture Notes in Computer Science, 2017, , 179-192.	1.0	3
9240	Inferring the State of the Aerosphere from Weather Radar. , 2017, , 311-343.		4
9241	Comparing MapReduce-Based k-NN Similarity Joins on Hadoop for High-Dimensional Data. Lecture Notes in Computer Science, 2017, , 63-75.	1.0	7
9242	Augmenting a Data Warehousing Curriculum with Emerging Big Data Technologies. Communications in Computer and Information Science, 2017, , 128-143.	0.4	2
9243	Feature Selection in High-Dimensional Dataset Using MapReduce. Communications in Computer and Information Science, 2018, , 101-115.	0.4	7
9244	Scalable Big Data Privacy with MapReduce. , 2019, , 1454-1462.		2
9245	Apache Flink. , 2019, , 51-58.		11
9246	Hadoop. Computer Communications and Networks, 2018, , 29-53.	0.8	3
9247	Implementing MapReduce with MSVL. Lecture Notes in Computer Science, 2018, , 148-167.	1.0	1
9248	Polygon Queries for Convex Hulls of Points. Lecture Notes in Computer Science, 2018, , 143-155.	1.0	1

#	ARTICLE	IF	CITATIONS
9249	StarMR: An Efficient Star-Decomposition Based Query Processor for SPARQL Basic Graph Patterns Using MapReduce. Lecture Notes in Computer Science, 2018, , 415-430.	1.0	3
9250	Communication-Efficient Distributed Optimization of Self-concordant Empirical Loss. Lecture Notes in Mathematics, 2018, , 289-341.	0.1	7
9254	Criteria to Compare Cloud Computing with Current Database Technology. Lecture Notes in Computer Science, 2008, , 114-126.	1.0	13
9255	Integrated Retrieval from Web of Documents and Data. Studies in Computational Intelligence, 2009, , 25-48.	0.7	3
9256	Data Parallel Bin-Based Indexing for Answering Queries on Multi-core Architectures. Lecture Notes in Computer Science, 2009, , 110-129.	1.0	5
9257	Beyond Conventional Data Warehousing – Massively Parallel Data Processing with Greenplum Database. Lecture Notes in Business Information Processing, 2009, , 89-96.	0.8	16
9258	Evaluating SPLASH-2 Applications Using MapReduce. Lecture Notes in Computer Science, 2009, , 452-464.	1.0	3
9259	Towards Efficient MapReduce Using MPI. Lecture Notes in Computer Science, 2009, , 240-249.	1.0	53
9260	A Skeletal Parallel Framework with Fusion Optimizer for GPGPU Programming. Lecture Notes in Computer Science, 2009, , 79-94.	1.0	23
9261	Programming Clouds. Lecture Notes in Computer Science, 2010, , 1-9.	1.0	7
9262	A Proposal for User-Defined Reductions in OpenMP. Lecture Notes in Computer Science, 2010, , 43-55.	1.0	5
9263	Hierarchical Link Analysis for Ranking Web Data. Lecture Notes in Computer Science, 2010, , 225-239.	1.0	34
9264	High Throughput Data-Compression for Cloud Storage. , 2010, , 1-12.		34
9265	The High-Activity Parallel Implementation of Data Preprocessing Based on MapReduce. Lecture Notes in Computer Science, 2010, , 646-654.	1.0	9
9266	Massive Graph Management for the Web and Web 2.0. Studies in Computational Intelligence, 2011, , 19-58.	0.7	5
9268	CnC-CUDA: Declarative Programming for GPUs. Lecture Notes in Computer Science, 2011, , 230-245.	1.0	12
9270	Distributed Attack Detection Using Agilis. , 2012, , 157-174.		1
9271	Scalability Challenges in Web Search Engines. The Kluwer International Series on Information Retrieval, 2011, , 27-50.	1.0	21

#	ARTICLE	IF	CITATIONS
9272	An Intermediate Algebra for Optimizing RDF Graph Pattern Matching on MapReduce. Lecture Notes in Computer Science, 2011, , 46-61.	1.0	53
9275	Rule-Based Distributed and Agent Systems. Lecture Notes in Computer Science, 2011, , 3-28.	1.0	12
9276	Database Foundations for Scalable RDF Processing. Lecture Notes in Computer Science, 2011, , 202-249.	1.0	18
9277	On the Benefits of Transparent Compression for Cost-Effective Cloud Data Storage. Lecture Notes in Computer Science, 2011, , 167-184.	1.0	15
9278	MapResolve. Lecture Notes in Computer Science, 2011, , 294-299.	1.0	10
9279	Comparing High Level MapReduce Query Languages. Lecture Notes in Computer Science, 2011, , 58-72.	1.0	31
9280	Cluster Computing, Recursion and Datalog. Lecture Notes in Computer Science, 2011, , 120-144.	1.0	6
9281	Parallelizing an Index Generator for Desktop Search. Lecture Notes in Computer Science, 2011, , 77-85.	1.0	1
9282	SpotMPI: A Framework for Auction-Based HPC Computing Using Amazon Spot Instances. Lecture Notes in Computer Science, 2011, , 109-120.	1.0	19
9283	Enhancing Query Support in HBase via an Extended Coprocessors Framework. Lecture Notes in Computer Science, 2011, , 75-87.	1.0	12
9284	An Approach for Processing Large and Non-uniform Media Objects on MapReduce-Based Clusters. Lecture Notes in Computer Science, 2011, , 172-181.	1.0	4
9285	High-Integrity MapReduce Computation in Cloud with Speculative Execution. Communications in Computer and Information Science, 2011, , 397-404.	0.4	6
9286	dipLODocus[RDF]â€”Short and Long-Tail RDF Analytics for Massive Webs of Data. Lecture Notes in Computer Science, 2011, , 778-793.	1.0	21
9287	Sorting, Searching, and Simulation in the MapReduce Framework. Lecture Notes in Computer Science, 2011, , 374-383.	1.0	114
9288	Loyalty-Based Resource Allocation Mechanism in Cloud Computing. Lecture Notes in Electrical Engineering, 2012, , 233-238.	0.3	4
9289	Confidant: Protecting OSN Data without Locking It Up. Lecture Notes in Computer Science, 2011, , 61-80.	1.0	18
9290	Resource Provisioning Framework for MapReduce Jobs with Performance Goals. Lecture Notes in Computer Science, 2011, , 165-186.	1.0	99
9291	Adaptive Parallelization of Queries to Data Providing Web Service Operations. Lecture Notes in Computer Science, 2012, , 49-69.	1.0	1

#	ARTICLE	IF	CITATIONS
9292	Generate, Test, and Aggregate. Lecture Notes in Computer Science, 2012, , 254-273.	1.0	11
9293	Pool-Based Distributed Evolutionary Algorithms Using an Object Database. Lecture Notes in Computer Science, 2012, , 446-455.	1.0	10
9294	Flex-GP: Genetic Programming on the Cloud. Lecture Notes in Computer Science, 2012, , 477-486.	1.0	23
9295	Bizard: An Online Multi-dimensional Data Analysis Visualization Tool. Lecture Notes in Computer Science, 2012, , 775-778.	1.0	2
9297	Scalable Sequence Similarity Search and Join in Main Memory on Multi-cores. Lecture Notes in Computer Science, 2012, , 13-22.	1.0	1
9298	A Panoramic View of 3G Data/Control-Plane Traffic: Mobile Device Perspective. Lecture Notes in Computer Science, 2012, , 318-330.	1.0	15
9299	PQL: A Purely-Declarative Java Extension for Parallel Programming. Lecture Notes in Computer Science, 2012, , 53-78.	1.0	6
9301	Eden â€“ Parallel Functional Programming with Haskell. Lecture Notes in Computer Science, 2012, , 142-206.	1.0	11
9302	Data Science and Distributed Intelligence: Recent Developments and Future Insights. Studies in Computational Intelligence, 2013, , 139-147.	0.7	2
9303	Enabling Cloud Interoperability with COMPSs. Lecture Notes in Computer Science, 2012, , 16-27.	1.0	12
9304	Cross-Language High Similarity Search Using a Conceptual Thesaurus. Lecture Notes in Computer Science, 2012, , 67-75.	1.0	15
9306	BOINC-MR: MapReduce in a Volunteer Environment. Lecture Notes in Computer Science, 2012, , 425-432.	1.0	3
9307	Automated and Transparent Model Fragmentation for Persisting Large Models. Lecture Notes in Computer Science, 2012, , 102-118.	1.0	18
9308	A Cloud Architecture with an Efficient Scheduling Technique. Lecture Notes in Computer Science, 2012, , 381-388.	1.0	1
9309	Exploring the Meaning behind Twitter Hashtags through Clustering. Lecture Notes in Business Information Processing, 2012, , 231-242.	0.8	14
9310	Applying MapReduce Framework to Peer-to-Peer Computing Applications. Lecture Notes in Computer Science, 2012, , 69-78.	1.0	4
9311	Computing Scientometrics in Large-Scale Academic Search Engines with MapReduce. Lecture Notes in Computer Science, 2012, , 609-623.	1.0	5
9312	CloudPack. Lecture Notes in Computer Science, 2012, , 374-393.	1.0	8

#	ARTICLE	IF	CITATIONS
9313	Budget Optimization for Online Campaigns with Positive Carryover Effects. Lecture Notes in Computer Science, 2012, , 86-99.	1.0	8
9314	CrowdLang: A Programming Language for the Systematic Exploration of Human Computation Systems. Lecture Notes in Computer Science, 2012, , 124-137.	1.0	38
9315	A Hybrid Anomaly Detection Framework in Cloud Computing Using One-Class and Two-Class Support Vector Machines. Lecture Notes in Computer Science, 2012, , 726-738.	1.0	21
9316	Web-Scale Job Scheduling. Lecture Notes in Computer Science, 2013, , 1-15.	1.0	8
9318	Scalable SAPRQL Querying Processing on Large RDF Data in Cloud Computing Environment. Lecture Notes in Computer Science, 2013, , 631-646.	1.0	4
9320	A Declarative Pipeline Language for Complex Data Analysis. Lecture Notes in Computer Science, 2013, , 17-34.	1.0	3
9321	When to Reach for the Cloud: Using Parallel Hardware for Link Discovery. Lecture Notes in Computer Science, 2013, , 275-289.	1.0	11
9322	Formal Concept Analysis via Atomic Priming. Lecture Notes in Computer Science, 2013, , 92-108.	1.0	8
9323	Fast Multi-fields Query Processing in Bigtable Based Cloud Systems. Lecture Notes in Computer Science, 2013, , 142-154.	1.0	1
9324	USTO.RE: A Private Cloud Storage Software System. Lecture Notes in Computer Science, 2013, , 452-466.	1.0	4
9326	Multi-level Parallel Query Execution Framework for CPU and GPU. Lecture Notes in Computer Science, 2013, , 330-343.	1.0	11
9327	A Benchmark for Multidimensional Statistical Data. Lecture Notes in Computer Science, 2013, , 358-371.	1.0	2
9328	A Novel Search Engine to Uncover Potential Victims for APT Investigations. Lecture Notes in Computer Science, 2013, , 405-416.	1.0	11
9329	MaRDiGraS: Simplified Building of Reachability Graphs on Large Clusters. Lecture Notes in Computer Science, 2013, , 83-95.	1.0	13
9330	Novel Client-Cloud Architecture for Scalable Instance-Intensive Workflow Systems. Lecture Notes in Computer Science, 2013, , 270-284.	1.0	6
9331	Fast Approximate Attribute Reduction with MapReduce. Lecture Notes in Computer Science, 2013, , 271-278.	1.0	4
9332	Diversity-Driven Widening. Lecture Notes in Computer Science, 2013, , 223-236.	1.0	10
9333	Quality Control for Crowdsourcing with Spatial and Temporal Distribution. Lecture Notes in Computer Science, 2013, , 169-182.	1.0	4

#	ARTICLE	IF	CITATIONS
9335	Virtualizing Platforms. Lecture Notes in Computer Science, 2013, , 203-226.	1.0	3
9337	Collective Corpus Weighting and Phrase Scoring for SMT Using Graph-Based Random Walk. Communications in Computer and Information Science, 2013, , 176-187.	0.4	1
9338	Big Data“Technologies and Potential. , 2014, , 107-124.		3
9339	Research of Scheduling Strategy Based on Fault Tolerance in Hadoop Platform. Communications in Computer and Information Science, 2013, , 509-517.	0.4	1
9340	Assured Cloud-Based Data Analysis with ClusterBFT. Lecture Notes in Computer Science, 2013, , 82-102.	1.0	7
9341	The Case for Multi-Engine Data Analytics. Lecture Notes in Computer Science, 2014, , 406-415.	1.0	4
9342	Implementing Graph Transformations in the Bulk“Synchronous“Parallel Model. Lecture Notes in Computer Science, 2014, , 325-339.	1.0	16
9343	Towards Reactive Programming for Object-Oriented Applications. Lecture Notes in Computer Science, 2014, , 227-261.	1.0	10
9344	K-mulus: Strategies for BLAST in the Cloud. Lecture Notes in Computer Science, 2014, , 237-246.	1.0	2
9345	Anomaly Detection and Structural Analysis in Industrial Production Environments. , 2017, , 91-95.		10
9346	Platform for Simulation and Improvement of Swarm Behavior in Changing Environments. Lecture Notes in Computer Science, 2014, , 121-129.	1.0	3
9347	Scalable Nonnegative Matrix Factorization with Block-wise Updates. Lecture Notes in Computer Science, 2014, , 337-352.	1.0	21
9348	Near-Constant-Time Distributed Algorithms on a Congested Clique. Lecture Notes in Computer Science, 2014, , 514-530.	1.0	21
9349	How to Store and Process Big Data: Are Today“s Databases Sufficient?. Lecture Notes in Computer Science, 2014, , 5-10.	1.0	3
9350	Parallel Ordinal Decision Tree Algorithm and Its Implementation in Framework of MapReduce. Communications in Computer and Information Science, 2014, , 241-251.	0.4	6
9351	Big Data“Technologies and Potential. , 2014, , 107-124.		9
9352	Efficient Processing of Multi-way Joins Using MapReduce. Communications in Computer and Information Science, 2015, , 73-80.	0.4	2
9354	Zukunft der globalen Geod“Asie und Fernerkundung aus Sicht des Deutschen GeoForschungsZentrum (GFZ), Potsdam. , 2017, , 443-497.		1

#	ARTICLE	IF	CITATIONS
9355	Efficient Level-Based Top-Down Data Cube Computation Using MapReduce. Lecture Notes in Computer Science, 2015, , 1-19.	1.0	5
9356	On the Computational Complexity of MapReduce. Lecture Notes in Computer Science, 2015, , 1-15.	1.0	10
9357	Pay-as-you-go Configuration of Entity Resolution. Lecture Notes in Computer Science, 2016, , 40-65.	1.0	5
9359	Apache Hadoop Yarn MapReduce Job Classification Based on CPU Utilization and Performance Evaluation on Multi-cluster Heterogeneous Environment. Advances in Intelligent Systems and Computing, 2016, , 35-44.	0.5	1
9360	An Optimization of the Delay Scheduling Algorithm for Real-Time Video Stream Processing. Lecture Notes in Electrical Engineering, 2016, , 173-183.	0.3	2
9361	A Density-Aware Similarity Join Query Processing Algorithm on MapReduce. Lecture Notes in Electrical Engineering, 2016, , 469-475.	0.3	2
9362	Quantum-Behaved Particle Swarm Optimization Using MapReduce. Communications in Computer and Information Science, 2016, , 173-178.	0.4	3
9363	Grouping-Aware Data Placement in HDFS for Data-Intensive Applications Based on Graph Clustering. Advances in Intelligent Systems and Computing, 2018, , 21-31.	0.5	2
9364	Using T-Drive and BerlinMod in Parallel SECONDO for Performance Evaluation of Geospatial Big Data Processing. Advances in Geographic Information Science, 2017, , 3-19.	0.3	1
9365	Virtual Networking with Azure for Hybrid Cloud Computing in Aneka. , 2017, , 93-114.		2
9366	A Privacy Preserving Platform for MapReduce. Communications in Computer and Information Science, 2017, , 88-99.	0.4	7
9367	Big Data Analysis in Cloud and Machine Learning. Studies in Big Data, 2019, , 51-85.	0.8	9
9368	A Vertex-Centric Graph Simulation Algorithm for Large Graphs. Communications in Computer and Information Science, 2018, , 238-254.	0.4	1
9369	A Comparison of Apache Spark Supervised Machine Learning Algorithms for DNA Splicing Site Prediction. Smart Innovation, Systems and Technologies, 2020, , 133-143.	0.5	3
9370	An Efficient Honey Bee Approach for Load Adjusting in Cloud Environment. Lecture Notes in Networks and Systems, 2020, , 3-12.	0.5	1
9371	Automatic Generation and Assessment of Student Assignments for Parallel Programming Learning. Communications in Computer and Information Science, 2020, , 184-194.	0.4	3
9372	Analysis and Forecast of Heart Syndrome by Intelligent Retrieval Approach. Lecture Notes in Networks and Systems, 2020, , 507-515.	0.5	1
9373	Process Model for Fog Data Analytics for IoT Applications. Studies in Big Data, 2020, , 175-198.	0.8	2

#	ARTICLE	IF	CITATIONS
9374	Thresholding Strategies for Deep Learning with Highly Imbalanced Big Data. <i>Advances in Intelligent Systems and Computing</i> , 2021, , 199-227.	0.5	12
9375	Big Data Computing and Mining in a Smart World. <i>Advances in Intelligent Systems and Computing</i> , 2021, , 15-27.	0.5	5
9376	A Study on Big Data Processing Frameworks: Spark and Storm. <i>Smart Innovation, Systems and Technologies</i> , 2020, , 415-424.	0.5	2
9377	An efficient distributed caching for accessing small files in HDFS. <i>Cluster Computing</i> , 2017, 20, 3579-3592.	3.5	10
9378	An Analysis of Public Clouds Elasticity in the Execution of Scientific Applications: a Survey. , 2016, 14, 193.		1
9379	Big Data: A Survey. , 0, .		1
9380	Multi-objective scheduling of MapReduce jobs in big data processing. <i>Multimedia Tools and Applications</i> , 2018, 77, 9979-9994.	2.6	28
9381	Online traffic-aware linked VM placement in cloud data centers. <i>Science China Information Sciences</i> , 2020, 63, 1.	2.7	5
9382	Fast execution of RDF queries using Apache Hadoop. <i>Advances in Computers</i> , 2020, 119, 1-33.	1.2	2
9383	A parallel model for SQL astronomical databases based on solid state storage. <i>Astronomy and Computing</i> , 2017, 20, 77-82.	0.8	1
9385	Towards privacy preserving AI based composition framework in edge networks using fully homomorphic encryption. <i>Engineering Applications of Artificial Intelligence</i> , 2020, 94, 103737.	4.3	32
9386	Mining maximal frequent patterns in transactional databases and dynamic data streams: A spark-based approach. <i>Information Sciences</i> , 2018, 432, 278-300.	4.0	53
9387	Aggregating and disaggregating packets with various sizes of payload in P4 switches at 100ÂGbps line rate. <i>Journal of Network and Computer Applications</i> , 2020, 165, 102676.	5.8	13
9388	Safe automated refactoring for intelligent parallelization of Java 8 streams. <i>Science of Computer Programming</i> , 2020, 195, 102476.	1.5	15
9391	Big Data over Networks. , 2016, , .		23
9392	Hadoop-based framework for big data analysis of synchronised harmonics in active distribution network. <i>IET Generation, Transmission and Distribution</i> , 2017, 11, 3930-3937.	1.4	19
9393	Fine-Grained Allocation Algorithm for Sharing Heterogeneous Resources in Data Center. <i>Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University</i> , 2020, 38, 589-595.	0.3	1
9394	MarDR: efficient MapReduce-based removal of duplicate DNA reads in the cloud. <i>Bioinformatics</i> , 2017, 33, 2762-2764.	1.8	12

#	ARTICLE	IF	CITATIONS
9406	Inferring Coflow Size Based on Broad Learning System in Data Center Network. , 2020, , .		2
9407	Serverless Straggler Mitigation using Error-Correcting Codes. , 2020, , .		4
9408	Conceptualization of social big data: Keynote. , 2017, , .		1
9409	A Fast PC Algorithm for High Dimensional Causal Discovery with Multi-Core PCs. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2019, 16, 1483-1495.	1.9	58
9411	Implementing Artificial Intelligence in Predicting Metrics for Characterizing Laser Propagation in Atmospheric Turbulence. Journal of Fluids Engineering, Transactions of the ASME, 2019, 141, .	0.8	3
9412	Rule Based Systems in a Distributed Environment: Survey. , 0, , .		5
9413	Parallel optimization of pixel purity index algorithm for massive hyperspectral images in cloud computing environment. Journal of Applied Remote Sensing, 2016, 10, 025024.	0.6	3
9415	Middleware for big data processing: test results. Physics of Particles and Nuclei Letters, 2017, 14, 1001-1007.	0.1	2
9416	Deterministic CUR for Improved Large-Scale Data Analysis: An Empirical Study. , 2012, , .		8
9417	Poster reception—Harnessing grid resources to enable the dynamic analysis of large astronomy datasets. , 2006, , .		3
9418	A domain-specific language for parallel and grid computing. , 2008, , .		5
9419	SIGPLAN programming language curriculum workshop. ACM SIGPLAN Notices, 2008, 43, 29-30.	0.2	2
9420	Estimating the impressionrank of web pages. , 2009, , .		14
9421	SLIPstream. , 2009, , .		18
9422	DEDUCE. , 2010, , .		37
9423	Centrality metric for dynamic networks. , 2010, , .		57
9424	Hybrid bulk synchronous parallelism library for clustered smp architectures. , 2010, , .		17
9425	ESQP. , 2010, , .		20

#	ARTICLE	IF	CITATIONS
9426	Comparing SQL and mapreduce to compute naive bayes in a single table scan. , 2010, , .		7
9427	Benchmarking cloud-based data management systems. , 2010, , .		35
9428	Towards extensible automatic image annotation with the bag-of-words approach. , 2010, , .		1
9429	Filtering harmful sentences based on three-word co-occurrence. , 2011, , .		4
9430	Scaling up machine learning. , 2011, , .		69
9431	Tarazu. ACM SIGPLAN Notices, 2012, 47, 61-74.	0.2	6
9432	Mapping a data-flow programming model onto heterogeneous platforms. ACM SIGPLAN Notices, 2012, 47, 61-70.	0.2	11
9433	Oolong. , 2012, , .		9
9434	Efficient task replication for fast response times in parallel computation. Performance Evaluation Review, 2014, 42, 599-600.	0.4	31
9435	SQLPhi. , 2014, , .		3
9436	Data-parallel finite-state machines. Computer Architecture News, 2014, 42, 529-542.	2.5	9
9437	Learning cloud computing and security through CloudSim simulation. , 2014, , .		1
9438	Modern Enterprises in the Bubble. Software Engineering Notes: an Informal Newsletter of the Special Interest Committee on Software Engineering / ACM, 2015, 40, 1-4.	0.5	21
9439	Declarative Networking. SIGMOD Record, 2014, 43, 5-16.	0.7	10
9440	A Survey on XML Fragmentation. SIGMOD Record, 2014, 43, 24-35.	0.7	8
9441	CheckCell. ACM SIGPLAN Notices, 2014, 49, 507-523.	0.2	17
9442	ASPIRE. ACM SIGPLAN Notices, 2014, 49, 861-878.	0.2	13
9443	FACADE. ACM SIGPLAN Notices, 2015, 50, 675-690.	0.2	10

#	ARTICLE	IF	CITATIONS
9444	Influence at Scale. , 2015, , .		36
9445	Generating performance portable code using rewrite rules: from high-level functional expressions to high-performance OpenCL code. , 2015, , .		84
9446	BwE. , 2015, , .		107
9447	Hopper. , 2015, , .		81
9448	Low Latency Geo-distributed Data Analytics. , 2015, , .		153
9449	Jupiter Rising. , 2015, , .		343
9450	FACADE. Computer Architecture News, 2015, 43, 675-690.	2.5	10
9451	CommGuard. Computer Architecture News, 2015, 43, 311-323.	2.5	2
9452	Special Issue on Performance and Resource Management in Big Data Applications. Performance Evaluation Review, 2015, 42, 2-2.	0.4	3
9453	Minimizing Interference and Maximizing Progress for Hadoop Virtual Machines. Performance Evaluation Review, 2015, 42, 62-71.	0.4	12
9454	Computable Bounds in Fork-Join Queueing Systems. Performance Evaluation Review, 2015, 43, 335-346.	0.4	11
9455	Cloud-Enabled Privacy-Preserving Truth Discovery in Crowd Sensing Systems. , 2015, , .		109
9456	SCANRAW. ACM Transactions on Database Systems, 2015, 40, 1-45.	1.5	20
9457	An iterative MapReduce framework for sports-based tweet clustering. , 2015, , .		3
9458	BwE. Computer Communication Review, 2015, 45, 1-14.	1.5	28
9459	Hopper. Computer Communication Review, 2015, 45, 379-392.	1.5	20
9460	Low Latency Geo-distributed Data Analytics. Computer Communication Review, 2015, 45, 421-434.	1.5	127
9461	Term-by-Term Query Auto-Completion for Mobile Search. , 2016, , .		5

#	ARTICLE	IF	CITATIONS
9462	Fast and Portable Locking for Multicore Architectures. ACM Transactions on Computer Systems, 2016, 33, 1-62.	0.6	16
9463	Resolve. , 2016, , .		17
9464	Distributed graph processing. , 2016, , .		2
9465	Virtualization technologies for the big data environment. , 2016, , .		1
9466	Generating performance portable code using rewrite rules: from high-level functional expressions to high-performance OpenCL code. ACM SIGPLAN Notices, 2015, 50, 205-217.	0.2	20
9467	Cartesian hoare logic for verifying k-safety properties. , 2016, , .		73
9468	MST in Log-Star Rounds of Congested Clique. , 2016, , .		43
9469	Contrasting Effects of Replication in Parallel Systems. Performance Evaluation Review, 2016, 44, 375-376.	0.4	5
9470	Ensemble based Parallel k means using Map Reduce for Aspect Based Summarization. , 2016, , .		6
9471	Cartesian hoare logic for verifying k-safety properties. ACM SIGPLAN Notices, 2016, 51, 57-69.	0.2	12
9472	LogMine. , 2016, , .		121
9473	Scaling Factorization Machines with Parameter Server. , 2016, , .		9
9474	Massively parallel algorithms for computing TIN DEMs and contour trees for large terrains. , 2016, , .		4
9475	User identification in cyber-physical space. , 2016, , .		11
9476	Enhanced fraud detection as a service supporting merchant-specific runtime customization. , 2017, , .		3
9477	CATS. , 2017, , .		2
9478	Fusing effectful comprehensions. , 2017, , .		4
9480	Fusing effectful comprehensions. ACM SIGPLAN Notices, 2017, 52, 17-32.	0.2	3

#	ARTICLE	IF	CITATIONS
9481	StreamQRE: modular specification and efficient evaluation of quantitative queries over streaming data. ACM SIGPLAN Notices, 2017, 52, 693-708.	0.2	11
9482	Effective Straggler Mitigation. Performance Evaluation Review, 2017, 45, 12-14.	0.4	30
9483	Automatically Distributing Eulerian and Hybrid Fluid Simulations in the Cloud. ACM Transactions on Graphics, 2018, 37, 1-14.	4.9	6
9484	Straggler Mitigation by Delayed Relaunch of Tasks. Performance Evaluation Review, 2018, 45, 224-231.	0.4	18
9485	IRIS. , 2018, , .		6
9486	Congested Clique Algorithms for the Minimum Cut Problem. , 2018, , .		9
9487	Massively Parallel Computation of Matching and MIS in Sparse Graphs. , 2019, , .		20
9488	Skyway. ACM SIGPLAN Notices, 2018, 53, 56-69.	0.2	10
9489	Dirichlet process mixture models made scalable and effective by means of massive distribution. , 2019, , .		4
9490	Data quality rule definition and discovery. , 2019, , .		1
9491	Machine learning and probabilistic data cleaning. , 2019, , .		1
9492	Beyond Analytics: The Evolution of Stream Processing Systems. , 2020, , .		31
9493	Causality-Guided Adaptive Interventional Debugging. , 2020, , .		12
9494	Probabilistic Data Structures in Adversarial Environments. , 2019, , .		29
9495	Gerenuk. , 2019, , .		13
9496	Lineage stash. , 2019, , .		23
9497	Connected Components on a PRAM in Log Diameter Time. , 2020, , .		5
9498	A Massively Parallel Algorithm for Minimum Weight Vertex Cover. , 2020, , .		5

#	ARTICLE	IF	CITATIONS
9499	Optimal Resource Allocation for Elastic and Inelastic Jobs. , 2020, , .		6
9500	Graph Sparsification for Derandomizing Massively Parallel Computation with Low Space. , 2020, , .		5
9501	Parallel approximate undirected shortest paths via low hop emulators. , 2020, , .		32
9502	Modeling of Request Cloning in Cloud Server Systems using Processor Sharing. , 2020, , .		3
9503	Pando. , 2019, , .		8
9504	Scalable Deep Learning on Distributed Infrastructures. ACM Computing Surveys, 2021, 53, 1-37.	16.1	99
9505	The Case for a Network Fast Path to the CPU. , 2019, , .		11
9506	PRISM. , 2019, , .		4
9507	Fundamental Limits of Approximate Gradient Coding. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2019, 3, 1-22.	1.4	29
9508	Rateless Codes for Near-Perfect Load Balancing in Distributed Matrix-Vector Multiplication. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2019, 3, 1-40.	1.4	51
9509	Big data challenges and achievements. , 2019, , .		32
9510	Coverage Pattern Mining Based on MapReduce. , 2020, , .		4
9511	PacketScope. , 2020, , .		17
9512	Fleet. , 2020, , .		24
9513	FlatStore: An Efficient Log-Structured Key-Value Storage Engine for Persistent Memory. , 2020, , .		101
9514	Remotely Controlled Manufacturing. , 2020, , .		3
9515	Blocking and Filtering Techniques for Entity Resolution. ACM Computing Surveys, 2021, 53, 1-42.	16.1	102
9516	A Survey on Automatic Parameter Tuning for Big Data Processing Systems. ACM Computing Surveys, 2021, 53, 1-37.	16.1	52

#	ARTICLE	IF	CITATIONS
9517	GiantVM. , 2020, , .		11
9518	Simple, Deterministic, Constant-Round Coloring in the Congested Clique. , 2020, , .		12
9519	Big Data Solutions Proposed for Cluster Computing Systems Challenges. , 2020, , .		9
9520	Internal Evaluation of Unsupervised Outlier Detection. ACM Transactions on Knowledge Discovery From Data, 2020, 14, 1-42.	2.5	14
9521	Packing R-trees with Space-filling Curves. ACM Transactions on Database Systems, 2020, 45, 1-47.	1.5	15
9522	DQEMU: A Scalable Emulator with Retargetable DBT on Distributed Platforms. , 2020, , .		3
9523	Fast Distributed k -NN Graph Construction Using Auto-tuned Locality-sensitive Hashing. ACM Transactions on Intelligent Systems and Technology, 2020, 11, 1-18.	2.9	7
9524	Particle. , 2020, , .		17
9525	Kappa. , 2020, , .		34
9526	Wukong. , 2020, , .		56
9527	Influence-based provenance for dataflow applications with taint propagation. , 2020, , .		7
9528	DiffStream: differential output testing for stream processing programs. , 2020, 4, 1-29.		12
9529	A modular cost analysis for probabilistic programs. , 2020, 4, 1-30.		15
9530	Optimal Load Balancing with Locality Constraints. Proceedings of the ACM on Measurement and Analysis of Computing Systems, 2020, 4, 1-37.	1.4	11
9531	An Evaluation of Serverless Data Processing Frameworks. , 2020, , .		5
9532	A Dynamic Users' Interest Discovery Model with Distributed Inference Algorithm. International Journal of Distributed Sensor Networks, 2014, 10, 280892.	1.3	9
9533	Big Data Visualization Collaborative Filtering Algorithm Based on RHadoop. International Journal of Distributed Sensor Networks, 2015, 2015, 1-9.	1.3	16
9534	A Group Mining Method for Big Data on Distributed Vehicle Trajectories in WAN. International Journal of Distributed Sensor Networks, 2015, 11, 756107.	1.3	4

#	ARTICLE	IF	CITATIONS
9535	Big Data Processing with Hadoop-MapReduce in Cloud Systems. International Journal of Cloud Computing and Services Science (IJ-CLOSER), 2012, 2, .	0.6	18
9536	Parallel Implementation of Classification Algorithms Based on Cloud Computing Environment. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2012, 10, .	0.1	20
9538	Comparisons Between MongoDB and MS-SQL Databases on the TWC Website. American Journal of Software Engineering and Applications, 2015, 4, 35.	0.1	8
9539	Processing Intensive Full-Waveform Aerial Laser Scanning Matlab Jobs through Condor. Internet of Things and Cloud Computing, 2013, 1, 5.	0.6	2
9540	Using weaker consistency models with monitoring and recovery for improving performance of key-value stores. Journal of the Brazilian Computer Society, 2019, 25, .	0.8	2
9541	A survey on data analysis on large-Scale wireless networks: online stream processing, trends, and challenges. Journal of Internet Services and Applications, 2020, 11, .	1.6	17
9542	An overview of recent distributed algorithms for learning fuzzy models in Big Data classification. Journal of Big Data, 2020, 7, .	6.9	15
9543	Efficient verification of parallel matrix multiplication in public cloud: the MapReduce case. Journal of Big Data, 2020, 7, .	6.9	2
9544	Comparison of sort algorithms in Hadoop and PCJ. Journal of Big Data, 2020, 7, .	6.9	3
9545	A comprehensive performance analysis of Apache Hadoop and Apache Spark for large scale data sets using HiBench. Journal of Big Data, 2020, 7, .	6.9	38
9546	Big Data Architecture for Climate Change and Disease Dynamics. , 2016, , 303-333.		3
9548	Toward Evolving Knowledge Ecosystems for Big Data Understanding. , 2013, , 3-55.		6
9550	Exploiting Big VGI to Improve Routing and Navigation Services. , 2014, , 191-206.		7
9551	Distributed and Parallel Computing. , 2014, , 15-44.		3
9552	MapReduce Family of Large-Scale Data-Processing Systems. , 2014, , 39-106.		5
9554	Big Data Research Analysis Based on Bibliometrics. Hans Journal of Data Mining, 2016, 06, 125-137.	0.2	3
9556	AutoAdmin: Automatic and Dynamic Resource Reservation Admission Control in Hadoop YARN Clusters. Scalable Computing, 2018, 19, .	0.7	3
9557	The Optimization Potential of Volunteer Computing for Compute or Data Intensive Applications. Journal of Communications, 2019, , 971-979.	1.3	3

#	ARTICLE	IF	CITATIONS
9558	An Approach for Processing Underground Spatial-Temporal Data by Cloud Computing. <i>Journal of Automation and Control Engineering</i> , 2013, 1, 164-165.	0.3	1
9559	A Parallel Oversampling Algorithm Based on NRSBoundary-SMOTE. <i>Journal of Information and Computational Science</i> , 2014, 11, 4655-4665.	0.1	15
9560	Task Scheduling Based on Multi-objective Genetic Algorithm in Cloud Computing. <i>Journal of Information and Computational Science</i> , 2015, 12, 1429-1438.	0.1	9
9561	Data Migration from Grid to Cloud Computing. <i>Applied Mathematics and Information Sciences</i> , 2013, 7, 399-406.	0.7	3
9562	Software Design and Implementation for MapReduce across Distributed Data Centers. <i>Applied Mathematics and Information Sciences</i> , 2013, 7, 85-90.	0.7	11
9564	SOA enabled ELTA: approach in designing business intelligence solutions in Era of Big Data. , 2015, 3, 49-63.		5
9565	Big Data Components for Business Process Optimization. <i>InformatikÄf EconomicÄf</i> , 2016, 20, 72-78.	0.2	2
9566	QUERY RESPONSE TIME COMPARISON NOSQLDB MONGODB WITH SQLDB ORACLE. <i>JUTI: Jurnal Ilmiah Teknologi Informasi</i> , 2015, 13, 95.	0.1	3
9567	Improved fair Scheduling Algorithm for Hadoop Clustering. <i>Oriental Journal of Computer Science and Technology</i> , 2017, 10, 194-200.	0.6	2
9568	Scalable P2P Botnet Detection with Threshold Setting in Hadoop Framework. <i>Journal of the Korea Institute of Information Security and Cryptology</i> , 2015, 25, 807-816.	0.1	5
9569	BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. <i>PLoS Computational Biology</i> , 2017, 13, e1005209.	1.5	218
9570	A fast and scalable framework for large-scale and ultrahigh-dimensional sparse regression with application to the UK Biobank. <i>PLoS Genetics</i> , 2020, 16, e1009141.	1.5	75
9571	Variant Callers for Next-Generation Sequencing Data: A Comparison Study. <i>PLoS ONE</i> , 2013, 8, e75619.	1.1	139
9572	GPU-BSM: A GPU-Based Tool to Map Bisulfite-Treated Reads. <i>PLoS ONE</i> , 2014, 9, e97277.	1.1	6
9573	Enabling Big Geoscience Data Analytics with a Cloud-Based, MapReduce-Enabled and Service-Oriented Workflow Framework. <i>PLoS ONE</i> , 2015, 10, e0116781.	1.1	37
9574	CLUSTOM-CLOUD: In-Memory Data Grid-Based Software for Clustering 16S rRNA Sequence Data in the Cloud Environment. <i>PLoS ONE</i> , 2016, 11, e0151064.	1.1	9
9575	Performance Optimization of Marine Science and Numerical Modeling on HPC Cluster. <i>PLoS ONE</i> , 2017, 12, e0169130.	1.1	3
9576	Halvade-RNA: Parallel variant calling from transcriptomic data using MapReduce. <i>PLoS ONE</i> , 2017, 12, e0174575.	1.1	17

#	ARTICLE	IF	CITATIONS
9577	An Effective Resource Management for Cloud Services using Clustering Schemes. , 2013, , .		3
9578	Enhancing Fault Tolerance based on Hadoop Cluster. International Journal of Database Theory and Application, 2014, 7, 37-48.	0.2	7
9579	A MapReduce Implementation of C4.5 Decision Tree Algorithm. International Journal of Database Theory and Application, 2014, 7, 49-60.	0.2	120
9580	Research of parallel DBSCAN clustering algorithm based on MapReduce. International Journal of Database Theory and Application, 2014, 7, 41-48.	0.2	5
9581	An Efficient Parallel Top-k Similarity Join for Massive Multidimensional Data Using Spark. International Journal of Database Theory and Application, 2015, 8, 57-68.	0.2	9
9582	The Concept of Pattern based Data Sharing in Big Data Environments. International Journal of Database Theory and Application, 2015, 8, 11-18.	0.2	8
9583	Efficient Pairwise Document Similarity Computation in Big Datasets. International Journal of Database Theory and Application, 2015, 8, 59-70.	0.2	10
9584	Bigdata Anonymization Using One Dimensional and Multidimensional Map Reduce Framework on Cloud. International Journal of Database Theory and Application, 2015, 8, 253-262.	0.2	5
9585	Comparative Study of Big Data Computing and Storage Tools: A Review. International Journal of Database Theory and Application, 2016, 9, 45-66.	0.2	31
9586	A Review of Sentiment Analysis in Twitter Data Using Hadoop. International Journal of Database Theory and Application, 2016, 9, 77-86.	0.2	30
9587	A Survey on SLA Management for Cloud Computing and Cloud-Hosted Big Data Analytic Applications. International Journal of Database Theory and Application, 2016, 9, 107-118.	0.2	16
9588	Improving MapReduce Performance by Data Prefetching in Heterogeneous or Shared Environments. International Journal of Grid and Distributed Computing, 2013, 6, 71-82.	0.8	13
9589	Collaborative Filtering Recommendation using Matrix Factorization: A MapReduce Implementation. International Journal of Grid and Distributed Computing, 2014, 7, 1-10.	0.8	3
9590	A Comprehensive Review of Straggler Handling Algorithms for MapReduce Framework. International Journal of Grid and Distributed Computing, 2014, 7, 139-148.	0.8	12
9591	Research for the Task Scheduling Algorithm Optimization based on Hybrid PSO and ACO for Cloud Computing. International Journal of Grid and Distributed Computing, 2014, 7, 87-96.	0.8	6
9592	An Architecture for Human Resource Information Management Using Cloud Computing. International Journal of Grid and Distributed Computing, 2014, 7, 181-190.	0.8	3
9593	A Parallel PageRank Algorithm with Power Iteration Acceleration. International Journal of Grid and Distributed Computing, 2015, 8, 273-284.	0.8	3
9594	Cloud Computing Environments Parallel Data Mining Policy Research. International Journal of Grid and Distributed Computing, 2015, 8, 135-144.	0.8	2

#	ARTICLE	IF	CITATIONS
9595	An Optimized Method of Translating SQL to More Efficient Map-reduce Tasks. International Journal of Grid and Distributed Computing, 2015, 8, 249-256.	0.8	1
9596	PCSP: A Preemptive Capacity Scheduler Policy for Scheduling Hadoop Jobs. International Journal of Grid and Distributed Computing, 2015, 8, 33-46.	0.8	1
9597	A Smart Strategy for Speculative Execution Based on Hardware Resource in a Heterogeneous Distributed Environment. International Journal of Grid and Distributed Computing, 2016, 9, 203-214.	0.8	5
9598	Efficient and Parallel Data Processing and Resource Allocation in the Cloud by using Nephelê's Data Processing Framework. International Journal of Grid and Distributed Computing, 2016, 9, 33-40.	0.8	1
9599	Building Efficient Resource Management Systems in the Cloud: Opportunities and Challenges. International Journal of Grid and Distributed Computing, 2016, 9, 157-172.	0.8	3
9600	Comparative Study of Multi-query Optimization Techniques using Shared Predicate-based for Big Data. International Journal of Grid and Distributed Computing, 2016, 9, 229-240.	0.8	14
9601	A Data Volume Aware Ant Colony Optimization Approach for Geographical Knowledge Cloud Service Composition. International Journal of Grid and Distributed Computing, 2016, 9, 103-116.	0.8	3
9602	Multi-Core Program Optimization: Parallel Sorting Algorithms in Intel Cilk Plus. International Journal of Hybrid Information Technology, 2014, 7, 151-164.	0.6	7
9603	BiShard Parallel Processor: A Disk-Based Processing Engine for Billion-Scale Graphs. International Journal of Multimedia and Ubiquitous Engineering, 2014, 9, 199-212.	0.3	6
9604	Efficient Data Replication Scheme based on Hadoop Distributed File System. International Journal of Software Engineering and Its Applications, 2015, 9, 177-186.	0.2	3
9605	Adaptive Consistency Approaches for Cloud Computing Platform. International Journal of Software Engineering and Its Applications, 2015, 9, 127-134.	0.2	3
9606	Data Science: Bigtable, MapReduce and Google File System. International Journal of Computer Trends and Technology, 2014, 16, 115-118.	0.1	4
9607	A Study on Big Data Integration with Data Warehouse. International Journal of Computer Trends and Technology, 2014, 9, 188-192.	0.1	13
9608	Big Data : Overview. International Journal of Computer Trends and Technology, 2014, 9, 266-268.	0.1	9
9609	An Enhanced Client-Server Assignment for Internet Distributed Systems. SSRG International Journal of Engineering Trends and Technology, 2014, 10, 198-201.	0.3	5
9610	Applications for ultrascale computing. Supercomputing Frontiers and Innovations, 2015, 2, .	0.5	1
9611	LASyM: A Learning Analytics System for MOOCs. International Journal of Advanced Computer Science and Applications, 2013, 4, .	0.5	26
9612	An Effective Storage Mechanism for High Performance Computing (HPC). International Journal of Advanced Computer Science and Applications, 2015, 6, .	0.5	3

#	ARTICLE	IF	CITATIONS
9613	A Novel Mapreduce Lift Association Rule Mining Algorithm (Mrlar) for Big Data. International Journal of Advanced Computer Science and Applications, 2016, 7, .	0.5	3
9614	HOP: Hardware makes Obfuscation Practical. , 2017, , .		30
9615	OBFUSCURO: A Commodity Obfuscation Engine on Intel SGX. , 2019, , .		25
9616	Modeling Structural Behaviour of Inhibitors of Cloud Computing: A TISM Approach. Transactions on Networks and Communications, 2014, 2, .	0.2	2
9617	Processing a trillion cells per mouse click. Proceedings of the VLDB Endowment, 2012, 5, 1436-1446.	2.1	59
9618	Piggybacking on social networks. Proceedings of the VLDB Endowment, 2013, 6, 409-420.	2.1	26
9619	JetScope. Proceedings of the VLDB Endowment, 2015, 8, 1680-1691.	2.1	10
9620	Quill. Proceedings of the VLDB Endowment, 2016, 9, 1623-1634.	2.1	8
9621	QUIS. Proceedings of the VLDB Endowment, 2017, 10, 1877-1880.	2.1	6
9622	Effective temporal dependence discovery in time series data. Proceedings of the VLDB Endowment, 2018, 11, 893-905.	2.1	8
9623	Chi. Proceedings of the VLDB Endowment, 2018, 11, 1303-1316.	2.1	48
9624	RHEEM: enabling cross-platform data processing. Proceedings of the VLDB Endowment, 2018, 11, 1414-1427.	2.1	32
9625	Streaming graph partitioning. Proceedings of the VLDB Endowment, 2018, 11, 1590-1603.	2.1	60
9626	Solving k-center clustering (with outliers) in MapReduce and streaming, almost as accurately as sequentially. Proceedings of the VLDB Endowment, 2019, 12, 766-778.	2.1	26
9627	AnalyticDB. Proceedings of the VLDB Endowment, 2019, 12, 2059-2070.	2.1	37
9628	Adapting TPC-C benchmark to measure performance of multi-document transactions in MongoDB. Proceedings of the VLDB Endowment, 2019, 12, 2254-2262.	2.1	20
9629	Shared arrangements. Proceedings of the VLDB Endowment, 2020, 13, 1793-1806.	2.1	9
9630	Profiling, what-if analysis, and cost-based optimization of MapReduce programs. Proceedings of the VLDB Endowment, 2011, 4, 1111-1122.	2.1	219

#	ARTICLE	IF	CITATIONS
9631	Tenzing a SQL implementation on the MapReduce framework. Proceedings of the VLDB Endowment, 2011, 4, 1318-1327.	2.1	74
9632	MapReduce programming and cost-based optimization?. Proceedings of the VLDB Endowment, 2011, 4, 1446-1449.	2.1	22
9633	Mosaic. Proceedings of the VLDB Endowment, 2020, 13, 2662-2675.	2.1	10
9634	Parallel graph algorithms in constant adaptive rounds. Proceedings of the VLDB Endowment, 2020, 13, 3588-3602.	2.1	10
9635	Query processing in multistore systems: an overview. International Journal of Cloud Computing, 2016, 5, 309.	0.3	5
9636	pvFPGA: paravirtualising an FPGA-based hardware accelerator towards general purpose computing. International Journal of High Performance Computing and Networking, 2017, 10, 179.	0.4	1
9637	Breaking news detection from the web documents through text mining and seasonality. International Journal of Knowledge and Web Intelligence, 2016, 5, 190.	0.2	2
9638	Systematizing Decentralization and Privacy: Lessons from 15 Years of Research and Deployments. Proceedings on Privacy Enhancing Technologies, 2017, 2017, 404-426.	2.3	39
9639	Adapting self-organizing Maps to the MapReduce programming paradigm. , 0, , .		3
9640	Stormgen - A Domain specific language to create ad-hoc Storm Topologies. , 0, , .		5
9641	Simulating Large-scale Aggregate MASs with Alchemist and Scala. , 0, , .		20
9642	Massively Parallel Feature Extraction Framework Application in Predicting Dangerous Seismic Events. , 0, , .		6
9643	Business Intelligence Platform for Big Data based on Scalable Distributed Two-Layer Data Store. , 0, , .		4
9644	Comparative Study Parallel Join Algorithms for MapReduce environment. Proceedings of the Institute for System Programming of RAS, 2012, 23, 285-306.	0.1	5
9648	Performing MapReduce on Data Centers with Hierarchical Structures. International Journal of Computers, Communications and Control, 2014, 7, 432.	1.2	2
9649	Hadoop Optimization for Massive Image Processing: Case Study Face Detection. International Journal of Computers, Communications and Control, 2014, 9, 664.	1.2	10
9650	Analysis on Buffer Occupancy of Quantized Congestion Notification in Data Center Networks. IEICE Transactions on Communications, 2016, E99.B, 2361-2372.	0.4	3
9651	A Sharper Generalization Bound for Divide-and-Conquer Ridge Regression. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 5305-5312.	3.6	3

#	ARTICLE	IF	CITATIONS
9652	SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. Journal of Artificial Intelligence Research, 0, 61, 863-905.	7.0	942
9653	From Frequency to Meaning: Vector Space Models of Semantics. Journal of Artificial Intelligence Research, 0, 37, 141-188.	7.0	1,557
9654	Parallel Model-Based Diagnosis on Multi-Core Computers. Journal of Artificial Intelligence Research, 0, 55, 835-887.	7.0	20
9655	AND/OR Branch-and-Bound on a Computational Grid. Journal of Artificial Intelligence Research, 0, 59, 351-435.	7.0	5
9656	Unique Sense: Smart Computing Prototype for Industry 4.0 Revolution with IOT and Bigdata Implementation Model. Indian Journal of Science and Technology, 2015, 8, .	0.5	12
9657	Efficiency of Stream Processing Engines for Processing BIGDATA Streams. Indian Journal of Science and Technology, 2016, 9, .	0.5	4
9658	Big Data in Bioinformatics. Mathematical Biology and Bioinformatics, 2017, 12, 102-119.	0.1	7
9659	Big Data in Bioinformatics. Mathematical Biology and Bioinformatics, 2018, 13, t1-t16.	0.1	5
9660	Big data analytics for proactive industrial decision support. Atp Magazin, 2016, 58, 62.	0.3	17
9661	Big Data Research in Information Systems: Toward an Inclusive Research Agenda. Journal of the Association for Information Systems, 2016, 17, I-XXXII.	2.4	348
9662	A New Parallel Item-Based Collaborative Filtering Algorithm Based on Hadoop. Journal of Software, 2015, 10, 416-426.	0.6	5
9663	Big data analysis in a geoinformatic problem of short-term traffic flow forecasting based on a k nearest neighbors method. Computer Optics, 2018, 42, 1101-1111.	1.3	29
9665	Optimizing Parallelism of Big Data Analytics at Distributed Computing System. International Journal on Advanced Science, Engineering and Information Technology, 2017, 7, 1716.	0.2	6
9666	Improved Map Reduce K Mean Clustering Algorithm for Hadoop Architecture. International Journal of Engineering and Computer Science, 0, , .	0.2	3
9667	Atm Plug-In for Distributed Text Mining in<i>R</i>. Journal of Statistical Software, 2012, 51, .	1.8	7
9668	Scalable Strategies for Computing with Massive Data. Journal of Statistical Software, 2013, 55, .	1.8	69
9669	BatchJobs and BatchExperiments: Abstraction Mechanisms for Using<i>R</i> in Batch Environments. Journal of Statistical Software, 2015, 64, .	1.8	30
9670	Inferring Concept Hierarchies from Text Corpora via Hyperbolic Embeddings. , 2019, , .		30

#	ARTICLE	IF	CITATIONS
9671	A Distant Supervision Approach to Semantic Role Labeling. , 2015, , .		3
9672	Distributed representation and estimation of WFST-based n-gram models. , 2016, , .		3
9674	Distributed Mining of Closed Patterns from Multi-Relational Data. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2015, 19, 804-809.	0.5	1
9675	gmm_diag and gmm_full: C++ classes for multi-threaded Gaussian mixture models and Expectation-Maximisation. Journal of Open Source Software, 2017, 2, .	2.0	4
9676	LiberTEM: Software platform for scalable multidimensional data processing in transmission electron microscopy. Journal of Open Source Software, 2020, 5, 2006.	2.0	26
9686	Scalable MCMC for Large Data Problems Using Data Subsampling and the Difference Estimator. SSRN Electronic Journal, 0, , .	0.4	2
9687	MapReduce Foundation of Big Data With Hadoop Environment. SSRN Electronic Journal, 0, , .	0.4	3
9688	Lambda Architecture “ Realtime Data Processing. SSRN Electronic Journal, 0, , .	0.4	2
9689	Cloud Computing Value Chains: Research from the Operations Management Perspective. SSRN Electronic Journal, 0, , .	0.4	1
9690	Knowledge management overview of feature selection problem in high-dimensional financial data: cooperative co-evolution and MapReduce perspectives. Problems and Perspectives in Management, 2019, 17, 340-359.	0.5	9
9691	An Efficient Distributed B-tree Index Method in Cloud Computing. Open Cybernetics and Systemics Journal, 2015, 8, 302-308.	0.3	3
9692	Bayesian Spam Filtering Mechanism Based on Decision Tree of Attribute Set Dependence in the MapReduce Framework. Open Cybernetics and Systemics Journal, 2014, 8, 435-441.	0.3	2
9693	An Improved Data Placement Strategy in a Heterogeneous Hadoop Cluster. Open Cybernetics and Systemics Journal, 2015, 9, 792-798.	0.3	2
9694	MapReduce-based Parallel Learning for Large-scale Remote Sensing Images. Open Automation and Control Systems Journal, 2014, 6, 1962-1974.	0.9	1
9695	SAMPLING BASED JOIN-AGGREGATE QUERY PROCESSING TECHNIQUE FOR BIG DATA. Indian Journal of Computer Science and Engineering, 2020, 11, 532-546.	0.2	14
9696	A robust combination technique. ANZIAM Journal, 0, 54, 394.	0.0	9
9697	Health Care and Precision Medicine Research: Analysis of a Scalable Data Science Platform. Journal of Medical Internet Research, 2019, 21, e13043.	2.1	68
9698	A New Improved Baum-Welch Algorithm for Unsupervised Learning for Continuous-Time HMM Using Spark. International Journal of Intelligent Engineering and Systems, 2020, 13, 214-226.	0.8	5

#	ARTICLE	IF	CITATIONS
9699	The efficient implementation of distributed indexing with Hadoop for digital investigations on Big Data. <i>Computer Science and Information Systems</i> , 2014, 11, 1037-1054.	0.7	9
9700	A distributed near-optimal LSH-based framework for privacy-preserving record linkage. <i>Computer Science and Information Systems</i> , 2014, 11, 745-763.	0.7	18
9701	Toward Effective Big Data Analysis in Continuous Auditing. <i>Accounting Horizons</i> , 2015, 29, 469-476.	1.1	133
9702	An REA Ontology-Based Model for Mapping Big Data to Accounting Information Systems Elements. <i>Journal of Information Systems</i> , 2017, 31, 45-61.	0.5	5
9703	Securing Big Data Provenance for Auditors: The Big Data Provenance Black Box as Reliable Evidence. <i>Journal of Emerging Technologies in Accounting</i> , 2016, 13, 17-36.	0.8	45
9704	Parallel Support Vector Machines on Multi-Core and Multiprocessor Systems. , 2011, , .		17
9705	Multi-layer Data Model for Transportation Logistics Solutions. , 2020, , .		3
9706	The RADStack: Open Source Lambda Architecture for Interactive Analytics. , 2017, , .		10
9707	ConTaaS: An Approach to Internet-Scale Contextualisation for Developing Efficient Internet of Things Applications. , 2017, , .		18
9708	A Seismic Data Processing System based on Fast Distributed File System. <i>International Journal of Computers & Technology</i> , 2015, 14, 5779-5788.	0.2	1
9709	Decision-Making Enhancement in a Big Data Environment: Application of the K-Means Algorithm to Mixed Data. <i>Journal of Artificial Intelligence and Soft Computing Research</i> , 2019, 9, 293-302.	3.5	23
9710	Cardinality Estimators do not Preserve Privacy. <i>Proceedings on Privacy Enhancing Technologies</i> , 2019, 2019, 26-46.	2.3	24
9711	Improving the Performance of Fair Scheduler in Hadoop. <i>Advances in Science, Technology and Engineering Systems</i> , 2017, 2, 1050-1058.	0.4	1
9712	Cooperation of Simulation and Data Model for Performance Analysis of Complex Systems. <i>International Journal of Simulation Modelling</i> , 2019, 18, 608-619.	0.6	37
9713	The Real Time Big Data Processing Framework Advantages and Limitations. <i>International Journal of Computer Sciences and Engineering</i> , 2017, 5, 305-312.	0.0	15
9714	Balance resource allocation for spark jobs based on prediction of the optimal resource. <i>Tsinghua Science and Technology</i> , 2020, 25, 487-497.	4.1	16
9715	Directions for Big Data Graph Analytics Research. <i>International Journal of Big Data</i> , 2015, 2, 15-27.	0.6	2
9716	A retweet prediction method of micro-blog big data users based on Map/Reduce. , 2016, , .		1

#	ARTICLE	IF	CITATIONS
9717	A Novel Parallel Approach of Cuckoo Search using MapReduce. , 2014, , .		11
9718	The Framework of Remote Sensing Image Map Service on Hadoop. , 2012, , .		5
9719	Research on the Small Files Problem of Hadoop. , 0, , .		1
9720	Locality-based Partitioning for Spark. , 2017, , .		2
9721	Spectral clustering algorithm based on Hadoop cloud platform research and implementation. , 2016, , .		3
9722	Improved Particle Optimization Algorithm Solving Hadoop Task Scheduling Problem. , 2015, , .		3
9723	Real-time Calculating Over Self-Health Data Using Storm. , 2015, , .		2
9724	Big Social Data Analytics: Opportunities, Challenges and Implications on Society. Online Journal of Communication and Media Technologies, 2016, 6, 17-32.	0.4	1
9725	Triplet lexicon models for statistical machine translation. , 2008, , .		14
9726	Scalable language processing algorithms for the masses. , 2008, , .		16
9727	Data-intensive text processing with MapReduce. , 2009, , .		135
9728	Coupling semi-supervised learning of categories and relations. , 2009, , .		37
9729	Distributed parse mining. , 2009, , .		4
9730	Web-scale distributional similarity and entity set expansion. , 2009, , .		125
9731	Using word-sense disambiguation methods to classify web queries by intent. , 2009, , .		7
9732	Polylingual Tree-Based Topic Models for Translation Domain Adaptation. , 2014, , .		32
9733	JoBimViz: A Web-based Visualization for Graph-based Distributional Semantic Models. , 2015, , .		9
9734	An Optimized Resource Scheduling Strategy for Hadoop Speculative Execution Based on Non-cooperative Game Schemes. Computers, Materials and Continua, 2020, 62, 713-729.	1.5	5

#	ARTICLE	IF	CITATIONS
9735	Big Data Anonymization in Cloud using k-Anonymity Algorithm using Map Reduce Framework. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 2019, , 50-56.	0.2	4
9736	Open-Source Federated Learning Frameworks for IoT: A Comparative Review and Analysis. Sensors, 2021, 21, 167.	2.1	64
9737	A Survey of Enabling Technologies for Smart Communities. Smart Cities, 2021, 4, 54-77.	5.5	33
9738	Storage of Mobile Sensor Data in Clouds using Information Classification Algorithms. International Journal of Advanced Networking and Applications, 2018, 10, 3893-3897.	0.2	2
9739	Cloud Computing: System Instances and Current Research. Ruan Jian Xue Bao/Journal of Software, 2010, 20, 1337-1348.	0.3	199
9740	Ontology Induction Based on Social Annotations. Jisuanji Xuebao/Chinese Journal of Computers, 2010, 33, 1823-1834.	0.3	1
9741	Large Scale Graph Data Processing on Cloud Computing Environments. Jisuanji Xuebao/Chinese Journal of Computers, 2011, 34, 1753-1767.	0.3	9
9742	Efficient Skyline Query Processing of Massive Data Based on Map-Reduce. Jisuanji Xuebao/Chinese Journal of Computers, 2011, 34, 1785-1796.	0.3	7
9743	Knowledge Reduction Algorithms in Cloud Computing. Jisuanji Xuebao/Chinese Journal of Computers, 2012, 34, 2332-2343.	0.3	11
9744	Real-Time Processing for High Speed Data Stream over Large Scale Data. Jisuanji Xuebao/Chinese Journal of Computers, 2012, 35, 477-490.	0.3	17
9745	A Security Protection Framework for Cloud Computing. Journal of Information Processing Systems, 0, , .	1.0	4
9746	A Hadoop-based Multimedia Transcoding System for Processing Social Media in the PaaS Platform of SMCCSE. KSII Transactions on Internet and Information Systems, 0, , .	0.7	5
9747	An Efficient Design and Implementation of an MdbULPS in a Cloud-Computing Environment. KSII Transactions on Internet and Information Systems, 2015, 9, 3182-3202.	0.7	1
9748	MissingFound: An Assistant System for Finding Missing Companions via Mobile Crowdsourcing. KSII Transactions on Internet and Information Systems, 2016, 10, .	0.7	2
9749	Improving Security of Parallel Algorithm Using Key Encryption Technique. Information Technology Journal, 2013, 12, 2398-2404.	0.3	4
9750	WBS Data Analysis Method Based on Information Supply Chain. Journal of Applied Sciences, 2013, 13, 2355-2358.	0.1	1
9751	Ontology-Based Scientific Service Instant Response System Design. Journal of Software Engineering, 2015, 9, 574-585.	0.2	2
9753	Performance analysis of large-scale parallel-distributed processing with backup tasks for cloud computing. Journal of Industrial and Management Optimization, 2014, 10, 113-129.	0.8	23

#	ARTICLE	IF	CITATIONS
9754	Performance analysis of backup-task scheduling with deadline time in cloud computing. Journal of Industrial and Management Optimization, 2015, 11, 867-886.	0.8	5
9755	InterSeg: a distributed image segmentation tool. , 0, , .		1
9756	Online Machine Learning. , 2013, , 27-54.		49
9757	A Cognitive Analytics Management Framework (CAM-Part 2). Advances in Logistics, Operations, and Management Science Book Series, 0, , 80-189.	0.3	2
9758	Parallel Data Reduction Techniques for Big Datasets. Advances in Data Mining and Database Management Book Series, 0, , 72-93.	0.4	10
9759	Large-Scale Sensor Network Analysis. Advances in Data Mining and Database Management Book Series, 0, , 314-347.	0.4	1
9760	Emergence of NoSQL Platforms for Big Data Needs. , 2014, , 787-798.		1
9761	Harnessing the Power of Big Data Analytics. , 2014, , 1107-1116.		3
9762	Driving Big Data with Hadoop Technologies. Advances in Data Mining and Database Management Book Series, 2014, , 232-262.	0.4	3
9763	Accessing Big Data in the Cloud Using Mobile Devices. Advances in Data Mining and Database Management Book Series, 2014, , 444-470.	0.4	25
9764	Rough Set Based Green Cloud Computing in Emerging Markets. , 2015, , 1078-1087.		4
9765	A Survey on Research Initiatives for Healthcare Clouds. Advances in Healthcare Information Systems and Administration Book Series, 2014, , 1-18.	0.2	5
9766	Perspectives of Big Data Analysis for Knowledge Generation in Project Management Contexts. Advances in IT Personnel and Project Management, 2015, , 1-18.	0.3	11
9767	Big Data Mining Based on Computational Intelligence and Fuzzy Clustering. Advances in Data Mining and Database Management Book Series, 2015, , 130-148.	0.4	5
9768	A Hybrid Approach to Big Data Systems Development. Advances in Public Policy and Administration, 2016, , 20-37.	0.1	1
9769	The Heterogeneity Paradigm in Big Data Architectures. Advances in Data Mining and Database Management Book Series, 2016, , 218-245.	0.4	1
9770	Challenges and Opportunities in Big Data Processing. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2016, , 1-24.	0.5	5
9771	A Review of RDF Storage in NoSQL Databases. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2016, , 210-229.	0.5	4

#	ARTICLE	IF	CITATIONS
9772	Parallel Data Reduction Techniques for Big Datasets. , 2016, , 734-756.		4
9773	Challenges and Opportunities in Big Data Processing. , 2016, , 2074-2097.		2
9774	Using Big Data in Healthcare. , 2016, , 889-899.		1
9775	Data Science and Big Data Analytics in Financial Services. Advances in Business Strategy and Competitive Advantage Book Series, 2016, , 396-423.	0.2	1
9776	Mastering Big Data in the Digital Age. Advances in Data Mining and Database Management Book Series, 0, , 104-129.	0.4	16
9777	Big Data Analytics Using Local Exceptionality Detection. Advances in Business Information Systems and Analytics Book Series, 2016, , 108-125.	0.3	4
9778	A Survey and Taxonomy of Energy Efficient Resource Management Techniques in Platform as a Service Cloud. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2017, , 410-454.	0.5	14
9779	Software as a Service, Semantic Web, and Big Data. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2017, , 264-285.	0.5	16
9780	Adaptive Multimedia Services in Next-Generation Broadband Wireless Access Network. Advances in Wireless Technologies and Telecommunication Book Series, 2017, , 1-31.	0.3	2
9781	In-Memory Analytics. , 2018, , 1806-1813.		1
9782	Cloud Computing Education Strategies. Advances in Data Mining and Database Management Book Series, 2018, , 43-54.	0.4	3
9783	Database Systems for Big Data Storage and Retrieval. Advances in Data Mining and Database Management Book Series, 2018, , 76-100.	0.4	2
9784	Visualization Tools for Big Data Analytics in Quantitative Chemical Analysis. Advances in Data Mining and Database Management Book Series, 2018, , 873-917.	0.4	1
9785	Big Data Analytics for Train Delay Prediction. Advances in Civil and Industrial Engineering Book Series, 2018, , 320-348.	0.2	2
9786	Big Data and Analytics. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2018, , 55-66.	0.5	1
9787	Security and Privacy Challenges in Big Data Environment. Advances in Information Security, Privacy, and Ethics Book Series, 2018, , 315-325.	0.4	7
9788	Crowd Computing. Advances in Logistics, Operations, and Management Science Book Series, 2019, , 166-198.	0.3	6
9789	Internet of Things and Its Impacts in Computing Intelligence. Advances in Civil and Industrial Engineering Book Series, 2019, , 103-136.	0.2	16

#	ARTICLE	IF	CITATIONS
9790	Improving Project Management Decisions With Big Data Analytics. Advances in Business Information Systems and Analytics Book Series, 2019, , 45-65.	0.3	3
9791	Role of Big Data in Internet of Things Networks. Advances in Data Mining and Database Management Book Series, 2019, , 273-299.	0.4	5
9792	Software-Defined Networking for Scalable Cloud-Based Services to Improve System Performance of Hadoop-Based Big Data Applications. , 2019, , 1460-1484.		2
9793	Towards Massive RDF Storage in NoSQL Databases. Advances in Data Mining and Database Management Book Series, 2019, , 263-284.	0.4	4
9794	Denial-of-Service and Botnet Analysis, Detection, and Mitigation. Advances in Wireless Technologies and Telecommunication Book Series, 2020, , 114-151.	0.3	2
9795	Adaptive Peer-to-Peer Social Networks for Distributed Content-Based Web Search. , 0, , 155-178.		2
9796	Big Data and Its Visualization With Fog Computing. , 2020, , 341-377.		1
9797	An Optimal Data Placement Strategy for Improving System Performance of Massive Data Applications Using Graph Clustering. International Journal of Ambient Computing and Intelligence, 2018, 9, 15-30.	0.8	7
9799	A Context-Based Performance Enhancement Algorithm for Columnar Storage in MapReduce with Hive. International Journal of Cloud Applications and Computing, 2013, 3, 38-50.	1.1	1
9800	ClouDET. International Journal of Cognitive Informatics and Natural Intelligence, 2016, 10, 12-31.	0.4	12
9801	The Impact of Virtualization on High Performance Computing Clustering in the Cloud. International Journal of Distributed Systems and Technologies, 2015, 6, 65-81.	0.6	9
9802	Towards Big Linked Data. International Journal of Data Warehousing and Mining, 2013, 9, 19-43.	0.4	2
9803	Software-Defined Networking for Scalable Cloud-based Services to Improve System Performance of Hadoop-based Big Data Applications. International Journal of Grid and High Performance Computing, 2016, 8, 1-22.	0.7	7
9804	Extracting Functional Dependencies in Large Datasets Using MapReduce Model. International Journal of Intelligent Information Technologies, 2014, 10, 19-35.	0.5	15
9805	Big Data and Healthcare Data. International Journal of Knowledge-Based Organizations, 2017, 7, 50-77.	0.3	6
9806	A Scalable MapReduce-enabled Glowworm Swarm Optimization Approach for High Dimensional Multimodal Functions. International Journal of Swarm Intelligence Research, 2016, 7, 32-54.	0.5	11
9807	A Fine-Grained Stateful Data Analytics Method Based on Resilient State Table. International Journal of Software Science and Computational Intelligence, 2018, 10, 66-79.	1.8	1
9808	A Big Data Test-bed for Analyzing Data Generated by an Air Pollution Sensor Network. International Journal of Web Services Research, 2016, 13, 19-35.	0.5	1

#	ARTICLE	IF	CITATIONS
9809	Towards Automated Data-Driven Model Creation for Cloud Computing Simulation. , 2015, , .		3
9810	Time Warp on the Go. , 2012, , .		4
9811	Scalable Stream Processing with Quality of Service for Smart City Crowdsensing Applications. EAI Endorsed Transactions on Mobile Communications and Applications, 2013, 1, e6.	0.5	2
9812	A Distributed QoS-Constraint Task Scheduling Scheme in Cloud Computing Environment: Model and Algorithm. Advances in Information Sciences and Service Sciences, 2012, 4, 283-291.	0.1	16
9813	Hbase - non SQL Database, Performances Evaluation. International Journal of Advancements in Computing Technology, 2010, 2, 42-52.	0.1	32
9814	How Computers Work: Computational Thinking for Everyone. Electronic Proceedings in Theoretical Computer Science, EPTCS, 0, 106, 1-19.	0.8	4
9815	Architecture Diagrams: A Graphical Language for Architecture Style Specification. Electronic Proceedings in Theoretical Computer Science, EPTCS, 0, 223, 83-97.	0.8	7
9816	Aggregate Graph Statistics. Electronic Proceedings in Theoretical Computer Science, EPTCS, 0, 264, 18-22.	0.8	2
9817	Resilient Blocks for Summarising Distributed Data. Electronic Proceedings in Theoretical Computer Science, EPTCS, 0, 264, 23-26.	0.8	3
9818	Research of Big Data Based on the Views of Technology and Application. American Journal of Industrial and Business Management, 2015, 05, 192-197.	0.4	12
9819	Big Data Analysis in Smart Manufacturing: A Review. International Journal of Communications, Network and System Sciences, 2017, 10, 31-58.	0.4	53
9820	Seismic Data Collection with Shakebox and Analysis Using MapReduce. Journal of Computer and Communications, 2015, 03, 94-101.	0.6	1
9821	Building a Productive Domain-Specific Cloud for Big Data Processing and Analytics Service. Journal of Computer and Communications, 2015, 03, 107-117.	0.6	3
9822	Hadoop-Based Similarity Computation System for Composed Documents. Journal of Computer and Communications, 2015, 03, 196-202.	0.6	2
9823	Big Data for Organizations: A Review. Journal of Computer and Communications, 2017, 05, 40-48.	0.6	10
9824	Optimized Homomorphic Scheme on Map Reduce for Data Privacy Preserving. Journal of Information Security, 2017, 08, 257-273.	0.4	1
9825	Data Modeling and Data Analytics: A Survey from a Big Data Perspective. Journal of Software Engineering and Applications, 2015, 08, 617-634.	0.8	25
9826	Hadoop and Its Role in Modern Image Processing. Open Journal of Marine Science, 2014, 04, 239-245.	0.3	12

#	ARTICLE	IF	CITATIONS
9827	Towards A Large-Scale Cognitive Radio Network Testbed: Spectrum Sensing, System Architecture, and Distributed Sensing. Journal of Communications, 2012, 7, .	1.3	14
9828	Multicores in Cloud Computing: Research Challenges for Applications. Journal of Computers, 2010, 5, .	0.4	11
9829	Energy-efficient Task Scheduling Model based on MapReduce for Cloud Computing using Genetic Algorithm. Journal of Computers, 2012, 7, .	0.4	18
9830	K-Means Method for Grouping in Hybrid MapReduce Cluster. Journal of Computers, 2013, 8, .	0.4	11
9831	A Scalable XSLT Processing Framework based on MapReduce. Journal of Computers, 2013, 8, .	0.4	4
9832	An Optimized Algorithm for Reduce Task Scheduling. Journal of Computers, 2014, 9, .	0.4	2
9833	A MapReduce based Parallel SVM for Email Classification. Journal of Networks, 2014, 9, .	0.4	18
9834	BIG-DATA and the Challenges for Statistical Inference and Economics Teaching and Learning. Multidisciplinary Journal for Education, Social and Technological Sciences, 2017, 4, 64.	0.8	5
9835	A Scientometric Analysis of Public Transport Research. Journal of Public Transportation, 2015, 18, 111-141.	0.3	26
9836	Modeling the Interrelationship between the Enablers of Cloud Computing. International Journal of Computer Applications, 2013, 67, 38-43.	0.2	7
9837	Parallel Data Processing in the Cloud using Nephele. International Journal of Computer Applications, 2013, 69, 1-8.	0.2	1
9838	Parallel Data Processing for Effective Dynamic Resource Allocation in the Cloud. International Journal of Computer Applications, 2013, 70, 1-4.	0.2	11
9839	FISM and FMICMAC Analysis on Enablers of Cloud Computing. International Journal of Computer Applications, 2014, 91, 9-16.	0.2	2
9840	New Design Principles for Effective Knowledge Discovery from Big Data. International Journal of Computer Applications, 2014, 96, 19-23.	0.2	1
9841	Research in Big Data and Analytics: An Overview. International Journal of Computer Applications, 2014, 108, 19-23.	0.2	8
9842	GeoProcessing Workflow Models for Distributed Processing Frameworks. International Journal of Computer Applications, 2015, 113, 33-38.	0.2	3
9843	Big Data Frameworks for Efficient Range Queries to Extract Interested Rectangular Sub Regions. International Journal of Computer Applications, 2015, 119, 36-39.	0.2	6
9844	Exploratory Implementation of Stream Clustering Algorithm using MongoDB. International Journal of Computer Applications, 2015, 121, 21-29.	0.2	2

#	ARTICLE	IF	CITATIONS
9845	Crawling the Hidden Web: An Approach to Dynamic Web Indexing. International Journal of Computer Applications, 2012, 55, 7-15.	0.2	6
9846	Use of Mapreduce for Data Mining and Data Optimization on a Web Portal. International Journal of Computer Applications, 2012, 56, 39-43.	0.2	8
9847	Comparative Study Load Balance Algorithms for Map Reduce Environment. International Journal of Computer Applications, 2014, 106, 41-50.	0.2	7
9848	An Outlook on Big Data and Big Data Analytics. International Journal of Computer Applications, 2015, 124, 37-41.	0.2	8
9849	Performance Analysis of Apriori Algorithm with Different Data Structures on Hadoop Cluster. International Journal of Computer Applications, 2015, 128, 45-51.	0.2	19
9850	Analysis of Map Reduce Performance using Prefetching Mechanism. International Journal of Computer Applications, 2015, 128, 11-13.	0.2	6
9851	A Survey on Big Data Management and Job Scheduling. International Journal of Computer Applications, 2015, 130, 41-49.	0.2	7
9852	Spark is superior to Map Reduce over Big Data. International Journal of Computer Applications, 2016, 133, 13-16.	0.2	1
9853	Big Data Analytics to Predict Breast Cancer Recurrence on SEER Dataset using MapReduce Approach. International Journal of Computer Applications, 2016, 150, 7-11.	0.2	3
9854	TOWARDS AN INTELLIGENT PLATFORM FOR BIG 3D GEOSPATIAL DATA MANAGEMENT. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 0, IV-4, 133-140.	0.0	3
9855	SIDELoading " INGESTION OF LARGE POINT CLOUDS INTO THE APACHE SPARK BIG DATA ENGINE. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XLI-B2, 343-348.	0.2	4
9856	BUILDING SPATIOTEMPORAL CLOUD PLATFORM FOR SUPPORTING GIS APPLICATION. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 0, II-4/W2, 55-62.	0.0	8
9857	HADOOP-BASED DISTRIBUTED SYSTEM FOR ONLINE PREDICTION OF AIR POLLUTION BASED ON SUPPORT VECTOR MACHINE. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XL-1/W5, 215-219.	0.2	13
9858	CLASSIFICATION OF BIG POINT CLOUD DATA USING CLOUD COMPUTING. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XL-3/W3, 553-557.	0.2	14
9859	PARALLEL CREATION OF VARIO-SCALE DATA STRUCTURES FOR LARGE DATASETS. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XL-4/W7, 1-9.	0.2	1
9860	A HADOOP-BASED ALGORITHM OF GENERATING DEM GRID FROM POINT CLOUD DATA. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XL-7/W3, 1209-1214.	0.2	6
9861	Big Data breaking barriers - first steps on a long trail. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XL-7/W3, 691-697.	0.2	7
9862	A ROBUST PARALLEL FRAMEWORK FOR MASSIVE SPATIAL DATA PROCESSING ON HIGH PERFORMANCE CLUSTERS. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XXXIX-B4, 213-217.	0.2	1

#	ARTICLE	IF	CITATIONS
9864	CACSS: TOWARDS A GENERIC CLOUD STORAGE SERVICE. , 2012, , .		4
9865	USING CLOUDS FOR SCIENCE, IS IT JUST KICKING THE CAN DOWN THE ROAD?. , 2012, , .		2
9866	DIFPL - Distributed Drone Flight Path Builder System. , 2015, , .		1
9867	Dynamic and Scalable Real-time Analytics in Logistics - Combining Apache Storm with Complex Event Processing for Enabling New Business Models in Logistics. , 2015, , .		1
9868	Big Data: A Survey - The New Paradigms, Methodologies and Tools. , 2015, , .		16
9869	MR-SAT: A MapReduce Algorithm for Big Data Sentiment Analysis on Twitter. , 2016, , .		4
9870	Benchmarking Hadoop Performance in the Cloud. , 2016, , .		9
9871	A Container-centric Methodology for Benchmarking Workflow Management Systems. , 2016, , .		13
9872	XIS-Reverse: A Model-driven Reverse Engineering Approach for Legacy Information Systems. , 2017, , .		3
9873	Adaptive Resource Management for Distributed Data Analytics based on Container-level Cluster Monitoring. , 2017, , .		2
9874	Towards a Scalable Architecture for Flight Data Management. , 2017, , .		3
9875	User-to-User Recommendation using the Concept of Movement Patterns: A Study using a Dating Social Network. , 2017, , .		1
9876	Parallel Privacy-preserving Record Linkage using LSH-based Blocking. , 2018, , .		14
9877	BRAID - A Hybrid Processing Architecture for Big Data. , 2018, , .		11
9878	Predicting SQL Query Execution Time with a Cost Model for Spark Platform. , 2020, , .		1
9879	GeoSimMR: A MapReduce Algorithm for Detecting Communities based on Distance and Interest in Social Networks. Data Science Journal, 2019, 18, .	0.6	3
9880	Jug: Software for Parallel Reproducible Computation in Python. Journal of Open Research Software, 2017, 5, 30.	2.7	24
9881	Analyzing Smart Grid Energy Data using Hadoop Based Big Data System. The Transactions of the Korean Institute of Electrical Engineers P, 2015, 64, 85-91.	0.0	3

#	ARTICLE	IF	CITATIONS
9882	Hadoop MapReduce Job Scheduling Algorithms Survey and Use Cases. Modern Applied Science, 2019, 13, 38.	0.4	4
9933	Conceptualization and theorization of the Big Data. International Journal of Innovation, 2016, 4, 23-41.	0.3	5
9934	Effects of Hypervisor on Distributed Big Data Processing in Virtualized Cluster Environment. KIISE Transactions on Computing Practices, 2016, 22, 89-94.	0.1	2
9935	A Preemptive Fair Scheduler Policy for Disco MapReduce Framework. , 0, , .		1
9937	High Performance Computation of Big Data: Performance Optimization Approach towards a Parallel Frequent Item Set Mining Algorithm for Transaction Data based on Hadoop MapReduce Framework. International Journal of Intelligent Systems and Applications, 2017, 9, 75-84.	0.9	11
9938	A Review on Large Scale Graph Processing Using Big Data Based Parallel Programming Models. International Journal of Intelligent Systems and Applications, 2017, 9, 49-57.	0.9	4
9939	A Novel and Efficient Method for Protecting Internet Usage from Unauthorized Access Using Map Reduce. International Journal of Information Technology and Computer Science, 2013, 5, 49-55.	0.8	11
9940	A Novel Approach for Identification of Hadoop Cloud Temporal Patterns Using Map Reduce. International Journal of Information Technology and Computer Science, 2014, 6, 37-42.	0.8	9
9941	Performance Analysis in Bigdata. International Journal of Information Technology and Computer Science, 2015, 7, 55-61.	0.8	6
9942	Cloud Implementation Of Agent-Based Simulation Model In Evacuation Scenarios. , 2016, , .		1
9943	THE DESIGN AND IMPLEMENTATION OF A GARBAGE TRUCK FLEET MANAGEMENT SYSTEM. South African Journal of Industrial Engineering, 2016, 27, .	0.2	4
9946	The Themes of Entrepreneurship Discourse: A Data Analytics Approach. Journal of Entrepreneurship, Management and Innovation, 2018, 14, 127-158.	0.6	2
9947	Volunteer Computing Simulation using Repast and Mason. Computer Science, 2013, 14, 153.	0.4	2
9948	CLUO: Web-Scale Text Mining System for Open Source Intelligence Purposes. Computer Science, 2013, 14, 45.	0.4	8
9949	Gene Sequence Input Formatting and MapReduce Computing. International Journal Bioautomation, 2019, , 233-246.	0.1	1
9950	MLitB: machine learning in the browser. PeerJ Computer Science, 0, 1, e11.	2.7	18
9951	A simple grid implementation with Berkeley Open Infrastructure for Network Computing using BLAST as a model. PeerJ, 2016, 4, e2248.	0.9	8
9952	Divide and Conquer (DC) BLAST: fast and easy BLAST execution within HPC environments. PeerJ, 2017, 5, e3486.	0.9	10

#	ARTICLE	IF	CITATIONS
9953	A Parallel Task scheduling Algorithm Based on Fuzzy Clustering in Cloud Computing Environment. International Journal of Machine Learning and Computing, 2014, 4, 437-444.	0.8	8
9954	Practical Datasets for Similarity Measures and Their Threshold Values. The Journal of Society for E-Business Studies, 2013, 18, 97-105.	0.5	4
9955	PARALLEL IMAGE DATABASE PROCESSING WITH MAPREDUCE AND PERFORMANCE EVALUATION IN PSEUDO DISTRIBUTED MODE. International Journal of Electronic Commerce Studies, 2012, 3, 211-228.	0.8	30
9956	A Method for Analyzing Web Log of the Hadoop System for Analyzing a Effective Pattern of Web Users. Journal of the Korea Society of IT Services, 2014, 13, 231-243.	0.0	2
9957	Fast Detection of Connected Components in Large Scale Graphs Using MapReduce. IOSR Journal of Engineering, 2014, 4, 35-42.	0.1	4
9958	MCF Tree-Based Clustering Method for Very Large Mixed-Type Data Set. IEEE Access, 2021, 9, 138580-138597.	2.6	0
9959	Simple, Deterministic, Constant-Round Coloring in Congested Clique and MPC. SIAM Journal on Computing, 2021, 50, 1603-1626.	0.8	4
9960	Query Optimization and Rebalancing Methods based on CMD. , 2021, , .		0
9961	Indexing on Healthcare Big Data. Advances in Intelligent Systems and Computing, 2021, , 271-283.	0.5	0
9962	TreeMerge: Efficient Generation of Minimal Hitting-Sets for Conflict Sets in Tree Structure for Model-Based Fault Diagnosis. IEEE Transactions on Reliability, 2021, 70, 1596-1610.	3.5	0
9963	An Apache Giraph Implementation of Distributed ADMM for Solving LASSO Problems. Advances in Intelligent Systems and Computing, 2021, , 547-556.	0.5	2
9964	System Stability Under Adversarial Injection of Dependent Tasks. IEEE Access, 2021, 9, 139516-139526.	2.6	0
9965	Tensors. Proceedings of the VLDB Endowment, 2021, 14, 1797-1804.	2.1	14
9966	An Efficient and Balanced Graph Partition Algorithm for the Subgraph-Centric Programming Model on Large-scale Power-law Graphs. , 2021, , .		1
9967	Distributed Data Processing for Large-Scale Simulations on Cloud. , 2021, , .		2
9968	Lazy-WL: A Wear-aware Load Balanced Data Redistribution Method for Efficient SSD Array Scaling. , 2021, , .		1
9969	High Performance Computing for CS-InSAR Data Processing. , 2021, , .		2
9970	Key Research Issues and Related Technologies in Crowdsourcing Data Collection. Wireless Communications and Mobile Computing, 2021, 2021, 1-13.	0.8	2

#	ARTICLE	IF	CITATIONS
9971	A Transfer Learning-Based Approach with Deep CNN for COVID-19- and Pneumonia-Affected Chest X-ray Image Classification. <i>SN Computer Science</i> , 2022, 3, 17.	2.3	35
9972	File System Support for Privacy-Preserving Analysis and Forensics in Low-Bandwidth Edge Environments. <i>Information (Switzerland)</i> , 2021, 12, 430.	1.7	6
9973	MuSe: a multi-level storage scheme for big RDF data using MapReduce. <i>Journal of Big Data</i> , 2021, 8, .	6.9	5
9974	Experience Paper: <i>sgx-dl.</i> , 2021, , .		2
9975	Scrooge. , 2021, , .		10
9976	Exploiting Reused-Based Sharing Work Opportunities in Big Data Multiquery Optimization with Flink. <i>Big Data</i> , 2021, 9, 454-479.	2.1	1
9977	Durable functions: semantics for stateful serverless. , 2021, 5, 1-27.		20
9978	RGen: Data Generator for Benchmarking Big Data Workloads. <i>Engineering Proceedings</i> , 2021, 7, 13.	0.4	0
9979	SmarT: Machine Learning Approach for Efficient Filtering and Retrieval of Spatial and Temporal Data in Big Data. <i>Journal of Information and Data Management</i> , 2021, 12, .	0.2	0
9980	Characterizing Microservice Dependency and Performance. , 2021, , .		73
9981	BlockHDFS: Blockchain-integrated Hadoop distributed file system for secure provenance traceability. <i>Blockchain: Research and Applications</i> , 2021, 2, 100032.	4.5	5
9982	The case for phase-aware scheduling of parallelizable jobs. <i>Performance Evaluation</i> , 2021, , 102246.	0.9	0
9983	Solving Large-Scale Granular Resource Allocation Problems Efficiently with POP. , 2021, , .		12
9984	Performance Optimization of a Parallel Error Correction Tool. <i>Engineering Proceedings</i> , 2021, 7, .	0.4	0
9985	Random Walks on Huge Graphs at Cache Efficiency. , 2021, , .		10
9986	Toward Responsible AI: An Overview of Federated Learning for User-centered Privacy-preserving Computing. <i>ACM Transactions on Interactive Intelligent Systems</i> , 2021, 11, 1-22.	2.6	7
9987	RICH: Strategy-proof and efficient coflow scheduling in non-cooperative environments. <i>Journal of Network and Computer Applications</i> , 2021, 196, 103233.	5.8	1
9990	Programming a Multicore Architecture without Coherency and Atomic Operations. , 2007, , .		0

#	ARTICLE	IF	CITATIONS
9991	Distributed, Large-Scale Latent Semantic Analysis by Index Interpolation. , 2008, , .		5
9992	Teaching large scale data processing. , 2008, , .		0
9993	Dynamic User-Defined Similarity Searching in Semi-Structured Text Retrieval. , 2008, , .		2
9994	Design and Anatomy of a Social Web Filtering Service. , 2008, , .		1
9997	Programming languages in a liberal arts education. ACM SIGPLAN Notices, 2008, 43, 45-49.	0.2	0
9998	An aspect-oriented approach to the undergraduate programming language curriculum. ACM SIGPLAN Notices, 2008, 43, 124-129.	0.2	1
9999	Synthesizing and Verifying Multicore Parallelism in Categories of Nested Code Graphs. Chapman & Hall/CRC Computational Science, 2008, , .	0.5	1
10001	Arabic cross-document coreference detection. , 2009, , .		2
10002	Enabling Technologies and Design Values for Building the Real World Web. Computational Methods in Science and Technology, 2009, 15, 83-93.	0.3	0
10003	Dynamic data services. , 2009, , .		0
10004	Towards a Theory of Universally Composable Cloud Computing. Lecture Notes in Computer Science, 2009, , 638-643.	1.0	1
10007	High-performance high-volume layered corpora annotation. , 2009, , .		0
10008	A MAPREDUCE FRAMEWORK FOR CHANGE PROPAGATION IN GEOGRAPHIC DATABASES. , 2009, , .		0
10009	Extending AspectJ for separating regions. , 2009, , .		3
10010	A Vision for Next Generation Query Processors and an Associated Research Agenda. Lecture Notes in Computer Science, 2009, , 1-11.	1.0	2
10012	Finding Related Web Pages in Parallel by Using Grouped Link Structures. Information Technology Journal, 2009, 8, 427-440.	0.3	0
10013	Analyzing Data Streams in Scientific Applications*. Chapman & Hall/CRC Computational Science, 2009, , .	0.5	0
10014	Accelerating Queries on Very Large Datasets. Chapman & Hall/CRC Computational Science, 2009, , .	0.5	5

#	ARTICLE	IF	CITATIONS
10017	RFID Event Data Processing: An Architecture for Storing and Searching. , 2010, , .		6
10018	IMPROVING THE WORKFLOW OF SEMANTIC WEB PORTALS USING M/R IN CLOUD PLATFORM. , 2010, , .		0
10019	Executing Multiple Group by Query Using MapReduce Approach: Implementation and Optimization. Lecture Notes in Computer Science, 2010, , 652-661.	1.0	3
10020	An IDA-Based Parallel Storage Scheme in the Scientific Data Grid. Data Science Journal, 2010, 9, 29-41.	0.6	0
10021	Practically Applicable Formal Methods. Lecture Notes in Computer Science, 2010, , 407-418.	1.0	5
10022	Characterising Effective Resource Analyses for Parallel and Distributed Coordination. Lecture Notes in Computer Science, 2010, , 67-83.	1.0	0
10024	Query Evaluation Techniques for Cluster Database Systems. Lecture Notes in Computer Science, 2010, , 351-362.	1.0	0
10025	Parallel Evolutionary Approach of Compaction Problem Using MapReduce. , 2010, , 361-370.		4
10026	Enhancing UNICORE Storage Management Using Hadoop Distributed File System. Lecture Notes in Computer Science, 2010, , 345-352.	1.0	1
10027	A Fully-Protected Large-Scale Email System Built on Map-Reduce Framework. Lecture Notes in Computer Science, 2010, , 662-669.	1.0	1
10028	Compress-and-conquer for optimal multicore computing. , 2010, , .		0
10030	Introducing Scalileo. , 2010, , .		0
10031	Merging File Systems and Data Bases to Fit the Grid. , 2010, , 13-25.		4
10032	Adaptive Conflict Unit Size for Distributed Optimistic Synchronization. Lecture Notes in Computer Science, 2010, , 547-559.	1.0	1
10033	Qualitative Analysis of Commercial Social Network Profiles. , 2010, , 95-113.		3
10034	Emerging Directions. Texts in Computer Science, 2010, , 189-205.	0.5	0
10035	A Performance Analysis Based on Hadoop Application's Characteristics in Cloud Computing. Journal of the Korea Society of Computer and Information, 2010, 15, 49-56.	0.0	2
10036	Block-Ranking: Content Similarity Retrieval Based on Data Partition in Network Storage Environment. International Journal of Digital Content Technology and Its Applications, 2010, 4, 85-94.	0.1	7

#	ARTICLE	IF	CITATIONS
10037	Large-Scale Data Processing. , 2010, , 89-138.		0
10038	Granules. , 2010, , 201-225.		1
10039	Granules: A Lightweight Runtime for Scalable Computing with Support for Map-Reduce SHRIDEEP PALLICKARA, JALIYA EKANAYAKE, AND GEOFFREY FOX. , 2010, , 215-240.		0
10040	Unsupervised discovery and training of maximally dissimilar cluster models. , 0, , .		18
10042	PROCESSING WIKIPEDIA DUMPS - A Case-study Comparing the XGrid and MapReduce Approaches. , 2011, , .		1
10043	Parallel Processing, Multiprocessors and Virtualization in Data-Intensive Computing. , 2011, , 235-248.		0
10044	Massive-Scale Analytics. , 2011, , 1089-1095.		0
10045	Data-Driven Modeling and Analysis of Online Social Networks. Lecture Notes in Computer Science, 2011, , 3-17.	1.0	3
10046	Energy Efficiency of Data Centers. , 2011, , 229-244.		3
10047	Parameter Sweep Job Submission to Clouds. , 2011, , 123-141.		1
10048	On the Processing of Extreme Scale Datasets in the Geosciences. , 2011, , 521-537.		1
10050	Framework for Development of Distributed Evolutionary Algorithms Based on MapReduce. Annals of DAAAM & Proceedings, 2011, , 1475-1476.	0.1	2
10051	OTPM: Failure Handling in Data-intensive Analytical Processing. , 2011, , .		2
10052	An Investigation on Applications of Cloud Computing in Scientific Computing. Communications in Computer and Information Science, 2011, , 201-206.	0.4	0
10053	Tree-Structured Data Processing Platform for Large-Scale Data Mining. Transactions of the Japanese Society for Artificial Intelligence, 2011, 26, 594-606.	0.1	0
10054	Semantic Technologies and Cloud Computing. , 2011, , 239-251.		5
10055	Evaluating the Performance and Scalability of MapReduce Applications on X10. Lecture Notes in Computer Science, 2011, , 46-57.	1.0	3
10056	A Parallel Method for Unpacking Original High Speed Rail Data Based on MapReduce. Advances in Intelligent and Soft Computing, 2011, , 59-68.	0.2	0

#	ARTICLE	IF	CITATIONS
10057	An Application for Processing Large and Non-Uniform Media Objects on MapReduce-Based Clusters. , 2011, , 667-683.		3
10058	Programming for Distributed Computing: From Physical to Logical Networks. , 2011, , 155-168.		0
10059	Adaptive Parallelization of Queries Calling Dependent Data Providing Web Services. Lecture Notes in Business Information Processing, 2011, , 132-154.	0.8	1
10060	Introducing a Novel Data Management Approach for Distributed Large Scale Data Processing in Future Computer Clouds. Lecture Notes in Computer Science, 2011, , 391-398.	1.0	0
10061	ComputErl â€“ Erlang-Based Framework for Many Task Computing. Lecture Notes in Computer Science, 2011, , 184-197.	1.0	0
10062	HANDLING DATA SKEW IN MAPREDUCE. , 2011, , .		19
10063	CONTENT-BASED RECOMMENDATION ALGORITHMS ON THE HADOOP MAPREDUCE FRAMEWORK. , 2011, , .		2
10064	Building a Faceted Browser in CouchDB Using Views on Views and Erlang Metaprogramming. Lecture Notes in Computer Science, 2011, , 104-121.	1.0	0
10065	A STATISTICAL APPROACH FOR IDENTIFYING MEMORY LEAKS IN CLOUD APPLICATIONS. , 2011, , .		1
10066	A New Industry-Centred Module on Structured Parallel Programming. , 2011, , 127-137.		0
10067	Many-Core Architecture Oriented Parallel Algorithm Design for Computer Animation. Lecture Notes in Computer Science, 2011, , 180-191.	1.0	0
10069	Massive Image Editing on the Cloud. , 2011, , .		1
10070	NEFOS: Rapid Cache-Aware Range Query Processing with Probabilistic Guarantees. Lecture Notes in Computer Science, 2011, , 62-77.	1.0	0
10072	MapReduce Platform for Parallel Machine Learning on Large-scale Dataset. Transactions of the Japanese Society for Artificial Intelligence, 2011, 26, 621-637.	0.1	0
10073	A Semantic Similarity Analysis for Data Mappings between Heterogeneous XML Schemas. Advances in E-Business Research Series, 2011, , 37-52.	0.2	1
10074	Cloud Computing: A Platform of Services for Services. Communications in Computer and Information Science, 2011, , 91-100.	0.4	1
10075	Modeling Virtual Footprints. International Journal of Agent Technologies and Systems, 2011, 3, 1-17.	0.1	0
10076	Towards Efficient Translation Memory Search Based on Multiple Sentence Signatures. , 0, , .		1

#	ARTICLE	IF	CITATIONS
10077	Realizing Secure Cloud Computing Environment by ECA Rules. International Journal of Advancements in Computing Technology, 2011, 3, 267-274.	0.1	0
10078	Survey of Parallel Data Processing in Context with MapReduce. , 2011, , .		3
10079	MapReduce for HITS Algorithm with Application to Chinese Word Networks. Journal of Convergence Information Technology, 2011, 6, 178-185.	0.1	3
10081	Entreposage et analyse en ligne dans les nuages avec Pig. Ingenierie Des Systemes D'Information, 2011, 16, 139-162.	0.5	0
10082	Evaluation of Effect of Network Energy Consumption in Load Distribution across Data Centers. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2012, , 501-517.	0.2	1
10083	Typing a Core Binary-Field Arithmetic in a Light Logic. Lecture Notes in Computer Science, 2012, , 19-35.	1.0	2
10084	LONG TERM BIOSIGNALS VISUALIZATION AND PROCESSING. , 2012, , .		2
10085	Challenges for Future Platforms, Services and Networked Applications. Notes on Numerical Fluid Mechanics and Multidisciplinary Design, 2012, , 19-30.	0.2	0
10086	Distributed Sampling Storage for Statistical Analysis of Massive Sensor Data. Lecture Notes in Computer Science, 2012, , 233-243.	1.0	0
10087	Scalable Subspace Logistic Regression Models for High Dimensional Data. Lecture Notes in Computer Science, 2012, , 685-694.	1.0	2
10088	Uncertain OLAP over Multidimensional Data Streams: State-of-the-Art Analysis and Research Perspectives. Lecture Notes in Computer Science, 2012, , 275-282.	1.0	1
10089	Abstract State Machines for Data-Parallel Computing. Lecture Notes in Computer Science, 2012, , 136-150.	1.0	0
10090	On-the-Fly Task Execution for Speeding Up Pipelined MapReduce. Lecture Notes in Computer Science, 2012, , 526-537.	1.0	0
10091	Service-Oriented Data Mining in Traditional Chinese Medicine. , 2012, , 75-86.		0
10092	Distributed Media Conversion System over Cloud Environment. Lecture Notes in Electrical Engineering, 2012, , 285-290.	0.3	2
10093	Design and Implementation of A Bicycle Parking System Based on Internet-of-things. , 2012, , .		2
10094	MapReduce-Based Data Stream Processing over Large History Data. Lecture Notes in Computer Science, 2012, , 718-732.	1.0	4
10095	Prediction of Protein-Protein Docking Sites Based on a Cloud-Computing Pipeline. International Journal of Machine Learning and Computing, 2012, , 798-801.	0.8	2

#	ARTICLE	IF	CITATIONS
10096	Hypervisor Agnostic Workload Characterization of Virtual Machines. , 2012, , .		0
10097	La visibilidad mediática de actores políticos en la red: La política gestión de contenidos digitales y la propiedad intelectual en España y en el Reino Unido (The Media Visibility of Political Actors on the Net: Tj ETQq1 1 0.784314 rgBT / Ov	0.4	0
10099	Enhancing Service-Oriented Computing with Software Mobility. Lecture Notes in Computer Science, 2012, , 487-501.	1.0	0
10100	Optimal Trust Mining and Computing on Keyed MapReduce. Lecture Notes in Computer Science, 2012, , 143-150.	1.0	0
10102	HIP: Information Passing for Optimizing Join-Intensive Data Processing Workloads on Hadoop. Lecture Notes in Computer Science, 2012, , 384-391.	1.0	0
10103	Grid, SOA and Cloud Computing. , 2012, , 12-51.		1
10104	Challenges in gigapixel multiscale image formation. , 2012, , .		2
10105	Second Level Parallelism using SIMD Accelerators on Heterogeneous MapReduce Clusters. , 2012, , .		0
10106	Scalable Distributed Architecture for Media Transcoding. Lecture Notes in Computer Science, 2012, , 288-302.	1.0	3
10107	The Computing Framework for Physics Analysis at LHCb. Lecture Notes in Computer Science, 2012, , 189-195.	1.0	0
10108	FLOW-BASED PROGRAMMING AS A SOLUTION FOR CLOUD COMPUTING REQUIREMENTS. , 2012, , .		0
10109	Parallel Phrase Scoring for Extra-large Corpora. Prague Bulletin of Mathematical Linguistics, 2012, , 98, .	0.7	2
10110	MINING ON THE CLOUD - K-means with MapReduce. , 2012, , .		3
10111	MRFM: An Efficient Approach to Spatial Join Aggregate. Lecture Notes in Computer Science, 2012, , 140-150.	1.0	1
10112	Flexible and Efficient Distributed Resolution of Large Entities. Lecture Notes in Computer Science, 2012, , 244-263.	1.0	2
10113	A New Framework for Join Product Skew. Lecture Notes in Computer Science, 2012, , 1-10.	1.0	2
10114	Survey of Public IaaS Cloud Computing API. IEEJ Transactions on Electronics, Information and Systems, 2012, 132, 173-180.	0.1	3
10115	Lightweight Semantics over Web Information Systems Content Employing Knowledge Tags. Lecture Notes in Computer Science, 2012, , 327-336.	1.0	3

#	ARTICLE	IF	CITATIONS
10116	MRKDSBC: A Distributed Background Modeling Algorithm Based on MapReduce. Lecture Notes in Computer Science, 2012, , 668-677.	1.0	1
10117	An Effective Approach to Build Optimal T-way Interaction Test Suites over Cloud Using Particle Swarm Optimization. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2012, , 193-198.	0.2	5
10118	RAVEN “ Boosting Data Analysis for the LHC Experiments. Lecture Notes in Computer Science, 2012, , 206-214.	1.0	2
10120	DISTRIBUTED WEB-SCALE INFRASTRUCTURE FOR CRAWLING, INDEXING AND SEARCH WITH SEMANTIC SUPPORT. Computer Science, 2012, 13, 5.	0.4	1
10121	Big Geospatial Data Processing Made Easy: A Working Guide to GeoSpark. , 2021, , 35-53.		0
10122	Architecting Distributed Domain Reducer in Cloud Environment. International Journal of Computer Applications, 2012, 40, 24-29.	0.2	7
10123	MoHPBGA: Multiobjective Hierarchical Population Balanced Genetic Algorithm using MapReduce. International Journal of Computer Applications, 2012, 40, 1-7.	0.2	1
10124	A Smart Proofreader for All Natural Languages: Achieving Semantic Understanding by Majority Vote. , 2012, 2012, 1-6.		1
10125	High-Dimensional Indexing for Video Retrieval. , 0, , .		2
10126	Parallel and Distributed Data Mining for Astronomy Applications. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series, 2012, , .	0.2	1
10127	mROSE: To Determine Tool Selection and to Understand Model-Driven Software Evolution. International Journal of Computer Applications, 2012, 42, 46-55.	0.2	2
10128	Big Data Management Challenges in a Meteorological Organisation. International Journal of E-Adoption, 2012, 4, 1-14.	1.0	2
10129	Resource Management for Data Intensive Tasks on Grids. , 0, , .		0
10130	Two Paradigms for Individuating Implementations. Journal of Cognitive Science, 2012, 13, 167-179.	0.2	2
10132	Distributed Commuting Augmented Shortest Path Finding for Geo Spatial Database. International Journal of Computer Applications, 2012, 52, 28-31.	0.2	0
10133	High Scalability of HDFS using Distributed Namespace. International Journal of Computer Applications, 2012, 52, 30-37.	0.2	2
10134	Applications and Algorithms for Inference of Huge Phylogenetic Trees: a Review. American Journal of Bioinformatics Research, 2012, 2, 21-26.	0.3	7
10137	Sort-Based Distributed Parallel Data Cube Computation Algorithm using MapReduce. Journal of the Institute of Electronics and Information Engineers, 2012, 49, 196-204.	0.0	1

#	ARTICLE	IF	CITATIONS
10138	Analysis and Learning Frameworks for Large-Scale Data Mining. , 0, , .		0
10139	Application of pretrained deep neural networks to large vocabulary speech recognition. , 0, , .		150
10140	The Method for Real-time Complex Event Detection of Unstructured Big data. Journal of Korea Spatial Information Society, 2012, 20, 99-109.	0.7	1
10141	EXPLOITING DYNAMIC RESOURCE ALLOCATION FOR EFFICIENT PARALLEL DATA PROCESSING IN CLOUD-BY USING NEPHELÊ™S ALGORITHM. International Journal of Computer Applications Technology and Research, 2012, 1, 79-81.	0.1	1
10142	Research on Parallel Classification Algorithms for Large-scale Data. Journal of Convergence Information Technology, 2012, 7, 333-340.	0.1	0
10143	A Reactive Scheduling Strategy Applied On MapReduce OLAM Operators System. Journal of Software, 2012, 7, .	0.6	2
10144	Research on Elastic Job Scheduling Model and Algorithm of Cloud Computing based on Hadoop. International Journal of Advancements in Computing Technology, 2012, 4, 473-483.	0.1	1
10145	The Creation and Placement of VMs and Tasks in Virtualized Hadoop Cluster Environments. Journal of Korea Multimedia Society, 2012, 15, 1499-1505.	0.1	0
10146	Optimization Techniques for Multiple Centrality Computations. Lecture Notes in Social Networks, 2013, , 21-47.	0.8	0
10147	Survey on Task Assignment Techniques in Hadoop. International Journal of Computer Applications, 2012, 59, 15-18.	0.2	3
10148	A Linked Data Reasoner in the Cloud. Lecture Notes in Computer Science, 2013, , 722-726.	1.0	1
10149	Scaling Out Recommender System for Digital Libraries with MapReduce. Lecture Notes in Computer Science, 2013, , 40-47.	1.0	0
10152	Cloud-TM. , 2013, , 192-224.		0
10153	A Middleware for Managing Big-Data Flows. Lecture Notes in Computer Science, 2013, , 410-424.	1.0	2
10154	Risk Aware Approach to Data Confidentiality in Cloud Computing. Lecture Notes in Computer Science, 2013, , 27-42.	1.0	1
10155	Collective Action and Human Computation. , 2013, , 421-445.		0
10156	Ad-Hoc Parallel Data Processing on Pay-As-You-Go Clouds with Nephelê. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2013, , 191-218.	0.5	0
10158	Non-uniform Region Based Clustering Routing Algorithm for Wireless Sensor Networks. , 0, , .		1

#	ARTICLE	IF	CITATIONS
10159	Provisioning and Maintenance. Advanced Topics in Science and Technology in China, 2013, , 1-20.	0.0	0
10160	StorageBIT - A Metadata-aware, Extensible, Semantic and Hierarchical Database for Biosignals. , 2013, , .		3
10161	Parallel Data Processing. , 2013, , 113-120.		0
10162	A Methodology for Identifying the Relationships Between Performance Factors for Cloud Computing Applications. Computer Communications and Networks, 2013, , 331-358.	0.8	0
10163	Large-Scale and High-Performance Computing. , 2013, , 1092-1105.		0
10164	The Challenge of Service Level Scalability for the Cloud. , 2013, , 37-48.		0
10165	Cloud Computing Diagnosis. , 2013, , 1-18.		1
10167	Data Clustering on Taiwan Crop Sales Under Hadoop Platform. , 2013, , 827-835.		3
10168	Social Network Path Analysis Based on HBase. Lecture Notes in Computer Science, 2013, , 770-779.	1.0	0
10169	Designing Similarity Indexes with Parallel Genetic Programming. Lecture Notes in Computer Science, 2013, , 294-299.	1.0	0
10170	Advanced Analytics with the SAP HANA Database. , 2013, , .		1
10171	SIG: Solution to TCP Incast in SDN Network Based Openflow Protocol. , 2013, , .		0
10172	Learning Cloud Computing and Cloud Security By Simulation. , 2013, , .		0
10173	Analysis Cloud - Running Sensor Data Analysis Programs on a Cloud Computing Infrastructure. , 2013, , .		1
10174	Heuristics-Based Workload Analysis for Relational DBMSs. Lecture Notes in Business Information Processing, 2013, , 25-36.	0.8	2
10179	Evaluation of a Self-adapting Method for Resource Classification in Folksonomies. Advances in Intelligent Systems and Computing, 2013, , 1-12.	0.5	0
10180	A Distributed Architecture for Multimedia File Storage, Analysis and Processing. Studies in Computational Intelligence, 2013, , 435-452.	0.7	2
10181	Querying Sentiment Development over Time. Lecture Notes in Computer Science, 2013, , 613-624.	1.0	2

#	ARTICLE	IF	CITATIONS
10184	Uncertain Subgraph Query Processing over Uncertain Graphs. Lecture Notes in Computer Science, 2013, , 132-139.	1.0	1
10185	MapReduce-Based Parallel Algorithm for Detecting and Resolving of Firewall Policy Conflict. Communications in Computer and Information Science, 2013, , 118-131.	0.4	0
10186	A Generic Tree-Like Index Framework in the Cloud. Lecture Notes in Computer Science, 2013, , 330-342.	1.0	0
10187	Modeling and Storing Complex Network with Graph-Tree. Advances in Intelligent Systems and Computing, 2013, , 305-315.	0.5	1
10188	Comparative Analysis of Andrew Files System and Hadoop Distributed File System. Lecture Notes on Software Engineering, 2013, , 122-125.	0.3	4
10189	Scaling Out the Performance of Service Monitoring Applications with BlockMon. Lecture Notes in Computer Science, 2013, , 253-255.	1.0	4
10190	The QUASIT Model and Framework for Scalable Data Stream Processing with Quality of Service. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2013, , 92-107.	0.2	5
10191	Can Relational DBMS Scale Up to the Cloud?. , 2013, , 317-328.		4
10192	A Platform to Support Object Database Research.. Journal of Object Technology, 2013, 12, 2:1.	0.8	0
10193	Personal Health Records System Based on App Engine. Lecture Notes in Electrical Engineering, 2013, , 653-659.	0.3	1
10194	Performance Enhancemnt of Map-Reduce Framework on GPU. , 2013, , .		1
10195	An Efficient Algorithm Based on MapReduce For Computing Frequent Item sets. , 2013, , .		0
10196	A Mechanism for Dependability Processing of Information Flow in Cloud Computing. , 0, , .		0
10197	Dependability Testing of MapReduce Systems. , 2013, , .		0
10198	Development Fundamental Technologies for the Multi-Scale Mass-Deployable Cooperative Robots. Journal of the Korean Society for Precision Engineering, 2013, 30, 11-17.	0.1	2
10199	Atmosphere: A Universal Cross-Cloud Communication Infrastructure. Lecture Notes in Computer Science, 2013, , 163-182.	1.0	1
10200	The Strategy of Classification Mining Based on Cloud Computing. , 0, , .		1
10201	Eliminating Homogeneous Cluster Setup for Efficient Parallel Data Processing. International Journal of Computer Applications, 2013, 64, 1-7.	0.2	2

#	ARTICLE	IF	CITATIONS
10202	A New Massive Data Processing Framework under Cloud Environment for Digital Community. Information Technology Journal, 2013, 12, 1079-1088.	0.3	0
10203	A SaaS Development Platform based on Cloud Computing. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013, 11, .	0.1	1
10204	A Tree Algorithm Based on Parallel Cloud Computing Model. Advances in Information Sciences and Service Sciences, 2013, 5, 645-653.	0.1	0
10205	Web Portal for Matching Loan Requests and Investment Offers in Peer-To-Peer Lending. International Journal of Web Portals, 2013, 5, 17-31.	1.1	0
10206	COMPUTING DOT – PRODUCT ON HETEROGENEOUS MASTER – WORKER PLATFORMS. International Journal of Pure and Applied Mathematics, 2013, 84, .	0.2	0
10208	Improved QoS by Clustering Identically Modulated WSN Nodes using Reprogrammable High Speed Architecture. International Journal of Computer Applications, 2013, 68, 28-33.	0.2	0
10211	Performance Tuning and Scheduling of Large Data Set Analysis in Map Reduce Paradigm by Optimal Configuration using Hadoop. International Journal of Computer Applications, 2013, 70, 37-41.	0.2	4
10212	Research on the Reconfigurable Network System Based on the Task Decomposition. Information Technology Journal, 2013, 12, 1873-1883.	0.3	1
10213	A Framework of High-Performance Computation Algorithms for Processing Satellite Remote Sensing Data. Journal of Information and Computational Science, 2013, 10, 2391-2397.	0.1	0
10214	Data-Replicas Scheduler for Heterogeneous MapReduce Cluster. Journal of Networks, 2013, 8, .	0.4	0
10215	Estudio sobre algoritmos genĂ©ticos en la nube y el modelo de programaci3n MapReduce. Enseñanza Y Aprendizaje De Ingenieria De Computadores, 2013, , .	0.4	0
10216	Cloud Offering Patterns. , 2014, , 79-150.		1
10217	Provisioning Mapreduce for Improving Security of Cloud Data. Journal of Artificial Intelligence, 2013, 6, 220-228.	0.7	1
10218	Computation in Hyperspectral Imagery (HSI) Data Analysis: Role and Opportunities. , 2014, , 1-27.		0
10219	A Transitive Recommendation System for Information Technology Communities. Acta Informatica Pragensia, 2013, 2, 1-17.	0.7	0
10220	Survey of Fault Tolerance Policy for Load Balancing Scheme in Distributed Computing. International Journal of Computer Applications, 2013, 74, 1-6.	0.2	2
10221	Optical OFDM-based Data Center Networks. Journal of Networks, 2013, 8, .	0.4	2
10223	Cross Cloud Mapreduce: A Result Integrity Check Framework on Hybrid Clouds. Services Transactions on Cloud Computing, 2013, 1, 26-39.	0.1	2

#	ARTICLE	IF	CITATIONS
10224	A Privacy Data-Oriented Hierarchical MapReduce Programming Model. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013, 11, .	0.1	3
10225	Research on Personalized Behaviors Recommendation System Based on Cloud Computing. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013, 12, .	0.1	4
10226	Assured Information Sharing (AIS)Using Private Clouds. , 2014, , 215-255.		1
10227	The Method of Analyzing Firewall Log Data using MapReduce based on NoSQL. Journal of the Korea Institute of Information Security and Cryptology, 2013, 23, 667-677.	0.1	2
10229	Using MapReduce Streaming for Distributed Life Simulation on the Cloud. , 0, , .		5
10230	UX Analysis for Mobile Devices Using MapReduce on Distributed Data Processing Platform. KIPS Transactions on Software and Data Engineering, 2013, 2, 589-594.	0.1	0
10231	Electricity Consumption Forecasting in the Age of Big Data. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013, 11, .	0.1	3
10232	Analysis of big data using Rhipe. Journal of the Korean Data and Information Science Society, 2013, 24, 975-987.	0.0	7
10233	Secure Collaborative Processing Architecture for MITB Attack Detection. International Journal of Network Security and Its Applications, 2013, 5, 83-93.	0.4	2
10235	Performance Analysis of Hadoop Map Reduce on Eucalyptus Private Cloud. International Journal of Computer Applications, 2013, 79, 10-13.	0.2	5
10236	Modeling Incast and its Empirical Validation. Springer Briefs in Electrical and Computer Engineering, 2014, , 31-65.	0.3	0
10237	Research of Business Engineering Process Optimization Based on MapReduce Framework. Lecture Notes in Electrical Engineering, 2014, , 233-242.	0.3	0
10238	The Efficient Method of Parallel Genetic Algorithm using MapReduce of Big Data. Journal of Korean Institute of Intelligent Systems, 2013, 23, 385-391.	0.0	3
10239	Research on Patient Privacy Protection for Medical Data in Cloud Computing. Journal of Networks, 2013, 8, .	0.4	3
10240	MCHITS: Monte Carlo based Method for Hyperlink Induced Topic Search on Networks. Journal of Networks, 2013, 8, .	0.4	0
10241	IT Architecture of Multiple Heterogeneous Data. Lecture Notes in Electrical Engineering, 2014, , 963-971.	0.3	0
10242	Machine Learning Algorithm Acceleration Using Hybrid (CPU-MPP) MapReduce Clusters. , 2014, , 129-153.		0
10243	Customer Analyst for the Telecom Industry. , 2014, , 101-127.		2

#	ARTICLE	IF	CITATIONS
10244	A Parallel Computing Method of Social Tag Cooccurrence Relation. Applied Mathematics and Information Sciences, 2013, 7, 2525-2531.	0.7	0
10247	Pre-arrangement Based Task Scheduling Scheme for Reducing MapReduce Job Processing Time. Journal of the Korea Society of Computer and Information, 2013, 18, 23-30.	0.0	0
10248	The Family of Map-Reduce. , 2014, , 1-39.		1
10249	A Novel Model for Home Media Streaming Service in Cloud Computing Environment. International Journal of Smart Home, 2013, 7, 265-274.	0.6	2
10250	A Platform for Massive Railway Information Data Storage. Lecture Notes in Electrical Engineering, 2014, , 799-808.	0.3	0
10252	CoCell: A Low-Diameter, High-Performance Data Center Network Architecture. Lecture Notes in Electrical Engineering, 2014, , 429-436.	0.3	0
10253	A New Data-Intensive Parallel Processing Framework for Spatial Data. Lecture Notes in Electrical Engineering, 2014, , 375-383.	0.3	2
10254	Research on Optimization of Resources Allocation in Cloud Computing Based on Structure Supportiveness. Lecture Notes in Electrical Engineering, 2014, , 849-858.	0.3	1
10255	Acceleration of Solving Non-Equilibrium Ionization via Tracer Particles and MapReduce on Eulerian Mesh. Lecture Notes in Computer Science, 2014, , 29-42.	1.0	2
10257	Embedding GPU Computations in Hadoop. International Journal of Networked and Distributed Computing, 2014, 2, 211.	1.3	5
10258	Cloud Computing and Big Data. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2014, , 118-132.	0.5	0
10259	Comparison of conjugate gradient method and Jacobi method algorithm on MapReduce framework. Applied Mathematical Sciences, 0, 8, 837-849.	0.0	2
10261	Parallel Data Processing. , 2014, , 117-124.		0
10262	User Management in Information Security Engineering Environment ISEE. Lecture Notes in Computer Science, 2014, , 18-34.	1.0	0
10263	Nonvolatile Memory Computing System. , 2014, , 131-180.		0
10264	Privacy Preserving OLAP Data Cubes. , 2014, , 1886-1897.		0
10265	Dimming Relations for the Efficient Analysis of Concurrent Systems via Action Abstraction. Lecture Notes in Computer Science, 2014, , 216-231.	1.0	0
10266	An Effective and Efficient Middleware for Supporting Distributed Query Processing in Large-Scale Cyber-Physical Systems. Lecture Notes in Computer Science, 2014, , 124-135.	1.0	0

#	ARTICLE	IF	CITATIONS
10268	InvarNet-X: A Comprehensive Invariant Based Approach for Performance Diagnosis in Big Data Platform. Lecture Notes in Computer Science, 2014, , 124-140.	1.0	0
10269	Automatic Self-Suspended Task for a MapReduce System on Cloud Computing. Lecture Notes in Computer Science, 2014, , 257-268.	1.0	0
10270	Equi-join for Multiple Datasets Based on Time Cost Evaluation Model. Lecture Notes in Computer Science, 2014, , 122-135.	1.0	0
10271	A Parallel and Scalable Framework for Non-overlapping Community Detection Algorithms. Lecture Notes in Computer Science, 2014, , 115-126.	1.0	0
10273	Understanding Vertical Scalability of I/O Virtualization for MapReduce Workloads: Challenges and Opportunities. Lecture Notes in Computer Science, 2014, , 3-12.	1.0	3
10274	Methodology to Determine Relationships between Performance Factors in Hadoop Cloud Computing Applications. , 2014, , .		0
10276	E-Business und Cloud Computing. , 2014, , 53-75.		0
10277	Ensuring Faultless Communication Behaviour in A Commercial Cloud. Lecture Notes in Computer Science, 2014, , 44-55.	1.0	0
10278	Visualization of Semantic Data Based on Selected Predicates. Lecture Notes in Computer Science, 2014, , 180-195.	1.0	0
10279	On the Complexity of List Ranking in the Parallel External Memory Model. Lecture Notes in Computer Science, 2014, , 384-395.	1.0	6
10280	Patent Literatures Translation System Based on Hadoop. Lecture Notes in Electrical Engineering, 2014, , 127-135.	0.3	2
10281	Parallel Ward Clustering for Chemical Compounds Using MapReduce. Communications in Computer and Information Science, 2014, , 258-267.	0.4	2
10282	Hadoop on a Low-Budget General Purpose HPC Cluster in Academia. Advances in Intelligent Systems and Computing, 2014, , 187-192.	0.5	0
10283	Desidoo, a Big-Data Application to Join the Online and Real-World Marketplaces. Advances in Intelligent Systems and Computing, 2014, , 169-175.	0.5	1
10285	Cloud Computing: Efficient Congestion Control in Data Center Networks. Computer Communications and Networks, 2014, , 143-177.	0.8	0
10286	A Novel Heuristic Scheme for Modeling and Managing Time Bound Constraints in Data-Intensive Grid and Cloud Infrastructures. Lecture Notes in Computer Science, 2014, , 172-191.	1.0	0
10287	Adaptive Affinity Propagation Clustering in MapReduce Environment. Lecture Notes in Computer Science, 2014, , 209-218.	1.0	0
10288	Semantics Based Intelligent Search in Large Digital Repositories Using Hadoop MapReduce. Lecture Notes in Computer Science, 2014, , 292-295.	1.0	0

#	ARTICLE	IF	CITATIONS
10289	MapReduce Performance Optimization Based on Block Aggregation. Advances in Intelligent Systems and Computing, 2014, , 853-861.	0.5	0
10290	Distributed Data Platform System Based on Hadoop Platform. Advances in Intelligent Systems and Computing, 2014, , 533-539.	0.5	0
10291	Mining of Web Server Logs in a Distributed Cluster Using Big Data Technologies. International Journal of Advanced Computer Science and Applications, 2014, 5, .	0.5	7
10292	A Visual Programming Approach to Big Data Analytics. Lecture Notes in Computer Science, 2014, , 393-404.	1.0	0
10293	Assessment, Design and Implementation of a Private Cloud for MapReduce Applications. Open Access Library Journal (oalib), 2014, 01, 1-10.	0.1	0
10294	Design and implementation of electronic authentication for web contents. Contemporary Engineering Sciences, 0, 7, 691-697.	0.2	0
10295	A Throughput Driven Task Scheduler for Batch Jobs in Shared MapReduce Environments. International Journal of Big Data, 2014, 1, 15-24.	0.6	0
10296	KOSHIK- A Large-scale Distributed Computing Framework for NLP. , 2014, , .		2
10297	Distributed and Parallel Computing. , 2014, , 1-30.		1
10298	Training Backpropagation Neural Network in MapReduce. , 2014, , .		3
10299	An Adaptable Infrastructure to Generate Training Datasets for Decompilation Issues. Advances in Intelligent Systems and Computing, 2014, , 85-94.	0.5	2
10300	A Micro-benchmark Suite for Evaluating HDFS Operations on Modern Clusters. Lecture Notes in Computer Science, 2014, , 129-147.	1.0	10
10301	Parallel Data Processing in Dynamic Hybrid Computing Environment Using MapReduce. Lecture Notes in Computer Science, 2014, , 1-14.	1.0	1
10302	Taxonomy Outline of Big Data Analytics Literature. , 2014, , 2456-2471.		0
10303	Design of Technology Value Analysis System Based on Patent Big Data. Lecture Notes in Computer Science, 2014, , 48-58.	1.0	0
10304	Efficient Reliability in Volunteer Storage Systems with Random Linear Coding. Lecture Notes in Computer Science, 2014, , 559-569.	1.0	0
10305	Technology Development for Satellite Data Warehouse Design. Izvestiya of Altai State University, 2014, , 172-175.	0.1	0
10306	An Efficient OLAP Query Algorithm Based on Dimension Hierarchical Encoding Storage and Shark. Lecture Notes in Computer Science, 2014, , 180-187.	1.0	0

#	ARTICLE	IF	CITATIONS
10307	Automatic Generation of Optimized Workflow for Distributed Computations on Large-Scale Matrices. Lecture Notes in Computer Science, 2014, , 79-92.	1.0	0
10309	Non-clairvoyant Reduction Algorithms for Heterogeneous Platforms. Lecture Notes in Computer Science, 2014, , 270-279.	1.0	0
10310	Overlapping Community Detection with a Maximal Clique Enumeration Method in MapReduce. Advances in Intelligent Systems and Computing, 2014, , 367-376.	0.5	1
10311	A System for Task Scheduling and Task Migration in Cloud Environment. IOSR Journal of Computer Engineering, 2014, 16, 115-118.	0.1	1
10312	A Distributed Approach For Chinese Micro-blog Hot Topic Detection. , 0, , .		0
10313	Big Data Processing Systems. , 2014, , 135-176.		2
10314	Energy Consumption Prediction by Using an Integrated Multidimensional Modeling Approach and Data Mining Techniques with Big Data. Lecture Notes in Computer Science, 2014, , 45-54.	1.0	3
10315	Parallelizing a CAD Model Processing Tool from the Automotive Industry. Lecture Notes in Computer Science, 2014, , 24-35.	1.0	7
10316	Scalable Distributed Genetic Algorithm for Data Ordering Problem with Inversion Using MapReduce. IFIP Advances in Information and Communication Technology, 2014, , 325-334.	0.5	1
10319	Association Rule Mining using Apriori Algorithm for Distributed System: a Survey. IOSR Journal of Computer Engineering, 2014, 16, 112-118.	0.1	0
10320	OLAP over Uncertain and Imprecise Data Streams. , 2014, , 1670-1679.		0
10321	An Improved Monte Carlo Ray Tracing for Large-Scale Rendering in Hadoop. , 0, , .		0
10322	Large-Scale Chinese Cross-Document Entity Disambiguation and Information Fusion. Lecture Notes in Computer Science, 2014, , 105-119.	1.0	0
10323	The FutureGrid Testbed for Big Data. , 2014, , 27-59.		0
10324	Compression Planner for Time Series Database with GPU Support. Lecture Notes in Computer Science, 2014, , 36-63.	1.0	4
10325	HDStore: An SSD/HDD Hybrid Distributed Storage Scheme for Large-Scale Data. Lecture Notes in Computer Science, 2014, , 209-220.	1.0	4
10326	Parallel Gene Clustering Using MapReduce. Lecture Notes in Computer Science, 2014, , 372-381.	1.0	1
10327	A Pattern Approach to Conquer the Data Complexity in Simulation Workflow Design. Lecture Notes in Computer Science, 2014, , 21-38.	1.0	4

#	ARTICLE	IF	CITATIONS
10328	Accelerate K-means Algorithm by Using GPU in the Hadoop Framework. Lecture Notes in Computer Science, 2014, , 177-186.	1.0	3
10329	Mobile Video Cloud Networks. Advances in Wireless Technologies and Telecommunication Book Series, 2014, , 157-182.	0.3	3
10330	Relational Data Access for Business Data Analytics. , 2014, , 2020-2027.		0
10331	An Insight to Data Stream Mining and other Emerging Learning Algorithms. IOSR Journal of Computer Engineering, 2014, 16, 19-24.	0.1	1
10332	Chrum: The Tool for Convenient Generation of Apache Oozie Workflows. Studies in Computational Intelligence, 2014, , 187-198.	0.7	0
10333	StreamMyRelevance!. Lecture Notes in Computer Science, 2014, , 272-289.	1.0	0
10334	Computing Power Utilization of Distributed Systems Using Distributed Compilation: A Clustered HPC Approach. British Journal of Mathematics & Computer Science, 2014, 4, 2884-2900.	0.3	0
10335	Data Intensive Technologies for Financial Regulatory Systems. SSRN Electronic Journal, 0, , .	0.4	0
10336	Research on Quickly Search in Massive Remote Sensing Images Based on Hbase. , 0, , .		0
10337	DSIMBench: A Benchmark for Microarray Data Using R. Lecture Notes in Computer Science, 2014, , 47-56.	1.0	2
10338	DIPAR: A Framework for Implementing Big Data Science in Organizations. Computer Communications and Networks, 2014, , 177-194.	0.8	1
10339	Visualization of Semantic Data Based on Selected Predicates. Lecture Notes in Computer Science, 2014, , 180-195.	1.0	1
10340	Synergic Impact of Inhibitors of Cloud Computing - FISM and FMICMAC Approach. International Journal of Computer Applications, 2014, 87, 24-29.	0.2	1
10341	Hadoop Mapreduce Framework in Big Data Analytics. International Journal of Computer Trends and Technology, 2014, 8, 115-119.	0.1	1
10342	Checkpoint and Replication Oriented Fault Tolerant Mechanism for MapReduce Framework. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013, 12, .	0.1	3
10343	Implementation and Analysis of Iterative MapReduce Based Heuristic Algorithm for Solving N-Puzzle. Journal of Computers, 2014, 9, .	0.4	1
10344	Supporting Practices in Professional Communities Using Mobile Cloud Services. , 2014, , 47-80.		1
10345	Job Scheduling in the Grid Computing using Criteria. International Journal of Computer Applications, 2014, 90, 5-9.	0.2	0

#	ARTICLE	IF	CITATIONS
10346	Supporting Practices in Professional Communities Using Mobile Cloud Services. , 2014, , 69-102.		0
10347	ChordMR: A P2P-based Job Management Scheme in Cloud. Journal of Networks, 2014, 9, .	0.4	5
10348	Measuring Hadoop Optimality by Lorenz Curve. Ungyong T'onggye Yon'gu = the Korean Journal of Applied Statistics, 2014, 27, 249-261.	0.0	0
10349	Data-parallel finite-state machines. ACM SIGPLAN Notices, 2014, 49, 529-542.	0.2	3
10350	Enhancing User-friendliness of the User Taste Prediction Service Using MapReduce Framework. International Journal of Multimedia and Ubiquitous Engineering, 2014, 9, 263-272.	0.3	0
10351	Preserving Privacy of Cloud Data Using Homomorphic Encryption in MapReduce. International Journal of Hybrid Information Technology, 2014, 7, 363-376.	0.6	0
10352	GEO DISTRIBUTED PARALLELIZATION PACTS IN MAP REDUCE FRAMEWORK IN MULTIPLE DATACENTER. International Journal of Research in Engineering and Technology, 2014, 03, 162-167.	0.1	0
10353	A Temporal Reasoning based Social Calendar Framework. International Journal of Multimedia and Ubiquitous Engineering, 2014, 9, 273-280.	0.3	0
10354	MapReduce Based Experimental Frame for Parallel and Distributed Simulation Using Hadoop Platform. , 2014, , .		4
10355	Multimedia Databases. , 2014, , 14-1-14-28.		0
10356	Extreme Big Data (EBD): Next Generation Big Data Infrastructure Technologies Towards Yottabyte/Year. Supercomputing Frontiers and Innovations, 2014, 1, .	0.5	6
10357	Parallel Clone Code Detector in MapReduce. Journal of Software, 2014, 9, .	0.6	0
10359	iMapReduce: Extending MapReduce for Iterative Processing. , 2014, , 107-126.		1
10360	A Distributed Phoenix++ Framework for Big Data Recommendation Systems. International Journal of Intelligent Computing Research, 2014, 5, 408-415.	0.5	0
10361	A Study on a Way to Utilize Big Data Analytics in the Defense Area. Journal of the Korean Operations Research and Management Science Society, 2014, 39, 1-19.	0.1	5
10362	Decombined Distributed Parallel VQ Codebook Generation Based on MapReduce. Journal of Digital Contents Society, 2014, 15, 365-371.	0.1	3
10363	An Extraction Method of Sentiment Infromation from Unstructured Big Data on SNS. Journal of Korea Multimedia Society, 2014, 17, 671-680.	0.1	6
10364	MapReduce Family of Large-Scale Data-Processing Systems. , 2014, , 54-121.		1

#	ARTICLE	IF	CITATIONS
10365	Distributed Programming for the Cloud: Models, Challenges, and Analytics Engines. , 2014, , 16-53.		0
10366	DESIGN OF FILE SYSTEM ARCHITECTURE WITH CLUSTER FORMATION ALONG WITH MOUNT TABLE. International Journal of Research in Engineering and Technology, 2014, 03, 418-422.	0.1	0
10368	END-TO-END BIG DATA PROCESSING PROTECTION IN CLOUD ENVIRONMENT USING BLACK BOXES - AN FPGA APPROACH. Services Transactions on Cloud Computing, 2014, 2, 14-27.	0.1	1
10369	Research on the Security based on Utility Theory in Cloud Computing Environment. International Journal of Security and Its Applications, 2014, 8, 321-328.	0.5	1
10370	Algorithm Optimization of Resources Scheduling Based on Cloud Computing. Journal of Multimedia, 2014, 9, .	0.3	0
10371	OLAPing and Mining Big Data: Large-Scale, Long-Running, Serendipitous Computations within Next-Generation Clouds. , 2014, , .		0
10372	Processing Method of Mass Small File Using Hadoop Platform. The Journal of Advanced Navigation Technology, 2014, 18, 401-408.	0.0	2
10373	Review on the Cloud Computing Programming Model. International Journal of Advanced Science and Technology, 2014, 70, 11-16.	0.3	2
10374	Backoff inspired features for maximum entropy language models. , 0, , .		5
10375	Improvisation to the R*-Tree kNN Join Principles in Distributed Environment. International Journal of Computer Applications, 2014, 101, 20-24.	0.2	0
10376	Performance Comparison of DW System Tajo Based on Hadoop and Relational DBMS. KIPS Transactions on Software and Data Engineering, 2014, 3, 349-354.	0.1	1
10377	Distributed Iterative Interference Cancellation and MapReduce Paradigm for Cooperative Network. Journal of Information and Computational Science, 2014, 11, 5447-5457.	0.1	0
10378	HBase-based Automatic Summary System using Twitter Trending Topics. Journal of Internet Computing and Services, 2014, 15, 63-72.	0.1	0
10379	Distributed Storage and Processing Method for Big Data Sensing Information of Machine Operation Condition. Journal of Software, 2014, 9, .	0.6	1
10380	Evaluation of Efficient Implementation of Big Data Switch for Iraqi Cellular Phone Service Providers. International Journal of Computer Applications, 2014, 104, 1-5.	0.2	0
10381	Approximate String Similarity Join using Hashing Techniques under Edit Distance Constraints. Journal of Software, 2014, 9, .	0.6	2
10382	Distinction of Discrete Transformations Applied to Hadoop's MapReduce. International Journal of Computer Applications, 2014, 104, 35-38.	0.2	0
10383	TecnologÃas <i>big data</i> para anÃ¡lisis y recuperaciÃ³n de imÃ¡genes web. Profesional De La Informacion, 2014, 23, 567-574.	2.7	2

#	ARTICLE	IF	CITATIONS
10384	Collaborative Defence for Distributed Attacks (Case Study of Palestinian Information Systems). International Journal of Network Security and Its Applications, 2014, 6, 31-37.	0.4	0
10386	A Distributed Cache Management Scheme for Efficient Accesses of Small Files in HDFS. The Journal of the Korea Contents Association, 2014, 14, 28-38.	0.0	0
10387	Functional programming for dynamic and large data with self-adjusting computation. ACM SIGPLAN Notices, 2014, 49, 227-240.	0.2	1
10388	Futuristic Smart Architecture for a Rapid Disaster Response. Lecture Notes in Social Networks, 2015, , 39-64.	0.8	0
10389	Securing MapReduce Result Integrity via Verification-based Integrity Assurance Framework. International Journal of Grid and Distributed Computing, 2014, 7, 53-70.	0.8	4
10390	A SURVEY ON NGS - SHORT READ ALIGNMENT IN HIGH PERFORMANCE COMPUTING. International Journal of Research in Engineering and Technology, 2014, 03, 84-88.	0.1	0
10391	Adaptive Resource Management Method base on ART in Cloud Computing Environment. Journal of the Korea Society for Simulation, 2014, 23, 111-119.	0.0	0
10392	An Algorithm of Association Rule Based on Cloud Computing. Open Automation and Control Systems Journal, 2014, 6, 1748-1753.	0.9	0
10393	Distributed REScala. ACM SIGPLAN Notices, 2014, 49, 361-376.	0.2	5
10394	Programming Models for Cloud Computing. , 2014, , 212-237.		0
10395	Research on Parallelization of KNN Locally Weighted Linear Regression Algorithm Based on MapReduce. Journal of Communications, 2015, , .	1.3	2
10397	Reliable Multipath Transmission in Data Center Network. Journal of Communications, 2015, , .	1.3	3
10398	Python in Data Science Research and Education. , 2015, , .		5
10400	Efficient Query Algorithm of Coallocation -Parallel-Hash-Join in the Cloud Data Center. Lecture Notes in Computer Science, 2015, , 306-320.	1.0	0
10401	Multi-Column Query Method Research and Optimization on HBase. Lecture Notes in Business Information Processing, 2015, , 414-421.	0.8	0
10403	NYSOL: A User-Centric Framework for Knowledge Discovery in Big Data. International Journal of Knowledge Engineering, 2015, 1, 214-218.	0.2	1
10404	Text Relevance Analysis Method over Large-Scale High-Dimensional Text Data Processing. Lecture Notes in Computer Science, 2015, , 371-379.	1.0	0
10405	Evaluating NoSQL Databases for Big Data Processing within the Brazilian Ministry of Planning, Budget, and Management. Advances in Web Technologies and Engineering Book Series, 2015, , 230-247.	0.4	0

#	ARTICLE	IF	CITATIONS
10406	Hybrid Computation Models for High Performance Biological Sequence Alignment on a Cloud System. , 2015, , 369-379.		0
10407	Parallel Programming Models and Systems for High Performance Computing. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2015, , 254-292.	0.5	0
10408	Optical Data Center Networks: Architecture, Performance, and Energy Efficiency. , 2015, , 351-391.		3
10409	Research on Design Features of Intelligent Product based on Big Data. , 0, , .		0
10410	Study on Cloud Storage based on the MapReduce for Big Data. , 2015, , .		0
10411	Constant time median filtering of extra large images using Hadoop. , 2015, , .		0
10412	Anomaly Detection in Cloud Environments. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2015, , 43-67.	0.5	0
10413	ReHRS: A Hybrid Redundant System for Improving MapReduce Reliability and Availability. Modeling and Optimization in Science and Technologies, 2015, , 187-209.	0.7	3
10414	A Result Verification Scheme for MapReduce Having Untrusted Participants. Advances in Intelligent Systems and Computing, 2015, , 11-19.	0.5	2
10415	Zukunft der globalen GeodÄsie und Fernerkundung aus Sicht des Deutschen GeoForschungsZentrum (GFZ), Potsdam. , 2015, , 1-55.		0
10416	The Design of Log Analysis Mechanism in SDN Quarantined Network System. Lecture Notes in Electrical Engineering, 2015, , 331-336.	0.3	0
10417	An Extreme Learning Approach to Fast Prediction in the Reduce Phase of a Cloud Platform. Lecture Notes in Computer Science, 2015, , 417-423.	1.0	0
10418	Large-Scale Web Traffic Log Analyzer on Hadoop Distributed File System. , 2015, , .		0
10419	Multi-Level Queue Dominant Resource Fairness in Cloud Computing. , 0, , .		0
10420	Research on security threats and Countermeasures for Cloud Computing. , 0, , .		0
10421	An Efficient MapReduce Framework for Intel MIC Cluster. Lecture Notes in Computer Science, 2015, , 129-139.	1.0	1
10422	DistDL: A Distributed Deep Learning Service Schema with GPU Accelerating. Lecture Notes in Computer Science, 2015, , 793-804.	1.0	3
10424	Parallelized Parameter Estimation of Biological Pathway Models. Lecture Notes in Computer Science, 2015, , 37-57.	1.0	4

#	ARTICLE	IF	CITATIONS
10425	Accessing Big Data in the Cloud Using Mobile Devices. , 2015, , 222-248.		0
10426	A Forensic-as-a-Service Delivery Platform for Law Enforcement Agencies. , 2015, , 2288-2306.		0
10427	Mapreduce Based Selective Neural Network Ensembles Using GA. Journal of Information and Computational Science, 2015, 12, 693-701.	0.1	0
10429	Research Overview of Big Data. , 2015, , .		0
10430	Research on Parallelized Stream Data Micro Clustering Algorithm. , 2015, , .		0
10431	An Efficient Data Extracting Method Based on Hadoop. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2015, , 87-97.	0.2	0
10432	A Research on Machine Learning Methods for Big Data Processing. , 2015, , .		1
10433	A Virtual Network Performance Optimization Strategy for Cloud-based Big Data Processing. , 2015, , .		0
10434	Linguistic Aggregation Functions using the MapReduce Paradigm. , 0, , .		0
10435	Distributed Stream Processing of RNN Query in Mobile Computing. , 0, , .		0
10436	Distributed Confidence-Weighted Classification on Big Data Platforms. Handbook of Statistics, 2015, 33, 145-168.	0.4	1
10437	Converting Instance Checking to Subsumption: A Rethink for Object Queries over Practical Ontologies. International Journal of Intelligence Science, 2015, 05, 44-62.	0.6	2
10438	Intelligent Management and Efficient Operation of Big Data. Advances in Data Mining and Database Management Book Series, 2015, , 105-129.	0.4	1
10439	A Proposal of a Big Web Data Application and Archive for the Distributed Data Processing with Apache Hadoop. Lecture Notes in Computer Science, 2015, , 285-294.	1.0	1
10441	On the Network Performance of Digital Evidence Acquisition of Small Scale Devices over Public Networks. Digital Forensics, Security and Law Journal, 2015, , .	0.0	0
10442	H.265 Video Streaming in Mobile Cloud Networks. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2015, , 238-277.	0.5	0
10443	Research on Chinese segmentation algorithm based on Hadoop cloud platform. , 0, , .		0
10444	Running Genome Wide Data Analysis Using a Parallel Approach on a Cloud Platform. Lecture Notes in Computer Science, 2015, , 188-192.	1.0	0

#	ARTICLE	IF	CITATIONS
10445	Analysis of the Big Data based on MapReduce. , 0, , .		0
10446	An Improved Frag-Shells Algorithm for Data Cube Construction Based on Irrelevance of Data Dispersion. Communications in Computer and Information Science, 2015, , 292-299.	0.4	0
10447	Datenbanken. , 2015, , 337-370.		0
10448	Getting Started with Spark. Computer Communications and Networks, 2015, , 73-99.	0.8	0
10449	Parallel Collaborative Filtering Recommendation Model Based on Two-Phase Similarity. Lecture Notes in Computer Science, 2015, , 1-10.	1.0	0
10451	Design and Implementation of an Efficient Web Services Data Processing Using Hadoop-Based Big Data Processing Technique. Journal of the Korea Academia-Industrial Cooperation Society, 2015, 16, 726-734.	0.0	2
10452	A Cloud-based Efficient On-line Analytical Processing System with Inverted Data Model. , 2015, , .		0
10453	Terabyte-Scale Image Similarity Search. Handbook of Statistics, 2015, , 279-301.	0.4	0
10454	Towards Truly Elastic Distributed Graph Computing in the Cloud. Lecture Notes in Computer Science, 2015, , 300-309.	1.0	0
10456	Contribution Resource Fairness with Time Constraints in Cloud Computing. , 0, , .		0
10457	Temporal Verification for Scientific Cloud Workflows: State-of-the-Art and Research Challenges. Communications in Computer and Information Science, 2015, , 59-74.	0.4	0
10459	Introduction: Data Science and BigData Computing. , 2015, , 3-15.		1
10460	Policy Driven Node Selection in MapReduce. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2015, , 55-72.	0.2	1
10461	Quality-of-Service in Data Center Stream Processing for Smart City Applications. , 2015, , 1047-1076.		0
10463	Exploring the Hamming Distance in Distributed Infrastructures for Similarity Search. Modeling and Optimization in Science and Technologies, 2015, , 1-28.	0.7	0
10464	HBelt: Integrating an Incremental ETL Pipeline with a Big Data Store for Real-Time Analytics. Lecture Notes in Computer Science, 2015, , 123-137.	1.0	1
10465	Cloud-TM. , 2015, , 749-781.		0
10466	A Survey on Research Initiatives for Healthcare Clouds. , 2015, , 170-186.		0

#	ARTICLE	IF	CITATIONS
10467	Efficient processing and utilizing Big Data in E-commerce. , 0, , .		0
10468	A Dynamic Extension and Data Migration Method Based on PVFS. Lecture Notes in Computer Science, 2015, , 540-552.	1.0	0
10469	Bringing Search Engines to the Cloud using Open Source Components. , 2015, , .		1
10470	Availability and Network-Aware MapReduce Task Scheduling over the Internet. Lecture Notes in Computer Science, 2015, , 214-227.	1.0	0
10471	MapReduce Job Scheduling Based on Remaining Job Sizes. IEICE Transactions on Communications, 2015, E98.B, 180-189.	0.4	0
10472	Dynamically Scheduling a Component-Based Framework in Clusters. Lecture Notes in Computer Science, 2015, , 129-146.	1.0	4
10473	Benchmarking Support Vector Machines Implementation Using Multiple Techniques. Advances in Intelligent Systems and Computing, 2015, , 227-238.	0.5	5
10474	Deadline-Oriented Task Scheduling for MapReduce Environments. Lecture Notes in Computer Science, 2015, , 359-372.	1.0	3
10476	A MapReduce-based Approach for Finding Inexact Patterns in Large Graphs. , 2015, , .		0
10477	Performance analysis for a cloud computing application. , 2015, , .		0
10478	Data Science and Distributed Intelligence. , 2015, , 1732-1740.		0
10479	Assessment of the Pillar 3 Financial and Risk Information Disclosures Usefulness to the Commercial Banks Users. Lecture Notes in Computer Science, 2015, , 429-440.	1.0	0
10480	Research and Application of Hadoop Monitoring based on JMX. , 0, , .		0
10481	Distributed Recommendation Algorithm Based on Matrix Decomposition on MapReduce Framework. Lecture Notes in Computer Science, 2015, , 447-457.	1.0	0
10482	Optimized-Memory Map-Reduce Algorithm for Mobile Learning. Lecture Notes in Computer Science, 2015, , 249-256.	1.0	1
10483	Query Execution Optimization Based on Incremental Update in Database Distributed Middleware. Lecture Notes in Computer Science, 2015, , 257-270.	1.0	1
10484	A Cloud-Friendly Communication-Optimal Implementation for Strassen's Matrix Multiplication Algorithm. IEICE Transactions on Information and Systems, 2015, E98.D, 1896-1905.	0.4	0
10485	Mobile Video Cloud Networks. , 2015, , 1000-1024.		0

#	ARTICLE	IF	CITATIONS
10486	Research On Software Testing Technology Under Big Data Background. , 2015, , .		0
10489	Design of a flow integrating information system based on cloud computing. , 0, , .		0
10490	Comparison of Machine Learning Algorithm on Map Reduction for Performance Improvement in Big Data. Indian Journal of Science and Technology, 2015, 8, 1-6.	0.5	3
10491	Selected Approaches and Frameworks to Carry out Genomic Data Analysis on the Cloud. Scalable Computing, 2015, 16, .	0.7	1
10492	A New HDFS Structure Model to Evaluate The Performance of Word Count Application on Different File Size. International Journal of Computer Applications, 2015, 111, 1-4.	0.2	2
10493	A Task Scheduling Algorithm Based on Potential Games in Cloud Computing Environment. International Journal of Future Generation Communication and Networking, 2015, 8, 247-260.	0.7	0
10494	Smart Travel Guide for Android-Enabled Devices. International Journal of Advanced Research in Computer and Communication Engineering, 2015, , 320-322.	0.1	1
10495	Non Relational Models for the Management of Large Amount of Astronomical Data. Acta Polytechnica CTU Proceedings, 2015, 2, 293-296.	0.3	0
10496	¾ Scalability and Cost Evaluation of Incremental Data Processing Using Amazon™s Hadoop Service. , 2015, , 58-75.		1
10497	Mining Infrequent Weighted Itemsets from Very Large Data based on MapReduce Model. , 2015, , .		0
10498	Data Processing for Big Data Applications using Hadoop Framework. International Journal of Advanced Research in Computer and Communication Engineering, 2015, , 177-180.	0.1	2
10499	The Network Factor in Proactive Digital Evidence Acquisition. International Journal of Intelligent Computing Research, 2015, 6, 517-526.	0.5	1
10500	Efficient Storage Management over Cloud using Data Compression without Losing Searching Capacity. International Journal of Computer Applications, 2015, 114, 1-6.	0.2	0
10501	L-INDEX: A Conceptual Index of Textual Data on MapReduce Platform. International Journal of Grid and Distributed Computing, 2015, 8, 213-234.	0.8	1
10502	Parallel Chameleon Clustering Based on MapReduce. Journal of Information and Computational Science, 2015, 12, 2053-2062.	0.1	1
10503	A New MapReduce Framework Based on Virtual IP Mechanism and Load Balancing Strategy. Open Cybernetics and Systemics Journal, 2015, 9, 253-261.	0.3	1
10504	A MapReduce-based kNN Join Query Processing Algorithm for Analyzing Large-scale Data. Journal of KIISE, 2015, 42, 504-511.	0.0	0
10505	SCALABLE ALGORITHM FOR THE SERVICE SELECTION PROBLEM. Services Transactions on Services Computing, 2015, 3, 10-24.	0.3	0

#	ARTICLE	IF	CITATIONS
10506	Review of several improved Apriori algorithms on Hadoop MapReduce environment. , 2015, , .		0
10507	Document Classification Using Distributed Machine Learning. , 2015, , .		2
10509	Quantitative Analysis of Document Stored Databases. International Journal of Computer Applications, 2015, 118, 37-41.	0.2	1
10510	CommGuard. ACM SIGPLAN Notices, 2015, 50, 311-323.	0.2	1
10511	Approach to Big Data Applications on Cloud System. , 2015, , .		0
10514	An Efficient Distributed Data Management Method based key Columns Partition Preprocessing. International Journal of Database Theory and Application, 2015, 8, 203-214.	0.2	0
10515	Identification of Similar Strings in a Dataset using Scalable Join. International Journal of Computer Applications, 2015, 120, 25-28.	0.2	0
10516	Survey on Data Processing and Scheduling in Hadoop. International Journal of Computer Applications, 2015, 119, 27-30.	0.2	0
10518	Cost Efficient Query Optimization in Mobile Environment. International Journal of Computer Applications, 2015, 120, 30-38.	0.2	5
10519	Applications Exploiting Multimedia Semantics. , 2015, , 177-200.		0
10520	Mining Interesting Infrequent Itemsets from Very Large Data based on MapReduce Framework. International Journal of Intelligent Systems and Applications, 2015, 7, 44-49.	0.9	4
10521	An Efficient Clustering Method based on Multi Centroid Set using MapReduce. KIISE Transactions on Computing Practices, 2015, 21, 494-499.	0.1	1
10522	Applying Mapreduce and Network Traffic Analysis to Control Access to Information Resources. St Petersburg State Polytechnical University Journal Computer Science Telecommunications and Control Systems, 2015, 217-222, 34-40.	0.1	0
10523	Research and implementation of K-means algorithm based on Hama. , 2015, , 37-40.		0
10524	¾ Big Data Management Challenges, Approaches, Tools, and Their Limitations. , 2015, , 65-78.		2
10525	Improving MapReduce Performance through Process Migration. International Journal of Engineering Research & Technology, 2015, V4, .	0.2	0
10526	Effective and Efficient Way of Reduce Dependency on Dataset with the Help of Mapreduce on Big Data. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2015, 15, .	0.1	3
10527	Handling Endogeneity Challenge in Big Astronomical Data. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, 8, 63-78.	0.2	0

#	ARTICLE	IF	CITATIONS
10528	Implementation and Performance Analysis of Academic_MapReduce Algorithm (AcdMR). International Journal of Computer Applications, 2015, 121, 17-20.	0.2	0
10529	Cuckoo Linear Algebra. , 2015, , .		1
10530	Efficient and Fault Tolerant Computation of Partially Idempotent Tasks. , 2015, , .		5
10531	An Algorithms for Tournament-based Big Data Analysis. Journal of Digital Contents Society, 2015, 16, 545-553.	0.1	0
10532	A Novel Method of Improving Cache Hit-rate in Hadoop MapReduce using SSD Cache. Journal of the Korea Society of Computer and Information, 2015, 20, 1-6.	0.0	0
10533	A Phoenix++ Based New Genetic Algorithm Involving Mechanism of Simulated Annealing. International Journal of Distributed Sensor Networks, 2015, 11, 806708.	1.3	0
10534	A Development Study of The VPT for the improvement of Hadoop performance. The Journal of the Korean Institute of Information and Communication Engineering, 2015, 19, 2029-2036.	0.1	0
10535	An Approach to Implement Map Reduce with NoSQL Databases. International Journal of Engineering and Computer Science, 0, , .	0.2	0
10536	Finding Frequent Route of Taxi Trip Events Based on MapReduce and MongoDB. KIPS Transactions on Software and Data Engineering, 2015, 4, 347-356.	0.1	0
10537	Customized Web User Interface for Hadoop Distributed File System. Advances in Intelligent Systems and Computing, 2016, , 567-576.	0.5	0
10538	Speculation-aware Cluster Scheduling. Performance Evaluation Review, 2015, 43, 42-44.	0.4	2
10539	A Comprehensive Survey on Centroid Selection Strategies for Distributed K-means Clustering Algorithm. International Journal of Computer Applications, 2015, 125, 36-42.	0.2	3
10540	Energy ICT convergence with big data services. Journal of the Korean Data and Information Science Society, 2015, 26, 1141-1154.	0.0	1
10541	EFFECTIVE AND EFFICIENT WAY OF REDUCE DEPENDENCY ON DATASET WITH THE HELP OF MAPREDUCE ON BIG DATA. International Journal of Students Research in Technology & Management, 2015, 3, 401-405.	0.1	0
10542	Towards big data processing in clouds: An online cost-minimization approach. Big Data & Information Analytics, 2015, 1, 15-29.	1.3	2
10543	A Secure Framework For Hadoop Communication. International Journal of Scientific and Engineering Research, 2015, 6, 1409-1416.	0.1	0
10544	The Design of Method for Efficient Processing of Small Files in the Distributed System based on Hadoop Framework. The Journal of the Korea Institute of Electronic Communication Sciences, 2015, 10, 1115-1122.	0.1	3
10545	Multi-Objective Bottleneck-Aware Allocation of Multiple Resources. , 2015, , .		0

#	ARTICLE	IF	CITATIONS
10547	Case Study of Genome Sequencing on an FPGA: Survey and a New Perspective. , 2015, , 73-94.		0
10548	Improving Hadoop Hive Query Response Times Through Efficient Virtual Resource Allocation. Advances in Intelligent Systems and Computing, 2016, , 215-225.	0.5	2
10549	Improve Constrained Recourse Fairness in Cloud Computing with Bottleneck-Aware Allocation. International Journal of U- and E- Service, Science and Technology, 2015, 8, 415-424.	0.1	0
10550	AES â€œ MR: A Novel Encryption Scheme for securing Data in HDFS Environment using MapReduce. International Journal of Computer Applications, 2015, 129, 12-19.	0.2	1
10551	A Cost Effective Virtual Cluster with Hadoop Framework for Big Data Analytics. International Journal of Database Theory and Application, 2015, 8, 199-214.	0.2	1
10552	A Fast and Scalable Image Retrieval Algorithms by Leveraging Distributed Image Feature Extraction on MapReduce. Journal of KIISE, 2015, 42, 1474-1479.	0.0	1
10553	Industrial-Scale Ad Hoc Risk Analytics Using MapReduce. Studies in Big Data, 2016, , 177-206.	0.8	0
10554	Monitoring and Notification System Based on Risk Analysis Using Map-Reduce Framework. International Journal of Computer Theory and Engineering, 2015, 7, 444-447.	3.2	1
10555	Real-Time Pedestrian Detection Using Apache Storm in a Distributed Environment. , 2015, , .		1
10556	A Queuing based Network I/O Scheduling with QoS Guarantee for Virtual Machine. International Journal of Future Generation Communication and Networking, 2015, 8, 123-134.	0.7	0
10557	Clustering Algorithm using the DFP-Tree based on the MapReduce. Journal of Internet Computing and Services, 2015, 16, 23-30.	0.1	0
10559	Research on Parallel Particle Filtering Target Tracking Algorithm Based on Hadoop. , 0, , .		1
10560	Big Data Paradigm for Healthcare Sector. , 2016, , 570-587.		1
10561	Data-Intensive Analytics for Cat Bonds by Considering Supply Chain Risks. Lecture Notes in Computer Science, 2016, , 136-147.	1.0	0
10562	Medical Image Dataset Processing over Cloud/MapReduce with Heterogeneous Architectures. , 2016, , 1-10.		0
10563	A Multi-commodity Network Flow Model for Cloud Service Environments. Communications in Computer and Information Science, 2016, , 186-197.	0.4	0
10564	TLCSim: A Large-Scale Two-Level Clustering Similarity Search with MapReduce. Lecture Notes in Computer Science, 2016, , 53-71.	1.0	0
10565	Research on the Advanced Computing Method for Supporting Large Data Quality Assessment and Improvement. , 2016, , .		0

#	ARTICLE	IF	CITATIONS
10566	Effective Task Scheduling for Large-Scale Video Processing. Lecture Notes in Computer Science, 2016, , 323-331.	1.0	1
10567	On the Dynamic Shifting of the MapReduce Timeout. Advances in Data Mining and Database Management Book Series, 2016, , 1-22.	0.4	0
10568	Information Technology Systems. , 2016, , 189-218.		2
10569	Research and Realization of AP Clustering Algorithm Based on Cloud Computing. , 0, , .		0
10570	An RDMA Middleware for Asynchronous Multi-stage Shuffling in Analytical Processing. Lecture Notes in Computer Science, 2016, , 61-74.	1.0	3
10572	Applying the K-Means Algorithm in Big Raw Data Sets with Hadoop and MapReduce. , 2016, , 1220-1243.		0
10573	Extracting Location Names from Unstructured Italian Texts Using Grammar Rules and MapReduce. Communications in Computer and Information Science, 2016, , 593-601.	0.4	1
10575	ChainMR Crawler: A Distributed Vertical Crawler Based on MapReduce. Lecture Notes in Computer Science, 2016, , 33-39.	1.0	1
10576	Fast Type Recognition of Missive Insulator Leakage Current Data Using Spark. Journal Electrical and Electronic Engineering, 2016, 4, 51.	0.7	1
10577	Improving Performance in Hadoop Using Automatic and Predictive Configuration. , 2016, , .		0
10578	Datenbanken. , 2016, , 337-370.		0
10579	Evaluating NoSQL Databases for Big Data Processing within the Brazilian Ministry of Planning, Budget, and Management. , 2016, , 1110-1128.		0
10580	Parallel Community Detection on Massive Graphs. , 2016, , .		0
10581	Partitioning Based N-Gram Feature Selection for Malware Classification. Lecture Notes in Computer Science, 2016, , 187-195.	1.0	3
10582	An Improved Content Based Image Retrieval System On Apache Spark. , 2016, , .		1
10583	Survey on Big Data Tools and Techniques. IOSR Journal of Computer Engineering, 2016, 01, 44-50.	0.1	0
10584	MapReduce Programming Platform. Integrated Series on Information Systems, 2016, , 99-119.	0.1	0
10585	H.265 Video Streaming in Mobile Cloud Networks. , 2016, , 906-944.		0

#	ARTICLE	IF	CITATIONS
10586	Scalable Data Mining, Archiving, and Big Data Management for the Next Generation Astronomical Telescopes. , 2016, , 2199-2225.		3
10588	A Review of RDF Storage in NoSQL Databases. , 2016, , 85-104.		0
10589	Using Canonical Correlation Analysis for Parallelized Attribute Reduction. Lecture Notes in Computer Science, 2016, , 433-445.	1.0	0
10590	RProtoBuf: Efficient Cross-Language Data Serialization in<i>R</i>. Journal of Statistical Software, 2016, 71, .	1.8	2
10591	Parallelization and Optimization of a CAD Model Processing Tool from the Automotive Industry to Distributed Memory Parallel Computers. Lecture Notes in Computer Science, 2016, , 377-388.	1.0	0
10592	Nearest Query on Distributed Binary Trees Starting from a Random Node. Communications in Computer and Information Science, 2016, , 257-271.	0.4	0
10593	CLASSIFICATION OF TEXT DATA FROM THE SOCIAL NETWORK TWITTER. , 2016, , .		1
10594	Maximum Satisfaction Scheduling algorithm Based on Hadoop Architecture. , 2016, , .		0
10595	Spatial Join with Hadoop. , 2016, , 1-5.		0
10596	Big Data Paradigm for Healthcare Sector. Advances in Public Policy and Administration, 2016, , 169-186.	0.1	2
10597	Mbench: Benchmarking a Multicore Operating System Using Mixed Workloads. Lecture Notes in Computer Science, 2016, , 50-63.	1.0	0
10598	Big Biological Data Management. Computer Communications and Networks, 2016, , 265-277.	0.8	0
10599	A Comparative Investigation of Sample Versus Normal Map for Effective BigData Processing. Advances in Intelligent Systems and Computing, 2016, , 749-755.	0.5	0
10600	Towards Conceptual Predictive Modeling for Big Data Framework. International Journal of Software Engineering and Its Applications, 2016, 10, 35-42.	0.2	0
10601	Application for Identifying and Monitoring Defaulters in Telecom Network. Advances in Intelligent Systems and Computing, 2016, , 569-585.	0.5	0
10602	Research on Naive Bayesian Spam SMS Filtering Based on MapReduce. Computer Science and Application, 2016, 06, 443-450.	0.0	0
10603	Harmony Search for Self-configuration of Fault-tolerant and Intelligent Grids. Lecture Notes in Computer Science, 2016, , 566-576.	1.0	0
10604	An ARM-Based Hadoop Performance Evaluation Platform: Design and Implementation. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2016, , 82-94.	0.2	2

#	ARTICLE	IF	CITATIONS
10605	IntegrityMR: Exploring Result Integrity Assurance Solutions for Big Data Computing Applications. International Journal of Networked and Distributed Computing, 2016, 4, 116.	1.3	1
10606	Trusted Heartbeat Framework for Cloud Computing. Journal of Information Security, 2016, 07, 103-111.	0.4	2
10607	Accelerating the Shuffle Phase to Speed up MapReduce Systems. , 2016, , .		0
10608	Adaptive Scheduling Strategy for Heterogeneous Spark Cluster. Computer Science and Application, 2016, 06, 692-704.	0.0	1
10609	Big Data Analysis of HPC Technologies From Scientific Computing. SSRN Electronic Journal, 0, , .	0.4	0
10610	Utility Computing and Its Utilization. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2016, , 1-21.	0.5	1
10611	Distributed File System. Integrated Series on Information Systems, 2016, , 79-97.	0.1	0
10612	Distributed Parameter Learning for Probabilistic Ontologies. Lecture Notes in Computer Science, 2016, , 30-45.	1.0	1
10613	Research on Relevance between Atmospheric Temperature and Surface Temperature in Forests Based on Hadoop. Hans Journal of Wireless Communications, 2016, 06, 73-80.	0.0	0
10614	A hybrid ant colony optimization algorithm based on MapReduce. , 2016, , .		1
10616	Mining Frequent Itemsets with Vertical Data Layout in MapReduce. Communications in Computer and Information Science, 2016, , 66-82.	0.4	1
10617	Accelerating Large-Scale Human Action Recognition with GPU-Based Spark. Lecture Notes in Computer Science, 2016, , 670-679.	1.0	0
10618	Open-Source Search Engines in the Cloud. Communications in Computer and Information Science, 2016, , 100-117.	0.4	0
10619	A kind of Prefetching Data Way to Hadoop MapReduce Environments. , 2016, , .		0
10620	Time-, Energy-, and Monetary Cost-Aware Cache Design for a Mobile-Cloud Database System. Lecture Notes in Computer Science, 2016, , 71-85.	1.0	1
10621	Using Managed High Performance Computing Systems for High-Throughput Computing. , 2016, , 61-79.		3
10622	Makespan Minimization for Batch Tasks in Data Centers. Lecture Notes in Computer Science, 2016, , 115-123.	1.0	1
10623	Distributed Programming via Safe Closure Passing. Electronic Proceedings in Theoretical Computer Science, EPTCS, 0, 203, 99-107.	0.8	0

#	ARTICLE	IF	CITATIONS
10624	Risk Identification Method for Cloud Computing Safety based on LSA-GCC and LSA-SAM. International Journal of Grid and Distributed Computing, 2016, 9, 227-244.	0.8	0
10625	Design and Implementation of HDFS Data Encryption Scheme Using ARIA Algorithms on Hadoop. KIPS Transactions on Computer and Communication Systems, 2016, 5, 33-40.	0.1	0
10627	CSR. , 2016, , .		1
10628	CSR. Operating Systems Review (ACM), 2016, 50, 773-787.	1.5	0
10629	Representative Information Retrieval Algorithm Based on PageRank Algorithm and MapReduce Model. International Journal of Database Theory and Application, 2016, 9, 25-36.	0.2	0
10630	A novel LTE network deployment scheme using telecom big data. , 2016, , 261-270.		9
10631	An overlay network for distributed exact and range search in one-dimensional space. Business Informatics, 2016, 2016, 26-36.	1.1	0
10632	Global Disease Monitoring and Forecasting with Wikipedia. Online Journal of Public Health Informatics, 2016, 8, .	0.4	1
10633	Una comparaci3n de rendimiento entre Oracle y MongoDB. Ciencia E IngenierAa Neogranadina, 2016, 26, 109.	0.1	3
10634	Optimizing hash partitioning for solid state drives. , 2016, , .		0
10635	Game theoretical allocation in a cloud datacenter. , 2016, , 67-88.		0
10637	Review on Privacy Preservation by Applying Scalable MapReduce BottomUp Generalization (MRBUG) Technique. International Journal of Computer Applications, 2016, 139, 17-21.	0.2	0
10638	Scalable Distributed Real-Time Processing for Large-Scale Data Streams. International Journal of Database Theory and Application, 2016, 9, 55-64.	0.2	0
10640	Evolution of Incremental Map Reduce Technique in Web Mining. International Journal of Computer Science and Engineering, 2016, 3, 43-45.	0.1	0
10641	Cloudflow - enabling faster biomedical pipelines with MapReduce and Spark. Scalable Computing, 2016, 17, .	0.7	0
10642	Data Optimization Techniques using Bloom Filter in Big Data. International Journal of Computer Applications, 2016, 142, 23-27.	0.2	6
10643	Study on Hadoop Cluster. IOSR Journal of Computer Engineering, 2016, 18, 80-83.	0.1	1
10644	P-Codec: Parallel Compressed File Decompression Algorithm for Hadoop. International Journal of Computers & Technology, 2016, 15, 6991-6998.	0.2	0

#	ARTICLE	IF	CITATIONS
10645	REMOTE ONLINE BIG DATA ANALYTICS ARCHITECTURE. International Journal of Research in Engineering and Technology, 2016, 05, 99-101.	0.1	0
10646	CSR. ACM SIGPLAN Notices, 2016, 51, 773-787.	0.2	0
10647	CHANGE DETECTION OF MOBILE LIDAR DATA USING CLOUD COMPUTING. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XLI-B3, 309-313.	0.2	6
10648	Improving Join Performance for SPARQL Query Processing in the Clouds. Journal of KIISE, 2016, 43, 700-709.	0.0	0
10649	An Empirical Evaluation Analysis of the Performance of In-memory Bigdata Processing Platform. Journal of the Korea Industrial Information Systems Research, 2016, 21, 13-19.	0.1	0
10650	Hadoop based Weblog Analysis: A Review. International Journal of Software Engineering and Its Applications, 2016, 10, 13-30.	0.2	4
10651	An Efficient Parallel Association Rule Mining Algorithm based on Map Reduce Framework. International Journal of Engineering Research & Technology, 2016, V5, .	0.2	1
10652	Bandwidth Optimization In Data Retrieval From Cloud Using Continuous Hive Language. International Journal of Engineering and Computer Science, 0, , .	0.2	0
10653	Improved Indexing & Searching Throughput. ACM SIGIR Forum, 2016, 50, 87-87.	0.4	0
10654	Contrasting Effects of Replication in Parallel Systems. , 2016, , .		9
10655	Heterogeneous Phase-Level Scheduling With Jobs Execution Scheduling Algorithm To Enhance Job Execution And Resource Management In Mapreduce. International Journal of Engineering and Computer Science, 0, , .	0.2	0
10656	Handling Streaming Data by Using Open Source Framework Storm in IoT Environment. KIPS Transactions on Software and Data Engineering, 2016, 5, 313-318.	0.1	0
10657	Internet Public Opinion Hot Spot Mining Based on Voting Mechanism and MapReduce Platform. International Journal of Database Theory and Application, 2016, 9, 83-94.	0.2	0
10658	Distributed Tasks: Introducing Distributed Computing to Programming Competitions. Olympiads in Informatics, 2016, 10, 177-194.	0.2	0
10659	CSR. Computer Architecture News, 2016, 44, 773-787.	2.5	1
10660	Matrix-based Filtering and Load-balancing Algorithm for Efficient Similarity Join Query Processing in Distributed Computing Environment. The Journal of the Korea Contents Association, 2016, 16, 667-680.	0.0	0
10661	JAS: JVM-Based Active Storage Framework for Object-based Storage Systems. International Journal of Database Theory and Application, 2016, 9, 121-134.	0.2	0
10662	Heterogeneous Data Processing using Hadoop and Java Map/Reduce. International Journal of Computer Applications, 2016, 146, 13-16.	0.2	5

#	ARTICLE	IF	CITATIONS
10663	External Merge Sorting in Tajo with Variable Server Configuration. Journal of KIISE, 2016, 43, 820-826.	0.0	0
10664	High Performance CDR Processing with MapReduce. Journal of ICT Research and Applications, 2016, 10, 95-109.	0.5	1
10665	Parallel Computing to Predict Breast Cancer Recurrence on SEER Dataset using Map-Reduce Approach. International Journal of Computer Applications, 2016, 149, 31-35.	0.2	0
10666	Design of Testbed for Agile Computing of MapReduce Applications by using Docker. International Journal of Contents, 2016, 12, 29-33.	0.1	0
10667	MapReduce Integrated Multi-algorithm for HPC Running State Analysis. Telkomnika (Telecommunication Computing Electronics and Control), 2016, 14, 1123.	0.6	0
10668	An object-based meta knowledge model for a distributed image interpretation system. , 0, , .		0
10669	Modular Architecture for Recommender Systems Applied in a Brazilian e-Commerce. Journal of Software, 2016, 11, 912-923.	0.6	2
10670	Clustering Algorithms for Big Data: A Survey. , 2016, , 143-162.		0
10671	Processing large-scale data with Apache Spark. Ungyong T'onggye Yon'gu = the Korean Journal of Applied Statistics, 2016, 29, 1077-1094.	0.0	2
10672	Analysis of Costing Issues in Big Data. , 2016, , 63-77.		0
10673	PISA. , 2016, , .		7
10675	HPC (High-Performance the Computing) for Big Data on Cloud: Opportunities and Challenges. International Journal of Computer Theory and Engineering, 2016, 8, 423-428.	3.2	8
10676	Grammatical Case Based IS-A Relation Extraction with Boosting for Polish. , 0, , .		0
10677	A Peta-Scale Data Movement and Analysis in Data Warehouse (APSDMADW). International Journal of Computer Applications, 2016, 151, 1-5.	0.2	5
10678	Approaching Symbolic Parallelization by Synthesis of Recurrence Decompositions. Electronic Proceedings in Theoretical Computer Science, EPTCS, 0, 229, 55-66.	0.8	0
10679	Framework Implementation of Image-Based Indoor Localization System Using Parallel Distributed Computing. The Journal of Korean Institute of Communications and Information Sciences, 2016, 41, 1490-1501.	0.0	0
10680	Declarative coordination of graph-based parallel programs. ACM SIGPLAN Notices, 2016, 51, 1-12.	0.2	1
10682	Application of the Big Data Technology for Massive Data of the Whole Life Cycle of EMU. Advances in Intelligent Systems and Computing, 2017, , 219-224.	0.5	0

#	ARTICLE	IF	CITATIONS
10683	A Memory Control Mechanism for Dynamic Scheduling in Hadoop. , 2016, , .		0
10684	Design of Processing Model for Connected Car Data Using Big Data Technology. Lecture Notes in Electrical Engineering, 2017, , 143-148.	0.3	1
10685	Obtaining Good Heuristics for Big Data Broadcasting Problem Using Novel Pipeline Approach.. International Journal of Engineering and Computer Science, 0, , .	0.2	0
10686	A Dynamical Load Balancing Method for Data Streaming and User Request in WebRTC Environment. Journal of Digital Contents Society, 2016, 17, 581-592.	0.1	0
10687	GEMs: shared-memory parallel programming for Node.js. ACM SIGPLAN Notices, 2016, 51, 531-547.	0.2	2
10688	A Map Reduce - based SVM Ensemble with Stochastic Gradient Descent. International Journal of Engineering Research & Technology, 2016, V5, .	0.2	0
10689	Significant Big Data Interpretation using Map Reduce Paradigm. International Journal of Computer Applications, 2016, 156, 7-11.	0.2	4
10691	Research on Optimization of Big data Storage Structure in Distributed System. , 0, , .		0
10692	Benchmarking Spark Distributed Data Structures: A Sequence Analysis Case Study. Communications in Computer and Information Science, 2017, , 77-88.	0.4	0
10693	Fast Plagiarism Detection in Large-Scale Data. Communications in Computer and Information Science, 2017, , 329-343.	0.4	0
10694	A Parallel and Resilient Frontend for High Performance Validation Suites. Lecture Notes in Computer Science, 2017, , 248-255.	1.0	0
10695	The Design of Electricity Sale Company Platform Based on qInternet Plusq. , 2017, , .		0
10696	Using the HashChain to Improve the Security of the Hadoop. , 2017, , .		0
10697	Big Data and Big Data Technologies. SpringerBriefs in Pharmaceutical Science & Drug Development, 2017, , 39-58.	0.4	2
10698	A Formal, Resource Consumption-Preserving Translation of Actors to Haskell. Lecture Notes in Computer Science, 2017, , 21-37.	1.0	1
10699	An Improved Approach for Analysis of Hadoop Data for All Files. International Journal of Computer Applications, 2017, 157, 15-20.	0.2	5
10700	Implementing MapReduce Applications in Dynamic Cloud Environments. Computer Communications and Networks, 2017, , 211-223.	0.8	1
10701	FP-MRBP: Fine-grained Parallel MapReduce Back Propagation Algorithm. Lecture Notes in Computer Science, 2017, , 680-687.	1.0	1

#	ARTICLE	IF	CITATIONS
10702	Customized Filesystem with Dynamic Stripe Strategies on Lustre-Based Hadoop. Communications in Computer and Information Science, 2017, , 541-553.	0.4	1
10703	A Correlation Network Model for Structural Health Monitoring and Analyzing Safety Issues in Civil Infrastructures. , 2017, , .		0
10704	An Approach for Cloud-Based Situational Analysis for Factories Providing Real-Time Reconfiguration Services. IFIP Advances in Information and Communication Technology, 2017, , 118-127.	0.5	0
10706	Scalable Multi-agent Simulation Based on MapReduce. Lecture Notes in Computer Science, 2017, , 364-371.	1.0	3
10707	Software Development Methodology for Cloud Computing and Its Impact. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2017, , 286-307.	0.5	1
10708	Bunching and Parallel Empowering Techniques for Hadoop Document Framework. International Journal of Engineering Research & Technology, 2017, V6, .	0.2	0
10711	Mining Perfectly Rare Itemsets on Big Data: An Approach Based on Apriori-Inverse and MapReduce. Advances in Intelligent Systems and Computing, 2017, , 508-518.	0.5	3
10712	A Review of Distributed Data Models for Learning. Lecture Notes in Computer Science, 2017, , 88-97.	1.0	3
10713	On the Use of ISO/IEC Standards to Address Data Quality Aspects in Big Data Analytics Cloud Services. Lecture Notes in Business Information Processing, 2017, , 149-164.	0.8	0
10714	Pragmatic Big Data and smart manufacturing. , 2017, , .		0
10715	An overall approach to achieve load balancing for Hadoop Distributed File System. International Journal of Web and Grid Services, 2017, 13, 448.	0.4	0
10716	Hierarchical Data Summarization. , 2017, , 1-7.		0
10717	Communication-aware virtual machine migration in cloud data centres. International Journal of High Performance Computing and Networking, 2017, 10, 372.	0.4	0
10718	Implement Of Network User Virtual Identity Merging Based on Spark. , 2017, , .		0
10719	WaybackVisor: Hypervisor-Based Scalable Live Forensic Architecture for Timeline Analysis. Lecture Notes in Computer Science, 2017, , 219-230.	1.0	3
10721	Time constraint influence maximization algorithm in the age of big data. International Journal of Computational Science and Engineering, 2017, 15, 165.	0.4	3
10723	Development and research of algorithms for clustering data of super-large volume. , 2017, , .		1
10724	The Experiential Utility. Lecture Notes in Computer Science, 2017, , 121-133.	1.0	0

#	ARTICLE	IF	CITATIONS
10725	Spatial Join with Hadoop. , 2017, , 1-5.		0
10726	Optimization Factor Analysis of Large-Scale Join Queries on Different Platforms. Lecture Notes in Computer Science, 2017, , 35-46.	1.0	1
10727	Big Geo-Data Handling Based on Parallel and Distributed Systemâ€™s Strategies. Geotechnologies and the Environment, 2017, , 85-109.	0.3	0
10728	Application of Batch and Stream Collaborative Computing in Urban Traffic Data Processing. Lecture Notes in Computer Science, 2017, , 725-734.	1.0	0
10729	High Performance Merging of Massive Data from Genome-Wide Association Studies. Lecture Notes in Computer Science, 2017, , 36-40.	1.0	0
10730	A New Distributed Histogram Equalization Processing Remote Sensing Images based on MapReduce Framework. , 2017, , .		0
10731	Map Reduce Autoscaling over the Cloud with Process Mining Monitoring. Communications in Computer and Information Science, 2017, , 109-130.	0.4	6
10733	A Prototype for Sentiment Analysis Using Big Data Tools. Communications in Computer and Information Science, 2017, , 103-117.	0.4	0
10734	Novel Innovation Model of E-Commerce Consumption Based on Big Data Processing and Analysis Technology. , 2017, , .		0
10735	Event Categorization and Key Prospect Identification from Storylines. Communications in Computer and Information Science, 2017, , 62-88.	0.4	0
10736	Model Calibration with Big Data. Conference Proceedings of the Society for Experimental Mechanics, 2017, , 315-322.	0.3	3
10737	From Group-by to Accumulation: Data Aggregation Revisited. , 2017, , .		0
10738	Beyond Map-Reduce: LATNODE â€™ A New Programming Paradigm for Big Data Systems. Lecture Notes in Electrical Engineering, 2017, , 69-75.	0.3	0
10739	Evolution of Cloud in Big Data With Hadoop on Docker Platform. Advances in Information Security, Privacy, and Ethics Book Series, 2017, , 41-62.	0.4	0
10740	A novel feature selection based on apriori property and correlation analysis for protein sequence classification using MapReduce. International Journal of Data Mining and Bioinformatics, 2017, 17, 255.	0.1	0
10741	A Novel Processing Model For Scds In ETL. , 2017, , .		0
10742	Huge Data Analysis and Processing Platform based on Hadoop. , 2017, , .		0
10743	Distributed and Connected Information in the Internet. Advanced Information and Knowledge Processing, 2017, , 153-174.	0.2	0

#	ARTICLE	IF	CITATIONS
10744	Experimental Study on Performance and Energy Consumption of Hadoop in Cloud Environments. Communications in Computer and Information Science, 2017, , 255-272.	0.4	2
10745	Cyber-Physical Systems in Vehicular Communications. Advances in Wireless Technologies and Telecommunication Book Series, 2017, , 477-497.	0.3	3
10746	Medical Image Dataset Processing over Cloud/MapReduce with Heterogeneous Architectures. , 2017, , 1206-1215.		0
10747	Generic and Concurrent Computation of Belief Combination Rules. Communications in Computer and Information Science, 2017, , 125-140.	0.4	0
10749	Simulating cheated results acceptance rates for gossip-based volunteer computing. International Journal of Mobile Network Design and Innovation, 2017, 7, 56.	0.1	0
10750	A reinforcement learning approach to map reduce auto-configuration under networked environment. International Journal of Security and Networks, 2017, 12, 135.	0.1	0
10751	Data mining of unstructured big data in cloud computing. International Journal of Business Intelligence and Data Mining, 2017, 12, 1.	0.2	6
10752	Max-flow Min-cut Algorithm in Spark with Application to Road Networks. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2017, , 12-22.	0.2	1
10757	Study and Experimental Setup of Testing of Performance Parameters on Various Distributed File Systems. International Journal of Computer Applications, 2017, 159, 23-27.	0.2	0
10759	Map-Reduce based Multiple Sub-Graph Enumeration Using Dominating-Set Graph Partition. International Journal of Information Engineering and Electronic Business, 2017, 9, 36-44.	1.0	0
10760	Cancelled: A New Efficient Resource Management Framework for Iterative MapReduce Processing in Large-Scale Data Analysis. IEICE Transactions on Information and Systems, 2017, E100.D, 704-717.	0.4	1
10761	Research on the Design of Parallel Hypernetworks Model Based on Hadoop Cloud Computing Platform. DEStech Transactions on Engineering and Technology Research, 2017, , .	0.0	0
10762	A testbed to enable comparisons between competing approaches for computational social choice. Big Data & Information Analytics, 2017, 1, 309-340.	1.3	0
10763	Extracting Parallel Paragraphs from Common Crawl. Prague Bulletin of Mathematical Linguistics, 2017, 107, 39-56.	0.7	1
10764	MatrixMap: Programming abstraction and implementation of matrix computation for big data analytics. Big Data & Information Analytics, 2017, 1, 349-376.	1.3	0
10765	25 \$\$\$ Years of Query Processing - From a Single, Stored Data Set to Big Data (and Beyond). Studies in Big Data, 2018, , 77-91.	0.8	0
10766	Analysing and Designing Of Traffic Data System Using Big Data. IOSR Journal of Computer Engineering, 2017, 19, 42-47.	0.1	0
10769	Large Graph Computing Systems. , 2017, , 347-362.		0

#	ARTICLE	IF	CITATIONS
10770	Managing Uncertainty in Large-Scale Inversions for the Oil and Gas Industry with Big Data. Studies in Big Data, 2018, , 149-173.	0.8	0
10771	Developing Cloud-Based Tools for Water Resources Data Analysis Using R and Shiny. Lecture Notes on Data Engineering and Communications Technologies, 2018, , 289-297.	0.5	0
10772	Big Data Analytics: A Perspective View. International Journal of Advanced Research in Computer Science and Software Engineering, 2017, 7, 32-41.	0.1	0
10773	The Research on Optimization Method of Hadoop MapReduce. DEStech Transactions on Engineering and Technology Research, 2017, , .	0.0	0
10774	Big Data Analysis for Aids Disease Detection System using Clustering Technique. International Journal of Computer Trends and Technology, 2017, 48, 85-92.	0.1	1
10775	A Review of Data Privacy Issues in Cloud Computing. International Journal of Computer Applications, 2017, 167, 5-7.	0.2	0
10776	BULUT BÄ°LÄ°ÅžÄ°M SÄ°STEMLERÄ°NDE VERÄ°NÄ°N FARKLI BOYUTLARI ÅœZERÄ°NE DERLEME. Å–mer Halisdemir Åœniversitesi MÄ°hendislik Bilimleri Dergisi, 2017, 6, 316-338.	0.2	4
10777	Asynchronous Page-Rank Computation in Spark. Advances in Intelligent Systems and Computing, 2018, , 567-573.	0.5	0
10778	Applying Probability Model to The Genetic Algorithm Based Cloud Rendering Task Scheduling. , 2017, , .		1
10779	SeeMore: A kinetic parallel computer sculpture for educating broad audiences on parallel computation. Journal of Parallel and Distributed Computing, 2017, 105, 183-199.	2.7	1
10780	PROCESSING BIG REMOTE SENSING DATA FOR FAST FLOOD DETECTION IN A DISTRIBUTED COMPUTING ENVIRONMENT. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XLII-4/W2, 137-138.	0.2	1
10781	Tools for the Storage and Analysis of Spatial Big Data. , 0, , .		0
10782	A Hadoop Extension for Analysing Spatiotemporally Referenced Events. Advances in Intelligent Systems and Computing, 2018, , 905-914.	0.5	0
10783	The Performance of Boolean Retrieval and Vector Space Model in Textual Information Retrieval. CommIT Journal, 2017, 11, 33.	0.2	3
10784	A Comprehensive Survey and Open Challenges of Mining Bigdata. Smart Innovation, Systems and Technologies, 2018, , 441-448.	0.5	2
10786	Geospatial Data Science: A Transdisciplinary Approach. , 2017, , 1-40.		0
10787	Viyojit. Computer Architecture News, 2017, 45, 613-626.	2.5	10
10788	Implementation of CouchDBViews. Advances in Intelligent Systems and Computing, 2018, , 241-251.	0.5	0

#	ARTICLE	IF	CITATIONS
10789	Coding for edge-facilitated wireless distributed computing with heterogeneous users. , 2017, , .		12
10790	A Brief Account of Iterative Big Data Clustering Algorithms. International Journal of Computer Sciences and Engineering, 2017, 5, 292-301.	0.0	2
10791	A Proposal for High Availability of HDFS Architecture based on Threshold Limit and Saturation Limit of the Namenode. International Journal of Information Engineering and Electronic Business, 2017, 9, 27-34.	1.0	2
10792	EDA3.0: Implications to Logic Synthesis. , 2018, , 1-20.		0
10793	DIFET: DISTRIBUTED FEATURE EXTRACTION TOOL FOR HIGH SPATIAL RESOLUTION REMOTE SENSING IMAGES. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 0, IV-4/W4, 209-213.	0.0	0
10794	Parallel Bat Algorithm Using MapReduce Model. International Journal of Information Technology and Computer Science, 2017, 9, 72-78.	0.8	2
10795	On Scalability of Predictive Ensembles and Tradeoff Between Their Training Time and Accuracy. Advances in Intelligent Systems and Computing, 2018, , 257-269.	0.5	0
10796	Efficient Graph Mining on Heterogeneous Platforms in the Cloud. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2018, , 12-21.	0.2	1
10797	A Review on Business Intelligence and Big Data. International Journal of Intelligent Systems and Applications in Engineering, 2017, 4, 206-215.	1.0	6
10798	Performance Observation of Apache Spark and Map Reduce Through K-Means. , 2017, , .		0
10799	BioHIPI: Biomedical Hadoop Image Processing Interface. Lecture Notes in Computer Science, 2018, , 540-548.	1.0	0
10800	Hadoop Based Parallel Deduplication Method for Web Documents. Lecture Notes in Electrical Engineering, 2018, , 499-504.	0.3	0
10801	A Survey of High-Performance Computing for Software Verification. Communications in Computer and Information Science, 2018, , 196-208.	0.4	2
10802	Data Locality and Dependency for MapReduce. , 2017, , 293-330.		0
10803	PLZMA: A Parallel Data Compression Method for Cloud Computing. Lecture Notes in Computer Science, 2018, , 504-518.	1.0	0
10804	Hierarchical Data Summarization. , 2018, , 1695-1701.		0
10805	Towards Distributed Model Analytics with Apache Spark. , 2018, , .		1
10806	On the Performance of Spark on HPC Systems: Towards a Complete Picture. Lecture Notes in Computer Science, 2018, , 70-89.	1.0	2

#	ARTICLE	IF	CITATIONS
10807	JSPRE: A Large-Scale Detection of Malicious JavaScript Code Based on Pre-filter. Lecture Notes in Computer Science, 2018, , 586-599.	1.0	1
10808	Trends and Challenges in Large-Scale HPC Network Analysis. Advances in Computer and Electrical Engineering Book Series, 2018, , 144-170.	0.2	0
10809	Hybrid algorithms for 3-SAT optimisation using MapReduce on clouds. International Journal of Innovative Computing and Applications, 2018, 9, 44.	0.2	0
10810	Iterative MapReduce. Advances in Computer and Electrical Engineering Book Series, 2018, , 117-132.	0.2	0
10811	Dichotomy of Fog Computing in the Realm of Cloud Computing: Exploring the Emerging Dimensions. , 2018, , 45-62.		0
10812	Information Retrieval and Search Engines. , 2018, , 259-304.		6
10813	An Efficient Partitioning Technique in SpatialHadoop. International Journal of Intelligent Computing and Information Sciences, 2018, 18, 1-13.	0.3	0
10814	Dynamic Case Bases and the Asymmetrical Weighted One-Mode Projection. Lecture Notes in Computer Science, 2018, , 385-398.	1.0	3
10815	Big Spectrum Data and Deep Learning Techniques for Cognitive Wireless Networks. Advances in Data Mining and Database Management Book Series, 2018, , 33-60.	0.4	0
10816	A Parallel Implementation of Information Gain Using Hive in Conjunction with MapReduce for Continuous Features. Lecture Notes in Computer Science, 2018, , 283-294.	1.0	0
10817	Big Data Processing on Cloud Computing Using Hadoop Mapreduce and Apache Spark. Advances in Business Information Systems and Analytics Book Series, 2018, , 224-250.	0.3	0
10818	A Big Data Approach for Traffic Classification using Data Mining. International Journal for Research in Applied Science and Engineering Technology, 2018, 6, 3293-3298.	0.1	0
10820	Mining Patterns Using Business Process Management. Advances in Web Technologies and Engineering Book Series, 2018, , 78-89.	0.4	2
10821	MapReduce Performance in MongoDB Sharded Collections. International Journal of Advanced Computer Science and Applications, 2018, 9, .	0.5	1
10822	Implementation Maximum Overall Coverage Constraint Non-negative Matrix Factorization for Hyperspectral Mixed Pixels Analysis Using MapReduce. Lecture Notes in Computer Science, 2018, , 417-426.	1.0	0
10823	Toward Performance Prediction for Multi-BSP Programs in ML. Lecture Notes in Computer Science, 2018, , 159-174.	1.0	1
10824	Parallel Processing of Graphs. Data-centric Systems and Applications, 2018, , 143-162.	0.2	0
10825	QKnob: A Knob-Based Fairness-Efficiency Scheduler for Cloud Computing with QoS Guarantees. Lecture Notes in Computer Science, 2018, , 837-853.	1.0	4

#	ARTICLE	IF	CITATIONS
10826	Two Rounds Are Enough for Reconstructing Any Graph (Class) in the Congested Clique Model. Lecture Notes in Computer Science, 2018, , 134-148.	1.0	1
10827	Subgroup Discovery. , 2018, , 71-98.		0
10828	Big Data and FinTech. Computational Social Sciences, 2018, , 139-163.	0.4	3
10829	Accelerating EGENMR Database Operations Using GPU Processing. Advances in Intelligent Systems and Computing, 2018, , 669-679.	0.5	1
10830	A Tag-Based Protection Method for Multi-tenant Data Security. Lecture Notes in Computer Science, 2018, , 553-565.	1.0	2
10831	Efficient spark analysis on incremental datasets. , 2018, , .		0
10832	BSP Programming Model. , 2018, , 1-5.		0
10833	Routing Optimization Using an Enhanced Protocol for Wireless Sensor Networks with Chain Structure. Journal of Communications, 2018, , 691-701.	1.3	1
10835	Reducing Stage Weight Estimation Error of Slow Task Detection in MapReduce Scheduling. Advances in Intelligent Systems and Computing, 2018, , 284-291.	0.5	0
10836	Mining Associations Between Two Categories Using Unstructured Text Data in Cloud. Advances in Intelligent Systems and Computing, 2018, , 545-550.	0.5	0
10837	Columnar Storage Formats. , 2018, , 1-6.		1
10838	The Refutation of Amdahl's Law and Its Variants. Lecture Notes in Computer Science, 2018, , 79-96.	1.0	0
10839	Cloud Computing Resource Scheduling Optimization Based on Chaotic Firefly Algorithm Based on the Tent Mapping. Communications in Computer and Information Science, 2018, , 118-128.	0.4	0
10840	Performance Comparison of Graph BFS Implemented in MapReduce and PGAS Programming Models. Lecture Notes in Computer Science, 2018, , 328-337.	1.0	1
10841	Reliability Evaluation of DCell Networks. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2018, , 213-225.	0.2	0
10842	Mining Multimodal Big Data. Advances in Data Mining and Database Management Book Series, 2018, , 674-702.	0.4	0
10843	Software Tools. Intelligent Systems Reference Library, 2018, , 157-171.	1.0	0
10844	Data Mining as a Cloud Service for Learning Artificial Intelligence. Lecture Notes in Computer Science, 2018, , 214-221.	1.0	0

#	ARTICLE	IF	CITATIONS
10846	Emerging trend of big data analytics in bioinformatics: a literature review. International Journal of Bioinformatics Research and Applications, 2018, 14, 144.	0.1	6
10848	Improving the Performance of Actor-Based Programs Using a New Actor to Thread Association Technique. Lecture Notes in Computer Science, 2018, , 122-136.	1.0	0
10850	Efficient Implementation of Hadoop MapReduce-Based Dataflow. Advances in Web Technologies and Engineering Book Series, 2018, , 372-385.	0.4	2
10851	In-Memory Distributed Training of Linear-Chain Conditional Random Fields with an Application to Fine-Grained Named Entity Recognition. Lecture Notes in Computer Science, 2018, , 155-167.	1.0	1
10852	Nesting Strategies for Enabling Nimble MapReduce Dataflows for Large RDF Data. , 2018, , 811-838.		0
10853	High Performance Storage for Big Data Analytics and Visualization. Advances in Data Mining and Database Management Book Series, 2018, , 254-275.	0.4	0
10854	Research on Algorithms for Planted (l,d) Motif Search. Communications in Computer and Information Science, 2018, , 136-143.	0.4	0
10855	MapReduce – The Scalable Distributed Data Processing Solution. , 2018, , 173-190.		0
10856	Mimir+: An Optimized Framework of MapReduce on Heterogeneous High-Performance Computing System. Lecture Notes in Computer Science, 2018, , 164-168.	1.0	1
10857	Distributed Collaborative Execution on the Edges and Its Application on AMBER Alert. SpringerBriefs in Computer Science, 2018, , 37-58.	0.2	1
10858	Evolving System Bottlenecks in the as a Service Cloud. SSRN Electronic Journal, 0, , .	0.4	0
10859	A Survey on Parallel Database Systems from a Storage Perspective: Rows Versus Columns. Communications in Computer and Information Science, 2018, , 5-20.	0.4	1
10860	Distributed Incremental View Maintenance. , 2018, , 1-9.		0
10861	Parallel Processing with Big Data. , 2018, , 1-7.		1
10862	Computing. Wireless Networks, 2018, , 35-49.	0.3	0
10863	Big Data in Massive Parallel Processing. Advances in Data Mining and Database Management Book Series, 2018, , 276-302.	0.4	1
10864	Efficient Aggregation Query Processing for Large-Scale Multidimensional Data by Combining RDB and KVS. Lecture Notes in Computer Science, 2018, , 134-149.	1.0	2
10865	The Assessing of Influence of Collective Intelligence on the Final Consensus Quality. Lecture Notes in Computer Science, 2018, , 15-24.	1.0	1

#	ARTICLE	IF	CITATIONS
10866	Cohesive Sub-network Mining in Protein Interaction Networks Using Score-Based Co-clustering with MapReduce Model (MR-CoC). <i>Advances in Intelligent Systems and Computing</i> , 2018, , 227-236.	0.5	3
10867	Aspect-Combining Functions for Modular MapReduce Solutions. <i>International Journal of Advanced Computer Science and Applications</i> , 2018, 9, .	0.5	0
10868	Research and Comparison of SQL Optimization Techniques based on MapReduce. <i>DEStech Transactions on Computer Science and Engineering</i> , 2018, , .	0.1	0
10869	Elephant Against Goliath: Performance of Big Data Versus High-Performance Computing DBSCAN Clustering Implementations. <i>Communications in Computer and Information Science</i> , 2018, , 251-271.	0.4	0
10870	Proving Equivalence Between Imperative and MapReduce Implementations Using Program Transformations. <i>Electronic Proceedings in Theoretical Computer Science, EPTCS</i> , 0, 268, 185-199.	0.8	1
10871	Big Data and Shipping-managing vessel performance. <i>International Journal on Informatics Visualization</i> , 2018, 2, 73-75.	0.5	2
10873	Importance of MapReduce for Big Data Applications: A Survey. <i>Asian Journal of Computer Science and Technology</i> , 2018, 7, 112-118.	0.1	1
10874	Parallel Optimal Grid-Clustering algorithm exploration on MapReduce Framework. <i>International Journal of Computer Applications</i> , 2018, 180, 35-39.	0.2	0
10875	In Search of Actionable Patterns of Lowest Cost - A Scalable Graph Method. <i>International Journal of Database Management Systems</i> , 2018, 10, 01-19.	0.2	2
10876	IOT SÄ°STEMLERÄ° Ä°Ä±Ä°N VERÄ°TABANI SÄ°STEM Ä–NERÄ°LERÄ°. <i>MuÄŸla Journal of Science and Technology</i> , 2018, 4, 46-52.		
10878	Product Usage Data Collection and Challenges of Data Anonymization. <i>Lecture Notes on Data Engineering and Communications Technologies</i> , 2019, , 117-136.	0.5	2
10879	Facilitating Dynamic RT-QoS for Massive-Scale Autonomous Cyber-Physical Systems. <i>IEICE Transactions on Communications</i> , 2018, E101.B, 1760-1767.	0.4	0
10880	CBHiveQL: Context-Based Hive Query Language for Large Data Analysis. <i>Studies in Computational Intelligence</i> , 2019, , 149-161.	0.7	0
10883	Treeplication: An Erasure Code that is Almost as Painless as Replication. , 2018, , .		0
10885	Enhancement of the K-Means Algorithm for Mixed Data in Big Data Platforms. <i>Advances in Intelligent Systems and Computing</i> , 2019, , 1025-1040.	0.5	4
10886	A Design of High Performance Simulation Platform. <i>DEStech Transactions on Computer Science and Engineering</i> , 2018, , .	0.1	0
10887	FUNCTIONAL STYLE PROGRAMMING IN JAVASCRIPT: TWO PROBABILITY PROBLEMS. <i>Informatics in School</i> , 2018, , 40-45.	0.1	0
10888	Performance of Map-Reduce Using Java-8 Parallel Streams. <i>Advances in Intelligent Systems and Computing</i> , 2019, , 723-736.	0.5	0

#	ARTICLE	IF	CITATIONS
10889	A Method to Reduce Processing Time by Parallelizing Generation of Voronoi Diagrams. , 2018, , .		0
10890	Post Event Investigation of Multi-stream Video Data Utilizing Hadoop Cluster. International Journal of Electrical and Computer Engineering, 2018, 8, 5089.	0.5	2
10891	Data Reliability-Aware and Cloud-Assisted Software Infrastructure for Body Area Networks. Internet of Things, 2019, , 303-318.	1.3	1
10892	An Analysis and Research of Network Log Based on Hadoop. Advances in Intelligent Systems and Computing, 2019, , 541-548.	0.5	0
10893	A Cooperative Approach for Distributed Intrusion Protection in PIS Systems. Lecture Notes on Data Engineering and Communications Technologies, 2019, , 888-894.	0.5	0
10894	Explore New Computing Environment for LHAASO Offline Data Analysis. , 2018, , .		0
10895	System Software for Data-Intensive Science. , 2019, , 99-120.		0
10896	Research Challenges in Big Data Solutions in Different Applications. SpringerBriefs in Applied Sciences and Technology, 2019, , 105-116.	0.2	1
10898	Building an Algorithmic Skeleton for Block Data Processing on Enterprise Desktop Grids. Communications in Computer and Information Science, 2019, , 678-689.	0.4	0
10900	Intelligent Management and Efficient Operation of Big Data. , 2019, , 1991-2016.		0
10902	Some Artificial Intelligence Driven Algorithms For Mobile Edge Computing in Smart City. Lecture Notes in Computer Science, 2019, , 110-119.	1.0	0
10904	Change Point Detection in Periodic Panel Data Using a Mixture-Model-Based Approach. Lecture Notes in Computer Science, 2019, , 139-150.	1.0	0
10905	Clustering Data in Secured, Distributed Datasets. Lecture Notes in Computer Science, 2019, , 557-572.	1.0	0
10906	Parallel Processing with Big Data. , 2019, , 1253-1259.		4
10907	Scalable Performance Modeling and Evaluation of MapReduce Applications. Communications in Computer and Information Science, 2019, , 441-458.	0.4	3
10908	A Hybrid Approach to Big Data Systems Development. , 2019, , 2271-2288.		0
10909	Data Science and Big Data Analytics in Financial Services. , 2019, , 1301-1329.		0
10910	Application of Big Data in Digital Epidemiology. Advances in Healthcare Information Systems and Administration Book Series, 2019, , 285-305.	0.2	0

#	ARTICLE	IF	CITATIONS
10911	A Comprehensive Study on Internet of Things Based on Key Artificial Intelligence Technologies and Industry 4.0. Advances in Computational Intelligence and Robotics Book Series, 2019, , 1-26.	0.4	1
10912	Voronoi-Based kNN Queries Using K-Means Clustering in MapReduce. Advances in Data Mining and Database Management Book Series, 2019, , 220-241.	0.4	0
10913	Using the bag-of-tasks model with centralized storage for distributed sorting of large data array. CEUR Workshop Proceedings, 2019, , 199-203.	2.3	0
10914	A Map-Reduce-Based Parallel Approach for Geospatial Data Interlinking in a Semantic Web. Geomatics Science and Technology, 2019, 07, 90-100.	0.2	0
10915	ADSM: Adaptive Data Scheduling Method for Hybrid Memories in Distributed System. IEEE Access, 2019, 7, 180218-180234.	2.6	0
10916	Fuzzy Models for Big Data Mining. Lecture Notes in Computer Science, 2019, , 257-260.	1.0	1
10917	Distributed Incremental View Maintenance. , 2019, , 682-690.		0
10918	A Two-Tiered Segmentation Approach for Transaction Data Warehousing. Advances in Data Mining and Database Management Book Series, 2019, , 1-27.	0.4	0
10919	Scalable Least Square Twin Support Vector Machine Learning. Lecture Notes in Computer Science, 2019, , 239-249.	1.0	1
10920	Big Data Processing and Big Analytics. Advances in Data Mining and Database Management Book Series, 2019, , 285-315.	0.4	0
10921	Evolution of Cloud in Big Data With Hadoop on Docker Platform. , 2019, , 1601-1622.		3
10922	Big Data Mining Based on Computational Intelligence and Fuzzy Clustering. , 2019, , 413-430.		0
10923	A Dynamic Evolutionary Scheduling Algorithm for Cloud Tasks with High Efficiency and High QoS Satisfactions. Communications in Computer and Information Science, 2019, , 131-142.	0.4	0
10924	CHAPTER 5. Making Big Data Available: Integrating Technologies for Toxicology Applications. Issues in Toxicology, 2019, , 166-184.	0.2	0
10925	Glowworm swarm optimization, Optimization, Swarm intelligence, Clustering. , 0, , .		0
10926	Deep Learning for Big Data Analytics. Advances in Computational Intelligence and Robotics Book Series, 2019, , 1-21.	0.4	0
10927	Baran. Advances in IT Standards and Standardization Research Series, 2019, , 124-161.	0.2	0
10928	H.265 Video Streaming in Mobile Cloud Networks. , 2019, , 1007-1047.		0

#	ARTICLE	IF	CITATIONS
10929	Shifting Perspectives. , 2019, , 554-574.		0
10930	An Enhanced Hybrid Feature Selection Approach for High Dimensional Data Processing. Journal of Testing and Evaluation, 2019, 47, 4161-4176.	0.4	0
10931	Anomaly Detection in Cloud Environments. , 2019, , 140-164.		0
10932	Big Data Indexing. , 2019, , 315-325.		1
10933	Research on Parallel Computing Method of Hydrological Data Cube. Lecture Notes in Computer Science, 2019, , 49-64.	1.0	0
10934	Survey on Various MapReduce Scheduling Algorithms. Advances in Computer and Electrical Engineering Book Series, 2019, , 499-515.	0.2	0
10935	A Framework for Classification of Data Stream Application in Vehicular Network Computing. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2019, , 57-67.	0.2	1
10936	Research and Improvement of Apriori Algorithm Based on Hadoop. International Journal of Advanced Network, Monitoring, and Controls, 2019, 3, 100-105.	0.2	1
10937	Architectures. , 2019, , 98-107.		0
10938	Caching for SQL-on-Hadoop. , 2019, , 431-435.		0
10939	Enabling Standard Geospatial Capabilities in Spark for the Efficient Processing of Geospatial Big Data. Communications in Computer and Information Science, 2019, , 133-148.	0.4	1
10940	Scalable Architectures for Big Data Analysis. , 2019, , 1446-1454.		1
10941	A New Spark Based K-Means Clustering with Data Removing Strategy. Lecture Notes in Business Information Processing, 2019, , 289-304.	0.8	0
10942	Analysis of Novel Hybrid Encryption Algorithm for Providing Security in Big Data. Communications in Computer and Information Science, 2019, , 158-169.	0.4	0
10943	Some Thoughts About Appealing Directions for the Future of Fuzzy Theory and Technologies Along the Path Traced by Lotfi Zadeh. Lecture Notes in Computer Science, 2019, , 240-243.	1.0	3
10944	Scalable Knowledge Graph Construction from Text Collections. , 2019, , .		10
10945	Columnar Storage Formats. , 2019, , 464-469.		0
10946	StreamMine3G: Elastic and Fault Tolerant Large Scale Stream Processing. , 2019, , 1643-1653.		0

#	ARTICLE	IF	CITATIONS
10947	In-Memory Analytics. Advances in Computer and Electrical Engineering Book Series, 2019, , 381-389.	0.2	0
10948	Data Citation Service for Wikipedia Articles. International Journal of Advanced Computer Science and Applications, 2019, 10, .	0.5	0
10950	BSP Programming Model. , 2019, , 377-381.		1
10951	Query Processing â€“ kNN. , 2019, , 1334-1338.		0
10952	Performance Evaluation of Big Data Analysis. , 2019, , 1265-1271.		0
10953	A Frequency Scaling Based Performance Indicator Framework for Big Data Systems. Lecture Notes in Computer Science, 2019, , 19-35.	1.0	0
10954	Automatically Configuring Parallelism for Hybrid Layouts. Communications in Computer and Information Science, 2019, , 120-125.	0.4	0
10955	Exploring the Core of Systems Engineering. SSRN Electronic Journal, 0, , .	0.4	1
10957	Performance Analysis and Optimization of Alluxio with OLAP Workloads over Virtual Infrastructure. Lecture Notes in Computer Science, 2019, , 319-330.	1.0	1
10958	Virtual Testbed: Ship Motion Simulation for Personal Workstations. Lecture Notes in Computer Science, 2019, , 717-728.	1.0	1
10959	Generation and Classification of Energy Load Curves Using a Distributed MapReduce Approach. Communications in Computer and Information Science, 2019, , 3-17.	0.4	1
10960	A novel entropy-based dynamic data placement strategy for data intensive applications in Hadoop clusters. International Journal of Big Data Intelligence, 2019, 6, 20.	0.4	0
10961	HPC Tools to Deal with Microarray Data. Methods in Molecular Biology, 2019, 1986, 227-243.	0.4	0
10962	The share Operator for Field-Based Coordination. Lecture Notes in Computer Science, 2019, , 54-71.	1.0	3
10963	Assessing the Dependability of Apache Spark System: Streaming Analytics on Large-Scale Ocean Data. Communications in Computer and Information Science, 2019, , 131-144.	0.4	0
10964	SparkDA: RDD-Based High-Performance Data Anonymization Technique for Spark Platform. Lecture Notes in Computer Science, 2019, , 646-662.	1.0	6
10965	Image Analysis at Scale for Finding the Links Between Structure and Biology. , 2019, , 425-443.		0
10966	Cost-Efficient Task Scheduling for Geo-distributed Data Analytics. Lecture Notes in Computer Science, 2019, , 95-104.	1.0	0

#	ARTICLE	IF	CITATIONS
10967	A Web Application for Characterizing Spontaneous Emotions Using Long EEG Recording Sessions. Intelligent Systems Reference Library, 2019, , 185-202.	1.0	1
10968	Big Data and the Internet of Things in Edge Computing for Smart City. Lecture Notes in Computer Science, 2019, , 99-109.	1.0	1
10969	Online Machine Learning in Big Data Streams: Overview. , 2019, , 1207-1218.		4
10970	Workflow Systems for Big Data Analysis. , 2019, , 1811-1816.		0
10971	Big Data Networks. , 2019, , 1-5.		0
10972	Hadoop. , 2019, , 913-917.		0
10973	Document Based RDF Storage Method for Efficient Parallel Query Processing. Communications in Computer and Information Science, 2019, , 13-25.	0.4	0
10974	SwitchAgg. , 2019, , .		12
10975	Dagstuhl Seminar on Big Stream Processing. SIGMOD Record, 2019, 47, 36-39.	0.7	0
10976	Efficient Dynamic Bloom Filter Method in Big Data Environment. The Journal of Korean Institute of Information Technology, 2019, 17, 13-20.	0.1	0
10977	Interoperable Convergence of Storage, Networking, and Computation. Lecture Notes in Networks and Systems, 2020, , 667-690.	0.5	1
10978	LDM: Lineage-Aware Data Management in Multi-tier Storage Systems. Lecture Notes in Networks and Systems, 2020, , 683-707.	0.5	1
10980	Parallelizing user-defined functions in the ETL workflow using orchestration style sheets. International Journal of Applied Mathematics and Computer Science, 2019, 29, 69-79.	1.5	2
10983	Clustering-Based Data Analysis With Hadoop. International Journal of Computer Trends and Technology, 2019, 67, 78-81.	0.1	0
10985	Elastic Multi-resource Network Slicing: Can Protection Lead to Improved Performance?. , 2019, , .		2
10986	Advantages of Giraph over Hadoop in Graph Processing. Engineering, Technology & Applied Science Research, 2019, 9, 4112-4115.	0.8	0
10988	Parallel and Distributed Pattern Mining. International Journal of Rough Sets and Data Analysis, 2019, 6, 1-17.	1.0	2
10989	Data Analysis and Visualization of Terrorist Attacks in the Philippines. International Journal of Simulation: Systems, Science and Technology, 0, , .	0.0	0

#	ARTICLE	IF	CITATIONS
10997	Customizable and scalable fuzzy join for big data. Proceedings of the VLDB Endowment, 2019, 12, 2106-2117.	2.1	4
10998	A Scalable and Reliable Model for Real-time Air Quality Prediction. , 2019, , .		2
10999	Tuplex. Proceedings of the VLDB Endowment, 2019, 12, 1958-1961.	2.1	0
11000	An Efficient Distributed Frequent Itemset Mining Algorithm Based on Spark for Big Data. International Journal of Intelligent Engineering and Systems, 2019, 12, 367-377.	0.8	2
11001	Automatic feature engineering from very high dimensional event logs using deep neural networks. , 2019, , .		1
11002	A Distributed Spatial Index on Smart Medical System. Computer Communications and Networks, 2020, , 287-303.	0.8	0
11003	MR-AntMiner: A Novel MapReduce Classification Rule Discovery with Ant Colony Intelligence. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2019, 23, 928-938.	0.5	0
11004	Towards Big Data Solutions for Industrial Tomography Data Processing. , 0, , .		1
11006	A Comparative Study on Different Approaches of Road Traffic Optimization Based on Big Data Analytics. Asset Analytics, 2020, , 119-126.	0.4	1
11007	An Empirical Study of Big Data: Opportunities, Challenges and Technologies. Advances in Intelligent Systems and Computing, 2020, , 49-65.	0.5	0
11008	Automated management of maritime container terminals using internet of things and big data technologies. , 2019, , .		0
11010	Comparative Analysis of Map Reduce Scheduling Algorithms. Journal of Advanced Research in Dynamical and Control Systems, 2019, 11, 20-31.	0.3	2
11011	A Conceptual Framework for Developing a Customized I 4.0 Education Scale: An Exploratory Research. Lecture Notes in Mechanical Engineering, 2020, , 203-216.	0.3	3
11012	Big Data Security Challenges and Preventive Solutions. Advances in Intelligent Systems and Computing, 2020, , 285-299.	0.5	1
11013	Hybrid Partitioning-Density Algorithm for K-Means Clustering of Distributed Data Utilizing OPTICS. International Journal of Data Warehousing and Mining, 2019, 15, 1-20.	0.4	2
11014	An Efficient Parallel Triangle Enumeration on the MapReduce Framework. IEICE Transactions on Information and Systems, 2019, E102.D, 1902-1915.	0.4	0
11015	Models of Data Storage Tables Connection Processes by MapReduce/Spark Technology. Herald of the Bauman Moscow State Technical University Series Instrument Engineering, 2019, , 79-94.	0.2	0
11016	Ultra-Scalable Simulations of Networks of Polarized Evolutionary Processors. , 2019, , .		0

#	ARTICLE	IF	CITATIONS
11017	An Energy Efficient and QoS Achieved Through MapReduce in Cloud Environment. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 781-792.	0.5	0
11018	A Sub-Linear Scalable MapReduce-based Apriori Algorithm. , 2019, , .		2
11019	A Sampling-Based System for Approximate Big Data Analysis on Computing Clusters. , 2019, , .		3
11020	MapReduce Hadoop Models for Distributed Neural Network Processing of Big Data Using Cloud Services. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 387-400.	0.5	0
11021	Soft Frequency Capping for Improved Ad Click Prediction in Yahoo Gemini Native. , 2019, , .		13
11022	Map-Reduce Based Generic Basis of Association Rules Mining from Big Bata. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 647-658.	0.5	0
11023	Workshop Summary. Software Engineering Notes: an Informal Newsletter of the Special Interest Committee on Software Engineering / ACM, 2019, 44, 48-52.	0.5	0
11024	Synchronization of Machine Learning into Electronic Health Records. <i>International Journal of Computer Applications</i> , 2019, 177, 40-47.	0.2	0
11025	Differential Approximation and Sprinting for Multi-Priority Big Data Engines. , 2019, , .		2
11026	SecDATAVIEW. , 2019, , .		5
11027	Prediction of Life Cycle based on EMR data of Companion Animals. <i>Journal of Digital Contents Society</i> , 2019, 20, 2505-2514.	0.1	0
11028	An Adequate Computation Offloading in Mobile Cloud Computing. <i>International Journal of Scientific Research in Computer Science Engineering and Information Technology</i> , 2019, , 208-216.	0.2	0
11029	The Library for Hadoop Deflate Compression Based on FPGA Accelerator with Load Balance. , 2019, , .		0
11030	Efficient Processing of Recursive Joins on Large-Scale Datasets in Spark. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 391-402.	0.5	0
11031	Distributed Graph Analytics. <i>Lecture Notes in Computer Science</i> , 2020, , 3-20.	1.0	0
11032	Twitter Sentimental Analytics Using Hive and Flume. <i>Algorithms for Intelligent Systems</i> , 2020, , 159-165.	0.5	0
11033	Healthcare Data Storage Based on HBase. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 199-205.	0.5	1
11034	The Birth of Digital – A Brief History of Digital Technologies. <i>Future of Business and Finance</i> , 2020, , 3-7.	0.3	1

#	ARTICLE	IF	CITATIONS
11035	Spatiotemporal Domain Decomposition for High Performance Computing: A Flexible Splits Heuristic to Minimize Redundancy. <i>Geotechnologies and the Environment</i> , 2020, , 27-50.	0.3	2
11037	Karst: Transactional Data Ingestion Without Blocking on a Scalable Architecture. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2022, 34, 2241-2253.	4.0	1
11038	Cloud Data Center Network Multi-User Multi-Service Concurrency Algorithm. <i>Computer Science and Application</i> , 2020, 10, 1422-1430.	0.0	0
11039	â€œIncrementalâ€•Evaluation for Genetic Crossover. <i>Lecture Notes in Computer Science</i> , 2020, , 396-404.	1.0	1
11040	Cyber-Physical Systems in Vehicular Communications. , 2020, , 411-431.		0
11041	JSON Schema Inference Approaches. <i>Lecture Notes in Computer Science</i> , 2020, , 173-183.	1.0	2
11042	Daten. , 2020, , 75-117.		0
11043	A Decentralized Data Processing Framework Based on PoUW Blockchain. <i>Communications in Computer and Information Science</i> , 2020, , 588-600.	0.4	0
11044	MrPC: Causal Structure Learning in Distributed Systems. <i>Communications in Computer and Information Science</i> , 2020, , 87-94.	0.4	0
11045	ANaN â€” ANalyse And Navigate: Debugging Compute Clusters with Techniques from Functional Programming and Text Stream Processing. <i>EPJ Web of Conferences</i> , 2020, 245, 01041.	0.1	0
11046	INSTalytics. <i>ACM Transactions on Storage</i> , 2019, 15, 1-30.	1.4	5
11047	Reconfigurable optical star network architecture for multicast media production data centres. <i>Optical Switching and Networking</i> , 2020, 36, 100556.	1.2	2
11049	Timon: A Timestamped Event Database for Efficient Telemetry Data Processing and Analytics. , 2020, , .		12
11051	Under the Concealing Surface: Detecting and Understanding Live Webcams in the Wild. <i>Proceedings of the ACM on Measurement and Analysis of Computing Systems</i> , 2020, 4, 1-25.	1.4	8
11052	A Framework for Total Quality Management of Diesel Generator Fuel Consumption using Machine Learning and Internet of Things (IoT). <i>International Journal of Computer Applications</i> , 2020, 176, 43-52.	0.2	0
11053	Towards Scalable UDTFs in Noria. , 2020, , .		0
11054	Near-Optimal Distributed Band-Joins through Recursive Partitioning. , 2020, 2020, 2375-2390.		4
11055	A Survey on Smart Analytics: Method, Tools, and Open Research Issues. <i>Jurnal Sistem Cerdas</i> , 2020, 3, 54-64.	0.1	0

#	ARTICLE	IF	CITATIONS
11056	Spur: Mitigating Slow Instances in Large-Scale Streaming Pipelines. , 2020, , .		4
11057	VizSciFlow: A Visually Guided Scripting Framework for Supporting Complex Scientific Data Analysis. Proceedings of the ACM on Human-Computer Interaction, 2020, 4, 1-37.	2.5	6
11058	Big Data and digital economy. Sistemnyj Analiz V Nauke I Obrazovanii, 2020, , 187-197.	0.0	0
11059	Compilers used in the Stratosphere Platform for Big Data Analytics. International Journal of Engineering Research & Technology, 2020, V9, .	0.2	0
11060	Substream-Centric Maximum Matchings on FPGA. ACM Transactions on Reconfigurable Technology and Systems, 2020, 13, 1-33.	1.9	5
11061	Enhancing Approach for Information Security in Hadoop. Polytechnic Journal, 2020, 10, 81-87.	0.1	1
11062	Fundamental Limits of Approximate Gradient Coding. , 2020, , .		7
11063	Thriving in a crowded and changing world: C++ 2006â€“2020. , 2020, 4, 1-168.		7
11064	Distributed Facial Feature Clustering Algorithm Based on Spatiotemporal Locality. Advances in Intelligent Systems and Computing, 2021, , 394-403.	0.5	1
11065	Boa Views. , 2020, , .		2
11066	Unconditional Lower Bounds for Adaptive Massively Parallel Computation. , 2020, , .		4
11067	OPS: Optimized Shuffle Management System for Apache Spark. , 2020, , .		4
11068	AGL. Proceedings of the VLDB Endowment, 2020, 13, 3125-3137.	2.1	37
11070	Magnet. Proceedings of the VLDB Endowment, 2020, 13, 3382-3395.	2.1	11
11072	Application of ELM-MapReduce Technique in Stock Market Forecasting. Smart Innovation, Systems and Technologies, 2021, , 469-476.	0.5	0
11073	Towards Elastic Data Warehousing by Decoupling Data Management and Computation. , 2020, , .		0
11074	DEADLINE AWARE OPTIMIZATION IN RESOURCE ALLOCATION FOR REDUCING MIGRATION COST. Advances in Mathematics: Scientific Journal (discontinued), 2020, 9, 6765-6775.	0.2	0
11075	An Approach to Implement MapReduce and Aggregation Pipeline Utilizing NoSql Technologies. Algorithms for Intelligent Systems, 2021, , 27-34.	0.5	0

#	ARTICLE	IF	CITATIONS
11076	Volt/Var Optimization for Active Power Distribution Systems on a Graph Computing Platform: An Paralleled PSO Approach. , 2020, , .		3
11077	TOWARDS INTELLIGENT GEO-DATABASE SUPPORT FOR EARTH SYSTEM OBSERVATION: IMPROVING THE PREPARATION AND ANALYSIS OF BIG SPATIO-TEMPORAL RASTER DATA. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 0, XLIII-B4-2020, 485-492.	0.2	1
11078	Analysis of Diabetes and Heart Disease in Big Data Using MapReduce Framework. Advances in Intelligent Systems and Computing, 2021, , 37-51.	0.5	3
11079	A Rack-Aware Pipeline Repair Scheme for Erasure-Coded Distributed Storage Systems. , 2020, , .		3
11080	Reducing Latency in Multi-Tenant Data Centers via Cautious Congestion Watch. , 2020, , .		5
11081	RESEARCH ON THE APPLICATION OF ARTIFICIAL INTELLIGENCE IN FUND MANAGER IDENTIFICATION. Singapore Economic Review, 0, , 1-35.	0.9	0
11082	Evaluation of an InfiniBand Switch: Choose Latency or Bandwidth, but Not Both. , 2020, , .		6
11083	An Effective Model for Handling the Big Data Streams Based on the Optimization-Enabled Spark Framework. Advances in Intelligent Systems and Computing, 2021, , 673-696.	0.5	4
11084	IMPROVING THE DATA TRANSMISSION SPEED IN CLOUD MIGRATION BY USING MAPREDUCE FOR BIGDATA. International Journal of Engineering Technology and Management Sciences, 2020, 4, 73-75.	0.0	0
11085	Establishment and Application of Big Data Processing Platform. , 2020, , .		0
11086	Real-Time Semantic Web Data Stream Processing Using Storm. , 2020, , .		0
11087	P-Cube: A New Two-Layer Topology for Data Center Networks Exploiting Dual-Port Servers. IEICE Transactions on Communications, 2020, E103.B, 940-950.	0.4	4
11088	A Survey of Multitier Programming. ACM Computing Surveys, 2021, 53, 1-35.	16.1	11
11089	Research on the Application of Big Data in Smart Food Safety. , 2020, , .		0
11091	Weapon-target assignment in unreliable peer-to-peer architecture based on adapted artificial bee colony algorithm. Frontiers of Computer Science, 2022, 16, 1.	1.6	4
11092	Introduction to Digital Pathology from Historical Perspectives to Emerging Pathomics. , 2022, , 1-22.		2
11093	OptDebug. , 2021, , .		1
11094	Clamor. , 2021, , .		0

#	ARTICLE	IF	CITATIONS
11096	Review of Parallel Processing Methods for Big Image Data Applications. Lecture Notes in Electrical Engineering, 2020, , 105-116.	0.3	1
11097	HSPP: Load-Balanced and Low-Latency File Partition and Placement Strategy on Distributed Heterogeneous Storage with Erasure Coding. Lecture Notes in Computer Science, 2020, , 200-214.	1.0	0
11098	High-Utility Itemset Mining in Big Dataset. Advances in Intelligent Systems and Computing, 2020, , 567-570.	0.5	0
11099	Distributed Ledger for Provenance Tracking of Artificial Intelligence Assets. IFIP Advances in Information and Communication Technology, 2020, , 411-426.	0.5	6
11100	Prediction Model of Traffic Flow Driven Based on Single Data in Smart Traffic Systems. , 2020, , 159-194.		2
11101	Sorting Process In Mapreduce Task. Impact Journal of Science and Technology, 2020, 14, .	0.0	0
11102	\$\$ LogRank^+ \$\$: A Novel Approach to Support Business Process Event Log Sampling. Lecture Notes in Computer Science, 2020, , 417-430.	1.0	0
11103	An Alternative to Power Measure for Fuzzy Rule-Based Classification Systems. Lecture Notes in Computer Science, 2020, , 231-244.	1.0	0
11104	Privacy-Preserving and Outsourced Density Peaks Clustering Algorithm. Communications in Computer and Information Science, 2020, , 538-552.	0.4	1
11105	Distributed Storage of Large Knowledge Graphs with Mobility Data. , 2020, , 181-211.		0
11106	Big Healthcare Data: Survey of Challenges and Privacy. International Journal on Informatics Visualization, 2020, 4, 184-190.	0.5	2
11107	SrSpark: Skew-resilient Spark based on Adaptive Parallel Processing. , 2020, , .		1
11108	Service Provisioning through High Level, Complexity Hiding Interfaces. , 2020, , .		0
11109	Connecting MapReduce Computations to Realistic Machine Models. , 2020, , .		1
11110	A Push/Pull-Based System Scheduling Framework for Large-Scale Cluster. , 2020, , .		0
11111	Real-Time Social Media Analytics with Deep Transformer Language Models: A Big Data Approach. , 2020, , .		4
11112	Intermediate Value Size Aware Coded MapReduce. , 2020, , .		2
11113	HeAT â€“ a Distributed and GPU-accelerated Tensor Framework for Data Analytics. , 2020, , .		6

#	ARTICLE	IF	CITATIONS
11114	DCache: A Distributed Cache Mechanism for HDFS based on RDMA. , 2020, , .		4
11115	Smart Cities Pilot Projects: An IoT Perspective. Lecture Notes in Intelligent Transportation and Infrastructure, 2021, , 231-255.	0.3	1
11116	SplitServe. , 2020, , .		6
11117	FeGAN. , 2020, , .		10
11118	On the Marriage of Asynchronous Many Task Runtimes and Big Data: A Glance. , 2020, , .		1
11119	Impact of Map-Reduce framework on Hadoop and Spark MR Application Performance. , 2020, , .		2
11120	Exposing data locality in HPC-based systems by using the HDFS backend. , 2020, , .		0
11121	Large-Scale Intelligent Microservices. , 2020, , .		5
11122	Performance Analysis of Multi-Node Hadoop Cluster Based on Large Data Sets. , 2020, , .		1
11123	Topological Coded Distributed Computing. , 2020, , .		6
11124	Memory Abstraction and Optimization for Distributed Executors. , 2020, , .		0
11125	Distributed Principal Component Analysis for Real-time Big Data Processing. , 2020, , .		0
11126	Distributing Active Learning Algorithms. , 2020, , .		0
11127	A Survey: Benchmarking and Performance Modelling of Data Intensive Applications. , 2020, , .		1
11128	OverSketched Newton: Fast Convex Optimization for Serverless Systems. , 2020, , .		6
11129	GuardSpark++: Fine-Grained Purpose-Aware Access Control for Secure Data Sharing and Analysis in Spark. , 2020, , .		4
11130	Fuzzy High-Utility Pattern Mining based on the Hadoop Framework. , 2020, , .		2
11131	Moving Compute towards Data in Heterogeneous multi-FPGA Clusters using Partial Reconfiguration and I/O Virtualisation. , 2020, , .		2

#	ARTICLE	IF	CITATIONS
11132	Time and Cost Prediction Models for Language Classification Over a Large Corpus on Spark. , 2020, , .		0
11133	Analysis of Big Data Analytics in Healthcare Sector: Applications and Tools. Journal of Computational and Theoretical Nanoscience, 2020, 17, 5605-5612.	0.4	0
11134	Parallel Approaches of Utility Mining for Big Data. Webology, 2020, 17, 31-43.	0.3	1
11135	ReSpark: Automatic Caching for Iterative Applications in Apache Spark. , 2020, , .		1
11136	Translation of Array-Based Loops to Spark SQL. , 2020, , .		2
11137	A Fast, Scalable, Universal Approach For Distributed Data Aggregations. , 2020, , .		3
11138	Memory and Performance Aware Scheduling Design for Hadoop MapReduce Framework. Bioscience Biotechnology Research Communications, 2020, 13, 242-246.	0.1	1
11139	Uncertain Spatial Data Management: An Overview. , 2021, , 355-397.		4
11140	Performance Optimization System for Hadoop and Spark Frameworks. Cybernetics and Information Technologies, 2020, 20, 5-17.	0.4	6
11141	Improved MapReduce Load Balancing through Distribution-Dependent Hash Function Optimization. , 2020, , .		1
11142	Crowdsourcing Answer Integration Algorithm For Big Data Environment. , 2020, , .		0
11143	Locality-Preserving Spatial Partitioning for Geo Big Data Analytics in Main Memory Frameworks. , 2020, , .		4
11144	Distributed Computing with Heterogeneous Servers. , 2020, , .		0
11145	Parallelizing Flow-Sensitive Demand-Driven Points-to Analysis. , 2020, , .		0
11147	Framework-Agnostic Optimization of Repeated Skewed Joins at Massive Scale. , 2020, , .		1
11148	Exploring Dynamic Task Loading in SGX-based Distributed Computing. IEEE Transactions on Services Computing, 2021, , 1-1.	3.2	3
11149	Improving Project Management Decisions With Big Data Analytics. , 2022, , 1765-1785.		0
11150	Role of Big Data in Internet of Things Networks. , 2022, , 336-363.		2

#	ARTICLE	IF	CITATIONS
11151	Application of Big Data in Digital Epidemiology. , 2022, , 958-978.		0
11152	Online and Distributed Robust Regressions with Extremely Noisy Labels. ACM Transactions on Knowledge Discovery From Data, 2022, 16, 1-24.	2.5	1
11153	Computing Graph Neural Networks: A Survey from Algorithms to Accelerators. ACM Computing Surveys, 2022, 54, 1-38.	16.1	87
11154	Datenbanken. , 2020, , 729-758.		0
11155	Federated Soft Gradient Boosting Machine for Streaming Data. Lecture Notes in Computer Science, 2020, , 93-107.	1.0	2
11157	Big Data Networks. , 2020, , 106-111.		0
11159	An Exploration on Cloud Computing Security Strategies and Issues. Lecture Notes in Networks and Systems, 2020, , 549-562.	0.5	4
11160	Trends and Technologies in Big Data Processing. Advances in Computational Intelligence and Robotics Book Series, 2020, , 17-42.	0.4	1
11161	An Insight into State-of-the-Art Techniques for Big Data Classification. , 2020, , 1742-1763.		0
11162	Parallel K-Prototypes Clustering with High Efficiency and Accuracy. Lecture Notes in Computer Science, 2020, , 380-395.	1.0	1
11163	A Taxonomy and Survey of Data Partitioning Algorithms for Big Data Distributed Systems. Lecture Notes in Networks and Systems, 2020, , 447-457.	0.5	2
11164	Conceptual Modeling of the Social Environment for Information Support in Management Processes. Communications in Computer and Information Science, 2020, , 138-151.	0.4	0
11165	AGINFRA PLUS: Running Crop Simulations on the D4Science Distributed e-Infrastructure. IFIP Advances in Information and Communication Technology, 2020, , 81-89.	0.5	3
11166	Big Data-Driven Detection of False Data Injection Attacks in Smart Meters. IEEE Access, 2021, 9, 144313-144326.	2.6	12
11167	CrashTuner. , 2019, , .		17
11168	A Comparative Approach to Secure Data Storage Model in Hadoop Framework. Advances in Intelligent Systems and Computing, 2020, , 135-143.	0.5	0
11169	A Scalable Synthesis of Multiple Models of Geo Big Data Interpretation. Journal of Software Engineering and Applications, 2020, 13, 104-128.	0.8	1
11170	Trade-Offs in Large-Scale Distributed Tuplewise Estimation And Learning. Lecture Notes in Computer Science, 2020, , 229-245.	1.0	0

#	ARTICLE	IF	CITATIONS
11172	Information Management: The Role of Information Technology in Treaty Verification. , 2020, , 285-297.		0
11173	Strark-H: A Strategy for Spatial Data Storage to Improve Query Efficiency Based on Spark. Lecture Notes in Computer Science, 2020, , 285-299.	1.0	0
11174	CLUST. , 2020, , .		2
11175	SparkNN: A Distributed In-Memory Data Partitioning for KNN Queries on Big Spatial Data. Data Science Journal, 2020, 19, .	0.6	3
11176	Batch Arrival Multiserver Queue with State-Dependent Setup for Energy-Saving Data Center. Infosys Science Foundation Series, 2020, , 421-440.	0.3	1
11177	Balancing Between Scalability and Accuracy in Time-Series Classification for Stream and Batch Settings. Lecture Notes in Computer Science, 2020, , 265-279.	1.0	3
11178	Extended Kalman Filter for Large Scale Vessels Trajectory Tracking in Distributed Stream Processing Systems. Lecture Notes in Computer Science, 2020, , 151-166.	1.0	1
11179	Scalable Sketch-Based Sport Video Retrieval in the Cloud. Lecture Notes in Computer Science, 2020, , 226-241.	1.0	1
11180	K-NN Classification for Large Dataset in Hadoop Framework. Lecture Notes in Electrical Engineering, 2020, , 495-501.	0.3	0
11181	Novel Scalable Deep Learning Approaches for Big Data Analytics Applied to ECG Processing. , 2020, , 633-653.		0
11182	An Intelligent Approval System for City Construction based on Cloud Computing and Big Data. , 2020, , 248-263.		0
11183	Efficient Imbalanced Multimedia Concept Retrieval by Deep Learning on Spark Clusters. , 2020, , 274-294.		0
11184	A Novel Approach With Dynamic Fuzzy Inference for Big Data Classification Problems. Advances in Computational Intelligence and Robotics Book Series, 2020, , 43-58.	0.4	0
11185	PhoeniQ: Failure-Tolerant Query Processing in Multi-node Environments. Lecture Notes in Computer Science, 2020, , 71-85.	1.0	1
11186	Spatial Association Pattern Mining Using In-Memory Computational Framework. Lecture Notes in Computer Science, 2020, , 239-246.	1.0	0
11187	Hybrid Warehouse Model and Solutions for Climate Data Analysis. Journal of Computer and Communications, 2020, 08, 75-98.	0.6	3
11188	Dynamic Estimation and Grid Partitioning Approach for Multi-objective Optimization Problems in Medical Cloud Federations. Lecture Notes in Computer Science, 2020, , 32-66.	1.0	0
11189	Near-Data Prediction Based Speculative Optimization in a Distribution Environment. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2020, , 111-122.	0.2	0

#	ARTICLE	IF	CITATIONS
11191	Hadoop Performance Acceleration by Effective Data and Job Placement. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 217-226.	0.5	0
11192	A Smart Data Approach for Automatic Data Analysis. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 691-699.	0.5	1
11193	High Order N-gram Model Construction and Application Based on Natural Annotation. <i>Lecture Notes in Computer Science</i> , 2020, , 321-328.	1.0	0
11194	Security Issues in Big Data. <i>Lecture Notes in Networks and Systems</i> , 2020, , 67-72.	0.5	1
11195	The Experimental Study of Performance Impairment of Big Data Processing in Dynamic and Opportunistic Environments. <i>Journal of Communications</i> , 2020, , 776-789.	1.3	2
11196	Open Source Software (OSS) for Big Data. <i>Advances in Computer and Electrical Engineering Book Series</i> , 2020, , 50-72.	0.2	1
11197	Decoding Big Data Analytics for Emerging Business Through Data-Intensive Applications and Business Intelligence. <i>Advances in Data Mining and Database Management Book Series</i> , 2020, , 66-80.	0.4	0
11198	Big Spectrum Data and Deep Learning Techniques for Cognitive Wireless Networks. , 2020, , 994-1015.		0
11199	Incremental Evaluation of Continuous Analytic Queries in HIFUN. <i>Communications in Computer and Information Science</i> , 2020, , 53-67.	0.4	0
11200	Enabling Hardware Affinity in JVM-Based Applications: A Case Study for Big Data. <i>Lecture Notes in Computer Science</i> , 2020, , 31-44.	1.0	1
11201	VEDZA: Kinect Based Virtual Shopping Assistant. <i>Lecture Notes in Networks and Systems</i> , 2020, , 465-474.	0.5	0
11202	Improved K-Means Algorithm for Optimizing Initial Centers. <i>Smart Innovation, Systems and Technologies</i> , 2020, , 213-219.	0.5	0
11203	High-performance computing for astronomical big data. , 2020, , 305-323.		0
11204	Cartographic Mapping Driven by High-Performance Computing: A Review. <i>Geotechnologies and the Environment</i> , 2020, , 159-172.	0.3	1
11205	Domain Application of High Performance Computing in Earth Science: An Example of Dust Storm Modeling and Visualization. <i>Geotechnologies and the Environment</i> , 2020, , 249-268.	0.3	0
11206	VEDL: A Novel Video Event Searching Technique Using Deep Learning. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 905-914.	0.5	8
11207	Action rules for sentiment analysis using Twitter. <i>International Journal of Social Network Mining</i> , 2020, 3, 35.	0.2	0
11208	Generalizing Streaming Pipeline Design for Big Data. <i>Advances in Intelligent Systems and Computing</i> , 2020, , 149-160.	0.5	2

#	ARTICLE	IF	CITATIONS
11209	A Learning-Based Approach for Evaluating the Capacity of Data Processing Pipelines. Lecture Notes in Computer Science, 2020, , 52-67.	1.0	0
11210	Hardware and software data processing system for research and scientific purposes based on Raspberry Pi 3 microcomputer. Proceedings of the Institute for System Programming of RAS, 2020, 32, 57-69.	0.1	1
11211	Efficient Distributed Clustering Algorithms on Star-Schema Heterogeneous Graphs. IEEE Transactions on Knowledge and Data Engineering, 2022, 34, 4781-4796.	4.0	3
11212	â€œSaksham Modelâ€•Performance Improvisation Using Node Capability Evaluation in Apache Hadoop. Advances in Data Mining and Database Management Book Series, 2020, , 206-230.	0.4	1
11213	Big Data Processing. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2020, , 111-132.	0.5	0
11214	Customer Satisfaction through Technological Integration. , 2020, , 1824-1858.		0
11215	A State of the Art on Big Data With Semantic Web Technologies. Advances in Computer and Electrical Engineering Book Series, 2020, , 195-216.	0.2	0
11216	A BSP Based Approach for NFAs Intersection. Lecture Notes in Computer Science, 2020, , 344-354.	1.0	1
11218	A Multi-layer System for Maritime Container Terminal Management Using Internet of Things and Big Data Technologies. Lecture Notes in Intelligent Transportation and Infrastructure, 2020, , 865-877.	0.3	0
11219	The Optimization of Resources Within the Implementation of a Big Data Solution. Advances in Intelligent Systems and Computing, 2020, , 439-446.	0.5	1
11221	Spark and dask performance analysis based on ARL image library. , 2020, , 325-345.		4
11222	A Lightweight Neural Network Combining Dilated Convolution and Depthwise Separable Convolution. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2020, , 210-220.	0.2	1
11224	A Parallel Compressed Data Cube Based on Hadoop. Advances in Intelligent Systems and Computing, 2020, , 838-845.	0.5	0
11225	Scalable Aggregation Service for Satellite Remote Sensing Data. Lecture Notes in Computer Science, 2020, , 184-199.	1.0	1
11226	Indexing. Computers in Health Care, 2020, , 181-223.	0.2	1
11227	DDSL: Efficient Subgraph Listing on Distributed and Dynamic Graphs. Lecture Notes in Computer Science, 2020, , 632-640.	1.0	0
11228	Resilient Distributed Collection Through Information Speed Thresholds. Lecture Notes in Computer Science, 2020, , 211-229.	1.0	2
11229	Data Localization Laws and Policy: The EU Data Protection International Transfers Restriction Through a Cloud Computing Lens. Script Ed, 2020, 17, 152-159.	0.4	0

#	ARTICLE	IF	CITATIONS
11230	An Efficient Approach for Improving Recursive Joins Based on Three-Way Joins in Spark. Communications in Computer and Information Science, 2020, , 566-578.	0.4	0
11231	An Open Source Project for Tuning and Analyzing MapReduce Performance in Hadoop and Spark. IEEE Software, 2022, 39, 61-69.	2.1	4
11232	Implementing a Language for Distributed Systems: Choices and Experiences with Type Level and Macro Programming in Scala. The Art Science and Engineering of Programming, 2020, 4, .	0.4	2
11234	Scalable Generalized Multitarget Linear Regression With Output Dependence Estimation. Lecture Notes in Computer Science, 2021, , 60-68.	1.0	0
11235	Fangorn. Proceedings of the VLDB Endowment, 2021, 14, 2972-2985.	2.1	0
11236	Napa. Proceedings of the VLDB Endowment, 2021, 14, 2986-2997.	2.1	8
11237	Hybrid Scalable Action Rule. , 2020, , .		0
11238	A FRAMEWORK FOR INCREMENTAL PARALLEL MINING OF INTERESTING ASSOCIATION PATTERNS FOR BIG DATA. International Journal of Computing, 0, , 106-117.	1.5	6
11239	On the Performance, Availability and Traffic Control Analysis of Virtualized Servers in Cloud Computing Environment. International Journal of Simulation: Systems, Science and Technology, 0, , .	0.0	0
11240	Resource-Aware MapReduce Runtime for Multi/Many-core Architectures. , 2020, , .		0
11241	Towards Performance Modeling of Speculative Execution for Cloud Applications. , 2020, , .		2
11242	Solving Billion-Scale Knapsack Problems. , 2020, , .		3
11243	HPTMT: Operator-Based Architecture for Scalable High-Performance Data-Intensive Frameworks. , 2021, , .		4
11244	A Cost-Efficient Resource Provisioning and Scheduling Approach for Deadline-Sensitive MapReduce Computations in Cloud Environment. , 2021, , .		2
11245	Large-Scale News Classification using BERT Language Model: Spark NLP Approach. , 2021, , .		20
11246	Intelligent Health Information Services Requirements Revisited from an Actor's Perspective. , 2021, , .		0
11247	SODA: A Semantics-Aware Optimization Framework for Data-Intensive Applications Using Hybrid Program Analysis. , 2021, , .		2
11248	DCQCN Advanced (DCQCN-A) : Combining ECN and RTT for RDMA Congestion Control. , 2021, , .		3

#	ARTICLE	IF	CITATIONS
11249	Optimizing Urban LiDAR Flight Path Planning Using a Genetic Algorithm and a Dual Parallel Computing Framework. Remote Sensing, 2021, 13, 4437.	1.8	5
11250	A platform for integrating heterogeneous data and developing smart city applications. Future Generation Computer Systems, 2022, 128, 552-566.	4.9	21
11251	Incremental mining of high utility sequential patterns using MapReduce paradigm. Cluster Computing, 2022, 25, 805-825.	3.5	5
11252	Efficient distributed path computation on RDF knowledge graphs using partial evaluation. World Wide Web, 0, , 1.	2.7	1
11253	An Enhanced Parallelisation Model for Performance Prediction of Apache Spark on a Multinode Hadoop Cluster. Big Data and Cognitive Computing, 2021, 5, 65.	2.9	3
11254	Image Data Handling in a Multinode Environment Using MPI. Lecture Notes in Networks and Systems, 2022, , 21-34.	0.5	0
11255	Scalable Time Series Classification in streaming and batch environments on Apache Spark. , 2020, , .		0
11256	Massively Parallel Algorithms for Minimum Cut. , 2020, , .		4
11257	A Scalable Data Processing Model for Big Data Analysis of Enterprise Technology Innovation. , 2020, , .		0
11258	Fundamental Limits of Approximate Gradient Coding. Performance Evaluation Review, 2020, 48, 21-22.	0.4	1
11259	Adaptive Asynchronous Parallelization of Graph Algorithms. ACM Transactions on Database Systems, 2020, 45, 1-45.	1.5	6
11261	Data Mining Methods for Optimizing Feature Extraction and Model Selection. , 2020, , .		1
11262	Health Data Analytics with an Opportunistic Big Data Algorithm. , 2020, , .		1
11263	MapReduce and Hadoop. , 0, , 197-215.		1
11264	Grid, SOA and Cloud Computing. , 0, , 12-51.		0
11265	Goal-Based Requirements Elicitation for Service Reuse in Cloud Computing. , 0, , 278-305.		0
11266	Modeling Virtual Footprints. , 0, , 96-113.		0
11267	A Domain-Specific Language for High-Level Parallelization. , 0, , 533-552.		1

#	ARTICLE	IF	CITATIONS
11268	Migrating a Legacy Web-Based Document-Analysis Application to Hadoop and HBase. , 0 , 226-247.		2
11269	Performance Analysis of a Distributed Execution Environment for JUnit Test Cases on a Small Cluster. Advances in Computer and Electrical Engineering Book Series, 0 , 96-112.	0.2	1
11270	A Forensic-as-a-Service Delivery Platform for Law Enforcement Agencies. , 0 , 313-330.		1
11271	Data Intensive Computing for Bioinformatics. , 0 , 287-321.		4
11272	Towards Energy Efficiency for Cloud Computing Services. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 0 , 306-328.	0.5	3
11273	Applying the K-Means Algorithm in Big Raw Data Sets with Hadoop and MapReduce. Advances in Data Mining and Database Management Book Series, 0 , 23-46.	0.4	1
11274	Compilation of References. , 0 , 0-0.		0
11275	Entity Resolution on Cloud. Advances in Data Mining and Database Management Book Series, 0 , 222-235.	0.4	0
11276	A Domain-Specific Language for High-Level Parallelization. , 0 , 276-295.		0
11277	Perspectives of Big Data Analysis for Knowledge Generation in Project Management Contexts. , 0 , 2019-2037.		0
11278	Predictive Analytics in Operations Management. Advances in Business Information Systems and Analytics Book Series, 0 , 68-92.	0.3	0
11279	Predictive Analytics in Operations Management. , 0 , 1205-1223.		0
11280	Overview of Big Data in Healthcare. Advances in Healthcare Information Systems and Administration Book Series, 0 , 360-384.	0.2	0
11281	StreamMine. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 0 , 260-283.	0.5	0
11282	Analyzing Concurrent Programs Title for Potential Programming Errors. Advances in Computer and Electrical Engineering Book Series, 0 , 380-415.	0.2	0
11283	Large Scale Graph Mining with MapReduce. Advances in Data Mining and Database Management Book Series, 0 , 299-314.	0.4	0
11284	Grid, SOA and Cloud Computing. , 0 , 45-85.		1
11285	Large Scale Graph Mining with MapReduce. , 0 , 66-78.		0

#	ARTICLE	IF	CITATIONS
11286	Economic, Political and Social Threats in the Information Age. , 0, , 17-30.		0
11287	Data Intensive Computing for Bioinformatics. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 0, , 207-241.	0.5	1
11288	Customizable Instance-Driven Webpage Filtering Based on Semi-Supervised Learning. , 2007, , .		5
11291	Big Data Technologies: A Comprehensive Survey. Lecture Notes in Networks and Systems, 2021, , 103-114.	0.5	1
11292	An Approach to Improved MapReduce and Aggregation Pipeline Utilizing NoSQL Technologies. Lecture Notes in Networks and Systems, 2021, , 187-195.	0.5	0
11293	Job-Aware File-Storage Optimization for Improved Hadoop I/O Performance. IEICE Transactions on Information and Systems, 2020, E103.D, 2083-2093.	0.4	2
11294	Lightweight image classifier using dilated and depthwise separable convolutions. Journal of Cloud Computing: Advances, Systems and Applications, 2020, 9, .	2.1	14
11295	Memory Management Approaches in Apache Spark: A Review. Advances in Intelligent Systems and Computing, 2021, , 394-403.	0.5	6
11296	Research on Multi-source Heterogeneous Big Data in Extra-large Enterprises. , 2020, , .		1
11297	Parallel Computing Framework Based on MapReduce and GPU Clusters. , 2020, , .		0
11299	Parallelizing Parallel Programs. , 2020, , .		0
11300	Distributed Approximation Algorithms for Steiner Tree in the CONGESTED CLIQUE. International Journal of Foundations of Computer Science, 2020, 31, 941-968.	0.8	5
11301	SparkBeagle. , 2020, , .		1
11302	Speeding up the data extraction of machine learning approaches: a distributed framework. , 2020, , .		0
11303	Scalable and serializable networked multi-actor programming. , 2020, 4, 1-30.		3
11304	Design and Implementation of Componentized Big Data Platform. , 2020, , .		0
11305	Security of AI Hardware Systems. , 2021, , 95-110.		0
11306	A big data architecture to a multiple purpose in healthcare surveillance. , 2020, , .		1

#	ARTICLE	IF	CITATIONS
11307	ESTemd: A Distributed Processing Framework for Environmental Monitoring based on Apache Kafka Streaming Engine. , 2020, , .		5
11309	Big Data Techniques, Tools, and Applications. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 0, , 185-212.	0.5	1
11311	UniDRM: Unified Data and Resource Management for Federated Vehicular Cloud Computing. IEEE Access, 2021, 9, 157052-157067.	2.6	3
11312	Big Data Architecture for Building Energy Management Systems. IEEE Transactions on Industrial Informatics, 2022, 18, 5738-5747.	7.2	6
11313	Evaluation Platform for DDM Algorithms With the Usage of Non-Uniform Data Distribution Strategies. International Journal of Information Technologies and Systems Approach, 2021, 15, 1-23.	0.8	6
11314	Amnis: Optimized stream processing for edge computing. Journal of Parallel and Distributed Computing, 2022, 160, 49-64.	2.7	11
11315	Scalable Mining of High-Utility Sequential Patterns With Three-Tier MapReduce Model. ACM Transactions on Knowledge Discovery From Data, 2022, 16, 1-26.	2.5	17
11316	XML2HBase: Storing and querying large collections of XML documents using a NoSQL database system. Journal of Parallel and Distributed Computing, 2022, 161, 83-99.	2.7	5
11317	A Context-based Sensed Data Search on Edge Computing for Finding Moving People. , 2021, , .		0
11318	Optimal Design of Collecting Module in Monitoring System. , 2021, , .		0
11319	Rateless Codes for Distributed Non-linear Computations. , 2021, , .		1
11320	Software Services Engineering Manifesto - A Cross-Cutting Declaration. , 2021, , .		2
11321	Effect of Monotonic Filtering on Graph Collection Dynamics. , 2021, , .		1
11322	What Distributed Systems Say: A Study of Seven Spark Application Logs. , 2021, , .		3
11323	Enabling Low-Redundancy Proactive Fault Tolerance for Stream Machine Learning via Erasure Coding. , 2021, , .		0
11324	Decoupling Control and Data Transmission in RDMA Enabled Cloud Data Centers. , 2021, , .		1
11325	Scalable Learning to Troubleshoot Query Performance Problems. , 2021, , .		2
11326	A Critical Analysis of Apache Hadoop and Spark for Big Data Processing. , 2021, , .		6

#	ARTICLE	IF	CITATIONS
11327	Scalable Global Mutual Information Based Feature Selection Framework for Large Scale Datasets. , 2021, , .		1
11328	A novel temporal frequent subgraph based mining algorithm using spark. Intelligent Decision Technologies, 2021, , 1-10.	0.6	0
11329	A Spark-based high utility itemset mining with multiple external utilities. Cluster Computing, 2022, 25, 889-909.	3.5	2
11330	Decentralized learning of randomization-based neural networks with centralized equivalence. Applied Soft Computing Journal, 2022, 115, 108030.	4.1	7
11332	Research on Stereo Library Based on Homogeneous Markov Chain. Journal of Physics: Conference Series, 2021, 2083, 032039.	0.3	0
11333	A Comparative Analysis of Big Data Frameworks: An Adoption Perspective. Applied Sciences (Switzerland), 2021, 11, 11033.	1.3	8
11334	The Collaborative Research Center FONDA. Datenbank-Spektrum, 2021, 21, 255-260.	1.2	3
11335	Large-scale interactive retrieval in art collections using multi-style feature aggregation. PLoS ONE, 2021, 16, e0259718.	1.1	3
11337	Extensions and Complements. , 2021, , 568-590.		0
11338	Accelerating temporal action proposal generation via high performance computing. Frontiers of Computer Science, 2022, 16, 1.	1.6	2
11339	Integrated benchmarking standard and decision support system for structured, semi structured, unstructured retail data. Wireless Networks, 0, , 1.	2.0	1
11340	Efficient inference models for classification problems with a high number of fuzzy rules. Applied Soft Computing Journal, 2022, 115, 108164.	4.1	5
11341	Genomic Data and Big Data Analytics. Lecture Notes in Networks and Systems, 2022, , 187-201.	0.5	0
11342	Practice of Streaming Processing of Dynamic Graphs: Concepts, Models, and Systems. IEEE Transactions on Parallel and Distributed Systems, 2023, 34, 1860-1876.	4.0	11
11343	Distributed Linearly Separable Computation. IEEE Transactions on Information Theory, 2022, 68, 1259-1278.	1.5	7
11344	Joint Radio Resource Allocation and Beamforming Optimization for Industrial Internet of Things in Software-Defined Networking-Based Virtual Fog-Radio Access Network 5G-and-Beyond Wireless Environments. IEEE Transactions on Industrial Informatics, 2022, 18, 4198-4209.	7.2	11
11345	Securing Big Data Scientific Workflows via Trusted Heterogeneous Environments. IEEE Transactions on Dependable and Secure Computing, 2022, 19, 4187-4203.	3.7	2
11346	Setting up Hadoop Environment and Running MapReduce Job for Teaching BDA in Universities â€œ An Experience. Communications in Computer and Information Science, 2021, , 57-71.	0.4	0

#	ARTICLE	IF	CITATIONS
11347	Hadoop Data Reduction Framework: Applying Data Reduction at the DFS Layer. IEEE Access, 2021, 9, 152704-152717.	2.6	1
11348	A Robust Service Mapping Scheme for Multi-Tenant Clouds. IEEE/ACM Transactions on Networking, 2022, 30, 1146-1161.	2.6	1
11349	Distributed Decoding for Coded Distributed Computing. IEEE Internet of Things Journal, 2022, 9, 12555-12562.	5.5	1
11351	Computational Polarization: An Information-Theoretic Method for Resilient Computing. IEEE Transactions on Information Theory, 2022, 68, 2211-2238.	1.5	2
11352	Ernie: Scalable Load-Balanced Multicast Source Routing for Cloud Data Centers. IEEE Access, 2021, 9, 168816-168830.	2.6	3
11353	A Network Load Perception based Task Scheduler for Parallel Distributed Data Processing Systems. IEEE Transactions on Cloud Computing, 2021, , 1-1.	3.1	0
11354	Stability and Optimization of Speculative Queueing Networks. IEEE/ACM Transactions on Networking, 2022, 30, 911-922.	2.6	1
11355	Massively Parallel and Dynamic Algorithms for Minimum Size Clustering. , 2022, , 1613-1660.		1
11356	An efficient skyline query algorithm in the distributed environment. Journal of Computational Science, 2022, 58, 101524.	1.5	2
11357	Distributed evidential clustering toward time series with big data issue. Expert Systems With Applications, 2022, 191, 116279.	4.4	12
11358	The Design of College Student Growth Management System Based on Data Mining. , 0, , .		0
11359	Evaluation of thread-local garbage collection. , 2020, , .		1
11360	An efficient Hadoop-based brain tumor detection framework using big data analytic. Software - Practice and Experience, 2022, 52, 805-818.	2.5	4
11361	Large-Scale Distributed Linkage of Records Containing Spatio-Temporal Information. , 2020, , .		0
11362	Incremental Type Migration Using Type Algebra. , 2020, , .		6
11364	Multitasking Elastic Scheduling Cluster in Tensorflow. , 2020, , .		0
11365	S3QLRDF: Property Table Partitioning Scheme for Distributed SPARQL Querying of large-scale RDF data. , 2020, , .		4
11366	MRSweep: Distributed In-Memory Sweep-line for Scalable Object Intersection Problems. , 2020, , .		0

#	ARTICLE	IF	CITATIONS
11367	Resource and Job Execution Context-Aware Hadoop Configuration Tuning. , 2020, , .		0
11368	On The Distributed Determinization Of Large NFAs. , 2020, , .		2
11369	Computation Model of Data Intensive Computing with MapReduce. , 2020, , .		1
11370	High Performance Data Engineering Everywhere. , 2020, , .		10
11371	Big Data Preprocessing: Needs and Methods. SSRG International Journal of Engineering Trends and Technology, 2020, 68, 100-104.	0.3	1
11372	Research on the Application of Transparent Encryption in Distributed File System HDFS. , 2020, , .		2
11373	Comparative Survey on Big data Security Applications, A Blink on Interactive Security Mechanism in Apache Ozone. , 2020, , .		1
11374	A Survey On Cloud-Based Distributed Computing System Frameworks. , 2020, , .		2
11375	Performance Analysis of Sales Big Data Processing using Hadoop and Hive in Cloud Environment. , 2020, , .		1
11376	Scalable and Hybrid Ensemble-Based Causality Discovery. , 2020, , .		2
11377	Design and Implementation of Technical Indicators Analysis System for Drilling Based on Microservices. , 2020, , .		1
11378	Coded Computing at Full Speed. , 2020, , .		1
11379	Performance Improvement of Hadoop ext4-based Disk I/O. , 2020, , .		0
11380	Trying to do it all in a single course: a surprisingly good idea. , 2020, , .		0
11381	SEIZE User Desired Moments: Runtime Inspection for Parallel Dataflow Systems. , 2020, , .		0
11382	Collage Inference: Using Coded Redundancy for Lowering Latency Variation in Distributed Image Classification Systems. , 2020, , .		5
11383	Parallel based Hiding of Sensitive Knowledge. , 2020, , .		4
11384	Edge-Stream: a Stream Processing Approach for Distributed Applications on a Hierarchical Edge-computing System. , 2020, , .		5

#	ARTICLE	IF	CITATIONS
11385	Scalable querying of nested data. Proceedings of the VLDB Endowment, 2020, 14, 445-457.	2.1	6
11386	Big Data Analytics in Online Structural Health Monitoring. International Journal of Prognostics and Health Management, 2016, 7, .	0.6	2
11387	D-colSimulation: A Distributed Approach for Frequent Graph Pattern Mining based on colSimulation in a Single Large Graph. , 2020, , .		0
11388	Basic Knowledge Construction Technique to Reduce The Volume of Low-Dimensional Big Data. , 2020, , .		1
11389	Introducing Big Data to Undergraduate Students: A Novel Approach in a Distance University. Revista Iberoamericana De Tecnologias Del Aprendizaje, 2020, 15, 291-298.	0.7	1
11390	The Subnetwork Investigation of Scale-Free Networks Based on the Self-Similarity. SSRN Electronic Journal, 0, , .	0.4	0
11391	Benchmarking Apache Spark and Hadoop MapReduce on Big Data Classification. , 2021, , .		5
11392	Edge-Based Sampling for Complex Network with Self-Similar Structure. , 2021, , .		1
11393	Back Propagation Neural Network with Feature Sensitivity Analysis in Railways: Short Term Prediction of Passenger Flow using Time, Environmental and Operational Factors. , 2021, , .		0
11394	NetANNS: A High-Performance Distributed Search Framework Based On In-Network Computing. , 2021, , .		0
11395	AHA: Adaptive Hadoop in Ad-hoc Cloud Environments. , 2021, , .		0
11396	AMG: An Auto-Tuning Method for Spark GraphX Running on GPU Cluster. , 2021, , .		0
11397	Index Matrix Representation of MapReduce Algorithm. , 2021, , .		0
11398	PSACS: Highly-Parallel Shuffle Accelerator on Computational Storage. , 2021, , .		2
11399	Energy-aware Coflow Scheduling for Sustainable Workload Management. , 2021, , .		1
11400	Expert System For Comparative Diagnosis Probability Of Lung Disease Using Nearest Neighbour Algorithm. , 2021, , .		0
11401	NetFC: Enabling Accurate Floating-point Arithmetic on Programmable Switches. , 2021, , .		13
11402	Enabling Large-Scale Federated Learning over Wireless Edge Networks. , 2021, , .		2

#	ARTICLE	IF	CITATIONS
11403	Halvade somatic: Somatic variant calling with Apache Spark. <i>GigaScience</i> , 2022, 11, .	3.3	2
11404	The role of big data analytics in the context of modeling design and operation of manufacturing systems. , 2022, , 243-275.		11
11405	Scheduling Parallel-Task Jobs Subject to Packing and Placement Constraints. <i>Operations Research</i> , 2022, 70, 3403-3419.	1.2	0
11406	Forecasting: theory and practice. <i>International Journal of Forecasting</i> , 2022, 38, 705-871.	3.9	256
11407	State-of-the-Art Survey on Web Search. <i>Studies in Big Data</i> , 2022, , 1-24.	0.8	1
11408	Efficient evolutionary computation model of closed high-utility itemset mining. <i>Applied Intelligence</i> , 2022, 52, 10604-10616.	3.3	16
11409	Programming big data analysis: principles and solutions. <i>Journal of Big Data</i> , 2022, 9, .	6.9	24
11411	BigFiRSt: A Software Program Using Big Data Technique for Mining Simple Sequence Repeats From Large-Scale Sequencing Data. <i>Frontiers in Big Data</i> , 2021, 4, 727216.	1.8	2
11413	Enabling Type Checking on Columns in Data Frame Libraries by Abstract Interpretation. <i>IEEE Access</i> , 2022, 10, 14418-14428.	2.6	0
11414	Edge Artificial Intelligence for 6G: Vision, Enabling Technologies, and Applications. <i>IEEE Journal on Selected Areas in Communications</i> , 2022, 40, 5-36.	9.7	206
11416	Metaheuristics for data mining: survey and opportunities for big data. <i>Annals of Operations Research</i> , 2022, 314, 117-140.	2.6	12
11417	M ² : Mixed Models With Preferences, Popularities and Transitions for Next-Basket Recommendation. <i>IEEE Transactions on Knowledge and Data Engineering</i> , 2023, 35, 4033-4046.	4.0	6
11418	A Multi-modal Ride Sharing Framework for Last Mile Connectivity. , 2022, , .		1
11419	A Case for Sampling-Based Learning Techniques in Coflow Scheduling. <i>IEEE/ACM Transactions on Networking</i> , 2022, 30, 1494-1508.	2.6	5
11420	A Topology-Aware Scheduling Strategy for Distributed Stream Computing System. <i>Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering</i> , 2022, , 132-147.	0.2	2
11422	Weather Forecast Using Sliding Window Algorithm Based on Hadoop and MapReduce. <i>Lecture Notes in Networks and Systems</i> , 2022, , 122-132.	0.5	2
11423	Equivalence classes and conditional hardness in massively parallel computations. <i>Distributed Computing</i> , 2022, 35, 165-183.	0.7	3
11424	A Survey on Spatio-temporal Data Analytics Systems. <i>ACM Computing Surveys</i> , 2022, 54, 1-38.	16.1	19

#	ARTICLE	IF	CITATIONS
11425	Speed Scaling on Parallel Servers With MapReduce Type Precedence Constraints. IEEE/ACM Transactions on Networking, 2022, 30, 1509-1524.	2.6	0
11426	Historical data based approach to mitigate stragglers from the Reduce phase of MapReduce in a heterogeneous Hadoop cluster. Cluster Computing, 0, , .	3.5	5
11427	Toward a reference architecture based science gateway framework with embedded <sc>learning</sc> support. Concurrency Computation Practice and Experience, 2023, 35, .	1.4	0
11428	Multi-objective optimization of data deployment and scheduling based on the minimum cost in geo-distributed cloud. Computer Communications, 2022, 185, 142-158.	3.1	7
11429	Service-based Analytics for 5G open experimentation platforms. Computer Networks, 2022, 205, 108740.	3.2	3
11430	NASCENT2: Generic Near-Storage Sort Accelerator for Data Analytics on SmartSSD. ACM Transactions on Reconfigurable Technology and Systems, 2022, 15, 1-29.	1.9	6
11431	Extending an asynchronous runtime system for high throughput applications: A case study. Journal of Parallel and Distributed Computing, 2022, 163, 214-231.	2.7	1
11432	Big data architecture and data mining analysis for market segment applications of differential global navigation satellite system (GNSS) services: case study of the analysis of the demand for navigation and agriculture. Journal of Navigation, 2022, 75, 421-436.	1.0	2
11433	HPTMT Parallel Operators for High Performance Data Science and Data Engineering. Frontiers in Big Data, 2021, 4, 756041.	1.8	1
11434	Application of cognitive computing in healthcare, cybersecurity, big data and IoT: A literature review. Information Processing and Management, 2022, 59, 102888.	5.4	61
11435	C4I System Security Architecture: A Perspective on Big Data Lifecycle in a Military Environment. Sustainability, 2021, 13, 13827.	1.6	3
11436	Block Storage Optimization and Parallel Data Processing and Analysis of Product Big Data Based on the Hadoop Platform. Mathematical Problems in Engineering, 2021, 2021, 1-14.	0.6	51
11439	IoT Cloud Computing Middleware for Crowd Monitoring and Evacuation. International Journal of Circuits, Systems and Signal Processing, 2021, 15, 1790-1802.	0.2	1
11450	A Unifying Framework for Parallel and Distributed Processing in R using Futures. R Journal, 2021, 13, 208.	0.7	44
11452	Recent Advances in Data Engineering for Networking. IEEE Access, 2022, 10, 34449-34496.	2.6	3
11454	Accelerated Distributed Approximate Newton Method. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34, 8642-8653.	7.2	3
11457	Efficient Large Scale NLP Feature Engineering with Apache Spark. , 2022, , .		14
11459	AutoFlow: Hotspot-Aware, Dynamic Load Balancing for Distributed Stream Processing. Lecture Notes in Computer Science, 2022, , 133-151.	1.0	2

#	ARTICLE	IF	CITATIONS
11461	Hierarchical Coded Matrix Multiplication in Heterogeneous Multihop Networks. IEEE Transactions on Communications, 2022, 70, 3597-3612.	4.9	2
11463	Recent Advancement and Challenges in Deep Learning, Big Data in Bioinformatics. Studies in Big Data, 2022, , 251-284.	0.8	1
11464	Optimizing Data Locality by Executor Allocation in Reduce Stage for Spark Framework. Lecture Notes in Computer Science, 2022, , 349-357.	1.0	0
11465	MapReduce Preprocess of Big Graphs for Rapid Connected Components Detection. , 2022, , .		15
11466	Using HW/SW Codesign for Deep Neural Network Hardware Accelerator Targeting Low-Resources Embedded Processors. IEEE Access, 2022, 10, 22274-22287.	2.6	4
11467	Toward a Systematic Survey for Carbon Neutral Data Centers. IEEE Communications Surveys and Tutorials, 2022, 24, 895-936.	24.8	28
11468	A User-based Collaborative Filtering System for Deck Recommendation in Game Clash Royale. , 2022, , .		0
11470	Feature Selection: From the Past to the Future. Learning and Analytics in Intelligent Systems, 2022, , 11-34.	0.5	4
11471	ProbSky: Efficient Computation of Probabilistic Skyline Queries over Distributed Data. IEEE Transactions on Knowledge and Data Engineering, 2022, , 1-1.	4.0	1
11472	Enabling pipeline parallelism in heterogeneous managed runtime environments via batch processing. , 2022, , .		1
11473	UACD: A Local Approach for Identifying the Most Influential Spreaders in Twitter in a Distributed Environment. Social Network Analysis and Mining, 2022, 12, 1.	1.9	3
11475	Effective and scalable legal judgment recommendation using pre-learned word embedding. Complex & Intelligent Systems, 2022, 8, 3199-3213.	4.0	7
11476	Big Data Integration Framework for Processing Petrophysical Data. , 2022, , .		0
11477	Cost Model Based Incremental Processing in Dynamic Graphs. Electronics (Switzerland), 2022, 11, 660.	1.8	1
11478	Personalized Liver Cancer Risk Prediction Using Big Data Analytics Techniques with Image Processing Segmentation. Computational Intelligence and Neuroscience, 2022, 2022, 1-11.	1.1	2
11479	A classification framework for straggler mitigation and management in a heterogeneous Hadoop cluster: A state-of-art survey. Journal of King Saud University - Computer and Information Sciences, 2022, 34, 7621-7644.	2.7	6
11480	Fast density peaks clustering algorithm in polar coordinate system. Applied Intelligence, 2022, 52, 14478-14490.	3.3	5
11481	PerFlow. , 2022, , .		3

#	ARTICLE	IF	CITATIONS
11482	Jiffy. , 2022, , .		15
11483	Parallel block-delayed sequences. , 2022, , .		3
11484	GraphBLAST: A High-Performance Linear Algebra-based Graph Framework on the GPU. ACM Transactions on Mathematical Software, 2022, 48, 1-51.	1.6	18
11485	ROME: All Overlays Lead to Aggregation, but Some Are Faster than Others. ACM Transactions on Computer Systems, 2021, 39, 1-33.	0.6	0
11486	Highlighting Current Issues in API Usage Mining to Enhance Software Reusability. WSEAS Transactions on Computer Research, 2022, 10, 29-34.	0.3	0
11487	Stream processing with dependency-guided synchronization. , 2022, , .		3
11488	Transactions across serverless functions leveraging stateful dataflows. Information Systems, 2022, 108, 102015.	2.4	3
11489	DOSP: an optimal synchronization of parameter server for distributed machine learning. Journal of Supercomputing, 2022, 78, 13865-13892.	2.4	1
11490	A comparative performance study of spark on kubernetes. Journal of Supercomputing, 2022, 78, 13298-13322.	2.4	2
11491	A Novel Consensus Algorithm Based on Segmented DAG and BP Neural Network for Consortium Blockchain. Security and Communication Networks, 2022, 2022, 1-16.	1.0	1
11493	A Survey of Biological Data in a Big Data Perspective. Big Data, 2022, 10, 279-297.	2.1	8
11494	Bioinformatic Challenges Detecting Genetic Variation in Precision Medicine Programs. Frontiers in Medicine, 2022, 9, 806696.	1.2	3
11495	Understanding the performance of erasure codes in hadoop distributed file system. , 2022, , .		0
11496	OSBN: architecture and control mechanism of optical switched satellite backbone network. Photonic Network Communications, 2022, 43, 165-176.	1.4	2
11497	You are what the permissions told me! Android malware detection based on hybrid tactics. Journal of Information Security and Applications, 2022, 66, 103159.	1.8	12
11498	B-EagleV: Visualization of Big Point Cloud Datasets in Civil Engineering Using a Distributed Computing Solution. Journal of Computing in Civil Engineering, 2022, 36, .	2.5	7
11499	GiantVM: A Novel Distributed Hypervisor for Resource Aggregation with DSM-aware Optimizations. Transactions on Architecture and Code Optimization, 2022, 19, 1-27.	1.6	2
11500	A Topological Approach to Partitioning Flow Networks for Parallel Simulation. Journal of Computing in Civil Engineering, 2022, 36, .	2.5	2

#	ARTICLE	IF	CITATIONS
11501	Stateful Serverless Computing with <scp>Crucial</scp>. ACM Transactions on Software Engineering and Methodology, 2022, 31, 1-38.	4.8	16
11504	Distributed Fuzzy Clustering Based Association Rule Mining: Design, Deployment and Implementation. , 2021, , .		1
11505	Data-centric Model Transformation Approach using Model2GraphFrame Transformations. Journal of Software Engineering Research and Development, 0, 9, .	1.0	0
11506	A Parallel Association Rule Mining Algorithm Based on Spark and its application. , 2021, , .		0
11507	T-Rank: Graph Data Analytics for Urban Traffic Modeling. , 2021, , .		1
11508	Receiver-Driven RDMA Congestion Control by Differentiating Congestion Types in Datacenter Networks. , 2021, , .		6
11509	R-AQM: Reverse ACK Active Queue Management in Multi-tenant Data Centers. , 2021, , .		1
11510	Incremental partitioning for efficient spatial data analytics. Proceedings of the VLDB Endowment, 2021, 15, 713-726.	2.1	3
11511	On Exploring the Optimum Configuration of Apache Spark Framework in Heterogeneous Clusters. , 2021, , .		0
11512	Big Data: Finding Frequencies of Faulty Multimedia Data. , 2021, , .		0
11514	A Scalable Short-Text Clustering Algorithm Using Apache Spark. , 2021, , .		0
11515	K-Cosine-Medoids Clustering Algorithm. , 2021, , .		0
11516	Weather Forecasting Framework for Time Series Data using Intelligent Learning Models. , 2021, , .		3
11517	Emerging Strategies to Big Data Analytics in Healthcare. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 2021, , 254-260.	0.2	0
11518	Training Data Reduction for Performance Models of Data Analytics Jobs in the Cloud. , 2021, , .		4
11519	Overview of Optimization Algorithms for Large-scale Support Vector Machines. , 2021, , .		1
11520	Design of Parallel Big Data Stream Transmission System Based on Chaotic Migration Learning. , 2021, , .		0
11521	Lightweight domain adaptation: A filtering pipeline to improve accuracy of an Automatic Speech Recognition (ASR) engine. , 2021, , .		1

#	ARTICLE	IF	CITATIONS
11522	Cherry: A Distributed Task-Aware Shuffle Service for Serverless Analytics. , 2021, , .		3
11523	Dynamic Length Factorization Machines for CTR Prediction. , 2021, , .		2
11524	Parallel Actors and Learners: A Framework for Generating Scalable RL Implementations. , 2021, , .		2
11525	Research on Automatic Online Analysis Method of Data Hotness in Big Data Scenario. , 2021, , .		0
11526	Scaling pair count to next galaxy surveys. Monthly Notices of the Royal Astronomical Society, 2022, 510, 3085-3097.	1.6	0
11527	Ernie: Data Center Multicast Source Routing. , 2021, , .		0
11528	Parallel Particle Swarm Optimization Based on Spark for Academic Paper Co-Authorship Prediction. Information (Switzerland), 2021, 12, 530.	1.7	2
11529	RStream: Simple and Efficient Batch and Stream Processing at Scale. , 2021, , .		2
11530	Rectilinear Range Query Processing on SpatialHadoop Platform. , 2021, , .		0
11531	A GPU Algorithm for Detecting Contextual Outliers in Multiple Concurrent Data Streams. , 2021, , .		6
11532	An Approach to Improve Distributed Computing. , 2021, , .		0
11533	Communication efficient distributed learning of neural networks in Big Data environments using Spark. , 2021, , .		1
11534	Translation of Array-Based Graph Programs to Spark SQL on Block Arrays. , 2021, , .		3
11535	Parallelization of Sequential Pattern Sampling. , 2021, , .		1
11536	Benchmarking and Performance Modelling of Dataflow with Cycles. , 2021, , .		0
11537	Tolerating Adversarial Attacks and Byzantine Faults in Distributed Machine Learning. , 2021, , .		2
11538	Large-Scale Closed High-Utility Itemset Mining. , 2021, , .		2
11539	Evaluating Serverless Architecture for Big Data Enterprise Applications. , 2021, , .		1

#	ARTICLE	IF	CITATIONS
11540	Greedy Nominator Heuristic: Virtual function placement on fog resources. Concurrency Computation Practice and Experience, 2022, 34, .	1.4	4
11541	A Proactive Data-Parallel Framework for Machine Learning. , 2021, , .		1
11542	MultiBlock: A Scalable Iterative Approach for Progressive Entity Resolution. , 2021, , .		1
11543	Cost-aware & Fault-tolerant Geo-distributed Edge Computing for Low-latency Stream Processing. , 2021, , .		0
11544	Scalable robust graph embedding with Spark. Proceedings of the VLDB Endowment, 2021, 15, 914-922.	2.1	2
11545	Chukonu. Proceedings of the VLDB Endowment, 2021, 15, 872-885.	2.1	0
11546	Parallel Population-Based Simulated Annealing for High-Dimensional Black-Box Optimization. , 2021, , .		0
11547	Multi-Level Over-the-Air Aggregation of Mobile Edge Computing Over D2D Wireless Networks. IEEE Transactions on Wireless Communications, 2022, 21, 8337-8353.	6.1	9
11548	An adaptive geographic meshing and coding method for remote sensing data. IOP Conference Series: Earth and Environmental Science, 2022, 1004, 012006.	0.2	0
11549	Towards Resilient and Efficient Big Data Storage: Evaluating a SIEM Repository Based on HDFS. , 2022, , .		1
11550	Technological Surveillance in Big Data Environments by using a MapReduce-based Method. Mobile Networks and Applications, 2022, 27, 1931-1940.	2.2	1
11551	Application Massive Data Processing Platform for Smart Manufacturing Based on Optimization of Data Storage. ACM Transactions on Management Information Systems, 2022, 13, 1-20.	2.1	1
11552	PRINS: scalable model inference for component-based system logs. Empirical Software Engineering, 2022, 27, 87.	3.0	2
11562	ServiceCombinatorsfor FarmingVirtualMachines. , 2008, , 22-22.		0
11620	Applications of Big Data Analytics in Investment Management. Journal of Database Management, 2022, 33, 1-32.	1.0	2
11621	Object Detection in Remote Sensing Images Using Picture Fuzzy Clustering and MapReduce. Computer Systems Science and Engineering, 2022, 43, 1241-1253.	1.9	2
11622	Networking Systems of AI: On the Convergence of Computing and Communications. IEEE Internet of Things Journal, 2022, 9, 20352-20381.	5.5	22
11623	Soft Computing and Machine Learning Techniques for e-Health Data Analytics. Studies in Computational Intelligence, 2022, , 83-104.	0.7	2

#	ARTICLE	IF	CITATIONS
11625	Joint Design of Shuffling and Function Assignment in Heterogeneous Coded Distributed Computing. IEEE Transactions on Signal Processing, 2022, 70, 2560-2575.	3.2	5
11626	Carbon-Aware Computing for Datacenters. IEEE Transactions on Power Systems, 2023, 38, 1270-1280.	4.6	24
11628	Distributed Business Process Discovery in Cloud Clusters. International Journal of Distributed Artificial Intelligence, 2022, 14, 0-0.	0.6	0
11629	Characterizing Java Streams in the Wild. , 2022, , .		2
11630	A comprehensive bibliometric analysis of Apache Hadoop from 2008 to 2020. International Journal of Intelligent Computing and Cybernetics, 2023, 16, 99-120.	1.6	22
11631	A Distributed N-FINDR Cloud Computing-Based Solution for Endmembers Extraction on Large-Scale Hyperspectral Remote Sensing Data. Remote Sensing, 2022, 14, 2153.	1.8	4
11632	Rateless codes for near-perfect load balancing in distributed matrix-vector multiplication. Communications of the ACM, 2022, 65, 111-118.	3.3	4
11633	Data Gathering, Processing, and Visualization for COVID-19. Advances in Data Mining and Database Management Book Series, 2022, , 1-27.	0.4	0
11634	MINING ONLINE REVIEWS TO SUPPORT CUSTOMERS'™ DECISION-MAKING PROCESS IN E-COMMERCE PLATFORMS: A NARRATIVE LITERATURE REVIEW. Journal of Organizational Computing and Electronic Commerce, 2022, 32, 69-97.	1.0	2
11635	Big Data Seismology. Reviews of Geophysics, 2022, 60, .	9.0	24
11636	Analyzing large-scale Data Cubes with user-defined algorithms: A cloud-native approach. International Journal of Applied Earth Observation and Geoinformation, 2022, 109, 102784.	0.9	7
11637	Efficient parallel processing of high-dimensional spatial kNN queries. Soft Computing, 2022, 26, 12291-12316.	2.1	1
11638	FPGA acceleration of deep reinforcement learning using on-chip replay management. , 2022, , .		3
11639	BBoxDB streams: scalable processing of multi-dimensional data streams. Distributed and Parallel Databases, 2022, 40, 559-625.	1.0	1
11640	A Near-Optimal Parallel Algorithm for Joining Binary Relations. Logical Methods in Computer Science, 0, Volume 18, Issue 2, .	0.4	0
11641	A Chinese Intelligent Teaching Platform for Colleges Based on Cloud Computing. Mobile Information Systems, 2022, 2022, 1-10.	0.4	1
11642	Controlled outputs, full data: A privacy-protecting infrastructure for <scp>MOOC</scp> data. British Journal of Educational Technology, 2022, 53, 756-775.	3.9	7
11643	Optimizing multicast flows in high-bandwidth reconfigurable datacenter networks. Journal of Network and Computer Applications, 2022, 203, 103399.	5.8	3

#	ARTICLE	IF	CITATIONS
11644	BP-SPARQL: A Query Language for Summarizing and Analyzing Big Process Data. , 2022, , 21-48.		4
11645	An Efficient and Privacy-Preserving Similarity Evaluation for Big Data Analytics. , 2015, , .		3
11646	Smart Shuffling in MapReduce: A Solution to Balance Network Traffic and Workloads. , 2015, , .		4
11648	Programmable Implementation and Blockchain Security Scheme Based on Edge Computing Firework Model. , 2022, , 480-503.		0
11649	A Distributed Collaborative Entrance Defense Framework Against DDoS Attacks on Satellite Internet. IEEE Internet of Things Journal, 2022, 9, 15497-15510.	5.5	9
11651	Distributed Learning With Dependent Samples. IEEE Transactions on Information Theory, 2022, 68, 6003-6020.	1.5	0
11652	Swallow swarm fuzzy C-Means algorithm (SSFCM) and distributed processing framework for big image data. AIP Conference Proceedings, 2022, , .	0.3	0
11653	Dynamic RAT Selection and Transceiver Optimization for Mobile-Edge Computing Over Multi-RAT Heterogeneous Networks. IEEE Internet of Things Journal, 2022, 9, 20532-20546.	5.5	0
11654	MRMR-GWICA: A hybrid gene selection and ensemble clustering framework for breast cancer gene expression data. AIP Conference Proceedings, 2022, , .	0.3	1
11655	User-defined operators. Proceedings of the VLDB Endowment, 2022, 15, 1119-1131.	2.1	7
11656	Overview of Distributed Storage of Power Grid Control Resources Data. , 2022, , .		1
11657	Harvesting the Aggregate Computing Power of Commodity Computers for Supercomputing Applications. Applied Sciences (Switzerland), 2022, 12, 5113.	1.3	1
11658	IMapC: Inner MAPping Combiner to Enhance the Performance of MapReduce in Hadoop. Electronics (Switzerland), 2022, 11, 1599.	1.8	3
11659	Indexing Metric Spaces for Exact Similarity Search. ACM Computing Surveys, 2023, 55, 1-39.	16.1	4
11660	DDS: integrating data analytics transformations in task-based workflows. Open Research Europe, 0, 2, 66.	2.0	1
11662	Improving the Performance of MapReduce for Small-Scale Cloud Processes Using a Dynamic Task Adjustment Mechanism. Mathematics, 2022, 10, 1736.	1.1	1
11663	A Review on Knowledge Discovery from Databases. Lecture Notes in Electrical Engineering, 2022, , 457-464.	0.3	2
11664	SCMiner: A Fast and Scalable GPU-Based Frequent Pattern Miner on SSDs. IEEE Access, 2022, 10, 62502-62519.	2.6	1

#	ARTICLE	IF	CITATIONS
11665	Cloud-Native Alternating Directions Solver for Isogeometric Analysis. SSRN Electronic Journal, 0, , .	0.4	0
11666	Big Data Analysis Using Hybrid Meta-Heuristic Optimization Algorithm and MapReduce Framework. Studies in Computational Intelligence, 2022, , 181-223.	0.7	6
11667	Optimizing MapReduce Task Scheduling on Virtualized Heterogeneous Environments Using Ant Colony Optimization. IEEE Access, 2022, 10, 55842-55855.	2.6	8
11669	A Scalable Similarity Join Algorithm Based on MapReduce and LSH. International Journal of Parallel Programming, 2022, 50, 360-380.	1.1	1
11670	Exploring core knowledge in business intelligence research. Internet Research, 2023, 33, 1179-1201.	2.7	3
11671	Deadline-aware Preemptive Job Scheduling in Hadoop YARN Clusters. , 2022, , .		3
11672	Collaborative Cluster Configuration for Distributed Data-Parallel Processing: AÂResearch Overview. Datenbank-Spektrum, 0, , .	1.2	1
11673	ALBERT: An automatic learning based execution and resource management system for optimizing Hadoop workload in clouds. Journal of Parallel and Distributed Computing, 2022, 168, 45-56.	2.7	3
11680	Distributed multimedia streaming service algorithms over cloud computing environment. WIT Transactions on Information and Communication Technologies, 2014, , .	0.0	0
11682	A Survey of Binary Code Fingerprinting Approaches: Taxonomy, Methodologies, and Features. ACM Computing Surveys, 2023, 55, 1-41.	16.1	12
11684	The Four Generations of Entity Resolution. Synthesis Lectures on Data Management, 2021, , .	0.6	34
11685	A Data-Driven Approach to Question Subjectivity Identification in Community Question Answering. Proceedings of the AAAI Conference on Artificial Intelligence, 2012, 26, 164-170.	3.6	9
11686	Parallel Computing Optimization for Ground-based TT&C Network Situational Data Processing. , 2021, , .		0
11687	BEDIM: Lateral Movement Detection In Enterprise Network Through Behavior Deviation Measurement. , 2021, , .		0
11688	Massively Digitized Power Grid: Opportunities and Challenges of Use-Inspired AI. Proceedings of the IEEE, 2023, 111, 762-787.	16.4	4
11689	On Distributed Computing With Heterogeneous Communication Constraints. IEEE/ACM Transactions on Networking, 2022, , 1-12.	2.6	3
11690	Big medical data, cloud computing, and artificial intelligence for improving diagnosis in healthcare. , 2022, , 139-150.		0
11692	Efficient Entity Resolution forÂBibliographic Data Using MapReduce. Communications in Computer and Information Science, 2022, , 192-200.	0.4	1

#	ARTICLE	IF	CITATIONS
11693	Coded Computation Across Shared Heterogeneous Workers With Communication Delay. IEEE Transactions on Signal Processing, 2022, 70, 3371-3385.	3.2	5
11694	Domain Adaptive Log Anomaly Prediction for Hadoop System. IEEE Internet of Things Journal, 2022, 9, 20778-20787.	5.5	4
11695	Extensible 3D Simulation of Aggregated Systems with AFCPP. Lecture Notes in Computer Science, 2022, , 55-71.	1.0	4
11696	Big Data Analytics. Springer Handbooks, 2022, , 107-118.	0.3	1
11699	æ°ħ°Žāˆ†â,fâ¼1/2'ç»œçš,,ç »æ•£æˆ;âžž'ä,žç»,â°â æ-1æ³•. Scientia Sinica Mathematica, 2022, , .	0.1	0
11700	Deterministic massively parallel connectivity. , 2022, , .		4
11701	Dew computing-assisted cognitive Intelligence-inspired smart environment for diarrhea prediction. Computing (Vienna/New York), 2022, 104, 2511-2540.	3.2	2
11702	Scaling Equi-Joins. , 2022, , .		1
11703	Mitigating TCP incast in data center networks using enhanced application layer technique. International Journal of Information Technology (Singapore), 2022, 14, 2523-2531.	1.8	1
11704	Correlation-Aware Flow Consolidation for Load Balancing and Beyond. Performance Evaluation Review, 2022, 49, 105-110.	0.4	1
11705	Äika: A Distributed Edge System for AI Inference. Big Data and Cognitive Computing, 2022, 6, 68.	2.9	3
11706	IOSIG: Declarative I/O-Stream Properties Using Pragmas. Datenbank-Spektrum, 2022, 22, 109-119.	1.2	1
11707	Multi-Tenant Cloud Data Services: State-of-the-Art, Challenges and Opportunities. , 2022, , .		4
11708	Near-data Prediction Based Speculative Optimization in a Distribution Environment. Mobile Networks and Applications, 2022, 27, 2339-2347.	2.2	2
11709	Link Load Correlation-Based Blocking Performance Analysis for Tree-Type Data Center Networks. Applied Sciences (Switzerland), 2022, 12, 6235.	1.3	0
11710	Accelerating Join of Distributed Datasets by a Given Criterion. , 2022, , .		0
11711	Perspectives on the integration between first-principles and data-driven modeling. Computers and Chemical Engineering, 2022, 166, 107898.	2.0	39
11712	Optimization Algorithms for Scalable Stream Batch Clustering with k Estimation. Applied Sciences (Switzerland), 2022, 12, 6464.	1.3	0

#	ARTICLE	IF	CITATIONS
11713	Actris 2.0: Asynchronous Session-Type Based Reasoning in Separation Logic. Logical Methods in Computer Science, 0, Volume 18, Issue 2, .	0.4	4
11714	Approximate computing, skeleton programming and run-time scheduling in an algorithm for process design and controllability in distributed and heterogeneous infrastructures. Computers and Chemical Engineering, 2022, 164, 107874.	2.0	1
11715	Linguistic Structure Prediction. Synthesis Lectures on Human Language Technologies, 2011, , .	2.3	32
11717	In-Network Computation for Large-Scale Federated Learning Over Wireless Edge Networks. IEEE Transactions on Mobile Computing, 2023, 22, 5918-5932.	3.9	2
11720	Cost-Based Virtual Machine Scheduling for Data-as-A-Service. SSRN Electronic Journal, 0, , .	0.4	0
11723	RCC: Enabling Receiver-Driven RDMA Congestion Control With Congestion Divide-and-Conquer in Datacenter Networks. IEEE/ACM Transactions on Networking, 2023, 31, 103-117.	2.6	0
11726	Polynomial Voting Rules. SSRN Electronic Journal, 0, , .	0.4	1
11727	In-Memory Indexed Caching for Distributed Data Processing. , 2022, , .		0
11728	Survey of Breast Cancer Detection Using Machine Learning Techniques in Big Data. , 2022, , 1381-1394.		0
11729	Colza: Enabling Elastic In Situ Visualization for High-performance Computing Simulations. , 2022, , .		3
11730	A unifying model for distributed data-intensive systems. , 2022, , .		2
11731	Analysis of market data with Noir. , 2022, , .		1
11732	Epidemiological challenges in pandemic coronavirus disease (<sc>COVID</sc>–19): Role of artificial intelligence. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2022, 12, .	4.6	3
11733	Comparative Analysis of Skew-Join Strategies for Large-Scale Datasets with MapReduce and Spark. Applied Sciences (Switzerland), 2022, 12, 6554.	1.3	2
11734	A Novel Inconsequential Encryption Algorithm for Big Data in Cloud Computing. Journal of Computer Sciences Institute, 0, 23, 140-144.	0.0	0
11735	Server-Side Workload Identification for HPC I/O Requests. , 2022, , .		1
11736	Python-Based Programming Framework for a Heterogeneous MapReduce Architecture. , 2022, , .		1
11737	<i>Gaia</i> Data Release 3. Astronomy and Astrophysics, 2023, 674, A2.	2.1	53

#	ARTICLE	IF	CITATIONS
11738	Programming Concurrent Systems. , 2022, , 27-46.		0
11739	An efficient Markov energy predictor for software defined wireless sensor networks. Wireless Networks, 2022, 28, 3391-3409.	2.0	2
11740	CyberGIS-Cloud: A unified middleware framework for cloud-based geospatial research and education. , 2022, , .		2
11741	Adaptive Massively Parallel Algorithms for Cut Problems. , 2022, , .		0
11742	CNS: Interactive Intelligent Analysis of Financial Management Software Based on Apriori Data Mining Algorithm. International Journal of Cooperative Information Systems, 2021, 30, .	0.6	10
11743	FPGA supported rough set reduct calculation for big datasets. Journal of Intelligent Information Systems, 2022, 59, 779-799.	2.8	3
11744	Content-Based Video Big Data Retrieval with Extensive Features and Deep Learning. Applied Sciences (Switzerland), 2022, 12, 6753.	1.3	11
11745	DiBB. , 2022, , .		0
11746	Speed-robust scheduling: sand, bricks, and rocks. Mathematical Programming, 0, , .	1.6	0
11747	Buffered Streaming Graph Partitioning. Journal of Experimental Algorithmics, 2022, 27, 1-26.	0.7	2
11748	Coded Parallel Transmission for Half-Duplex Distributed Computing. Information (Switzerland), 2022, 13, 342.	1.7	0
11749	Massively Parallel Algorithms for b-Matching. , 2022, , .		0
11750	Multi-Objective reliability-based workflow scheduler: An elastic and persuasive task scheduler based upon <scp>modifiedâ€flower</scp> pollination algorithm in cloud environment. Concurrency Computation Practice and Experience, 2022, 34, .	1.4	1
11751	The subnetwork investigation of scale-free networks based on the self-similarity. Chaos, Solitons and Fractals, 2022, 161, 112140.	2.5	1
11752	An efficient format-independent watermarking framework for large-scale data sets. Expert Systems With Applications, 2022, 208, 118085.	4.4	4
11753	Parallelized Hitting Set Computation for Model-Based Diagnosis. Proceedings of the AAAI Conference on Artificial Intelligence, 2015, 29, .	3.6	9
11754	CrowdMR: Integrating Crowdsourcing with MapReduce for AI-Hard Problems. Proceedings of the AAAI Conference on Artificial Intelligence, 2015, 29, .	3.6	1
11755	Winner Determination in Huge Elections with MapReduce. Proceedings of the AAAI Conference on Artificial Intelligence, 2017, 31, .	3.6	7

#	ARTICLE	IF	CITATIONS
11756	Cross-Domain Ranking via Latent Space Learning. Proceedings of the AAAI Conference on Artificial Intelligence, 2017, 31, .	3.6	2
11757	Data-Parallel Computing Meets STRIPS. Proceedings of the AAAI Conference on Artificial Intelligence, 2013, 27, 474-480.	3.6	1
11758	An Optimal Task Assignment Policy and Performance Diagnosis Strategy for Heterogeneous Hadoop Cluster. Proceedings of the AAAI Conference on Artificial Intelligence, 2013, 27, 1664-1665.	3.6	1
11759	Comparative Study, Applications and Challenges of Big Data Analysis Technologies. , 2019, 15, 66-77.		0
11760	DCoflow: Deadline-Aware Scheduling Algorithm for Coflows in Datacenter Networks. , 2022, , .		1
11761	Early straggler tasks detection by recurrent neural network in a heterogeneous environment. Applied Intelligence, 2023, 53, 7369-7389.	3.3	1
11762	Massively Parallel Computation in a Heterogeneous Regime. , 2022, , .		0
11763	Large-scale multivariate sparse regression with applications to UK Biobank. Annals of Applied Statistics, 2022, 16, .	0.5	3
11764	Modernization of Legacy Systems. , 2015, , .		3
11767	Uncovering Hidden Structure through Parallel Problem Decomposition. Proceedings of the AAAI Conference on Artificial Intelligence, 2014, 28, .	3.6	0
11768	Converting Instance Checking to Subsumption: A Rethink for Object Queries over Practical Ontologies. Proceedings of the AAAI Conference on Artificial Intelligence, 2014, 28, .	3.6	0
11769	Fused Feature Representation Discovery for High-Dimensional and Sparse Data. Proceedings of the AAAI Conference on Artificial Intelligence, 2014, 28, .	3.6	0
11770	Entity Resolution in a Big Data Framework. Proceedings of the AAAI Conference on Artificial Intelligence, 2015, 29, .	3.6	4
11771	Parallel Algorithms for Operations on Multi-Valued Decision Diagrams. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32, .	3.6	0
11772	An alternative C++-based HPC system for Hadoop MapReduce. Open Computer Science, 2022, 12, 238-247.	1.3	3
11773	Renting Servers in the Cloud: The Case of Equal Duration Jobs. SSRN Electronic Journal, 0, , .	0.4	0
11774	Novel Decentralized Security Architecture for the Centralized Storage System in Hadoop using Blockchain Technology. , 2022, , .		8
11775	A Locality-aware Cooperative Distributed Memory Caching for Parallel Data Analytic Applications. , 2022, , .		0

#	ARTICLE	IF	CITATIONS
11776	VeloxDFS: Streaming Access to Distributed Datasets to Reduce Disk Seeks. , 2022, , .		0
11777	Scape: Scalable Collaborative Analytics System on Private Database with Malicious Security. , 2022, , .		2
11778	Stragglers' Detection in Big Data Analytic Systems: The Impact of Heartbeat Arrival. , 2022, , .		1
11779	SparkFlow: Towards High-Performance Data Analytics for Spark-based Genome Analysis. , 2022, , .		3
11780	On Performance Modeling and Prediction for Spark-HBase Applications in Big Data Systems. , 2022, , .		1
11781	ML-ECN: Multi-Level ECN Marking for Fair Datacenter Traffic Forwarding. , 2022, , .		0
11782	A Serverless Engine for High Energy Physics Distributed Analysis. , 2022, , .		4
11783	From Batch Processing to Real Time Analytics: Running Presto® at Scale. , 2022, , .		6
11784	Heterogeneous Coded Distributed Computing with Nonuniform Input File Popularity. , 2022, , .		0
11785	Parallel Hybrid BBO Search Method for Twitter Sentiment Analysis of Large Scale Datasets Using MapReduce. , 2022, , 1219-1236.		0
11786	DoF of a Cooperative X-Channel with an Application to Distributed Computing. , 2022, , .		2
11787	SMART: Speedup Job Completion Time by Scheduling Reduce Tasks. Journal of Computer Science and Technology, 2022, 37, 763-778.	0.9	1
11788	Layered Contention Mitigation for Cloud Storage. , 2022, , .		0
11789	Latency-based Vector Scheduling of Many-task Applications for a Hybrid Cloud. , 2022, , .		2
11790	Sophisticated methods for noise filtering, subgroup discovery, and classification in big data analysis. Journal of Intelligent and Fuzzy Systems, 2022, , 1-17.	0.8	0
11791	A Massively Parallel Modularity-Maximizing Algorithm with Provable Guarantees. , 2022, , .		1
11792	Parallel Weighted Random Sampling. ACM Transactions on Mathematical Software, 2022, 48, 1-40.	1.6	1
11793	Ensemble of Random and Isolation Forests for Graph-Based Intrusion Detection in Containers. , 2022, , .		4

#	ARTICLE	IF	CITATIONS
11794	Improving MapReduce heterogeneous performance using KNN fair share scheduling. Robotics and Autonomous Systems, 2022, 157, 104228.	3.0	4
11795	Mining high utility itemsets with time-aware scheduling using Apache Spark. Concurrency Computation Practice and Experience, 2022, 34, .	1.4	0
11796	Sequential motifs in observed walks. Journal of Complex Networks, 2022, 10, .	1.1	1
11797	Sparx: Distributed Outlier Detection at Scale. , 2022, , .		0
11798	Toward data lakes as central building blocks for data management and analysis. Frontiers in Big Data, 0, 5, .	1.8	2
11799	A dynamic feature selection and intelligent model serving for hybrid batch-stream processing. Knowledge-Based Systems, 2022, 256, 109749.	4.0	3
11800	Enhanced-Sweep: Communication Cost Efficient Top-K Best Region Search. Arabian Journal for Science and Engineering, 2023, 48, 2121-2132.	1.7	1
11801	Scalable Differentially Private Clustering via Hierarchically Separated Trees. , 2022, , .		1
11802	Analyzing Online Transaction Networks with Network Motifs. , 2022, , .		2
11803	Cloud-based storage and computing for remote sensing big data: a technical review. International Journal of Digital Earth, 2022, 15, 1417-1445.	1.6	11
11805	Flash. , 2022, , .		1
11806	Formal Concept Analysis Applications in Bioinformatics. ACM Computing Surveys, 2023, 55, 1-40.	16.1	3
11807	A review of non-cognitive applications for neuromorphic computing. Neuromorphic Computing and Engineering, 2022, 2, 032003.	2.8	13
11808	Parallel incremental efficient attribute reduction algorithm based on attribute tree. Information Sciences, 2022, 610, 1102-1121.	4.0	11
11809	MapReduce accelerated attribute reduction based on neighborhood entropy with Apache Spark. Expert Systems With Applications, 2023, 211, 118554.	4.4	9
11810	Data Management in Machine Learning Systems. Synthesis Lectures on Data Management, 2019, , .	0.6	10
11818	Distributed Clustering Approach by Apache Pyspark Based on Seer for Clinical Data. International Journal of Pattern Recognition and Artificial Intelligence, 0, , .	0.7	0
11819	Mutation Operators for Java Streams. , 2022, , .		0

#	ARTICLE	IF	CITATIONS
11820	V-Matrix: A wave theory of value creation for big data. International Journal of Accounting Information Systems, 2022, 47, 100575.	2.6	3
11821	Situation and Prospects of Big Data Application on Drilling and Completion of Bohai Oilfield. Journal of Engineering Studies, 2020, 12, 136-141.	0.0	0
11822	Big Data Distributed Storage and Processing Case Studies. Lecture Notes in Networks and Systems, 2022, , 826-837.	0.5	0
11823	Alps: An Adaptive Load Partitioning Scaling Solution for Stream Processing System on Skewed Stream. Lecture Notes in Computer Science, 2022, , 17-31.	1.0	0
11824	An Optimized K-means Clustering Approach on Top of MapReduce. Lecture Notes in Electrical Engineering, 2022, , 19-25.	0.3	0
11825	Distributed Computation for Marginal Likelihood based Model Choice. Bayesian Analysis, 2023, 18, .	1.6	0
11826	MapChain: A Blockchain-Based Verifiable Healthcare Service Management in IoT-Based Big Data Ecosystem. IEEE Transactions on Network and Service Management, 2022, 19, 3896-3907.	3.2	18
11827	Random Partition Based Adaptive Distributed Kernelized SVM for Big Data. IEEE Access, 2022, 10, 95623-95637.	2.6	1
11828	Disaster Analysis on Government Data. Advanced Technologies and Societal Change, 2022, , 19-29.	0.8	2
11829	Levelized Multiple Workflow Allocation Strategy Under Precedence Constraints With Task Merging in IaaS Cloud Environment. IEEE Access, 2022, 10, 92809-92827.	2.6	19
11830	Deep Reinforcement Learning for System-on-Chip: Myths and Realities. IEEE Access, 2022, 10, 98048-98064.	2.6	3
11831	KNN Normalized Optimization and Platform Tuning Based on Hadoop. IEEE Access, 2022, 10, 81406-81433.	2.6	5
11832	Multi-agent Reinforcement Learning for Task Allocation in Cooperative Edge Cloud Computing. Lecture Notes in Computer Science, 2022, , 283-297.	1.0	0
11833	A Temporal Ontology Guided Clustering Methodology with a Case Study on Detection and Tracking of Artificial Intelligence Topics. SSRN Electronic Journal, 0, , .	0.4	0
11834	Machine Learning for Software-Defined Networking. SpringerBriefs in Computer Science, 2022, , 1-6.	0.2	0
11835	Development of a framework for adaptive productivity management for edge computing based IoT applications. AIP Conference Proceedings, 2022, , .	0.3	2
11836	Enabling In-Network Floating-Point Arithmetic for Efficient Computation Offloading. IEEE Transactions on Parallel and Distributed Systems, 2022, 33, 4918-4934.	4.0	1
11837	CORE-BCD-mAI: A Composite Framework for Representing, Querying, and Analyzing Big Clinical Data by Means of Multidimensional AI Tools. Lecture Notes in Computer Science, 2022, , 175-185.	1.0	4

#	ARTICLE	IF	CITATIONS
11838	Scalable Distributed Data Anonymization for Large Datasets. IEEE Transactions on Big Data, 2023, 9, 818-831.	4.4	0
11839	Cascaded Coded Distributed Computing Schemes Based on Symmetric Designs. IEEE Transactions on Communications, 2022, 70, 7179-7190.	4.9	1
11840	R-AQM: Reverse ACK Active Queue Management in Multitenant Data Centers. IEEE/ACM Transactions on Networking, 2023, 31, 526-541.	2.6	1
11841	Serving Hybrid-Cloud SQL Interactive Queries at Twitter. Lecture Notes in Computer Science, 2022, , 3-21.	1.0	1
11842	Flexible Distributed Matrix Multiplication. IEEE Transactions on Information Theory, 2022, 68, 7500-7514.	1.5	0
11843	Worker Assignment for Multiple Masters to Speed up Coded Distributed Computing in Heterogeneous Clusters. IEEE Transactions on Services Computing, 2022, , 1-16.	3.2	0
11844	CoK: A Survey of Privacy Challenges in Relation to Data Meshes. Lecture Notes in Computer Science, 2022, , 85-102.	1.0	2
11845	An Optimized Straggler Mitigation Framework for Large-Scale Distributed Computing Systems. IEEE Access, 2022, 10, 97075-97088.	2.6	1
11846	A Comprehensive Survey on the Process, Methods, Evaluation, and Challenges of Feature Selection. IEEE Access, 2022, 10, 99595-99632.	2.6	9
11847	Cyber Threats and National Security: The Use and Abuse of Artificial Intelligence. , 2022, , 679-700.		1
11848	Big Spatial Data Systems - A review. , 2022, , .		0
11849	Survey of Breast Cancer Detection Using Machine Learning Techniques in Big Data. , 2022, , 371-385.		0
11850	A survey on cloud computing scheduling algorithms. Multiagent and Grid Systems, 2022, 18, 119-148.	0.5	2
11851	YARN Schedulers for Hadoop MapReduce Jobs: Design Goals, Issues and Taxonomy. Recent Advances in Computer Science and Communications, 2022, 16, .	0.5	0
11852	Improving Aerospace Big Data Infrastructure and Applications with Distributed File System and Massive Parallel Calculation. , 2022, , .		0
11853	Towards practical multikernel Oses with MySyS. , 2022, , .		1
11854	CDI-E. Proceedings of the VLDB Endowment, 2022, 15, 3319-3331.	2.1	0
11855	Sherlock: an open-source data platform to store, analyze and integrate Big Data for computational biologists. F1000Research, 0, 10, 409.	0.8	0

#	ARTICLE	IF	CITATIONS
11856	SemSwap. , 2022, , .		2
11858	Carbon emission and cost of blockchain mining in a case of peer-to-peer energy trading. <i>Frontiers in Built Environment</i> , 0, 8, .	1.2	1
11859	A survey on knowledge-aware news recommender systems. <i>Semantic Web</i> , 2024, 15, 21-82.	1.1	4
11860	Forkâ€“join and redundancy systems with heavy-tailed job sizes. <i>Queueing Systems</i> , 2023, 103, 131-159.	0.6	2
11861	Power to the Learner: Towards Human-Intuitive and Integrative Recommendations with Open Educational Resources. <i>Sustainability</i> , 2022, 14, 11682.	1.6	5
11862	A systematic review of data-driven approaches to fault diagnosis and early warning. <i>Journal of Intelligent Manufacturing</i> , 2023, 34, 3277-3304.	4.4	24
11863	Augmentation of literature review of COVID-19 radiology. <i>World Journal of Radiology</i> , 0, 14, 342-351.	0.5	0
11864	Cloud based evaluation of databases for stock market data. <i>Journal of Cloud Computing: Advances, Systems and Applications</i> , 2022, 11, .	2.1	0
11865	Performance Evaluation of Query Plan Recommendation with Apache Hadoop and Apache Spark. <i>Mathematics</i> , 2022, 10, 3517.	1.1	1
11866	Hyperfast second-order local solvers for efficient statistically preconditioned distributed optimization. <i>EURO Journal on Computational Optimization</i> , 2022, 10, 100045.	1.5	2
11867	QoS-Aware Co-Scheduling for Distributed Long-Running Applications on Shared Clusters. <i>IEEE Transactions on Parallel and Distributed Systems</i> , 2022, 33, 4818-4834.	4.0	9
11868	Visual-based Deep Learning for Clothing from Large Database. , 2015, , .		6
11869	Extreme k-Center Clustering. <i>Proceedings of the AAAI Conference on Artificial Intelligence</i> , 2021, 35, 3941-3949.	3.6	2
11870	Designing a Hybrid AI Residency. <i>Proceedings of the AAAI Conference on Artificial Intelligence</i> , 2021, 35, 15640-15646.	3.6	0
11871	Application of Association Rule Mining in Preventing Cyberattacks. , 2021, 67, 25-41.		0
11872	FaST. , 2022, , .		3
11873	Evaluating Deep Packet Inspection in Large-scale Data Processing. , 2022, , .		0
11876	On Generalizing Divide and Conquer Parallel Programming Pattern. <i>Mathematics</i> , 2022, 10, 3925.	1.1	0

#	ARTICLE	IF	CITATIONS
11877	PARROT: pattern-based correlation exploitation in big partitioned data series. VLDB Journal, 0, , .	2.7	0
11878	Modeling Different Deployment Variants of a Composite Application in a Single Declarative Deployment Model. Algorithms, 2022, 15, 382.	1.2	1
11880	Optimized scheduling of multi-user MapReduce jobs in heterogeneous environment. Concurrency Computation Practice and Experience, 2022, 34, .	1.4	1
11881	Cloud-native alternating directions solver for isogeometric analysis. Future Generation Computer Systems, 2022, , .	4.9	0
11882	Research on Diverse Feature Fusion Network Based on Video Action Recognition. Lecture Notes in Networks and Systems, 2023, , 526-537.	0.5	0
11883	Evaluation of a scientific data search infrastructure. Concurrency Computation Practice and Experience, 2022, 34, .	1.4	1
11884	Resampling-Free Bootstrap Inference for Quantiles. Lecture Notes in Networks and Systems, 2023, , 548-562.	0.5	0
11885	ForkMV: Mean-and-Variance Estimation of Fork-Join Queuing Networks for Datacenter Applications. , 2022, , .		2
11886	Artificial intelligence-based approaches for traditional fermented alcoholic beverages™ development: review and prospect. Critical Reviews in Food Science and Nutrition, 0, , 1-11.	5.4	2
11887	Parallel Computation of Sovereign Default Models. Computational Economics, 0, , .	1.5	0
11888	The essence of online data processing. , 2022, 6, 899-928.		0
11889	A Provenance-based Execution Strategy for Variant GPU-accelerated Scientific Workflows in Clouds. Journal of Grid Computing, 2022, 20, .	2.5	0
11890	Scalability and Performance of LiDAR Point Cloud Data Management Systems: A State-of-the-Art Review. Remote Sensing, 2022, 14, 5277.	1.8	2
11891	EAST discharge prediction without integrating simulation results. Nuclear Fusion, 2022, 62, 126060.	1.6	6
11892	OpenHardwareVC. , 2022, , .		2
11893	High-Performance Statistical Computing in the Computing Environments of the 2020s. Statistical Science, 2022, 37, .	1.6	4
11894	Scalability Challenges in Web Search Engines. Synthesis Lectures on Information Concepts, Retrieval, and Services, 2016, , .	0.6	9
11895	ARTigo: Building an Artwork Search Engine With Games and Higher-Order Latent Semantic Analysis. , 0, 1, 15-20.		6

#	ARTICLE	IF	CITATIONS
11896	Initial Results on Generating Macro Actions from a Plan Database for Planning on Autonomous Mobile Robots. , 0, 27, 498-503.		12
11897	Scalable Architectures for Big Data Analysis. , 2022, , 1-9.		0
11898	Semantic Mining of Social Networks. Synthesis Lectures on Data, Semantics and Knowledge, 2015, , .	3.9	1
11899	Scalable Processing of Spatial-Keyword Queries. Synthesis Lectures on Data Management, 2019, , .	0.6	1
11900	RibsNet: A Scalable, High-Performance, and Cost-Effective Two-Layer Based Cloud Data Center Network Architecture. IEEE Transactions on Network and Service Management, 2022, , 1-1.	3.2	1
11901	Multi-Tenancy- and Redundancy-Aware In-Network Aggregation using Programmable Switches. IEEE Network, 2022, , 1-8.	4.9	1
11902	Elastic Resource Allocation for Coded Distributed Computing Over Heterogeneous Wireless Edge Networks. IEEE Transactions on Wireless Communications, 2023, 22, 2636-2649.	6.1	2
11903	Graph Computing Systems and Partitioning Techniques: A Survey. IEEE Access, 2022, 10, 118523-118550.	2.6	9
11904	A High-Performantal and Server-Centric Based Data Center Network. IEEE Transactions on Network Science and Engineering, 2023, 10, 592-605.	4.1	2
11905	Load Balancing in Compute Clusters With Delayed Feedback. IEEE Transactions on Computers, 2023, 72, 1610-1622.	2.4	0
11906	JHTD: An Efficient Joint Scheduling Framework Based on Hypergraph for Task Placement and Data Transfer Across Geographically Distributed Data Centers. IEEE Access, 2022, 10, 116302-116316.	2.6	0
11907	Security Issues and Challenges in Vehicular Big Data Network (VBDN): A Survey. Communications in Computer and Information Science, 2022, , 26-35.	0.4	0
11908	A Utility-Based Distributed Pattern Mining Algorithm With Reduced Shuffle Overhead. IEEE Transactions on Parallel and Distributed Systems, 2023, 34, 416-428.	4.0	3
11909	Visualization Analysis on the Hotspots and Frontier Evolution of International Big Data Research. Journal of Engineering Studies, 2014, 06, 282-293.	0.0	0
11910	Cioppino: Multi-Tenant Crowd Management. , 0, 5, 41-50.		0
11911	Fast and Scalable Memristive In-Memory Sorting with Column-Skipping Algorithm. , 2022, , .		1
11912	Novel Abstraction and Offload Mechanisms for High Performance Cloud-native Distributed Object Stores. , 2022, , .		0
11913	Deep learning-Based Real-time malicious network traffic detection system for Cyber-Physical Systems. , 2022, , .		2

#	ARTICLE	IF	CITATIONS
11914	Tuning Target Delay for RTT-based Congestion Control. , 2022, , .		2
11915	Transwarp Data Cloud: A Stable, Efficient, and Intelligent Data Application Cloud Platform. , 2022, , .		0
11916	ESCHER. , 2022, , .		2
11917	Data-Efficient Performance Modeling for Configurable Big Data Frameworks by Reducing Information Overlap Between Training Examples. Big Data Research, 2022, 30, 100358.	2.6	1
11918	See it to believe it?. , 2022, , .		1
11919	Detecting Missing-Permission-Check Vulnerabilities in Distributed Cloud Systems. , 2022, , .		1
11920	Improving ML-based information retrieval software with user-driven functional testing and defect class analysis. , 2022, , .		0
11921	Flexible, highly scalable and cost-effective network structures for data centers. Journal of Network and Computer Applications, 2023, 210, 103542.	5.8	2
11922	Efficient Sorting, Duplicate Removal, Grouping, and Aggregation. ACM Transactions on Database Systems, 2022, 47, 1-35.	1.5	3
11923	Efficient flow scheduling in distributed deep learning training with echelon formation. , 2022, , .		0
11924	Automating network heuristic design and analysis. , 2022, , .		1
11925	Robust and Efficient Sorting with Offset-value Coding. ACM Transactions on Database Systems, 2023, 48, 1-23.	1.5	1
11926	Key Technologies and Practice of Big Data Platform for Scientific Research. Journal of Engineering Studies, 2014, 06, 266-274.	0.0	0
11927	A Brief Analysis of Challenges and Opportunities of Space Remote Sensing System in the Big Data Age. Journal of Engineering Studies, 2014, 06, 244-258.	0.0	1
11928	Inâ€‘memory Data Analytics. Synthesis Lectures on Emerging Engineering Technologies, 2017, , 111-142.	0.2	0
11929	Research on Reverse Skyline Query Algorithm Based on Decision Set. Journal of Database Management, 2022, 33, 1-28.	1.0	0
11930	Comprehensive techniques for multi-tenant deep learning framework on a Hadoop YARN cluster. Cluster Computing, 0, , .	3.5	0
11931	<scp>Oasis</scp> : Controlling Data Migration in Expansion of Object-based Storage Systems. ACM Transactions on Storage, 2023, 19, 1-22.	1.4	0

#	ARTICLE	IF	CITATIONS
11932	Accelerating automation of digital health applications via cloud native approach. , 2022, , .		1
11933	Mining High Utility Time Interval Sequences Using MapReduce Approach: Multiple Utility Framework. IEEE Access, 2022, 10, 123301-123315.	2.6	1
11934	Overview of Big Data Analytics Technologies in Smart Grid. International Journal of Recent Technology and Engineering, 2023, 11, 36-45.	0.2	0
11935	Spatial-aware data partition for distributed memory parallelization of ANN search in multimedia retrieval. Parallel Computing, 2023, 115, 102992.	1.3	2
11936	mpi4py.futures: MPI-Based Asynchronous Task Execution for Python. IEEE Transactions on Parallel and Distributed Systems, 2023, 34, 611-622.	4.0	5
11937	Blockchain-escorted distributed deep learning with collaborative model aggregation towards 6G networks. Future Generation Computer Systems, 2023, 141, 555-566.	4.9	5
11938	Cost-Sensitive Feature Extraction and Selection in Genre Classification. , 2009, 24, 57-72.		2
11939	A Robust and Fast Data Management System for Machine-Learning Research of Tokamaks. IEEE Transactions on Plasma Science, 2022, 50, 4980-4986.	0.6	1
11940	Exploring Memory Access Similarity to Improve Irregular Application Performance for Distributed Hybrid Memory Systems. IEEE Transactions on Parallel and Distributed Systems, 2023, 34, 797-809.	4.0	0
11941	Training and Serving Machine Learning Models at Scale. Lecture Notes in Computer Science, 2022, , 669-683.	1.0	0
11942	Research on the Integration of Market Supervision Big Data from the Perspective of Life Cycle – Take the Jiangsu Provincial Market Supervision Bureau as an Example. , 2022, , 113-126.		0
11943	Query Processing: kNN. , 2022, , 1-6.		0
11944	A Scalable Analytical Framework for Complex Event Episode Mining With Various Domains Applications. IEEE Access, 2022, 10, 130672-130685.	2.6	0
11945	Three-Way Decisions on Streaming Computing Platforms Supporting Decision-Making in Complex Large Real-World Environments. IEEE Access, 2022, 10, 122314-122328.	2.6	1
11946	MAPMAP.JS: A DATA-DRIVEN WEB MAPPING API FOR THEMATIC CARTOGRAPHY. Revista Brasileira De Cartografia, 2015, 67, .	0.1	3
11947	ENHANCING APPROACH USING HYBRID PAILLER AND RSA FOR INFORMATION SECURITY IN BIGDATA. , 2019, 15, 63-74.		6
11948	A DISTRIBUTED ALGORITHM FOR PROTEIN IDENTIFICATION FROM TANDEM MASS SPECTROMETRY DATA. , 2022, 18, 16-27.		0
11949	Aggregate Drone Monitoring of Wildfires. , 2022, , .		0

#	ARTICLE	IF	CITATIONS
11968	Autonomous learning for fuzzy systems: a review. Artificial Intelligence Review, 2023, 56, 7549-7595.	9.7	6
11969	ISort: SSD Internal Sorting Algorithm for Big Data. Mobile Information Systems, 2022, 2022, 1-9.	0.4	0
11970	Deciphering clinical abbreviations with a privacy protecting machine learning system. Nature Communications, 2022, 13, .	5.8	6
11971	Analysis and Prediction of Severity of United States Countrywide Car Accidents Based on Machine Learning Techniques. , 2022, , .		1
11973	Data-Driven Task Offloading Method for Resource-Constrained Terminals via Unified Resource Model. IEEE Internet of Things Journal, 2023, 10, 9703-9715.	5.5	3
11974	Survey on Exact kNN Queries over High-Dimensional Data Space. Sensors, 2023, 23, 629.	2.1	15
11975	Workflow Systems for Big Data Analysis. , 2022, , 1-7.		0
11976	Programming a Multicore Architecture without Coherency and Atomic Operations. , 2014, , .		0
11977	Efficient parallel viterbi algorithm for big data in a spark cloud computing environment. Procedia Computer Science, 2022, 215, 937-946.	1.2	0
11978	Selected Aspects of Interactive Feature Extraction. Lecture Notes in Computer Science, 2022, , 121-287.	1.0	3
11979	Image-Based Structural Health Monitoring: A Systematic Review. Applied Sciences (Switzerland), 2023, 13, 968.	1.3	8
11980	Detecting Anomalous Multivariate Time-Series via Hybrid Machine Learning. IEEE Transactions on Instrumentation and Measurement, 2023, 72, 1-11.	2.4	5
11981	Sherlock: an open-source data platform to store, analyze and integrate Big Data for computational biologists. F1000Research, 0, 10, 409.	0.8	0
11982	Dynamic Straggler Mitigation for Large-Scale Spatial Simulations. ACM Transactions on Spatial Algorithms and Systems, 2023, 9, 1-34.	1.1	1
11983	Data Analysis-Oriented Stochastic Scheduling for Cost Efficient Resource Allocation in NFV Based MEC Network. IEEE Transactions on Vehicular Technology, 2023, 72, 6695-6708.	3.9	2
11984	Evolution of Hadoop and Big Data Trends in Smart World. , 2023, , 99-127.		1
11985	MSstatsShiny: A GUI for Versatile, Scalable, and Reproducible Statistical Analyses of Quantitative Proteomic Experiments. Journal of Proteome Research, 0, , .	1.8	2
11986	Distributed Convex Relaxation for Heterogeneous Task Replication in Mobile Edge Computing. IEEE Transactions on Mobile Computing, 2023, , 1-16.	3.9	2

#	ARTICLE	IF	CITATIONS
11987	A Broker-Based Task-Scheduling Mechanism Using Replication Approach for Cloud Systems. Intelligent Automation and Soft Computing, 2023, 36, 2217-2232.	1.6	1
11988	Advanced Pattern-Mining System for Fake News Analysis. IEEE Transactions on Computational Social Systems, 2023, 10, 2949-2958.	3.2	1
11989	Skyline recomputation in Big Data. Information Systems, 2023, 114, 102164.	2.4	2
11990	Speed-based Load Balancer for Scheduling Reduce Tasks to Process Intermediate Data of MapReduce Applications on Cloud Computing. , 2015, , .		1
11991	An Efficient Astronomical Cross-matching model Based on MapReduce Mechanism. , 2015, , .		0
11992	Parallel Instance Filtering for Malware Detection. , 2022, , .		0
11993	HSP: Hybrid Synchronous Parallelism for Fast Distributed Deep Learning. , 2022, , .		1
11994	Accelerating the Task Activation and Data Communication for Dataflow Computing. , 2022, , .		0
11995	A Comparative Study of Large Automata Distributed Processing. , 2022, , .		0
11996	Methodology to Evaluate the Performance of Hadoop MapReduce on a Hyper-V Cluster using SAN Storage. , 2022, , .		0
11997	Robust Distributed Bayesian Learning with Stragglers via Consensus Monte Carlo. , 2022, , .		0
11998	Bias Protected Attributes Data Balancing using Map Reduce. , 2022, , .		1
11999	MR-KClust: An efficient Map Reduce based clustering Technique. , 2022, , .		0
12000	Reinforcement Learning based Scheduling for Spark Jobs in Cloud Environment. , 2022, , .		0
12001	Optimal Job Scheduling and Bandwidth Augmentation in Hybrid Data Center Networks. , 2022, , .		1
12002	Data Locality in High Performance Computing, Big Data, and Converged Systems: An Analysis of the Cutting Edge and a Future System Architecture. Electronics (Switzerland), 2023, 12, 53.	1.8	3
12003	MARbLE: Multi-Agent Reinforcement Learning at the Edge for Digital Agriculture. , 2022, , .		4
12004	A Distributed Evolutionary Framework for Large-scale SNP-SNP Interaction Detection. , 2022, , .		0

#	ARTICLE	IF	CITATIONS
12005	Taming Hybrid-Cloud Fast and Scalable Graph Analytics at Twitter. , 2022, , .		1
12018	Continuous Distributed Processing of Software Defined Radar. , 2021, , .		0
12019	The Lannion report on Big Data and Security Monitoring Research. , 2022, , .		0
12020	Performance comparison of Dask and Apache Spark on HPC systems for neuroimaging. Concurrency Computation Practice and Experience, 2023, 35, .	1.4	1
12021	S3QLRDF: distributed SPARQL query processing using Apache Sparkâ€™a comparative performance study. Distributed and Parallel Databases, 2023, 41, 191-231.	1.0	2
12022	Managing the supply chain. , 2023, , 101-152.		3
12023	ShakeFlow: Functional Hardware Description with Latency-Insensitive Interface Combinators. , 2023, , .		0
12024	Big Data Integration for Industry 4.0. , 2023, , 247-268.		0
12025	System Log Parsing: A Survey. IEEE Transactions on Knowledge and Data Engineering, 2023, , 1-20.	4.0	6
12027	Information Spaces and Efficient Information Accumulation in Calibration Problems. Lecture Notes on Data Engineering and Communications Technologies, 2023, , 53-62.	0.5	0
12028	A Generic Service to Provide In-Network Aggregation for Key-Value Streams. , 2023, , .		3
12029	Survey of Distributed Computing Frameworks for Supporting Big Data Analysis. Big Data Mining and Analytics, 2023, 6, 154-169.	7.5	7
12030	FeatureBox: Feature Engineering on GPUs for Massive-Scale Ads Systems. , 2022, , .		2
12031	Comparison of Meta-Heuristic Algorithms for Task Scheduling in Distributed Stream Processing. , 2022, , .		0
12032	Events as a Basis for Workflow Scheduling. , 2022, , .		1
12033	A parallel compression framework for fractal images using DCT block classification. , 2022, , .		0
12034	Performance Impact of NVMe-Over-TCP on HDFS Workloads. , 2022, , .		0
12035	Dynamic Replication Policy on HDFS Based on Machine Learning Clustering. IEEE Access, 2023, 11, 18551-18559.	2.6	3

#	ARTICLE	IF	CITATIONS
12036	Big data processing with 1D-Crosspoint Arrays. International Journal of Parallel, Emergent and Distributed Systems, 0, , 1-26.	0.7	0
12037	Mining Skyline Patterns from Big Data Environments based on a Spark Framework. Journal of Grid Computing, 2023, 21, .	2.5	1
12038	Efficient query evaluation techniques over large amount of distributed linked data. Information Systems, 2023, 115, 102194.	2.4	2
12039	Efficient deterministic MapReduce algorithms for parallelizable problems. Journal of Parallel and Distributed Computing, 2023, 177, 28-38.	2.7	0
12040	Scalable Text Index Construction. Lecture Notes in Computer Science, 2022, , 252-284.	1.0	0
12041	Speedy Management of Data Using MapReduce Approach. , 2022, , 286-297.		1
12042	Automatic Decomposition of a Sequential Algorithm for MapReduce Frameworks. , 2022, , .		0
12043	Hybrid Cloud and HPC Approach to High-Performance Dataframes. , 2022, , .		0
12044	Partitionable Programs Using Tyro V2. Communications in Computer and Information Science, 2022, , 185-199.	0.4	0
12045	Credit Card Fraud Detection using Big Data Technologies. International Journal of Advanced Research in Science, Communication and Technology, 0, , 783-786.	0.0	0
12046	BSP Programming Model. , 2022, , 1-5.		0
12047	DenKv: Addressing Design Trade-offs of Key-value Stores for Scientific Applications. , 2022, , .		0
12048	Apache Flink. , 2022, , 1-9.		7
12049	Challenges and Triumphs Teaching Distributed Computing Topics at a Small Liberal Arts College. , 2022, , .		0
12050	A Distributed Framework for Predictive Analytics Using Big Data and MapReduce Parallel Programming. Mathematical Problems in Engineering, 2023, 2023, 1-10.	0.6	1
12051	Real-Time Simulation of Ground Vehicles on Deformable Terrain. Journal of Computational and Nonlinear Dynamics, 2023, 18, .	0.7	1
12052	Scheduling distributed multiway spatial join queries: optimization models and algorithms. International Journal of Geographical Information Science, 0, , 1-32.	2.2	0
12053	Skyline Pattern Mining by Quantity-Utility Constraints in Large-Scale Databases. , 2022, , .		0

#	ARTICLE	IF	CITATIONS
12054	Erasure Codes for Cold Data in Distributed Storage Systems. Applied Sciences (Switzerland), 2023, 13, 2170.	1.3	0
12055	Leveraging State-of-the-Art Engines for Large-Scale Data Analysis in High Energy Physics. Journal of Grid Computing, 2023, 21, .	2.5	2
12056	Research Hotspots and Frontier Evolution of Foreign Big Data Management—A Visual Analysis Based on Web of Science Database. , 2023, , 237-245.		0
12057	IRONEDGE: Stream Processing Architecture for Edge Applications. Algorithms, 2023, 16, 123.	1.2	0
12058	Parallelism-Optimizing Data Placement for Faster Data-Parallel Computations. Proceedings of the VLDB Endowment, 2022, 16, 760-771.	2.1	0
12059	Machine Learning Empowered Intelligent Data Center Networking. SpringerBriefs in Computer Science, 2023, , 15-99.	0.2	1
12061	A scalable and real-time system for disease prediction using big data processing. Multimedia Tools and Applications, 2023, 82, 30405-30434.	2.6	2
12062	A Scalable Distributed Query Framework for Unstructured Big Clinical Data: A Case Study on Diabetic Records. Lecture Notes in Networks and Systems, 2023, , 691-699.	0.5	0
12063	On the big data processing algorithms for finding frequent sequences. Concurrency Computation Practice and Experience, 2023, 35, .	1.4	2
12064	End-to-end Delay Analysis for Coded Computing with Multi-Task Arrival. , 2022, , .		0
12065	Reproducible and Portable Big Data Analytics in the Cloud. IEEE Transactions on Cloud Computing, 2023, , 1-17.	3.1	2
12066	Canary: Fault-Tolerant FaaS for Stateful Time-Sensitive Applications. , 2022, , .		2
12067	State-of-the-Art of Artificial Intelligence and Big Data Analytics Reviews in Five Different Domains: A Bibliometric Summary. Sustainability, 2023, 15, 4026.	1.6	21
12068	Solid-state 360° optical beamforming for reconfigurable multicast optical wireless communications. Optics Express, 2023, 31, 10070.	1.7	2
12069	An investigation of distributed computing for combinatorial testing. Software Testing Verification and Reliability, 0, , .	1.7	0
12070	A Comprehensive Review on Data Partitioning and Sampling Techniques for Processing Big Data. , 2022, , .		0
12071	Network Coding Approaches for Distributed Computation over Lossy Wireless Networks. Entropy, 2023, 25, 428.	1.1	0
12072	Sparse FCM-Based Map-Reduce Framework for Distributed Parallel Data Clustering in E-Khool Learning Platform. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2023, 31, 1-23.	0.9	1

#	ARTICLE	IF	CITATIONS
12073	Affinity-aware resource provisioning for long-running applications in shared clusters. Journal of Parallel and Distributed Computing, 2023, 177, 1-16.	2.7	2
12074	A Power Grid Load Forecasting Method Based on MapReduce Improved Deep Boltzmann Machine. Lecture Notes in Electrical Engineering, 2023, , 958-964.	0.3	0
12075	Scaling Distributed Database Joins by Decoupling Computation and Communication. International Journal of Database Management Systems, 2023, 15, 1-18.	0.2	0
12076	Towards elastic in situ analysis for high-performance computing simulations. Journal of Parallel and Distributed Computing, 2023, 177, 106-116.	2.7	2
12077	ASP-Based Declarative Reasoning in Data-Intensive Enterprise and IoT Applications. Algorithms, 2023, 16, 159.	1.2	1
12078	Context-aware Task Offloading and Resource Allocation in Cloud-Edge System for Electric Data Processing. , 2022, , .		0
12079	Cloud Computing Value Chains: Research from the Operations Management Perspective. Manufacturing and Service Operations Management, 2023, 25, 1338-1356.	2.3	3
12080	Friend Recommendation System Using Map-Reduce and Spark: A Comparison Study. , 2023, , .		1
12081	Ontology Mediated Document Retrieval for Exploratory Big Data Analytics. , 2023, , .		0
12082	A Distributed Cache Mechanism of HDFS to Improve Learning Performance for Deep Reinforcement Learning. , 2022, , .		0
12083	Maximizing Bigdata Retrieval: Block as a Value for NoSQL over SQL. , 2022, , .		0
12084	Efficient Distributed Secure Memory with Migratable Merkle Tree. , 2023, , .		1
12085	IRPDP_HT2: a scalable data pre-processing method in web usage mining using Hadoop MapReduce. Soft Computing, 2023, 27, 7907-7923.	2.1	0
12086	Sparse and Outlier Robust Extreme Learning Machine Based on the Alternating Direction Method of Multipliers. Neural Processing Letters, 2023, 55, 9787-9809.	2.0	1
12087	A cache sharing mechanism based on RD MA. , 2022, , .		0
12088	Toward Federated Learning With Byzantine and Inactive Users: A Game Theory Approach. IEEE Access, 2023, 11, 34138-34149.	2.6	1
12089	The Memory-Bounded Speedup Model and Its Impacts in Computing. Journal of Computer Science and Technology, 2023, 38, 64-79.	0.9	1
12090	A Framework for Mapping DRL Algorithms With Prioritized Replay Buffer Onto Heterogeneous Platforms. IEEE Transactions on Parallel and Distributed Systems, 2023, 34, 1816-1829.	4.0	1

#	ARTICLE	IF	CITATIONS
12091	An Overview on the Challenges and Limitations Using Cloud Computing in Healthcare Corporations. Big Data and Cognitive Computing, 2023, 7, 68.	2.9	5
12092	Deep Web Search Log Mining Using Spark for Efficient and Scalable Web Information Retrieval. Advances in Intelligent Systems and Computing, 2023, , 19-31.	0.5	0
12093	DDS: integrating data analytics transformations in task-based workflows. Open Research Europe, 0, 2, 66.	2.0	0
12094	Large-scale correlation network construction for unraveling the coordination of complex biological systems. Nature Computational Science, 2023, 3, 346-359.	3.8	5
12095	An Advanced Fitness Function Optimization Algorithm for Anomaly Intrusion Detection Using Feature Selection. Applied Sciences (Switzerland), 2023, 13, 4958.	1.3	1
12096	RSCAT: Towards zero touch congestion control based on actor-critic reinforcement learning and software-defined networking. Journal of Network and Computer Applications, 2023, 215, 103639.	5.8	2
12097	Film and television art innovation in network environment by using collaborative filtering recommendation algorithm. Soft Computing, 2023, 27, 7579-7589.	2.1	0
12098	Opportunistic Monitoring of Multithreaded Programs. Lecture Notes in Computer Science, 2023, , 173-194.	1.0	3
12099	C-BEDIM and S-BEDIM: Lateral movement detection in enterprise network through behavior deviation measurement. Computers and Security, 2023, 130, 103267.	4.0	1
12103	Reinforcement Learning-Based Task Scheduling Algorithm for On-Satellite Data Analysis. Lecture Notes in Computer Science, 2023, , 398-409.	1.0	0
12109	Implementation of load balancing algorithm based on flink cluster. , 2022, , .		1
12113	Scheduling Many-Task Applications on Multi-clouds and Hybrid Clouds. Lecture Notes in Computer Science, 2023, , 65-78.	1.0	0
12115	A Unified Stream and Batch Graph Computing Model for Community Detection. Communications in Computer and Information Science, 2023, , 110-124.	0.4	0
12119	SQL Query Optimization in Distributed NoSQL Databases for Cloud-Based Applications. Lecture Notes in Computer Science, 2023, , 21-41.	1.0	0
12123	Deterministic Massively Parallel Symmetry Breaking for Sparse Graphs. , 2023, , .		0
12124	On Parallel k-Center Clustering. , 2023, , .		0
12126	Implicit Directed Acyclic Graphs (DAGs) for Parallel Outlier/Anomaly Detection Ensembles. IFIP Advances in Information and Communication Technology, 2023, , 3-15.	0.5	0
12127	High-Performance and Flexible Parallel Algorithms for Semisort and Related Problems. , 2023, , .		0

#	ARTICLE	IF	CITATIONS
12130	Hierarchical Bloom Filter. , 2023, , 79-86.		0
12132	Distributed Training. , 2023, , 95-101.		0
12134	Towards Accelerating Data Intensive Application's Shuffle Process Using SmartNICs. , 2023, , .		0
12136	Application of Time Series Database for IoT Smart City Platform. , 2023, , .		0
12144	CMLCompiler: A Unified Compiler for Classical Machine Learning. , 2023, , .		0
12146	Rational Social Edge Computing. , 2023, , 29-48.		0
12149	Basic theory and technology of reasoning engine. , 2023, , 113-127.		0
12150	System Architecture. , 2023, , 159-191.		0
12159	AI/ML Data Pipelines for Edge-Cloud Architectures. , 2023, , 159-181.		0
12164	Smart Urban Metabolism: A Big-Data and Machine Learning Perspective. , 2023, , 325-344.		3
12165	Stock Market Predictions Based on Time Series Prophet. , 2022, , .		0
12167	Dimensions and Hadoop of Big Data. A Review. , 2023, , 323-333.		0
12169	Skadi: Building a Distributed Runtime for Data Systems in Disaggregated Data Centers. , 2023, , .		0
12170	Security Framework for Big Data Usage in Cloud-based e-Learning Application. Advanced Sciences and Technologies for Security Applications, 2023, , 193-242.	0.4	0
12171	Coded Distributed Computing Schemes with Fewer Files. , 2023, , .		0
12175	Multi-User Distributed Computing Via Compressed Sensing. , 2023, , .		0
12178	The Role of Wireless Body Area Networks in Smart Healthcare System in the Context of Big Data and AI. Advances in Systems Analysis, Software Engineering, and High Performance Computing Book Series, 2023, , 185-201.	0.5	0
12179	Application of Mathematical Models in Linear Algebra to the Metaverse Ecosystem. Advances in Computational Intelligence and Robotics Book Series, 2023, , 255-278.	0.4	0

#	ARTICLE	IF	CITATIONS
12180	Context Swap: Multi-PIM System Preventing Remote Memory Access for Large Embedding Model Acceleration. , 2023, , .		0
12184	A More Efficient System for Peer-to-Peer Botnet Detection. , 2023, , .		0
12185	Efficiency Assessment of MapReduce Algorithm on a Serverless Platform. , 2023, , .		0
12186	Imbalanced Word Counting Using MapReduce in Serverless Platform. , 2023, , .		0
12187	A lightweight distributed processing computer system. , 2023, , .		0
12188	Drill: Log-based Anomaly Detection for Large-scale Storage Systems Using Source Code Analysis. , 2023, , .		0
12190	D2D Network: Implementation of Blockchain Based Equitable Cognitive Resource Sharing System. , 2023, , .		0
12191	An Enhanced Machine Learning Approach to Identify Noise and Detect Relevant Structures for Predictive Modeling. , 2023, , .		0
12192	The Demand for Big Data Skills in China. Lecture Notes in Networks and Systems, 2023, , 711-727.	0.5	0
12193	Efficient Parallel Mining of High-utility Itemsets on Multicore Processors. , 2023, , .		0
12194	SChain: Scalable Concurrency over Flexible Permissioned Blockchain. , 2023, , .		0
12196	The Metaverse Data Deluge: What Can We Do About It?. , 2023, , .		2
12197	InferTurbo: A Scalable System for Boosting Full-graph Inference of Graph Neural Network over Huge Graphs. , 2023, , .		0
12198	Efficient Parallel Mining of High-utility Itemsets on Multicore Processors. , 2023, , .		0
12201	SPOAHA: Spark Program Optimizer Based on Artificial Hummingbird Algorithm. Lecture Notes in Computer Science, 2023, , 317-331.	1.0	0
12204	Towards a Flexible and Scalable Data Stream Algorithm in FCA. Lecture Notes in Computer Science, 2023, , 104-117.	1.0	0
12205	ClassLib: An Library for Solving Linear Classification Problems on Apache Spark. , 2023, , .		0
12206	Systemarchitektur. , 2023, , 179-215.		0

#	ARTICLE	IF	CITATIONS
12208	Selecting Efficient Cluster Resources for Data Analytics: When and How to Allocate for In-Memory Processing?. , 2023, , .		0
12209	Scalable RDMA Transport with Efficient Connection Sharing. , 2023, , .		0
12210	Analyzing Data Format Interoperability in API Ecosystems Using Big Data Architecture. , 2023, , .		0
12212	Exoshuffle: An Extensible Shuffle Architecture. , 2023, , .		0
12214	SparkACO: A Spark-Based Ant Colony Optimization for Path Planning of Large-Scale Complex Road Networks. , 2023, , .		0
12217	Computational methods for collisional stellar systems. Living Reviews in Solar Physics, 2023, 9, .	5.0	2
12219	Artificial Intelligence in Marketing and Organizational Decision-Making: Some Challenges and Concerns. Internet of Things, 2023, , 9-23.	1.3	0
12220	A Generic Efficient Scientific Workflow Engine for the Optimizations of Run-Time Execution. , 2023, , .		0
12224	Effective Management of Time Series Data. , 2023, , .		0
12225	Detection of Faulty node with Hybrid Machine Learning using SVM model. , 2023, , .		0
12226	SIM: Secure Interval Membership Testing and Applications to Secure Comparison. , 2023, , .		0
12228	Invited Paper: On the Cost-Optimal Parallel Solution of the Majority Problem. , 2023, , .		0
12229	On the Implications of Heterogeneous Memory Tiering on Spark In-Memory Analytics. , 2023, , .		0
12230	Exploit Approximation to Support Fault Resiliency in MPI-based Applications. , 2023, , .		0
12231	A Novel Spark-Based Algorithm for Mining Frequent Utility Patterns. , 2023, , .		0
12233	Fundamental Limits of Distributed Linearly Separable Computation under Cyclic Assignment. , 2023, , .		1
12235	Distributed k-Means with Outliers in General Metrics. Lecture Notes in Computer Science, 2023, , 474-488.	1.0	0
12237	Ten Key Problems of Pre-trained Models: An Outlook of Representation Learning. , 2023, , 491-521.		0

#	ARTICLE	IF	CITATIONS
12239	A Study on Distributed Machine Learning Techniques for Large-Scale Weather Forecasting. Advances in Computational Intelligence and Robotics Book Series, 2023, , 44-64.	0.4	0
12241	Lightwave Fabrics: At-Scale Optical Circuit Switching for Datacenter and Machine Learning Systems. , 2023, , .		1
12242	In Support of Push-Based Streaming for the Computing Continuum. Lecture Notes in Computer Science, 2023, , 339-350.	1.0	0
12243	Data Engineering Challenges in AI automation. , 2023, , .		0
12244	Domain-Oriented Transformation Method for Big Data Analysis Process Model. , 2023, , .		0
12246	Toward Fair and Efficient Congestion Control: Machine Learning Aided Congestion Control (MLACC). , 2023, , .		0
12247	Domain-Specific Framework for Performance Analysis. , 2023, , 227-254.		0
12249	Multithreaded Algorithms for Lossless Intra Compression of Point Cloud Geometry Based on the Silhouette 3d Coder. , 2023, , .		0
12251	DAG-Aware Optimization for Geo-Distributed Data Analytics. , 2023, , .		0
12252	Dynamic Data-Driven Application Systems for Reservoir Simulation-Based Optimization: Lessons Learned and Future Trends. , 2023, , 287-330.		0
12253	To err is AI. , 2023, , .		0
12255	Stragglers in Distributed Matrix Multiplication. Lecture Notes in Computer Science, 2023, , 74-96.	1.0	0
12259	Exploratory Data Analysis of Walmart Outlets Sales using Data Analytics Techniques. , 2023, , .		1
12261	Enabling Highly-Efficient DNA Sequence Mapping via ReRAM-based TCAM. , 2023, , .		0
12263	Association Rule Mining in Distributed Environment: A Survey. Lecture Notes in Networks and Systems, 2023, , 113-120.	0.5	0
12266	NVMe-Driven Lazy Cache Coherence for Immutable Data with NVMe over Fabrics. , 2023, , .		0
12267	The Case for the Anonymization of Offloaded Computation. , 2023, , .		0
12268	Contract-Driven Design of Scientific Data Analysis Workflows. , 2023, , .		0

#	ARTICLE	IF	CITATIONS
12269	Making Your Program Oblivious: A Comparative Study for Side-channel-Safe Confidential Computing. , 2023, , .		1
12272	Improve the Daily Societal Operations Using Credit Fraud Detection: A Big Data Classification Solution. , 2023, , 97-110.		0
12275	Comparison of Data Processing Performance of Hadoop and Spark Based on Huawei Cloud and Serverless. Lecture Notes in Electrical Engineering, 2023, , 195-204.	0.3	0
12277	A Distributed Alignment-free Pipeline for Human SNPs Genotyping. , 2023, , .		0
12279	Two-Stage Coded Distributed Learning: A Dynamic Partial Gradient Coding Perspective. , 2023, , .		0
12284	Unified Shared Memory: Friend or Foe? Understanding the Implications of Unified Memory on Managed Heaps. , 2023, , .		0
12285	G-Meta: Distributed Meta Learning in GPU Clusters for Large-Scale Recommender Systems. , 2023, , .		0
12287	Coded Reactive Stragglers Mitigation in Distributed Computing Systems. , 2023, , .		0
12288	Data Volume-Aware Computation Task Scheduling for Smart Grid Data Analytic Applications. , 2023, , .		0
12294	Phases, Modalities, Spatial and Temporal Locality: Domain Specific ML Prefetcher for Accelerating Graph Analytics. , 2023, , .		0
12295	EvLog: Identifying Anomalous Logs over Software Evolution. , 2023, , .		0
12296	SQCC: Stable Queue Congestion Control. , 2023, , .		0
12297	The Role of Machine Learning in Big Data Analytics: Current Practices and Challenges. Transactions on Computational Science and Computational Intelligence, 2024, , 47-74.	0.3	0
12299	Scalable Computation of Fuzzy Joins Over Large Collections of JSON Data. , 2023, , .		0
12301	Peachy Parallel Assignments (EduHPC 2023). , 2023, , .		0
12302	Distributed Data Locality-Aware Job Allocation. , 2023, , .		1
12303	Towards A Massive-scale Distributed Neighborhood Graph Construction. , 2023, , .		0
12304	Mars Attacks!. , 2023, , .		0

#	ARTICLE	IF	CITATIONS
12305	Scalable Centroid Based Fuzzy Min-Max Neural Network Ensemble Classifier Using MapReduce. Lecture Notes in Computer Science, 2023, , 597-607.	1.0	0
12306	Straggler Exploitation in Distributed Computing Systems with Task Grouping. , 2023, , .		0
12307	Design and Implementation of External Storage Large-Scale Graph Computing System. , 2023, , .		0
12308	A Novel Distributed K-Means Clustering Algorithm for Big Text Data. , 2023, , .		0
12310	Clustering big data for novel health care system. AIP Conference Proceedings, 2023, , .	0.3	0
12311	Accelerating I/O in Distributed Data Processing Systems with Apache Arrow CHFS. , 2023, , .		0
12312	Mobile Cloud Offloading: An Analysis of Methodologies, Limitations, and Research Problems. , 2023, , .		0
12315	A Big Data Analytics Architecture Framework for Oilseeds and Textile Industry Production and International Trade for Sub-Saharan Africa (SSA). , 2023, , 106-139.		0
12316	Dynamic Priority Coflow Scheduling in Optical Circuit Switched Networks. Lecture Notes in Electrical Engineering, 2024, , 215-226.	0.3	0
12317	Analysis of Arabs Publications In the Field of Computer Networks. , 2023, , .		0
12321	Learning to Drive Software-Defined Solid-State Drives. , 2023, , .		0
12323	The Hardness of Optimization Problems on the Weighted Massively Parallel Computation Model. Lecture Notes in Computer Science, 2024, , 106-117.	1.0	1
12324	Scalable Actor Networks with CAL. , 2023, , .		0
12326	Heterogeneous Secure Coded Matrix Multiplication: Straggler Problem versus Information Leakage. , 2023, , .		0
12328	Optimizing Parallelization Strategies for the Big-Means Clustering Algorithm. Communications in Computer and Information Science, 2024, , 17-32.	0.4	0
12331	Detecting Violations of Conditional Functional Dependencies in Distributed Database. , 2023, , .		0
12332	Research on Imbalanced Data Ensemble Algorithm Based on Density Peaks Clustering and Fitness. , 2023, , .		0
12333	We are on the Same Side. Alternative Sieving Strategies for the Number Field Sieve. Lecture Notes in Computer Science, 2023, , 138-166.	1.0	0

#	ARTICLE	IF	CITATIONS
12335	A City-level High-performance Spatio-temporal Mobility Simulation System. , 2023, , .		0
12340	SimplePIM: A Software Framework for Productive and Efficient Processing-in-Memory. , 2023, , .		0
12347	Understanding and Optimizing Workloads for Unified Resource Management in Large Cloud Platforms. , 2023, , .		2
12354	UCAPF: A Unified Processing Platform for Large-scale Virtual Screening. , 2023, , .		0
12356	Big Data analysis. , 2024, , 259-276.		0
12357	ChimeraPy: A Scientific Distributed Streaming Framework for Real-time Multimodal Data Retrieval and Processing. , 2023, , .		0
12358	Lazy Read: Asynchronous Execution of Synchronous File I/O. , 2023, , .		0
12359	Renting Servers in the Cloud: Parameterized Analysis of FirstFit. , 2024, , .		0
12360	TensorBank: Tensor Lakehouse for Foundation Model Training. , 2023, , .		0
12362	Minimap2 parallelization method based on distributed computing. , 2023, , .		0
12363	Preliminary Exploration on Node-To-Node Fault Tolerance Coordination in Distributed System. , 2023, , .		0
12364	Secure Data Sharing and Mining in Distributed Environments. , 2023, , .		0
12366	Bridging Between Active Objects: Multitier Programming for Distributed, Concurrent Systems. Lecture Notes in Computer Science, 2024, , 92-122.	1.0	0
12367	A Survey of Actor-Like Programming Models for Serverless Computing. Lecture Notes in Computer Science, 2024, , 123-146.	1.0	0
12368	High Availabale Hadoop Deployment Modes with Enterprise-Level capabilities. , 2023, , .		0
12370	FS: A Scalable and Elastic Distributed File System Metadata Service using Serverless Functions. , 2023, , .		0
12372	PRISPARK: Differential Privacy Enforcement for Big Data Computing in Apache Spark. , 2023, , .		0
12373	Map-Reduce (Hadoop) based Data Clustering for BigData: A Survey. , 2023, , .		0

#	ARTICLE	IF	CITATIONS
12377	Efficient In-Memory Point Cloud Query Processing. Lecture Notes in Geoinformation and Cartography, 2024, , 267-286.	0.5	0
12380	Data-Driven Pathways to Sustainable Energy Solutions. , 2024, , 1-31.		0
12383	Beibu Gulf Marine Ranch: Utilizing BeiDou Grid Code and Multi-system Integration for Modernized Management and Monitoring. Communications in Computer and Information Science, 2024, , 247-257.	0.4	0
12387	Fractional Bamboo Trimming and Distributed Windows Scheduling. Lecture Notes in Computer Science, 2024, , 69-81.	1.0	0
12389	High Performance Computing and Its Application in Computational Biomimetics. Series in Bioengineering, 2024, , 21-46.	0.3	0
12390	Infrastructures for High-Performance Computing: Cloud Computing. , 2024, , .		0
12391	Infrastructures for High-Performance Computing: Cloud Computing Development Environments. , 2024, , .		0
12395	Expansion of AI and ML Breakthroughs in HPC with Shift to Edge Computing in Remote Environments. Series in Bioengineering, 2024, , 129-145.	0.3	0