

Hong Jiang

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3103059/publications.pdf>

Version: 2024-02-01

137
papers

2,649
citations

394286

19
h-index

360920

35
g-index

138
all docs

138
docs citations

138
times ranked

1476
citing authors

| # | ARTICLE | IF | CITATIONS |
|----|--|------|-----------|
| 1 | Performance impact and interplay of SSD parallelism through advanced commands, allocation strategy and data granularity. , 2011, , . | | 276 |
| 2 | A Comprehensive Study of the Past, Present, and Future of Data Deduplication. Proceedings of the IEEE, 2016, 104, 1681-1710. | 16.4 | 200 |
| 3 | Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage. IEEE Transactions on Services Computing, 2017, 10, 701-714. | 3.2 | 175 |
| 4 | MAD2: A scalable high-throughput exact deduplication approach for network backup services. , 2010, , . | | 64 |
| 5 | AE: An Asymmetric Extremum content defined chunking algorithm for fast and bandwidth-efficient data deduplication. , 2015, , . | | 64 |
| 6 | HBA: Distributed Metadata Management for Large Cluster-Based Storage Systems. IEEE Transactions on Parallel and Distributed Systems, 2008, 19, 750-763. | 4.0 | 60 |
| 7 | DEBAR: A scalable high-performance de-duplication storage system for backup and archiving. , 2010, , . | | 58 |
| 8 | Ddelta: A deduplication-inspired fast delta compression approach. Performance Evaluation, 2014, 79, 258-272. | 0.9 | 58 |
| 9 | SAM: A Semantic-Aware Multi-tiered Source De-duplication Framework for Cloud Backup. , 2010, , . | | 54 |
| 10 | AA-Dedupe: An Application-Aware Source Deduplication Approach for Cloud Backup Services in the Personal Computing Environment. , 2011, , . | | 52 |
| 11 | Read-Performance Optimization for Deduplication-Based Storage Systems in the Cloud. ACM Transactions on Storage, 2014, 10, 1-22. | 1.4 | 50 |
| 12 | Application-Aware Local-Global Source Deduplication for Cloud Backup Services of Personal Storage. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 1155-1165. | 4.0 | 49 |
| 13 | Using High-Bandwidth Networks Efficiently for Fast Graph Computation. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1170-1183. | 4.0 | 47 |
| 14 | PUD-LRU: An Erase-Efficient Write Buffer Management Algorithm for Flash Memory SSD. , 2010, , . | | 42 |
| 15 | The Design of Fast Content-Defined Chunking for Data Deduplication Based Storage Systems. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 1717-1731. | 4.0 | 42 |
| 16 | Similarity and Locality Based Indexing for High Performance Data Deduplication. IEEE Transactions on Computers, 2015, 64, 1162-1176. | 2.4 | 41 |
| 17 | A Scalable Inline Cluster Deduplication Framework for Big Data Protection. Lecture Notes in Computer Science, 2012, , 354-373. | 1.0 | 39 |
| 18 | CABdedupe: A Causality-Based Deduplication Performance Booster for Cloud Backup Services. , 2011, , . | | 38 |

| # | ARTICLE | IF | CITATIONS |
|----|--|-----|-----------|
| 19 | Combining Deduplication and Delta Compression to Achieve Low-Overhead Data Reduction on Backup Datasets. , 2014, , . | | 37 |
| 20 | GCaR. , 2016, , . | | 34 |
| 21 | POD: Performance Oriented I/O Deduplication for Primary Storage Systems in the Cloud. , 2014, , . | | 31 |
| 22 | DARE: A Deduplication-Aware Resemblance Detection and Elimination Scheme for Data Reduction with Low Overheads. IEEE Transactions on Computers, 2016, 65, 1692-1705. | 2.4 | 31 |
| 23 | Supporting Scalable and Adaptive Metadata Management in Ultralarge-Scale File Systems. IEEE Transactions on Parallel and Distributed Systems, 2011, 22, 580-593. | 4.0 | 30 |
| 24 | Application-Aware Big Data Deduplication in Cloud Environment. IEEE Transactions on Cloud Computing, 2019, 7, 921-934. | 3.1 | 30 |
| 25 | An efficient fault-tolerant scheduling algorithm for real-time tasks with precedence constraints in heterogeneous systems. , 0, , . | | 29 |
| 26 | P-Dedupe: Exploiting Parallelism in Data Deduplication System. , 2012, , . | | 29 |
| 27 | Proactive Data Migration for Improved Storage Availability in Large-Scale Data Centers. IEEE Transactions on Computers, 2015, 64, 2637-2651. | 2.4 | 26 |
| 28 | GRAID: A Green RAID Storage Architecture with Improved Energy Efficiency and Reliability. , 2008, , . | | 25 |
| 29 | Semantic-Aware Metadata Organization Paradigm in Next-Generation File Systems. IEEE Transactions on Parallel and Distributed Systems, 2012, 23, 337-344. | 4.0 | 25 |
| 30 | A Fast Asymmetric Extremum Content Defined Chunking Algorithm for Data Deduplication in Backup Storage Systems. IEEE Transactions on Computers, 2016, , 1-1. | 2.4 | 25 |
| 31 | Public Auditing for Trusted Cloud Storage Services. IEEE Security and Privacy, 2019, 17, 10-22. | 1.5 | 25 |
| 32 | Accelerating content-defined-chunking based data deduplication by exploiting parallelism. Future Generation Computer Systems, 2019, 98, 406-418. | 4.9 | 25 |
| 33 | An adaptive steganography scheme for voice over IP. , 2009, , . | | 24 |
| 34 | STEM: Spatiotemporal Management of Capacity for Intra-core Last Level Caches. , 2010, , . | | 23 |
| 35 | SAR: SSD Assisted Restore Optimization for Deduplication-Based Storage Systems in the Cloud. , 2012, , . | | 23 |
| 36 | Improving Storage Availability in Cloud-of-Clouds with Hybrid Redundant Data Distribution. , 2015, , . | | 23 |

| # | ARTICLE | IF | CITATIONS |
|----|--|-----|-----------|
| 37 | A Novel Weighted-Graph-Based Grouping Algorithm for Metadata Prefetching. IEEE Transactions on Computers, 2010, 59, 1-15. | 2.4 | 22 |
| 38 | HPDA: A hybrid parity-based disk array for enhanced performance and reliability. , 2010, , . | | 21 |
| 39 | IOFollow: Improving the performance of VM live storage migration with IO following in the cloud. Future Generation Computer Systems, 2019, 91, 167-176. | 4.9 | 20 |
| 40 | FAST: Near Real-Time Searchable Data Analytics for the Cloud. , 2014, , . | | 19 |
| 41 | SnapMig: Accelerating VM Live Storage Migration by Leveraging the Existing VM Snapshots in the Cloud. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1416-1427. | 4.0 | 19 |
| 42 | EdgeDB: An Efficient Time-Series Database for Edge Computing. IEEE Access, 2019, 7, 142295-142307. | 2.6 | 19 |
| 43 | AMP: An Affinity-Based Metadata Prefetching Scheme in Large-Scale Distributed Storage Systems. , 2008, , . | | 18 |
| 44 | Communication-Aware Load Balancing for Parallel Applications on Clusters. IEEE Transactions on Computers, 2010, 59, 42-52. | 2.4 | 18 |
| 45 | Improving Availability of RAID-Structured Storage Systems by Workload Outsourcing. IEEE Transactions on Computers, 2011, 60, 64-79. | 2.4 | 18 |
| 46 | ANTELOPE: A Semantic-Aware Data Cube Scheme for Cloud Data Center Networks. IEEE Transactions on Computers, 2014, 63, 2146-2159. | 2.4 | 18 |
| 47 | A communication-reduced and computation-balanced framework for fast graph computation. Frontiers of Computer Science, 2018, 12, 887-907. | 1.6 | 18 |
| 48 | Improving Performance for Flash-Based Storage Systems through GC-Aware Cache Management. IEEE Transactions on Parallel and Distributed Systems, 2017, 28, 2852-2865. | 4.0 | 17 |
| 49 | Transparency-Orientated Encoding Strategies for Voice-over-IP Steganography. Computer Journal, 2012, 55, 702-716. | 1.5 | 16 |
| 50 | Improving Restore Performance in Deduplication-Based Backup Systems via a Fine-Grained Defragmentation Approach. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 2254-2267. | 4.0 | 16 |
| 51 | Modeling parallel applications performance on heterogeneous systems. , 0, , . | | 15 |
| 52 | Hierarchical Bloom filter arrays (HBA): a novel, scalable metadata management system for large cluster-based storage. , 0, , . | | 14 |
| 53 | Efficiently Representing Membership for Variable Large Data Sets. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 960-970. | 4.0 | 14 |
| 54 | A Fast Filtering Mechanism to Improve Efficiency of Large-Scale Video Analytics. IEEE Transactions on Computers, 2020, 69, 914-928. | 2.4 | 14 |

| # | ARTICLE | IF | CITATIONS |
|----|---|-----|-----------|
| 55 | A dynamic load balancing scheme for I/O-intensive applications in distributed systems. , 0, , . | | 13 |
| 56 | SANE: Semantic-Aware Namespace in Ultra-Large-Scale File Systems. IEEE Transactions on Parallel and Distributed Systems, 2014, 25, 1328-1338. | 4.0 | 13 |
| 57 | Nexus: a novel weighted-graph-based prefetching algorithm for metadata servers in petabyte-scale storage systems. , 2006, , . | | 11 |
| 58 | SAFE: A Source Deduplication Framework for Efficient Cloud Backup Services. Journal of Signal Processing Systems, 2013, 72, 209-228. | 1.4 | 11 |
| 59 | Underprovisioning the Grid Power Infrastructure for Green Datacenters. , 2015, , . | | 11 |
| 60 | Improving Overall Performance of TLC SSD by Exploiting Dissimilarity of Flash Pages. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 332-346. | 4.0 | 11 |
| 61 | A Black-Box Fork-Join Latency Prediction Model for Data-Intensive Applications. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 1983-2000. | 4.0 | 11 |
| 62 | Comparison of Mesh and Hierarchical Networks for Multiprocessors. , 1994, , . | | 10 |
| 63 | Improving the performance of I/O-intensive applications on clusters of workstations. Cluster Computing, 2006, 9, 297-311. | 3.5 | 10 |
| 64 | PFP: Improving the Reliability of Deduplication-based Storage Systems with Per-File Parity. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 2117-2129. | 4.0 | 10 |
| 65 | Modeling of domain pinning effect in polycrystalline ferroelectric ceramics. Ferroelectrics, 1996, 182, 61-68. | 0.3 | 9 |
| 66 | JOR: A Journal-guided Reconstruction Optimization for RAID-Structured Storage Systems. , 2009, , . | | 9 |
| 67 | Lessons Learned from Comprehensive Deployments of Multiagent CSCL Applications I-MINDS and ClassroomWiki. IEEE Transactions on Learning Technologies, 2011, 4, 47-58. | 2.2 | 9 |
| 68 | GC-ARM: Garbage Collection-Aware RAM Management for Flash Based Solid State Drives. , 2012, , . | | 9 |
| 69 | DDOps: dual-direction operations for load balancing on non-dedicated heterogeneous distributed systems. Cluster Computing, 2014, 17, 503-528. | 3.5 | 9 |
| 70 | An Improved Decoding Algorithm for Generalized RDP Codes. IEEE Communications Letters, 2016, 20, 632-635. | 2.5 | 9 |
| 71 | EDC: Improving the Performance and Space Efficiency of Flash-Based Storage Systems with Elastic Data Compression. IEEE Transactions on Parallel and Distributed Systems, 2018, 29, 1261-1274. | 4.0 | 9 |
| 72 | PA-SSD. , 2018, , . | | 9 |

| # | ARTICLE | IF | CITATIONS |
|----|--|-----|-----------|
| 73 | GC-Aware Request Steering with Improved Performance and Reliability for SSD-Based RAIDs. , 2018, , . | | 9 |
| 74 | Improving Flash Memory Performance and Reliability for Smartphones With I/O Deduplication. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2019, 38, 1017-1027. | 1.9 | 9 |
| 75 | A Hybrid Update Strategy for I/O-Efficient Out-of-Core Graph Processing. IEEE Transactions on Parallel and Distributed Systems, 2020, 31, 1767-1782. | 4.0 | 9 |
| 76 | Improved read performance in a cost-effective, fault-tolerant parallel virtual file system (CEFT-PVFS). , 2003, , . | | 8 |
| 77 | RoLo: A Rotated Logging Storage Architecture for Enterprise Data Centers. , 2010, , . | | 8 |
| 78 | DBA: A Dynamic Bloom Filter Array for Scalable Membership Representation of Variable Large Data Sets. , 2011, , . | | 8 |
| 79 | Improving Hybrid FTL by Fully Exploiting Internal SSD Parallelism with Virtual Blocks. Transactions on Architecture and Code Optimization, 2015, 11, 1-19. | 1.6 | 8 |
| 80 | A comparative study of parallel and distributed Java projects for heterogeneous systems. , 2002, , . | | 7 |
| 81 | Data Grids: Supporting Data-Intensive Applications in Wide-Area Networks. , 2006, , 481-494. | | 7 |
| 82 | Adaptive Consistency Guarantees for Large-Scale Replicated Services. , 2008, , . | | 7 |
| 83 | ForkTail. , 2018, , . | | 7 |
| 84 | A highly cost-effective task scheduling strategy for very large graph computation. Future Generation Computer Systems, 2018, 89, 698-712. | 4.9 | 7 |
| 85 | LiteTE: Lightweight, Communication-Efficient Distributed-Memory Triangle Enumerating. IEEE Access, 2019, 7, 26294-26306. | 2.6 | 7 |
| 86 | Systematic Erasure Codes with Optimal Repair Bandwidth and Storage. ACM Transactions on Storage, 2017, 13, 1-27. | 1.4 | 7 |
| 87 | Distributed systems middleware architecture from a software engineering perspective. , 0, , . | | 6 |
| 88 | False Rate Analysis of Bloom Filter Replicas in Distributed Systems. , 0, , . | | 6 |
| 89 | Implementation and Evaluation of a Popularity-Based Reconstruction Optimization Algorithm in Availability-Oriented Disk Arrays. , 2007, , . | | 6 |
| 90 | GreenGear. , 2016, , . | | 6 |

| # | ARTICLE | IF | CITATIONS |
|-----|--|-----|-----------|
| 91 | Real-Time Semantic Search Using Approximate Methodology for Large-Scale Storage Systems. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1212-1225. | 4.0 | 6 |
| 92 | Customizable SLO and Its Near-Precise Enforcement for Storage Bandwidth. ACM Transactions on Storage, 2017, 13, 1-25. | 1.4 | 6 |
| 93 | GreenSprint: Effective Computational Sprinting in Green Data Centers. , 2018, , . | | 6 |
| 94 | Performance and configuration of hierarchical ring networks for multiprocessors. , 0, , . | | 5 |
| 95 | I-MINDS: an application of multiagent system intelligence to on-line education. , 0, , . | | 5 |
| 96 | Performance and cost effectiveness of a cluster of workstations and MD-GRAPE 2 for MD simulations. , 0, , . | | 5 |
| 97 | A distributed shared object model based on a hierarchical consistency protocol for heterogeneous clusters. , 0, , . | | 5 |
| 98 | Exploiting redundancy to boost performance in a RAID-10 style cluster-based file system. Cluster Computing, 2006, 9, 433-447. | 3.5 | 5 |
| 99 | TRIP: Temporal Redundancy Integrated Performance Booster for Parity-Based RAID Storage Systems. , 2010, , . | | 5 |
| 100 | Elastic-RAID: A New Architecture for Improved Availability of Parity-Based RAIDs by Elastic Mirroring. IEEE Transactions on Parallel and Distributed Systems, 2016, 27, 1044-1056. | 4.0 | 5 |
| 101 | HUS-Graph. , 2018, , . | | 5 |
| 102 | SPA-SSD: Exploit Heterogeneity and Parallelism of 3D SLC-TLC Hybrid SSD to Improve Write Performance. , 2019, , . | | 5 |
| 103 | On some architectural issues of optical hierarchical ring networks for shared-memory multiprocessors. , 0, , . | | 4 |
| 104 | DuoModel: Leveraging Reduced Model for Data Reduction and Re-Computation on HPC Storage. IEEE Letters of the Computer Society, 2018, 1, 5-8. | 1.1 | 4 |
| 105 | PUSH THE BOTTLENECK OF STREAMING MEDIA SYSTEM FROM STREAMING MEDIA SERVER TO NETWORK. International Journal of Image and Graphics, 2005, 05, 859-869. | 1.2 | 3 |
| 106 | A framework for efficient inconsistency detection in a grid and internet-scale distributed environment. , 0, , . | | 3 |
| 107 | Adaptive Load Balancing for Long-Range MD Simulations in A Distributed Environment. , 2006, , . | | 3 |
| 108 | Detecting Duplicates over Sliding Windows with RAM-Efficient Detached Counting Bloom Filter Arrays. , 2011, , . | | 3 |

| # | ARTICLE | IF | CITATIONS |
|-----|--|-----|-----------|
| 109 | Factors affecting scalability of multithreaded Java applications on manycore systems. , 2015, , . | | 3 |
| 110 | Storage Sharing Optimization Under Constraints of SLO Compliance and Performance Variability. IEEE Transactions on Services Computing, 2019, 12, 58-72. | 3.2 | 3 |
| 111 | TriangleKV: Reducing Write Stalls and Write Amplification in LSM-Tree Based KV Stores With Triangle Container in NVM. IEEE Transactions on Parallel and Distributed Systems, 2022, 33, 4339-4352. | 4.0 | 3 |
| 112 | Statistics and analysis tools for a computer-supported collaborative learning system. Proceedings - Frontiers in Education Conference, FIE, 2007, , . | 0.0 | 2 |
| 113 | Optimal Encoding and Decoding Algorithms for the RAID-6 Liberation Codes. , 2020, , . | | 2 |
| 114 | EaD: ECC-Assisted Deduplication With High Performance and Low Memory Overhead for Ultra-Low Latency Flash Storage. IEEE Transactions on Computers, 2023, 72, 208-221. | 2.4 | 2 |
| 115 | How to Realize Efficient and Scalable Graph Embeddings via an Entropy-Driven Mechanism. IEEE Transactions on Big Data, 2023, 9, 358-371. | 4.4 | 2 |
| 116 | Performance properties of combined heterogeneous networks. , 0, , . | | 1 |
| 117 | Flexible mechanisms for performance enhancements of cluster networks. , 0, , . | | 1 |
| 118 | Improving the performance of communication-intensive parallel applications executing on clusters. , 0, , . | | 1 |
| 119 | Intelligent Collaborating Agents to Support Teaching and Learning. , 0, , . | | 1 |
| 120 | Design and evaluation of a new and effective fairness scheme for multicasting in internet-scale distributed systems. , 0, , . | | 1 |
| 121 | Adaptive Load-Balancing for Force-Decomposition Based 3-Body Molecular Dynamics Simulations in A Heterogeneous Distributed Environment with Variable Number of Processors. Parallel Processing (ICPP), Proceedings of the International Symposium, 2007, , . | 0.0 | 1 |
| 122 | Accurate Performance Modeling and Guidance to the Adoption of an Inconsistency Detection Framework. , 2008, , . | | 1 |
| 123 | 2-Hopper: Accurately Estimate Individual and Social Attributes of Social Networks With Fewer Repeats via Random Walk. IEEE Access, 2019, 7, 139827-139838. | 2.6 | 1 |
| 124 | A-Cache: Asymmetric Buffer Cache for RAID-10 Systems Under a Single-Disk Failure to Significantly Boost Availability. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41, 723-736. | 1.9 | 1 |
| 125 | DedupHR: Exploiting Content Locality to Alleviate Read/Write Interference in Deduplication-based Flash Storage. IEEE Transactions on Computers, 2021, , 1-1. | 2.4 | 1 |
| 126 | GraphCP: An I/O-Efficient Concurrent Graph Processing Framework. , 2021, , . | | 1 |

| # | ARTICLE | IF | CITATIONS |
|-----|---|-----|-----------|
| 127 | A reconfigurable optical bus structure for shared memory multiprocessors with improved performance. , 0, , . | | 1 |
| 128 | ComboTree: A Persistent Indexing Structure With Universal Operational Efficiency and Scalability. IEEE Transactions on Parallel and Distributed Systems, 2022, 33, 2277-2290. | 4.0 | 1 |
| 129 | Understanding and Exploiting the Full Potential of SSD Address Remapping. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41, 5112-5125. | 1.9 | 1 |
| 130 | Reconfigurable optical bus and performance optimization. Journal of Computer Science and Technology, 1996, 11, 296-312. | 0.9 | 0 |
| 131 | The role of partitioning in Time Warp simulation. , 0, , . | | 0 |
| 132 | Comparisons between mesh and hierarchical ring networks for shared-memory multiprocessors. Canadian Journal of Electrical and Computer Engineering, 1998, 23, 119-125. | 1.5 | 0 |
| 133 | I/O performance of an RAID-10 style parallel file system. Journal of Computer Science and Technology, 2004, 19, 965-972. | 0.9 | 0 |
| 134 | DSFS: Decentralized security for large parallel file systems. , 2010, , . | | 0 |
| 135 | SASLO: Support User-Customized SLO Policy via Programmable End-to-End VM-Oriented IO Control. , 2015, , . | | 0 |
| 136 | LCC-Graph: A high-performance graph-processing framework with low communication costs. , 2016, , . | | 0 |
| 137 | A Renewable Energy Driven Approach for Computational Sprinting. IEEE Transactions on Parallel and Distributed Systems, 2019, 30, 1449-1463. | 4.0 | 0 |