Suren Byna

Professor
The Department of Computer Science and Engineering
The Ohio State University
Columbus, OH, USA

Phone: (630) 487-1230 E-Mail: *byna.1@osu.edu* Homepage: https://sbyna.github.io

About me

Scientific discoveries depend heavily on efficient management of data life cycle. As a computer scientist, my research interests are concerned with the development of methods and algorithms for data management on high performance computing (HPC) systems to be autonomous, transparent, and efficient so that domain scientists spend more time on their science and less on managing data. My R&D technologies and algorithms in achieving this goal have been published at premier conferences and workshops. The technologies developed in my work have also been integrated into well-known I/O software libraries, such as HDF5 that are used by a vast number of scientific applications in HPC and beyond. I've mentored more than 30 students and postdoctoral scholars from diverse backgrounds in this process. I currently lead multiple projects that include researchers and engineers from various institutions including US Department of Energy (DOE) national labs, industry, and academia. I collaborate with a number of professors at different universities and leaders in the industry. I strive to achieve the goal of efficient and impactful data management and high-performance computing.

Highlights

- Publications: More than 160 conference, workshop, and journal publications
  - Citations: 3168, h-index: 32, i-10-index: 81 on 04/10/2023 - Google Scholar
- Funding as a Principal Investigator (PI):
  - Current projects: $13,224 k
  - Completed projects: $9,186 k
- Project team leadership:
  - ECP ExaIO (~28 members from 5 institutes, 2019-2023)
  - PDC (11 members from 4 institutes, 2016-2022)
  - EOD-HDF5 (9 members from 3 institutes, 2017-2022)
  - ExaHDF5 (16 members from 3 institutes, 2014-2019)
- Mentoring:
  - 6 post-doctoral scholars
  - 5 early career staff members at LBNL
  - > 30 students as summer interns over 7 years at LBNL
- ~45 invited speeches at workshops, conferences, or seminars
- Selected professional service activities:
  - Organizing committee, an upcoming DOE workshop on Storage and I/O
  - Breakout group co-lead, DOE workshop on Extreme Heterogeneity, 2018
  - Organizing member of HDF5 User Group (HUG) meetings 2020 and 2021
  - Technical Advisory Board of The HDF Group
  - IPDPS 2022 Workshops, Vice chair
  - ISC High Performance – Tutorials: vice chair in 2022 and chair 2023
  - PC vice chair or area chair for multiple CCGrid and HiPC conferences
  - Research Posters chair at SC20
  - Keynote speaker at the Workshop on Challenges and Opportunities of Efficient and Performant Storage Systems (CHEOPS), at EuroSys 2021

- Associate Editor of the IEEE Transactions on Big Data (IEEE TBD) journal between 2014 and 2017
- Program committee member of more than 75 conferences from 2008

- Awards: CUG 2016 Best Paper award, HPC 2016 Best Paper Award, HPDC 2015 Best Paper Finalist, CUG 2013 Best Paper Award, SC08 Best Paper Finalist, SC03 Best Poster Award

Research Interests

- High performance data management: Enhancing parallel I/O libraries for exascale, characterizing and optimizing of I/O performance, and designing next generation storage systems and I/O at exascale
- AI and ML: Making data, metadata, and provenance ready for AI by applying FAIR principles, and efficiently managing experimental and observational data (EOD) from large-scale experiments and edge computing
- Supporting science: Extreme climate event detection and analysis in satellite and climate model data, simulation and analysis of trillion particle-scale applications, I/O support for Exascale Computing Project (ECP) applications

Job History

01/2023 – Present	Professor The Ohio State University
01/2023 – Present	Visiting Faculty Scientist, Lawrence Berkeley National Laboratory
07/2022 – 12/2022	Senior Scientist, Lawrence Berkeley National Laboratory
04/2015 – 06/2022	Computer Staff Scientist, LBNL
11/2010 – 03/2015	Computer Research Scientist, LBNL
04/2009 – 10/2010	Researcher, NEC Labs America Inc., Princeton
01/2007 – 03/2009	Research Assistant Professor, IIT, Chicago
1998 – 1999	Software Engineer, XL Softech, India

Education

Doctor of Philosophy, Computer Science Dec. 2006
Illinois Institute of Technology, Chicago, IL
Server-based Data Push Architecture for Data Access Performance Optimization

Master of Science, Computer Science May 2001
Illinois Institute of Technology, Chicago, IL
Performance Database and Its Applications in High Performance Computing

Bachelor of Technology, Electronics & Communication Engineering, Aug. 1997
Jawaharlal Nehru Technological University, Anantapur, India

Awards & Honors

- Associate Editor of the IEEE Transactions on Big Data (IEEE TBD) journal 2014-2017
- Best Paper Award winner, Cray User Group meeting, 2016 (CUG 2016)
- Best Paper Award winner, HPC 2016 conference
- Best paper finalist, HPDC 2015
- Best Paper Award winner, Cray User Group meeting, 2013 (CUG 2013)
- Best Paper Award Nominee at Supercomputing 2008 Conference (SC08)
- Best Poster Award winner for "Improving the Performance of MPI Derived Datatypes by Optimizing Memory-Access Cost," Supercomputing 2003 Conference (SC '03)

Patents

- "Energy-aware Task Consolidation on Graphics Processing Unit (GPU)", US Publication number: US8643656 B2

- "Data-aware Heterogeneous Scheduling", US Patent publication number: US8375392 B2

Software Development

- *Proactive Data Containers (PDC)*, a runtime system for managing data as objects to manage data transparently in heterogeneous and distributed storage devices
- *Async I/O VOL Connector* – A HDF5 Virtual Object Layer (VOL) connector for performing asynchronous I/O by overlapping I/O latency with computation phases
- *GPU I/O Virtual File Driver* – A HDF5 Virtual File Driver (VFD) for utilizing NVIDIA's GPU Direct Storage (GDS) package
- *H5Prov* – A HDF5 Virtual Object Layer (VOL) for collecting provenance by intercepting the HDF5 API
- *Data Elevator*, an engine for moving data transparently between different hierarchical storage layers
- *ArrayUDF*, an analysis framework for executing user defined functions on array data structures
- *h5bench* – A benchmark suite to exercise a diverse set of I/O patterns in HDF5
- *BD-CATS-IO*, an I/O routine for big data clustering algorithm
- *VPIC IO Utilities*, a set of I/O routines for VPIC simulation code
- *PIOK (Parallel I/O Kernel) Suite*, a collection of parallel I/O kernels extracted from real scientific applications
- *IO-SIG*: Generates IO-Signature, a pre-defined notation that can provide simple and clear representation of parallel I/O access patterns
- Software Designer of the Scientific Data Services (SDS) framework
- *HDF5 Readers and writers* for efficient parallel I/O
- *TECA: Toolkit for Extreme Climate event Analysis* – Atmospheric River detection

Collaborations and Leadership

- Lead the ECP ExaIO project team containing 24 members that include researchers and engineers from LBNL, The HDF Group, ANL, LLNL, and ORNL
- Lead the ECP HDF5 project team containing 17 members that include researchers and engineers from LBNL, The HDF Group, and ANL
- Lead the PDC team of 9 members that include researchers and engineers from LBL, The HDF Group, and ANL
- Lead the EOD-HDF5 team, which includes researchers, professors, engineers, and students from LBL and the HDF Group, Texas Tech Univ., and Ohio State University
- Lead ExaHDF5 project that involves researchers from The HDF5 Group and Argonne National Laboratory
- Lead postdocs and students in the TOKIO project, and worked as a Co-PI collaborating with ANL and Northeastern University
- Lead LBNL researchers and collaborate with Prof. Nagiza Samatova from NCSU on "In situ AMR data indexing and querying" project
- Collaborated with scientists from different domains at LBNL – climate, astronomy, ice sheet modeling, accelerator modeling
- Lead the effort of producing and analyzing 1 trillion plasma physics particle data (400 TB in size) with researchers from UC San Diego, Los Alamos National Lab, the HDF Group, and National Tsinghua University
- Worked with researchers from UCSD, LANL, LBNL, and Intel in analyzing trillion particle datasets from cosmology and plasma physics using novel data clustering algorithms

Mentoring

- Mentoring one postdoc in the SDM group (from 2023)

- Mentored one postdoc in the SDM group from 2021, who is now a regular staff
- Mentored five postdocs and several researchers and engineers in the SDM group at LBNL
- Supervised one engineer at NERSC
- Mentoring one Ph.D. student from North Carolina State University (John Ravi)
- Mentoring two Ph.D. students from Texas Tech University (Wei Zhang and Chenxu Niu)
- Mentored a Ph.D. student from UIUC (Babak Behzad) in developing an auto-tuning framework for parallel I/O optimizations
- Mentored at least five students each in 2022, 2020, and 2019, and three summer students each in 2018, 2017, and 2016, six summer students in 2015, three summer interns in 2014, two summer interns in 2013
- Served on thesis committee of two students (Wei Zhang and Jialin Liu) from Texas Tech University
- Serving on thesis committee of a student from UIUC (Margaret Lawson)
- Guided a postdoc at LBNL (Bin Dong), who is now a regular staff member
- Mentoring five summer students in 2021
- Mentored four interns at NEC Laboratories America, Inc.
- Assisted Prof. Xian-He Sun in various courses on parallel computing and programming.
- Mentored 6 students at Illinois Tech on their MS projects or thesis research
- Resulted in several research publications from the mentoring efforts with postdocs and students

Selected Invited Talks

- Keynote speech at The Thirteenth International Conference on Cloud Computing, GRIDs, and Virtualization (CLOUD COMPUTING) 2022 – "Towards Autonomous Storage and I/O Systems for Scientific Data", Apr 2022.
- Invited speech at the SIAM Conference on Parallel Processing for Scientific Computing (PP22) – "*A Runtime System for Orchestrating Data Movement in Heterogeneous Memory and Storage Hierarchies",* Feb 2022
- Keynote speech at the Workshop on Re-envisioning Extreme-Scale I/O for Emerging Hybrid HPC Workloads 2021 at IEEE Cluster 2021 - "*Trends, Challenges, and Potential Solutions in Scalable Storage and I/O*"
- Keynote speech at the Workshop on Challenges and Opportunities of Efficient and Performant Storage Systems (CHEOPS) at EuroSys 2021 - "*Parallel I/O at Exascale with HDF5 and Proactive Data Containers*"
- Invited speech at Dagstuhl Seminar on "Understanding I/O behavior in scientific and data-intensive computing" – "*IOMiner – A multi-level analysis to detect root causes of I/O bottlenecks*", Aug 2021
- Invited speech at Argonne Training Program on Extreme-Scale Computing (ATPESC) 2021 - "*ExaIO HDF5 features and application use cases*"
- Invited speech at PASC 2021 Mini-symposium on Data Movement Orchestration on HPC Systems – "*Automatic and Transparent Scientific Data Management with Object Abstractions*"
- Invited speech at ISC High Performance 2021 Focus Session panel on "Is Memory Becoming Indistinguishable from Storage" – "*Memory and storage continuum - A path with Proactive Data Containers (PDC)*"
- Invited speech at Oak Ridge Leadership Computing Facility (OLCF) I/O and Data Seminar Series (Jun 2019) – "*Efficient Parallel I/O with HDF5 and Proactive Data Containers (PDC)*"
- Invited speech at Indian Institute of Sciences, Bangalore (Dec 2019) – "*Efficient data management with Object-centric data abstractions*"

- Invited speech at Indian Institute of Sciences, Bangalore (Dec 2019) – "*ExaHDF5 – Readying HDF5 for exascale computing*"
- Invited speech at Inria – Rennes, France (Sep 2019) - "*Efficient data management with Object-centric data abstractions*"
- Invited speech at IMT Atlantique, Nantes, France (Sep 2019) – "ExaHDF5 - *Efficient Parallel I/O with HDF5*"
- Panelist on software upstream contributions at the ECP Annual Meeting, 2019
- Invited speech at DOE Data Day (D3) 2019 – "*Proactive Data Containers: An Intelligent Object-Centric Data Management System for HPC*"
- Invited speech at Dagstuhl Seminar on "Challenges and Opportunities of User-Level File Systems for HPC" – "*Proactive Data Containers – User-level data management runtime system*", May 2017

Publications

Google Scholar Profile: https://goo.gl/iTz2pS
(Citations: 3168, h-index: 32, i-10-index: 81 on 04/10/2023)

Book

1. Bin Dong, John Wu, Suren Byna, "*User-Defined Tensor Data Analysis*", Springer Briefs in Computer Science, 2021 (https://link.springer.com/book/10.1007/978-3-030-70750-7).

Journal Articles

2. Sunggon Kim, Alex Sim, Kesheng Wu, Suren Byna, Teng Wang, Yongseok Son, and Hyeonsang Eom, "*Design and Implementation of Dynamic I/O Control Scheme for Large Scale Distributed File Systems*", Journal of Cluster Computing, 25, pages 4423–4438, 2022. https://doi.org/10.1007/s10586-022-03640-0.
3. Bin Dong, Alex Popescu, Verónica Rodríguez Tribaldos, Suren Byna, Jonathan Ajo-Franklin, Kesheng Wu, "*Real-time and Post-hoc Compression for Data from Distributed Acoustic Sensing*", Computers & Geosciences, Volume 166, 2022, 105181, ISSN 0098-3004, https://doi.org/10.1016/j.cageo.2022.105181.
4. Houjun Tang, Quincey Koziol, John Ravi, and Suren Byna, "*Transparent Asynchronous Parallel I/O using Background Threads*", IEEE Transactions on Parallel and Distributed Systems (TPDS), vol. 33, no. 4, pp. 891-902, 1 April 2022, doi: 10.1109/TPDS.2021.3090322.
5. Jingqing Mu, Jerome Soumagne, Suren Byna, Quincey Koziol, Houjun Tang, and Richard Warren, "*Interfacing HDF5 with A Scalable Object-centric Storage System on Hierarchical Storage*", Journal of Concurrency and Computation: Practice and Experience, 2020.
6. Suren Byna, M. Scot Breitenfeld, Bin Dong, Quincey Koziol, Elena Pourmal, Dana Robinson, Jerome Soumagne, Houjun Tang, Venkatram Vishwanath, and Richard Warren, "*ExaHDF5: Delivering Efficient Parallel I/O on Exascale Computing Systems*," Journal of Computer Science and Technology, 2020, 35(1): 145-160. DOI: 10.1007/s11390-020-9822-9 (**JCST 2022 Highly Cited Paper Award**),
7. Babak Behzad, Suren Byna, Prabhat and Marc Snir, "*Optimizing I/O Performance of HPC Applications with Autotuning*", ACM Transactions on Parallel Computing, 2019
8. Beytullah Yildiz, Kesheng Wu, Suren Byna, and Arie Shoshani, "*Parallel membership queries on very large scientific data sets using bitmap indexes*", Concurrency and Computation: Practice and Experience, 2019

9. Soyoung Jeon, Prabhat, Suren Byna, Junmin Gu, William Collins, and Michael Wehner, "*Characterization of Extreme Precipitation within Atmospheric River Events over California*", Journal of Advances in Statistical Climatology, Meteorology and Oceanography, 1, 45-57, doi:10.5194/ascmo-1-45-2015, 2015.
10. Surendra Byna and Xian-He Sun, "*Editorial, Special issue on Data Intensive Computing*," Journal of Parallel and Distributed Computing (JPDC), Feb 2011.
11. Surendra Byna, Yong Chen, Xian-He Sun, "*Taxonomy of Data Prefetching Strategies for Multicore Processors*," Journal of Computer Science and Technology (JCST), Volume 24, Number 3 / May, pp. 405-417, 2009.
12. Xian-He Sun, Surendra Byna, and Yong Chen, "*Server-based Data Push Architecture for Multi-processor Environments*," Journal of Computer Science and Technology (JCST), 22(5): 641~652, September 2007.
13. Surendra Byna, Kirk W. Cameron and Xian-He Sun, "*Isolating Costs in Shared Memory Communication Buffering*," Parallel Processing Letters, Vol. 15, No. 4 (2005) 357-365, December 2005.

Conference/Workshop Publications

14. Hammad Ather, Jean Luca Bez, Boyana Norris, and Suren Byna, "Illuminating the I/O Optimization Path of Scientific Applications", ISC HPC 2023.
15. John Ravi, Suren Byna, Quincey Koziol, Houjun Tang, and Michela Becchi, "*Understanding Asynchronous Parallel I/O on HPC Systems*", 37th IEEE International Parallel and Distributed Processing Symposium (IPDPS) 2023.
16. John Ravi, Suren Byna, and Michela Becchi, "Runway: In-transit Data Compression on Heterogeneous HPC Systems", The 23rd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid) 2023.
17. Md Kamal Hossain Chowdhury, Houjun Tang, Jean Luca Bez, Purushotham Bangalore, and Suren Byna, "Efficient Asynchronous I/O with Request Merging", 4th Workshop on Extreme-Scale Storage and Analysis (ESSA) 2023.
18. Jean Luca Bez, Hammad Ather, and Suren Byna, "*Drishti: Guiding End-Users in the I/O Optimization Journey*", PDSW 2022 (in conjunction with SC22)
19. Rajeev Jain, Houjun Tang, Akash Dhruv, J. Austin Harris, and Suren Byna, "*Accelerating Flash-X Simulations with Asynchronous I/O*", PDSW 2022 (in conjunction with SC22)
20. Sian Jin, Dingwen Tao, Houjun Tang, Sheng Di, Suren Byna, Zarija Lukic, and Franck Cappello, "*Accelerating Parallel Write via Deeply Integrating Predictive Lossy Compression with HDF5*", Accepted to appear at SC22.
21. Runzhou Han, Suren Byna, Houjun Tang, Bin Dong, and Mai Zheng, "*PROV-IO: An I/O-Centric Provenance Framework for Scientific Data on HPC Systems*", HPDC 2022.
22. Jean Luca Bez, Ahmad Maroof Karimi, Arnab K. Paul, Bing Xie, Suren Byna, Philip Carns, Sarp Oral, Feiyi Wang, and Jesse Hanley, "*Access Patterns and Performance Behaviors of Multi-layer Supercomputer I/O Subsystems under Production Load*", HPDC 2022.
23. Suren Byna, "*Understanding Parallel I/O Performance and Tuning*", Keynote Speech Abstract, In Fifth International Workshop on Systems and Network Telemetry and Analytics (SNTA '22).
24. John Wu, Bin Dong, and Suren Byna, "*FasTensor: Data-Parallel Analysis for Large Remote Sensing Data*", IEEE GRSS workshop on Remote Sensing Data Management Technologies in GeoScience 2022 (RSDM-GeoSci 2022)
25. Huihuo Zheng, Venkatram Vishwanath, Quincey Koziol, Houjun Tang, John Ravi, John Mainzer and Suren Byna, "*HDF5 Cache VOL: Efficient and Scalable Parallel I/O through Caching Data on*

Node-local Storage", The 22nd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGrid) 2022.

26. Amit Bashyal, Peter Van Gemmeren, Saba Sehrish, Kyle Knoepfel, Suren Byna, and Qiao Kang, "*Data Storage for HEP Experiments in the Era of High-Performance Computing*", the Proceedings of the US Community Study on the Future of Particle Physics (Snowmass 2021)
27. Sian Jin, Sheng Di, Jiannan Tian, Suren Byna, Dingwen Tao, and Franck Cappello, "*Improving Prediction-Based Lossy Compression Dramatically via Ratio-Quality Modeling*", ICDE 2022.
28. Suren Byna, Stratos Idreos, Terry Jones, Kathryn Mohror, Rob Ross, Florin Rusu, "*Basic Research Needs for Management and Storage of Scientific Data*", A brochure on priority research directions discussed in the DOE ASCR workshop on Management and Storage of Scientific Data, 2022. Web. doi:10.2172/1845705
29. Runzhou Han, Suren Byna, and Mai Zheng, "*Towards A Practical Provenance Framework for Scientific Data on HPC Systems*", Poster, FAST '22
30. Suren Byna, Stratos Idreos, Terry Jones, Kathryn Mohror, Rob Ross, Florin Rusu, "*Position Papers for the ASCR Workshop on the Management and Storage of Scientific Data*", 2022. Web. doi:10.2172/1843500 (*Editor*)
31. Qiao Kang, Scot Breitenfeld, Kaiyuan Hou, Wei-keng Liao, Robert Ross, and Suren Byna, "*Optimizing Performance of Parallel I/O Accesses to Non-contiguous Blocks in Multiple Array Variables*", IEEE BigData 2021 conference, Dec 2021
32. Jiwoo Bang, Chungyong Kim, Kesheng Wu, Alex Sim, Suren Byna, Hanul Sung, Hyeonsang Eom , "*An In-Depth I/O Pattern Analysis in HPC Systems*", 28th IEEE International Conference on High Performance Computing, Data, & Analytics (HiPC 2021)
33. Houjun Tang, Suren Byna, N. Anders Petersson, and David Mccallen, "*Tuning Parallel Data Compression and I/O for Large-scale Earthquake Simulation*", The Second International Workshop on Big Data Reduction (IWBDR) 2021, held with 2021 IEEE Big Data
34. Wei Zhang, Suren Byna, Hyogi Sim, SangKeun Lee, Sudharshan Vazhkudai, Yong Chen, "*Exploiting User Activeness for Data Retention in HPC Systems*", Supercomputing 2021 (SC21)
35. J. Bez, H. Tang, B. Xie, D. Williams-Young, R. Latham, R. Ross, S. Oral, and S. Byna, "*I/O Bottleneck Detection and Tuning: Connecting the Dots using Interactive Log Analysis*", the 6th International Parallel Data Systems Workshop (PDSW'21), in conjunction with Supercomputing 2021 (SC21)
36. Houjun Tang, Bing Xie, Suren Byna, Phillip Carns, Quincey Koziol, Sudarsan Kannan, Jay Lofstead, and Sarp Oral, "*SCTuner: An Autotuner Addressing Dynamic I/O Needs on Supercomputer I/O Subsystems*", the 6th International Parallel Data Systems Workshop (PDSW'21), in conjunction with Supercomputing 2021 (SC21)
37. Cong Xu, Martin Flotin, Suparna Bhattacharya, Suren Byna, "*Data-Aware Storage Tiering for Deep Learning*", the 6th International Parallel Data Systems Workshop (PDSW'21), in conjunction with Supercomputing 2021 (SC21)
38. Bo Fang, Daoce Wang, Sian Jin, Quincey Koziol, Zhao Zhang, Qiang Guan, Suren Byna, Sriram Krishnamoorthy, Dingwen Tao, "*Characterizing Impacts of Storage Faults on HPC Applications: A Methodology and Insights*", 2021 IEEE International Conference on Cluster Computing, September 7–10, 2021.
39. Tonglin Li, Suren Byna, Quincey Koziol, Houjun Tang, Jean Luca Bez, and Qiao Kang, "*h5bench: HDF5 I/O Kernel Suite for Exercising HPC I/O Patterns*", Cray User Group meeting (CUG) 2021.
40. Bing Xie, Houjun Tang, Suren Byna, Jesse Hanley, Quincey Koziol, Tonglin Li, Sarp Oral, "*Battle of the Defaults: Extracting Performance Characteristics of HDF5 under Production Load*", CCGrid 2021

41. Suren Byna, Quincey Koziol, Houjun Tang, Wei Zhang, and Yong Chen, "*Searching metadata stored in self-describing file formats efficiently*", AGU 2020 Abstract in the "Recent Advances and Future Directions in Hydro-informatics: From Data to Decision-Making" session.
42. John Ravi, Suren Byna, Quincey Koziol, "*GPU Direct I/O with HDF5*", 5th International Parallel Data Systems Workshop (PDSW) 2020, in conjunction with SC20
43. Bjoern Enders, Debbie Bard, Cory Snavely, Lisa Gerhardt, Jason Lee, Becci Totzke, Katie Antypas, Suren Byna, Ravi Cheema, Shreyas Cholia, Mark Day, Aditi Gaur, Annette Greiner, Taylor Groves, Mariam Kiran, Quincey Koziol, Kelly Rowland, Chris Samuel, Ashwin Selvarajan, Alex Sim, David Skinner, Rollin Thomas, Gabor Torok, "*Cross-facility science with the Superfacility Project at LBNL*", IEEE/ACM 2nd Annual Workshop on Extreme-scale Experiment-in-the-Loop Computing (XLOOP), 2020, in conjunction with SC20
44. Debbie Bard, Cory Snavely, Lisa Gerhardt, Jason Lee, Becci Totzke, Katie Antypas, Suren Byna, Ravi Cheema, Shreyas Cholia, Mark Day, Aditi Gaur, Annette Greiner, Taylor Groves, Mariam Kiran, Quincey Koziol, Kelly Rowland, Chris Samuel, Ashwin Selvarajan, Alex Sim, David Skinner, Rollin Thomas, Gabor Torok, "*The Superfacility project: automated pipelines for experiments and HPC*", Extended Abstract for State of the Practice talk series, SC20 (2020)
45. Kai-yuan Hou, Quincey Kozio, and Suren Byna, "*TaskWorks: A Task Engine for Empowering Asynchronous Operations in HPC Applications*", SC20 Research Posters, 2020
46. Chenxu Niu, Wei Zhang, Suren Byna, and Yong Chen, "*Semantic Search for Self-Describing Scientific Data Formats*", SC20 Research Posters, 2020
47. Houjun Tang, Suren Byna, Bin Dong, and Quincey Koziol, "*Parallel Query Service for Object-centric Data Management Systems*", The 6th IEEE International Workshop on High-Performance Big Data and Cloud Computing (HPBDC) 2020, in conjunction with IPDPS 2020.
48. Donghe Kang, Oliver Rübel, Suren Byna, and Spyros Blanas, "*Predicting and Comparing the Performance of Array Management Libraries*", IPDPS 2020
49. Bin Dong, Veronica Rodriguez Tribaldos, Xin Xing, Suren Byna, Jonathan Ajo-Franklin, and Kesheng Wu, "*DASSA: Parallel DAS Data Storage and Analysis for Subsurface Event Detection*", IPDPS 2020
50. Jiwoo Bang, Chungyong Kim, Kesheng Wu, Alex Sim, Suren Byna, Sunggon Kim, and Hyeonsang Eom, "*HPC Workload Characterization Using Feature Selection and Clustering*", 3rd International Workshop on System and Network Telemetry and Analytics (SNTA'20), 2020
51. Tirthak Patel, Suren Byna, Glenn K. Lockwood, Nicholas J. Wright, Philip Carns, Rob Ross, and Devesh Tiwari, "*Uncovering Access, Reuse, and Sharing Characteristics of I/O-Intensive Files on Large-Scale Production HPC Systems*", FAST '20
52. Wei Zhang, Suren Byna, Chenxu Niu, and Yong Chen, "*Exploring Metadata Search Essentials for Scientific Data Management*", 26th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC) 2019
53. Houjun Tang, Suren Byna, Stephen Bailey, Zarija Lukic, Jialin Liu, Quincey Koziol, and Bin Dong, "*Tuning Object-centric Data Management Systems for Large Scale Scientific Applications*", 26th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC) 2019
54. Richard Warren, Jerome Soumagne, Jingqing Mu, Houjun Tang, Suren Byna, Bin Dong, and Quincey Koziol, "*Analysis in the Data Path of an Object-centric Data Management System*", 26th IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC) 2019
55. Houjun Tang, Quincey Koziol, Suren Byna, John Mainzer, and Tonglin Li, "*Enabling Transparent Asynchronous I/O using Background Threads*", PDSW 2019, in conjunction with SC19.

56. Glenn K. Lockwood, Shane Snyder, Suren Byna, Philip Carns, and Nicholas J. Wright, "*Understanding Data Motion in the Modern HPC Data Center*", PDSW 2019, in conjunction with SC19.
57. Megha Agarwal, Divyansh Singhvi, Preeti Malakar, and Suren Byna, "*Active Learning-based Automatic Tuning and Prediction of Parallel I/O Performance*", PDSW 2019, in conjunction with SC19.
58. Tirthak Patel, Suren Byna, Glenn K. Lockwood, and Devesh Tiwari, "*Revisiting I/O Behavior in Large-Scale Storage Systems: The Expected and the Unexpected*", SC19
59. Wei Zhang, Suren Byna, Houjun Tang, Brody Williams, and Yong Chen, "*MIQS: Metadata Indexing and erying Service for Self-Describing File Formats*", SC19
60. John Mainzer, Neil Fortner, Gerd Heber, Elena Pourmal, Quincey Koziol, Suren Byna, and Marc Paterno, "*Sparse Data Management in HDF5*", 2019 IEEE/ACM 1st Annual Workshop on Large-scale Experiment-in-the-Loop Computing (XLOOP), 2019 in conjunction with SC19, pp. 20-25, doi: 10.1109/XLOOP49562.2019.00009.
61. Donghe Kang, Oliver Rübel, Suren Byna, and Spyros Blanas, "*Comparison of Array Management Library Performance - A Neuroscience Use Case*", SC19 Poster
62. Bin Dong, Kesheng Wu, Suren Byna, and Houjun Tang, "*SLOPE: Structural Locality-aware Programming Model for Composing Array Data Analysis*", ISC High Performance 2019.
63. Jingqing Mu, Jerome Soumagne, Suren Byna, Quincey Koziol, Houjun Tang, and Richard Warren, "*Interfacing HDF5 with A Scalable Object-centric Storage System on Hierarchical Storage*", CUG 2019
64. Tonglin Li, Quincey Koziol, Houjun Tang, Jialin Liu, and Suren Byna, "*I/O Performance Analysis of Science Applications Using HDF5 File-level Provenance*", CUG 2019
65. Teng Wang, Suren Byna, Glenn Lockwood, Philip Carns, Shane Snyder, Sunggon Kim, and Nicholas Wright, "*A Zoom-in Analysis of I/O Logs to Detect Root Causes of I/O Performance Bottlenecks*", IEEE/ACM CCGrid 2019
66. Sunggon Kim, Alex Sim, Kesheng Wu, Suren Byna, Teng Wang, Yongseok Son and Hyeonsang Eom, "*DCA-IO: A Dynamic I/O Control Scheme for Parallel and Distributed File System*", IEEE/ACM CCGrid 2019
67. Bin Dong, Teng Wang, Houjun Tang, Quincey Koziol, Kesheng Wu, and Suren Byna, "*ARCHIE: Data Analysis Acceleration with Array Caching in Hierarchical Storage*", IEEE International Conference on Big Data (IEEE BigData) 2018
68. J.S. Vetter, R. Brightwell, M. Gokhale, P. McCormick, R. Ross, J. Shalf, K. Antypas, D. Donofrio, A. Dubey, T. Humble, C. Schuman, B. Van Essen, S. Yoo, A. Aiken, D. Bernholdt, S. Byna, K. Cameron, F. Cappello, B. Chapman, A. Chien, M. Hall, R. Hartman-Baker, Z. Lan, M. Lang, J. Leidel, S. Li, R. Lucas, J. Mellor-Crummey, P. Peltz, Jr., T. Peterka, M. Strout, and J. Wilke, "*Extreme Heterogeneity 2018: DOE ASCR Basic Research Needs Workshop on Extreme Heterogeneity*", US Department of Energy, Office of Science, Advanced Scientific Computing Research, 2018, doi:10.2172/1473756.
69. Suren Byna, Quincey Koziol, Venkatram Vishwanath, Jerome Soumagne, Houjun Tang, Kimmy Mu, Richard Warren, François Tessier, Bin Dong, Teng Wang, and Jialin Liu, "*Proactive Data Containers (PDC): An object-centric data store for large-scale computing systems*", AGU Fall Meeting 2018
70. Wei Zhang, Houjun Tang, Suren Byna, and Yong Cheng, "*DART: Distributed Adaptive Radix Tree for Efficient Affix-based Keyword Search on HPC Systems*", The 27th International Conference on Parallel Architectures and Compilation Techniques (PACT'18)
71. Jialin Liu, Quincey Koziol, Gregory Butler, Neil Fortner, Mohamad Chaarawi, Houjun Tang, Suren Byna, Glenn Lockwood, Ravi Cheema, Kristy Kallback-Rose, Damian Hazen, and Prabhat, "*Evaluation of HPC Application I/O on Object Storage Systems*", 3rd Joint International Workshop

on Parallel Data Storage and Data Intensive Scalable Computing Systems (PDSW-DISCS), 2018 (Held in conjunction with SC18)

72. Fahim Chowdhury, Jialin Liu, Quincey Koziol, Thorsten Kurth, Steven Farrell, Suren Byna, Prabhat, Weikuan Yu, "*Initial Characterization of I/O in Large-Scale Deep Learning Applications*", Work in Progress paper, 3rd Joint International Workshop on Parallel Data Storage and Data Intensive Scalable Computing Systems (PDSW-DISCS), 2018 (Held in conjunction with SC18)
73. Teng Wang, Suren Byna, Bin Dong, and Houjun Tang, "*UniviStor: Integrated Hierarchical and Distributed Storage for HPC*", IEEE Cluster 2018.
74. Teng Wang, Suren Byna, Glenn Lockwood, Nicholas Wright, Phil Carns, and Shane Snyder, "*IOMiner: Large-scale Analytics Framework for Gaining Knowledge from I/O Logs*", IEEE Cluster 2018.
75. Kimmy Mu, Jerome Soumagne, Houjun Tang, Suren Byna, Quincey Koziol, and Richard Warren, "*A Server-managed Transparent Object Storage Abstraction for HPC*", IEEE Cluster 2018.
76. Glenn Lockwood, Shane Snyder, Teng Wang, Suren Byna, Phil Carns, and Nicholas Wright, "*A Year in the Life of a Parallel File System*", 2018 International Conference for High Performance Computing, Networking, and Storage (SC'18)
77. Haoyuan Xing, Sofoklis Floratos, Spyros Blanas, Suren Byna, Prabhat, Kesheng Wu, Paul Brown, "*ArrayBridge: Interweaving declarative array processing in SciDB with imperative HDF5-based programs*", 34th IEEE International Conference on Data Engineering (ICDE) 2018
78. Houjun Tang, Suren Byna, Francois Tessier, Teng Wang, Bin Dong, Jingqing Mu, Quincey Koziol, Jerome Soumagne, Venkatram Vishwanath, Jialin Liu, and Richard Warren, "*Toward Scalable and Asynchronous Object-centric Data Management for HPC*", 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid) 2018
79. Bharti Wadhwa, Suren Byna, Ali R. Butt, "*Toward Transparent Data Management in Multi-layer Storage Hierarchy for HPC Systems*", IEEE International Conference on Cloud Engineering 2018 (IC2E 2018)
80. Glenn Lockwood, Shane Snyder, Wucherl Yoo, Kevin Harms, Zachary Nault, Suren Byna, Philip Carns, Nicholas Wright, "*UMAMI: A Recipe for Generating Meaningful Metrics through Holistic I/O Performance Analysis*", 2nd Joint International Workshop on Parallel Data Storage & Data Intensive Scalable Computing Systems (PDSW-DISCS), 2017 (Held in conjunction with SC17)
81. Houjun Tang, Suren Byna, Bin Dong, Jialin Liu, and Quincey Koziol, "*SoMeta: Scalable Object-centric Metadata Management for High Performance Computing*", IEEE Cluster 2017
82. Bin Dong, Kesheng Wu, Suren Byna, Jialin Liu, Weijie Zhao, and Florin Rusu, "*ArrayUDF: User-Defined Scientific Data Analysis on Arrays*", The ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC) 2017
83. Suren Byna, Mohamad Chaarawi, Quincey Koziol, John Mainzer, and Frank Willmore, "*Tuning HDF5 subfiling performance on parallel file systems*", Cray User Group Conference 2017 (CUG 2017)
84. Jialin Liu, Quincey Koziol, Houjun Tang, François Tessier, Wahid Bhimji, Brandon Cook, Brian Austin, Suren Byna, Bhupender Thakur, Glenn Lockwood, Jack Deslippe, and Prabhat, "*Understanding the I/O Performance Gap Between Cori KNL and Haswell*", Cray User Group Conference 2017 (CUG 2017)
85. Cong Xu, Shane Snyder, Omkar Kulkarni, Vishwanath Venkatesan, Philip Carns, Suren Byna, Robert Sisneros, and Kalyana Chadalavada, "*DXT: Darshan eXtended Tracing*", Cray User Group Conference 2017 (CUG 2017)
86. Bin Dong, Suren Byna, Kesheng Wu, Prabhat, Hans Johansen, Jeffrey N. Johnson, and Noel Keen, "*Data Elevator: Low-contention Data Movement in Hierarchical Storage System*", The 23rd annual IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC), 2016.
87. Wenzhao Zhang, Houjun Tang, Stephen Ranshous, Suren Byna, Daniel F. Martin, Kesheng Wu, Bin Dong, Scott Klasky, and Nagiza F. Samatova, "*Exploring memory hierarchy and network topology for runtime AMR data sharing across scientific applications*", Big Data Conference 2016
88. Houjun Tang, Suren Byna, Steve Harenberg, Wenzhao Zhang, Xiaocheng Zou, Daniel F. Martin, Bin Dong, Dharshi Devendran, Kesheng Wu, David Trebotich, Scott Klasky, and Nagiza Samatova,

"*In situ Storage Layout Optimization for AMR Spatio-temporal Read Accesses*", The International Conference on Parallel Processing (ICPP), 2016, to be held in Philadelphia, Aug 16-19, 2016

89. Bin Dong, Suren Byna, and Kesheng Wu, "*SDS-Sort: Scalable Dynamic Skew-aware Parallel Sorting*", The ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC) 2016
90. Md. Mostofa Ali Patwary, Nadathur Satish, Narayanan Sundaram, Jialin Liu, Peter Sadowski, Evan Racah, Suren Byna, Wahid Bhimji, Craig Tull, Prabhat, Predeep Dubey, "*PANDA: Extreme Scale Parallel K-Nearest Neighbor on Distributed Architectures*", 30th International Parallel & Distributed Processing Symposium, 2016.
91. Houjun Tang, Suren Byna, Steven Harenberg, Xiaocheng Zou, Wenzhao Zhang, Kesheng Wu, Bin Dong, Oliver Rubel, Kristofer Bouchard, Scott Klasky and Nagiza Samatova, "*Usage Pattern-Driven Dynamic Data Layout Reorganization*", 16th International Symposium on Cluster, Cloud and Grid Computing (CCGrid) 2016.
92. Wenzhao Zhang, Houjun Tang, Steven Harenberg, Suren Byna, Xiaocheng Zou, Dharshi Devendran, Daniel Martin, Kesheng Wu, Bin Dong, Scott Klasky and Nagiza Samatova, "*AMRZone: A Runtime AMR Data Sharing Framework for Scientific Applications*", 16th International Symposium on Cluster, Cloud and Grid Computing (CCGrid) 2016.
93. Dharshi Devendran, Suren Byna, Bin Dong, Brian van Straalen, Hans Johansen, and Noel Keen, "*Collective I/O Optimizations for Adaptive Mesh Refinement (AMR) Data Writes on Lustre File System*", Cray User Group meeting (CUG) 2016.
94. Cong Xu, Suren Byna, Vishwanath Venkatesan, Omkar Kulkarni, and Kalyana Chadalavada, "*LIOProf: Exposing Lustre File System Behavior for I/O Middleware*", Cray User Group meeting (CUG) 2016.
95. Wahid Bhimji, Debbie Bard, Melissa Romanus, David Paul, Andrey Ovsyannikov, Brian Friesen, Matt Bryson, Joaquin Correa, Glenn K. Lockwood, Vakho Tsulaia, Suren Byna, Steve Farrell, Doga Gursoy, Chris Daley, Vince Beckner, Brian Van Straalen, Nicholas Wright, Katie Antypas, and Prabhat, "*Accelerating Science with the NERSC Burst Buffer Early User Program*", CUG 2016, [**Best Paper Award**]
96. Jialin Liu, Alex Gittens, Suren Byna, Quincey Koziol, Lisa Gerhardt, Evan Racah, Michael F. Ringenburg, and Prabhat, "*H5Spark: Bridging the I/O Gap between Spark and Scientific Data Formats on HPC Systems*", Cray User Group meeting (CUG) 2016.
97. X. Zou, D. Boyuka II, D. Desai, D. F. Martin, S. Byna, K. Wu, K. Bansal, B. Dong, W. Zhang, H. Tang, D. Devendran, D. Trebotich, S. Klasky, H. Johansen, and N. Samatova, "*AMR-aware In Situ Indexing and Scalable Querying*", 24th High Performance Computing Symposium (HPC) 2016. [**HPC 2016 Best Paper Award, Spring Simulation Multi-Conference 2016 (SpringSim'16) Best Paper Runner up**]
98. Burlen Loring, Suren Byna, Prabhat, Junmin Gu, Hari Krishnan, Michael Wehner, and Oliver Ruebel, "*TECA an Extreme Event Detection and Climate Analysis Package for High Performance Computing*", The AMS (American Meteorological Society) 96th Annual Meeting, Jan 2016
99. Harinarayan Krishnan, Burlen Loring, Suren Byna, Michael F. Wehner, Travis A. O'Brien, Prabhat, Chris Paciorek, and Daithi Stone, "*Enabling End-to-End Climate Science Workflows in High Performance Computing Environments*", The AMS (American Meteorological Society) 96th Annual Meeting, Jan 2016
100. Md. Mostofa Ali Patwary, Suren Byna, Nadathur Satish, Narayanan Sundaram, Zarija Lukic, Vadim Roytershteyn, Michael Anderson, Yushu Yao, Prabhat, and Pradeep Dubey, "*BD-CATS: Big Data Clustering at Trillion Particle Scale*", ACM/IEEE Supercomputing (SC15), 2015.
101. Bin Dong, Suren Byna, and Kesheng Wu, "*Heavy-tailed Distribution of Parallel I/O System Response Time*", 10th Parallel Data Storage Workshop (PDSW) 2015, to be held in conjunction with SC15, November 2015
102. Babak Behzad, Suren Byna, Prabhat and Marc Snir, "*Pattern-driven Parallel I/O Tuning*", 10th Parallel Data Storage Workshop (PDSW) 2015, in conjunction with SC15, November 2015

103. Shane Snyder, Philip Carns, Robert Latham, Misbah Mubarak, Chris Carothers, Babak Behzad, Huong Vu Thanh Luu, Suren Byna, and Prabhat, "*Techniques for Modeling Large-scale HPC I/O Workloads*", the 6th International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems (PMBS15), in conjunction with SC15, November 2015
104. Jinoh Kim, Bin Dong, Suren Byna, and Kesheng Wu, "*Security for the Scientific Data Service Framework*", 2nd International Workshop on Privacy and Security of Big Data (PSBD 2015), in conjunction with IEEE BigData 2015
105. Jialin Liu, Yong Chen and Suren Byna, "*Collective Computing for Scientific Big Data Analysis*", Eighth International Workshop on Parallel Programming Models and Systems Software for High-End Computing (P2S2), 2015, in conjunction with ICPP 2015
106. Bin Dong, Suren Byna, Kesheng Wu, "*Spatially Clustered Join on Heterogeneous Scientific Data Sets*", accepted for publication at IEEE Big Data 2015
107. Babak Behzad, Suren Byna, Stefan Wild, Prabhat and Marc Snir, "*Dynamic Model-driven Parallel I/O Performance Tuning*", IEEE Cluster 2015
108. H. Luu, M. Winslett, W. Gropp, R. Ross, P. Carns, K. Harms, Prabhat, S. Byna, and Y. Yao, "*A Multi-platform Study of I/O Behavior on Petascale Supercomputers*", The 24th ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC) 2015, [Best paper finalist]
109. X. Zou, K. Wu, D.Boyuka II, D. Martin, S. Byna, H. Tang, K. Bansal, T. Ligocki, H. Johansen, and N. Samatova, "*Parallel In Situ Detection of Connected Components in Adaptive Mesh Refinement Data*", IEEE/ACM CCGrid 2015
110. Suren Byna and Brian Austin, "*Evaluation of Parallel I/O Performance and Energy Consumption with Frequency Scaling on Cray XC30*", Cray Users Group (CUG) meeting 2015
111. Suren Byna, Kalyana Chadalavada, Rob Sisneros, and Quincey Koziol, "*Tuning Parallel I/O on Blue Waters for Writing 10 Trillion Particles*", Cray Users Group (CUG) meeting 2015
112. S. Jeon, C. Paciorek, Prabhat, S. Byna, W. Collins, and M. Wehner, "*Uncertainty Quantification for Characterizing Spatial Tail Dependence under Statistical Framework*", Poster @ American Geophysical Union (AGU) - Fall Meeting 2014
113. M. S. Breitenfeld, K. Chadalavada, R. Sisneros, S. Byna, Q. Koziol, N. Fortner, Prabhat, and V. Vishwanath, "*Recent Progress in Tuning Performance of Large-scale I/O with Parallel HDF5*", Work-in-Progress presentation @ the 9th Parallel Data Storage Workshop (PDSW) held in conjunction with SC14, 2014
114. Bin Dong, Suren Byna, and John Wu, "*Parallel Query Evaluation as a Scientific Data Service*", IEEE Cluster 2014
115. S. Blanas, K. Wu, S. Byna, B. Dong, and A. Shoshani, "*Parallel Data Analysis Directly on Scientific File Formats*", ACM SIGMOD 2014
116. B. Behzad, S. Byna, S. Wild, Prabhat, and M. Snir, "*Improving Parallel I/O Autotuning with Performance Modeling*", The 23rd ACM International Symposium on High-Performance Parallel and Distributed Computing (HPDC) 2014
117. Hsuan-Te Chiu, Jerry Chou, Venkat Vishwanath, Suren Byna, and Kesheng Wu, "*Simplifying Index File Structure to Improve I/O Performance of Parallel Indexing*", The 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS 2014)
118. Ted Habermann, Andrew Collette, Steve Vincena, Werner Benger, Jay Jay Billings, Matt Gerring, Konrad Hinsen, Pierre de Buyl, Mark Könnecke, Filipe Maia, and Suren Byna, "*The Hierarchical Data Format (HDF): A Foundation for Sustainable Data and Software*", 2nd Workshop on Sustainable Software for Science: Practice and Experiences (WSSSPE2), in conjunction with Supercomputing 2014 (SC14)
119. J. Liu, S. Byna, B. Dong, K. Wu, and Y. Chen, "*Model-driven Data Layout Selection for Improving Read Performance*", High Performance Data Intensive Computing (HPDIC2014) workshop in conjunction with IEEE IPDPS 2014

120. B. Behzad, H. Luu, J. Huchette, S. Byna, Prabhat, R. Aydt, Q. Koziol, and M. Snir, "*Taming Parallel I/O Complexity with Auto-Tuning*", ACM/IEEE SuperComputing Conference (SC13), November 2013.
121. B. Dong, S. Byna, and K. Wu, "*SDS: A Framework for Scientific Data Services*", 8th Parallel Data Storage Workshop (PDSW) held in conjunction with SC13, 2013
122. B. Dong, S. Byna, and K. Wu, "*Expediting Scientific Data Analysis with Reorganization*", IEEE Cluster 2013
123. J.L. Liu, S. Byna, and Y. Chen, "*Segmented Analysis for Reducing Data Movement*", IEEE International Conference on Big Data (BigData 2013)
124. K-W. Lin, J. Chou, S. Byna, and K. Wu, "*Optimizing FastQuery Performance on Lustre File System*", 25th International Conference on Scientific and Statistical Database Management (SSDBM) 2013.
125. B. Behzad, J. Huchette, H. Luu, R. Aydt, S. Byna, Y.Yao, Q. Koziol, and Prabhat, "*A Framework for Auto-tuning HDF5 Application*", HPDC 2013 Short Paper.
126. S. Jeon, Prabhat, S. Byna, W. Collins, and M. F. Wehner, "*Event Detection and Spatial Analysis for Characterizing Extreme Precipitation*", Poster @ American Geophysical Union (AGU) - Fall Meeting 2013
127. Prabhat, S. Byna, V. Vishwanath, W. Bethel, W. Collins, and M. F. Wehner, "*TECA: Extreme Climate Analytics on Petascale Platforms*", Poster @ American Geophysical Union (AGU) - Fall Meeting 2013
128. S. Byna, A. Uselton, Prabhat, D. Knaak, and Y. He, "*Trillion Particles, 120,000 cores, and 350 TBs: Lessons Learned from a Hero I/O Run on Hopper*", Cray User Group meeting 2013, [Best Paper Award]
129. E. Wes Bethel, Prabhat, Suren Byna, Oliver Rübel, K. John Wu, and Michael Wehner, "*Why High Performance Visual Data Analytics is both Relevant and Difficult*", Proceedings of Visualization and Data Analysis 2013, IS&T/SPIE Electronic Imaging 2013, San Francisco CA, USA.
130. S. Byna, J. Chou, O. Rubel, Prabhat, H. Karimabadi, W. Daughton, V. Roytershteyn, W. Bethel, M. Howison, K-H. Hsu, K-W. Lin, A. Shoshani, A. Uselton, K. Wu, "*Parallel I/O, Analysis, and Visualization of a Trillion Particle Simulation*", in the Proc. of the ACM/IEEE SuperComputing Conference (SC'12), November 2012.
131. B. Behzad, J. Huchette, H. Luu, R. Aydt, S. Byna, M. Chaarawi, Q. Koziol, Prabhat, and Y. Yao, "*POSTER: Auto-tuning of Parallel I/O Parameters for HDF5 Applications*", Poster at the ACM/IEEE SuperComputing Conference (SC'12), November 2012.
132. Prabhat, O. Rubel, S. Byna, K. Wu, M. Wehner, W. Bethel, "*TECA: A Parallel Toolkit for Extreme Climate Analysis*," Third Worskhop on Data Mining in Earth System Science (DMESS 2012), June 2012.
133. Michael Wehner, Suren Byna, Prabhat, Thomas Yopes, and John Wu, "*Atmospheric rivers in the CMIP3/5 historical and projection simulations*", Poster at the WCRP workshop on CMIP5 Model Analysis, March 5-9 2012. https://escholarship.org/uc/item/7s91q21f
134. Y. Yin, S. Byna, H. Song, X.H. Sun, and R. Thakur, "*Boosting Application-Specific Parallel I/O Optimization Using IOSIG*," IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid) 2012.
135. S. Byna, J. Chou, O. Rübel, Prabhat, H. Karimabadi, W. S. Daughton, V. Roytershteyn, E. W. Bethel, M. Howison, K-J. Hsu, K-W. Lin, A. Shoshani, A. Uselton, and K. Wu, "*Parallel Data, Analysis, and Visualization of a Trillion Particles*", Lightning Talk and Poster Presentation at the 6th Extremely Large Databases Conference (XLDB), 2012.
136. Prabhat, Surendra Byna, Chris Paciorek, Gunther Weber, Kesheng Wu, Thomas Yopes, Michael Wehner, William Collins, George Ostrouchov, Richard Strelitz, E. Wes Bethel, "*Pattern Detection and Extreme Value Analysis on Large Climate Data,*" American Geophysical Union (AGU) Fall Meeting 2011, San Francisco, December 2011.
137. E. W. Bethel, R. Ross, W-K. Liao, Prabhat, K. Schuchardt, P-T. Bremer, O. Rübel, S. Byna, K. Wu, F. Li, M. Wehner, J. Patchett, H-W. Shen, D. Pugmire, and D. Williams, "*Recent Advances in Visual Data Exploration and Analysis of Climate Data*", Poster at the SciDAC 3 Principal Investigator Meeting, Rockville, MD, September 2012.

138. E. W. Bethel, Prabhat, O. Rübel, S. Byna, K. Wu, F. Li, M. Wehner, J. Chou, M. Howison, J. Qiang, R. D. Ryne, C. G. R. Geddes, and E. Cormier-Michel, "*Big Data Analysis and Visualization: What do Linacs and Troplical Storms have in Common?*", Invited paper at the 11th International Computational Accelerator Physics Conference (ICAP) 2012.
139. K. Wu, S. Byna, D. Rotem, and A. Shoshani, "*Scientific Data Services - A High-Performance I/O System with Array Semantics*," 1st Workshop on High-Performance Computing meets Databases (HPCDB 2011), Co-located with Supercomputing 2011, Seattle WA.
140. S. Byna, Prabhat, M. Wehner, and K. Wu, "*Detecting Atmospheric Rivers in Large Climate Datasets*," 2nd International Workshop on Petascale Data Analytics: Challenges, and Opportunities (PDAC-11), Co-located with Supercomputing 2011, Seattle WA.
141. D. Li, S. Byna, S. Chakradhar, "*Energy-Aware Workload Consolidation on GPU,*" Accepted for publication at 7th International Workshop on Scheduling and Resource Management for Parallel and Distributed Systems, 2011 (in conjunction with ICPP 2011).
142. M. Becchi, S. Byna, S. Cadambi, and S. Chakradhar, "*Data-Aware Scheduling of Legacy Kernels on Heterogeneous Platforms with Distributed Memory*," to appear in 22nd ACM Symposium on Parallelism in Algorithms and Architectures (SPAA), Santorini, Greece, June 2010.
143. S. Byna, J. Meng, A. Raghunathan, S. Chakradhar, and S. Cadambi, "*Best Effort Semantic Document Search on GPUs*," Third Workshop on General-Purpose Computation on Graphics Procesing Units, Pittsburgh, PA, Mar 2010.
144. J. Meng, A. Raghunathan, S. Chakradhar, and S. Byna, "*Exploiting the Forgiving Nature of Applications for Scalable Parallel Execution*," in proceedings of the IEEE International Parallel & Distributed Processing Symposium (IPDPS), 2010.
145. X.-H. Sun, S. Byna, D. Holmgren, "*Modeling Data Access Contention in Multicore Architectures,*" in Proceedings of the Fifteenth International Conference on Parallel and Distributed Systems (ICPADS'09), Dec. 2009.
146. Zhibin Fang, Xian-He Sun, Yong Chen, and Surendra Byna, "*Core-aware Memory Access Scheduling Schemes*," IEEE International Parallel & Distributed Processing Symposium (IPDPS), 2009.
147. Suren Byna, Yong Chen, Xian-He Sun, Rajeev Thakur, and William Gropp, "*Parallel I/O Prefetching Using MPI File Caching and I/O Signatures,*" in proceedings of the IEEE/ACM Supercomputing (SC08), Austin, Nov 2008
148. Yong Chen, Suren Byna, Xian-He Sun, Rajeev Thakur, and William Gropp, "*Hiding I/O Latency with Pre-execution Prefetching for Parallel Applications*," in proceedings of the IEEE/ACM Supercomputing (SC08), Austin, Nov 2008. [**Best Paper and Best Student Paper Awards Nominee**]
149. Suren Byna, Yong Chen, and Xian-He Sun, "*A Taxonomy of Data Prefetching Mechanisms,*" in Proceedings of the International Symposium on Parallel Architectures, Algorithms, and Networks (ISPAN), Sydney, Australia, 2008.
150. Xian-He Sun, Yong Chen and Suren Byna, "*Scalable Computing in Multicore Era*," in Proceedings of the International Symposium on Parallel Algorithms, Architectures and Programming (PAAP '08), 2008.
151. Yong Chen, Suren Byna, Xian-He Sun, Rajeev Thakur, and William Gropp, "*Exploring Parallel I/O Concurrency with Pre-execution Prefetching*," in Proceedings of the 37th International Conference on Parallel Processing (ICPP), September 2008.
152. Suren Byna, Yong Chen, William Gropp, Xian-He Sun, and Rajeev Thakur, "*POSTER: The Server-Push I/O Architecture for High-End Computing*," presentation in the International Conference for High Performance Computing, Networking, Storage and Analysis, (Supercomputing 2007), November 2007.
153. Yong Chen, Surendra Byna, and Xian-He Sun, "*Data Access History Cache and Associated Data Prefetching Mechanisms*," in Proceedings of the International Conference on Supercomputing (SC '07), November 2007.
154. Xian-He Sun, Surendra Byna, and Yong Chen, "*Improving Data Access Performance with Server Push Architecture*," in Proceedings of the NSF Next Generation Software Program Workshop (in conjunction with IPDPS '07), March 2007.

155. Surendra Byna, Xian-He Sun, Rajeev Thakur and William Gropp, "*Automatic Memory Optimizations for Improving MPI Derived Datatype Performance,*" in Proceedings of the 13th European PVM/MPI Users' Group Meeting, 2006 (Euro PVM/MPI '06), Bonn, Germany, Sep 2006.
156. Surendra Byna, Xian-He Sun, and Ryan Nakhoul, "*Memory Servers: A Scope of SOA for High-End Computing,*" in Proceedings of the IEEE International Conference on Services Computing, 2006 (SCC '06), Chicago, September 2006.
157. Surendra Byna, Xian-He Sun, William Gropp and Rajeev Thakur, "*Predicting the Memory-Access Cost Based on Data Access Patterns,*" in Proceedings of the IEEE International Conference on Cluster Computing, San Diego, September 2004.
158. Surendra Byna, William Gropp, Xian-He Sun and Rajeev Thakur, "*Improving the Performance of MPI Derived Datatypes by Optimizing Memory-Access Cost*," in Proceedings of the IEEE International Conference on Cluster Computing, Hong Kong, December 2003.
159. Surendra Byna, William Gropp, Xian-He Sun and Rajeev Thakur, "*POSTER: Improving the Performance of MPI Derived Datatypes by Optimizing Memory-Access Cost*," in Proceedings of the ACM/IEEE Supercomputing Conference (SC '03), November 2003. [**Best Poster Award Winner**]
160. Surendra Byna, Kirk W. Cameron and Xian-He Sun, "*Memory-Aware Communication -An Experimental Study with MPI,*" in Proceedings of the 1st International Workshop on Hardware/Software Support for Parallel and Distributed Scientific and Engineering Computing (SPDSEC '02), 2002.

Book Chapters

1. Suren Byna, Prabhat, Homa Karimabadi, and William Daughton, "*Parallel I/O for a Trillion-Particle Plasma Physics Simulation*", Book chapter in "High Performance Parallel I/O", October 23, 2014 by Chapman and Hall/CRC, edited by Prabhat and Quincey Koziol
2. Mark Howison, Prabhat, and Suren Byna, "*IOTA – I/O Trace and Analysis toolkit*", Book chapter in "High Performance Parallel I/O", October 23, 2014 by Chapman and Hall/CRC, edited by Prabhat and Quincey Koziol
3. Surendra Byna, Kirk Cameron, and Xian-He Sun, "*Quantification of memory communication*," Book Chapter in High Performance Scientific and Engineering Computing- Hardware/Software Support, Kluwer Academic Publishers, Chapter 3, pp: 31-44, 2003.

Technical Reports

1. Haoyuan Xing, Sofoklis Floratos, Spyros Blanas, Suren Byna, Prabhat, Kesheng Wu, Paul Brown, "*ArrayBridge: Interweaving declarative array processing with high-performance computing*", arXiv preprint, arXiv:1702.08327, 2017
2. Spyros Blanas and Suren Byna, "Towards Exascale Scientific Metadata Management", arXiv preprint, arXiv:1503.08482, 2015
3. S. Cadambi, A. Majumdar, M. Becchi, S. Byna, S.T. Chakradhar, H.P. Graf, "*A Massively Parallel, Smart Memory-based Accelerator for Data Analytics*", NEC Labs TR2009-L120 November 2009.
4. Surendra Byna and Xian-He Sun, "*Data-access Memory Servers for Multi-processor Environments*," IIT CS TR-2005-001, Nov 2005.
5. Kirk W. Cameron, Xian-He Sun, Surendra Byna, and Rong Ge, "*Predicting and Evaluating Memory Communication Performance*," USC CSCE TR-2003-021. Sep 2003.
6. Surendra Byna, William Gropp, Xian-He Sun, and Rajeev Thakur, "*Improving the Performance of MPI Derived Datatypes by Optimizing Memory-Access Cost*," Preprint ANL/MCS-P1045-0403, Math. and Computer Science Division, Argonne National Laboratory, Apr 2003.

Other contributions (Software DOIs and project reports from OSTI)

1. Bin Dong, John Wu, and Suren Byna, "H5TurboPFor v0.0.1", Computer software. https://www.osti.gov//servlets/purl/1820937. USDOE. 15 Sep. 2021. Web. doi:10.11578/dc.20210917.7.
2. Jean Luca Bez and Suren Byna, "DXT Explorer v0.1." Computer software. https://www.osti.gov//servlets/purl/1819264. Aug. 2021. Web. doi:10.11578/dc.20210908.3.

3. Tonglin Li, Suren Byna, Houjun Tang, Quincey Koziol, "H5bench: a benchmark suite for parallel HDF5 (H5bench) v0.6. Computer software. https://www.osti.gov//servlets/purl/1798322. USDOE; Oak Ridge National Laboratory (ORNL). 15 Apr. 2021. Web. doi:10.11578/dc.20210624.4.
4. Wu, Kesheng, Byna, Suren, Dong, Bin, and USDOE. FasTensor (FT) v0.0.1. Computer software. https://www.osti.gov//servlets/purl/1777894. USDOE. 10 Feb. 2021. Web. doi:10.11578/dc.20210417.1.
5. Koziol, Quincey, Byna, Suren, Hou, Kai-Yuan, and USDOE. TaskWorks v1.0. Computer software. https://www.osti.gov//servlets/purl/1773602. USDOE. 21 Sep. 2020. Web. doi:10.11578/dc.20210401.2.
6. Tang, Houjun, Koziol, Quincey, Byna, Suren, and USDOE. Asynchronous I/O VOL Connector (AsyncVOL) v0.1. Computer software. https://www.osti.gov//servlets/purl/1648028. USDOE. 1 Aug. 2020. Web. doi:10.11578/dc.20200818.23.
7. Suren Byna et al., "ExaIO project" in the "ECP Software Technology Capability Assessment Report", edited by Heroux, Michael A., McInnes, Lois Curfman, Thakur, Rajeev, Vetter, Jeffrey S., Li, Xiaoye Sherry, Aherns, James, Munson, Todd, and Mohror, Kathryn, 2020. Web. doi:10.2172/1760096.
8. Suren Byna et al. (ExaIO team), contributions to the "2020 Exascale Computing Project Annual Meeting (Executive Summary Report)" edited by Marta Garcia, and Todd Munson, https://doi.org/10.2172/1649172.
9. Byna, Suren, Chen, Yong, Zhang, Wei, USDOE, National Science Foundation, National Science Foundation, and National Science Foundation. MIQS v0.6. Computer software. https://www.osti.gov//servlets/purl/1772232. USDOE; National Science Foundation (NSF). 26 Aug. 2020. Web. doi:10.11578/dc.20210322.3.
10. Byna, Suren, Wright, Nicholas, Lockwood, Glenn, Wang, Teng, and USDOE. IOMiner v0.3. Computer software. https://www.osti.gov//servlets/purl/1773093. USDOE. 7 Jul. 2020. Web. doi:10.11578/dc.20210330.4.
11. Byna, Suren, Soumagne, Jerome, Warren, Richard, Mu, Jingqing, Vishwanath, Venkat, Tessier, Francois, Koziol, Quincey, Tang, Houjun, Wang, Teng, Dong, Bin, and Liu, Jialin. Final Technical Report - Proactive Data Containers for Scientific Storage. United States: N. p., 2019. Web. doi:10.2172/1577855.
12. Suren Byna et al., "Proactive Data Containers: An Intelligent Object-Centric Data Management System for HPC", in the "DOE Data Day 2019 Report", edited by Gaylord, Jessie, Ruppert, Stanley, Laney, Daniel, and Abdulla, Ghaleb, 2019. https://www.osti.gov/servlets/purl/1579608, doi:10.2172/1579608
13. Wu, Kesheng, Byna, Surendra, Dong, Bin, and USDOE. VPIC IO utilities. Computer software. https://www.osti.gov//servlets/purl/1487266. USDOE. 7 Dec. 2018. Web. doi:10.11578/dc.20181218.4.
14. Byna, Surendra. BD-CATS-IO. Computer software. https://www.osti.gov//servlets/purl/1459439. Vers. 00. USDOE. 27 Apr. 2017. Web.
15. Wu, Kesheng, Prabhat, Mr, Byna, Surendra, Dong, Bin, and USDOE. Data Elevator. Computer software. https://www.osti.gov//servlets/purl/1564138. USDOE. 30 Apr. 2016. Web. doi:10.11578/dc.20181217.2.
16. Byna, Suren, Dong, Bin, Tang, Houjun, Koziol, Quincey, Mu, Jingqing, Soumagne, Jerome, Vishwanath, Venkat, Warren, Richard, Tessier, François, and USDOE. Proactive Data Containers (PDC) v0.1. Computer software. https://www.osti.gov//servlets/purl/1772576. USDOE. 11 May. 2017. Web. doi:10.11578/dc.20210325.1.
17. Wu, Kesheng, Chou, Jerry, Byna, Suren, Howison, Mark, Bethel, E. Wes, Koziol, Quincey, Cao, Peter, Charawi, Mohamad, Chilan, Christian, McGreevy, Mike, Schuchardt, Karen, Palmer, Bruce, and Prabhat, Mr., "Technical Highlights from the ExaHDF5 project", United States: N. p., 2012. https://www.osti.gov/servlets/purl/1196785

Ph.D. Thesis

Suren Byna, "*Server-based Data Push Architecture for Data Access Performance Optimization*", A Dissertation in the Department of Computer Science, Illinois Institute of Technology, Chicago, IL, December 2006, Advisor: Xian-He Sun