Call for papers: PDSW'24

The 9th International Parallel Data Systems Workshop
http://www.pdsw.org/

November 17, 2024 9:00 AM - 5:30 PM

Held in conjunction with SC24, Atlanta, GA

General Chair
Bing Xie, Microsoft, USA

Program Co-Chairs
Suren Byna, The Ohio State University, USA
Anthony Kougkas, Illinois Institute of Technology, USA

Reproducibility Co-Chairs
Jean Luca Bez, Lawrence Berkeley National Laboratory, USA
Radita Liem, RWTH Aachen University, Germany

Publicity Chair
Qian Gong, Oak Ridge National Laboratory, USA

Web Chair
Joan Digney, Carnegie Mellon University

Important Dates - Paper and AD Deadlines Extended

Regular Papers and Reproducibility Study Papers

Submissions due:	Aug 9th, 2024, 11:59 PM AoE
AD due:	Aug 16th, 2024, 11:59 PM AoE
Paper Notification:	Sep 6th, 2024, 11:59 PM AoE
Camera ready due:	Sep 27th, 2024, 11:59 PM AoE
Final AD/AE due:	Oct 15, 2024, 11:59 PM AoE

Work in Progress (WIP)

Submissions due:	Sept 13th, 2024, 11:59PM AoE
WIP Notification:	On or before Sept 21st, 2024

We are excited to announce the 9th International Parallel Data Systems Workshop (PDSW'24), to be held in conjunction with SC24: The International Conference for High Performance Computing, Networking, Storage, and Analysis, in Atlanta, GA. PDSW'24 builds upon the rich legacy of its predecessor workshops, the Petascale Data Storage Workshop (PDSW, 2006–2015) and the Data Intensive Scalable Computing Systems (DISCS, 2012–2015) workshop. Since their successful merger in 2016, the joint workshop has drawn an average of 200 attendees annually.

The increasing importance of efficient data storage and management continues to drive scientific productivity across traditional simulation-based HPC environments and emerging Cloud, AI/ML, and Big Data analysis frameworks. Challenges are compounded by the rapidly expanding volumes of experimental and observational data, the growing disparity between computational and storage hardware performance, and the rise of novel data-driven algorithms in machine learning. This workshop aims to advance research and development by addressing the most pressing challenges in large-scale data storage and processing.

We invite the community to contribute original research manuscripts that introduce and evaluate novel algorithms or architectures, share significant scientific case studies or workloads, or assess the reproducibility of previously published work. We emphasize the importance of community collaboration for problem identification, workload capture, solution interoperability, standardization, and shared tools. Authors are encouraged to provide comprehensive experimental environment details (software versions, benchmark configurations, etc.) to promote transparency and facilitate collaborative progress.

Topics of Interest:
- **Scalable Architectures**: Distributed data storage, archival, and virtualization.
- **New Data Processing Models and Algorithms**: Application of innovative data processing models and algorithms for parallel computing and analysis.

- **Performance Analysis**: Benchmarking, resource management, and workload studies.
- **Cloud and Container-Based Models**: Enabling cloud and container-based frameworks for large-scale data analysis.
- **Storage Technologies**: Adaptation to emerging hardware and computing models.
- **Data Integrity**: Techniques to ensure data integrity, availability, reliability, and fault tolerance.
- **Programming Models and Frameworks**: Big data solutions for data-intensive computing.
- **Hybrid Cloud Data Processing**: Integration of hybrid cloud and on-premise data processing.
- **Cloud-Specific Opportunities**: Data storage and transit opportunities specific to cloud computing.
- **Storage System Programmability**: Enhancing programmability in storage systems.
- **Data Reduction Techniques**: Filtering, compression, and reduction techniques for large-scale data.
- **File and Metadata Management**: Parallel file systems, metadata management at scale.
- **In-Situ and In-Transit Processing**: Integrating computation into the memory and storage hierarchy for in-situ and in-transit data processing.
- **Alternative Storage Models**: Object stores, key-value stores, and other data storage models.
- **Productivity Tools**: Tools for data-intensive computing, data mining, and knowledge discovery.
- **Data Movement**: Managing data movement between compute and data-intensive components.
- **Cross-Cloud Data Management**: Efficient data management across different cloud environments.
- **AI-enhanced Systems**: Storage system optimization and data analytics using machine learning.
- **New Memory and Storage Systems**: Innovative techniques and performance evaluation for new memory and storage systems.

Regular Paper Submissions

All submissions to the PDSW'24 will undergo a rigorous double-anonymous peer review process overseen by the workshop program committee. Successful submissions will be published in the SC24 Workshop Proceedings and featured on the workshop website alongside associated talk slides.

Template and Submission

- A full paper up to **6 pages** in length, excluding references and AD/AE appendices.
- Artifact Description (AD) Appendix is mandatory and Artifact Evaluation (AE) Appendix is optional.
  - Submissions with AD and AE Appendix will be considered favorably for the PDSW Best Paper award.
- Papers must adhere to the IEEE proceedings template. Download it here.
- Submit your papers by **Aug 2nd, 2024 , 11:59 PM AoE** to: https://submissions.supercomputing.org/

Reproducibility Initiative

Aligned with the SC24 Reproducibility Initiative, we encourage detailed and structured artifact descriptions (AD) using the SC24 format. The AD should include a field for one or more links to data (zenodo, figshare, etc.) and code (Github, GitLab, Bitbucket, etc.) repositories. For the artifacts that will be placed in the code repository, we encourage authors to follow the PDSW 2024 Reproducibility Addendum on how to structure the artifact, as it will make it easier for the reviewing committee and readers of the paper in the future.

Work in Progress (WIP) Session

The WIP session will showcase brief 5-minute presentations on ongoing work that may not yet be ready for a full paper submission. WIP papers will not be included in the proceedings. A one-page abstract is required for participation. WIP Submissions Due: September 13th, 2024, 11:59 PM AoE