SC22
Dallas, TX | hpc accelerates.

PDSW 2022

7th International Parallel Data Systems Workshop

Kento Sato, General Chair
Amelie Chi Zhou, Program Co-Chair
Bing Xie, Program Co-Chair

Welcome !

- The goal of PDSW is to facilitate research that addresses the most critical challenges in scientific data storage, data processing and data analytics

- In 2016, PDSW (Parallel Data Systems Workshop) was established by combining two predecessor workshops
  - The Petascale Data Storage Workshop (PDSW, 2006-2015)
  - The Data Intensive Scalable Computing Systems workshop (DISCS 2012-2015)

- This joint workshop, PDSW, brings together experts from several overlapping communities from HPC, Big data and Data analytics
  - It has been continuing for 6 years from 2016 to 2022

What's new in PDSW 2022 ?

- PDSW2022 extends the scope to align with new technologies:
  - The application of new data processing models and algorithms
  - Hybrid cloud/on-premise data processing, storage and its management
  - Storage system optimization and data analytics with AI
  - Innovative techniques and performance evaluation for new memory and storage systems

This wouldn’t be possible without the combined efforts of this year’s workshop team:

Kento Sato:	General Chair
Amelie Chi Zhou:	Program Co-Chair
Bing Xie:	Program Co-Chair
Carlos Maltzahn and Alexandru Uta:	Reproducibility Co-Chairs
Jean Luca Bez:	Publicity Chair
Joan Digney:	Web and Publications Chair
Jay Lofstead:	Steering Committee Chair
Dean Hildebrand:	Steering Committee Vice Chair

And of course: thank you to everyone who contributed research papers and WIP presentations for sharing your work with the community!

Program highlights
(Full program at pdsw.org)

- Invited Talk
  - Splinters - Distributed IO Sampling for Cloud Data Centers – Design and Applications

    - by Arif Merchant, Research Scientist with the Storage Analytics group at Google

- Technical presentations
  - 5 full paper presentations (10 submissions; 50% of acceptance rate)
    - 4 reviewers are assigned to each paper and reviewed
    - Final decisions are made during online PC meeting
  - 3 work in progress (WIP) presentations (3 submissions; 100% of acceptance rate)
    - WIP abstracts are reviewed by program co-chair

- Virtual participation available through SC22 Digital Experience and Slido

We also owe a big thanks to the program committee:

These subject matter experts are not just gatekeepers for PDSW. They provide constructive guidance to our community to help make our research stronger.

Jalil Boukhobza, University of Western Brittany, France

Suren Byna, Lawrence Berkeley National Laboratory

Yong Chen, Texas Tech University

Wei Der Chen, University of Edinburgh

Dong Dai, University of North Carolina at Charlotte

Matthieu Dorier, Argonne National Laboratory (ANL)

Bogdan Ghit, Databricks

Qian Gong, Oak Ridge National Laboratory

Luanzheng Guo, Pacific Northwest National Laboratory

Shadi Ibrahim, Inria

Tanzima Islam, Texas State University

Youngjae Kim, Sogang University

Johann Lombardi, Intel Corporation

Xiaoyi Lu, University of California, Merced

Xiaosong Ma, Qatar Computing Research Institute

Kathryn Mohror, Lawrence Livermore National Laboratory

Diana Moise, Hewlett Packard Enterprise

Sarah Neuwirth, Habilitation Candidate at Goethe University

M. Mustafa Rafique, Rochester Institute of Technology

Raghunath Raja Chandrasekar, Stealth Startup

Michael Schöttner, Duesseldorf University

Vasily Tarasov, IBM Corporation

Qing Zheng, Los Alamos National Lab

We are also grateful to

Rio Yokota and Tjerk P. Straatsma:	Workshop Chairs
Leah Glick and Taylor Carr:	Support team behind Linklings
Patrick Kellenberger:	IEEE Computer Society

Invited Talk

7th International Parallel Data Systems Workshop

Arif Merchant,
Google

Splinters - Distributed IO Sampling for Cloud Data Centers – Design and Applications