



SC18

Dallas, TX | **hpc**
inspires.

The International Conference for
High Performance Computing,
Networking, Storage and Analysis
Dallas, TX
November 11 - 16, 2018

Call for papers: PDSW-DISCS'18

**3rd Joint International Workshop
on Parallel Data Storage & Data
Intensive Scalable Computing Systems**

Monday, November 12, 2018

9:00am - 6:00pm

Held in conjunction with SC18

*Program Co-Chairs: Suzanne McIntosh, NYU;
Raghunath Raja Chandrasekar, Amazon Web Services
General Chair: Kathryn Mohror, LLNL*

Workshop Abstract:

We are pleased to announce that the 3rd Joint International Workshop on Parallel Data Storage and Data Intensive Scalable Computing Systems (PDSW-DISCS'18) will be hosted at SC18: The International Conference for High Performance Computing, Networking, Storage and Analysis. This one day joint workshop combines two overlapping communities to better promote and stimulate researchers' interactions to address some of the most critical challenges for scientific data storage, management, devices, and processing infrastructure for both traditional compute intensive simulations and data-intensive high performance computing solutions. Special attention will be given to issues in which community collaboration can be crucial for problem identification, workload capture, solution interoperability, standards with community buy-in, and shared tools.

Many scientific problem domains continue to be extremely data intensive. Traditional high performance computing (HPC) systems and the programming models for using them such as MPI were designed from a compute-centric perspective with an emphasis on achieving high floating point computation rates. But processing, memory, and storage technologies have not kept pace and there is a widening performance gap between computation and the data management infrastructure. Hence data management has become the performance bottleneck for a significant number of applications targeting HPC systems. Concurrently, there are increasing challenges in meeting the growing demand for analyzing experimental and observational data.

In many cases, this is leading new communities to look towards HPC platforms. In addition, the broader computing space has seen a revolution in new tools and frameworks to support Big Data analysis and machine learning. There is a growing need for convergence between these two worlds. Consequently, the U.S. Congressional Office of Management and Budget has informed the U.S. Department of Energy that new machines beyond the first exascale machines must address both traditional simulation workloads as well as data intensive applications. This coming convergence prompted the integration of the PDSW and DISCS workshops into a single entity to address the common challenges. The scope of the new joint PDSW-DISCS workshop is summarized as:

- Scalable storage architectures, archival storage, storage virtualization, emerging storage devices and techniques
- Performance benchmarking, resource management, and workload studies from production systems including both traditional HPC and data-intensive workloads
- Programmability, APIs, and fault tolerance of storage systems
- Parallel file systems, metadata management, and complex data management, object and key-value storage, and other emerging data storage/retrieval techniques
- Programming models and frameworks for data intensive computing including extensions to traditional and nontraditional programming models, asynchronous multi-task programming models, or to data intensive programming models
- Techniques for data integrity, availability and reliability
- Productivity tools for data intensive computing, data mining and knowledge discovery
- Application or optimization of emerging "big data" frameworks towards scientific computing and analysis
- Techniques and architectures to enable cloud and container-based models for scientific computing and analysis
- Techniques for integrating compute into a complex memory and storage hierarchy facilitating in situ and in transit data processing
- Data filtering/compressing/reduction techniques, maintaining scientific validity for large scale compute-intensive workloads
- Tools and techniques for managing data movement among compute and data intensive components both solely within the computational infrastructure as well as incorporating the memory/storage hierarchy

Paper Submissions:

This year, we are soliciting two categories of papers, regular papers and reproducibility study papers. Both will be evaluated by a competitive peer review process under the supervision of the workshop program committee. Selected papers and associated talk slides will be made available on the workshop web site. The papers will also be published in the digital libraries of the IEEE and ACM.

Regular Paper Submissions

We invite regular papers which may optionally undergo validation of experiment results by providing reproducibility information. Papers successfully validated earn a badge in the ACM DL in accordance with ACM's artifact evaluation policy.

NEW! Reproducibility Study Paper Submissions

We also call for reproducibility studies that for the first time reproduce experiments from papers previously published in PDSW-DISCS or in other peer-reviewed conferences with similar topics of interest. Reproducibility study submissions are selected by the same peer-reviewed competitive process as regular papers, except these papers undergo validation of the reproduced experiment and must include reproducibility information that can be evaluated by a provided automation service. Successful validation earns the original publication a badge in the ACM DL in accordance with ACM's artifact evaluation policy.

Guidelines for Regular Papers and Reproducibility Study Papers

Submit a not previously published paper as a PDF file, indicate authors and affiliations. Papers must be at least 8 pages long and no more than 12 pages long (including appendices and references). Download the IEEE conference paper template. Download the IEEE template at: <https://www.ieee.org/conferences/publishing/templates.html>. Details on reproducibility will be available on the website by July 1, 2018.

Work-in-progress (WIP) Submissions

There will be a WIP session where presenters provide brief 5-minute talks on their on-going work, with fresh problems/solutions. WIP content is typically material that may not be mature or complete enough for a full paper submission. A one-page abstract is required.

Important Dates

Regular Papers and Reproducibility Study Papers due:

Submissions due: Sep. 2, 2018, 11:59 PM AoE

Paper Notification: Sep. 30, 2018

Camera ready due: Oct. 5, 2018

Slides due: Nov. 9, 2018, 3:00 pm CST

Work in Progress (WIP) submissions due:

Submissions due: Nov. 1, 2018, 11:59 PM AoE

WIP Notification: Nov. 7, 2018