

HPC I/O in the Data Center Workshop (HPC-IODC)

Julian M. Kunkel, Jay Lofstead, Jean-Thomas Acquaviva

¹ Georg-August-Universität Göttingen
Göttingen, Germany

`julian.kunkel@gwdg.de`

² Center for Computing Research
Sandia National Laboratories
Albuquerque, USA

`gflofst@sandia.gov`

³ Data Direct Networks, France
`jtacquaviva@ddn.com`

1 Introduction

Many public and privately funded data centers host supercomputers for running large scale simulations and analyzing experimental and observational data. These supercomputers run usually tightly coupled parallel applications that require hardware components that deliver the best performance. In contrast, commercial data centers, such as Facebook and Google, execute loosely coupled workloads with a broad assumption of regular failures. The dimension of the data centers is enormous. A 2013 article summarizes commercial data centers' dimensions [4]. It estimates, for example, that Facebook hosts around 100 PB of storage, and Google and Microsoft manage around 1 million servers each – although the hardware is split among several physical data centers – a modus operandi not suitable for HPC centers. With the increasing importance of using machine learning to reveal underlying patterns in data, the data storage rates are accelerating to feed these additional use cases. Combining traditional modeling and simulation with ML workloads yields both a write and read-intensive workload for a single workflow.

Management of the huge amount of data is vital for the effective use of the contained information. However, with limited budgets, it is a daunting task for data center operators, especially as the design and storage system required hardware depends heavily on the executed workloads. A co-factor of the increasing difficulty is the increase in complexity of the storage hierarchy with the adoption of SSD and memory class storage technology. The US Department of Energy recognizes the importance of data management, listing it among the top 10 research challenges for Exascale [3].

There are several initiatives, consortia and special tracks in conferences that target RD&E audiences. Examples are the Storage Networking Industry Association (SNIA) for enterprises, the Big Data and Extreme-Scale Computing

(BDEC) initiative¹, the Parallel Data Systems Workshop (PDSW) and the HEC FSIO workshop [1].

There are many I/O workloads studies and performance analysis reports for parallel I/O available. Additionally, many surveys of enterprise technology usage include predictions of analysis for future storage technology and the storage market [2]. However, the analysis conducted for HPC typically focuses on applications and not on the data center perspective. Information about data center operational aspects is usually described in file system-specific user groups and meetings or described partially in research papers as part of the evaluation environment.

In the HPC IODC workshop, we bring together I/O experts from data centers and application workflows to share current practices for scientific workflows, issues, and obstacles for both hardware and the software stack, and RD&E to overcome these issues.

Due to the COVID-19 crisis, the ISC HPC conference changed to a digital edition. We preserved the nature of the workshop and organized it as a virtual full-day meeting on the regular workshop day with minimal changes to the agenda. The morning session was served in BigBlueButton while the afternoon session was part of the official ISC HPC conference program and served with Zoom.

2 Organization of the Workshop

The workshop was organized by

- Julian Kunkel (*Georg-August-Universität Göttingen, Germany*)
- Jay Lofstead (*Sandia National Labs, USA*)
- Jean-Thomas Acquaviva (*DDN*)

The workshop is supported by the Centre of Excellence in Simulation of Weather and Climate in Europe (ESiWACE), the Virtual Institute for I/O (VI4IO)² and the Journal of High-Performance Storage (JHPS)³.

The workshop covered the following tracks:

- **Research paper presentations** – authors needed to submit a paper regarding relevant state-of-the-practice or research for I/O in the datacenter.
- **Talks from I/O experts** – authors needed to submit a rough outline for the talk related to the operational aspects of the data center.
- **Student Mentoring Session** – students need to submit an abstract for their PhD topic.
- Moderated **discussion** for hot topics to identify key issues and potential solutions in the community.

¹ <http://www.exascale.org/bdec/>

² <http://vi4io.org>

³ <https://jhps.vi4io.org/>

This year, we broadened our scope by including the student mentoring session. To foster the next generation of data-related practitioners and researchers, students are encouraged to submit an abstract aligned with the workshop topics. At the workshop, the students give a lightning talk about what they are working on followed by feedback of the community about how to further the work, what the impact could be, alternative research directions, and other topics to help the students progress in their studies.

The CFP has been issued at the beginning of January. Important deadlines were:

- Submission deadline: 2021-03-24 AoE
- Author notification: 2021-04-24
- Workshop: 2021-07-02

2.1 Programm Committee

- Thomas Boenisch (*High-Performance Computing Center Stuttgart*)
- Suren Byna (*Lawrence Berkeley National Laboratory*)
- Matthew Curry (*Sandia National Laboratories*)
- Philippe Deniel (*CEA*)
- Sandro Fiore (*University of Trento*)
- Wolfgang Frings (*Juelich Supercomputing Centre*)
- Javier Garcia (*Blas Carlos III University*)
- Stefano Gorini (*Swiss National Supercomputing Centre*)
- Adrian Jackson (*The University of Edinburgh*)
- Ivo Jimenez (*University of California, Santa Cruz*)
- Anthony Kougkas (*Illinois Institute of Technology*)
- Glenn Lockwood (*Lawrence Berkeley National Laboratory*)
- Carlos Maltzahn (*University of California, Santa Cruz*)
- George S. Markomanolis (*Oak Ridge National Laboratory*)
- Sandra Mendez (*Barcelona Supercomputing Center*)
- Robert Ross (*Argonne National Laboratory*)
- Feiyi Wang (*Oak Ridge National Laboratory*)
- Xue Wei (*Tsinghua University*)
- Bing Xie (*Oak Ridge National Lab*)

3 Workshop Summary

Over the full-day program, about 35 attendees were constantly connected to the virtual session. In spite of the workshop being held online squarely on European Time zone making attending from North America more difficult. This is in line with the in-person attendance at previous instances and included many North American attendees, including some from the American west coast.

We had a good mix of talks from I/O experts, data center relevant research, and two discussion sessions. A short summary of the presentations is given in

the following. The slides and video recordings of the presentations are available on the workshop's webpage: <https://hps.vi4io.org/events/2021/iodc>.

The morning session covered the three accepted research papers and the four expert talks. For the latter, systems and challenges were discussed for the institutions in CERN, EPCC, CSC and LuxProvide. These research papers and expert talks addressed a wide scope of storage and data issue. For instance new data format such as CERN's RNTuple or the difficulties related to performance evaluation. LuxProvide with its newly established Meluxina system received numerous questions about its architecture, hinting that this system, the greenest in Europe, is triggering a lot of interest.

In the afternoon session, a panel about "The impact of HPC and Cloud convergence on storage" involved five speakers from Amazon, MetOffice, National University of Singapore, Linksfoundation and ECMWF. The panel was structured around individual presentations followed by live discussion based on audience's questions. The panelists acknowledged the on-going convergence between HPC and Cloud, but nevertheless underscored some key differences. Among these differences the workloads diversity in Cloud which brings additional value for heterogeneity in large Cloud Data Center. The notion of Cloud and HPC reflect as well the self-perception of data center: interestingly MetOffice starts a new partnership with a major cloud provider, in order to focus on its core task, weather prediction, without being diverted by the need to host a service. Whereas ECMWF see self hosting as a key part of their success.

The session was followed by the new student mentoring session. In this session, four students presented their PhD works. We believe the feedback from attendees were constructive and received positive feedback from the students and attendees about the session. Therefore, we will include it in the next year's workshop again.

The major distinguishing feature for this workshop compared to other venues is the discussion rounds. The opportunity for themed, open discussions about issues both pressing and relevant to the data center community facilitates sharing experiences, solutions, and problems.

Albeit the workshop was virtual, the discussion covered aspects around use-cases and reasons for doing I/O, standardization of APIs as alternatives for POSIX, i.e., would S3 be sufficient or an object storage API with KV. We also continued the discussion around HPC and cloud convergence.

References

1. Bancroft, M., Bent, J., Felix, E., Grider, G., Nunez, J., Poole, S., Ross, R., Salmon, E., Ward, L.: Hec fsio 2008 workshop report. In: High End Computing Interagency Working Group (HECIWG), Sponsored File Systems and I/O Workshop HEC FSIO (2009)
2. IDC: Enterprise storage services survey. <http://www.idc.com/getdoc.jsp?containerId=254468>

3. Lucas, R., committee members: Top ten exascale research challenges. Online <http://science.energy.gov/~media/ascr/ascac/pdf/meetings/20140210/Top10reportFEB14.pdf> (Feb 2014)
4. StorageServers Blog: Facts and stats of world's largest data centers. Online <https://storageservers.wordpress.com/2013/07/17/facts-and-stats-of-worlds-largest-data-centers/> (July 2013)