HPC

I/ODC

I/O in the Data Center

Workshop

HPS
https://hps.vi4io.org
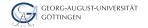
Julian M. Kunkel (University of Göttingen/GWDG)
Jay Lofstead (Sandia National Laboratories)
Jean-Thomas Acquaviva (DDN)

9th HPC I/O in the Data Center Workshop

Sponsoring

The workshop is powered by

The Virtual Institute for IO

The Journal of
High-Performance Storage

Julian M. Kunkel (University of Göttingen/GWDG) , Jay Lofstead (Sandia National Laboratories), Jean-Thomas Acquaviva (DDN)

3/10

The Virtual Institute for I/O

Goals

- Provide a platform for I/O enthusiasts for exchanging information
- Foster training and collaboration in the field of high-performance I/O
- Track and encourage the deployment of large storage systems by hosting information about high-performance storage systems

https://www.vi4io.org

The Journal of High-Performance Storage (ISSN 2748-7814)

Features

■ Open reviews, i.e., anyone can provide feedback

■ Living papers, i.e., can improve over time

■ Digital replicability (of analysis/experiments)

■ Free open access

https://jhps.vi4io.org

Status

■ Two issues with 3 papers published

■ If you'd be interested to publish? Why not?

Motivation for the Workshop

- I/O perspective of centers is often ignored
- Data centers aim to provide optimal service and performance

Providing a good storage strategy is challenging

- Zoo of emerging storage solutions/technology
- Variety of HPC file systems have pro/cons
- Management of large volume/file numbers (PByte/Billions of files)
- Middleware to fix file system issues present in all file systems
- Data center needs to consider IO interaction of applications/workflows

Understanding Systems and Users

Knowing the behavior would allow to provide a better system

- A perfect understanding of usage and efficiency would allow for
  - ▶ selection of the right storage technology
  - ▶ gearing optimization effort towards mostly used I/O libraries
  - ▶ understanding the requirements for the procurement
  - ▶ optimizing the data center's efficiency as a whole
- But users often don't know their I/O patterns (and workflows)
- The I/O stack is challenging even for experts

Maybe I/O experts from data centers can make a difference

- From **individual** activity towards **community** effort and ultimately useful **conventions**

Julian M. Kunkel (University of Göttingen/GWDG) , Jay Lofstead (Sandia National Laboratories), Jean-Thomas Acquaviva (DDN)

7/10

About the HPC-IODC Workshop

Goal: Bring together I/O experts from data centers

- Regardless of file system
- Foster information exchange
- Opportunity for networking

Topics of interest

- Scientific workload
- Usage characteristics (file, folders, scientific libraries)
- System perspective
- Architecture
- Performance aspects and monitoring
- Issues during production and potential solutions

Workshop

Dissimination of results

■ We are working on a collaborative notebook https://s.gwdg.de/XAtczP

■ We will write a preface and summarize the workshop results

Interactivity

■ Critical discussion is welcome and expected from attendees

■ Discussion time slots: open topics, everyone can raise/discuss issues
   ▶ Ultimately controlled by a moderator

Agenda

Morning session

9:00 **Welcome**

9:15 **Storage semantics**

11:00 *Coffee break*

11:30 **Expert talks**

13:00 *Lunch Break*

Afternoon

14:00 **Expert talks**

16:00 *Coffee break*

16:30 **Expert talks**

17:30 *Discussion of hot topics*

18:00 *Farewell*