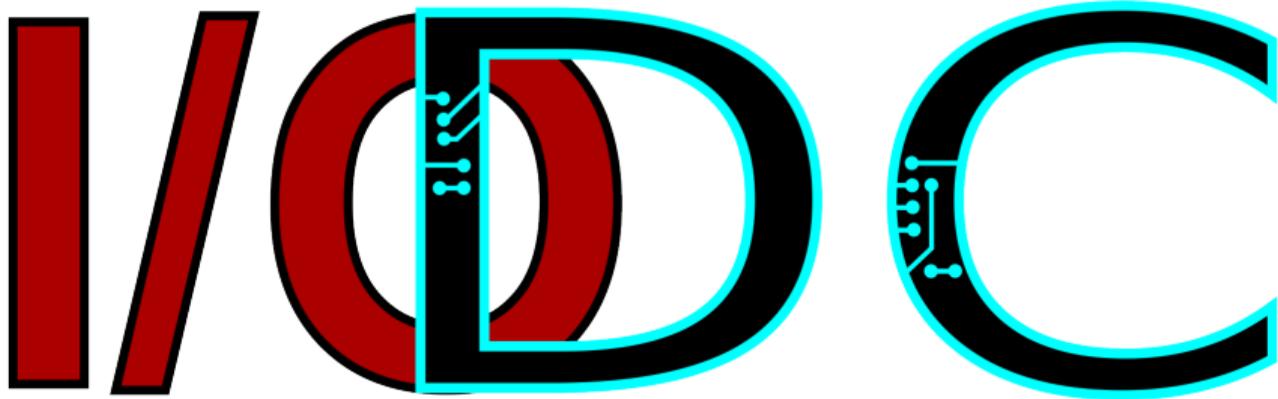


HPC



I/O in the Data Center

Workshop

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

## 8th HPC I/O in the Data Center Workshop

**HPS**

<https://hps.vi4io.org>



# Sponsoring

The workshop is powered by



esiwace

CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER  
AND CLIMATE IN EUROPE

vi4io

The Virtual Institute for IO

J∞HPS

The Journal of  
High-Performance Storage

## EU funded Project: ESiWACE

### The Centre of Excellence in Simulation of Weather and Climate in Europe

- Representing the European community for
  - ▶ climate modelling and numerical weather simulation
- Goals in respect to HPC environments:
  - ▶ Improve efficiency and productivity
  - ▶ Supporting the end-to-end workflow of global Earth system modelling
  - ▶ Establish demonstrator simulations that run at highest affordable resolution
- Funding via the European Union's Horizon 2020 program (grant #823988)

<http://esiwace.eu>



**esiwace**

CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER  
AND CLIMATE IN EUROPE

# 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>



# Outlook for VI4IO

## Selection of ongoing activities

- Collaboration with the HPC Certification Forum
  - ▶ <https://hpc-certification.org>
  - ▶ Ensure that IO-related competences are properly represented in the skill-tree
- Capturing file-system specific characteristics properly
  - ▶ For the high-performance storage list and the IO500
  - ▶ Discussion at the later talk

## The Goals of the HPC Certification Forum (HPCCF)

- Fine-grained standardizing HPC knowledge representation
- Establishing international certificates attesting knowledge
- Supporting an ecosystem around the HPC competences

# 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 be on the editorial board or to publish?

# 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**

# 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 **Speed introduction: 1-2 sentences + expectations**

9:30 **Systems**

11:00 *Coffee break*

11:30 **Expert talks**

13:00 *Lunch Break*

## Afternoon

14:00 **Performance**

16:00 *Coffee break*

16:30 **Student mentoring session /  
Research talks**

17:30 *Discussion of hot topics*

18:00 *Farewell*

The ESiWACE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No **823988**



*Disclaimer: This material reflects only the author's view and the EU-Commission is not responsible for any use that may be made of the information it contains*