Understanding I/O Performance Behavior (UIOP) Workshop

Julian M. Kunkel

Deutsches Klimarechenzentrum GmbH (DKRZ)

2017-03-22

DKRZ

About DKRZ

German Climate Computing Center

DKRZ – Partner for Climate Research

High Performance Computing.
Sophisticated Data Management.
Competent Service.

Group Wissenschaftliches Rechnen (Scientific Computing)

Composed of DKRZ research division and Universität Hamburg research group

Research

- Analysis of parallel I/O
- I/O & energy tracing tools
- Middleware optimization

- Alternative I/O interfaces
- Data reduction techniques
- Cost & energy efficiency

Goal of the Workshop

Support development of tools to identify (in-)efficient usage of I/O resources

- From the perspective of users and data centers

Approach of the workshop:

1. Understanding future storage architectures and implications on workflows
2. Introduce monitoring tools to enable optimization of system/applications
   - Identify and quantify I/O inefficiencies of **observed workloads**
3. Benchmarking provides (optimal) baselines to understand storage better
   - Also enables comparison between storage systems and sites
   - Mandatory tool for procurements and steering of vendor effort

Support

This workshop is supported by:

And powered by:

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 #675191)

http://esiwace.eu

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER
AND CLIMATE IN EUROPE

Introduction: VI4IO

Goals of the Virtual Institute for I/O

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

Web page: `http://www.vi4io.org`

Introduction

Philosophical cornerstones of the institute

- To allow participation of everybody without a membership fee
- To treat every member and participant equally
- To be an independent organization
    - Independent of vendors and research facilities

Open Organization

- The organization uses a wiki as central hub
    - Everybody (registered users) can edit the content
    - Mayor changes should be discussed (see below)
    - The wiki uses tag clouds to link between similar entities
- Supported by mailing lists
    - Call-for-papers
    - Announce list for relevant information
    - Contribute list to discuss and steer organizational issues
    - IO-500 (development of a benchmark for the IO-500 list)
- Mayor changes should be discussed on the contribute mailing list
- Members can vote for changes

Everybody is welcome to participate

The High-Performance Storage List (HPSL)

- Aims to track storage characteristics over time
  - Uses a component model that covers: site, supercomputers and storage
- Provides means to explore data
- Is community maintained

Site characteristics

Key	Value	Rf
Institution	DKRZ	
Nationality	DE	
Web page	🌐 https://www.dkrz.de/	
Energy consumption	2.00 MW	
Power usage effectiveness	1.04	

System architecture

DKRZ hosts the Mistral supercomputer which is tightly coupled with two Lustre file systems. We have some small compute and supporting infrastructure that may access the storage of Mistral and a large HPSS system.

Supercomputers

- Mistral

Storage systems

- HPSS
- Lustre02
- Lustre01

2017

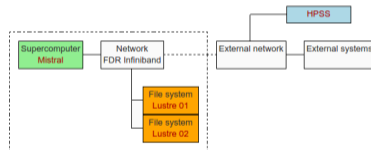| #	Site		Supercomputer			Storage	
	Name	nationality	Name	compute_peak	memory_capacity	Name	capacity ↑
				in PFLOPs	in TiB		in PiB
1	LANL	US	Trinity	11.08	1.00	Lustre	72.83
2	DKRZ	DE	Mistral	3.12	204.00	HPSS Lustre02 Lustre01	52.00
3	LLNL	US	Sequoia	20.10	1.00	Lustre	48.85
4	RIKEN	JP	K Computer	10.62	1.00	Lustre FEFS	39.77
5	NERSC	US	Cori Phase I	4.90	204.00	Lustre	30.00
6	ORNL	US	Titan	27.10	645.74	Lustre	28.00
7	NCSA	US	Blue Waters	13.40	1,500.00	Lustre HPSS	26.40
8	ANL	US	Mira	10.00	698.49	GPFS	21.32
9	JSC	DE	Juqueen	5.90	407.45	HPSS JUST	20.30
10	JAMSTEC	JP	Earth Simulator	1.31	291.00	Archive Home Data Work	19.62

...

| 30 | CSCS | CHE | Piz Daint | 7.79 | 153.70 | Lustre | 2.22 |

Tracking Data Enables Analysis

- On 30 systems that are currently in the list
- Corr. storage capacity &
  - memory capacity = 0.63
  - compute peak = 0.13
- Mean(storage/mem capacity) = 45.6
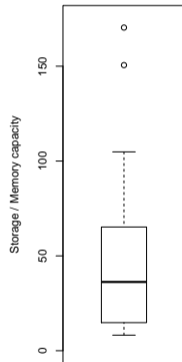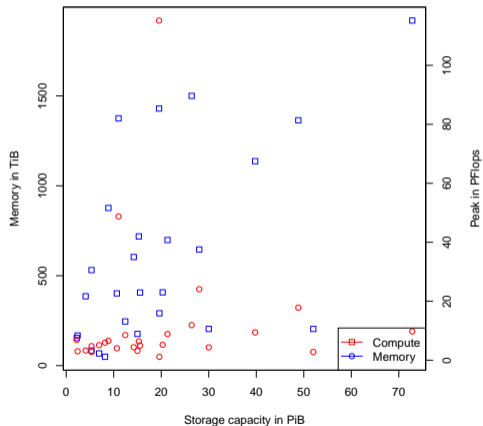
Agenda

Wednesday

- Architectures
- Discussion: Architectures
- *Social Event*

Thursday

- Monitoring
- Discussion: Monitoring
- *Lunch*
- Benchmarking
- Discussion: Benchmarking
- *Farewell*

Social Event

- Location: Rickmer Rickmers at the Harbor (see webpage for the location)
- When:
    - 19:00 reception with finger food
    - 19:30 buffet
    - Museum attendance covered (opens until 22:00)
    - You can attend the museum even at 18:30
- How: We will depart at 18:30 from DKRZ main entry using U3 (Schlump)
    - Towards Landungsbrücken
- Tickets: You have to pay for yourself...
    - Single short trip ticket: 1.60 Euro
    - Group daily ticket (5 people): 11.60 Euro
    - We will wait at Schlump until you have bought your tickets

Discussion

- A discussion slot follows every topic block
  - The appointed moderator leads the discussion
- Add questions and relevant issues to our Google Doc
  https://docs.google.com/document/d/
  1fPMzSRepwwSgNBh7SGcmUTaN1mYR0oPVhd-mRKjY81U/edit
- Do not be shy ... simply add stuff !

Results of the Workshop

A PDF with the results will be published on the web page

- Summarizing the presentations
- Summarizing the discussions
- We will use a simple Google Doc
    - Everyone is welcome to summarize the own talk / add something
    - I will later convert it to LaTeX and put it on the web page
    - Yes, it is embedded at the end of the discussion document !
- Feel free to edit the document during the workshop
    - I will make notes in the document during the workshop

Further Advertisement

Upcomming events

- 2017-06-22 **HPC I/O in the Data Center Workshop**
  https://wr.informatik.uni-hamburg.de/events/2017/iodc
  During the ISC-HPC workshop day in Frankfurt !
  Submission deadline: 2017-03-28 (extension possible)

- BoF: The Virtual Institute of I/O and the I/O 500
  During ISC-HPC, Frankfurt, 2017-06-18

- 2017-09-25/26 Workshop: **Exascale I/O for Unstructured Grids (EIUG)**
  https://wr.informatik.uni-hamburg.de/events/2017/eiug
  Will take place at DKRZ