OPEN SOURCE SOFTWARE FOR SCIENTIFIC AND PARALLEL COMPUTING

Open Source SG

Virtual meeting starting shortly...

11 June 2020

Future Meetings

Thursday 16th July

FPGA evening

Monday 20th July

London RISC-V Virtual Meetup

Thursday 20th August (To be confirmed)

Open Source Facial Recognition

Thursday 17th September

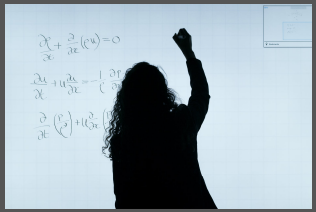
Women in Open Source

Computer Science Department

HPS

University of Reading

Open Source Software for Scientific and Parallel Computing



Limitless Storage
Limitless Possibilities

https://hps.vi4io.org

Julian M. Kunkel

BCS Open Source Specialist Group Meeting

2020-06-11

LIMITLESS **POTENTIAL** | LIMITLESS **OPPORTUNITIES** | LIMITLESS **IMPACT**

Computational Science

University of Reading

- *"Applied computer science and mathematics that uses advanced computing capabilities to understand and solve complex problem"* [Wikipedia]
- Modeling and simulation of the laws of nature within computer systems
  - Serves experiment and fuels the refinements of theory in the scientific method
- The computer system is the vehicle for conducting the scientific experiment
  - Some experiments are too expensive, dangerous, infeasible in real world
  - Example: climate models, protein folding, fusion reaction, car engines, ...
- Enables the analysis of observational data to mine for knowledge
  - Examples: Cern/LHC, Square Kilometer Array (SKA), sensor networks
- Users: Scientists, PhD candidates, engineers

High-Performance Computing (HPC)

University of Reading

- HPC: Field providing massive compute resources for a computational task
  - Task needs too much memory or time for a normal computer
  - ⇒ Enabler of scientific computing
- Supercomputer: aggregates power of many compute devices
  - Operated in a data center (public often funds centers)

Example Supercomputer

Oak Ridge National Lab: Summit

- Compute: 4,608 dual socket nodes
  - 6 GPUs per Node
  - Linpack: 148 Petaflop/s
- Storage: 200 Petabyte
  - NVM: 1.6 TB/Node
- Energy: 13 MW

Open Source for Supercomputers & Data Centers

University of Reading

- Supercomputers run with Linux
- Majority of software to manage supercomputers is OS
  - Data center has typically support contracts with vendors
- Scientific software uses primarily OS libraries
- Scientific software is often OS (may also be restricted to scientists)
- Some independent software vendors primarily for industry applications

Claim

- OS enables High-Performance Computing and accelerates science

Agenda

- Open Source Software in High-Performance Computing
  Shane Canon, Lawrence Berkeley National Laboratory/NERSC, USA
- Lessons learned from creating/using the Ceph open-source storage system
  Carlos Maltzahn, UC Santa Cruz, USA
- High Performance Computing in a world of Data Science
  Martin Callaghan, University of Leeds, UK