Scientific computing at MPP

MPP Project review 2024, Stefan Kluth, 10.12.2024

G3, MPI für Physik und Astrophysik, 1966

HPC System Viper, MPCDF, 2024

People

- MPP IT Fachabteilung (T. Hahn)
  - Manuel Krämer, Konstantinos Kiriakidis, Katrin Krebs, Uwe Leupold, Yaozhi Pan
- MPP at MPCDF (SK)
  - Cesare Delle Fratte, Meisam Tabriz
- MPP Computing Commission
  - M. Kado, T. Hahn, SK (chair), D. Paneque, O. Schulz, S. Stonjek
  - Meetings (generally) public

Overview

Activity

O(1000 or more) batch jobs,
GPU support

O(100) Batch jobs

Programming, interactive work,
E-mail, web, documents, etc

Compute systems

MPCDF systems: VIPER, RAVEN
MPP cluster at MPCDF

MPP cluster at MPCDF
MPP condor batch jobs

MPP desktop PCs, few developer
machines w/ GPU and large RAM,
MPCDF RVS

MPCDF

VIPER: 768 nodes dual AMD EPYC 9554 (128c), 512 (768, 1024, 2304) GB RAM, 340 nodes dual AMD MI300A GPUs, Infiniband 200/400 Gb/s

RAVEN: 1592 nodes Xeon 8360Y (72c), 256 (512, 2048) GB RAM, +192 nodes four Nvidia A100 GPUs, Infiniband 100/200 Gb/s

MPP cluster at MPCDF

- ~100 (+special) nodes, >3500 cores, 4 GB/core
  - Node groups at, bt, login nodes mppui1, 2, 3, 4
  - 12 nodes dual AMD EPYC 9554 (128c), 768 GB RAM
- Large RAM nodes zt1, zt2 (Henn group):
  - zt1: 6 TB, 192 cores, zt2: 3 TB, 36 cores
- More than 5 PB storage
  - /ptmp/mpp (gpfs), dCache
- CentOS7 → ALMA9, Slurm batch, apptainer, /cvmfs

MPP desktop PCs

- All groups
  - > 200 PCs ≥8 core, 2-4 GB/core, SSD
  - opensuse tumbleweed, condor, /remote/ceph, /cvmfs, apptainer
- Common
  - CEPH storage > 2 PB
  - Few servers with 512-1024 GB RAM, Nvidia GPUs (ODSL, ATLAS, BAT/LEGEND), R&D, local jobs

MPP 1 Linux

- Desktop PCs common Linux OS
  - Rolling release (opensuse tumbleweed)
  - Gnome, KDE, ..., apptainer, /cvmfs, /remote/ceph, condor
- Rolling release Linux
  - Guarantee support of modern hardware
  - Decent and recent desktop software
  - Scientific workflows in apptainer containers, data in /remote/ceph, batch jobs with condor or MPCDF

MPP cluster at MPCDF

New procurements 2024

GPU server (ATLAS):
dual AMD EPYC 9554 (128c), 1536 GB RAM
8 Nvidia H200 GPUs, dual 100 GBE NIC
Delivered, installed, to be configured

CPU servers (Henn and Zanderighi groups):
40 dual AMD EPYC 9754 (256c), 1536 GB RAM, 3.84 TB SSD, dual 100 GBE NIC
1 dual AMD EPYC 9754 (256c), 6144 GB RAM, 3.84 SSD, dual 100 GBE NIC
Delivery 18.12.

Storage

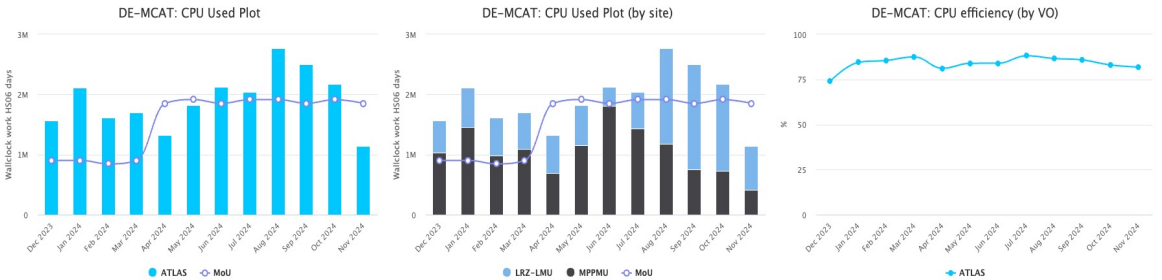
- CEPH at MPP full, performance loss
  - Upgrade existing old nodes: new SSDs, more RAM
  - Join CEPH and bring to normal state
  - Move CEPH in steps to MPCDF
  - Evaluate EOS (CERN) as CEPH replacement
  - Plan for new system BAR application 2025
- Storage at MPCDF (gpfs, dCache)
  - new hardware installation and configuration was slow
  - now up+running, needs configuration
  - plan to migrate /ptmp to new gpfs volume

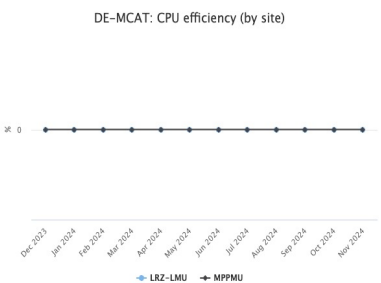
WLCG on MPP cluster @ MPCDF

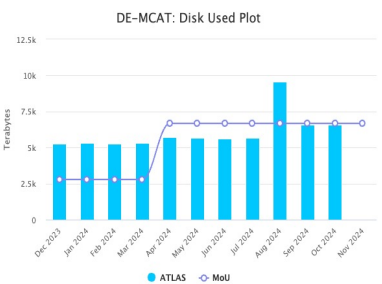
WLCG Accounting

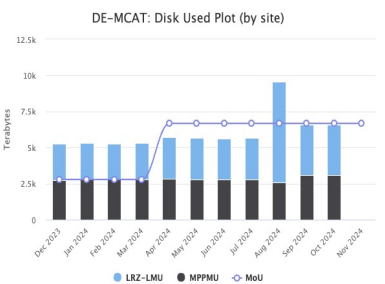
Dec 2023 - Nov 2024

Centre: DE-MCAT









Notes

* Wallclock Work — Sum of wallclock time used by jobs as reported by the batch system during the referenced month multiplied by benchmarked HEPSPEC06 of the CPU resource and by number of processors.

Page 6 DE-MCAT Thu, 05 Dec 2024 11:35:02 GMT

DE-MCAT:

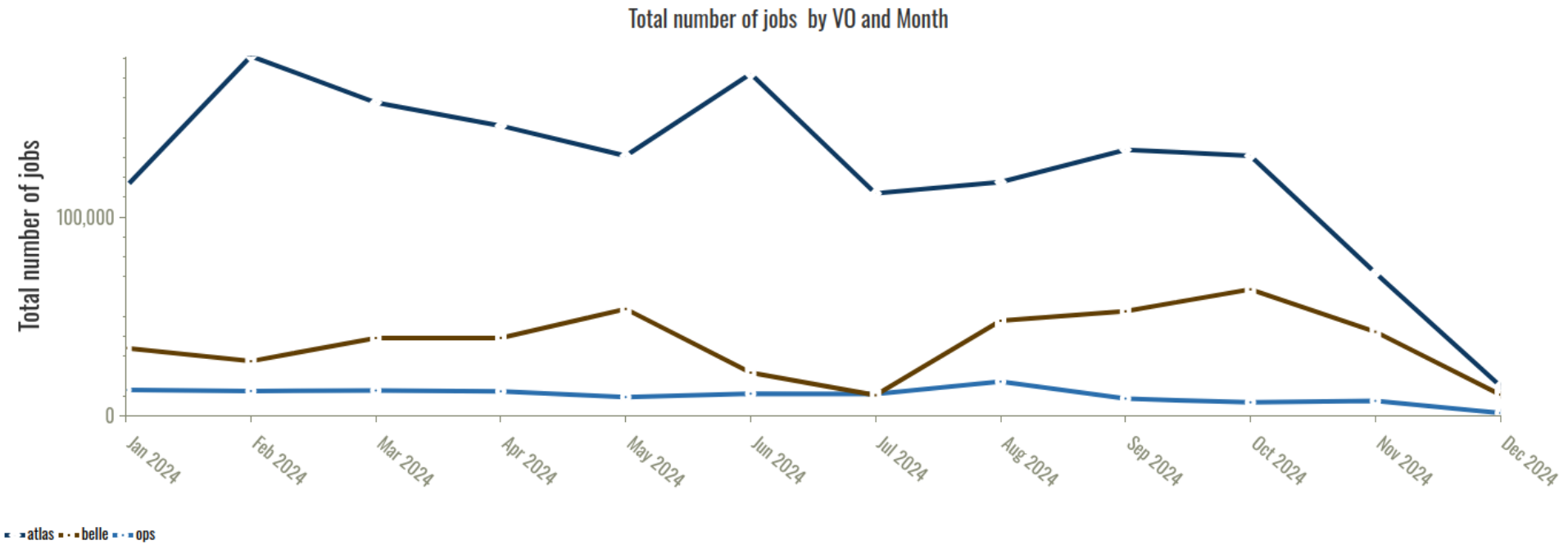
Federated ATLAS T2 MPP+LMU at MPCDF+LRZ

MoU = pledged

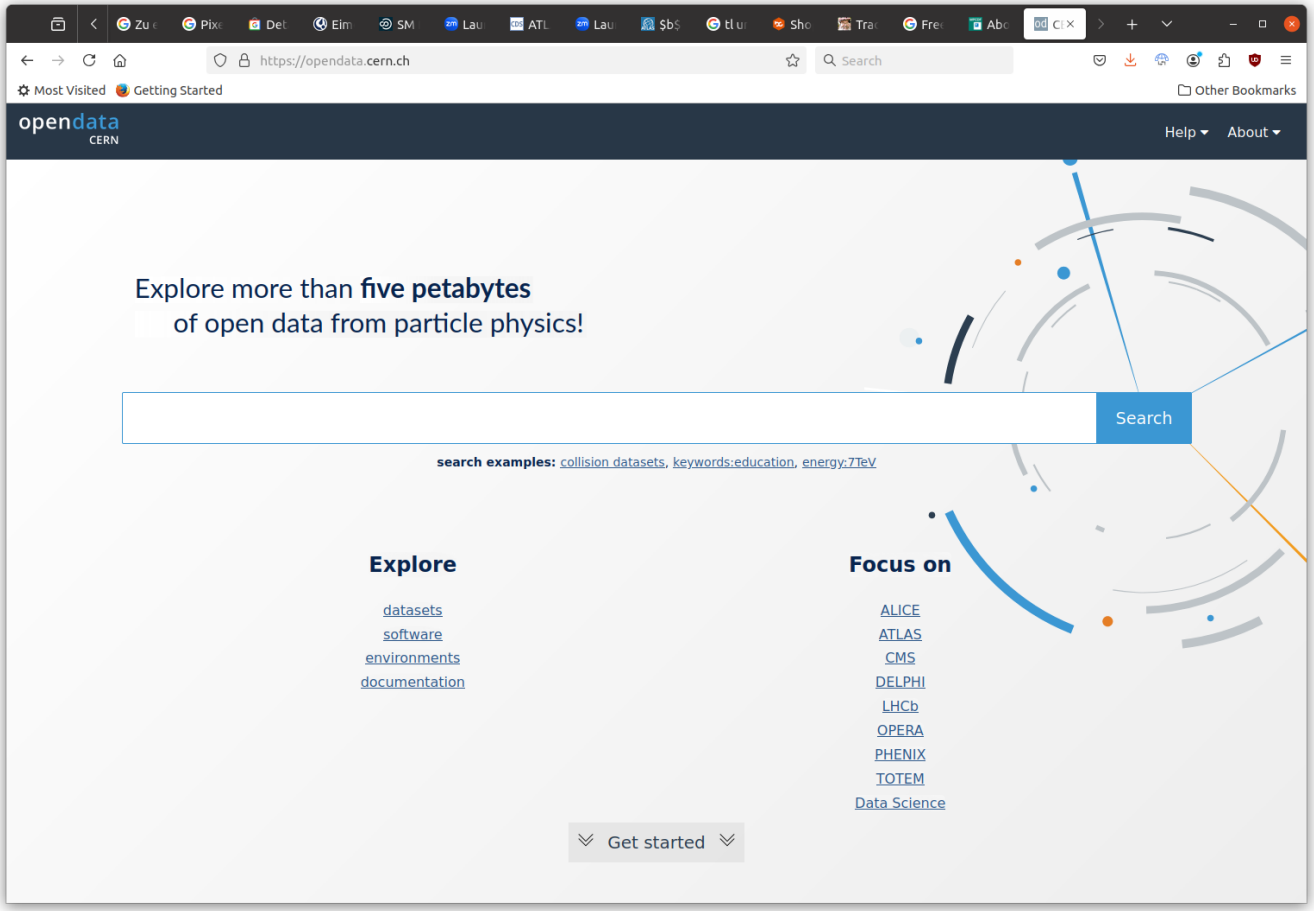
Pledges exceeded (integral)

WLCG on MPP cluster @ MPCDF

https://accounting.egi.eu/egi/site/MPPMU/njobs/VO/DATE/2024/01/2024/12/egi/onlyinfrajobs/



CERN opendata



Data, software, Documentation

JADE now “public” after agreement of collaboration, to be integrated

OPAL opendata process ongoing

SW deployment: COPR

COPR: ”Community projects” (Andrii Verbytskyi)

copr.fedorainfracloud.org/coprs/averbytys/HEPrpms

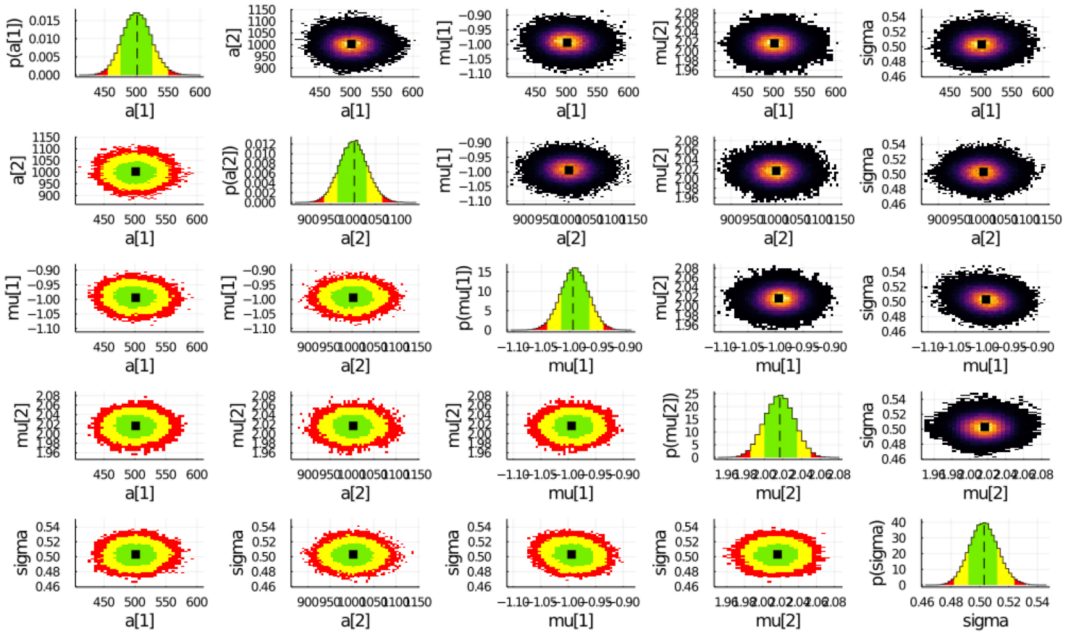
Supports EPEL8/9, Fedora, Tumbleweed

Basis for stand-alone apptainer/Docker containers w/o LCG over /cvmfs, or for Github/Gitlab CI

Scientific software: BAT.jl

Bayesian Analysis Toolkit

BAT.jl plotting: Posterior projections



Bayesian Analysis Toolkit in julia, led by Oliver Schulz

MCMC sampling: MH, HMC, Cuba, ...

Integration with statistics and ML packages in julia

KATRIN, MADMAX, GERDA, COSINUS, ZEUS pdfs, ...

Scientific computing at MPP

- Central for many theory and exptl results
- Increasing demand for CPU (GPU) and storage
- MPP: 1 Linux, /remote/ceph service, condor
- MPCDF: new CPU / GPU servers, new > 5 PB storage
- Linux containers (apptainer) on MPP PCs, MPP cluster, and on MPCDF HPC systems
- Container builds supported by COPR HEPrpms