

# 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  $\geq 8$  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

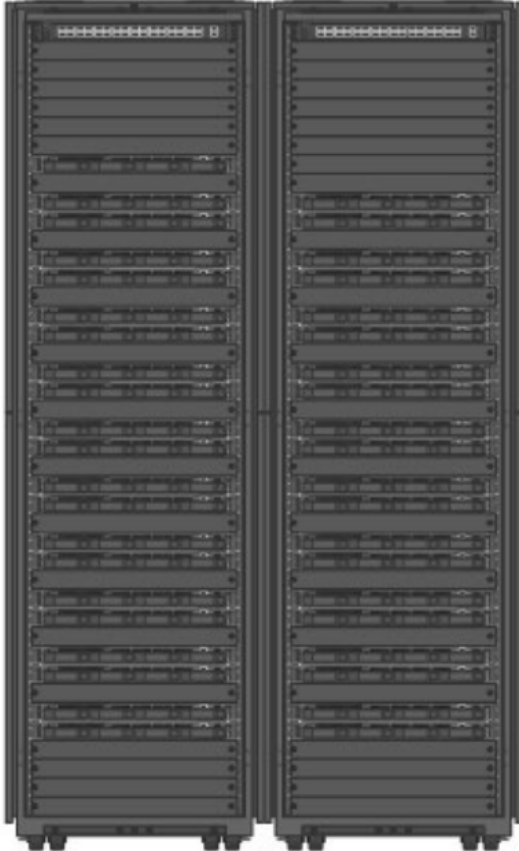
- Desktop PCs common Linux OS
  - Rolling release (opensuse tumbleweed)
  - Gnome, KDE, ..., apptainer, /cvmfs, /remote/ceph, condor
- Rolling release Linux
  - Guarantee support of modern hardware
  - Decent and recent desktop software
  - Scientific workflows in apptainer containers, data in /remote/ceph,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

## Racklayout HPC 41 Nodes

- MPI f. Physik -

21 Nodes (eth)

20 Nodes (eth)



# 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

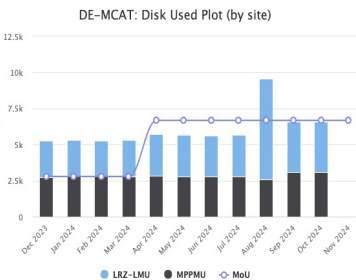
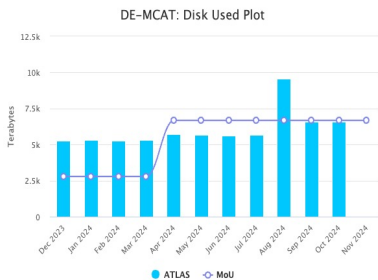
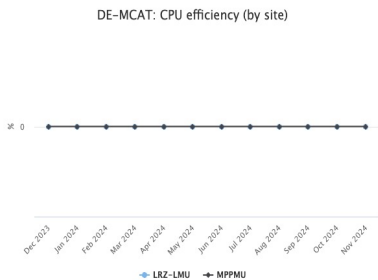
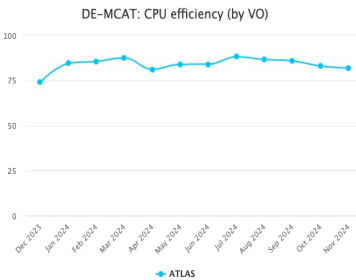
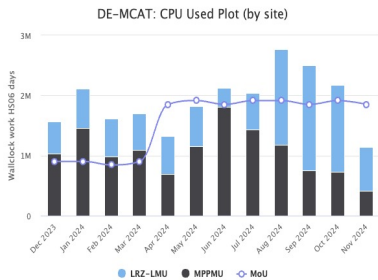
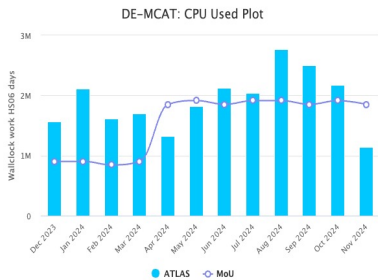
# WLCG on MPP cluster @ MPCDF



## WLCG Accounting

Dec 2023 - Nov 2024

Centre: DE-MCAT



DE-MCAT:

Federated ATLAS T2  
MPP+LMU at  
MPCDF+LRZ

MoU = pledged

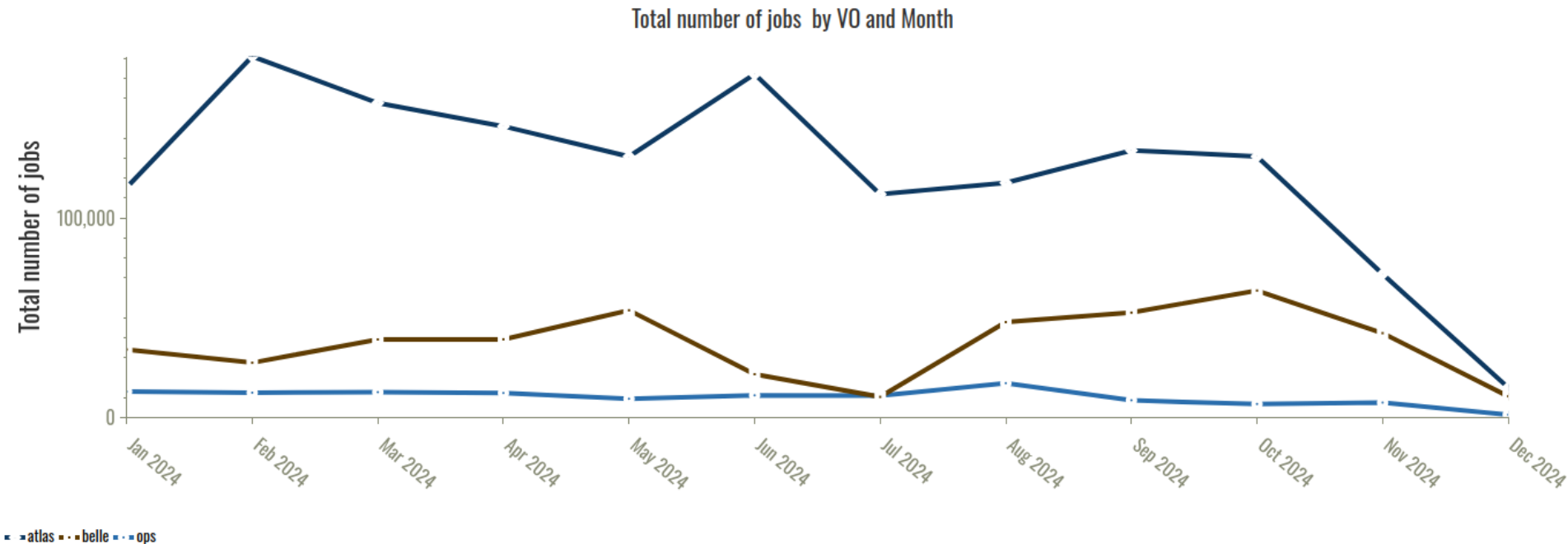
Pledges exceeded  
(integral)

**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.

# WLCG on MPP cluster @ MPCDF

<https://accounting.egi.eu/egi/site/MPPMU/njobs/VO/DATE/2024/01/2024/12/egi/onlyinfracjobs/>



# CERN opendata

The screenshot shows the CERN opendata website in a browser. The URL is <https://opendata.cern.ch>. The page features a dark blue header with the 'opendata CERN' logo and 'Help' and 'About' links. The main content area has a light gray background with a large search bar and a blue 'Search' button. Above the search bar, it says 'Explore more than five petabytes of open data from particle physics!'. Below the search bar, there are search examples: 'collision\_datasets', 'keywords:education', and 'energy:7TeV'. On the left, under 'Explore', there are links for 'datasets', 'software', 'environments', and 'documentation'. On the right, under 'Focus on', there is a list of experiments: ALICE, ATLAS, CMS, DELPHI, LHCb, OPERA, PHENIX, TOTEM, and Data Science. At the bottom, there is a 'Get started' button with a dropdown arrow.

Data, software,  
Documentation

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

OPAL opendata  
process ongoing

The screenshot shows the COPR web interface for the 'HEPrpms' project. The page displays a list of project packages with columns for Name, Last Build Version, Last Build Submitted, Last Build Status, Automatic Build, and Actions. The status for all listed packages is 'succeeded'.

Name	Last Build Version	Last Build Submitted	Last Build Status	Automatic Build	Actions
apfel	3.1.1-1004	a month ago	succeeded	Disabled	-
applgrid	1.6.36-1	a month ago	succeeded	Disabled	-
ariadne	4.12-8	a month ago	succeeded	Disabled	-
binder	1.4.2-1	a month ago	succeeded	Disabled	-
blackhat	0.9.9-7	a month ago	succeeded	Disabled	-
cascade	3.3.3-3	a month ago	succeeded	Disabled	-
cernlib	2024.06.12.0-2	a month ago	succeeded	Disabled	-
CGAL	5.2-1	a month ago	succeeded	Disabled	-
chaplín	1.2-3	a month ago	succeeded	Disabled	-
clhep	2.4.7.1-1	a month ago	succeeded	Disabled	-
COCOA	0.1.1-2	a month ago	succeeded	Disabled	-
collier	1.2.8-1	a month ago	succeeded	Disabled	-
cuba	4.2.2-3	a month ago	succeeded	Disabled	-
Delphes	3.5.1pre10-1	a month ago	succeeded	Disabled	-
EvtGen	2.2.3-1	a month ago	succeeded	Disabled	-
f90cache	0.99c-2	a month ago	succeeded	Disabled	-
fastjet	3.4.2-1001	a month ago	succeeded	Disabled	-
fastnlo	2.5.0.2826-5	a month ago	succeeded	Disabled	-
FeynHiggs	2.19.0-2	5 months ago	succeeded	Disabled	-
fjcontrib	1.056-1	a month ago	succeeded	Disabled	-
form	4.3.1-2	a month ago	succeeded	Disabled	-
geant4	11.2.1-1	a month ago	succeeded	Disabled	-

# SW deployment: COPR

COPR: "Community projects" (Andrii Verbytskyi)

[copr.fedorainfracloud.org/coprs/averbyts/HEPrpms](https://copr.fedorainfracloud.org/coprs/averbyts/HEPrpms)

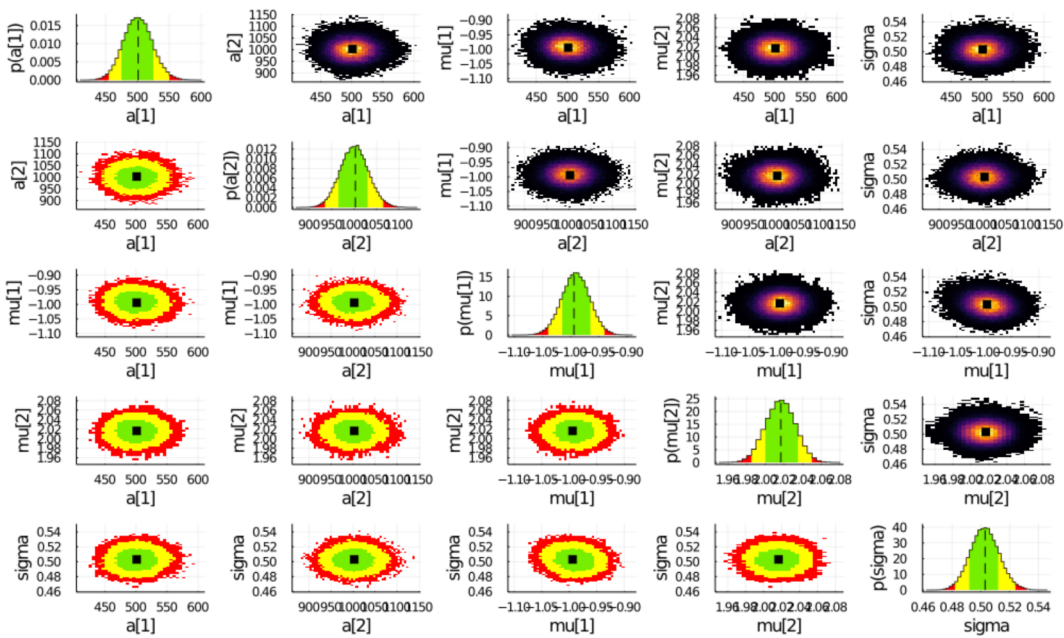
Supports EPEL8/9, Fedora, Tumbleweed



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



## 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