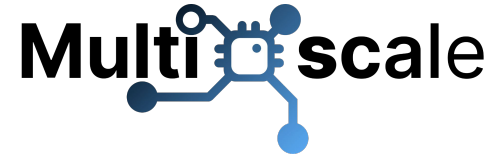




CernVM
File system



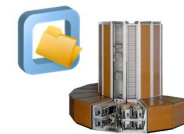
Best Practices for CernVM-FS in HPC



online tutorial
Mon 4 Dec 2023



Acknowledgements



CernVM
File system



Jakob Blomer (CERN)

Valentin Völkl (CERN)

Laura Promberger (CERN)



Bob Dröge (Univ. of Groningen)

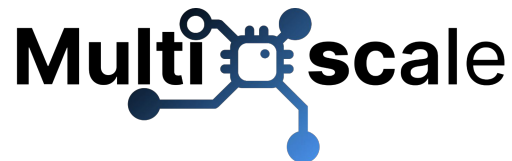
Lara Peeters (HPC-UGent)

Kenneth Hoste (HPC-UGent)

Thomas Röblitz (Univ. of Bergen)

Alan O'Cais (CECAM)

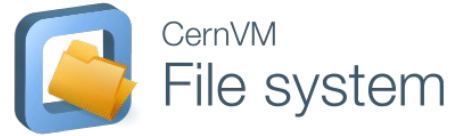
Caspar van Leeuwen (SURF)



EuroHPC
Joint Undertaking

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and countries participating in the project under grant agreement No 101093169.

A word from the CernVM-FS team



- CVMFS is developed at CERN, by a small dedicated team
 - *(Valentin Volkl, Laura Promberger, Jakob Blomer)*
 - Goes back almost 15 years, with origins in early cloud computing, hence the name "CernVM-File System"!
- Long term perspective - will be supported through HL-LHC (--2030) and beyond
- The High-Energy Physics community typically does High-Throughput computing, but already uses HPC sites (NERSC, Mare Nostrum ...) - CVMFS is running there already!
- We know that HPC sites tend to be unique in their capabilities and constraints
 - Look forward to discussing your setup
- Questions or issues? Find us online* (**after this tutorial, use EESSI slack for immediate questions*)!
 - <https://github.com/cvmfs/cvmfs> (open source and issue tracker)
 - <https://cernvm-forum.cern.ch> (Forum for general questions and discussion)
 - <https://mattermost.web.cern.ch/cernvm> (chat, requires lightweight CERN account)

- Collaboration between scientific partners (**CECAM** nodes) who deliver 3 pilot use cases, and technical partners (**EESSI** members) who provide the tools to allow application software to be seamlessly used on any available hardware.
- MultiXscale targets improving the three Ps:
 - **Productivity** of scientists who develop and/or use open source codes
 - **Performance** of those codes on EuroHPC hardware (and beyond)
 - **Portability** of the codes and workflows from laptop to server to cloud to HPC
- The focus of MultiXscale is **multi-scale modelling** and leveraging the opportunities that EuroHPC offers
- Advances the transition towards **use of exascale resources** for the community



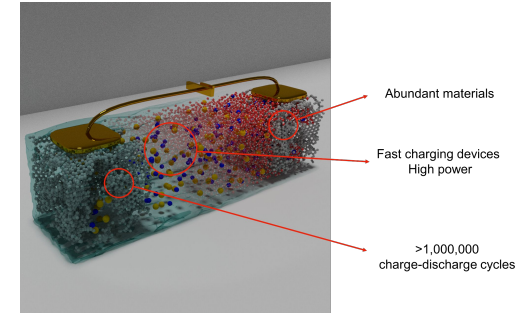
EuroHPC
Joint Undertaking



Co-funded by
the European Union

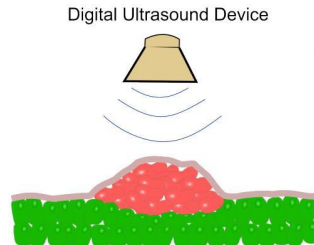
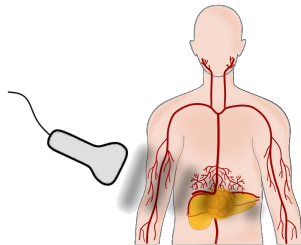
Pilot use cases:

- Batteries for sustainable energy
- Biomedical Applications (digital twin for ultrasound process)
- Design of innovative helicopters for civil applications



Each pilot features phenomena that manifest at **multiple scales**, and as a result each utilizes **multiple methods** and a **variety of software applications**.

The challenge is to **couple** these, and make sure they can efficiently **use of large scale resources**.

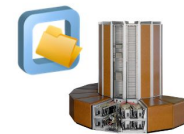


Practical information



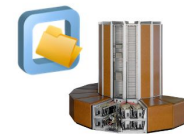
- See also email that was sent to registered attendees this morning
- Tutorial website: <https://multixscale.github.io/cvmfs-tutorial-hpc-best-practices>
- **Zoom link is only to be shared with registered attendees**
- Be aware that session is being recorded, will be available via [MultiXscale YouTube channel](#)
- Dedicated channel in [EESSI Slack](#) for Q&A: [#cvmfs-best-practices-hpc](#) (use threads!)
- Hands-on demos will be done throughout the tutorial
- All necessary information to work through hands-on yourself **later** is included in the tutorial website

Agenda



- [13:35-14:00 CET] Introduction to CernVM-FS (incl. flagship repositories)
- [14:00-14:20 CET] European Environment for Scientific Software Installations (EESSI)
- [14:20-15:00 CET] Accessing CernVM-FS repositories (incl. **hands-on demo**)
- [15:00-15:30 CET] *(coffee break)*
- [15:30-15:50 CET] Configuring CernVM-FS on HPC infrastructure (incl. **hands-on demo**)
- [15:50-16:10 CET] Troubleshooting for CernVM-FS (incl. **hands-on demo**)
- [16:10-16:30 CET] Performance aspects of CernVM-FS (incl. **hands-on demo**)
- [16:30-16:40 CET] Containers and CernVM-FS (incl. **hands-on demo**)
- [16:40-16:50 CET] Creating a CernVM-FS repository (birds-eye view)
- [16:50-17:00 CET] Q&A

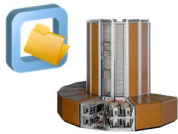
Introduction to CernVM-FS



multixscale.github.io/cvmfs-tutorial-hpc-best-practices/cvmfs

- What is CernVM-FS?
- Technical details
- Flagship repositories
- Appendix - CernVM-FS terminology:
repository, client, proxy server, Stratum 1 replica server, ...
multixscale.github.io/cvmfs-tutorial-hpc-best-practices/appendix/terminology

European Environment for Scientific Software Installations



multixscale.github.io/cvmfs-tutorial-hpc-best-practices/eessi



<https://eessi.io>

E E S S I
EUROPEAN ENVIRONMENT FOR
SCIENTIFIC SOFTWARE INSTALLATIONS

(hands-on demo)

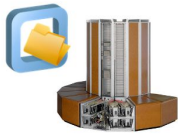
```
$ ls /cvmfs/software.eessi.io
README.eessi host_injections versions

$ source /cvmfs/software.eessi.io/versions/2023.06/init/bash
...

[EESSI 2023.06] $ module load TensorFlow/2.13.0
[EESSI 2023.06] $ python -c 'import tensorflow as tf; print(tf.__version__)'
2.13.0
```

[14:00-14:20 CET]

Accessing CernVM-FS repositories



multixscale.github.io/cvmfs-tutorial-hpc-best-practices/access

- Setting up a CernVM-FS client system (**hands-on demo**)
- Setting up a proxy server (**hands-on demo**)
- Setting up a Stratum 1 replica server (**hands-on demo**)
- Alternative ways to access CernVM-FS repositories (**hands-on demo**)
 - `cvmfsexec`
 - Apptainer with `--fusemount`
 - Alien cache

Coffee break (15:00-15:30 CET)



[15:00-15:30 CET]

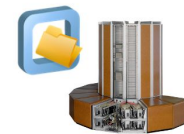
Configuring CernVM-FS on HPC infrastructure



multixscale.github.io/cvmfs-tutorial-hpc-best-practices/configuration_hpc

- **Diskless worker nodes**
 - In-memory client cache
 - Loopback on shared filesystem
 - Alien cache
- **Offline worker nodes**
 - Squid proxy in local network
 - Private Stratum 1 replica server
 - Alien cache
- **Worker nodes without CernVM-FS**
 - Syncing a to another filesystem
 - Alternative access mechanisms

Troubleshooting for CernVM-FS



multixscale.github.io/cvmfs-tutorial-hpc-best-practices/troubleshooting

```
$ ls /cvmfs/software.eessi.io
ls: cannot access '/cvmfs/software.eessi.io': No such file or directory

$ ls /cvmfs/
cvmfs-config.cern.ch

$ ls /cvmfs/cvmfs-config.cern.ch
etc

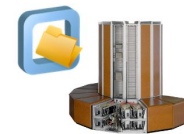
$ ls /cvmfs/software.eessi.io
README.eessi  host_injections  versions
```



(hands-on demo)

[15:50-16:10 CET]

Performance aspects of CernVM-FS

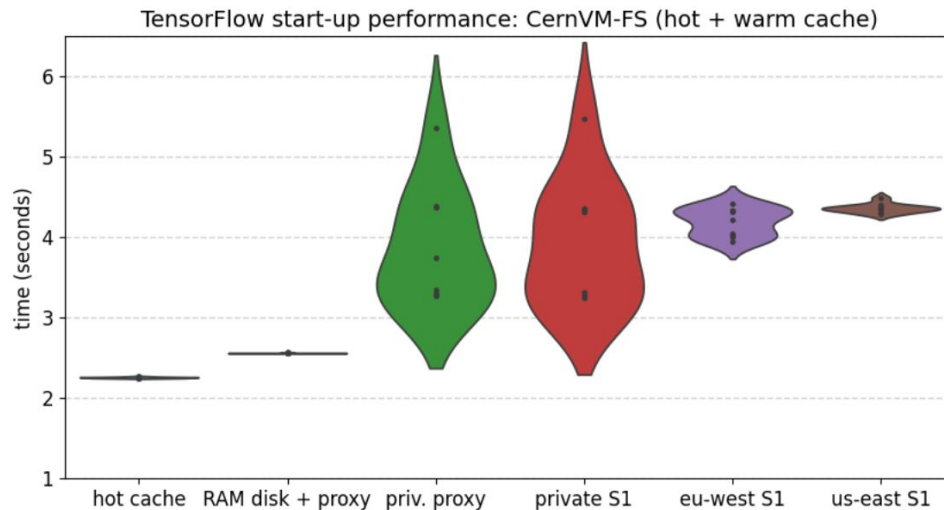
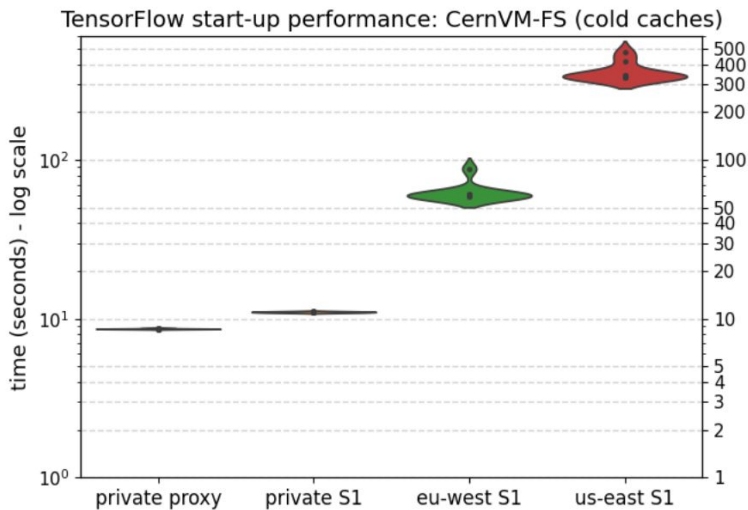


multiscale.github.io/cvmfs-tutorial-hpc-best-practices/performance

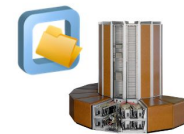
Evaluating software start-up performance with TensorFlow

```
python -c 'import tensorflow'
```

(hands-on demo)



Containers and CernVM-FS



multixscale.github.io/cvmfs-tutorial-hpc-best-practices/containers

- Accessing containers hosted in CernVM-FS via Apptainer
- Ingesting container images
- Using /cvmfs inside containers

Q&A



Time for questions, and a (Zoom) group picture!

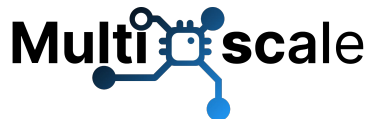


CernVM
File system

cernvm.cern.ch/fs | [docs](#) | [GitHub](#) | [forum](#) | [Mattermost](#) (chat)



eessi.io | [docs](#) | [GitHub](#) | [Slack](#) | [YouTube](#) | [support](#) | [Twitter/X](#)



www.multixscale.eu | [YouTube](#) | [Twitter/X](#) | [LinkedIn](#)