EESSI

EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

Multi scale

Extending Arm's Reach by Going EESSI

ISC'24 - Hamburg - 15 May 2024

Kenneth Hoste (Ghent University, Belgium)

kenneth.hoste@ugent.be

whoami

GHENT UNIVERSITY

VLAAMS SUPERCOMPUTER CENTRUM

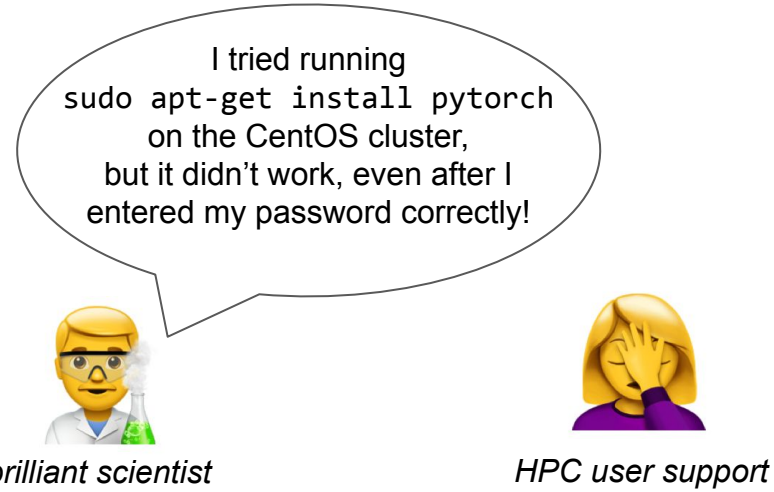
ugent.be/hpc vscentrum.be

Kenneth Hoste: Computer scientist from Belgium

- HPC system administrator + user support at Ghent University since Oct 2010
- BDFL of **EasyBuild** - tool to build & install **scientific software on HPC clusters**
- Active contributor to **EESSI**, partner in MultiXscale EuroHPC CoE
- Fan of open source software (FOSS), beers, and stickers
- "How To Make Package Managers Cry" talk at FOSDEM'18 (if you haven't seen it, you should!)

Scientific software is a different breed

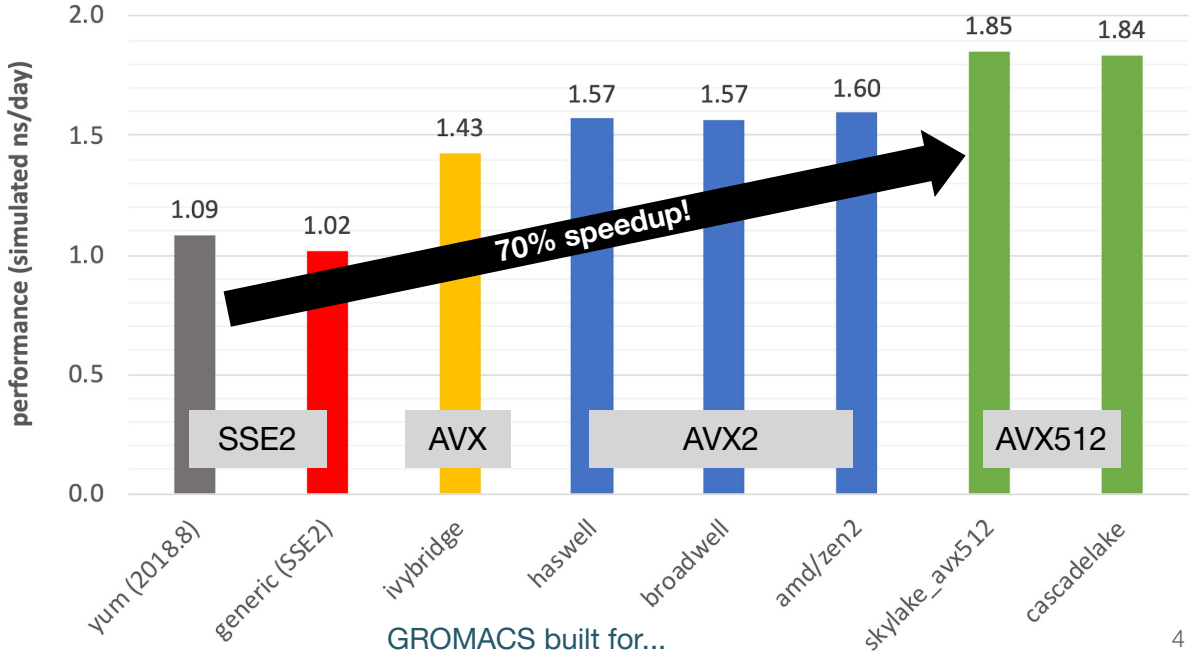
- Scientists may not be trained software engineers (and that's OK)
- They often need help to get the software installed properly (especially on HPC systems)
- **Standard packaging tools do not suffice**, we ~~want~~ need to **build from source**
- Scientific software should be compiled for system on which it will be used (performance!)



Optimized scientific software installations

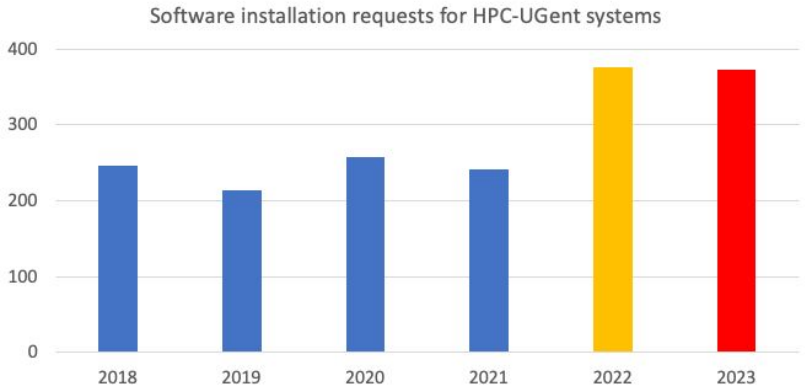
- Software should be optimized for the system it will run on
- Impact on performance is often significant for scientific software

- Example: GROMACS 2020.1 (PRACE benchmark, Test Case B)
- Metric: (simulated) ns/day, higher is better
- Test system: dual-socket Intel Xeon Gold 6420 (Cascade Lake, 2x18 cores)
- **Performance of different GROMACS binaries, on exact same hardware/OS**



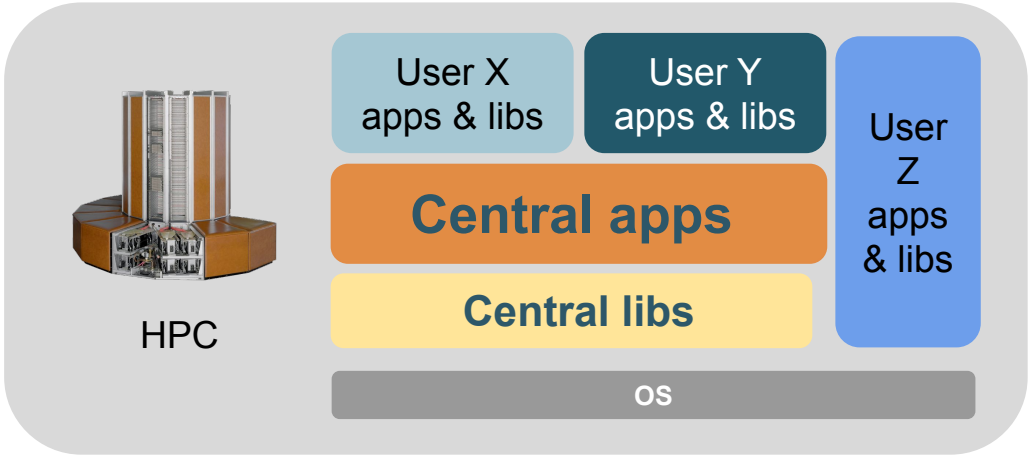
Landscape of scientific computing is changing

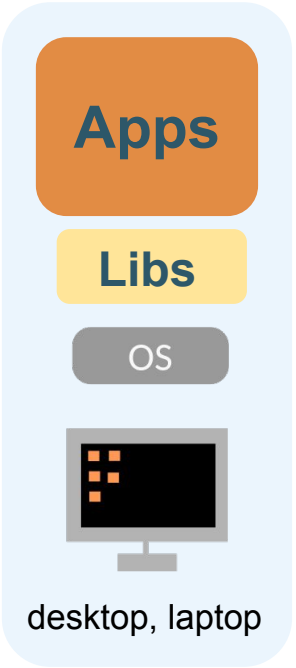
- **Explosion of available scientific software** applications (bioinformatics, AI, …)
- Increasing interest in **cloud** for scientific computing (flexibility!)
- **Increasing variety in processor (micro)architectures** beyond Intel & AMD:
  Arm is ~~coming~~ already here (see Fugaku, JUPITER, …), RISC-V is coming (soon?)
- Broader adoption of **accelerated computing**, beyond NVIDIA GPUs
- In strong contrast: available (wo)manpower **in HPC support teams is (still) limited**…

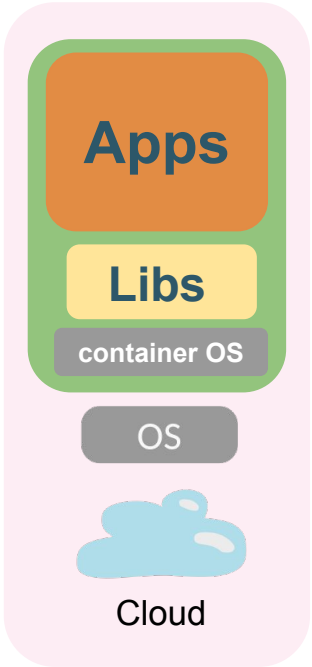


We need to collaborate more…

- **Too much software** for a single support team to handle
- Different systems (CPU, GPU, OS, …) => different problems
- Existing tools (EasyBuild, Spack) are **not sufficient anymore…**
- Lots of **duplicate work** across HPC sites and scientists
- **Diverse software stacks** across different platforms







European Environment for Scientific Software Installations

- **Public repository of (optimized!) scientific software *installations***
- Avoid duplicate work by collaborating on a shared software stack
- Uniform way of providing software to users, regardless of the system they use!
- Should work on any Linux OS (incl. WSL) and system architecture
  - From laptops and personal workstations to HPC clusters and cloud
  - Support for different CPUs (AMD, Intel, Arm, RISC-V), interconnects, GPUs, etc.
- **Focus on performance, automation, testing, collaboration**

EESSI

EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

https://eessi.io

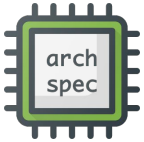
https://eessi.io/docs

ReFrame

Testing

Software layer

Optimized applications + dependencies



EasyBuild

Lmod

Compatibility layer

Levelling the ground across client OSs

gentoo

Filesystem layer

Distribution of the software stack

CernVM-FS

Host OS provides network & GPU drivers, resource manager (Slurm), ...

Host operating system

intel

AMD

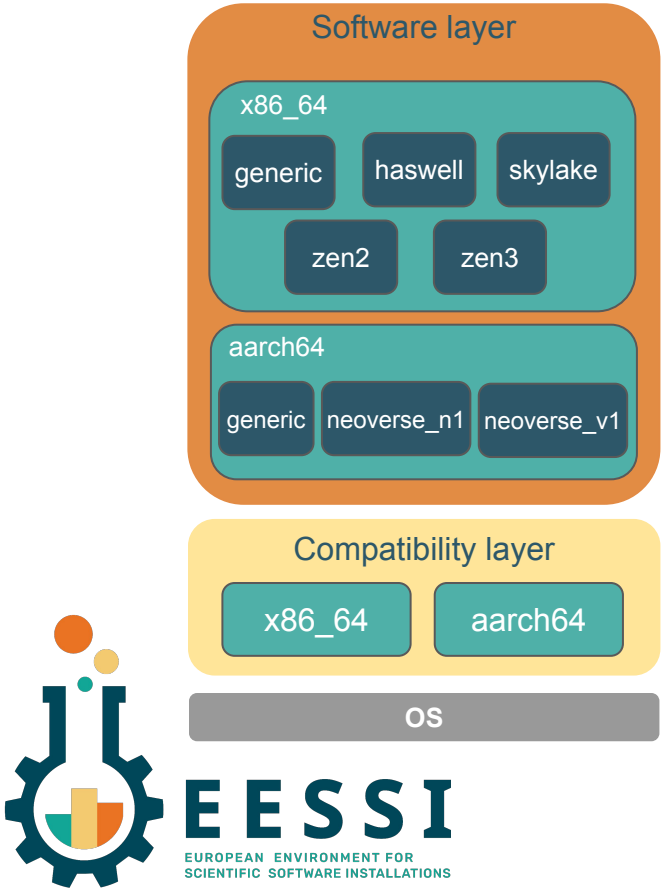
ARM

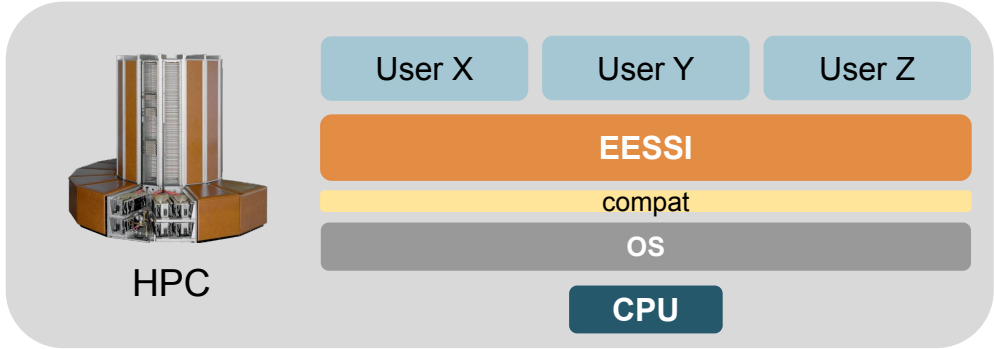
RISC-V

EESSI

EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

EESSI to the rescue!





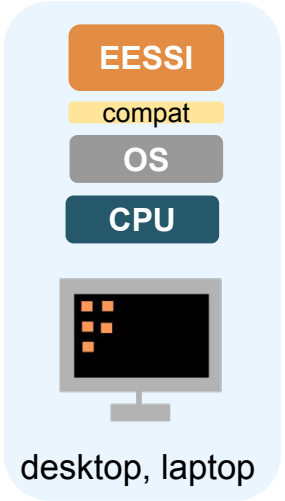
Shared repository of **(optimized)** scientific software **installations**

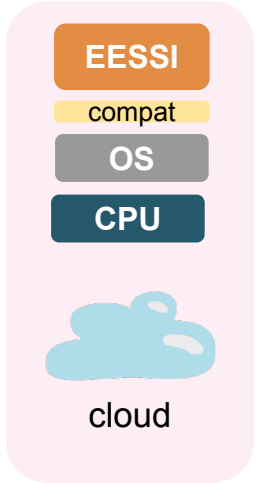
Same software stack everywhere!

eessi.io

eessi.io/docs

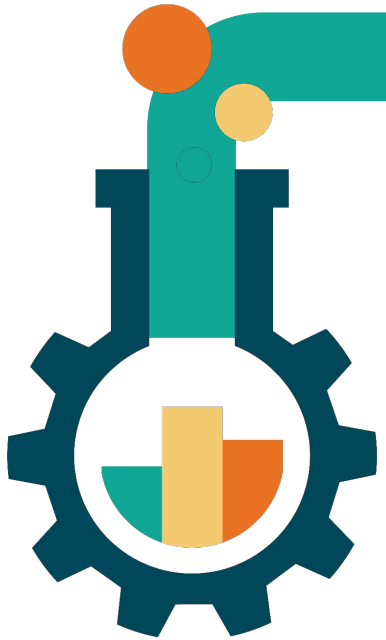
eessi.io/docs/support





EESSI
EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

EESSI ingredients: filesystem layer



EESSI

EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

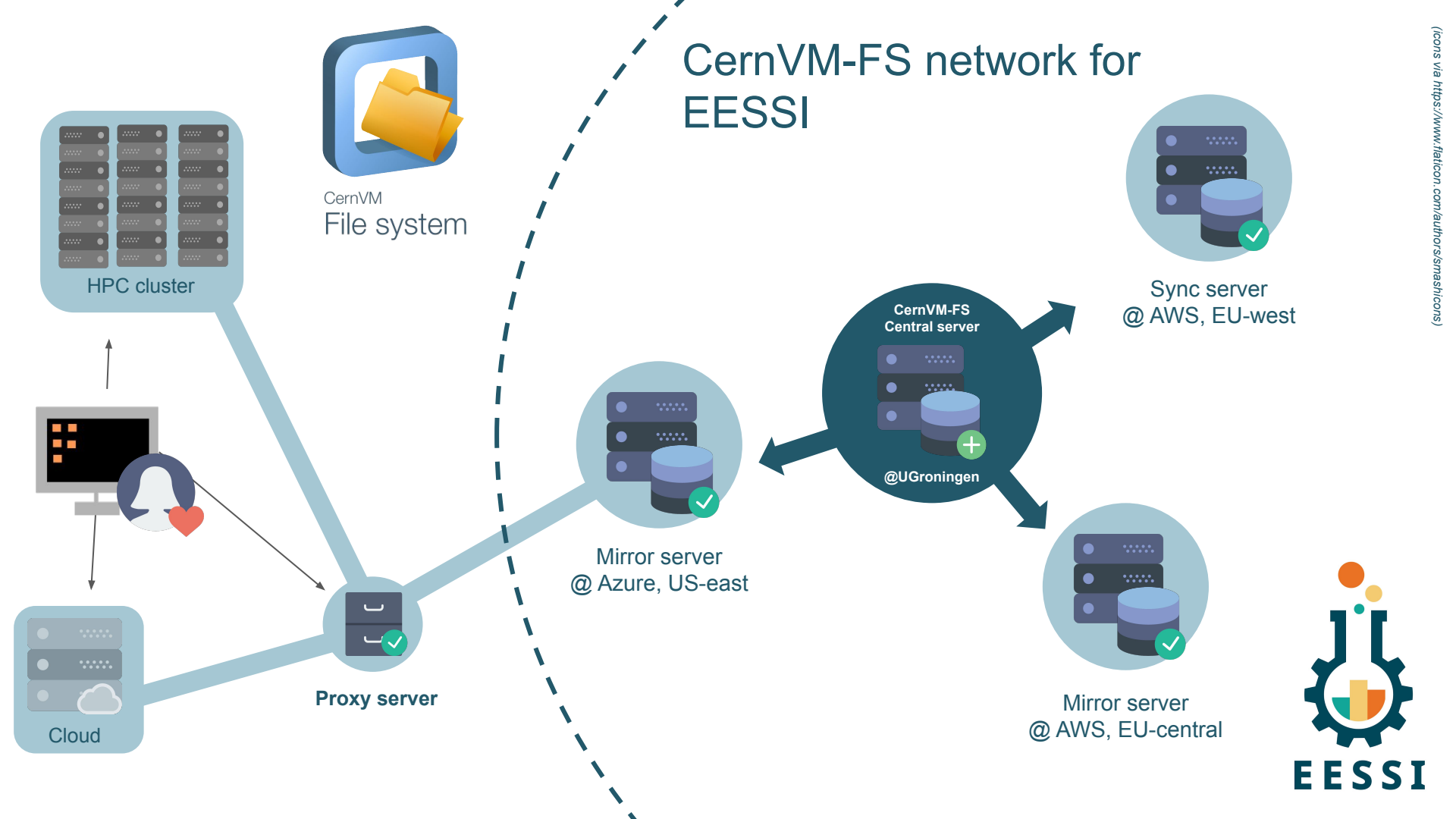
CernVM File system

Global distribution of software installations

https://cernvm.cern.ch/fs

Read-only network filesystem
"Streaming Scientific Software"

CernVM-FS network for EESSI

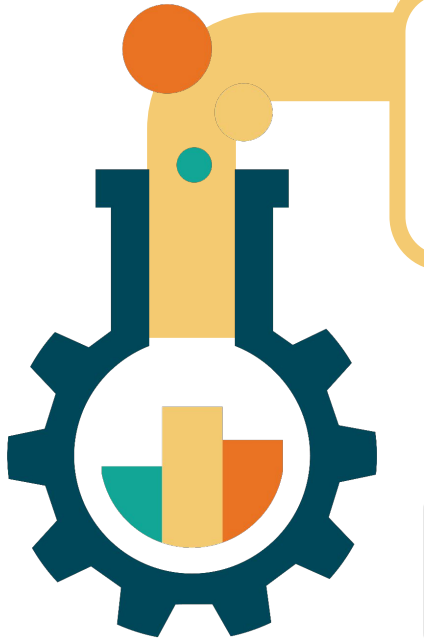


(icons via https://www.flaticon.com/authors/smashicons)

Tutorial “Best Practices for CernVM-FS in HPC”

- https://multixscale.github.io/cvmfs-tutorial-hpc-best-practices
- Held online on 4 Dec 2023 (~3 hours), **recorded & available on YouTube**
- Over 200 registrations, ~125 attending the meeting
- Lecture + hands-on demos
- Topics:
  - Introduction to CernVM-FS + EESSI
  - Configuring CernVM-FS: client, Stratum 1 mirror server, proxy server
  - Troubleshooting problems
  - Benchmarking of start-up performance

EESSI ingredients: compatibility layer



gentoo linux™

Abstraction away from host operating system

Similar to container images

"Containers without the containing"

CernVM File system

Global distribution of software installations

EESSI

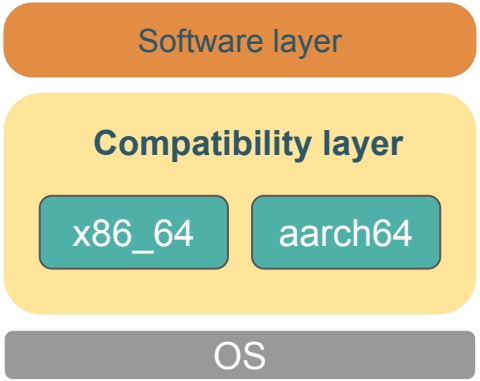
EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS

EESSI compatibility layer

github.com/EESSI/compatibility-layer

- *“Containers without the containing”*
- **Minimal collection of tools and libraries** (incl. glibc, bash, Python, Lmod, …)
- **Built from source per CPU family** (`x86_64`, `aarch64`, …) with Gentoo Prefix
- Installations included in software layer **only link to compat layer** (RPATH)
- Ensures **compatibility** with any client system running Linux

```
$ ls /cvmfs/software.eessi.io/versions/2023.06/compat/linux/aarch64/
bin  etc  lib  lib64  opt  reprod  run  sbin  stage1.log  stage2.log
stage3.log  startprefix  tmp  usr  var
```



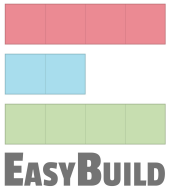
EESSI ingredients: software layer

Abstraction away from host operating system

Global distribution of software installations

EESSI

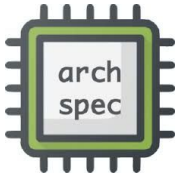
EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS



Optimized software installations for specific CPU microarchitectures

Intuitive user interface

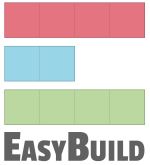
```
module avail,
module load, ...
```



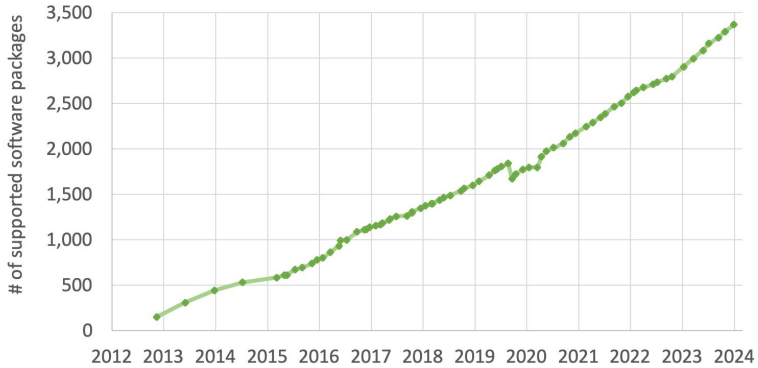
Automatic selection of best suited part of software stack for host CPU microarchitecture

EasyBuild (in a nutshell)

https://easybuild.io



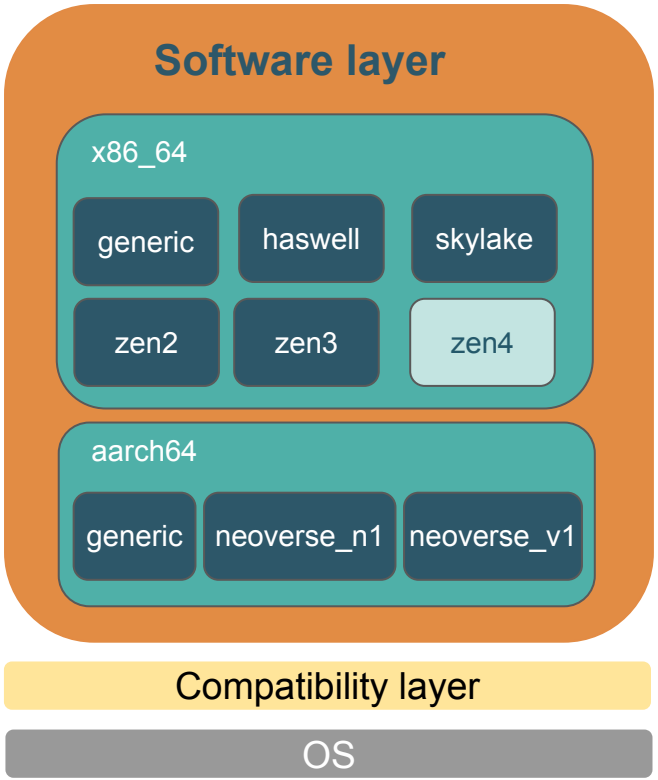
- EasyBuild is a tool to make **installing (scientific) software on HPC systems** easier
- **Building from source** + optimizing for specific CPU is strongly preferred (performance!)
- Created in 2009 by HPC-UGent team, open source (GPLv2) since 2012
- Now a **world-wide community** (> 400 unique contributors, close to 3,000 PRs per year)
- ~3,500 different software projects supported (excl. versions) + ~2,000 “extensions”



Software layer

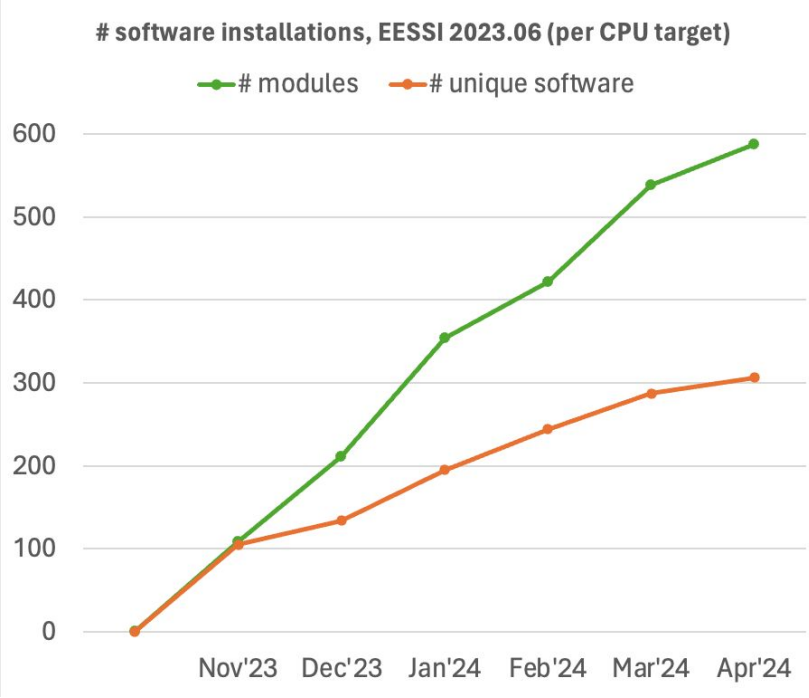
github.com/EESSI/software-layer

- Installations of scientific software applications
- **Optimized for specific CPU targets**
- Works on any client system running Linux, since we only link to libraries in compat layer
- Built using EasyBuild
- Environment modules as user interface (via Lmod)
- Detection of host CPU via archspec (Python) or archdetect (bash)
- **Best subset of software installations for host CPU Is automatically selected**



Overview of installed software

- ~600 software software installations available per CPU target, and increasing every day
  - Including ESPResSo, GROMACS, LAMMPS, OpenFOAM, PyTorch, R, QuantumESPRESSO, TensorFlow, WRF
  - eessi.io/docs/available_software (coming soon!)
- Includes (**but is not limited to!**) applications specific to the MultiXscale CoE
- Focus on recent compiler toolchains: currently targeting foss/2023a and foss/2023b



Getting access to EESSI

- Native installation of CernVM-FS *(requires admin privileges)*
  eessi.io/docs/getting_access/native_installation
- Using a container (via Apptainer)
  eessi.io/docs/getting_access/eessi_container
- Via cvmfsexec
  github.com/cvmfs/cvmfsexec

To check whether you have access to EESSI:

```
ls /cvmfs/software.eessi.io
```

Accessing EESSI via CernVM-FS

```
# Native installation
# Installation commands for RHEL-based distros
# like CentOS, Rocky Linux, Almalinux, Fedora, …

# install CernVM-FS
sudo yum install -y
https://ecsft.cern.ch/dist/cvmfs/cvmfs-release/cvmfs-release-latest.noarch.rpm
sudo yum install -y cvmfs

# create client configuration file for CernVM-FS
# (no proxy, 10GB local CernVM-FS client cache)
sudo bash -c "echo 'CVMFS_CLIENT_PROFILE="single"' > /etc/cvmfs/default.local"
sudo bash -c "echo 'CVMFS_QUOTA_LIMIT=10000' >> /etc/cvmfs/default.local"

# Make sure that EESSI CernVM-FS repository is accessible
sudo cvmfs_config setup
```

Alternative ways of accessing EESSI are available, via a container image, via cvmfsexec, …
eessi.io/docs/getting_access/native_installation - eessi.io/docs/getting_access/eessi_container

Using EESSI

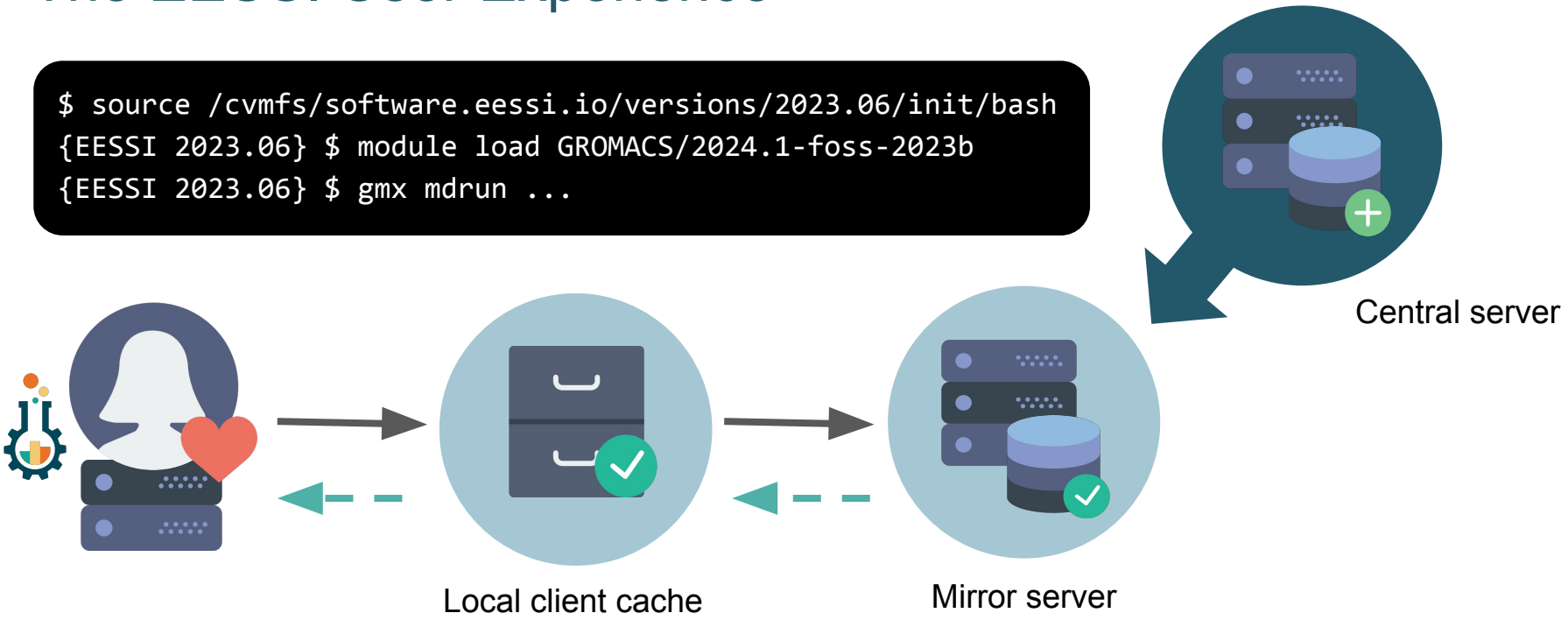
eessi.io/docs/using_eessi/eessi_demos

```
/cvmfs/software.eessi.io/versions/2023.06/software
`-- linux
    |-- aarch64
    |   |-- generic
    |   |-- neoverse_n1
    |   `-- neoverse_v1
    |       |-- modules
    |       `-- software
    `-- x86_64
        |-- amd
        |   |-- zen2
        |   |-- zen3
        |   `-- zen4
        |-- generic
        `-- intel
            |-- haswell
            `-- skylake_avx512
```

```
$ source /cvmfs/software.eessi.io/versions/2023.06/init/bash
Found EESSI pilot repo @
/cvmfs/software.eessi.io/versions/2023.06!
archdetect says aarch64/neoverse_v1
Using aarch64/neoverse_v1 as software subdirectory
...
Environment set up to use EESSI pilot software stack, have fun!

{EESSI 2023.06} $ module load R/4.3.2-gfbf-2023a

{EESSI 2023.06} $ which R
/cvmfs/software.eessi.io/versions/2023.06/software/linux/aarch64
/neoverse_v1/software/R/4.3.2-gfbf-2023a/bin/R

{EESSI 2023.06} $ R --version
R version 4.3.2
```

The EESSI User Experience

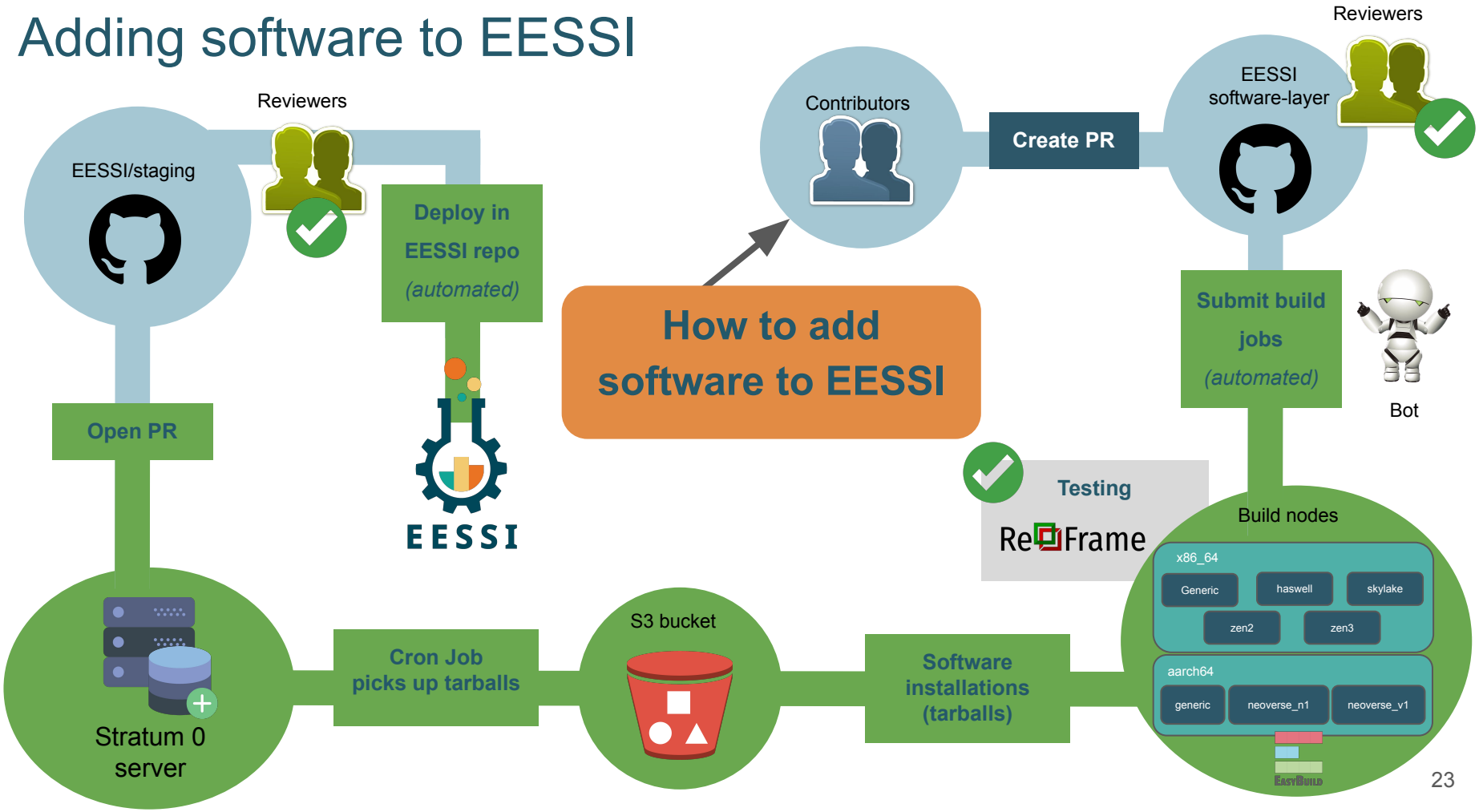
```
$ source /cvmfs/software.eessi.io/versions/2023.06/init/bash
{EESSI 2023.06} $ module load GROMACS/2024.1-foss-2023b
{EESSI 2023.06} $ gmx mdrun ...
```



EESSI provides **on-demand streaming**

of (scientific) software (like music, TV-series, …)

Adding software to EESSI



EESSI test suite

ReFrame

eessi.io/docs/test-suite

EESSI

- Ensure quality of the software installations provided by EESSI
- A suite of **portable tests** for scientific software, using ReFrame
- Version 0.2.0 includes tests for GROMACS, TensorFlow, OSU Micro Benchmarks
- WIP: tests for LAMMPS, QuantumESPRESSO, CP2K, OpenFOAM, ESPResSo, PyTorch
- See also MultiXscale deliverable D1.2 “**Plan for the design of a portable test suite**” https://zenodo.org/records/10451718

Problems we have found through software testing (1/2)

Hang/crash in Open MPI's `smcuda` Byte Transport Layer (BTL) component

- See https://gitlab.com/eessi/support/-/issues/41
- Upstream issue: https://github.com/open-mpi/ompi/issues/12270
- Causes hanging or failing tests for FFTW, OpenFOAM, ESPResSo, …
- Only happens on Arm Neoverse V1 (AWS Graviton 3)
- Fixed by Luke Robinson (AWS), see https://github.com/open-mpi/ompi/pull/12338
- Fix will be included in upcoming Open MPI release (v4.1.7+)
- Patch already applied in Open MPI installations included in EESSI 2023.06

Problems we have found through software testing (2/2)

Failing tests in GROMACS test suite

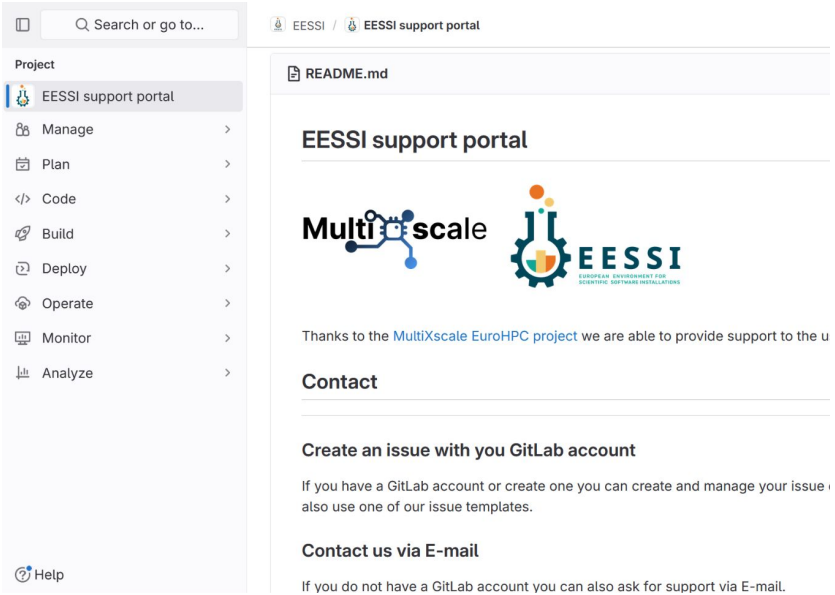
- See https://gitlab.com/eessi/support/-/issues/47
- Filesystem race when running tests concurrently (GROMACS PR #4066)
- **Bug in SVE support, leading to (very) wrong results for several tests**
  - See https://gitlab.com/gromacs/gromacs/-/issues/5057
  - Works fine on A64FX (512-bit SVE), but **problem on Neoverse V1 (256-bit SVE)**
  - WIP fix in https://gitlab.com/gromacs/gromacs/-/merge_requests/4299
  - Will be fixed in upcoming GROMACS release (2024.2?)
  - Workaround for GROMACS 2024.1 in EESSI: use `-DGMX_SIMD=ARM_NEON_ASIMD`

Getting support for EESSI

GitLab

gitlab.com/eessi/support

- Via GitLab, or via email: `support@eessi.io`
- Report problems
- Ask questions
- Request software
- Get help with contributing
- Suggest features
- Confidential tickets possible (security issues, …)



Dedicated support team, thanks to EuroHPC Centre-of-Excellence MultiXscale

Multiscale

www.multixscale.eu

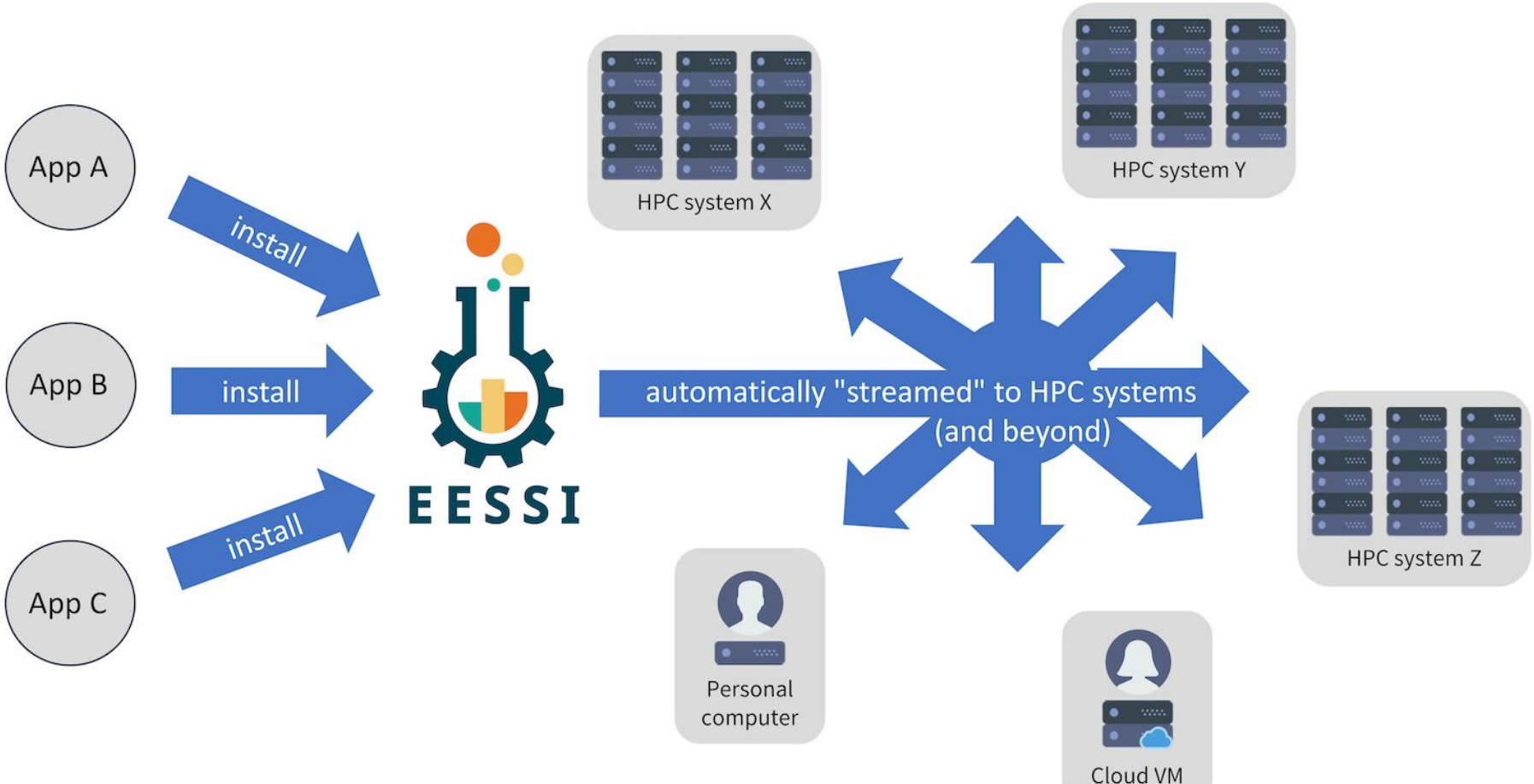
EuroHPC Joint Undertaking

- EuroHPC Centre of Excellence: 4 year project (2023-2026), €6M budget (50% for EESSI)
- Collaboration between EESSI and CECAM: total of 16 partners (academic + industry)
- EESSI focuses on technical aspects: providing a shared stack of scientific software
- Scientific target: multiscale simulations with 3 key use cases
  - Helicopter design and certification for civil transport
  - Battery applications to support the sustainable energy transition
  - Ultrasound for non-invasive diagnostics and biomedical applications

Kemijski inštitut · Universitat de Barcelona · Universität Stuttgart · Universitas Bergensis · Sorbonne Université · Université de Toulouse · Consiglio Nazionale delle Ricerche · Max-Planck-Gesellschaft · HPC Now! · University of Groningen · SURF · Jülich Forschungszentrum · Ghent University · BSC Barcelona Supercomputing Center · Leonardo · iit Istituto Italiano di Tecnologia

The role of EESSI in MultiXscale

EESSI as a shared software stack



EESSI in a nutshell

- **On-demand streaming** of **optimized** scientific software installations
- **Works on any Linux distribution** thanks to EESSI compat layer
- **Uniform software stack** across various systems: laptop, HPC, cloud, …
- Community-oriented: **let's tackle the challenges we see together!**

Acknowledgements

- 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.

- Thanks to Amazon Web Services (AWS) and Microsoft Azure for generously sponsoring the EESSI project with cloud credits, feedback, and guidance.

Website: eessi.io

GitHub: github.com/eessi

Documentation: eessi.io/docs

YouTube channel: youtube.com/@eessi_community

Paper (open access): doi.org/10.1002/spe.3075

EESSI support portal: gitlab.com/eessi/support

Monthly online meetings (first Thursday, 2pm CEST)

Join our mailing list & Slack channel