Introduction to EESSI

EESSI Community Meeting @ Amsterdam

14 Sept 2022

Caspar van Leeuwen (SURF)

Welcome to the online audience!

Join the EESSI Slack for discussions https://www.eessi-hpc.org/join

Use dedicated channel for this meeting: #community-meeting-2022

Q&A, discussions, feedback on talks, …

About me

- Caspar van Leeuwen
- High Performance Computing, Machine Learning
- Joined SURF 5 years ago

SURF

- Collaborative organization for IT in Dutch Education & Research
- Hosts Snellius, the Dutch national supercomputer

  (among other things)

 https://www.surf.nl/en/research-ict/compute-services

This talk…

… covers the 'why' and 'how' for people new to EESSI!

Everyone else: bear with us, we'll go more in-depth as this meeting progresses! :)

Getting Scientific Software Installed

```
─── INSTALL.SH ───
#!/bin/bash

pip install "$1" &
easy_install "$1" &
brew install "$1" &
npm install "$1" &
yum install "$1" & dnf install "$1" &
docker run "$1" &
pkg install "$1" &
apt-get install "$1" &
sudo apt-get install "$1" &
steamcmd +app_update "$1" validate &
git clone https://github.com/"$1"/"$1" &
cd "$1";./configure;make;make install &
curl "$1" | bash &
```
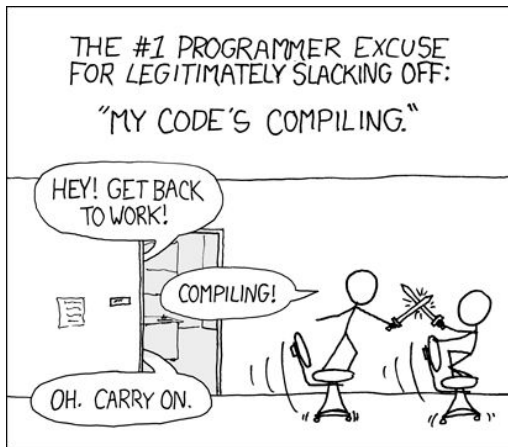
How to make package managers cry
(or)
How to piss off package managers
(pick one)

Kenneth Hoste
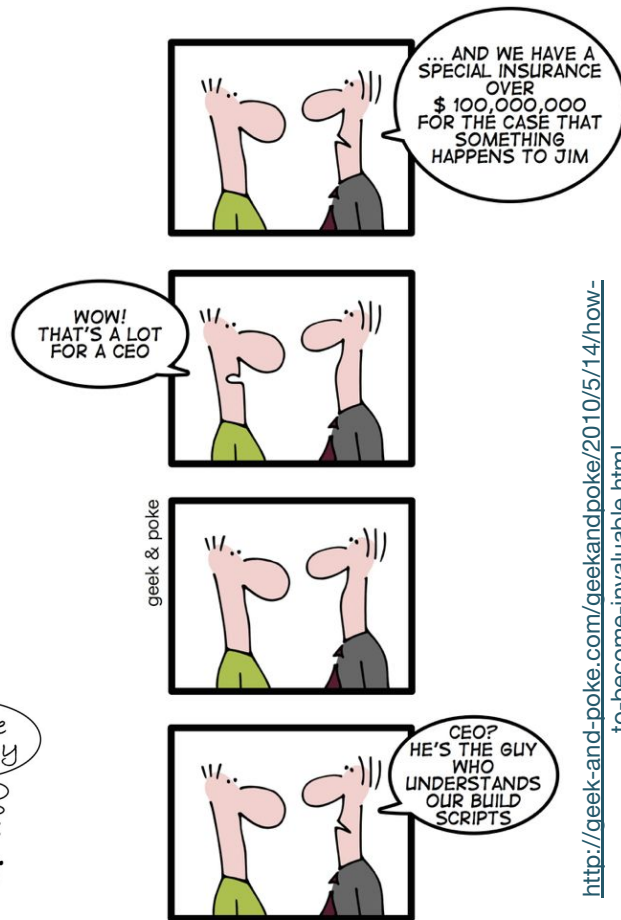
kenneth.hoste@ugent.be   GitHub: @boegel   Twitter: @kehoste

⚙ FOSDEM 2018
Package Management devroom
Feb 3rd 2018, Brussels (Belgium)


https://xkcd.com/303

5

The issue: software deployment

- More scientific disciplines run (large) computations
  - More users
  - Less experience in software deployment, optimization, etc
- Explosion of open source scientific software in recent years
- Increasing variety in hardware
  - CPUs: Intel, AMD, Arm, POWER, RISC-V, …
  - Accelerators: NVIDIA, AMD, Intel Xe, …
- Rise of the cloud: Microsoft Azure, Amazon EC2, Google, Oracle, …
- In stark contrast: available manpower in HPC support teams…

Image: freesvg.org

The issue: software deployment

- How can we properly support users in this increasingly complex environment?

- Can we deploy software in a smarter way?
  - Meeting organized by Dell with interested parties

- The advent of EESSI…

Image: Wikimedia Commons
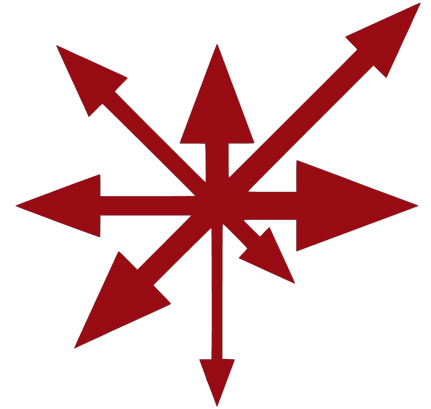
How the EESSI project was started

- May'19: Meeting with Univ. of Groningen, TU Delft, TU Eindhoven, VU Amsterdam at Univ. of Cambridge, organised by Dell Technologies

- Outcome: informal agreement to work together on something that benefits all parties…

- Founding members created *Scientific Software Repository for Compute Clusters (SSR4CC)*

- Jan'20: Follow-up meeting in Groningen

- 5 March'20: Meeting in Delft, joined by Kenneth Hoste (HPC-UGent) to present EasyBuild

- Joint effort was started, project name was changed to EESSI

- Introductory talk on EESSI at HPCKP'20 conference by Kenneth fueled a lot of interest

- Utter lack of meetings with opportunity to have beers together since then… until now!

EESSI partners & interested parties

Founding partners:

Extensive interest from HPC *and* cloud community:

EESSI: Scope & goals

- *European Environment for Scientific Software Installations (EESSI)*

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

- Avoid duplicate work across (HPC) sites by collaborating on a shared software stack

- Uniform way of providing software to users, regardless of system they use!

- Should work on any Linux OS (+ WSL, and possibly macOS) and system architecture
  - From laptops and personal workstations to HPC clusters and cloud
  - Support for different CPUs, interconnects, GPUs, etc.

- **Focus on performance, automation, testing, collaboration**

EESSI
EUROPEAN ENVIRONMENT FOR
SCIENTIFIC SOFTWARE INSTALLATIONS

https://www.eessi-hpc.org

https://eessi.github.io/docs (try out the pilot setup!)

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**
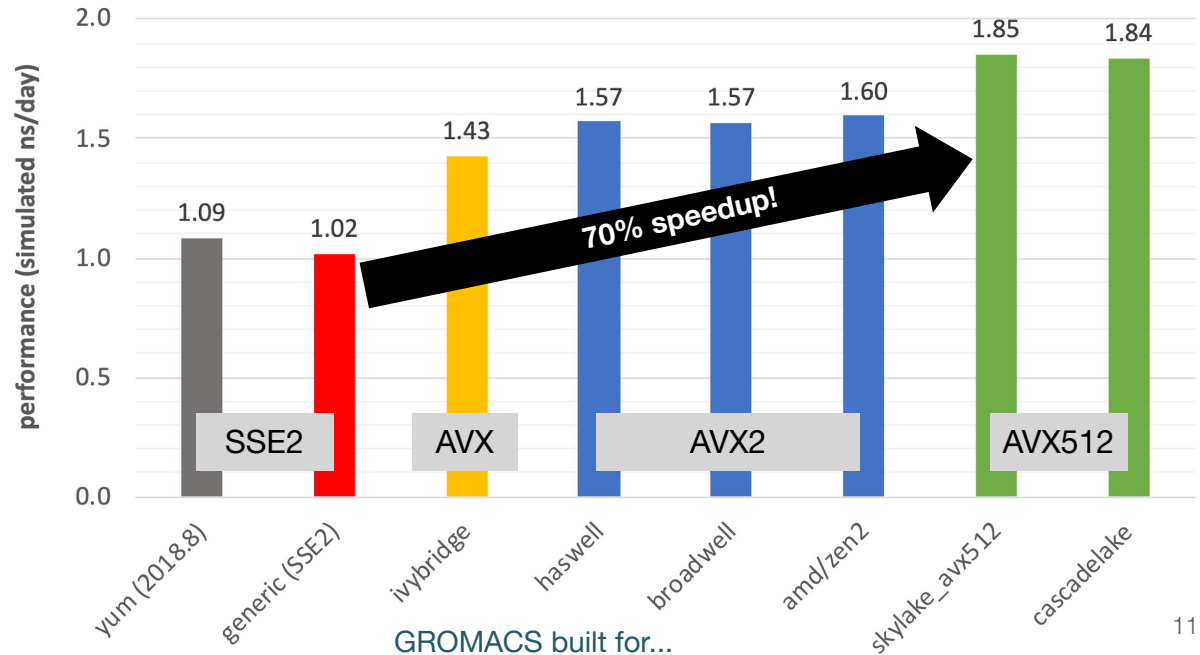
Avoid duplicate work

Current situation

● Use build tools (e.g. EasyBuild, Spack)

  ○ Build procedures can be shared, but each site still installs their own stack

  ○ Build procedures do not always work 'out of the box' …

EESSI

● All contribute to *one* shared software stack

Image: Wikimedia Commons

Provide uniform software stack

Current situation

- Moving from one system to another (e.g. laptop, cloud, HPC cluster) takes effort!
  - Moving data
  - Recreating software environment

Using EESSI

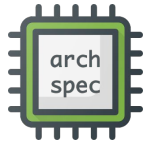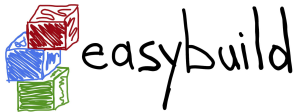- Software environment is identical
- Only move data

Image: rawpixel.com

Sounds great, but… how? We need…

- A way to get the software distributed globally
- Abstraction from the OS (like a container)
- Optimized builds for a large range of hardware architectures
- Selection of the right optimization at runtime (ideally automatic)

Sounds great, but… how? We need…

CernVM-FS

gentoo

- A way to get the software distributed globally

- Abstraction from the OS (like a container)

- Optimized builds for a large range of hardware architectures

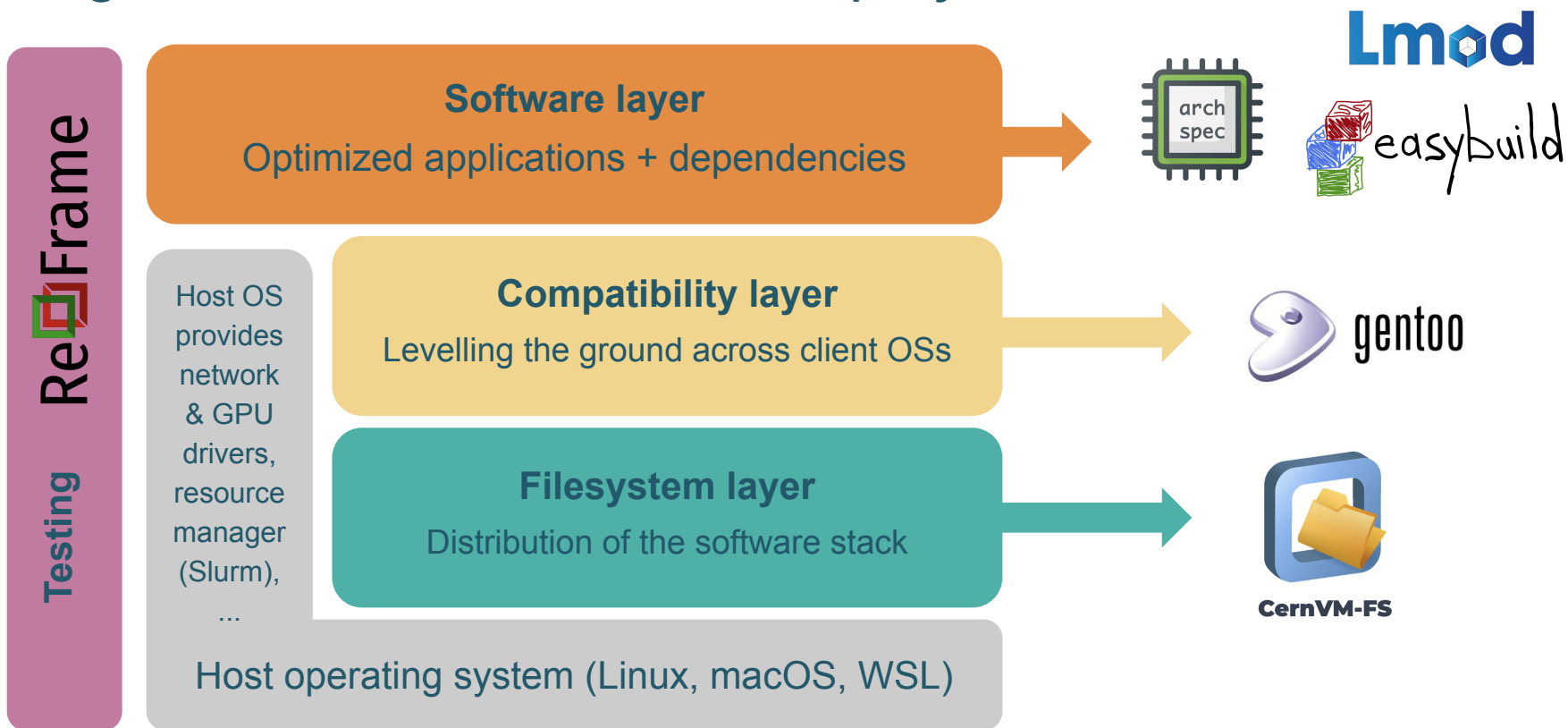- Selection of the right optimization at runtime
(ideally automatic)

easybuild

arch
spec

High-level overview of EESSI project

ReFrame

Testing

Software layer
Optimized applications + dependencies

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

Compatibility layer
Levelling the ground across client OSs

Filesystem layer
Distribution of the software stack

Host operating system (Linux, macOS, WSL)

arch spec

Lmod

easybuild

gentoo

CernVM-FS

intel

AMD

ARM

POWER9

RISC-V

Filesystem layer

CernVM-FS

https://cvmfs.readthedocs.io

HPC cluster X

Cloud A

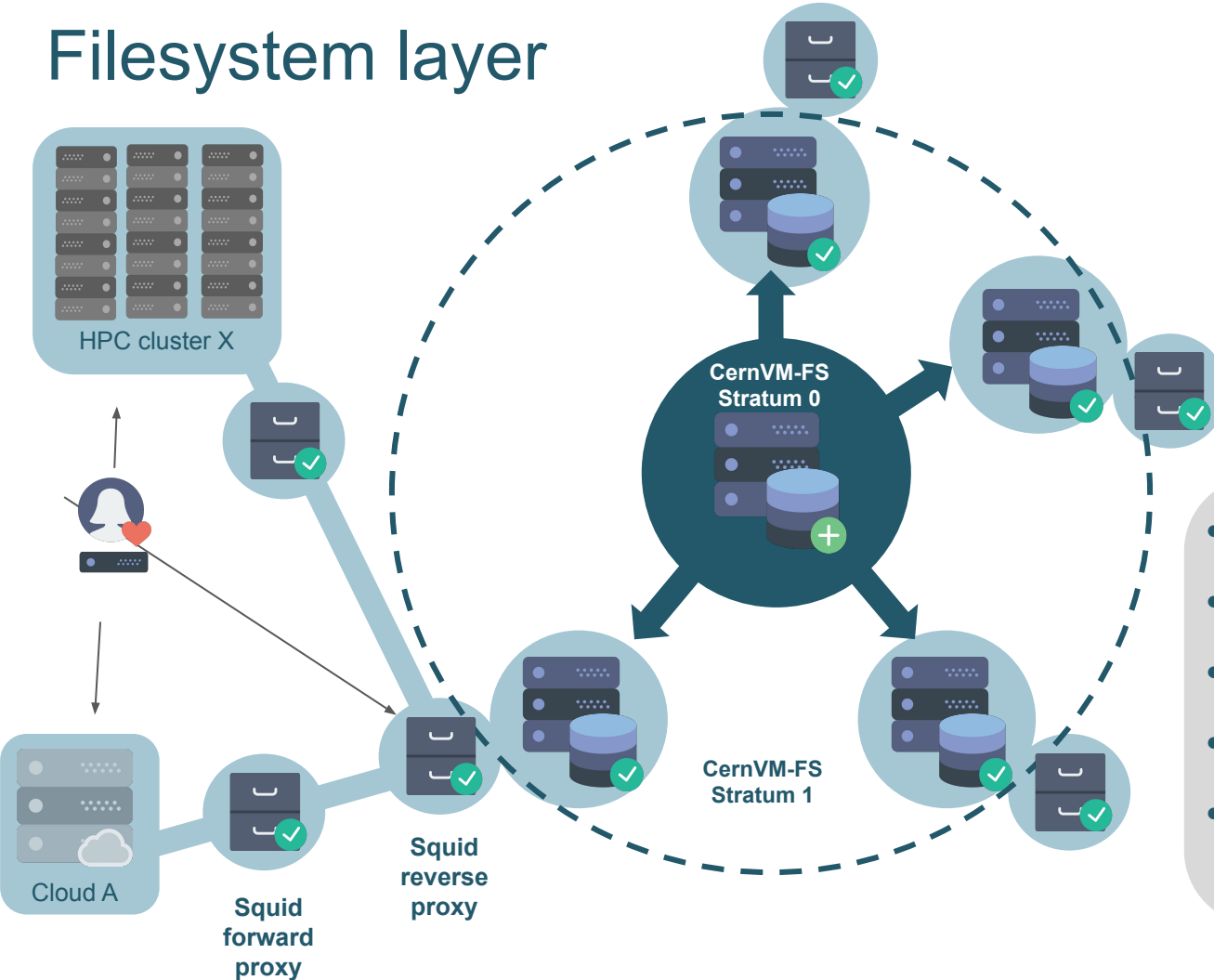**Squid forward proxy**

Squid reverse proxy

**CernVM-FS Stratum 0**

CernVM-FS Stratum 1

- Global distribution of software installations

- Centrally managed software stack

- Redundant network of "mirrors"

- Multiple levels of caching

- **Same software stack everywhere**:

 laptops, HPC clusters, cloud VMs, …

17

Compatibility layer

- **Gentoo Prefix** installation (in `/cvmfs/.../compat/<os>/<arch>/`)

- Set of tools & libraries installed in non-standard location

- Limited to low-level stuff, incl. glibc (no Linux kernel or drivers)
  - Similar to the OS layer in container images

- Only targets a supported processor **family** (`aarch64, ppc64le, x86_64`)

- **Levels the ground for different client operating systems** (Linux distros, later also macOS?)

- Currently in pilot repository:

  `/cvmfs/pilot.eessi-hpc.org/versions/2021.12/compat/linux/aarch64`

  `/cvmfs/pilot.eessi-hpc.org/versions/2021.12/compat/linux/ppc64le`

  `/cvmfs/pilot.eessi-hpc.org/versions/2021.12/compat/linux/x86_64`

E E S S I

powered by

gentoo

|	
Software layer	
Compatibility layer	
Filesystem layer	
host OS	
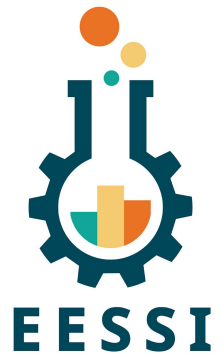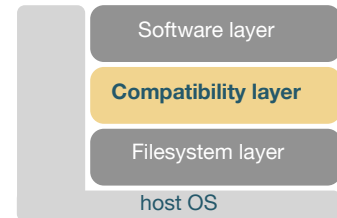
Software layer

- Provides scientific software applications, libraries, and dependencies

- **Optimized for specific CPU microarchitectures** (Intel Haswell, ...)

  - Separate subdirectory/tree for each (in `/cvmfs/.../software/...`)

- **Leverages libraries** (like glibc) **from compatibility layer** (not from host OS)

- Installed with EasyBuild, incl. environment module files

- Lmod environment modules tool is used to access installations

- **Best subdirectory for host is selected automatically** via archspec

E E S S I

powered by

easybuild

Lmod

arch spec

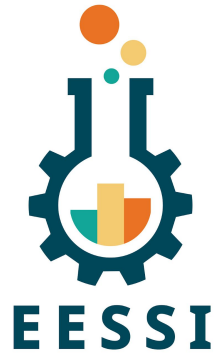| Software layer |
| Compatibility layer |
| Filesystem layer |
| host OS |
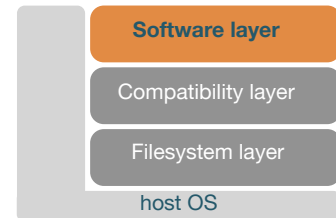
Testing

Software testing is always important, but if everyone uses *the same stack*,

even more so!

- Development of *portable* test suite with ReFrame

- ReFrame: designed to run tests on HPC clusters

  - Can interact with batch systems

  - Allows testing of applications at scale

  - *Also* allows simple tests on a single node (laptop, cloud node)

- More on this in "Testing in software layer", Friday, 9:30 CEST

Current status: pilot repository 2021.12

- Working **proof of concept**

- Ansible playbooks, scripts, docs at `https://github.com/eessi`

- CernVM-FS: Stratum 0 @ Univ. of Groningen + four Stratum 1 servers

- Software (CPU-only): Bioconductor, GROMACS, OpenFOAM, R, TensorFlow, Spark, IPython, Horovod, QuantumESPRESSO, ReFrame, …

- Hardware targets:
    - `{aarch64,ppc64le,x86_64}/generic`
    - `intel/{haswell,skylake_avx512}, amd/{zen2,zen3},`
      `aarch64/{graviton2,graviton3), ppc64le/power9le`

https://eessi.github.io/docs/pilot

Benefits

For HPC support teams:

- Share software deployment effort globally

- Knowledge sharing within the EESSI community

For scientists:

- Easily transition from local laptop, to cloud node, to HPC cluster

- More easily adopt new architectures

- Reproducibility: can run with *the exact same software installation* as another scientist

For hardware providers / funding agencies

- More (scientific) work done with the same hardware (due to optimization)

… And many, many more (see "EESSI use cases" talk @ 14:00)

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

Website: https://www.eessi-hpc.org

Join our mailing list & Slack channel
https://www.eessi-hpc.org/join

Documentation: https://eessi.github.io/docs

GitHub: https://github.com/eessi

Twitter: @eessi_hpc

YouTube channel (brand new!)

Monthly online meetings (first Thursday, 2pm CEST)