



EESSI meeting

3 Mar 2022

<https://github.com/EESSI/meetings/wiki>

Agenda



1. Quick introduction by new people
2. EESSI-related meetings in last month
3. Progress update per EESSI layer (+ GitHub App, MPItrampoline)
4. Monitoring EESSI infrastructure
5. 2021.12 version of pilot repository
6. AWS/Azure sponsorship update
7. Update on EESSI paper
8. S4 project proposal for NeIC call
9. Past/upcoming events
10. Q&A

Quick introduction by new people



New people on the call: feel free to introduce yourself!

- Who are you, where do you work, on what?
- Why are you interested in the EESSI project?
- Are you planning to actively contribute,
and if so, to which aspect(s) of the project?

EESSI-related meetings (1/2)



Feb 9th: CernVM-FS coordination meeting

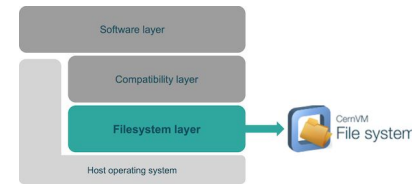
- CernVM workshop (Sept 12-14 2022, Amsterdam) will be in-person with remote possibilities
 - Registration is open: <https://indico.cern.ch/e/cvm22>
(you can still change from in-person to remote and vice versa later on)
 - Suggestions for talks are welcome: EESSI talk? Study using CernVM-FS for MPI workloads?
 - **Maybe an interesting opportunity for an EESSI community meetup (Sept 15-16)?**
- CernVM-FS v2.9.1 release expected soon, nightly builds are available
 - Includes a fix for the GEO API + Python 3 issue

EESSI-related meetings (2/2)



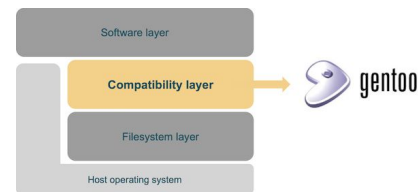
- Feb 18th: EESSI/Azure sync
 - Limited attendance: only Alan + Bob + Kenneth (EESSI) + Davide (MS Azure)
 - Nice discussion on way to support NVIDIA GPUs in EESSI
 - Notes available at <https://github.com/EESSI/meetings/wiki/Azure-meeting-Feb-18-2022>

Progress update: filesystem layer



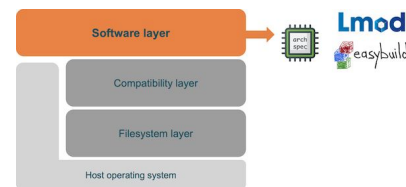
- Installed a nightly build of CernVM-FS v2.9.1 on the Stratum 1 servers running (RH)EL 8
 - This has fixed the issue with GeoAPI that was found during the hackathon
 - Fix was immediately detected by the monitoring page: <https://monitoring.eessi-infra.org> \o/
- How to properly set up a CernVM-FS data repository for hosting data files for testing?
 - Requested by Hugo for, initially, WRF; these files can be very large
 - CernVM-FS docs have recommendations: <https://cvmfs.readthedocs.io/en/stable/cpt-large-scale.html>
 - An external web server, optimised for large files, is required
 - The CernVM-FS client will be redirected to that external server when accessing the file in the repository

Progress update: compatibility layer



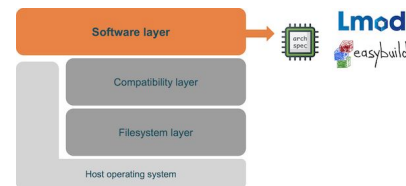
- No security updates required for 2021.06 and 2021.12 compat layers (as reported by Gentoo's `glsa-check` tool)
 - There were some new GLSA's, but only for web browsers
- Project idea submitted for Google Summer of Code in Gentoo project
 - *"RISC-V support for Gentoo Prefix"*
 - Main goal: make it possible to bootstrap and use a Gentoo prefix system on RISC-V
 - https://wiki.gentoo.org/wiki/Google_Summer_of_Code/2022/Ideas/RISC-V_support_for_Gentoo_Prefix

Progress update: software layer (1/3)



- [PR #166](#) by Hugo: proposal for `eessi-init` command to simplify setting up EESSI environment
- [PR #168](#) by Kenneth: flesh out script to (re)generate Lmod cache
 - Motivated by (manually resolved) problem with incorrect Lmod cache (see [issue #165](#))
 - Lmod cache for all target CPUs can be generated on any system (for example on Stratum-0)
- [PR #170](#) by Hugo: proposal for config file and script for managing datasets
- [Issue #169](#): ReFrame no longer available through software layer, only via compat layer (old version)
 - ReFrame should probably be provided (only) through software layer (easier to keep up-to-date)
 - ReFrame that is needed for testing compat layer can be installed on-the-fly when running tests

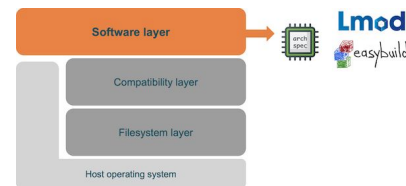
Progress update: software layer (2/3)



Progress towards a GitHub App for processing pull requests to add software to EESSI

- PyGHee (“piggy”) is a Python library that simplifies implementing GitHub Apps
- Currently (v0.0.1) takes care of logging activity + event data, verifying incoming events, collecting event info, calling event handler (if one is defined)
- Includes CI workflow to verify that PyGHee functionality works as designed
- Available at <https://pypi.org/project/PyGHee> (GPLv2), install with `pip3 install PyGHee`
- Basic documentation available in README at <https://github.com/boegel/pyghee>
- **Excellent starting point for implementing a bot for EESSI software layer GitHub repo!**

Progress update: software layer (3/3)



MPItrampoline - wrapper library to dynamically switch between different MPI implementations
(see [EasyBuild Tech Talk V](#))

- Toolchain support merged into EasyBuild develop branch (see [PR #3971](#))
- Tested with OSU benchmarks, works as advertised (see [WIP easyconfigs PR#15018](#))
- Tested with GROMACS: no noticeable performance impact in initial testing
- Lots of development via collaboration with developer (7 releases in a week)
 - Some software requires patching due to not respecting the MPI standard
- Fortran seems to be current weak point
 - Successful compilation of CP2K but lot's of RUNTIME FAIL in test suite
- MPIwrapper will require extensive testing (but we can be very clever in the easyblock)

EESSI pilot repository

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

**NOT FOR
PRODUCTION USE!**



- 2021.06: considered “final” (no further changes, except security updates in compat layer if needed)
- Current status for 2021.12 (mostly unchanged since EESSI monthly meeting of Feb’22)
 - Compatibility layer: in place for `aarch64` / `ppc64le` / `x86_64`
 - List of installed packages is pretty much identical \o/
 - Software layer:
 - Software installations included in 2021.06 also in place for 2021.12, incl. GROMACS, OpenFOAM, TensorFlow + Horovod, R + Bioconductor, QuantumESPRESSO
 - Additional software (vs 2021.06): SciPy-bundle with foss/2021a (excl. `ppc64le`), WRF
 - **Problem with incorrect Lmod cache files fixed (see [issue #165](#))**
 - TODO:
 - Move “latest” symlink from 2021.06 to 2021.12
 - Update documentation at <https://eessi.github.io/docs/pilot>

Progress update: infrastructure (1/2)



- Now using <https://runatlantis.io> for Terraform pull request automation
 - Running on EC2
 - Atlantis, PRs, and `terraform plan` caused the infrastructure repo to go private
- Refactored a lot of Terraform code to modularize common tasks
 - Create nodes (automates DNS, images, backup...)
 - AMI image selecting
 - Configuration of backup on AWS
- Proper and automated backup for long-life persistent nodes (on AWS)
 - Uses AWS Backup Vault
 - Nodes created with the `aws-core-node` module automatically get backed up via labels

Progress update: infrastructure (2/2)



- A proper testing environment
 - Available through Terraform / Atlantis
 - Its own Terraform structure and state
 - Its own DNS subdomain (`testing.eessi-infra.org`)
 - Automated with the new modules (creating a node within the testing environment changes a lot of settings, does not enable backup, etc.)
- Ansible roles for EESSI (@ [terjekv/ansible-eessi-roles](https://github.com/terjekv/ansible-eessi-roles))
 - Still being tested
 - Needs Ansible 2.10 and Jinja 3.*, but note that updating Ansible does not update Jinja...
 - Works if you're not Bob. Or at least if you're Terje. Ansible is a fickle beast.
 - More testing this month, easier with a proper testing environment

Progress update: monitoring (1/4)





- Monitoring v2, Grafana + Prometheus + [cvmfs-server-scraper](#)
 - This solution does away with `cvmfs-servermon` and `blackbox-exporter` hackery
 - `cvmfs-server-scraper` is a Python library that scrapes the public API of CVMFS servers
 - Not EESSI/Prometheus specific, but offers a Prometheus exporter
 - Prometheus gets raw data, not `cvmfs-servermon` status data, so one can thus create alerts based on metrics of one's own choosing
 - All checks from `cvmfs-servermon` are implemented, as well as a few new ones:
 - Track CernVM-FS repo versions across all nodes
 - List differences between S0 and S1s
 - Track last timestamps and repo ttls.

Progress update: monitoring (2/4)



Status page: <http://status.eessi-infra.org> (<https://monitoring.eessi-infra.org>)


EESSI :: Status 



Normal service

EESSI services operating without issues.


✓ Normal service ⚠ Degraded ⚠ Warning ✖ Failed ➔ Maintenance



Stratum0

✓


- cvmfs-config : 31
- pilot : 356



Repositories

✓

	Version	Snapshot
pilot	✓	✓



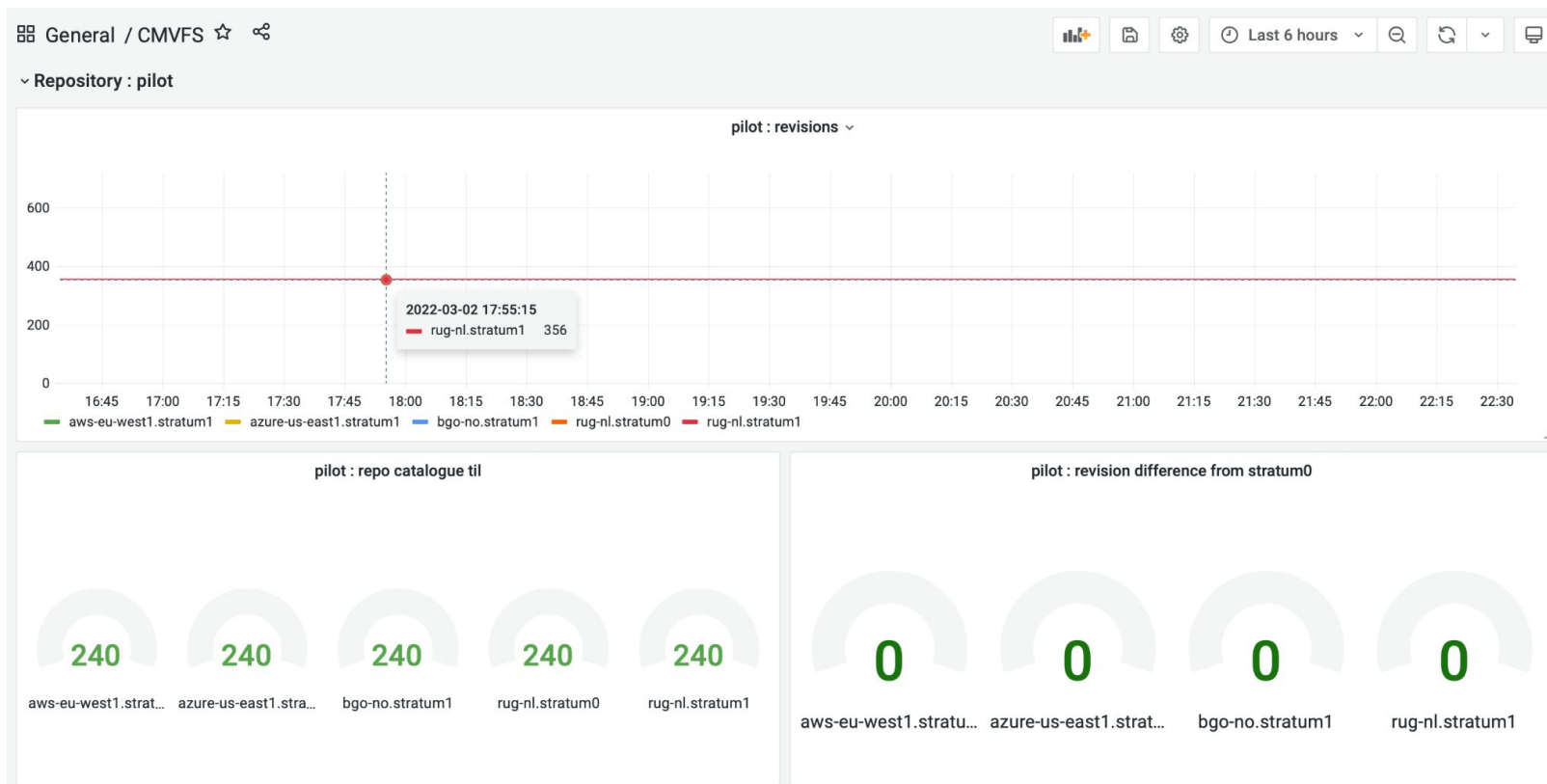
Stratum1s

✓

	Updates	GeoAPI
aws-eu-west1	✓	✓
azure-us-east1	✓	✓
bgo-no	✓	✓
rug-nl	✓	✓

Last updated 2022-03-03T12:50:01 | contact@eessi-hpc.org

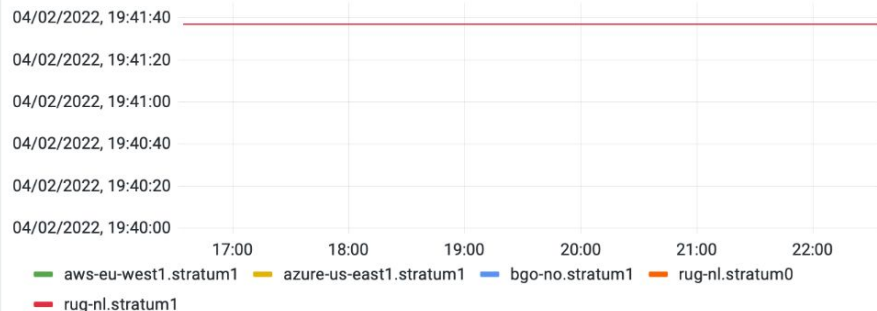
Progress update: monitoring (3/4)



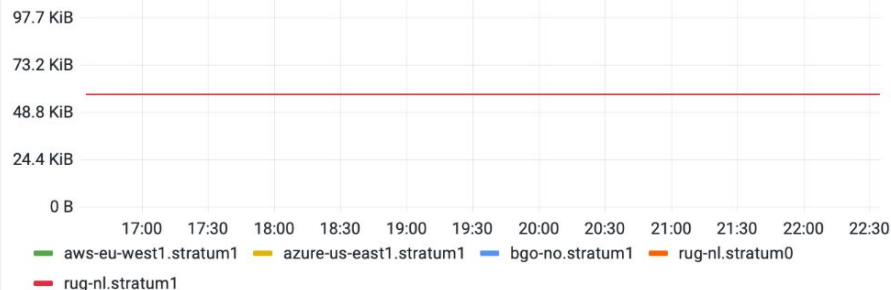
Progress update: monitoring (4/4)



pilot : last timestamp



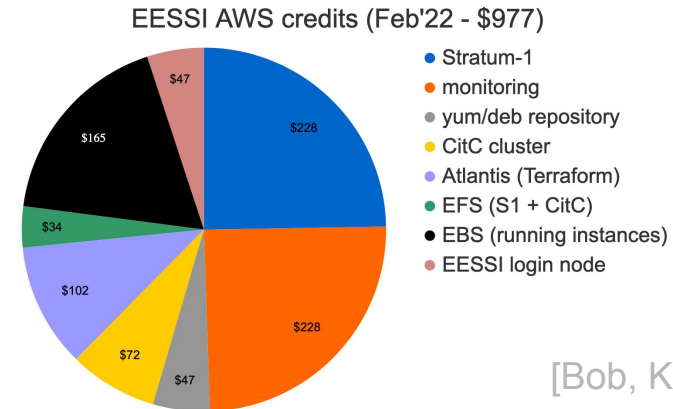
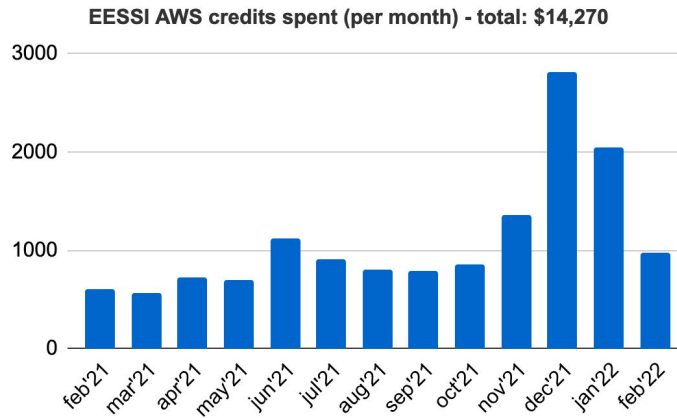
pilot : repo root size



Usage of sponsored AWS credits



- **Ask in #aws-resources Slack channel to get access!**
- Original batch of \$25,000 worth of sponsored credits expired on Jan 31, 2022!
- Request for new credits is WIP, extra \$10k worth of credits already received to bridge the gap
- In February '22: ~\$977 worth of credits spent on Stratum-1 server, monitoring node, ...
- ~\$14,270 worth of credits spent in total so far (since Feb'21), all covered by sponsored credits
- Christian left AWS, Ollly Perks (ex-Arm, now AWS) is likely to become new contact

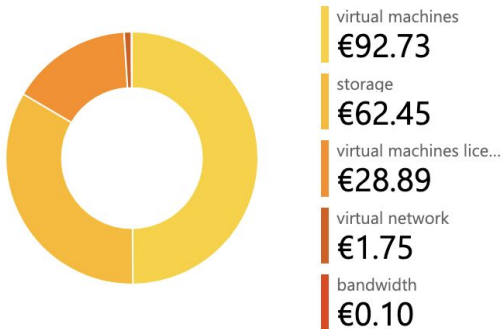


Azure sponsorship

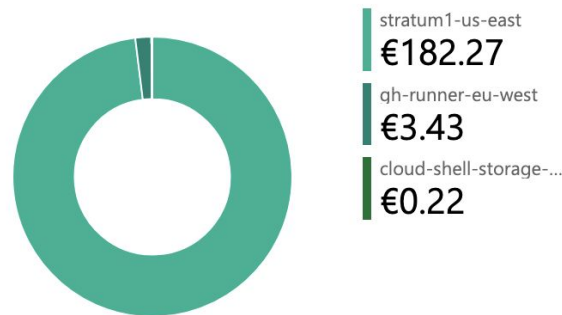


- Sponsored credits (€40,000) are being put to good use!
- **Ask in #azure-resources Slack channel to get access!**
- In Feb'22: ~€ 185 worth of credits spent, pretty much all on Stratum-1
- ~€2,655 worth of credits spent in total (since Sept'21)

Service name ▾



Resource group name ▾



Update on EESSI paper: published!



- ***EESSI: A cross-platform ready-to-use optimised scientific software stack***
- Authors: Bob Dröge (Univ. of Groningen), Victor Holanda Rusu (CSCS), Kenneth Hoste (HPC-UGent), Caspar van Leeuwen (SURF), Alan O'Cais (JSC - CECAM), Thomas Röblitz (Univ. of Bergen)
- Published in:
 - Journal of Software: Practice and Experience (<https://onlinelibrary.wiley.com/journal/1097024x>)
 - Special issue *New Trends in High-Performance Computing: Software Systems and Applications*
- Original submission on 31 May 2021, revised on 19 Nov, accepted on 24 Jan, **published 16 Feb 2022**
- Thanks to everyone who contributed!
- **Open-access paper available at** <https://dx.doi.org/10.1002/spe.3075>
- Spread the word!

S4 project proposal for NeIC call



- New proposal was submitted for NeIC call: **Scientific Software Stacks as a Service (S4)**
- <https://neic.no/news/2022/01/12/invitation-for-development-projects> (deadline 1 March 2022)
- Initiative and most of writing in 10-page proposal by Thomas Röblitz (Univ. of Bergen)
- 3 activities:
 - EESSI development & deployment
 - Community specific software stacks & support for application developers
 - Training & outreach
- Project partners:
 - Nordic: Aalto Univ, KTH, Sigma2, Umeå Univ, Univ. of Oslo, Univ. of Tartu
 - Non-Nordic (min. involvement): CERN, CSCS, SURF, RUG, UGent + VUB (VSC)
- Total effort of ~5 FTEs for 2 years, ~75% for Nordic partners
- Decision expected September '22, start of project in summer 2023 (if accepted)

[Thomas, Kenneth]

Past/upcoming events



- Presentation on EESSI for Simula Research Lab (simula.no) by Thomas
 - On Wed 16 March 2022
 - Simula does research on HPC topics, and has testbed with different hardware
- Call for project proposals for BioHackathon Europe 2022
 - <https://biohackathon-europe.org> - <https://biohackathon-europe.org/projects.html>
 - 7-11 November 2022 in Paris
 - Maybe an interesting venue to promote EESSI?
 - Submission are lightweight (max. 300 words) - deadline: 8 April 2022