EESSI

EUROPEAN ENVIRONMENT FOR
SCIENTIFIC SOFTWARE INSTALLATIONS

7 Sep 2023

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

Agenda

1. Quick introduction by new people

2. EESSI-related meetings and events in last month

3. Progress update per EESSI layer (incl. build-and-deploy bot + test suite)

4. EESSI pilot repository (2023.06)

5. Outlook to new EESSI repository (*.eessi.io)

6. EESSI support portal

7. AWS/Azure sponsorship update

8. Update on MultiXscale EuroHPC project

9. 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

- *(8 Aug'23)* MultiXscale sync meeting (WP1+WP5) ([notes](#))

- Sync meetings on EESSI test suite (notes: [9 Aug'23](#), [25 Aug'23](#), [6 Sept'23](#))

- *(10 Aug'23)* EESSI/AWS sync meeting ([notes](#))

- *(21 Aug'23)* EESSI/Azure sync meeting ([notes](#))

- *(5 Sept'23)* Sync meeting on building software for EESSI 2023.06 ([notes](#))

- *(5 Sept'23)* Call with CernVM-FS developers on Best Practices in HPC tutorial ([notes](#))

Progress update: filesystem layer
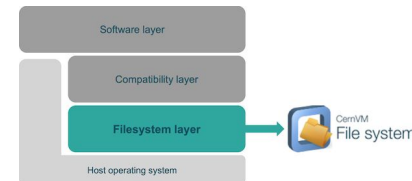
- New Stratum-0 server for EESSI at Univ. of Groningen
    - Plan is to use this for `*.eessi.io` CernVM-FS repositories (`software.eessi.io`?)
    - Hardware is up and running - needs work on RAID, network, firewall cfg, yubikeys (Bob)
        - Expected to be done in the next couple of days
    - Determine access rules: who can help administer it?

- Still need to figure out performance issues with Stratum-1 @ RUG (issue #151)

- Automated ingestion is having (performance) issues due to the increasing number of tarballs
    - Hitting a GitHub API rate limit

- CernVM-FS 2.11.0 was released (see release notes)
    - Performance improvements, telemetry support, …

[Bob, Kenneth]

Progress update: compatibility layer
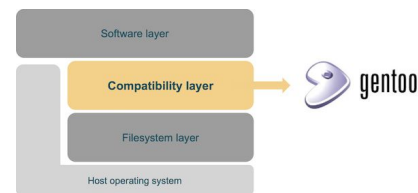
- 2023.06 version is available in EESSI pilot repository (`pilot.eessi-hpc.org`)

  - Was built with bot (see [PR #188](#))

  - Deployment was still done manually, needs more work to automate (see [issue #189](#))

- **2023.04 version of compat layer will be removed from pilot repository**

  - There will be no software layer on top of 2023.04 compat layer,

    too many problems with OpenSSL 3.x (see issues [#257](#), [#258](#), [#259](#))

- First steps to build compat layer for `software.eessi.io` repo (version 2023.09) [Thomas]

  - Basically same setup as in EESSI pilot 2023.06, but with repo under `eessi.io`

  - May be a nice opportunity to also let bot take care of deployment (not just building it)

[Bob, Thomas, Kenneth]

Progress update: software layer
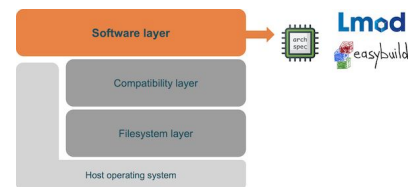
- Incorrect CPU detection by archspec for `zen3` ([issue #322](#)) and `neoverse_v1` ([issue #320](#))

- Fix for `archdetect` ([PR #264](#)), so we can use it by default in EESSI init script

- Script to inspect build job (WIP) ([PR #317](#))

- EasyBuild PRs relevant to EESSI got merged, to be included with EasyBuild v4.8.1

  - Correctly determine path to active binutils in TensorFlow easyblock ([PR #2218](#))

  - Patch Java binaries/libraries to ensure correct glibc is used with alternate sysroot (PR [#2557](#) + [#2995](#))

  - Strip out hardcoded `-march=native` used by RapidJSON 1.1.0 ([PR #18725](#))

- Supporting containers in EESSI ([issue #323](#)) like QIIME2 (cfr. approach by The Alliance)

- Blacklisting applications that use the `Conda` easyblock ([issue #324](#)) like QIIME2, FSL, …

- Software layer for EESSI pilot 2023.06 is gradually being populated (details in next slides)

[Kenneth, Thomas, Lara, Caspar, Satish, Xin]

Bot for building + deploying software layer

Progress on implementation of build-and-deploy bot

https://github.com/EESSI/eessi-bot-software-layer

- Existing bot implementation has been working as designed for software-layer PRs
    - Bot takes care of building + deploying of software installations
    - Log for failing builds is copied to shared directory on AWS CitC cluster for inspection

- 14 PRs (all merged) to clean up existing code base in preparation of v0.1.0 release
    - Cleaning up code style, comments, docstrings - no functional changes
    - Issues opened for more intrusive changes that can be tackled later

- Failing to apply PR patch should be reported in PR (issue #212)

[Thomas, Kenneth, Bob]

EESSI pilot repository (1/3)

https://eessi.github.io/docs/pilot
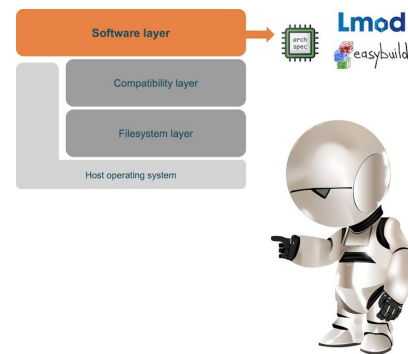
NOT FOR PRODUCTION USE!

- 2021.12 version is "frozen", no more changes planned there, but it's still the default ("`latest`") version

- 2023.06 is being populated via PRs to software-layer repo + build-and-deploy bot
  - Supported CPU targets: see http://www.eessi.io/docs/software_layer/cpu_targets
  - Recently added software installations (for all supported CPU targets):
    - `SciPy-bundle/2021.10-foss-2021b` (PR #306) + issue #318 (numpy test suite fails on `neoverse_v1`)
    - `GROMACS/2021.5-foss-2021b` (PR #304)
    - `foss/2022b` (PR #309) + issue #314 (additional LAPACK test suite failures on `neoverse_v1`)
    - `Arrow/6.0.0-foss-2021a` (PR #316) + `EasyBuild/4.8.0` (PR #319)
    - `Java/11` (PR #327), incl. fix for `create_tarball.sh` script to correctly handle `.modulerc.lua`
    - Fix installation of ReFrame 4.2.0 (include hpctestlib) (PR #311)
    - Don't filter cURL as dependency in EasyBuild (PR #326)
      - To fix problems observed with building R + TensorFlow
      - `CMake` and `git` modules were reinstalled, since they depend on cURL

[Kenneth]

EESSI pilot repository (2/3)

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

NOT FOR PRODUCTION USE!

- 2023.06 is being populated via PRs to software-layer repo + build-and-deploy bot

    - WIP:

        - `TensorFlow/2.7.1-foss-2021b` (20 missing installations, build running) via PR #321

        - `foss/2022a` via PR #310 (trouble with FFTW tests (issue #325) + OpenBLAS on aarch64 (WIP))

        - WRF/4.3-foss-2021a via PR #290 (custom hook for WRF configure needs to be updated)

        - R 4.1.0 + RStudio-Server via PR #299 (failed because of cURL, should be retried)

        - Fix for `archdetect` (PR #264), so we can use it by default in EESSI init script

        - Script to inspect build job (PR #317)

    - TODO:

        - OpenFOAM, R + Bioconductor (to bring on par with EESSI 2021.12)

        - ESPResSo, waLBerla, LAMMPS (in context of MultiXscale)

        - Ensure that Lmod cache update is done correctly, includes *all* available modules

        - Come up with a better approach to replace existing software installations

[Kenneth]

EESSI pilot repository (3/3)

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

NOT FOR PRODUCTION USE!

Policy to deal with limited number of failing tests that only occur on a particular CPU target

- Assess whether failing tests really seem to signal a serious problem
    - Very small number of (additional) failing tests => probably not a reason to block installation
    - Only for a specific CPU target: `neoverse_v1` is Arm 64-bit with SVE vector instructions support

- Put workaround in place to ignore failing tests (yet retain test suite result in EasyBuild installation log)
    - Should be relatively easy to do via the custom EasyBuild hooks we use (`eb_hooks.py`)

- **Open tracker issue** in EESSI/software-layer repo that provides all details + allows following up

- **Report the problem upstream**, try and get feedback to assess severity of the problems + how to fix

- Examples: OpenBLAS, FFTW, numpy on `aarch64/neoverse_v1` (issues #314, #318, #325)

- TODO: Properly document this policy, should probably be part of contribution policy (docs PR #108)

[Kenneth]

Contribution policy (proposal)

- **Proposal** for contribution policy for adding software to EESSI ([docs PR #108](#))
  - Preview available [here](#) - will be reworked based on feedback - **more feedback welcome!**
  - Initial policy - can be revised later as needed
- Summary:
  - Only open source software (we should verify this by requiring [SPDX license IDs](#))
  - Software must be built by the bot (no manual builds)
  - Software must be supported by latest EasyBuild release (can be relaxed later);

    `--from-pr` and `--include-easyblocks-from-pr` should only be used for *merged* PRs
  - A compiler toolchain still supported by latest EasyBuild release must be used
  - Ideally all software is installed for all supported CPU targets (exceptions allowed)
  - Recent software versions and toolchains should be preferred
  - There should be a way of testing the installations - ideally via the [EESSI test suite](#)
- To add: policy being applied when limited number of tests fail for a specific CPU target

[Kenneth, Bob]

EESSI test suite

Pull requests:

- TensorFlow bug fix (PR #79, merged)

- Updated Vega config for ReFrame 4.X performance formatting (PR #78, merged)
  - Past runs did not properly log performance numbers. Now ~30 days logging data.

- Make automatic module discovery more specific (PR #81, merged)
  - Generate TensorFlow test for *TensorFlow/2.6.0-foss-2021a-CUDA-11.3.*1 but not for *Horovod/0.22.1-foss-2021a-CUDA-11.3.1-TensorFlow-2.6.0*

- OSU point-to-point works (CPU and GPU), collectives: WIP (PR #54)

Working towards v0.1.0 release (almost there)

- Restructure example configuration files to define common logging behaviour (PR #80)

- Add up-to-date documentation on installing + using EESSI test suite (WIP, see docs WIP PR #110)

- `pip install` of test suite not working with RHEL 8 OS pip/python (issue #82)

[Caspar, Sam, Satish, Kenneth]
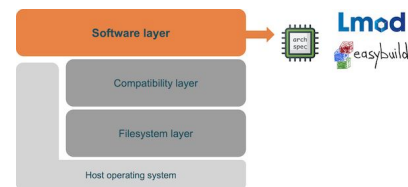
Infrastructure + monitoring

http://status.eessi-infra.org

- Update the CVMFS scraper to modern typed Python (completed)

- Add proper CI and tests for the scraper (completed, but more tests desired)

- Clean up internal scraper structure (first pass completed, possible migration to pydantic)

- Add proper documentation (not done)

- Add support for CVMFS infrastructures to the scraper (not done)

- Release new version of the scraper (not done)

- Use the new scraper release with the status page generator (not done)

- Overhaul the status page generator with functionality from the new scraper

- Release new status page generator and redo the status page delivery (own service)

E E S S I

[Terje]

Support portal for EESSI (MultiXscale task 5.1)

- **EESSI support portal** set up at https://gitlab.com/eessi/support
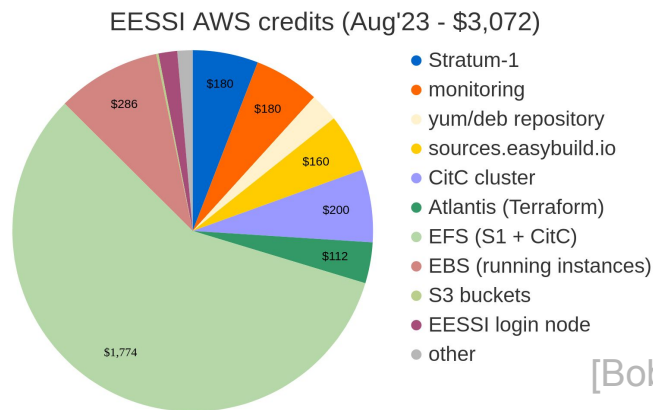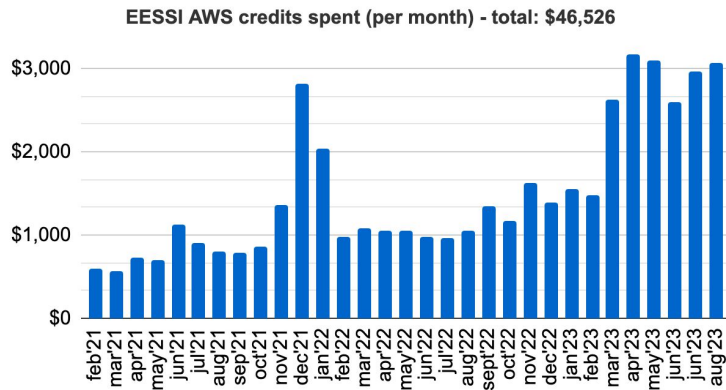 - Tickets can be created by opening an issue in GitLab, or sending an email to support@eessi.io
  - Labels for issues are updated (https://gitlab.com/eessi/support/-/labels)
  - Repository updated (templates for issues, replies, …)
  - Wiki updated (not public, internal docs for EESSI support team members)
  - Open PR #109 to add software portal page to EESSI docs with draft of initial support policy for EESSI
  - Proposal for the support rotation in the wiki (https://gitlab.com/eessi/support/-/wikis/Proposal-for-rotation)
- **TODO**
  - Feedback on initial support policy for EESSI (what is supported, what is not, …)
  - Set up and document support portal
    - Add page on "Getting support" to EESSI documentation (see PR #109)
  - Feedback on proposal for support rotation (only relevant for MultiXscale partners involved in that task)
  - Make sure that EESSI support team members have access to the repository
  - Schedule regular sync meetings + hand-over between support rotations

[Lara, Kenneth, Alan, …]

Sponsored AWS credits

- Ask in #aws-resources Slack channel to get access!

- Currently ~$7,230 worth of sponsored credits left (valid until 30 Nov'23)

- ~$3,072 "spent" in Aug'23 on Stratum-1, monitoring, sources.easybuild.io, **Slurm cluster (build bot)**

- ~$46.5k worth of credits spent in total so far (since Feb'21), all covered by sponsored credits

- **Increase in consumed credits due to extensive activity with build-and-deploy bot**

- Growing large cost for Elastic File System (EFS) service - mostly due to huge tarballs created by the bot?

- Monthly sync meetings with Brendan/Angel/Matt/Francesco (AWS) every 2nd Thursday of the month

EESSI AWS credits spent (per month) - total: $46,526

EESSI AWS credits (Aug'23 - $3,072)

- Stratum-1 — $180
- monitoring — $180
- yum/deb repository
- sources.easybuild.io — $160
- CitC cluster — $200
- Atlantis (Terraform) — $112
- EFS (S1 + CitC) — $1,774
- EBS (running instances) — $286
- S3 buckets
- EESSI login node
- other

[Bob, Kenneth]

Sponsored Azure credits

- Sponsored credits (€40,000) are being put to good use!

- **Ask in #azure-resources Slack channel to get access!**

- In Aug'23: ~€1,447 worth of credits spent

- ~€18.3k worth of (sponsored) credits spent in total (since Sept'21)

- Used for: Stratum-1, GitHub Runners, heterogeneous Slurm cluster, Ampere Altra build node

- Virtual Slurm cluster in Azure

  - Current setup using Azure Cyclecloud is WIP - should start over with Magic Castle instead

EESSI Azure credits spent (per month) - total: €18,346
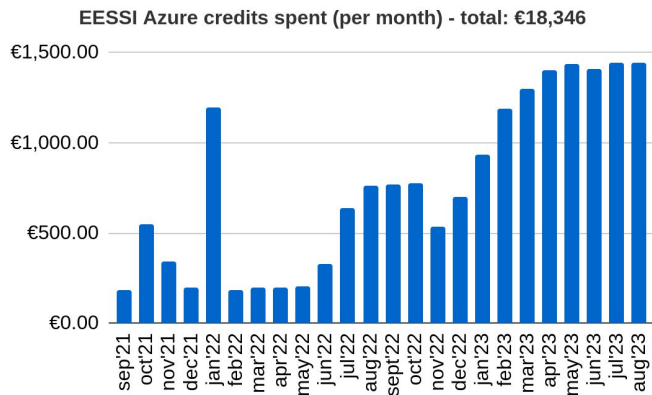
Service name ⌄	Resource group name ⌄
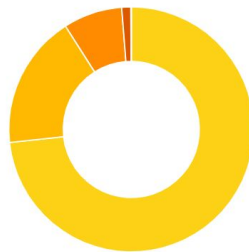
Virtual Machines €1,060.62
Storage €255.67
Virtual Machines Lice... €112.41
Virtual Network €17.34
Bandwidth €0.89

kh-build-altra-rhel86 €541.82
kh-cyclecloud-slurm... €448.21
gh-runner-eu-west €236.68
stratum1-us-east €217.60
azure-mc_resource_... €2.12

Aug'23

[Bob, Kenneth]

Multi✕scale

www.multixscale.eu   github.com/multixscale

- Presentation by Alan to NCCs about MultiXscale training program (which has EESSI focus)

- CI/CD collaboration via CASTIEL2

  - Pathway to making EESSI available on different EuroHPC JU systems
  - **EESSI already available on Vega (Slovenia), Karolina (Czech Republic)**
  - Discussion with Meluxina (Luxembourg) ongoing - access available already
  - Deucalion (Portugal) debuted last week, contacted them for follow-up

- "Best Practices for CernVM-FS on HPC systems" training event

  - Being developed in https://github.com/multixscale/cvmfs-tutorial-hpc-best-practices
  - In collaboration with CernVM-FS developers & experts
  - Date to be determined (tentative Mon 4th Dec'23) - most likely fully virtual event

- Deliverables due end of 2023

  - D1.1: Report on shared software stack prototype
  - D1.2: Plan for design of a portable test suite
  - D5.1: Community contribution policy and GitHub App
  - D5.2 : Support portal for EESSI

[Alan, Kenneth, Lara, Bob, Thomas, Caspar, Satish, Elisabeth, …]

EESSI @ PackagingCon'23
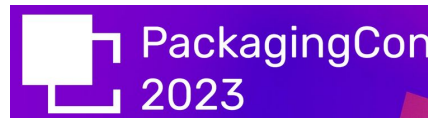
- 26-28 Oct'23 in Berlin (Germany), hybrid event - https://packaging-con.org

- Conference on software package management (in broad sense)

- Talk submitted on EESSI by Lara & Kenneth - **was accepted \o/**

  "Streaming optimized scientific software installations on any Linux distro with EESSI"

- In process of getting approval for attending PackagingCon'23

  - Talk can be pre-recorded in case only remote attendance is required

- Nice opportunity to present EESSI, CernVM-FS, MultiXscale to broad audience

  people with extensive experience in software packaging and installation

[Kenneth]

EESSI @ Supercomputing'23?

- 12-17 Nov'23 in Denver (US) - https://sc23.supercomputing.org

- Who is planning to attend? (confirmed: Alan, HPCNow!)

- Planned activities:
    - Magic Castle tutorial accepted (Alan)
    - Submission on EESSI for HPC User Support Tools (HUST) workshop (Alan+Kenneth)
    - Booth talks @ Microsoft + AWS?
        - No lightning talks this year at AWS booth, but Azure should be possible?
    - MultiXscale presence in EuroHPC booth?
        - Video only (work in progress by HPCNow!)
        - No swag…but if we bring some stickers maybe they will not complain?

[Alan, Kenneth]