EESSI meeting

Dec 2nd 2021

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

Agenda

1. Quick introduction by new people
2. EESSI-related meetings in last month [Kenneth, Bob, Caspar]
3. Progress update per EESSI layer [Kenneth, Bob, Caspar]
4. 2021.12 version of pilot repository [Kenneth, Bob]
5. AWS/Azure sponsorship update [Kenneth, Bob]
6. Update on EESSI journal paper [Thomas]
7. EESSI hackathon(s) [Kenneth, Bob]
8. Upcoming events: Computing Insight UK [Jörg]
9. 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

- Nov 9th: monthly CernVM-FS coordination meeting
- Nov 15th: brainstorm meeting about 2021.12 pilot
- Nov 16th: lightning talk on EESSI at HPC System Testing BoF session at SC'21
- Nov 24th: brainstorm meeting about software deployment and testing
- Nov 26th: brainstorm meeting about GitHub App for software-layer repo

[Bob, Kenneth, Caspar]

EESSI-related meetings

Nov 9th: monthly CernVM-FS coordination meeting

- Version 2.9 coming soon (already available in testing repository; publicly released on Nov 25th)
- CernVM workshop 2022 (Sept 12-14, Amsterdam) will be in-person with remote possibilities
- Support for different EL8 distributions
  - Packages are expected to work on all derivative distros (Rocky Linux, AlmaLinux, CentOS Stream)
- Discussion about the need/wish for a higher frequency of CernVM-FS releases
- Possibly more Arm packages in the future, built on an Apple M1
  - Kenneth gave some pointers to using fosshost.org, AWS (via EESSI credits?), OSU OSL
- CernVM-FS package builds and tests are done with Jenkins + Openstack VMs
  - Full integration test run takes >24 hours, smaller one ~30 minutes
  - Building and testing with QEMU is too slow

[Bob, Kenneth]

EESSI-related meetings

Nov 15th: brainstorm meeting about 2021.12 EESSI pilot repository

- https://github.com/EESSI/meetings/wiki/Brainstorm-meeting-2020.12-pilot-Nov-15th-2021
- GitHub Project: https://github.com/orgs/EESSI/projects/7
- Most important actions:
  - Add CUDA support
  - Install more software, and use more recent toolchains
  - Facilitate building your own software/modules on top of it
  - Add more automation for the software build pipeline
- Can we make this a Long-Term Supported (LTS) release?
  - Note: **still pilot!**

[Bob]

EESSI-related meetings

Nov 16th: lightning talk on EESSI at HPC System Testing Birds-of-a-Feather @ SC'21

- Invited via ReFrame developers
- By Caspar (SURF)
- See also https://olcf.github.io/hpc-system-test-wg/events/sc21bof
- **Recording available at https://www.youtube.com/watch?v=ALpZhUDHjls**

[Caspar]

EESSI-related meetings

Nov 24th: brainstorm meeting about software deployment and testing (1/2)

- CI tests software layer
  - Only run test for software that was updated
  - On build node: smoke tests (e.g. rerun `eb --sanity-check-only`), small application runs (single node, < 10 min), multiple OS-es
  - On test cluster: multi node application runs (multinode, < 10 min)
- CI tests compat layer
  - One set of tests, run all at every change
- Monitoring
  - Single & multi node application runs, multiple test cases per software

[Caspar]

EESSI-related meetings

Nov 24th: brainstorm meeting about software deployment and testing (2/2)

- Test selection
  - Tags: tests will be tagged depending on where they will be run
    (example: ‘build’ for build node tests, ‘daily/weekly/monthly’ for monitoring tests)
  - Correct subset of tests can then be run using ‘reframe -t <tags>’
  - Hardware: tests should implement mechanisms to be skipped,
    for example if GPU test is run on non-GPU node
- CI triggers & infrastructure
  - Github App, listens to “events”
  - Run on VM clusters (Azure, AWS, CSCS?)

https://github.com/EESSI/meetings/wiki/Brainstorm-meeting-software-deployment-Nov-24th-2021

[Caspar]

EESSI-related meetings

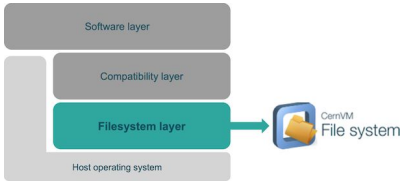
Nov 26th: brainstorm meeting about GitHub App for software-layer repo

https://github.com/EESSI/meetings/wiki/Brainstorm-meeting-github-app-software-layer-Nov-26th-2021

- Discuss implementation of a “bot” to allow contributions to EESSI repo
  - In preparation of EESSI hackathon…
  - Which GitHub pull request events should trigger an action?
  - How should implementation be structured?
  - How can work be split up across multiple people working in parallel?
  - TODO (Bob, Kenneth):
    - Set up EESSI/eessi-bot repository (better name?)
    - Flesh out structure + set up basic CI
    - Collect event data that can be replayed to test implementation

[Bob, Kenneth]

Progress update: filesystem layer

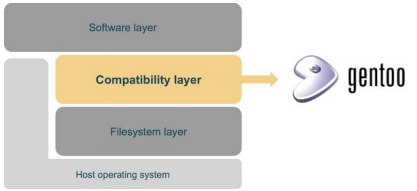


- Directory structure tweaked a bit, main changes:
  - Add a `versions` directory to the root, move pilot version subdirectories in there
  - Use variant symlinks `modules` and `software` in the root, which can be set by a client
  - More details: https://github.com/EESSI/filesystem-layer/issues/32
  - This will be used/changed for the next EESSI pilot (2021.12)

```
.
└── cvmfs/
    └── pilot.eessi-hpc.org/
        ├── README
        ├── modules -> 'site_production/$(EESSI_SITE_ARCH)/modules'
        ├── software -> 'site_production/$(EESSI_SITE_ARCH)/software'
        ├── latest -> versions/2021.12
        ├── site_production -> 'versions/$(EESSI_SITE_PRODUCTION_VERSION)'
        ├── 2021.06     # we can't (re)move this
        └── versions/
            ├── 2021.06 -> ../2021.06
            └── 2021.12
```

[Bob]

Progress update: compatibility layer (1/2)



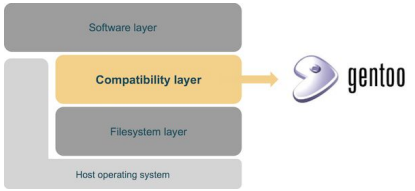
- No security updates required for 2021.06 compat layer
  (as reported by Gentoo's `glsa-check` tool)

- Lots of progress in preparation of EESSI pilot 2021.12
  - 11 merged PRs in `compatibility-layer`
  - 20 merged PRs in `gentoo-overlay`

- Playbook to build compatibility layer has been improved *significantly*
  - More control over Prefix bootstrap using fixed commit of Gentoo repo (instead of Portage snapshot)
  - Stick to glibc 2.33 + GCC 9.4.0 as system compiler (to avoid trouble with CUDA, etc.)
  - Ensure that `ld.gold` (binutils) is installed (expected by EasyBuild, avoid using host `ld.gold`)

[Bob]

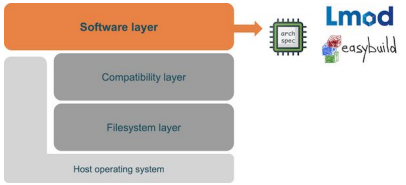
Progress update: compatibility layer (2/2)



- Changes for packages included in compat layer (cfr. `EESSI/gentoo-overlay` repository)
  - Updated package set 2021.12 (see eessi_sets.yml)
  - Added `rich` (Python library), `bash-completion`
  - Using `rdma-core` from Gentoo (rather than our own `ebuild`)
  - Removed `pkgconfig` (not really needed, removed in Gentoo in favor of `pkgconf`)
- Enabled running of Portage tests for Lmod, ReFrame, archspec
- Add Repoman to GitHub Actions workflow (used as linter for our own `ebuild` files)
- ReFrame tests updated to take into account changed directory layout (`versions` subdirectory)

[Bob]

Progress update: software layer



- Init script will need some love to be compatible with archspec v0.1.3 (used in 2021.12 compat layer)
  - See https://github.com/EESSI/software-layer/issues/142
- Update build script for EasyBuild v4.5.0 (WIP, being tested)
- Starting extending set of software included in EESSI…
  - More recent toolchains, more applications, etc.
  - Perhaps only once automated contribution workflow is in place? (cfr. hackathon); see also https://github.com/EESSI/software-layer/issues/149
  - More organised module tree (don't mix modules from different "generations"/toolchains)?

[Kenneth]

EESSI pilot repository

NOT FOR PRODUCTION USE!

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

- 2021.06: considered “final”
  - No updates or changes planned, will stay in place for now (to be archived into container?)
  - Except for security updates in compat layer, if needed…

- Current status for 2021.12:
  - New path: `/cvmfs/pilot.eessi-hpc.org/` **`versions`**`/2021.12`
  - Compatibility layer:
    - Ansible playbook works on `aarch64`/`ppc64le`/`x86_64`
    - List of installed packages is pretty much identical \o/
    - Will be ingested into EESSI pilot repository soon…
  - Software layer: making sure that installing software from 2021.06 still works (WIP)

[Kenneth, Bob]

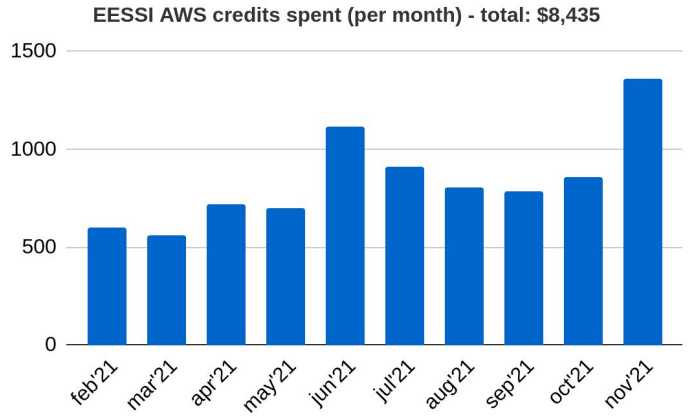
Progress update: testing

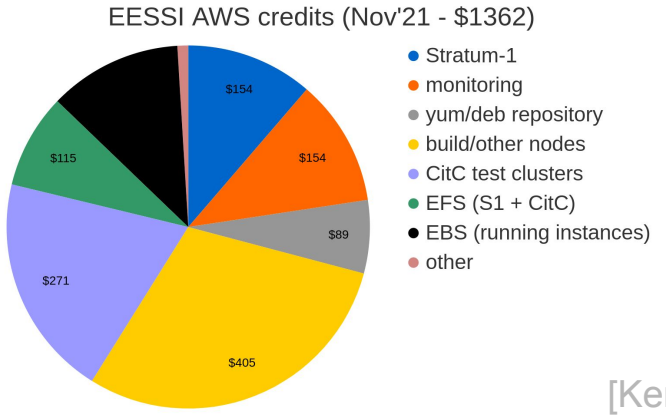
- Use ‘library test’ format of ReFrame
  - Test-specific things in library test (command line, input files, sanity pattern, etc)
    - Leverage CSCS “library of tests”
    - See https://github.com/eth-cscs/reframe/tree/v3.9.2/hpctestlib
  - “System-specific” in test implementation (#tasks, modules to load, etc)
    - EESSI implementation (still generic, e.g. ‘one task per core’)
- EESSI hackathon
  - Create more library tests (Victor, ...)
  - Make more EESSI implementations (Caspar, …)
  - Create Github App (Kenneth, Bob, …)
  - Set up resources to run CI

[Caspar]

Usage of sponsored AWS credits

- Sponsored credits ($25,000) are being put to good use!
- **Ask in #aws-resources Slack channel to get access!**
- In November ‘21: ~$1362 worth of credits spent on Stratum-1 server, monitoring node, repository node, lots of build nodes, aarch64 VM for CernVM-FS;
- ~$8,435 worth of credits spent in total
- Credits will expire on Jan 31, 2022! (new sponsored credits will be requested)

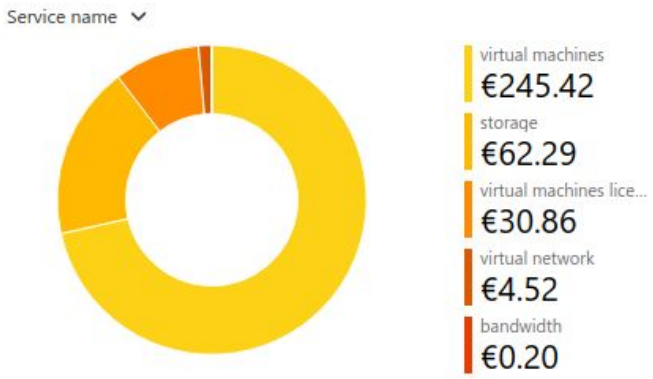


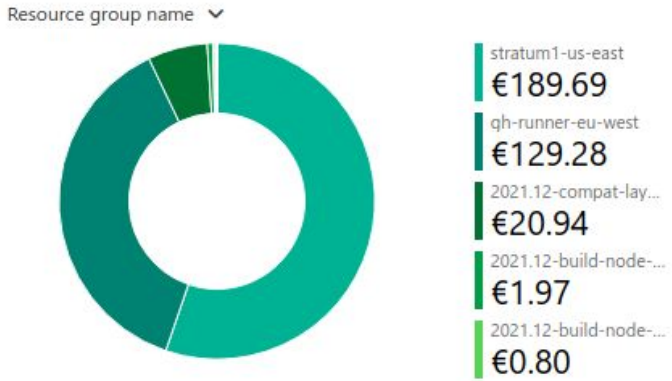


[Kenneth, Bob]

Azure sponsorship

- **Ask in #azure-resources Slack channel to get access!**
- In November ‘21: ~€ 343 worth of credits spent
- On Stratum 1, (stopped) GitHub Runners, and compatibility layer build nodes
- ~€1,074 worth of credits spent in total





[Kenneth, Bob]

Update on EESSI journal paper

"**EESSI: A Cross-Platform Ready-To-Use Optimized Scientific Software Stack**"

- Submitted to special issue *"New Trends in HPC: Software Systems and Applications"* in *"Software: Practice and Experience"* journal (Wiley)
- Authors: Bob, Kenneth, Victor, Alan, Caspar, Thomas (proof-reading by Adam)
- Received reviews in mid August (overall fairly positive)
- Valuable comments and suggestions for improvements
- **Revised version has been submitted! (Nov 17th)** 🎉
- **Patiently awaiting decision w.r.t. publication…**
- **If you would like to read the draft paper, let us know!**

[Thomas]

EESSI hackathon(s)

- Focused effort (part time) on making good progress on various tasks
  - Self-organising small groups (2-5 people) working on a (single) specific task
  - With support from “experienced” EESSI contributors
  - Week of **13-17 December 2021** (+ 17-21 January 2022)
  - 3 Zoom sessions each week: kickoff on Monday, sync on Wednesday, show & tell on Friday
  - More information: https://github.com/EESSI/meetings/wiki/EESSI-hackathon-Dec'21
- Make some good use of our remaining AWS (and Azure) credits
  - Current AWS credits expire end of Jan’22, still ~$15k left!
- Dedicated channel in EESSI Slack: #hackathon
- **If you plan to participate, please register and select tasks you are interested in!**
  - https://doodle.com/poll/xha7h6pawwuk5xc2

[Kenneth, Bob]

EESSI hackathon: overview of tasks (1/2)

(see also https://hackmd.io/L763hQgRS5Shn04rAbVmWA)

1. Nice overview of EESSI software stack (documentation)
2. Installing software on top of EESSI
3. **Workflow to propose additions to EESSI software stack**
4. Expand EESSI software stack
5. GPU support
6. EESSI test suite (ReFrame)
7. Monitoring
8. Setting up a (private) Stratum-1

[Kenneth, Bob]

EESSI hackathon: overview of tasks (2/2)

(see also https://hackmd.io/L763hQgRS5Shn04rAbVmWA)

9. Risk analysis (continued)
10. Performance evaluation
11. Distribute building of software across multiple nodes
12. EasyBuild issues & PRs related to EESSI
13. Document resources available to EESSI
14. More user-friendly documentation on accessing EESSI
15. Set up autoscaling self-hosted GitHub Runners
16. Export a version of the EESSI stack to a tarball and/or container image

[Kenneth, Bob]

Upcoming events

- Computing Insight UK (CIUK) 2021:
  - https://www.scd.stfc.ac.uk/Pages/CIUK2021.aspx
  - December 9-10, 2021, Manchester (and online)
  - Jörg's submission on EESSI was not accepted, but it is kept as a reserve…

[Jörg]