EESSI meeting

Sept 2nd 2021

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

Agenda

1. Quick introduction by new people

2. EESSI-related meetings in last month [Kenneth, Bob]

3. Progress update per EESSI layer [Kenneth, Bob]

4. 2021.06 version of pilot repository [Kenneth]

5. AWS/Azure sponsorship update [Kenneth, Bob, Henk-Jan]

6. Update on EESSI journal paper + S4 NeIC project proposal [Thomas]

7. 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

- July 13th: monthly CVMFS coordination meeting [Bob]

  - Workaround for the in-place update corruption bug: publish file with direct I/O flag
    - Disables kernel cache for that file; performance hit
    - Clients and server both need the 2.9 release, older version will ignore it
    - Alternative (with better performance): symlink

  - Google Summer of Code Project: client preload functionality
    - Improve startup performance of apps (e.g. Python)
    - Manually configure bundles on the Stratum 0
    - Bundle gets downloaded on first access of a file from that bundle

  - Discussion about using S3 for the Stratum 0 (or 1)
    - May be attractive if you have this at your site anyway
    - "Makes sense to keep a good relationship with your S3 admins"
    - Makes the Stratum 0 sort of stateless: can be redeployed in minutes

[Kenneth, Bob]

EESSI-related meetings

- Aug 27th: meeting with Martin Brandt (SURF) [Bob, Kenneth, Caspar]

  - Discuss practical setup of leveraging sponsored Azure credits for EESSI

  - Separate subscription for EESSI

  - Dedicated account for API access (for Terraform, etc.)

  - Martin has lots of experience with Terraform/Ansible/Azure/AWS

  - Monthly follow-up meetings with EESSI + SURF + Microsoft Azure

  - Martin will hopefully join next EESSI meeting to present SURF Research Cloud...

[Kenneth, Bob]

Progress update: filesystem layer
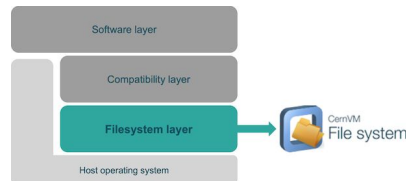
- Automated ingestion of tarballs (WIP: PR #90) almost ready for some real testing...

  - Script runs on the Stratum 0

  - Fetches tarballs (and automatically generated metadata files) from the AWS S3 bucket

  - Add metadata file to special GitHub repository, open PR for ingestion approval

  - After approval, the tarball gets ingested to the CernVM-FS repository

A new tarball has been staged.
Please review the contents of this tarball carefully.
Merging this PR will lead to automatic ingestion of the tarball.

▼ Details
Overview of the contents of the tarball:

```
software/linux/x86_64/amd/zen2/.lmod/cache
software/linux/x86_64/amd/zen2/.lmod/cache/spiderT.lua
software/linux/x86_64/amd/zen2/.lmod/cache/spiderT.luac_5.1
software/linux/x86_64/amd/zen2/.lmod/cache/spiderT.old.lua
software/linux/x86_64/amd/zen2/.lmod/cache/spiderT.old.luac_5.1
software/linux/x86_64/amd/zen2/.lmod/cache/timestamp
software/linux/x86_64/amd/zen2/modules/all/Arrow/0.17.1-foss-2020a-Python-3.8.2.lua
software/linux/x86_64/amd/zen2/modules/all/Bazel/3.6.0-GCCcore-9.3.0.lua
software/linux/x86_64/amd/zen2/modules/all/Boost/1.72.0-gompi-2020a.lua
software/linux/x86_64/amd/zen2/modules/all/CGAL/4.14.3-gompi-2020a-Python-3.8.2.lua
```

[Bob]

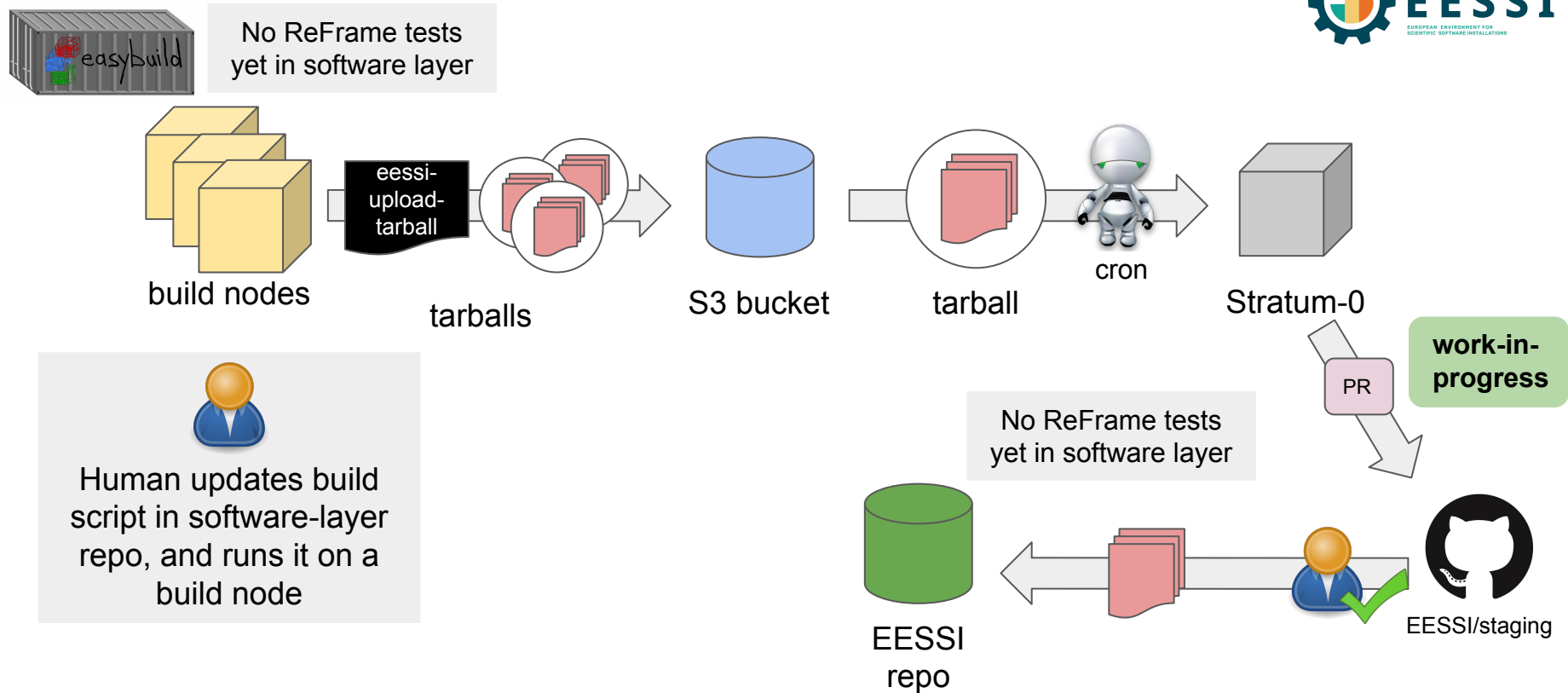# Progress on automating software deployment

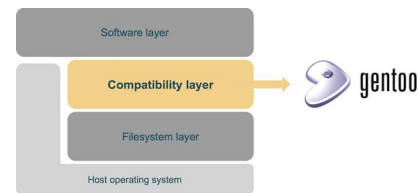Goal: automated, with human oversight
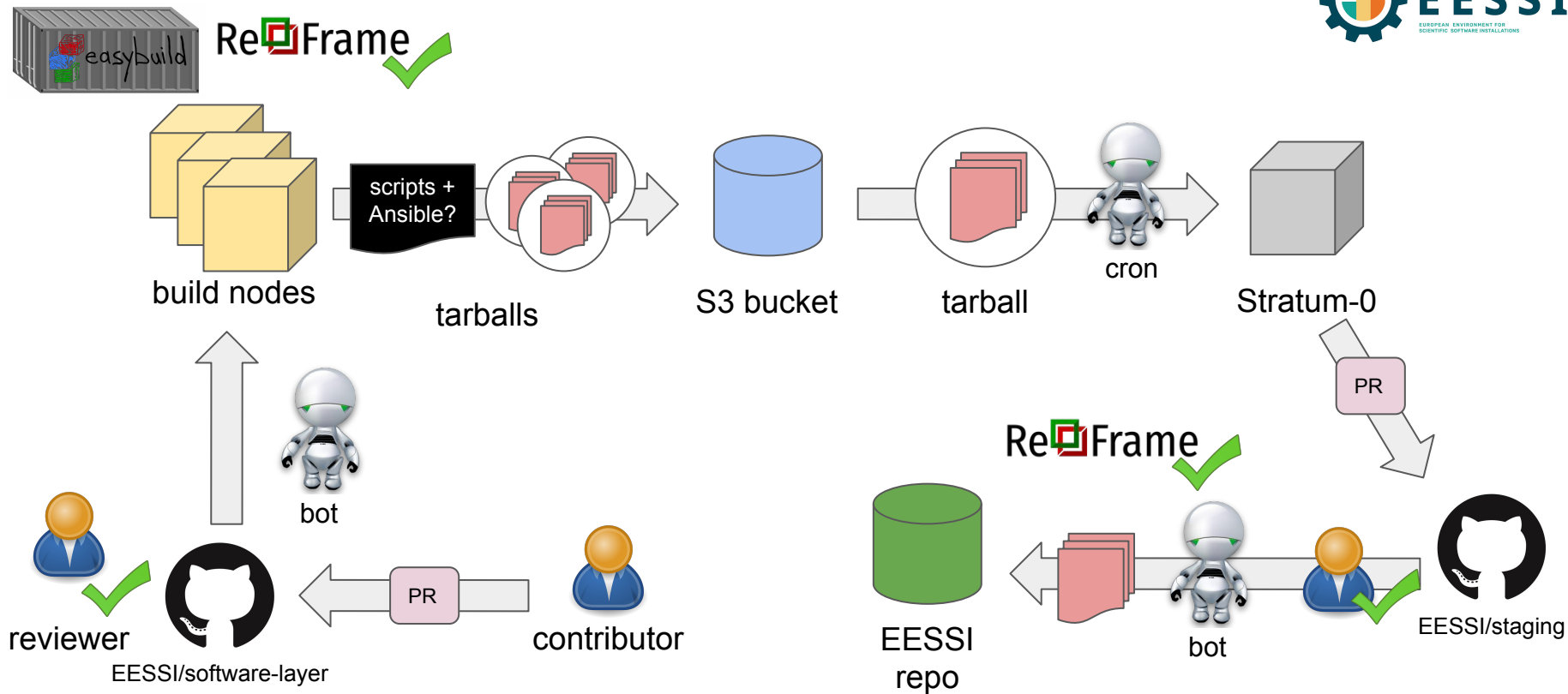
Progress on automating software deployment


Current status (or close to it)

Progress update: compatibility layer

- ReFrame tests for compat layer are now run daily,

  see https://github.com/EESSI/compatibility-layer/actions

- `patchelf` included (required to ensure that binary installations honor compat layer)

- No security updates required (automated check didn't raise any issues)

- PR for `rdma-core` (with Prefix support) has been merged into the Gentoo repository

[Bob]

Progress update: software layer
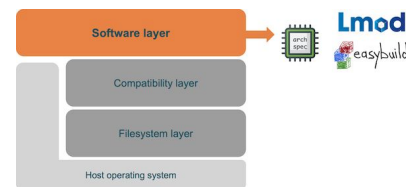
- More modular organisation of init scripts (used by build script + `init/bash`)
  - Script to set up minimal EESSI environment (`init/minimal_eessi_env`)
  - Script to configure EasyBuild (`configure_easybuild`)

- Updates to build script for EasyBuild v4.4.1 + fix ReFrame installation (+ cleanup)

- Java problem fixed: using `patchelf` to patch installation to pick up glibc from compat layer
  - EasyBuild PR: https://github.com/easybuilders/easybuild-easyblocks/pull/2557

- Injection of path into RPATH section to override OpenMPI library via variant symlink at `/cvmfs/pilot.eessi-hpc.org/host_injections` (see PR #116)

- Script to easily start build node environment (using Singularity) - see PR #129

    `./start_build_node_env.sh /tmp/$USER/EESSI`

  Replaces most of the documentation at eessi.github.io/docs/software_layer/build_nodes...

[Kenneth]

EESSI pilot repository

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

2021.06 version of pilot software stack

- Target CPUs:
  - `{aarch64,x86_64,ppc64le}/generic`
  - `intel/{haswell,skylake_avx512},amd/zen2,aarch64/graviton2, ppc64le/power9le`

- Software: Bioconductor, GROMACS, OpenFOAM, R, TensorFlow, Spark, IPython, Horovod, QuantumESPRESSO, ReFrame, … (some stuff excluded for `ppc64le`…)

Current status:

- Compatibility layer in place for `x86_64` + `aarch64` + `ppc64le`
- Software layer built for majority of targets (`aarch64` + `x86_64` done, `ppc64le` is work-in-progress)
- Binaries linked to OpenMPI will pick up on libraries in

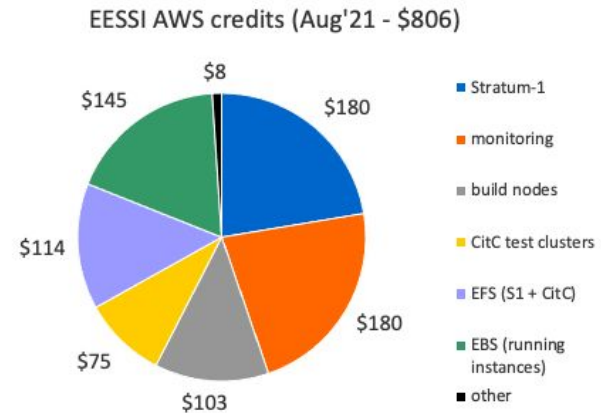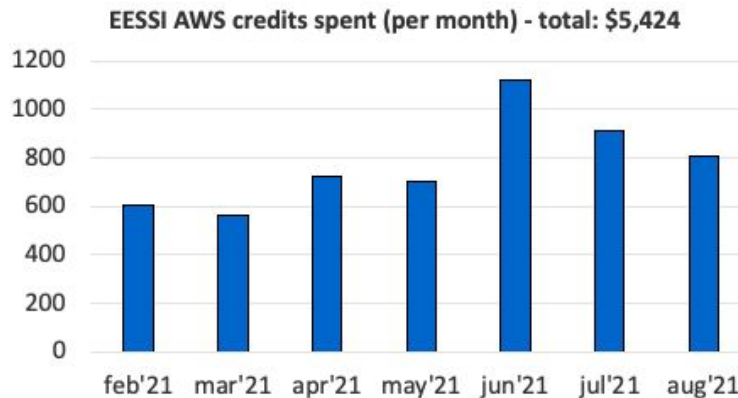 `host_injections/rpath_overrides/OpenMPI/system/lib`

To do:
- Ingest prepared tarballs (opportunity to test Bob's semi-automated ingestion procedure!)
- Init scripts (trivial)
- GPU installations: on hold...

[Kenneth, Alan, Bob]

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 August '21: ~$806 worth of credits spent

- on Stratum-1, monitoring node, build nodes, ...

- ~$5,424 worth of credits spent in total

EESSI AWS credits spent (per month) - total: $5,424

EESSI AWS credits (Aug'21 - $806)

$8
$145
$180
$114
$180
$75
$103

- Stratum-1
- monitoring
- build nodes
- CitC test clusters
- EFS (S1 + CitC)
- EBS (running instances)
- other

[Kenneth, Henk-Jan]

Azure sponsorship

- **Sponsored credits ($40,000) available for EESSI project \o/**

- Provided through SURF Research Cloud (separate subscription)

- API access available (for Terraform, etc.)

- Initial setup done: (admin) accounts for Bob, Kenneth, Terje

- Additional accounts can be created EESSI admins (self-managed)

- Goals: build nodes, (large scale) testing, demos, Stratum 1, etc.

- Monthly follow-up meetings planned with Azure + SURF
  - every 3rd Friday at 3pm CEST (starting Sept'21)

[Kenneth, Bob]

Update on EESSI journal paper

- Title: "**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

- Discussed necessary changes and distributed work

- Revised version due mid November

- No guarantee of final approval

[Thomas]

Coming soon...

- Feedback on S4 project NeIC proposal

  - "Scientific Software Stacks as a Service (S4)", ~100PMs over 2 years

  - Partners: UoIceland, UoEstonia, Sigma2 (NO), SNIC/Umeå (SE),
    CERN, CSCS, UGent, RUG

  - 11 project applications were sent,
    https://neic.no/news/2021/03/18/open-call-applications

  - NeIC board meets Sept 22nd, notification expected shortly after...

[Thomas]