CVMFS users workshop report + Plans for the future

Nov 13 2024, OTF
Valentin Völkl for the CVMFS Development Team
CERN

CernVM Users Workshop   indico.cern.ch/e/cvm24

- 16th - 18th September at CERN
  - Marked handover of project leadership from Jakob to me

- **60 registrations** in total: 30 in person 30 remote
- Affiliated with **25 institutes**

-

- Monday afternoon: **CERN team presentation**
- Tuesday morning: **External / Keynote speakers**
- Tuesday afternoon: **Experiment/site reports**
- Wednesday morning: **Varnish Hands-on**

- [30 Contributions](#)

CernVM Workshop 2024
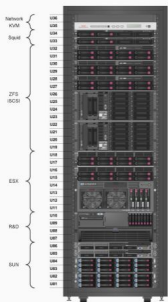
16-18 September 2024
CERN, Geneva

As CVMFS turns 15, the workshop aims to bring together users and developers to discuss the current status of the CernVM ecosystem and the future directions, with a fresh look onto the landscape of cloud technology and software delivery.

This is an in-person event with the possibility of joining remotely. If you plan to attend, either in-person or remotely, please register!

https://indico.cern.ch/e/cvm24

Selected Highlights

- Initial technology choices leading to CVMFS 2.0:
  Fuse, C++, HTTP CDN, SQLite file catalogs, content-addressed storage
- CernVM infrastructure (including CVMFS storage, release managers etc.) operated from building 157
  - Fully virtualized with VMware ESX
  - Storage using Solaris/ZFS: our initial solution for snapshotting & replication
- Presented at CHEP 2010 in Taipei
  (15 million files under management)
- Growing interest in using CernVM-FS on the Grid outside the VM
  (virtualization came back later a few years later with OpenStack, Docker, k8s)
  - to address shortcomings of AFS, NFS, Grid installation jobs

Distributed File Systems

CernVM Software Appliance

COD

Carnegie Mellon

http://www.coda.cs.cmu.edu/

Chirp/Parrot/Grow-FS

Chirp! HDFS! XRootD! iRODS! FTP!

https://ccl.cse.nd.edu/

OpenAFS

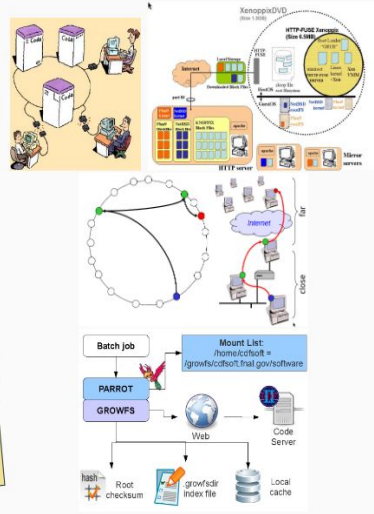https://www.openafs.org/

IgorFS

http://doi.ieeecomputersociety.org/10.1109/P2P.2008.19

- Technological look back on 15 years of CVMFS by Jakob Blomer
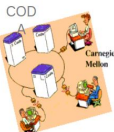
- Part of the CernVM R&D project on virtualization
  ▸ Predrag's 2018 Talk: CernVM 10 years after
- Decouple the experiment software from the virtual machine image using a **global network file system**
- Looked into several existing options
  - Coda: AFS with offline mode
  - HTTP-Fuse: on-demand bootable Linux image
  - Igor-FS: file system w...
  - GROWFS: CernVM-F... the Parrot system call... toolkit

CVMFS v1 was a GROW-FS Fuse frontend written by Leandro Franco
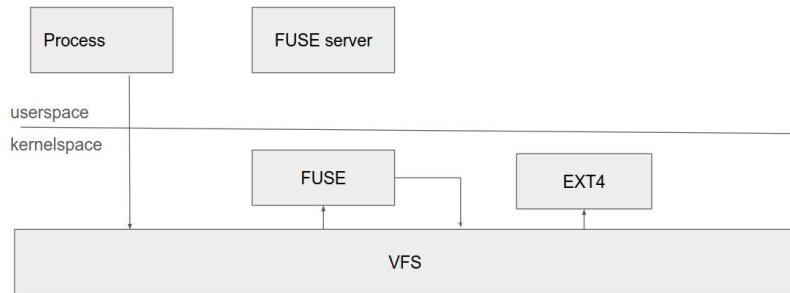
4

FUSE developer talks

- Very fruitful discussion on technological foundation of CVMFS
- We'll profit from some ongoing developments - passthrough and io_uring
- Could deposit our particular needs and wishes

Miklos Szeredi
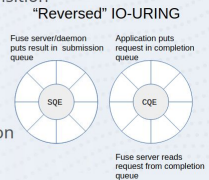
Bernd Schubert

EESSI: European HPC software distribution using CVMFS
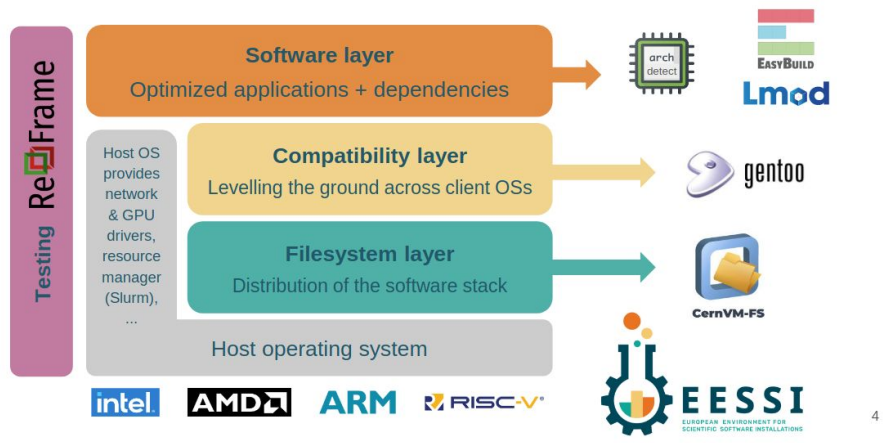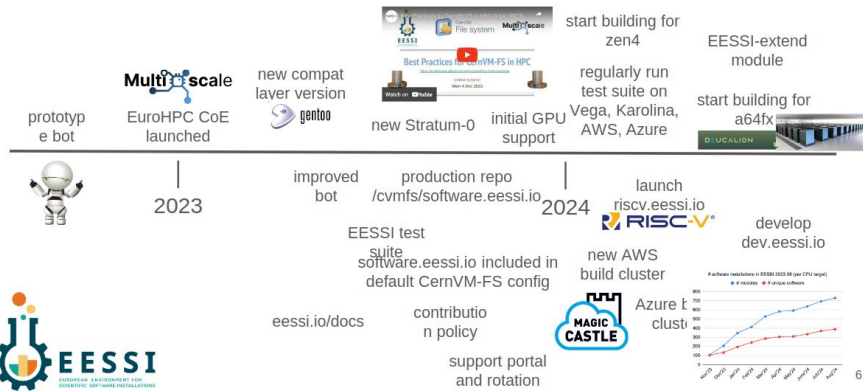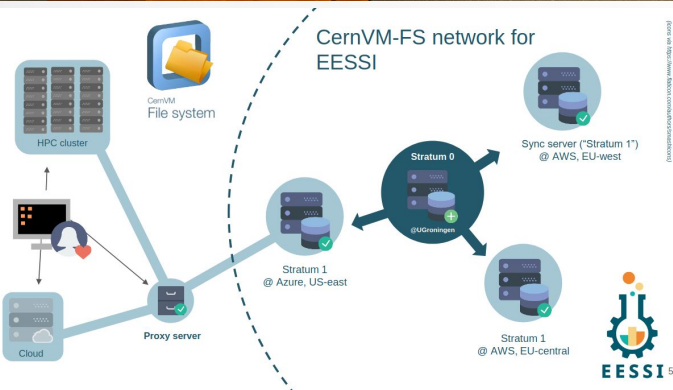
CVMFS Snapshotter Benchmarking: Workshop + CHEP

New usecases of CVMFS at CMS
Andrea Valenzuela Ramirez

CernVM-FS Use Cases at CMS

CMS Offline & Computing deploys to CernVM-FS under different use cases:
- Distribution of experiment **production software** (CMSSW).
- Distribution of **Integration Builds** (IBs).
- **Continuous Integration** (CI) purposes.

Repository Name	Size	Garbage Collection	Parallel Setup	Publishing (ops/day)	Year
/cvmfs/cms.cern.ch	23 TB	No	No	$\sim$ 5-30	2009
/cvmfs/cms-ib.cern.ch	3.77 TB	Yes (weekly)	Yes	$\sim$ 40	2016
/cvmfs/cms-ci.cern.ch	883 GB	Yes (weekly)	No	$\sim$ 1-40	2020

Table: CMS main repositories and their characteristics in terms of size, garbage collection frequency, publication setup, number of commits and year of creation.

- Distribution of **CMSSW environment images** in `unpacked.cern.ch`.

▶ CernVM-FS Workshop 2022

Usage of HPC Resources

- Access to AMD GPUs at LUMI (Finland) through the project *Exploring the Use of AMD GPUs for High-Performance Computing in the CMS Reconstruction.*

- Access to cvmfs using `singcvmfs exec`.

- CMS container images deployed to the LUMI user node from Dockerhub.

- Use `SINGCVMFS_REPOSITORIES` to indicate which repositories to load.
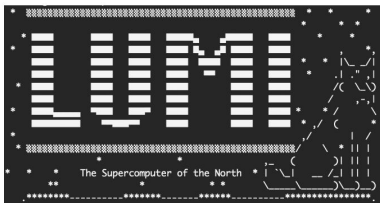
```
export SINGCVMFS_REPOSITORIES=cms.cern.ch, cms-ib.cern.ch, cms-ci.cern.ch,
grid.cern.ch, unpacked.cern.ch, patatrack.cern.ch
```

Distribution of Gridpacks

- CMS high precision analyses require very precise Monte Carlo generators. For example, to guarantee Next to Leading Order (NLO) calculations.
- MadGraph generates the outcomes of particle interactions, which can be latter used to speed up computations.
- Concretely, MadGraph produces the so-called Gridpacks.
- Gridpacks are "pre-computed diagrams" used speed-up Monte Carlo generation.
- Distributed in tarballs, they are uncompressed for every generator job on local disk. **Many sites do not support such operation**.
- The proposed solution was serving already-untarred Gridpacks via CernVM-FS.

It is a new use-case of distribution of lookup files at CMS.

- At the moment, content is synchronized using `rsync` from `/eos` to `/cvmfs`, but it seems a nice use-case for `cvmfs_server ingest` utility.

8

Varnish Hands-On

- **Varnish** was represented with several developers at workshop
  - Possible replacement of SQUID as proxy server technology
  - In first order, addresses concerns about maintainability of SQUID
  - CVMFS usecase seems very well covered, FRONTIER to be investigated

Outlook

Takeaways

- Many good suggestions and ideas that will form the plan of work 2025
- Some clear potential for improvement in Stratum-1 operations and efficiencies
  - Incremental, interruptible garbage collection will be high priority item
- As well as still painpoints in usage HPC sites
  - Although EESSI is helping our community a lot
- Evolution of hardware towards many-core machines
  - poses challenges for CVMFS - usage of memory, fds ...
- Some interesting technology investigations
  - Varnish as proxy
  - New libfuse developments
- Community interest in data-distribution-over-CVMFS
  - Only makes sense for low-throughput,non-performance critical applications
  - Will need some dedicated RnD, prioritize software distribution

What is the maximum CPU core count on machines at your sites?

ATLAS grid site survey
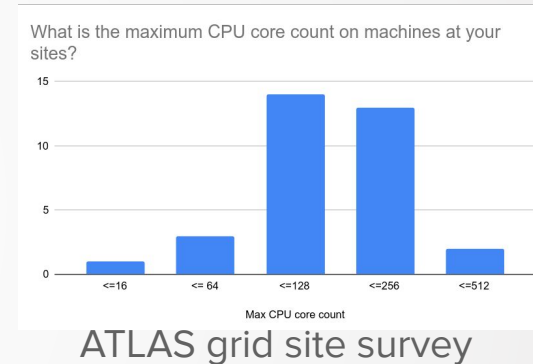
Container tools

- Vital for smooth integration of CVMFS for new users/sites
  - CVMFS snapshooter in particular has potential to be very efficient while having little friction
- Often still need a bit of polish
  - Should be much improved in 2.12
  - Hope for a closer integration of snapshotter in CERN-IT/gitlab/kubernetes
- Containerisation+kubernetes could address issue of scaling publishing infrastructure - **"Elastic publishing"**
- unpacked.cern.ch needs some measures to keep growth sustainable
  - Started cleanup campaign, better monitoring and possible auto-removal of non-critical images after 1 year

Number of files in /cvmfs over time

lhcb.cern.ch
cms.cern.ch
atlas.cern.ch
alice.cern.ch
unpacked.cern.ch
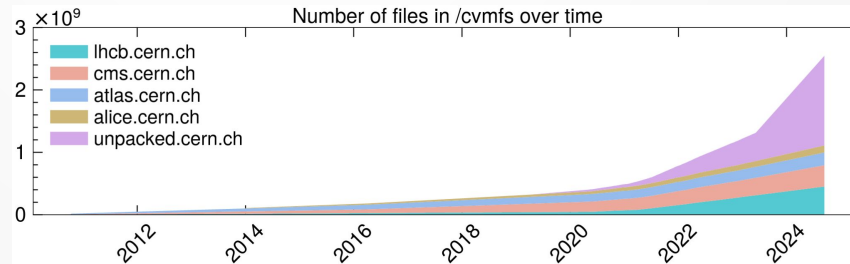
Plans to be discussed

- Stop Support of Centos 7
  - Will still build last round of packages for upcoming 2.12 release
- Possible Deprecation of x509 authenticated CVMFS repositories
  - LIGO/VIRGO is the only user we know of, currently phasing out their CVMFS authenticated repositories, removing them altogether in 2025
  - Will help slim down codebase and free up developers
- Importance of Hotpatching capabilities of CVMFS?
  - Update currently possible without remount or stop of jobs

- Clients on 2.11.4 should updated ASAP
- Issue call for action to improve usage of cvmfs-testing repository for staged rollouts

Conclusion

- Very fruitful event, rich technological program
- Next workshop planned in 2026, after CHEP
- Outlook on next program of work:

1. Carry-over tasks: ZSTD compression, gateway performance and functionality improvements
2. Many-core performance engineering
3. Technology exploration of VARNISH as a proxy server
4. Integration in HPC sites
5. True offline mode for laptop usecase

CernVM Workshop 2024

16-18 September 2024 CERN, Geneva

As CVMFS turns 15, the workshop aims to bring together users and developers to discuss the current status of the CernVM ecosystem and the future directions, with a fresh look onto the landscape of cloud technology and software delivery.

This is an in-person event with the possibility of joining remotely. If you plan to attend, either in-person or remotely, please register!

https://indico.cern.ch/e/cvm24

Backup

History

- CernVM 2022 Workshop at Nikhef [event page](#)
- CernVM 2021 Virtual Workshop [event page](#)
- CernVM 2019 Workshop at CERN [event page](#)
- CernVM 2018 Workshop at CERN [event page](#)
- CernVM 2016 Workshop at RAL [event page](#)
- CernVM 2015 Workshop at CERN [event page](#)
- 2010 Workshop on adapting applications and computing services to multi-core and virtualization [event page](#)
- 2009 Workshop on adapting applications and computing services to multi-core and virtualization [event page](#)
- 2008 Workshop on virtualization and multi-core technologies for LHC [event page](#)

"Hi Valentin,

Thank you for the invitation, I have fond memories of my time at [the last workshop at] CERN.... "

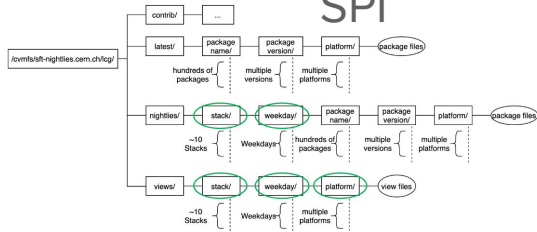New overlay FS features in cvmfs_server; FUSE-T on macOS

Yuriy Belikov
CernVM FS Team,
Sep 16th 2024

Zstd:
A new compression algorithm for CVMFS

Laura Promberger
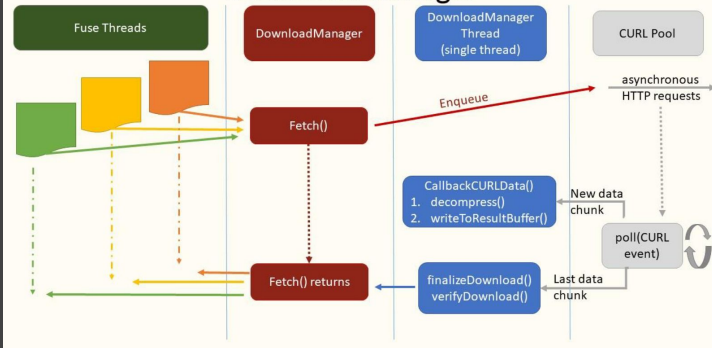CernVM Workshop 2024

Implementation

Tim Ehmann, SPI

- Parallelisation for view creation
- Parallelisation for installation in nightlies between different stacks (dev3, dev4, dev3cuda, dev4cuda, …)

16/09/24   Tim Ehmann   10

DownloadManager

EOS Open Storage
eosxd & **eos**cfsd FUSE filesystems

CernVM Workshop 2024
16–18 Sept 2024
CERN
Europe/Zurich timezone

Elvin Alin Sindrilaru
& Andreas-Joachim Peters
for the EOS Project - CERN IT - Storage Group

EOS Fuse Outlook

• most relevant performance limits in **eos**xd are addressable in MGM + FST implementation - FUSE bottlenecks play only minor role in our environment

• **eos**xd would benefit from XATTR caching in kernel - which does not exist

• An architecture where **meta-data** caches can be directly shared between user-space and kernel would be desirable and simplify the user-space implementation - had a look at extFUSE

/eos - EOS via FUSE at CERN

read all protocols > 1 TB/s

write all protocols > 300 GB/s

read FUSE traffic peak 292 GB/s
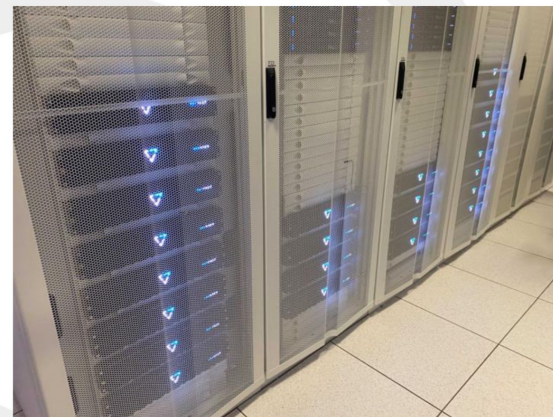
write FUSE traffic peak 6 GB/s

Large scale data processing with CVMFS

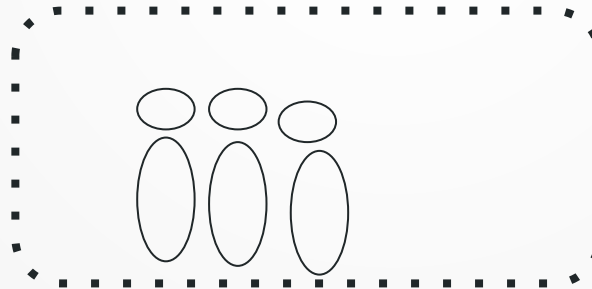CVMFS Workshop, September 2024

Data Archive

- Realtime-updated repository time-series market data
- Contains raw data and derivative products for end-users

Conclusion

- **Workshop was very rewarding**
- Many fruitful discussions, and encounters with users we don't interact with that often
  - Learned about interesting new possible applications of CVMFS in bioinformatics and fusion research
- Nice social program for external participants
  - Workshop dinner at Bains des Paquis
  - Guided Visit to AD (Thanks to Jacopo Fanini and Siara Fabbri!)
  - CERN 70 festivities
- Next workshop planned in 2 years

Group photo

… next time!

CernVM Workshop 2024

16-18 September 2024
CERN, Geneva

As CVMFS turns 15, the workshop aims to bring together users and developers to discuss the current status of the CernVM ecosystem and the future directions, with a fresh look onto the landscape of cloud technology and software delivery.

This is an in-person event with the possibility of joining remotely. If you plan to attend, either in-person or remotely, please register!

https://indico.cern.ch/e/cvm24