CernVM Workshop Summary

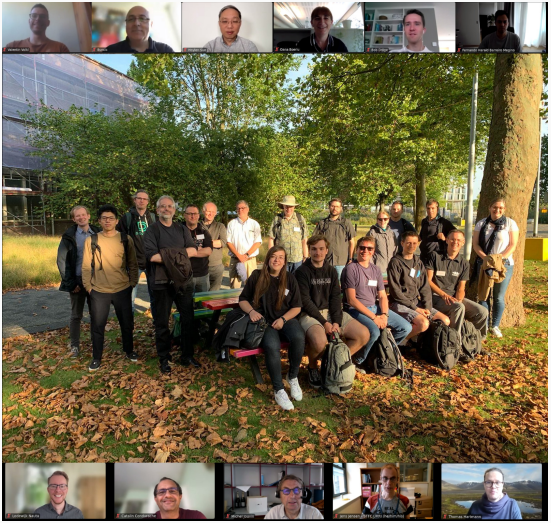
Laura Promberger (CERN), Jakob Blomer (CERN)

GDB, 12 October 2022

Overview

- 2-day event in Nikhef, Amsterdam: 12-13 September 2022
  - Indico Page
- Organizers
  - Dennis Van Dok, Nikhef
  - Jakob Blomer, CERN
  - Mary Hester, Nikhef
- 22 talks from 21 speakers
  - Productive *Open Discussion* Session on Monday morning
- Total registrations: 72
  - 25 people on-site
- On Monday evening: Walking tour and dinner

Group Picture



Grouping of the Talks

- CVMFS-Related News and Status
- CVMFS Users - Scientific Community
- CVMFS Infrastructure Providers - Scientific Community
- CVMFS Users - Industry and HPC

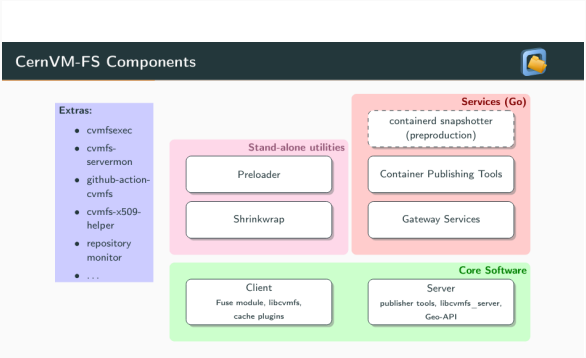
CVMFS-Related News and Status – Overview

	Talk	Speaker
1)	CernVM-FS: Status and Plans	*Jakob Blomer*
2)	CernVM-FS publisher on Kubernetes	*Andrea Valenzuela Ramirez*
3)	CernVM 5: A fully containerized CernVM	*Jakob Karl Eberhardt*
4)	Harbor Registry at CERN: Status and Enhancements	*Ricardo Rocha*
5)	CSI Driver for CernVM-FS	*Robert Vasek*

CVMFS-Related News and Status I

1) CernVM-FS: Status and Plans – *Jakob Blomer*

- Version 2.10 will be soon released (currently pre-production testing)
- Focus on scalability and client performance (e.g. work on better kernel cache utilization, network optimization such as proxy sharding)
- Growing container deployment on the grid through unpacked.cern.ch service



2) CernVM-FS publisher on Kubernetes – *Andrea Valenzuela Ramirez*

- Proof of Concept: Ephemeral shells/containers for cvmfs publishing service
- Serverless model
  - Any client can be a publisher if it has the correct keys

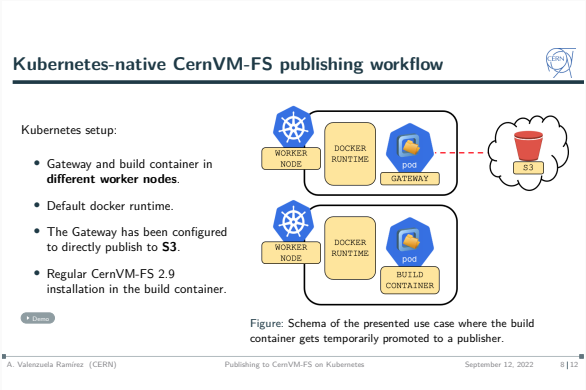


Figure: Schema of the presented use case where the build container gets temporarily promoted to a publisher.

3) CernVM 5: A fully containerized CernVM – *Jakob Karl Eberhardt*

- Until now: CernVM was a full VM → CernVM5 extends it to containers
- Phase-out *CernVM online* – CernVM 3 and 4 stay available
- CernVM 5 features: derivable image, resiliant to environment changes by experiment stacks
- Planned release date: Q4/2022

CernVM 5: Resulting Base Layer Image

	CernVM 5 Base Layer Image	Naively Derived Image	Naively Derived Image + System Applications
Volume (uncompressed)	805MB	1030MB	1450MB
Volume (compressed)	284MB	382MB	523MB
Installed Packages	457	502	612
Image Layers	single	multiple	multiple
Standard Derivability	yes	yes (but not a true base layer)	yes (but not a true base layer)

+ Better control over installed packages → smaller image, easier to distribute
+ Improved build time and caching
+ Single layered → less complex, faster to construct, true to concept of base layers
+ Compared to ≤ 4: Derivable image

CVMFS-Related News and Status IV

4) Harbor Registry at CERN: Status and Enhancements – *Ricardo Rocha*

- Allows for sophisticated container image management
- Mature project with vendor support
- CVMFS unpacked.cern.ch takes advantage of proxying

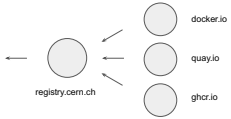
Proxy Caches

Pull-through cache for other registries

Already enabled a docker.io pull-through cache to cover for recent API restrictions

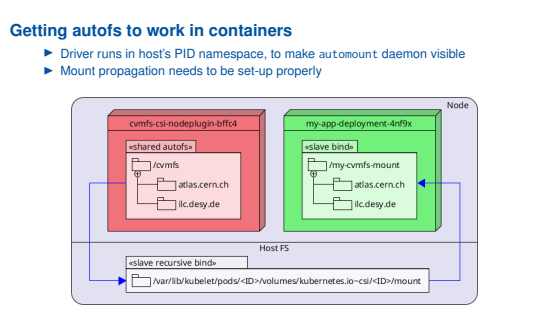
https://kubernetes.docs.cern.ch/docs/registry/quickstart/#pull-through-caches

Optimized access, also helps enforcing CVE/Vulnerability checks



5) CSI Driver for CernVM-FS – *Robert Vasek*

- *Container Storage Interface* is used for external storage in container orchestrators
- `Cvmfs-csi` exposes cvmfs as kubernetes-native external storage
- `Cvmfs-csi` version 2
  - Supports: `autofs`, client configuration during runtime, and any type of local cache volume



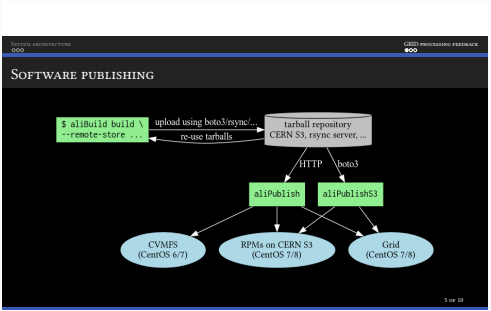
CVMFS Users – Scientific Community – Overview

	Talk	Speaker
1)	ALICE's software infrastructure	*Timo Wilken*
2)	CVMFS usage and performance for LHCb	*Ben Couturier*
3)	CVMFS at GSI	*Sören Lars Gerald Fleischer*
4)	CMS deployments on CernVM-FS	*Andrea Valenzuela Ramirez*
5)	ATLAS Installations on CVMFS	*Oana Vicky Boeriu*
6)	ATLAS Cloud R&D, Kubernetes and CVMFS	*Fernando Harald Barreiro Megino*
7)	CVMFS for KM3NeT	*Mieke Bouwhuis*
8)	Software infrasructure of XENONnT	*Joran Angevaare*

CVMFS Users - Scientific Community I

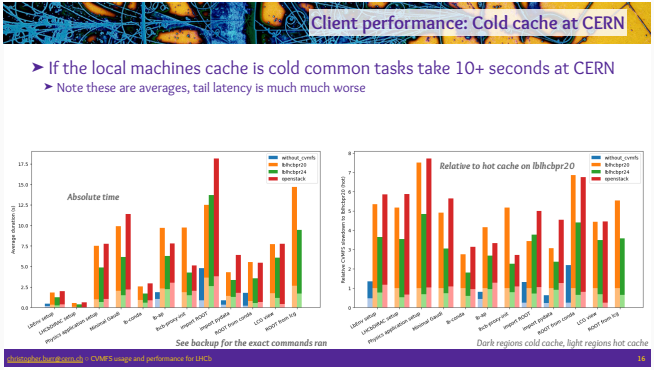
1) ALICE's software infrastructure – *Timo Wilken*

- Summary of the ALICE build & test system
- CVMFS challenge
  - ALICE software framework uses many *processes*, thus puts more load on the cvmfs client



2) CVMFS usage and performance for LHCb – *Ben Couturier*

- 3 CVMFS repositories
  - Software (8.9 TB), Nightlies (publish around 500 GB/day), Detector condition data
- Large Python stacks cause long startup delays with CVMFS cold cache



4) CMS deployments on CernVM-FS – *Andrea Valenzuela Ramirez*

- 3 repositories (each has 1 single release manager)
  - CMSSW, continuous integration and integration builds
- Has 500+ CMSSW container images in unpacked.cern.ch
- About to switch to parallel publishing/gateway (as well as ATLAS nightlies)

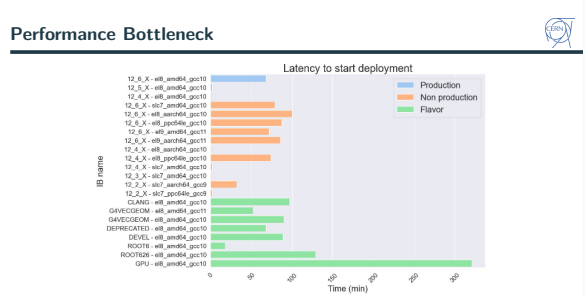
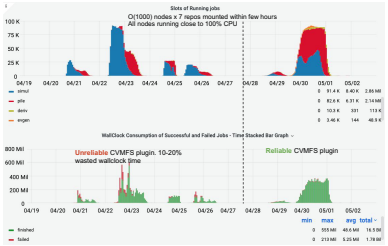


Figure: Waiting time of deployment jobs to `cms-ib.cern.ch` on 28/09 11pm IBs.

6) ATLAS Cloud R&D, Kubernetes and CVMFS – *Fernando Harald Barreiro Megino*

- Harvester kubernetes integration for batch allows K8S to run jobs like any Grid site
- 3 commonly used "blueprints" for cvmfs on k8s
  - CERN CSI driver, CERN Sciencebox driver, PRP OSG driver
- Case for an official CVMFS deployment for k8s

CVMFS plugin stress test: fast scale out on GCP

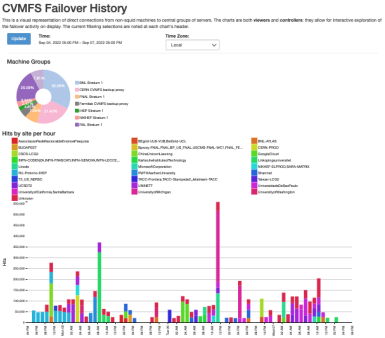


CVMFS Infrastructure

	Talk	Speaker
1)	WLCG Stratum 1 Status and Plans	*Dave Dykstra*
2)	CVMFS Mix at CERN	*Fabrizio Furano*
3)	A new UK StashCache at Edinburgh for DUNE	*Wenlong Yuan*
4)	CVMFS in Canadian Advanced Research Computing	*Ryan Taylor*

CVMFS Infrastructure I

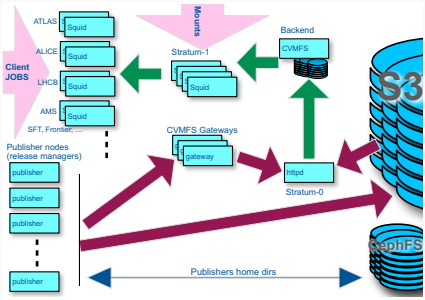
1) **WLCG Stratum 1 Status and Plans** – *Dave Dykstra*
   - Full replica needs 40 TB; raid-10 for optimal write speed
   - HA-pair preferred and reverse-frontier squid proxy for monitoring
   - Cloudfare CDN: openhtc.io



CVMFS Infrastructure II

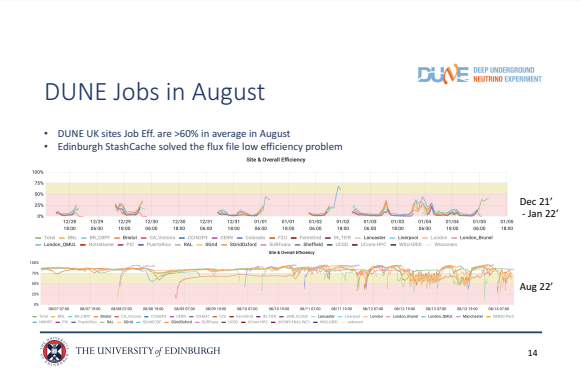
2) CVMFS Mix at CERN – *Fabrizio Furano*

- 4.5 billion files, >200 TB of data; 63 release managers and repositories
- CVMFS data on S3; release managers home are on CephFS



3) A new UK StashCache at Edinburgh for DUNE – *Wenlong Yuan*

- Use case for cvmfs as a namespace for large data repositories
- Set up a new UK Stash Cache instance at Edinburgh to reduce latency & I/O wait
- Increases efficiency of DUNE HTC grid jobs in the UK from <10% to >60%



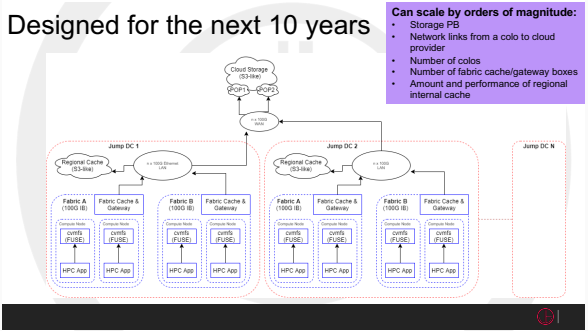
CVMFS Users – Industry and HPC

	Talk	Speaker
1)	Exascale data processing with CVMFS	*Matt Harvey*
2)	First experiences using CVMFS in Microsoft Azure	*Hugo Meiland*
3)	Lazy Container Image Distribution With eStargz And P2P Image Sharing on IPFS	*Kohei Tokunaga*
4)	Decarbonizing Scientific Computing	*Andrew Grimshaw*
5)	The European Environment for Scientific Software Installations (EESSI)	*Kenneth Hoste*

CVMFS Users - Industry and HPC I

1) Exascale data processing with CVMFS – *Matt Harvey*

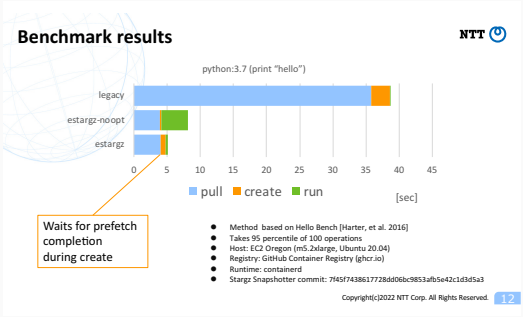
- Jump trading: privately-owned trading firm
- Data archive: Create 200 TB per week, increases 2x per year
- Before: `rsync` and GPFS – needed support of growth size >10
- Uses cvmfs as metadata storage that tracks where the actual data is stored



CVMFS-Related News and Status I

3) Lazy Container Image Distribution With eStargz And P2P Image Sharing on IPFS – *Kohei Tokunaga*

- eStargz images and stargz-snapshotter allow for lazy pulling
- Foundation for the cvmfs containerd snapshotter
- Potentially very big impact on container image distribution

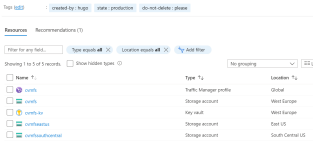


CVMFS Users - Industry and HPC II

2) First experiences using CVMFS in Microsoft Azure – *Hugo Meiland*

- Ongoing work to make cvmfs available to Azure HPC OnDemand clouds (#10 in Top500)
- Contribution for Azure Blob instead of S3 in CVMFS
- Close connection to EESSI project

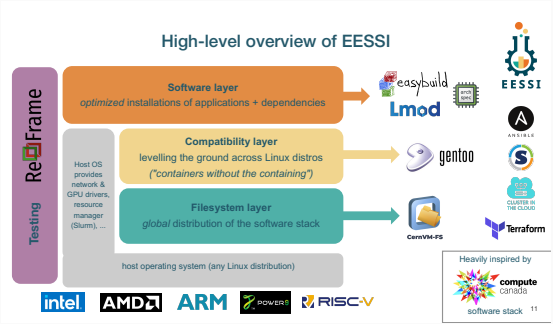
CVMFS code



- Added support for Azure Blob next to S3 (okt 2021?)
- Cvmfs_server can build stratum0 directly on Azure Blob
  - With keys in keyvault, build machine is expandable
- Using traffic manager i.s.o. geo-ip
- Sync containers i.s.o. stratum0 -> stratum1

5) EESSI – *Kenneth Hoste*

- Installing scientific software is (still) a tremendous problem for HPC sites
- EESSI project consists of three main layers
  - Global distribution of EESSI software stack via cvmfs
  - Compatibility layer using Gentoo Prefix
  - Software layer, hosting optimized installations along with required dependencies



Summary

- First in-person CernVM workshop since 2019
- Very productive and pleasant event
  Huge thanks to NIKHEF!
- General themes:
  - Significantly growing load to cvmfs client and server due to more supported architectures, more agile software development, more complex software stacks and more cores per box
  - Ongoing efforts on client performance improvements and parallelized publishing
  - Ongoing efforts on integration with container ecosystem: easier k8s support, scaling of the container image conversion service (unpacked.cern.ch)