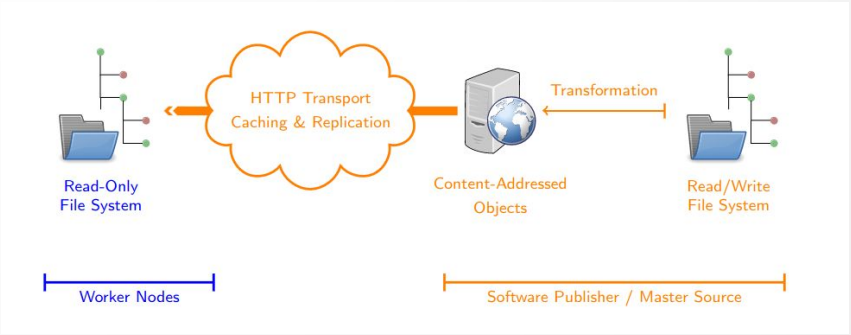
CernVM FileSystem (CVMFS)

Introduction from the development team -
Hepix Spring Workshop 2024

Apr 16th 2024, Hepix
Valentin Völkl for the CVMFS Development Team at CERN

What is CVMFS?

- Global, **read-only filesystem** for **software distribution**
  - with a user experience similar to an on-demand streaming service (... but for scientific software)
- implemented as a filesystem in userspace, via *libfuse*
  - allows client to be installed flexibly on all workernodes
- Optimized for storing and distributing software
  - Content-adressable storage allows **De-duplication**
  - Multi-level **caching**, use of HTTP transport
  - **Compression** of data
  - Verification of data **integrity**
  - **...**



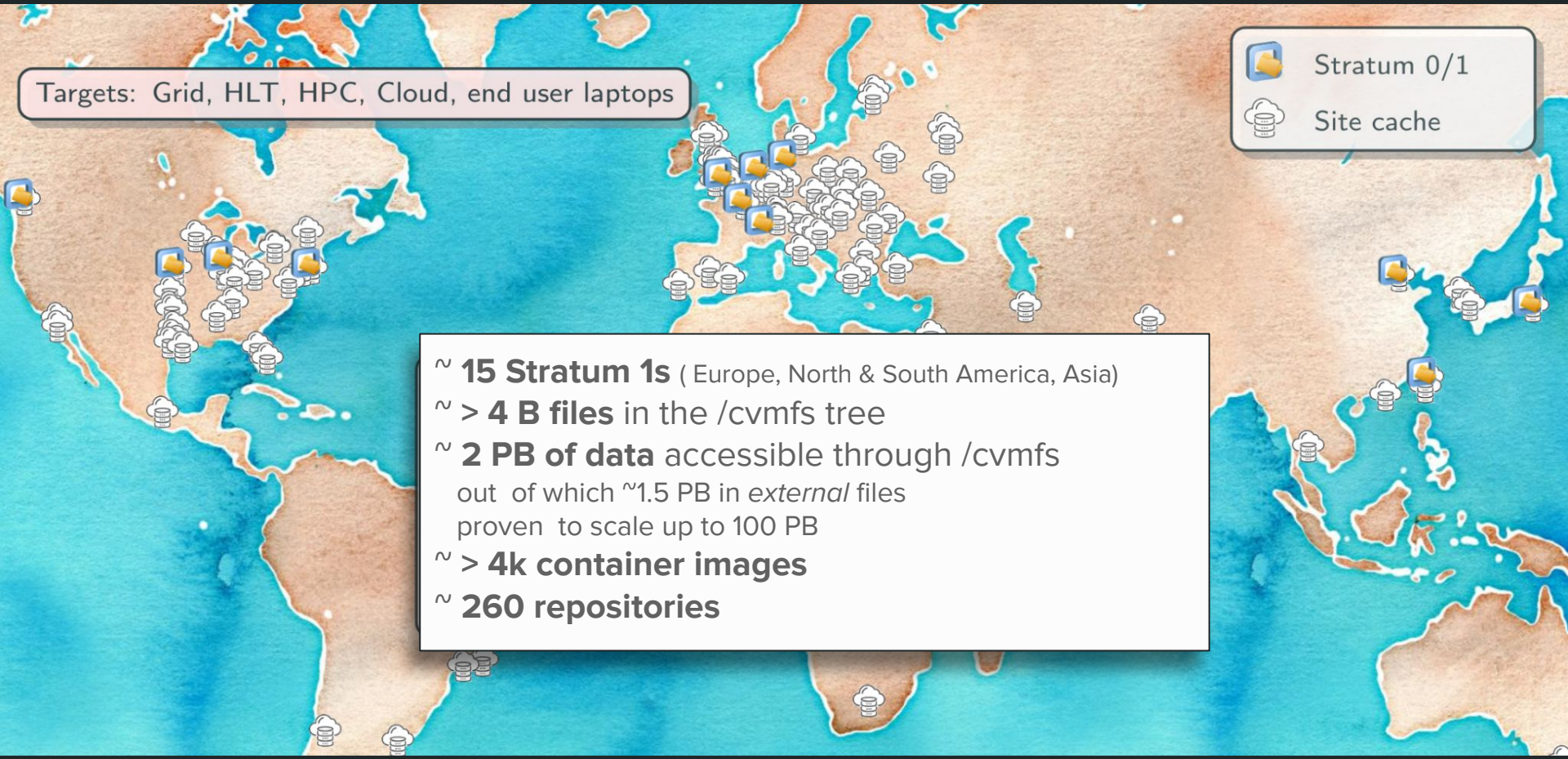
Key users:

- LHC & smaller CERN experiments
- Euclid, **Jump Trading** (contractual partners)
- Other scientific communities & industry (e.g., EESSI, LIGO, SKA, LSST, Roche, etc.)

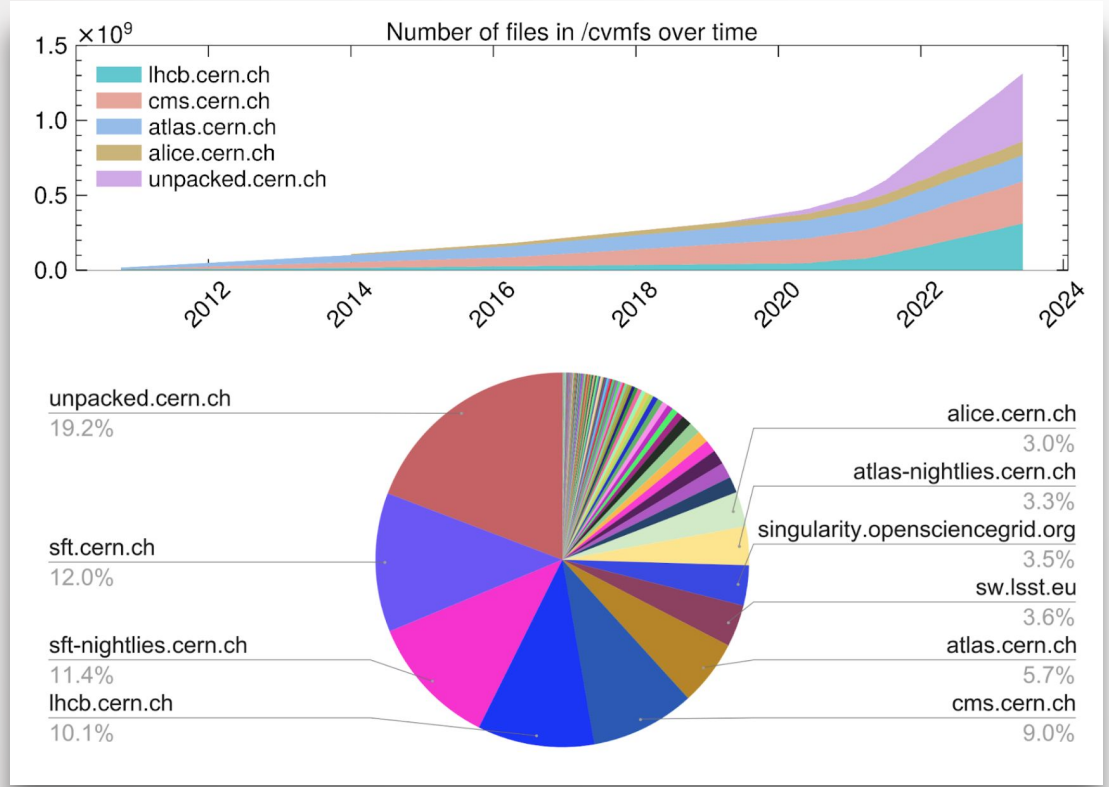
Key stakeholders:

- Experiments & end users: producers and consumers of data
- Site operators: focus on smooth operations, low-maintenance effort
- Stratum 1 operators: donate resources to the WLCG/cvmfs operations
- Developers: SFT, Jump Trading, Fermilab, community (“cvmfs-contrib”)

***See Matt Harveys talk in this contribution!**


Targets: Grid, HLT, HPC, Cloud, end user laptops
Stratum 0/1
Site cache
~ 15 Stratum 1s (Europe, North & South America, Asia)
~ > 4 B files in the /cvmfs tree
~ 2 PB of data accessible through /cvmfs
out of which ~1.5 PB in external files
proven to scale up to 100 PB
~ > 4k container images
~ 260 repositories

CVMFS in numbers



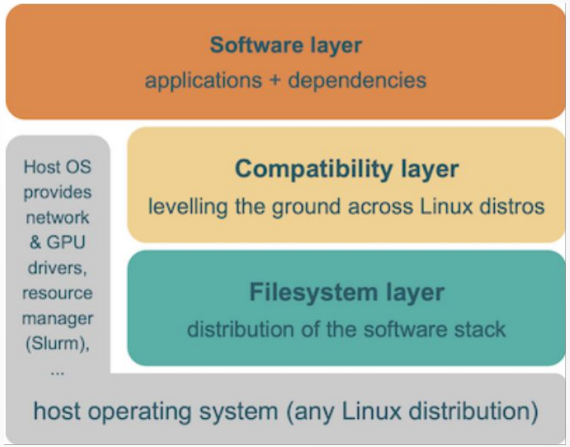
~ **15 Stratum 1s**
~ **> 4 B files** in the /cvmfs tree
~ **2 PB of data** accessible through /cvmfs
out of which ~1.5 PB in *external* files
proven to scale up to 100 PB
~ **> 4k container images**
~ **260 repositories**

- Backed by S3(+CEPH) or local storage
- Thanks to IT-Storage and the operators who expertly manage this infrastructure!

HPC sites can be a particular challenge, with many restrictions.

The CVMFS development team supports the EESSI project, which provides unified software installations to European HPC sites on CVMFS.

EESSI
EUROPEAN ENVIRONMENT FOR SCIENTIFIC SOFTWARE INSTALLATIONS



Containers

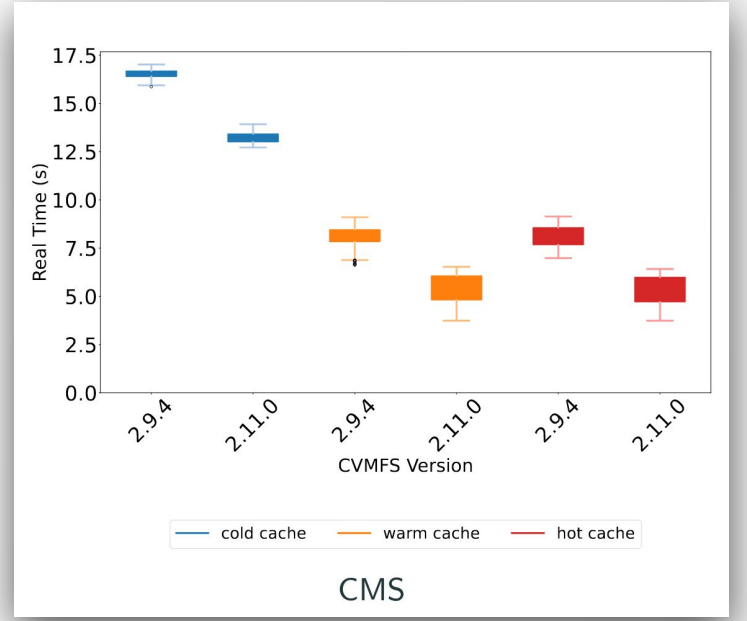
- CVMFS provides tooling to unpack, store and distribute containers, with *unpacked.cern.ch* being the biggest repository:

```
~$ ls /cvmfs/unpacked.cern.ch/registry.hub.docker.com/cmssw/cs8\:x86_64-d20211124
afs   build  dev          etc   lib64   mnt   proc  sbin         sys  var
bin   cvmfs  environment  home  lost+found  opt   root  singularity  tmp
boot  data   eos          lib   media   pool  run   srv          usr
```

- *Apptainer* can directly launch the container from this root file system.
- The same benefits from using CVMFS apply! Leading to:
  - Drastically faster container **startup** times
  - Automatic **cache management** of container images on the worker nodes

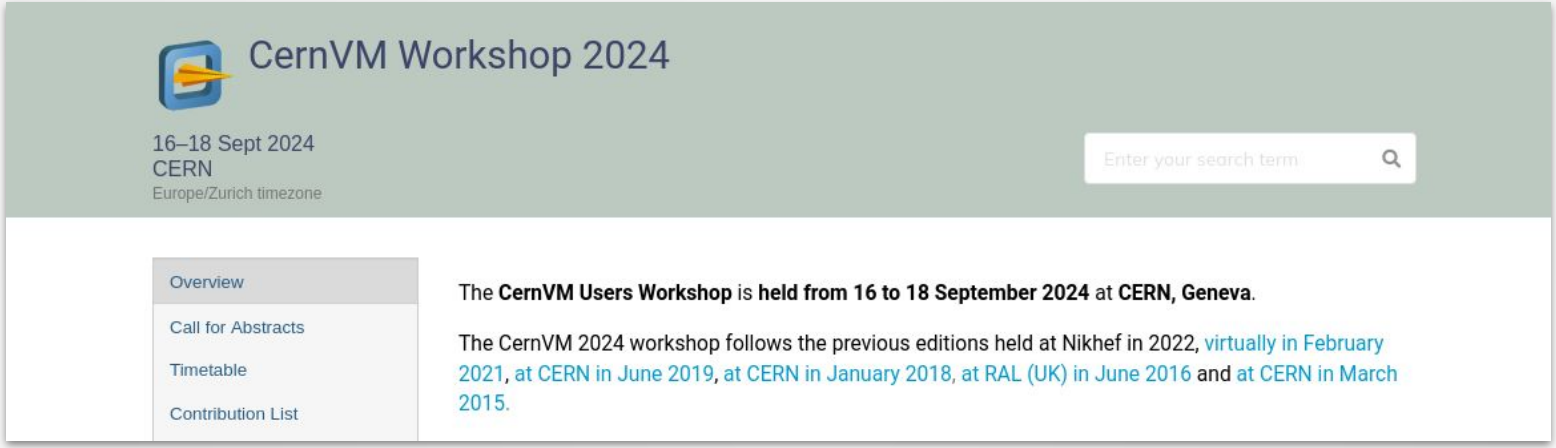
Performance engineering & recent developments

- Jump Trading has pushed the technical limits of CVMFS in key areas - highly parallelized workloads and high-frequency publications
  - Provided more than 40 PRs, as well as feedback and testing!
  - See following talk by M. Harvey
- Parallel decompression of objects after download
- Support for Zstd compression



L. Promberger, client speedup by better use of kernel page cache, see CHEP 2023 for more details

CernVM Workshop



Stay tuned & Register!

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

Thank you!