

# CernVM FileSystem (CVMFS)

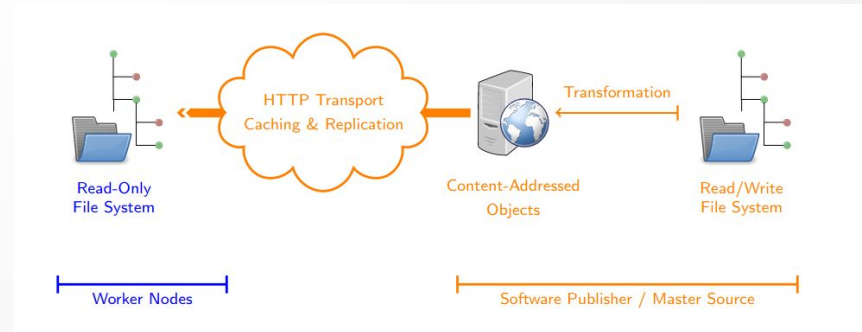
Introduction from the development team -  
Hepix Spring Workshop 2024



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-addressable storage allows **De-duplication**
  - Multi-level **caching**, use of HTTP transport
  - **Compression** of data
  - Verification of data **integrity**
  - ...



A world map with a light beige and blue color scheme. Numerous lightbulb icons are scattered across the map, primarily concentrated in Europe and North America. A semi-transparent white box is overlaid on the map, containing text.

## Key users:

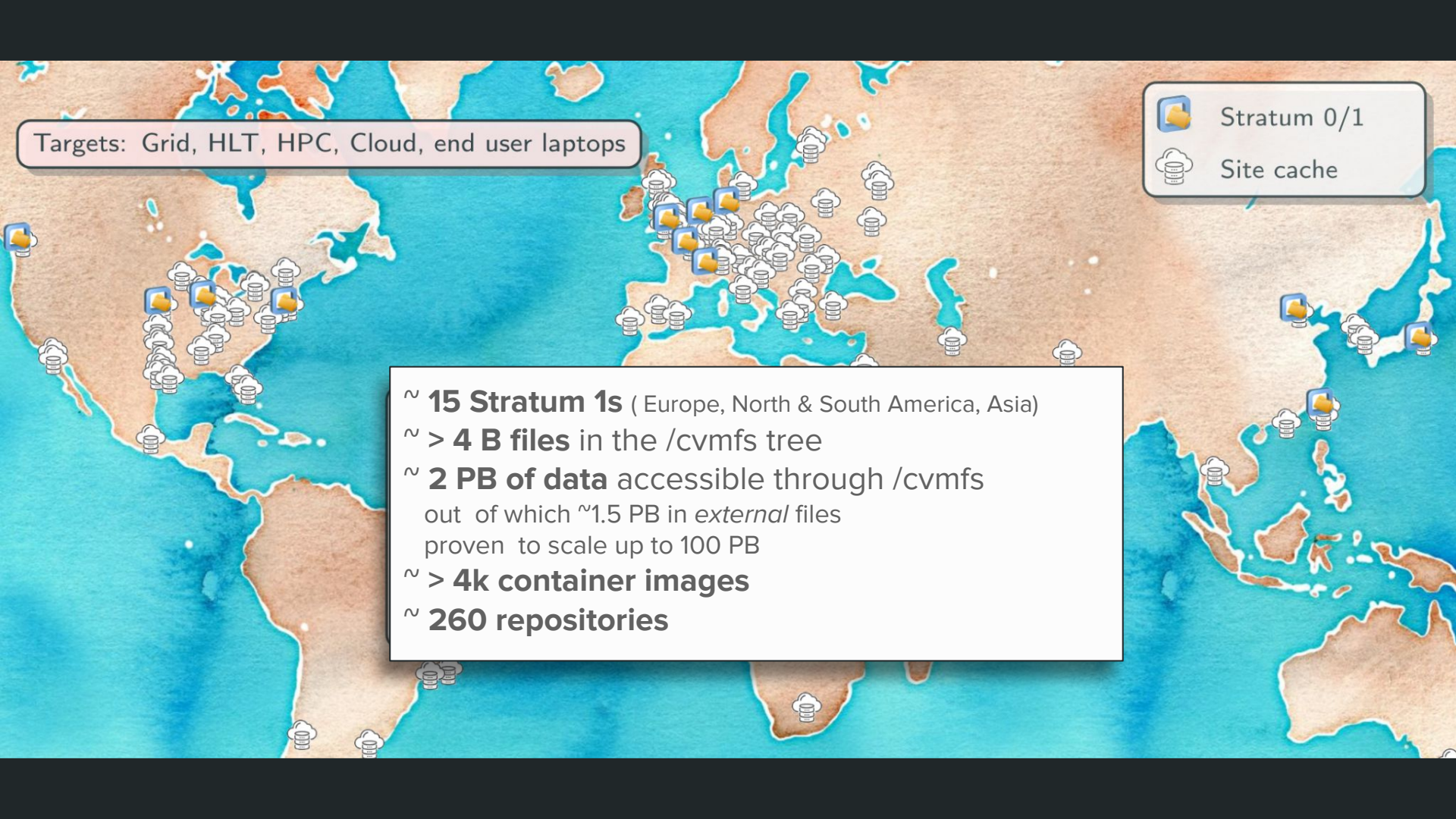
- LHC & smaller CERN experiments
- Euclid, **Jump Trading** (contractual partners)
- Other scientific communities & industry (e.g., EESSI, LIGO, SKA, LSST, Roche, etc.)

## 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!**



A world map with a watercolor-style background. The map is populated with two types of icons: blue folders with yellow arrows (representing Stratum 0/1) and white server racks (representing Site cache). The Stratum 0/1 icons are concentrated in North America, Europe, and East Asia. Site cache icons are more widely distributed across all major landmasses. A legend in the top right corner identifies these icons. A text box in the top left lists targets: Grid, HLT, HPC, Cloud, and end user laptops. A large text box in the center provides statistics for the Stratum 1s.

Targets: Grid, HLT, HPC, Cloud, end user laptops

 Stratum 0/1

 Site cache

~ **15 Stratum 1s** ( Europe, North & South America, Asia)

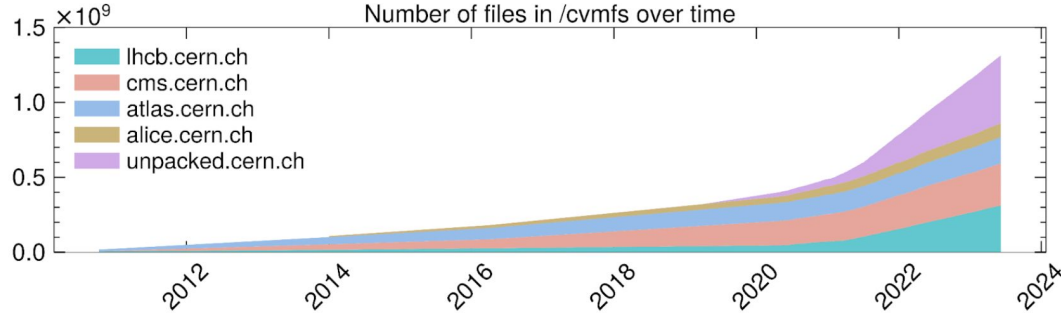
~ **> 4 B files** in the /cvmfs tree

~ **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



unpacked.cern.ch

19.2%

sft.cern.ch

12.0%

sft-nightlies.cern.ch

11.4%

lhcb.cern.ch

10.1%

alice.cern.ch

3.0%

atlas-nightlies.cern.ch

3.3%

singularity.opensciencegrid.org

3.5%

sw.lsst.eu

3.6%

atlas.cern.ch

5.7%

cms.cern.ch

9.0%

~ **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

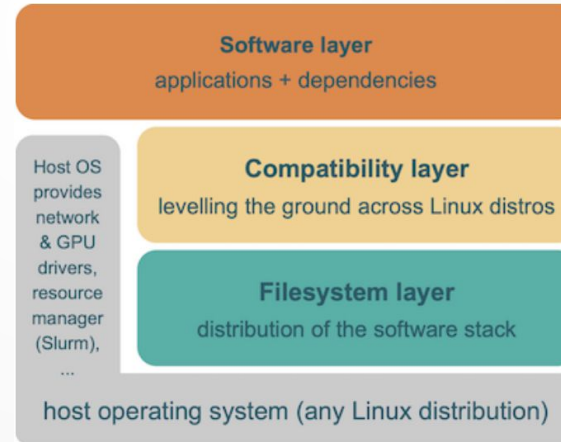
~ **> 4k container images**

~ **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.



# Containers

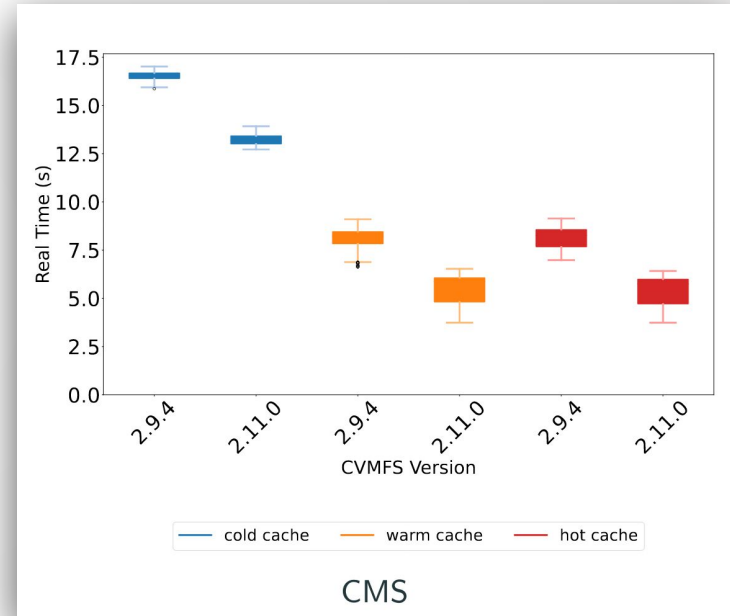
- CVMFS provides tooling to unpack, store and distribute containers, with *unpacked.cern.ch* being the biggest repository:

```
~$ 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



The screenshot shows the Indico event page for the CernVM Workshop 2024. The header features the event logo (a blue square with a yellow pencil) and the title "CernVM Workshop 2024". Below the title, the dates "16–18 Sept 2024" and location "CERN" are listed, along with the time zone "Europe/Zurich timezone". A search bar is present on the right side of the header. The main content area has a left sidebar with navigation links: "Overview" (highlighted), "Call for Abstracts", "Timetable", and "Contribution List". The main text area contains the following information:

The **CernVM Users Workshop** is held from **16 to 18 September 2024** at **CERN, Geneva**.

The CernVM 2024 workshop follows the previous editions held at Nikhef in 2022, [virtually in February 2021](#), [at CERN in June 2019](#), [at CERN in January 2018](#), [at RAL \(UK\) in June 2016](#) and [at CERN in March 2015](#).

**Stay tuned & Register!**

**<https://indico.cern.ch/e/cvm24>**

**[Thank you!](#)**