CernVM-FS

With material from Steve Traylen, Jakob Blomer

Andrea Dotti (andrea.dotti@cern.ch)
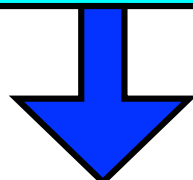
What is it?

- CernVM-FS: **Cern V**irtual **M**achine **F**ile **S**ystem
  - An **HTTP file system** based on FUSE to distribute software
  - Originally used as a way to distribute LHC experiments' sw to Virtual Machines
  - Now independent of VM
  - Being **deployed at GRID sites** around the world (for ATLAS, LHCb, AMS)
- **Geant4 repository is available**
  - For example, go to lxplus and do: ls /cvmfs/geant4.cern.ch
 - We use this system since 2011 for GRID validation

Operating System & Applications

CernVM-FS

OS Kernel

Fuse

HTTP Content Distribution Network

File System Buffers

CernVM-FS Hard Disk Cache

CernVM-FS "Repository" (All Releases Available)
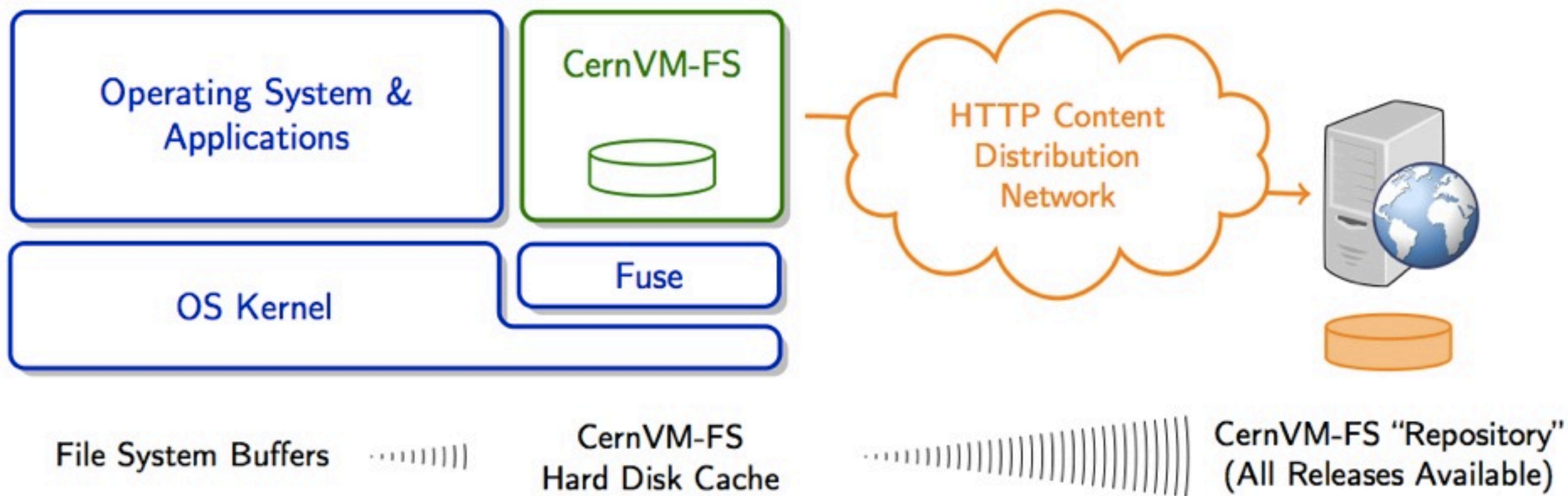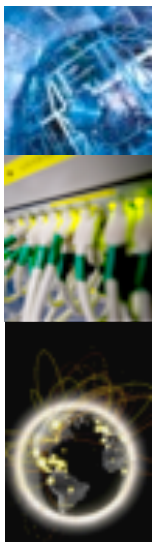
Stratum 0

/cvmfs/repo/MyFile — Shadow Tree: the one write location.

cvmfs_sysc, operates on all new files in repo, e.g MyFile.

/repo/A345....de43b — Public tree: contains hashed compressed files.

Stratum 0 Web Server - only Stratum 1s ever connect.
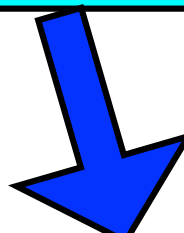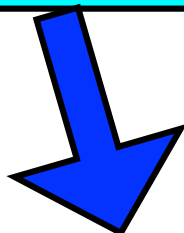
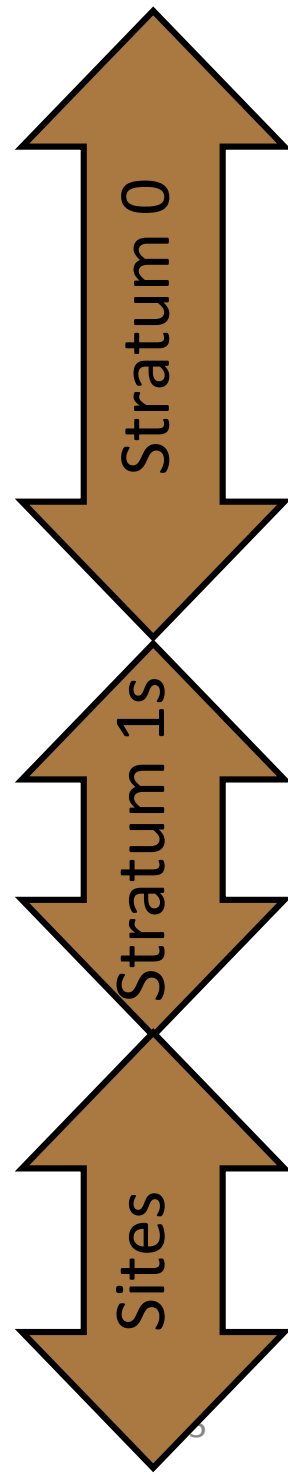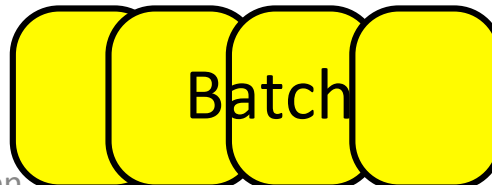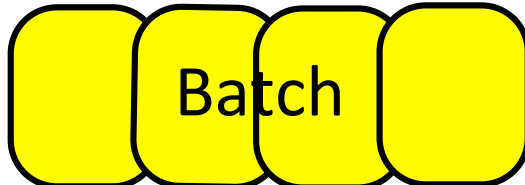Stratum ones copy all new data with "cvmfs_replicate".

Stratum 1s

Stratum 1 Full Copy

Stratum 1 Full Copy

Geo separated and fully redundant.

Sites

SiteA OnDemand Cache

SiteB Partial Cache — Squids

Batch

Batch

Steve Traylen, CERN IT.

4

Current work

⊣ Re-organizing our repository to **match "AFS" style**

⊣ Installing tools for **GRID usage**

⊣ Repository layout:

  ⊣ **/cvmfs/geant4.cern.ch** : Top level directory

    ⊣ .../geant4/9.5.ref07 : Specific release

    ⊣ .../geant4/9.5.ref07/x86_64-slc5-gcc43-opt : Binaries

    ⊣ .../geant4/9.5.ref07/share : GNUMakefile, CMake, include files

    ⊣ .../opt/ : Additional software (GRID validation, application binaries)

Benefits

- **Very similar to AFS**: CervnVM-FS client for Linux (very easy to install), MAC-OSX coming soon
  - KEK colleagues installed it on their GRID cluster in few days!
- Differently from AFS it's becoming **available on the GRID**
  - Wherever ATLAS runs it's very probably CVMFS is already there, enabling of G4 repository is very little configuration extra
- On client side it is read-only: no worry about corrupting repository
- **Very efficient**: client-side local cache, hashing of files
  - Network usage is limited to needed files
  - Once cache if populated access to files is very fast, and it works even if network is down (robust against network hiccups)

To Consider

- Central repositories are four "Stratum1"
  - "Stratum 0" is R/W, but clients never connect to it directly
  - **For US or Asian colleagues it would be wise to have our own local mirror (reduce latency)**
    - Simple to do: since it's HTTP needs one (or more) SQUID servers, very standard technology
- **Delay** between repository update and client
  - It takes few hours (will be reduced in the future) for an update of the central repository to be propagated world-wide
- No support (yet?) for file access permission
  - Whoever installs the client can read our repositories
  - For this reason **we distribute only binaries**

Conclusions

- CernVM-FS is a very attractive way to distribute Geant4 binaries
  - Linux installation (SLC5) already in place
  - MacOSX could follow
- It is used by large LHC experiments: can perfectly deal with our software size
  - Our experience so far for GRID validation is very positive
- We would like to advertise its usage in our community
  - Each release will be made available on repository
  - Link from Geant4 download area?