
# CernVM-FS – Beyond LHC Computing

Ian Collier, Catalin Condurache

STFC RAL Tier 1

HEPiX Ann Arbor November 1st 2013

# What is CVMFS?

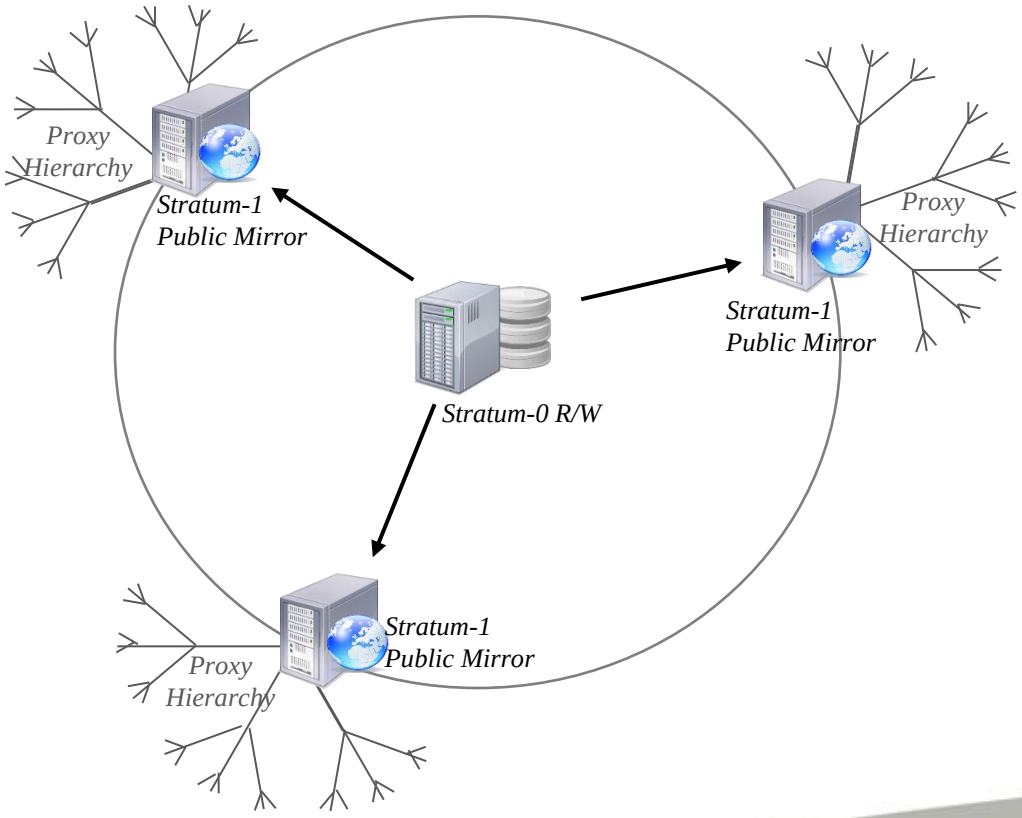
- Read-only, distributed filesystem, originally developed to get frequently changing VO software to VMs that might not have access to software servers.
- Data integrity and validity are ensured by the signed file catalog and access authentication for software server updates (done by Software Grid Manager or other privileged member of the VO).
- Built using standard technologies (fuse, sqlite, http, squid proxies and caches).
- Removes the need for local installation jobs and conventional software servers at sites & helps standardise the computing environment across the Grid.
- Once the signed catalog has been downloaded and mounted, metadata operations require no further network access. Together with the file based de-duplication this makes CernVM-FS efficient in terms of disk usage and network traffic.
- The software needs one single installation and then is available at any site with CernVM-FS client installed

# CernVM-FS at RAL Tier-1

- Summer 2010 - RAL was the first Tier-1 centre to test CernVM-FS at scale and worked towards getting it accepted and deployed within WLCG.
- February 2011 – first CernVM-FS Stratum-1 replica for LHC VOs in operation outside CERN.
- December 2012 – first CernVM-FS Stratum-0 services deployed for mice and na62 VOs supported by the GridPP UK Project.
- January – June 2013 – CernVM-FS Stratum-0 service extended to international small VOs (hone, phys.vo.ibergrid.eu, enmr.eu).
- Ongoing involvement in verification of CernVM-FS client software at scale before it is released into production.
- RAL Tier 1 now leading EGI task force to establish network of Stratum 0 & Stratum 1s across EGI sites.

# CernVM-FS WLCG deployment

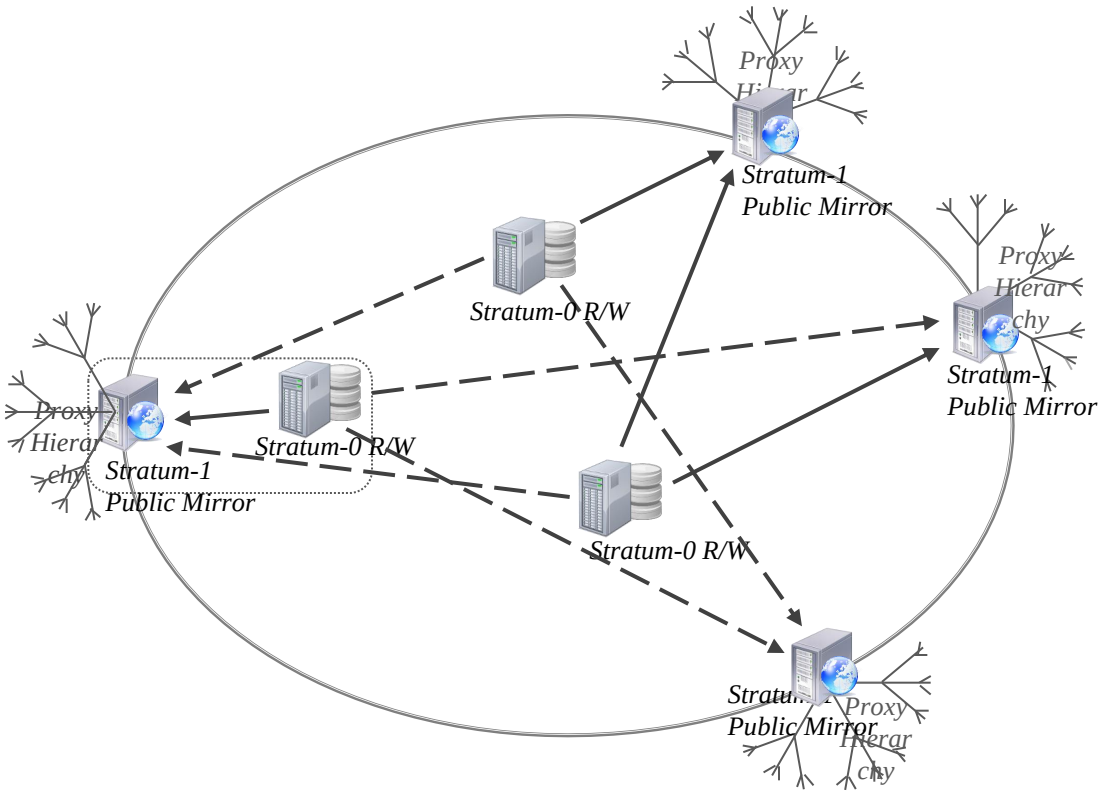
- Software is installed by LHC VOs at Stratum-0 hosted at CERN and replicated to Stratum-1 hosted by WLCG Tier-1 sites
- CernVM-FS clients connect to one of the Stratum-1 services (via local squid caches)
- Client manages transparent failover to other Stratum-1 in case of connection problems




Science & Technology Facilities Council

# CernVM-FS EGI deployment

- Stratum-0 (source repositories) and Stratum-1 (replicas) can be geographically co-located, or not
- Stratum-1 can replicate a whole Stratum-0 (solid), or can partially replicate (dotted) – the ***'relaxed'*** model




Science & Technology Facilities Council

# CernVM-FS Task Force

- Set up by EGI to establish a CernVM-FS infrastructure that allows EGI VOs to use it as a standard method of distribution of their software across the EGI computing resources by:
  - promoting the use of CernVM-FS technology by VOs
  - creating a network of sites providing CernVM-FS services (Stratum-0, Stratum-1, squid)
- Collaboration and cross replication with Open Science Grid (OSG) and WLCG, as well as sharing configurations & monitoring tools

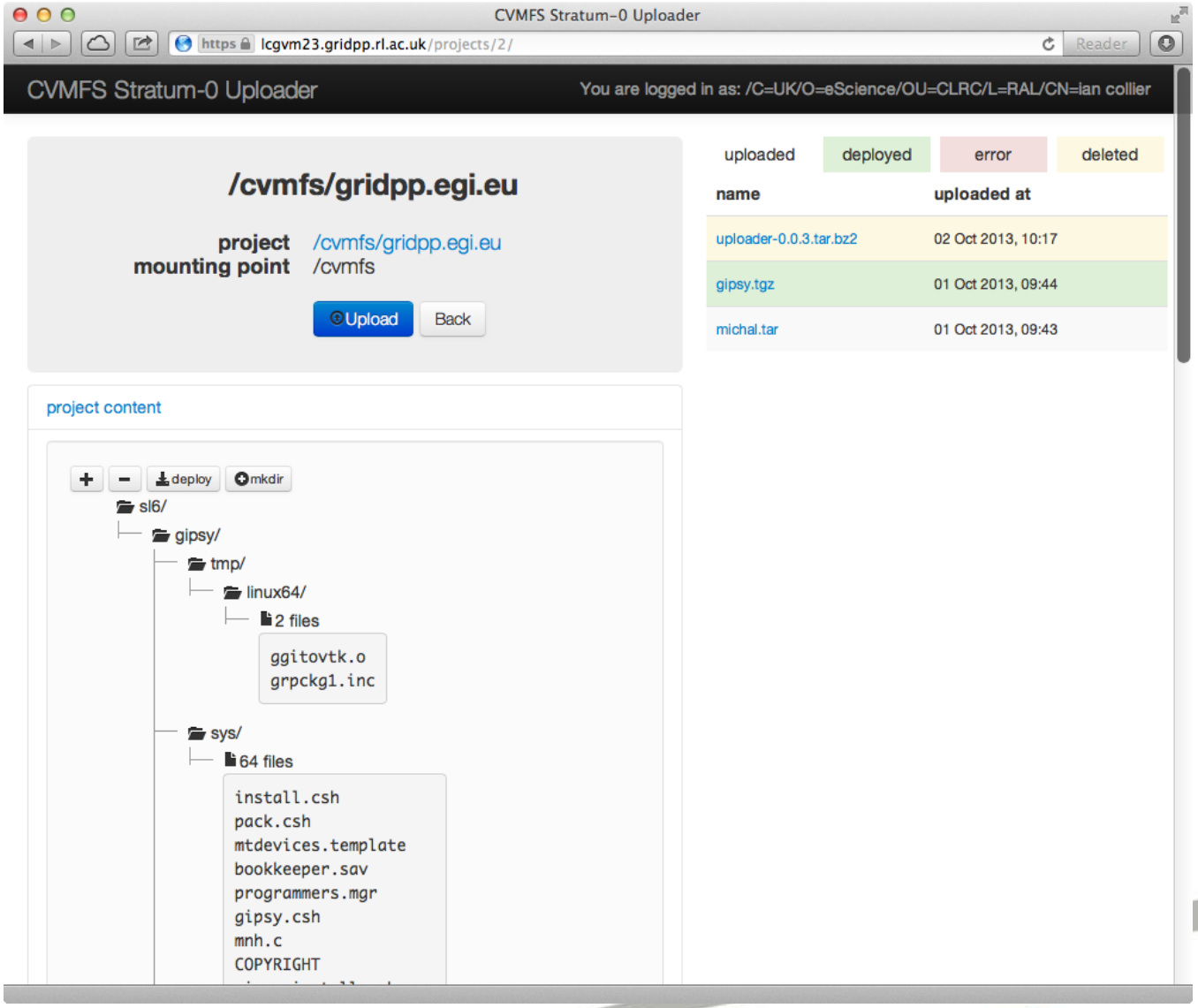
# *enmr.eu* CernVM-FS Stratum-0 at RAL Tier-1

- WeNMR - EGI supported VO providing a platform integrating the computational approaches necessary for NMR and SAXS data analysis and structural modelling. Its VO – enmr.eu –largest in Life Sciences.
- Over 570 registered users and a steady growth.
- More than 25 sites -> software distribution a large amount of work.
- CernVM-FS Stratum-0 repository for enmr.eu at RAL Tier-1 has reduced the effort required at supporting sites maintaining multiple NFS software areas.
- Bringing new sites on line much easier
- Installation jobs run by the enmr.eu/Role=Manager at RAL Tier-1 batch farm upload and configure new software releases under /cvmfs/wenmr.egi.eu directory.
- Now consoldiating the enmr.eu Stratum-0 and to move toward a CernVM-FS only environment.

Science & Technology Facilities Council

# CernVM-FS Stratum-0 Web Frontend

- Web application for CernVM-FS Stratum-0 uploads used as an alternative to installation jobs or ‘power users’.
- Developed by a student on an Erasmus Programme placement at RAL-Tier 1 UK.
- Users can upload tarballs and unpack them within the /cvmfs/<repo_name> ‘space’, followed by synchronization with the real CernVM-FS Stratum-0 repository.
- Authenticates with X509 certificates (managed by a web server)
  - Further authentication mechanisms can be added
- Removes need for privileged roles and jobs at sites




Science & Technology Facilities Council

# In Summary

- 3 years ago cvmfs was a promising idea  to solve
- Now the way WLCG experimets distribute their software across grid
- Growing number of local services & small support for non-LHC Vos
- Now entire alternative infrastructures
  - OSG established last year
  - EGI led by RAL
  - Characterised by sharing ideas & resources
- Benefit already been seen in new communities

Science & Technology Facilities Council

# Questions??