The EGI CernVM-FS Infrastructure
Evolution Towards a Global Facility and Latest Developments

Catalin Condurache

STFC RAL UK

ISCG, Taipei, March 2017

Outline

- Introduction
- Brief history
- EGI CernVM-FS infrastructure
- About the users
- Recent developments
- Plans

Introduction - CernVM-FS ?

- Read-only network file system based on HTTP that is designed to deliver scientific software onto virtual machines and physical worker nodes in a fast, scalable and reliable way
- Built using standard technologies (fuse, sqlite, http, squid and caches)

Introduction - CernVM-FS ?

- Files and directories are hosted on standard web servers and get distributed through a hierarchy of caches to individual nodes
- Mounted in the universal ***/cvmfs*** namespace at client level
- Software needs one single installation, then it is available at any site with CernVM-FS client installed and configured

Introduction - CernVM-FS ?

- The method to distribute HEP experiment software within WLCG, also adopted by other computing communities outside HEP
- Can be used everywhere (because of http and squid) i.e. cloud environment, local clusters (not only grid)
- Add CernVM-FS client to a VM image =>

  /cvmfs space automatically available

Brief History

- Following success of using CernVM-FS as primary method of distribution of experiment software and conditions data to WLCG sites…
- …Sep 2012 – non-LHC Stratum-0 service at RAL Tier1
  - supported by GridPP UK project
  - '*gridpp.ac.uk*' name space
- …Aug 2013 – expansion to EGI level
  - initiative to establish a CernVM-FS infrastructure that allowed EGI VOs to use it as a standard method of distribution of their software at grid sites
- *'egi.eu'* new space name for repositories

EGI CernVM-FS Infrastructure

- Stratum-0 service @ RAL
  - maintains and publishes the current state of the repositories
  - 32GB RAM, 12TB disk, 2x E5-2407 @ 2.20GHz
  - cvmfs-server v2.3.2 (includes the CernVM-FS toolkit)
  - 31 repositories – 780 GB
  - *egi.eu*
    - *auger, biomed, cernatschool, chipster, comet, config-egi*
    - *dirac, extras-fp7, galdyn, ghost, glast, hyperk, km3net*
    - *ligo, lucid, mice, neugrid, pheno, phys-ibergrid, pravda*
    - *researchinschools, snoplus, supernemo, t2k, wenmr, west-life*
  - *gridpp.ac.uk*
    - *londongrid, scotgrid, northgrid, southgrid, facilities*

Science & Technology Facilities Council

EGI CernVM-FS Infrastructure

- CVMFS Uploader service @ RAL
  - in-house implementation that provides upload area for *egi.eu* (and *gridpp.ac.uk*) repositories
  - currently 1.28 TB – repo master copies
  - GSI-OpenSSH interface (gsissh, gsiscp, gsisftp)
    - similar to standard OpenSSH tools with added ability to perform X.509 proxy credential authentication and delegation
    - DN based access, also VOMS Role possible
  - rsync mechanism between Stratum-0 and Uploader

EGI CernVM-FS Infrastructure

- Stratum-1 service
  - standard web server (+ CernVM-FS server toolkit) that creates and maintains a mirror of a CernVM-FS repository served by a Stratum-0 server
  - worldwide network of servers (RAL, NIKHEF, TRIUMF, ASGC, IHEP) replicating the *egi.eu* repositories
  - RAL – 2-node HA cluster (cvmfs-server v2.2.3)
    - each node – 64 GB RAM, 55 TB storage, 2xE5-2620 @2.4GHz
    - it replicates 65 repositories – total of 16 TB of replica

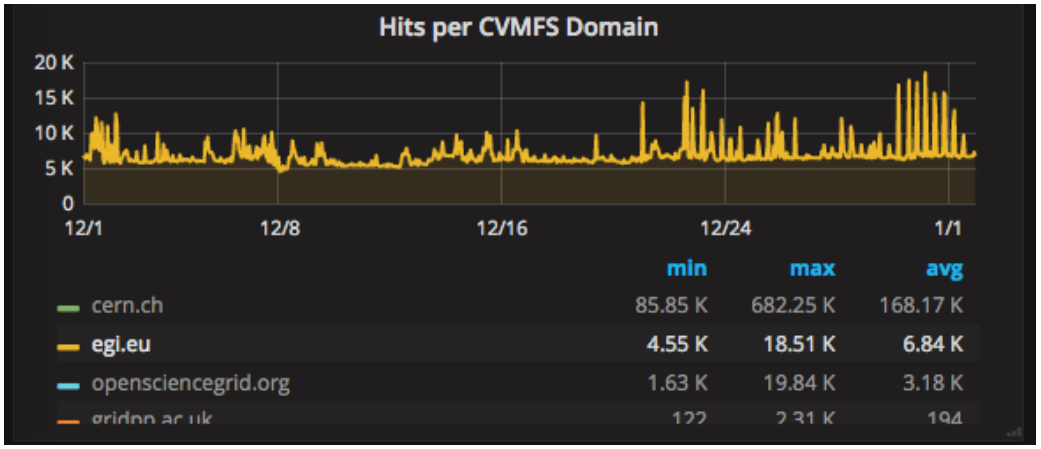
      - *egi.eu, gridpp.ac.uk* and *nikhef.nl* domains

      - also many *cern.ch, opensciencegrid.org* and *desy.de* repositories

Science & Technology Facilities Council

EGI CernVM-FS Infrastructure

- Stratum-1 service – plots, statistics
  - RAL - ~400 reqs/min, 350 MB/s
    - *egi.eu* - 2 - 4 reqs/s and 25 - 35 kB/s



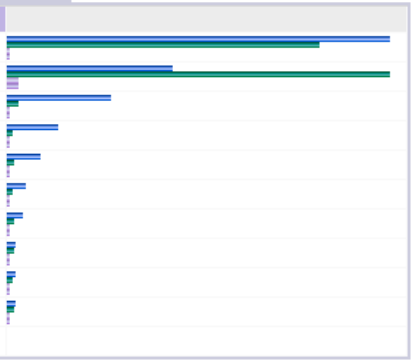
Science & Technology Facilities Council

EGI CernVM-FS Infrastructure

- Stratum-1 service – plots, statistics
  - TRIUMF – *egi.eu* only
    - up to 2 reqs/s
    - up to 3 kB/s

Pages-URL (Top 10) - Full list - Entry - Exit

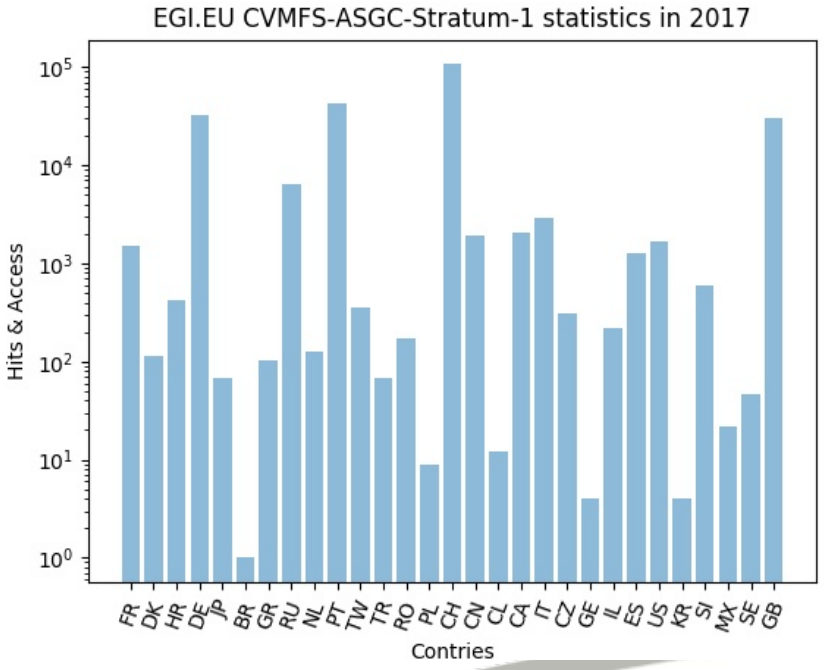
26 different pages-url	Viewed	Average size	Entry	Exit
cvmfs/snoplus.egi.eu	341596	33.92 KB	337	336
cvmfs/dirac.egi.eu	147548	41.43 KB	9942	9964
cvmfs/auger.egi.eu	92726	1.23 KB	600	597
cvmfs/phys-ibergrid.egi.eu	45200	578 Bytes	618	610
cvmfs/biomed.egi.eu	29859	795 Bytes	1332	1411
cvmfs/wenmr.egi.eu	16586	587 Bytes	784	770
cvmfs/t2k.egi.eu	14135	770 Bytes	178	188
cvmfs/mice.egi.eu	7631	653 Bytes	213	69
cvmfs/pheno.egi.eu	7379	646 Bytes	246	247
cvmfs/cernatschool.egi.eu	6883	672 Bytes	39	33
Others	88136	806 Bytes	837	838



Science & Technology Facilities Council

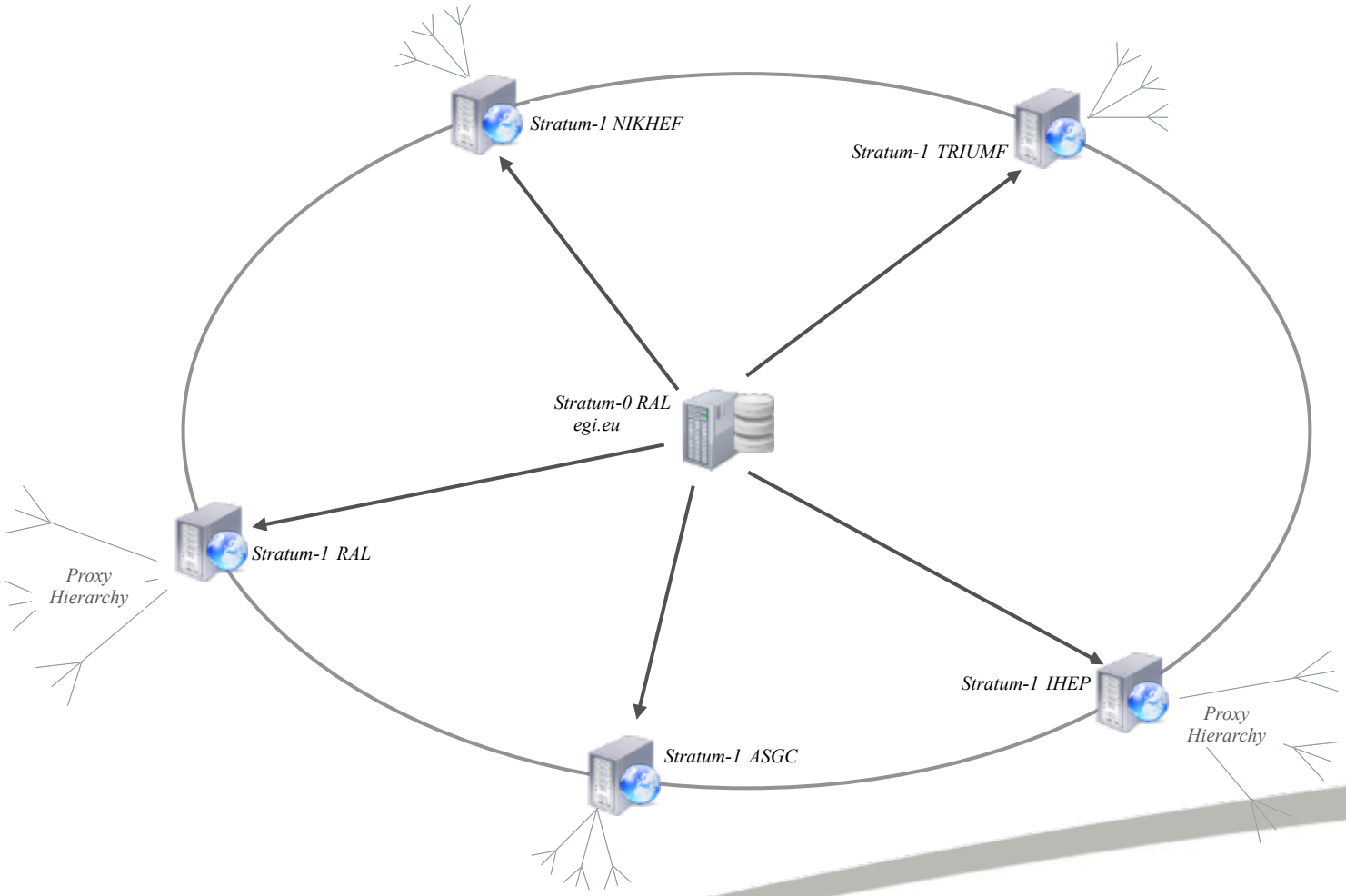
EGI CernVM-FS Infrastructure

- Stratum-1 service – plots, statistics
  - NIKHEF – *egi.eu* – 1 req/s, 12 kB/s
  - ASGC



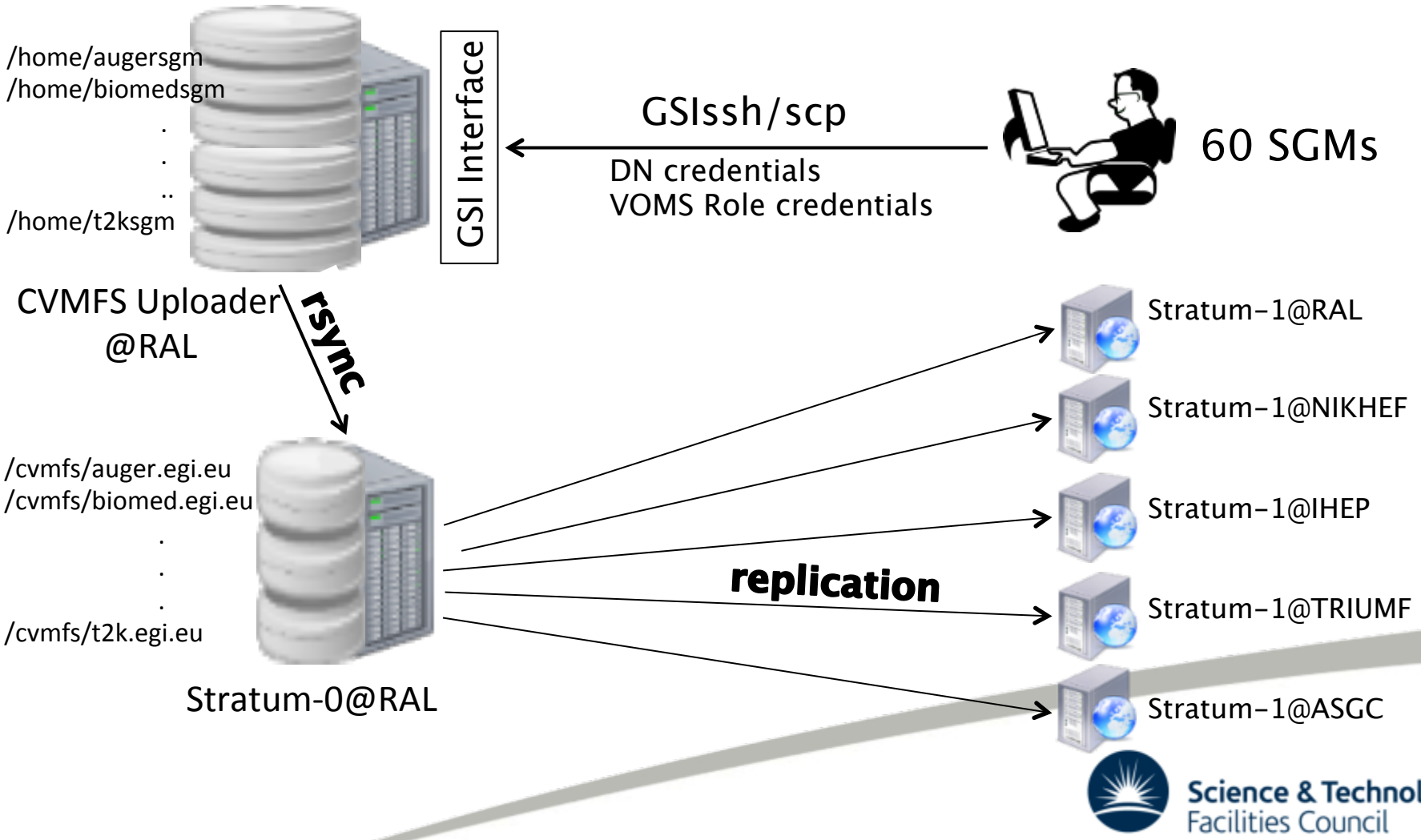
Science & Technology Facilities Council

EGI CernVM-FS Infrastructure Topology



Science & Technology Facilities Council

Repository Uploading Mechanism @ RAL



Science & Technology Facilities Council

Who Are the Users?

- Broad range of HEP and non-HEP communities
- High Energy Physics
  - *comet, hyperk, mice, t2k, snoplus*
- Medical Sciences
  - *biomed, neugrid*
- Physical Sciences
  - *cernatschool, comet, pheno*
- Space and Earth Sciences
  - *auger, glast, extras-fp7*
- Biological Sciences
  - *chipster, enmr*

The Users - What Are They Doing?

Grid Environment

- snoplus.snolab.ca VO
  - uses CernVM-FS for MC production (also ganga.cern.ch)
- cernatschool.org VO
  - educational purpose, young users get used with grid computing
  - software unit tests maintained in the repository
- dirac.egi.eu
  - repository maintained by the DIRAC interware developers
  - contains the DIRAC clients, environment settings for various DIRAC services (France Grilles, GridPP, DIRAC4EGI)
  - repository is therefore accessed by any user submitting to a DIRAC service

The Users - What Are They Doing?

Grid Environment

- auger VO
  - simulations for the Pierre Auger Observatory at sites using the same software environment provisioned by the repository
- pheno VO
  - maintain HEP software – Herwig, HEJ
  - daily automated job that distributes software to CVMFS
- other VOs
  - software provided by their repositories at each site ensures similar production environment

The Users - What Are They Doing?
Cloud Environment

- chipster
  - the repository distributes several genomes and their application indexes to ‘chipster’ servers
  - without the repo the VMs would need to be updated regularly and become too large
  - four VOs run ‘chipster’ in EGI cloud (test, pilot level)
- enmr.eu VO
  - use DIRAC4EGI to access VM for GROMACS service
  - repository mounted on VM
- other VOs
  - mount their repo on the VM and run specific tasks (sometime CPU intensive)

EGI CernVM-FS Service

Recent Developments

- Operations Level Agreement for Stratum-0
  - between STFC and EGI.eu
  - provisioning, daily running and availability of service
  - service to be advertised through the EGI Service Catalog
- Two EGI Operational Procedures
  - process of enabling the replication of CernVM-FS spaces across OSG and EGI CernVM-FS infrastructures - https://wiki.egi.eu/wiki/PROC20
  - process of creating a repository within the EGI CernVM-FS infrastructure for an EGI VO – https://wiki.egi.eu/wiki/PROC22

EGI CernVM-FS Service Developments

'Protected' CernVM-FS Repositories

- Repositories natively designed to be public with non-authenticated access
  - one needs to know only minimal info - access to the public signing key and repository URL
- Widespread usage of technology (beyond LHC and HEP) led to use cases where software needed to be distributed was not public-free
  - software with specific license for academic use
  - communities with very specific rules about data access
- Questions raised at STFC and within EGI about availability of this feature/posibility for couple of years

EGI CernVM-FS Service Developments

‘Protected’ CernVM-FS Repositories

- Work done within OSG on “Accessing Data Federations with CVMFS” (CHEP 2016 https://indico.cern.ch/event/505613/contributions/2230923/) added the possibility to introduce and manage authorization and authentication using security credentials such as X.509 proxy certificate
- We took the opportunity and looked to make use of this new feature by offering 'secure' CernVM-FS to interested user communities

EGI CernVM-FS Service Developments

'Protected' CernVM-FS Repositories

- Working prototype at RAL
  - Stratum-0 with mod_gridsite, https enabled
    - 'cvmfs_server publish' operation incorporates an authorization info file (DNs, VOMS roles)
    - access based on .gacl (Grid Access Control List) file in *<repo>/data/* directory that has to match the required DNs or VOMS roles
  - CVMFS client + cvmfs_helper package (enforces authz to the repository)
    - obviously 'root' can always see the namespace and the files in the client cache
  - Client connects directly to the Stratum-0
    - no Stratum-1 or squid in between - caching is not possible for HTTPS

EGI CernVM-FS Service Developments

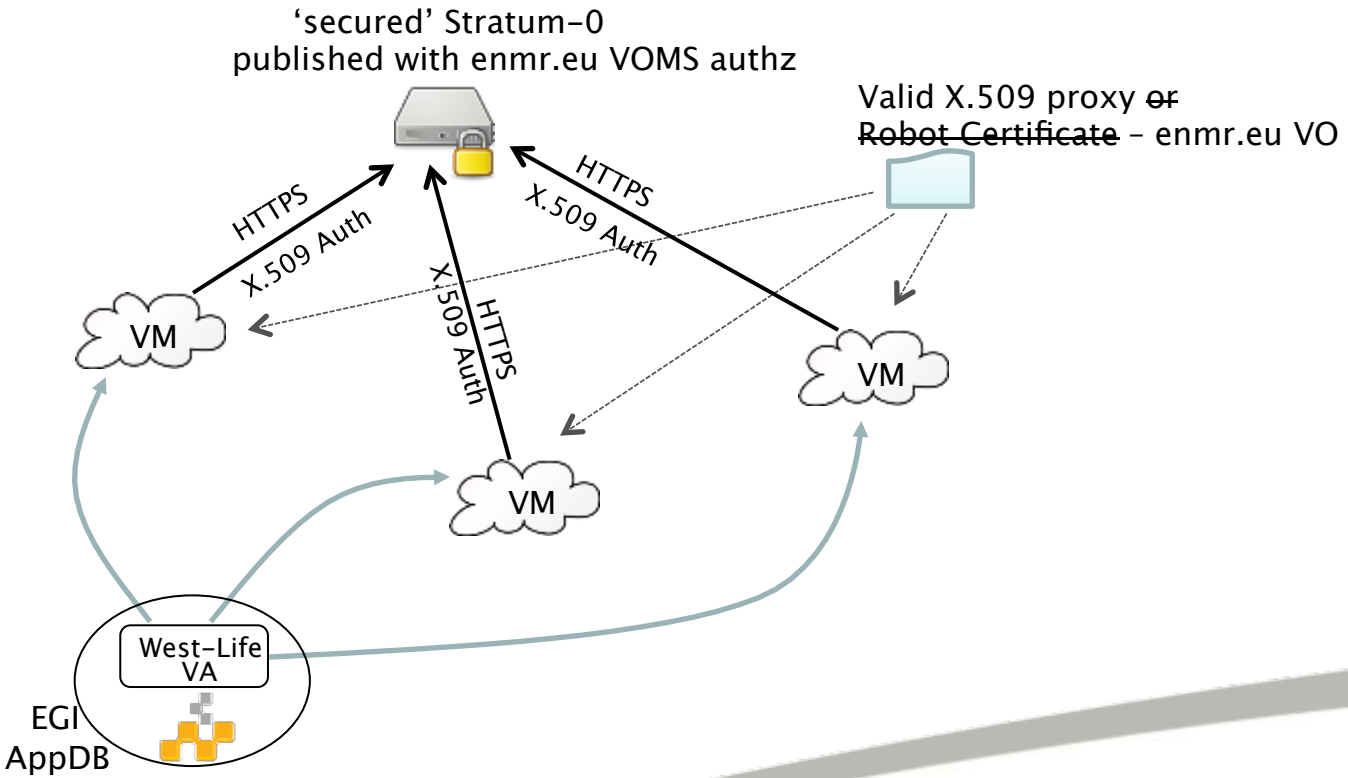
'Protected' CernVM-FS Repositories

- Cloud environment - good starting point for a use case
  - multiple VMs instantiated at various places and accessing the 'secure' repositories provided by a Stratum-0
  - a VM is not shared usually, it has a single user (which has root privileges as well)
  - the user downloads a certificate, creates a proxy and starts accessing the 'secure' repo
  - process can automated by using 'robot' certificates
    - and better downloading valid proxies
- Another possible use case
  - access from shared UIs, worker nodes

EGI CernVM-FS Service Developments

'Protected' CernVM-FS Repositories

- West-Life (H2020) project – 1st use case at STFC



Science & Technology Facilities Council

EGI CernVM-FS Service Developments

Configuration Repository

- Standard mountable CernVM-FS repo that resembles the directory structure of */etc/cvmfs*
  - set by *CVMFS_CONFIG_REPOSITORY=config-egi.egi.eu*
  - */cvmfs/config-egi.egi.eu/etc/cvmfs/...*
- Can be used to centrally maintain the public keys and configuration of repos that are not distributed with the static packages
- New *cvmfs-config-egi* RPM to replace *cvmfs-config-default* at EGI sites
  - similarly *cvmfs-config-osg* for OSG sites
  - 'orphan' sites still to use *cvmfs-config-default*

EGI CernVM-FS Service Developments

Configuration Repository

- All non-local configs to be moved there
- WN / VM at RAL (or EGI) with cvmfs-config-egi RPM
  - *egi.eu* configs will be installed in */etc/cvmfs*
  - *cern.ch, opensciencegrid.org* etc configs will be provided via */cvmfs/config-egi.egi.eu/etc/cvmfs*
  - easier to ban a domain, repository that has been corrupted or compromised

```
[root@lcg1765 ~]# echo $CVMFS_CONFIG_REPOSITORY
config-egi.egi.eu
[root@lcg1765 ~]# ls -l /cvmfs/config-egi.egi.eu/etc/cvmfs/
total 2
drwxr-xr-x 2 cvmfs cvmfs 88 Jan  3 13:40 config.d
drwxr-xr-x 2 cvmfs cvmfs 30 Jan  6 10:40 domain.d
drwxr-xr-x 5 cvmfs cvmfs 29 Jan  6 10:41 keys
```

EGI CernVM-FS Service Developments

Configuration Repository

- Support for the *africa-grid.org* CernVM-FS namespace
  - part of CODE-RADE project (South Africa) – COntinuous DElivery of Research Applications in a Distributed Environment
  - http://www.africa-grid.org/CODE-RADE/
  - Stratum-0 for *code-rade.africa-grid.org* repository in NGI_ZA
  - Stratum-1 in NGI_ZA, another one possibly in EGI
  - with configurations provisioned by */cvmfs/config-egi.egi.eu* any EGI CernVM-FS client will be able to access */cvmfs/code-rade.africa-grid.org/* filesystem
  - if same configurations available in */cvmfs/config-osg.opensciencegrid.org* then worldwide access to the repository

Science & Technology Facilities Council

EGI CernVM-FS Service Plans

Proxy Auto Configuration

- CernVM-FS supports Web Proxy Auto Discovery (WPAD) protocol and Proxy Auto Configuration (PAC)
- Proxy settings can be automatically gathered through WPAD and loaded from a PAC file
- Information about available proxies is maintained at CERN for WLCG and can also be used by EGI
- See “Web Proxy Auto Discovery for WLCG” (CHEP 2016 http://indico.cern.ch/event/505613/contributions/2230709/)

EGI CernVM-FS Service Developments

Proxy Auto Configuration

- Very useful when CernVM-FS is used within FedCloud
- A single Virtual Appliance instantiated at multiple places might not have access to the info about a local proxy (contextualization might provide it though…)
- Experience showed that VMs were usually accessing Stratum-1
- Work ongoing for a mechanism to discover the closest available squid to be integrated into CernVM-FS client configuration

Acknowledgements

- Stratum-1 administrators (Dennis van Dok, Di Qing, Felix Lee)
- CernVM-FS developers (Jakob Blomer, Dave Dykstra, Brian Bockelman, Rene Meusel)
- Colleagues at RAL

Science & Technology Facilities Council

- Thank you!
- Questions?