



CernVM-FS

Status and Latest Developments from RAL Tier-1 Perspective

Catalin Condurache

STFC RAL UK

- CernVM-FS ?
- CernVM-FS infrastructure and service @RAL
 - Topology, components and current status
- Recent developments and plans @RAL
 - ‘Confidential’ repositories
 - Large-Scale CVMFS
 - ‘Configuration Repository’ rollout
 - S3 + CEPH backend – Stratum-1

- CernVM-FS ?
- CernVM-FS infrastructure and service @RAL
 - Topology, components and current status
- Recent developments and plans @RAL
 - ‘Confidential’ repositories
 - Large-Scale CVMFS
 - ‘Configuration Repository’ rollout
 - S3 + CEPH backend – Stratum-1

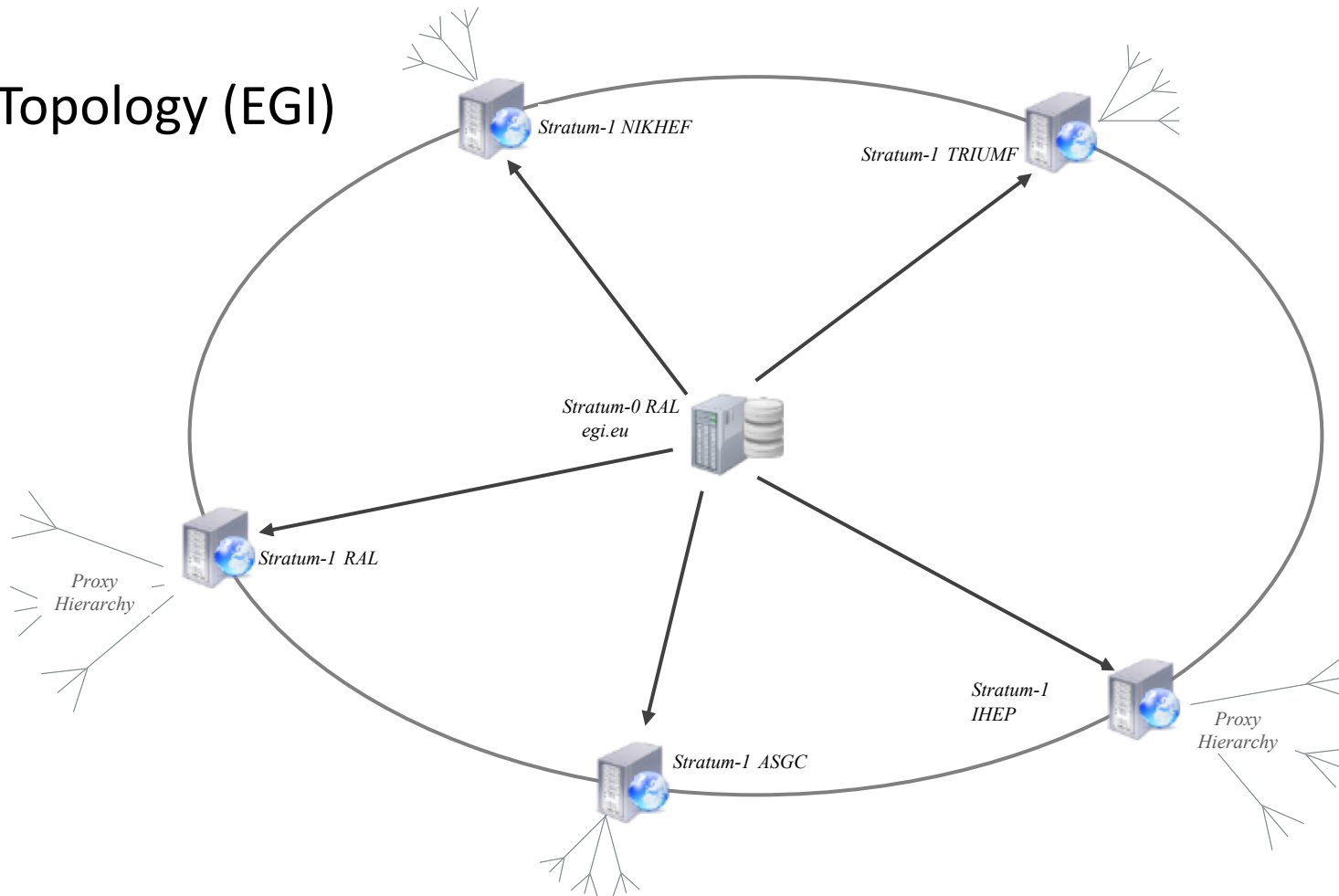
CernVM File System ?

- Read-only filesystem that provides a low-maintenance, reliable software distribution service for HEP and non-HEP communities
- Built using standard technologies (http, sqlite, fuse, squid)
- Data is aggressively cached, de-duplicated across different SW releases and transported via HTTP and proxy servers for the fastest possible retrieval of files
- Digitally signed repositories ensure data integrity

- CernVM-FS ?
- **CernVM-FS infrastructure and service @RAL**
 - Topology, components and current status
- Recent developments and plans @RAL
 - ‘Confidential’ repositories
 - Large-Scale CVMFS
 - ‘Configuration Repository’ rollout
 - S3 + CEPH backend – Stratum-1

CernVM-FS Infrastructure @RAL

- Topology (EGI)

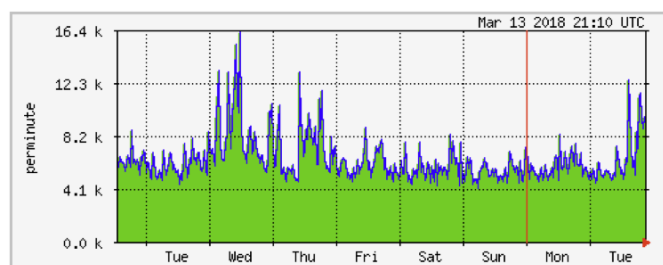


- Stratum-0 service (EGI, STFC)
 - Maintains and publishes the current state of the repositories
 - 32GB RAM, 12TB disk, 2x E5-2407 @2.20GHz
 - cvmfs-server v2.4.4
 - 35 repositories (egi.eu and gridpp.ac.uk) – 2.5 TB – 16.77 million files
 - Repository stats
 - largest number of files: mice (8×10^6), t2k (1.2×10^6), snoplus (1.0×10^6)
 - largest total file size: chipster (640 GB), mice (280 GB), t2k (100 GB)
 - largest revision number: pheno (722), t2k (215), auger (173)

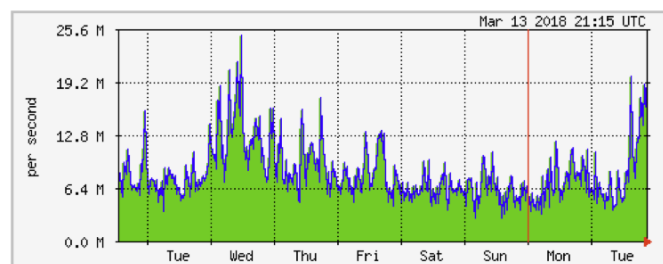
- Stratum-1 service (WLCG, EGI, STFC)
 - Part of the worldwide network of servers (RAL, NIKHEF, TRIUMF, ASGC, IHEP) replicating the *egi.eu* repositories
 - RAL - 2-node HA cluster (cvmfs-server v2.4.4)
 - each node – 64 GB RAM, 55 TB storage, 2xE5-2620 @2.4GHz
 - dual stack IPv4/IPv6 since September 2017
 - it replicates 78 repositories – total of 33 TB of replica
 - *egi.eu*, *gridpp.ac.uk* and *nikhef.nl* domains
 - also many *cern.ch*, *opensciencegrid.org*, *desy.de*, *africa-grid.org*, *ihep.ac.cn* and *in2p3.fr* repositories

CernVM-FS Infrastructure @RAL

- Stratum-1 service – plots, statistics



	Max	Average	Current
HTTP reqs	16.2 kreq/min	6324.0 req/min	9422.0 req/min
HTTP fetches	16.2 kreq/min	6324.0 req/min	9422.0 req/min



	Max	Average	Current
Total	24.7 MB/s	8277.0 kB/s	14.4 MB/s
Fetches	24.7 MB/s	8269.0 kB/s	14.4 MB/s

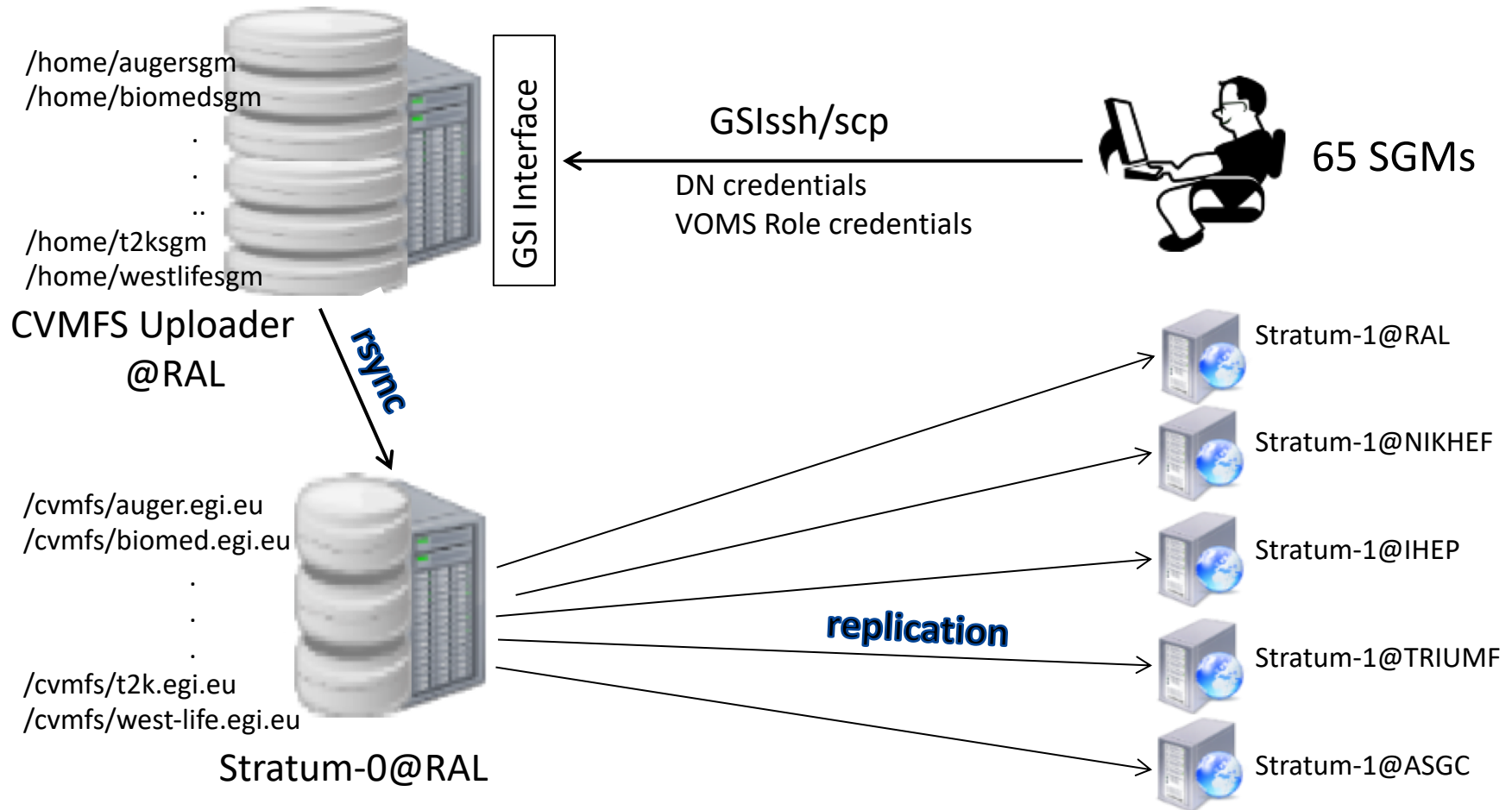
Day	Number of visits	Pages	Hits	Bandwidth
01 Mar 2018	1,384	7,351,021	7,351,021	1046.45 GB
02 Mar 2018	1,511	7,896,639	7,896,639	966.50 GB
03 Mar 2018	1,481	6,107,397	6,107,399	764.17 GB
04 Mar 2018	1,483	3,879,398	3,879,398	515.22 GB
05 Mar 2018	1,530	4,594,225	4,594,225	692.04 GB
06 Mar 2018	2,023	5,200,364	5,200,366	797.09 GB
07 Mar 2018	1,494	8,543,074	8,543,074	1221.23 GB
08 Mar 2018	1,752	6,561,422	6,561,423	882.27 GB
09 Mar 2018	1,725	4,818,352	4,818,352	772.66 GB
10 Mar 2018	1,941	4,834,857	4,834,860	605.06 GB
11 Mar 2018	2,045	4,327,273	4,327,273	580.88 GB
12 Mar 2018	1,125	2,434,125	2,434,125	322.34 GB

CernVM-FS Infrastructure @RAL

- CVMFS Uploader service @ RAL (EGI, STFC)
 - In-house implementation that provides upload area for *egi.eu* (and *gridpp.ac.uk*) repositories
 - Currently 1.7 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

CernVM-FS Infrastructure @RAL

- Repository uploading mechanism



- Operations Level Agreement for Stratum-0 service
 - Between STFC and EGI.eu
 - Annually renewed
 - Provisioning, daily running and availability of service
 - Service to be advertised through the EGI Service Catalog
- The EGI Staged Rollout
 - RAL is an early Adopter for cvmfs client and server, also frontier-squid
 - Followed by upload to UMD-4 repo
- Ongoing campaign at EGI sites to use cvmfs-config-egi RPM
 - As replacement for cvmfs-config-default at client level
 - Makes use of “Config Repository” (`CVMFS_CONFIG_REPOSITORY` param)

- CernVM-FS ?
- CernVM-FS infrastructure and service @RAL
 - Topology, components and current status
- **Recent developments and plans @RAL**
 - ‘Confidential’ repositories
 - Large-Scale CVMFS
 - ‘Configuration Repository’ rollout
 - S3 + CEPH backend – Stratum-1

Developments – ‘Confidential’ 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 specific rules on data access
- Questions raised at STFC and within EGI about availability of this feature/possibility in the last 3 years

Developments – ‘Confidential’ CernVM-FS Repositories

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

Developments – ‘Confidential’ 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_x509_helper* package (enforces authz to the repository)
 - also some *globus-** packages need installed!
 - 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

Developments – ‘Confidential’ CernVM-FS Repositories

- First proof of concept – 15 December 2017
 - Work with Tom Whyntie (former cernatschool.org), now working on UK Biobank project (www.ukbiobank.ac.uk)
 - `/cvmfs/gridmi.egi.eu`
 - 2 use cases
 - to host FMRIB Software Library (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki>) and make access available to a user group with well-defined access rights via a CVMFS repository
 - distribute securely proprietary SW and/or data (i.e. gradient distortion correction information) used by MRI scanners in order to process and analyse scan information



Developments – ‘Confidential’ CernVM-FS Repositories

- Also good progress with CCP4 collaboration
 - To host CCP4 software (<http://www.ccp4.ac.uk>)
 - (probably) `/cvmfs/ccp4.egi.eu`
- Exploratory discussions with ELI-NP
 - Extreme Light Infrastructure – Nuclear Physics Project (<http://www.eli-np.ro/>)
 - `/cvmfs/eli-np.egi.eu`

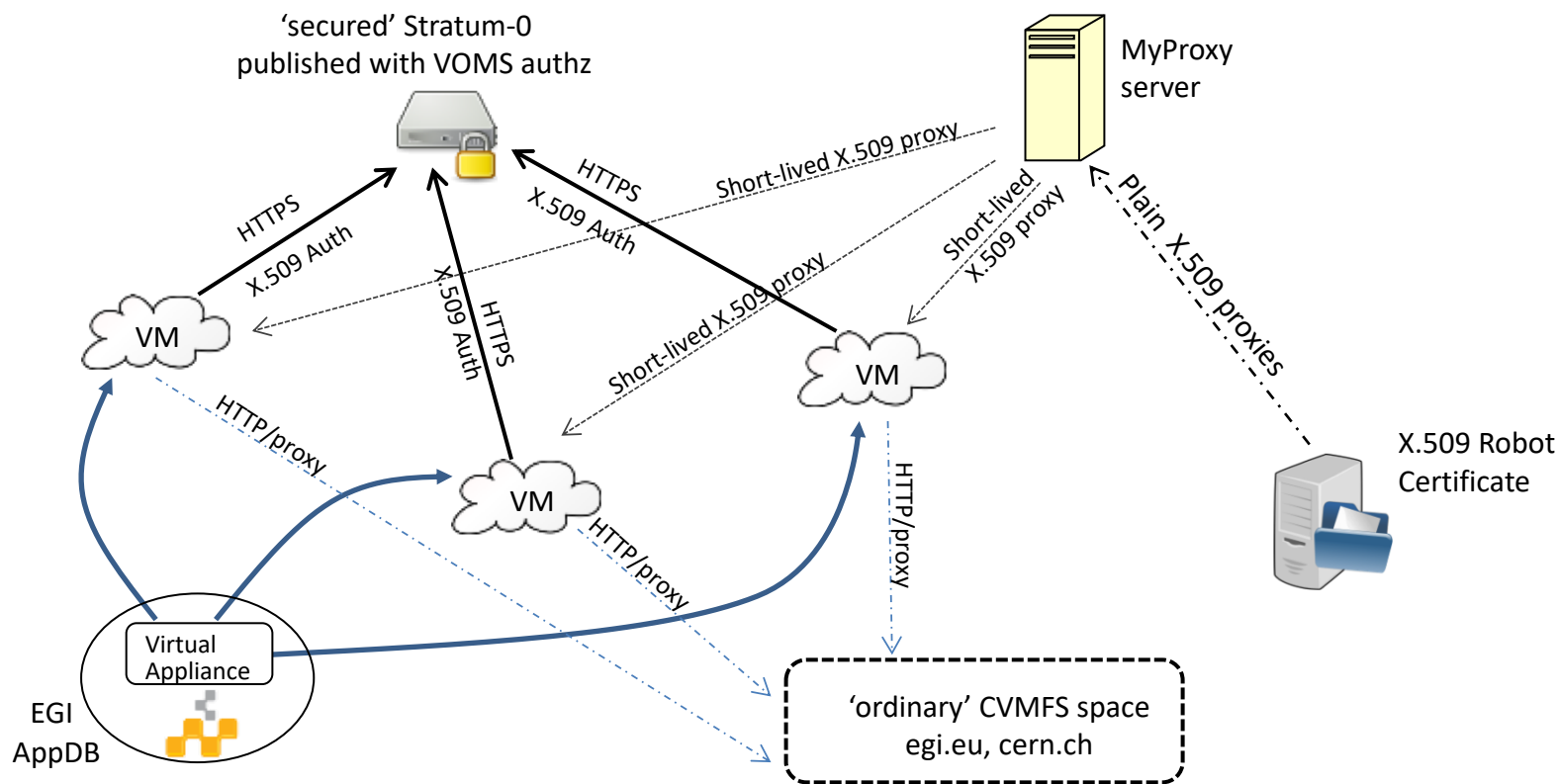


Plans – ‘Confidential’ CernVM-FS Repositories

- 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 be automated by using ‘robot’ certificates
 - and better by downloading valid proxies

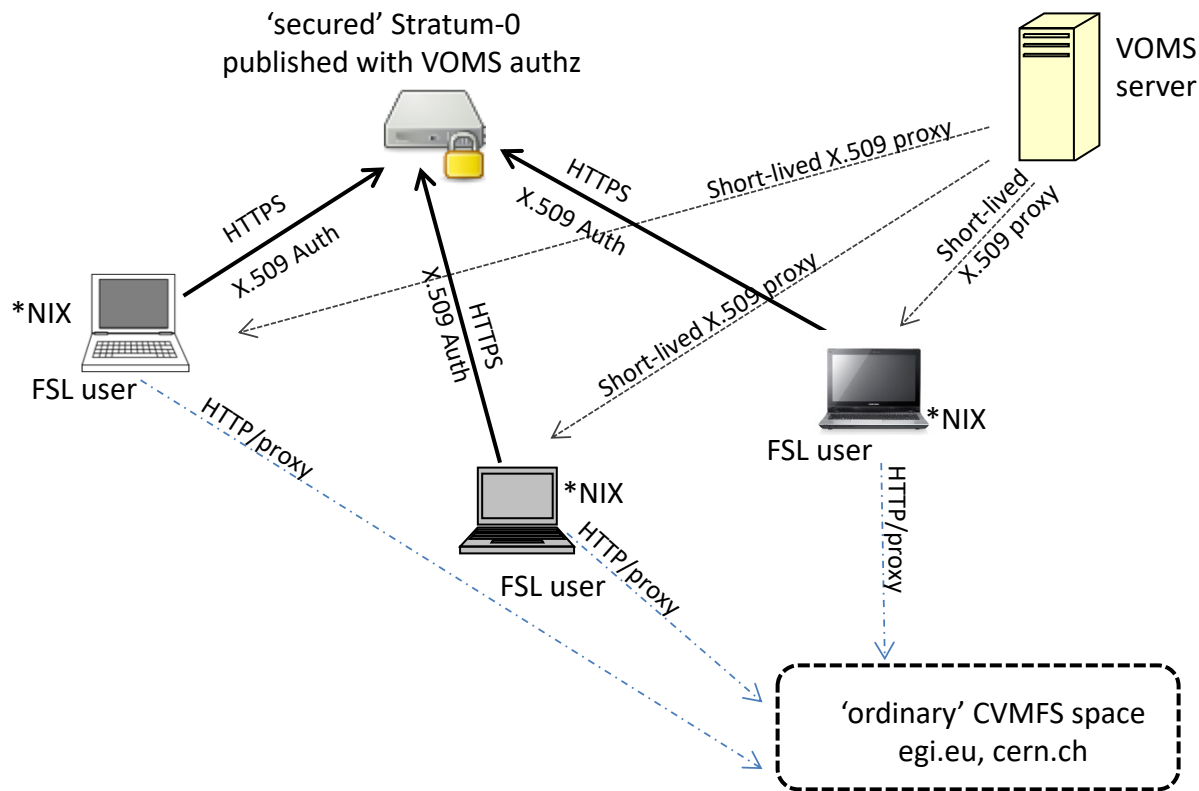
Plans – ‘Confidential’ CernVM-FS Repositories

- EGI FedCloud use case



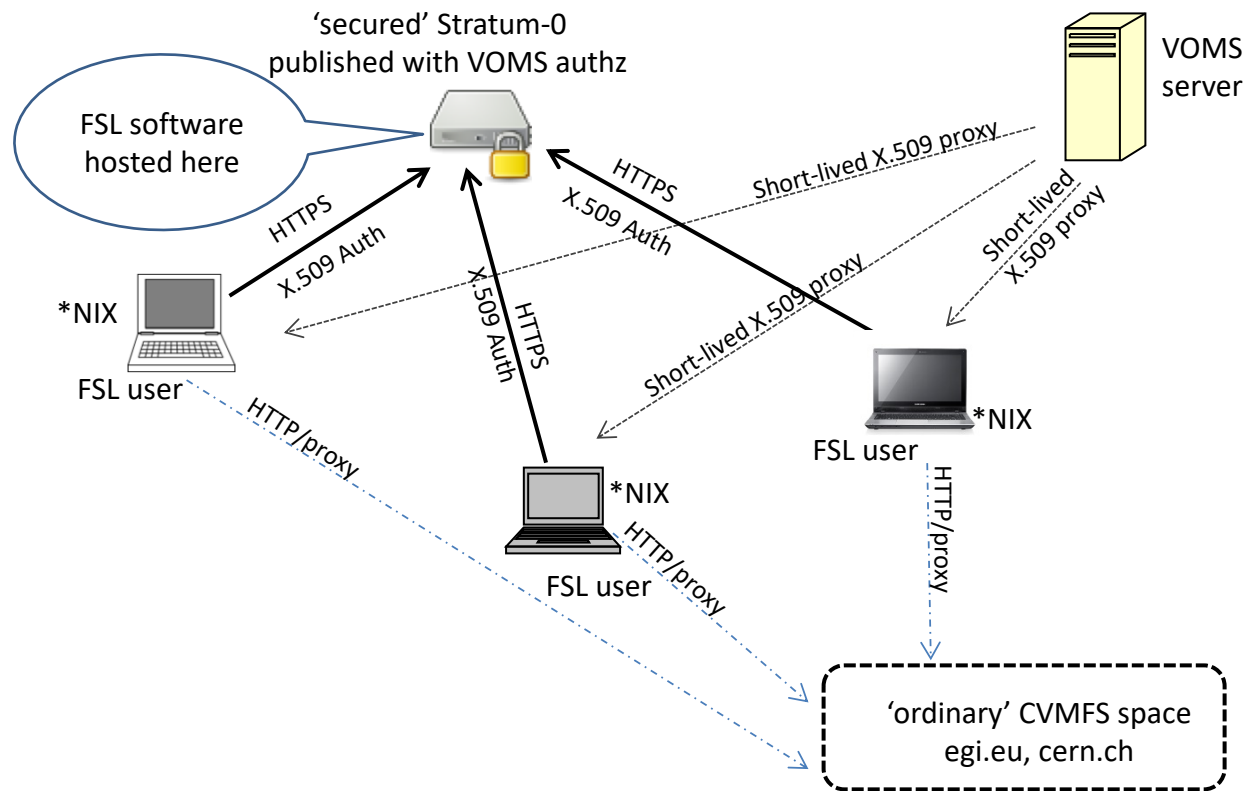
Plans – ‘Confidential’ CernVM-FS Repositories

- FSL use case



Plans – ‘Confidential’ CernVM-FS Repositories

- FSL use case



Developments – Large-Scale CVMFS

- CVMFS primarily developed for distributing large software stacks (GB)
- Colleagues from OSG developed extensions to CVMFS software that permit distribution of large, non-public datasets (TB to PB)
- Data is not stored within the repository - only checksums and the catalogs
 - Data is externally stored – Dynafed interface to CEPH storage
 - CVMFS clients are configured to be pointed at a non-CVMFS data
 - i.e. external XROOT storage can be referred by a CVMFS repository and accessed in a POSIX-like manner ('ls', 'cp' etc)
- Work in early stage at RAL (for LIGO – incl X.509 read-access authorization)

Developments – ‘Config Repository’ Rollout

- 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 (also can ban compromised repos)
- Ongoing campaign at EGI sites to use *cvmfs-config-egi* RPM
 - As replacement for *cvmfs-config-default* at client level

Developments – ‘Config Repository’ Rollout

- 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, also replica at RAL
 - With configurations provisioned by */cvmfs/config-egi.egi.eu* any EGI CernVM-FS client is 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

Plans - S3 + CEPH backend on Stratum-1

- Use S3 interface to CEPH backend storage for Stratum-1
 - As an upgrade plan when disk space an issue on existing HA 2-node cluster
 - Initial tests successful
 - What about HA topology?
 - Some investigations needed

Acknowledgements

- CernVM-FS developers (Jakob Blomer, Dave Dykstra, Brian Bockelman, Radu Popescu, Derek Wetzel)
- Colleagues at RAL
- Operators of other Stratum-1s

- Thank you!
- Questions?