Science & Technology Facilities Council

Recent Developments and Plans on CernVM-FS at RAL

Catalin Condurache

Rutherford Appleton Laboratory

STFC UK Research and Innovation

ISGC, Taipei, April 2019

Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - ‘confidential’ repositories
  - Repository Gateway and Release Managers

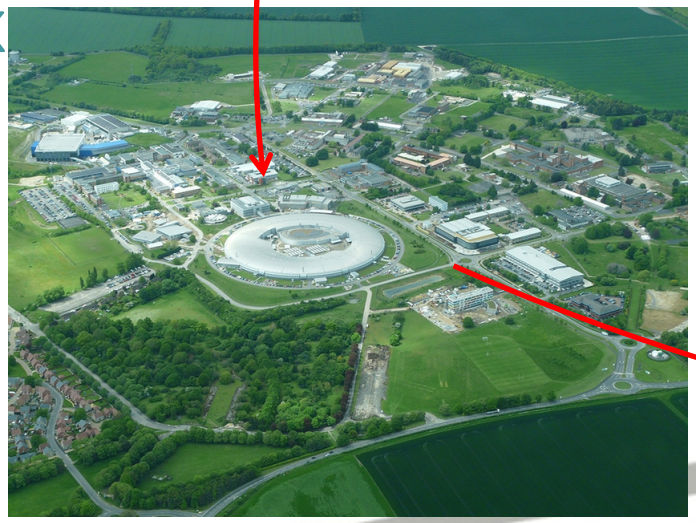
Science & Technology Facilities Council

Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - 'confidential' repositories
  - Repository Gateway and Release Managers

UK LHC Tier-1 @RAL

- Hosted and run by Scientific Computing Dept at Rutherford Appleton Laboratory
- RAL part of Science and Technology Facilities Council – part of UK Research and Innovation
- 15 miles south of Oxford on Harwell Campus



Science & Technology Facilities Council

Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - 'confidential' repositories
  - Repository Gateway and Release Managers

Science & Technology Facilities Council

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

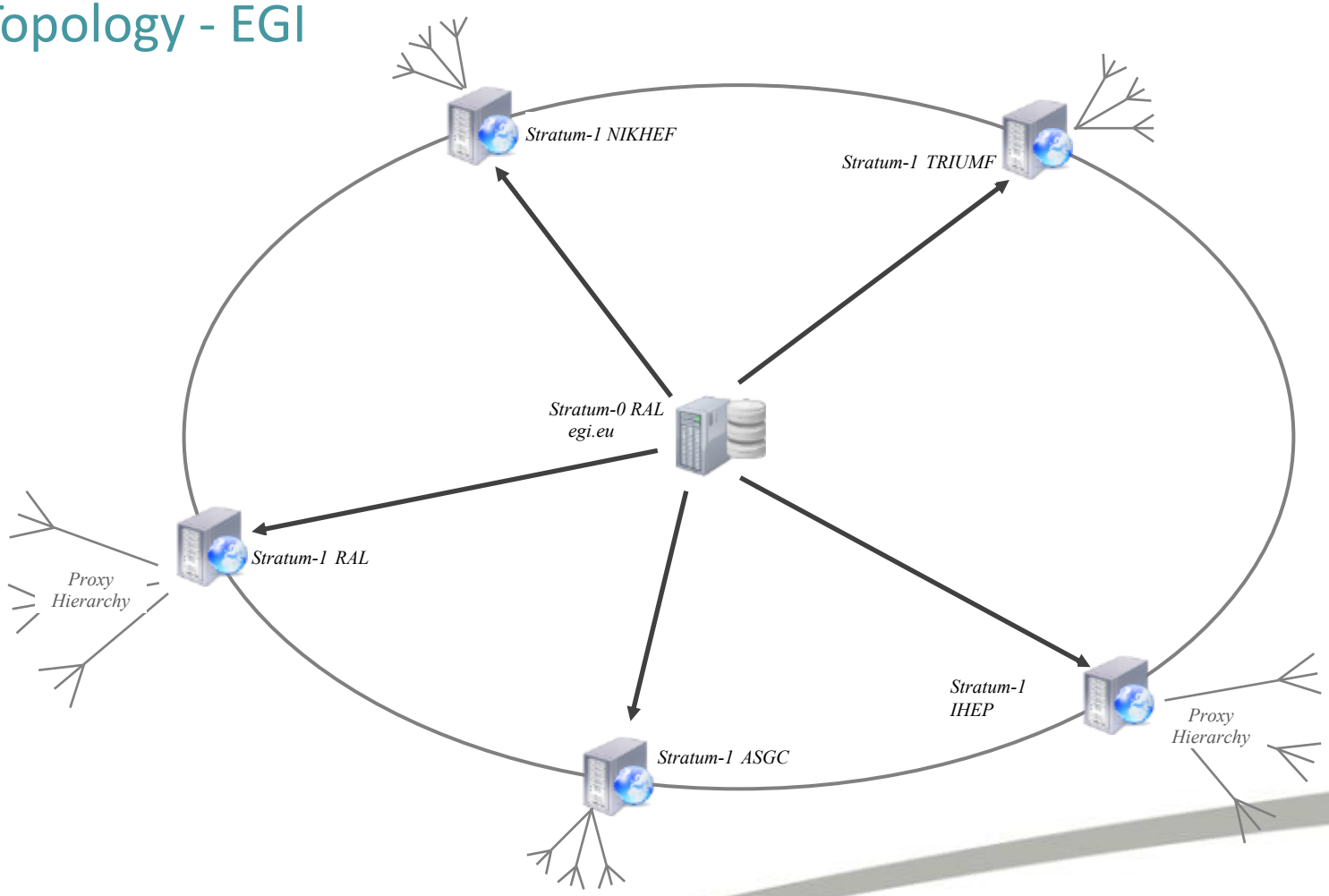
Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - 'confidential' repositories
  - Repository Gateway and Release Managers

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

Topology - EGI



Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

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 v2.5.2
- 35 repositories (egi.eu and gridpp.ac.uk) – 2.7 TB – 24.4 million files
- Repository stats
  - largest number of files: mice (9.2×10^6), facilities (2.3×10^6), t2k (1.9×10^6)
  - largest total file size: chipster (644 GB), mice (346 GB), t2k (151 GB)
  - largest revision number: pheno (735), t2k (256), auger (178)

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

Stratum-1 service (WLCG, EGI, STFC)

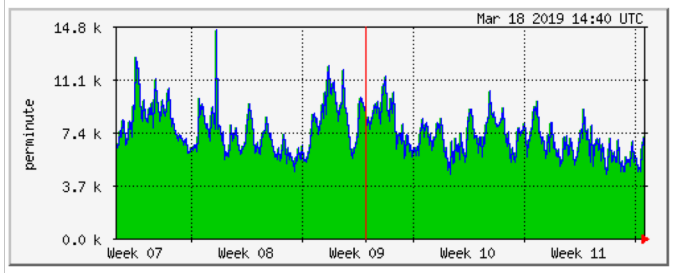
- Part of the worldwide network of servers (RAL, NIKHEF, TRIUMF, ASGC, IHEP) replicating the egi.eu repositories
- RAL - 2-node HA cluster (cvmfs-server v2.4.4)
- each node – 64 GB RAM, 55 TB storage, 2xE5-2620 @2.4GHz
- dual stack IPv4/IPv6 since September 2017
- it replicates 92 repositories – total 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

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

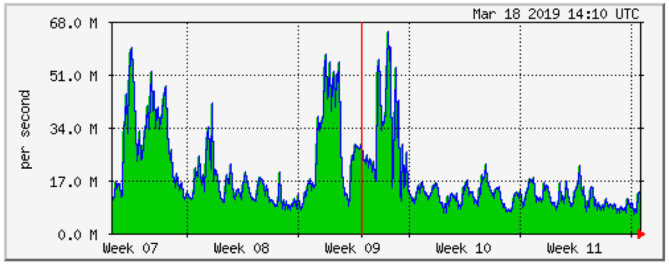
Stratum-1 service – plots, stats

`Monthly' Graph (2 Hour Average)



	Max	Average	Current
HTTP reqs	14.5 kreq/min	7276.0 req/min	7106.0 req/min
HTTP fetches	14.5 kreq/min	7258.0 req/min	7089.0 req/min

`Monthly' Graph (2 Hour Average)



	Max	Average	Current
Total	64.2 MB/s	18.3 MB/s	12.6 MB/s
Fetches	64.2 MB/s	18.3 MB/s	12.6 MB/s

Day	Number of visits	Pages	Hits	Bandwidth
01 Mar 2019	442	9,368,815	9,368,820	2115.91 GB
02 Mar 2019	491	9,877,127	9,877,128	3664.11 GB
03 Mar 2019	424	6,159,908	6,159,914	2163.84 GB
04 Mar 2019	473	6,847,322	6,847,327	1152.84 GB
05 Mar 2019	547	6,897,867	6,897,876	1154.80 GB
06 Mar 2019	570	5,069,423	5,069,424	936.55 GB
07 Mar 2019	456	7,028,231	7,028,234	1103.76 GB
08 Mar 2019	456	7,779,084	7,779,087	1299.55 GB
09 Mar 2019	389	7,899,510	7,899,511	1055.11 GB
10 Mar 2019	360	6,238,252	6,238,254	869.23 GB
11 Mar 2019	453	7,928,587	7,928,590	1196.16 GB
12 Mar 2019	454	6,543,791	6,543,794	1109.86 GB
13 Mar 2019	548	5,620,998	5,620,999	1013.65 GB
14 Mar 2019	458	6,627,068	6,627,069	1213.42 GB
15 Mar 2019	571	5,739,162	5,739,162	896.55 GB
16 Mar 2019	469	4,815,913	4,815,913	812.83 GB
17 Mar 2019	291	2,247,836	2,247,836	377.77 GB

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

Uploader service (EGI, STFC)

- In-house implementation that provides upload area for *egi.eu* (and *gridpp.ac.uk*) repositories
- Currently 2.0 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

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

Uploader service (EGI, STFC)

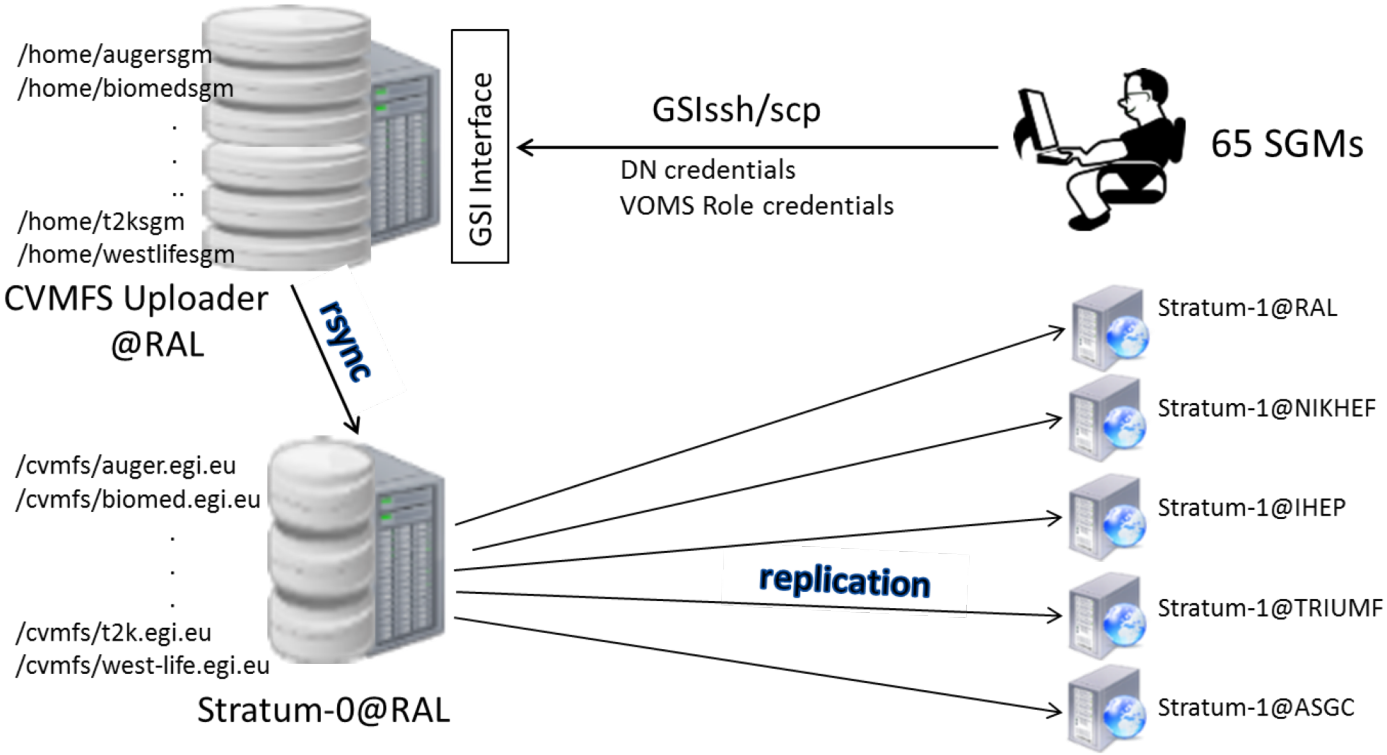
- In-house implementation that provides upload area for *egi.eu* (and *gridpp.ac.uk*) repositories
- Currently 2.0 TB – repo master copies
- GSI-OpenSSH interface (gsissh, gsiscp, gsisftp)
  - similar to standard OpenSSH tools with added ability to perform X.509 proxy credential authentication and delegation
  - DN based access, also VOMS Role possible
- *rsync* mechanism between Stratum-0 and Uploader

- **Soon to be changed – see 'Plans' slides**

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

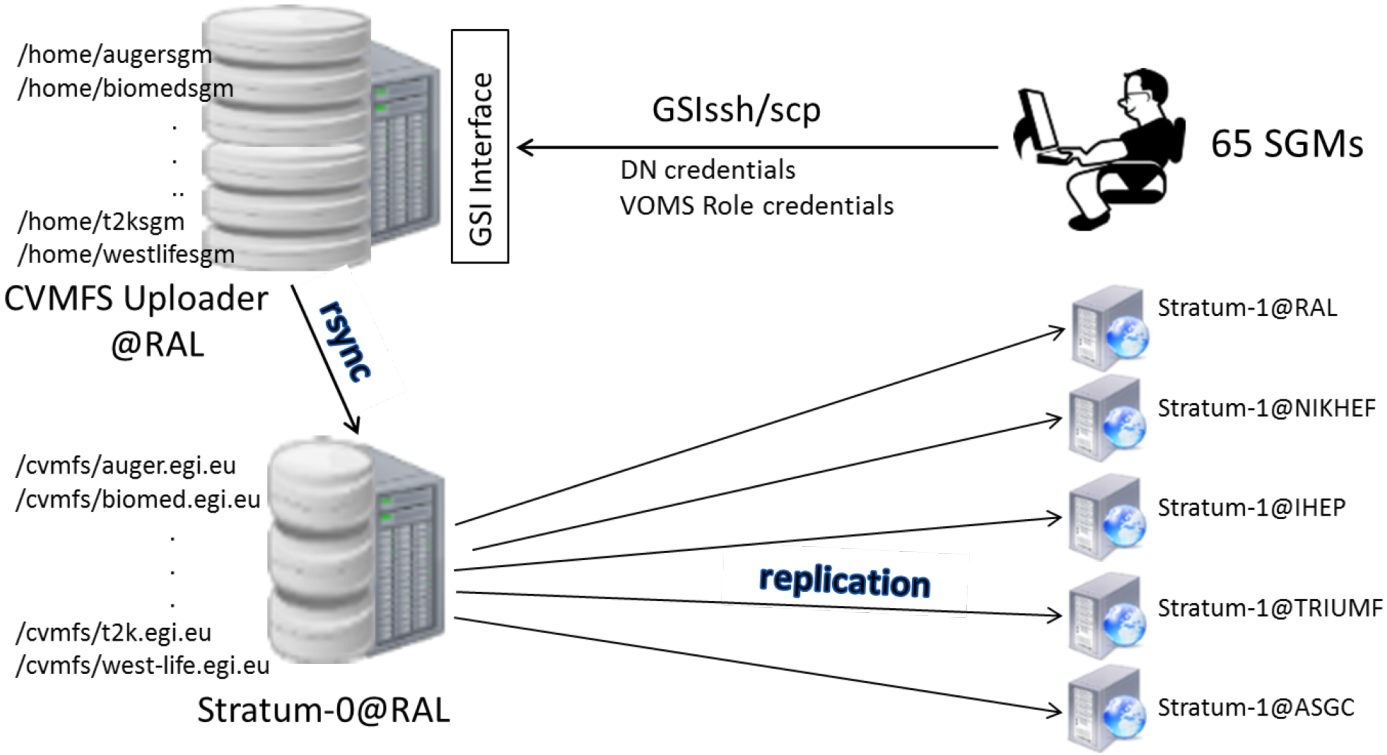
Repository uploading mechanism



Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

Repository uploading mechanism



Soon to be changed – see 'Plans' slides

Science & Technology Facilities Council

CernVM-FS Infrastructure @RAL

Squid service (UK-wide)

- Battery of 5 frontier-squid nodes under *cvmfs-squid.gridpp.rl.ac.uk* alias
- Currently frontier-squid v3.5.28 – plans for frontier-squid-4
- Dual stack IPv4/6
- Serves requests from **.ac.uk* hosts
- Also used by ATLAS and CMS to access the Frontier launchpads

- *atlas-squid* and *cms-squid* aliases

Science & Technology Facilities Council

Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - ‘confidential’ repositories
  - Repository Gateway and Release Managers

Science & Technology Facilities Council

EGI CernVM-FS service

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*

Operations Level Agreement for Stratum-0

- Between STFC and EGI.eu
- Provisioning, daily running and availability of service
- Service advertised through the EGI Service Catalogue

The EGI Staged Rollout

- RAL is an early Adopter for cvmfs client and server, also frontier-squid
- Once reports are accepted, upload to UMD-4 repo

Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - ‘confidential’ repositories
  - Repository Gateway and Release Managers

Science & Technology Facilities Council

Recent Developments and Plans

‘Confidential’ CernVM-FS repositories

- Repositories natively designed to be public with non-authenticated access – minimal info needed: 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 recent years
- 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

‘Confidential’ CernVM-FS Repositories

Working prototype at RAL

- Stratum-0 with *mod_gridsite*, *https* enabled
  - ‘*cvmfs_server publish*’ operation incorporates an authz info file
  - Access based on .gacl (Grid Access Control List) file in *<repo>/data/* directory that has to match the approved DNs or VOMS roles
- CVMFS client + *cvmfs_x509_helper* to enforce authz to the repository
  - Also some globus-* packages need installed!
  - *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 not possible for HTTPS

Science & Technology Facilities Council

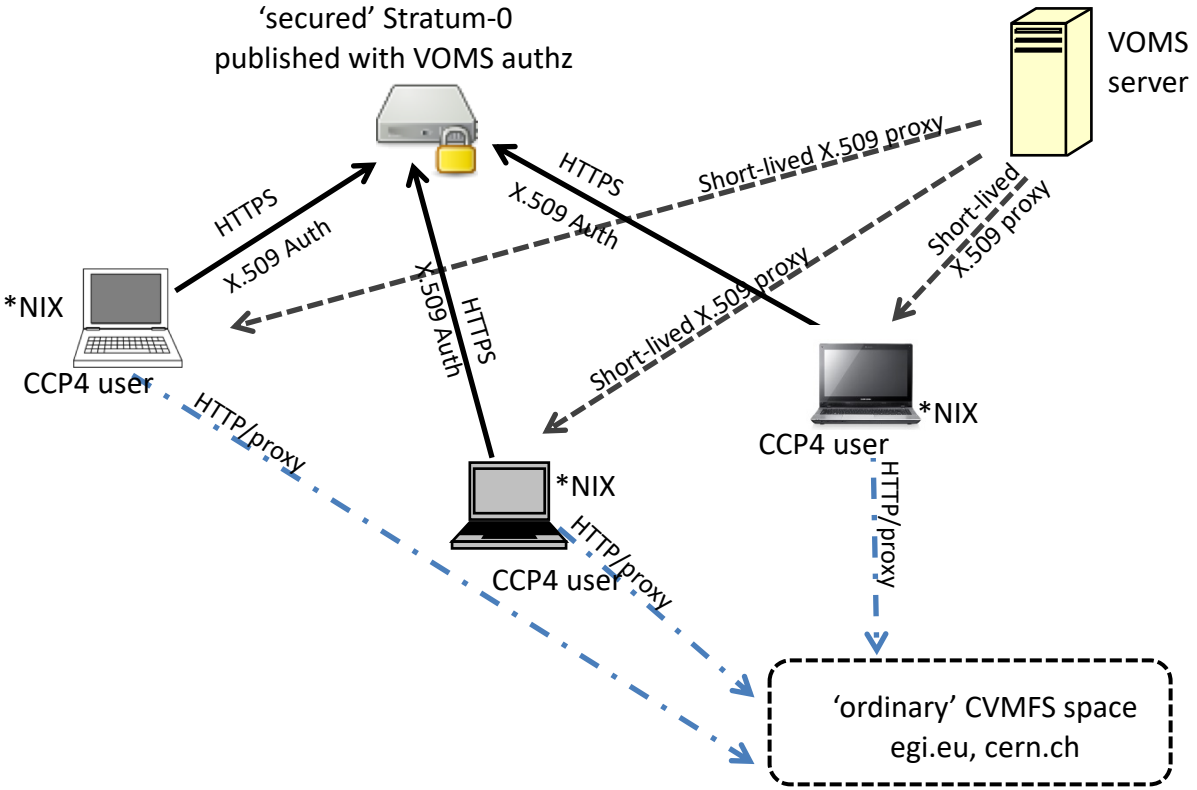
‘Confidential’ CernVM-FS Repositories

Working prototype at RAL

- Repository for Collaborative Computational Project No.4
  - www.ccp4.ac.uk
  - To host CCP4 Software for Macromolecular X-Ray Crystallography
  - (probably) /cvmfs/ccp4-sw.stfc.uk
- Also exploratory discussions with ELI-NP
  - Extreme Light Infrastructure – Nuclear Physics Project www.eli-np.ro

‘Confidential’ CernVM-FS Repositories

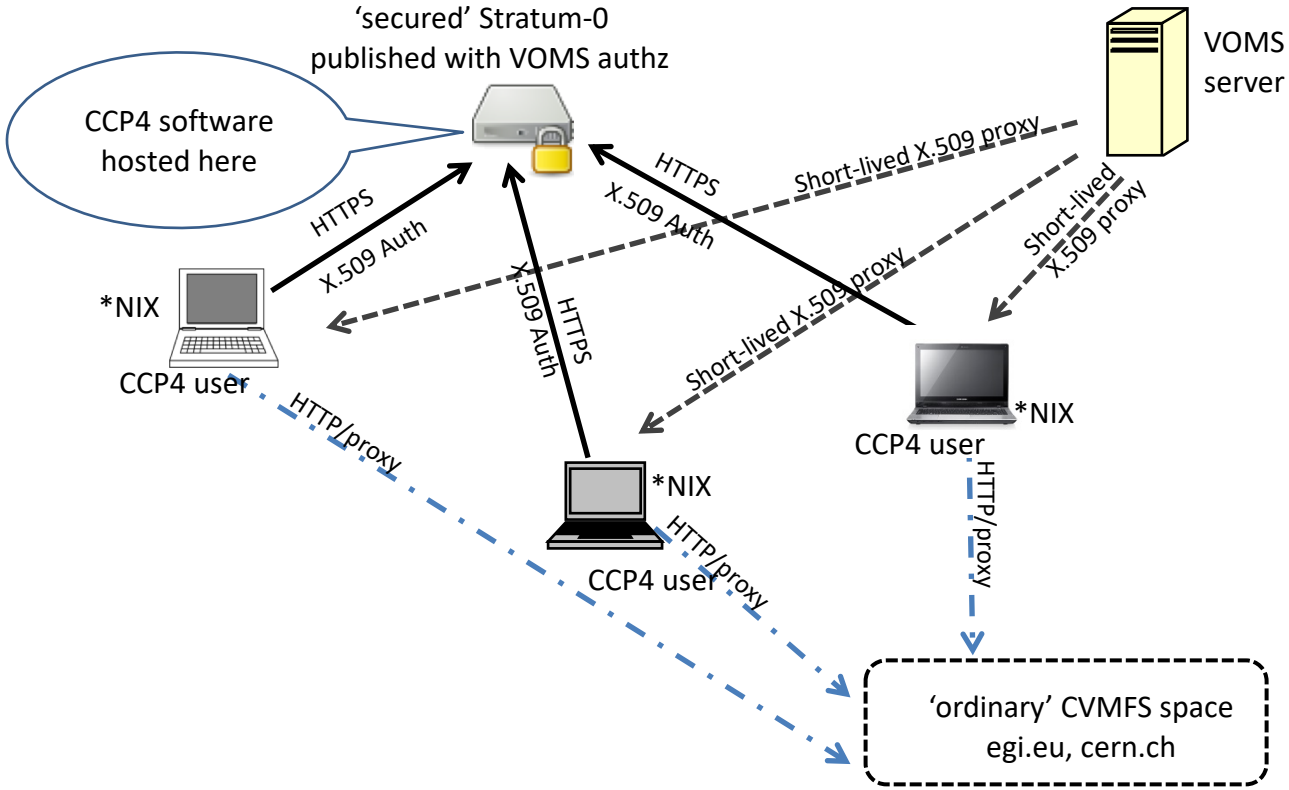
CCP4 use case



Science & Technology Facilities Council

‘Confidential’ CernVM-FS Repositories

CCP4 use case



Science & Technology Facilities Council

Outline

- UK LHC Tier-1 @RAL
- CernVM-FS ?
- CernVM-FS infrastructure @RAL
  - Topology, components and current status
- EGI CernVM-FS service
  - Operational Procedures (PROC20, PROC22)
  - OLA
  - Staged Rollout Reports (cvmfs, frontier-squid)
- Recent developments and plans
  - 'confidential' repositories
  - Repository Gateway and Release Managers

Science & Technology Facilities Council

Plans

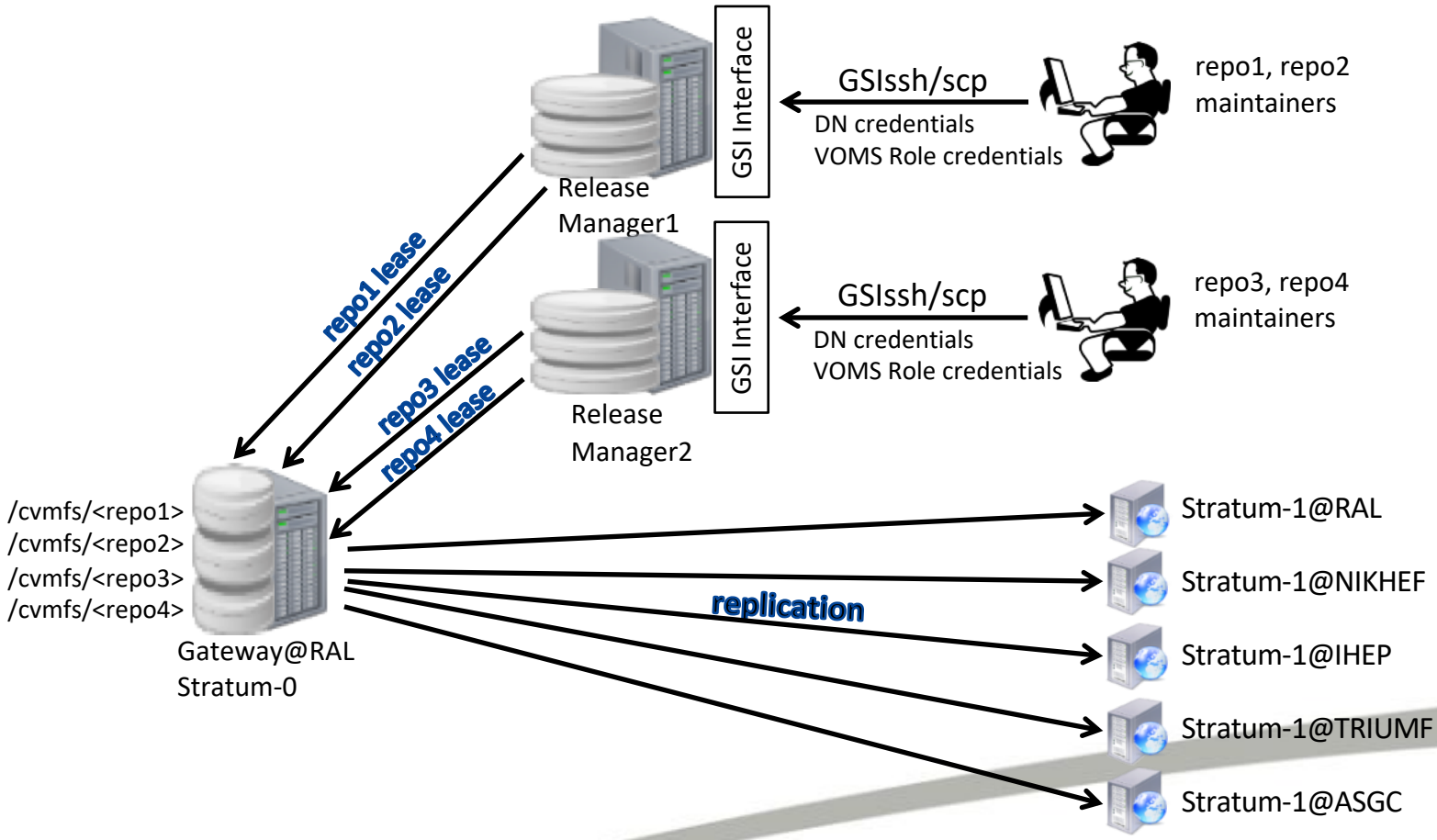
Repository Gateway and Release Managers

- Replacement for the existing Uploader service
- Makes use of recently introduced CernVM-FS features
- Gateway (GW)
  - Machine running an instance of the CVMFS repository gateway – it has access to the authoritative storage of the managed repositories
  - *cvmfs*, *cvmfs-server*, *cvmfs-gateway* packages
- Release manager (RM)
  - Machine running CVMFS server tools – can request leases from a GW and publish changes to different repositories
  - *cvmfs*, *cvmfs-server* packages

Science & Technology Facilities Council

Plans

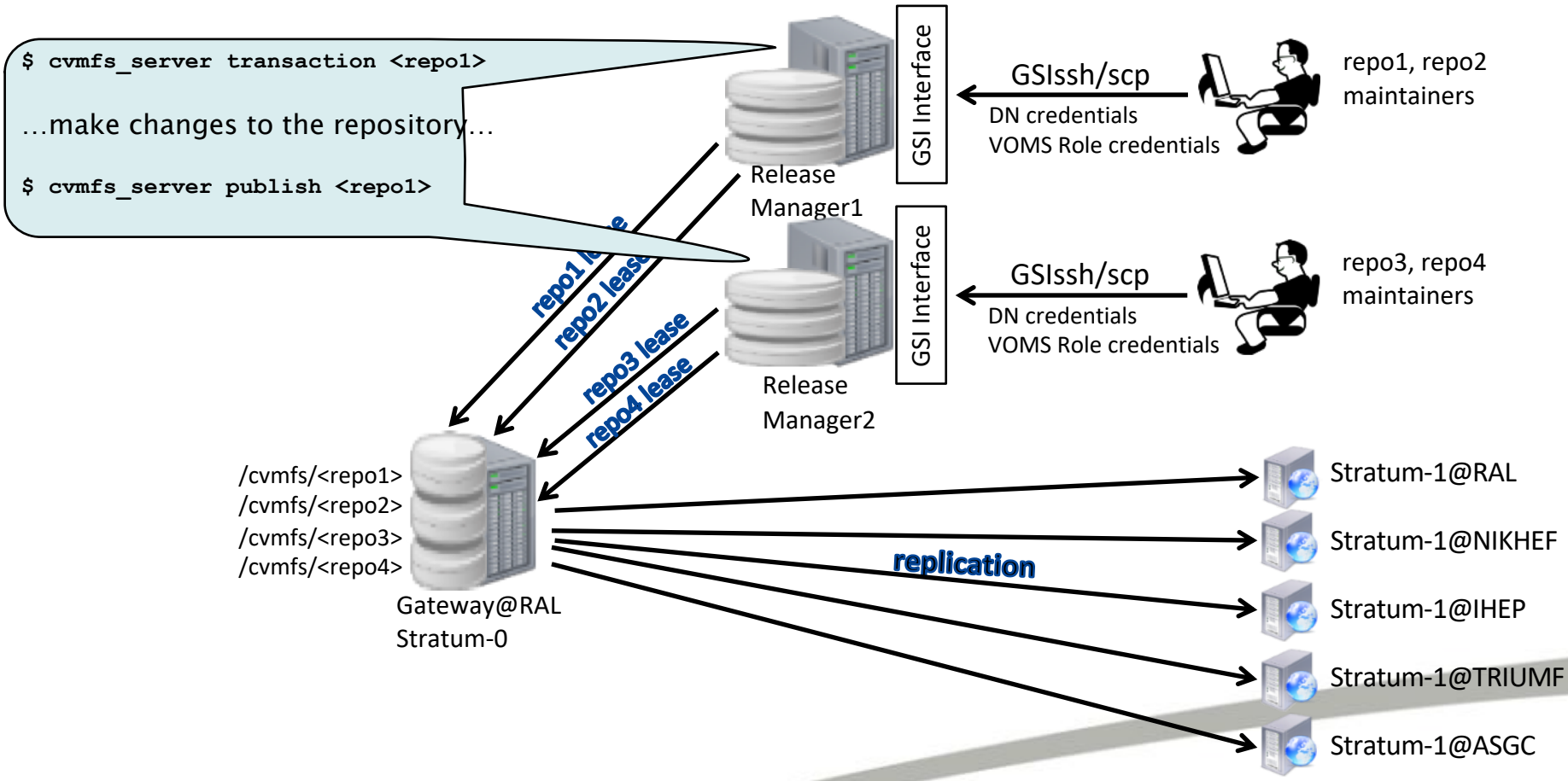
Repository Gateway and Release Managers



Science & Technology Facilities Council

Plans

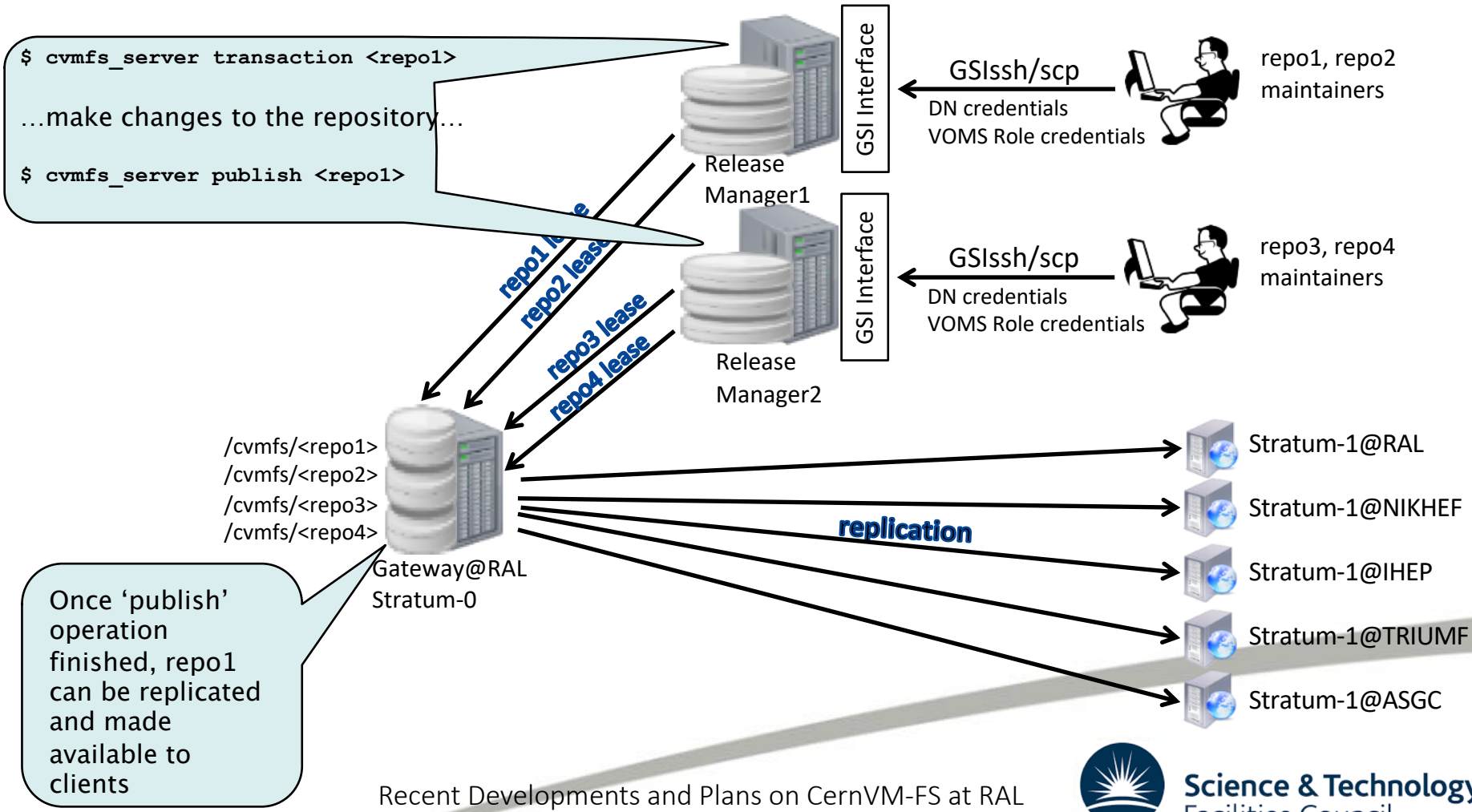
Repository Gateway and Release Managers



Science & Technology Facilities Council

Plans

Repository Gateway and Release Managers



Science & Technology Facilities Council

Plans

(GW + RM) vs Uploader service

- With the Uploader…
  - 60min+ delay between last repo maintenance operation and publishing at Stratum-0 level (*rsync* run by hourly cronjob)
  - Repo maintainers do not have to use cvmfs related commands
- With the GW + RM…
  - Repo maintainers need to use the *cvmfs_server {transaction,publish,abort}* commands
  - Repo is published and ready for replication once *cvmfs_server publish <repo>* is executed
- Transition period
  - Existing repos need to be gradually migrated from old Uploader (+ Stratum-0) mechanism to new GW + RM configuration
  - Reverse proxy configuration for each migrated repo to be used on Stratum-0 server - migration will be transparent at replication level
  - Initially *config-egi.egi.eu* and **.gridpp.ac.uk* repos

Science & Technology Facilities Council

Thank you!
Questions?