Feedback from EGI

EGI CernVM-FS Infrastructure, Status and Plans

Catalin Condurache

STFC RAL Tier-1

CernVM Users Workshop, CERN, 5-6 March 2015

Outline

- Brief history
- EGI CernVM-FS Task Force
- EGI CernVM-FS infrastructure
- EGI CernVM-FS status, plans, deployment issues
- Software installation mechanism at RAL

Brief History

- Following proven success in experiments software distribution using CernVM-FS to WLCG sites…
- …RAL Tier-1 started non-LHC Stratum-0 service
  - September 2012
  - '*gridpp.ac.uk*' CernVM-FS domain
- Supported by GridPP UK project
- Initially for UK VOs, then extended

EGI CernVM-FS Task Force

- EGI initiative to establish a CernVM-FS infrastructure that allows EGI VOs to use it as a standard method of distribution of their software at grid sites

- Kick-off meeting – August 2013

- Members: reps from NGIs and other collaborations, VO SGMs, system admins at sites

EGI CernVM-FS Task Force

- To promote the use of CernVM-FS technology among user communities

- To create a network of sites providing CernVM-FS services (Stratum-0, Stratum-1, Squid)

- To encourage cooperation with other organizations (OSG, WLCG, regional grid infrastructures) by cross-replicating repositories for VOs supported by multiple collaborations

EGI CernVM-FS Infrastructure

- The model follows the CernVM-FS for WLCG topology
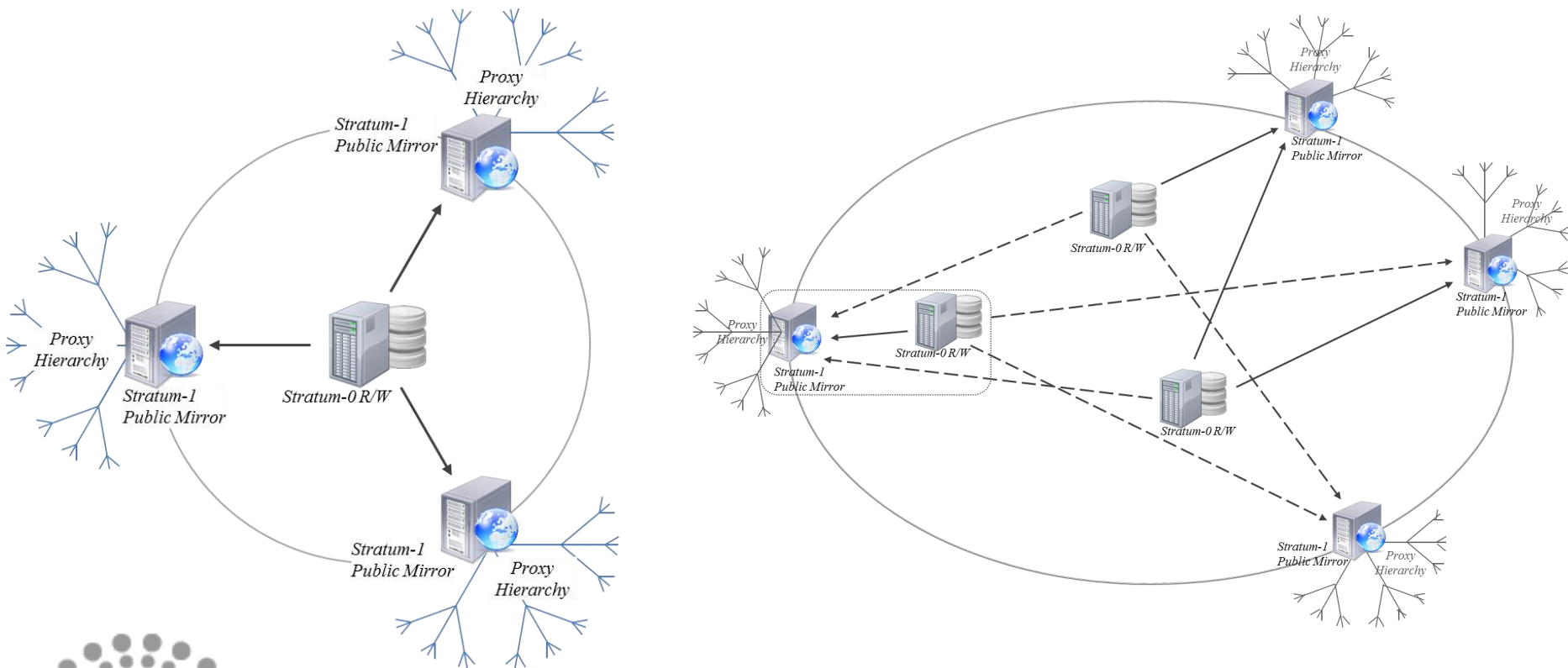
EGI CernVM-FS Infrastructure

- *'relaxed'* topology

- Stratum-0 are disjoint and represent the source repositories where software is installed by VOs

- Stratum-0, Stratum-1 can be geographically co-located or not

- Stratum-1 can replicate an entire Stratum-0 or can only partially replicate – the *'relaxed'* model

- It partially makes use of the existent hierarchy of proxy servers used for LHC software distribution
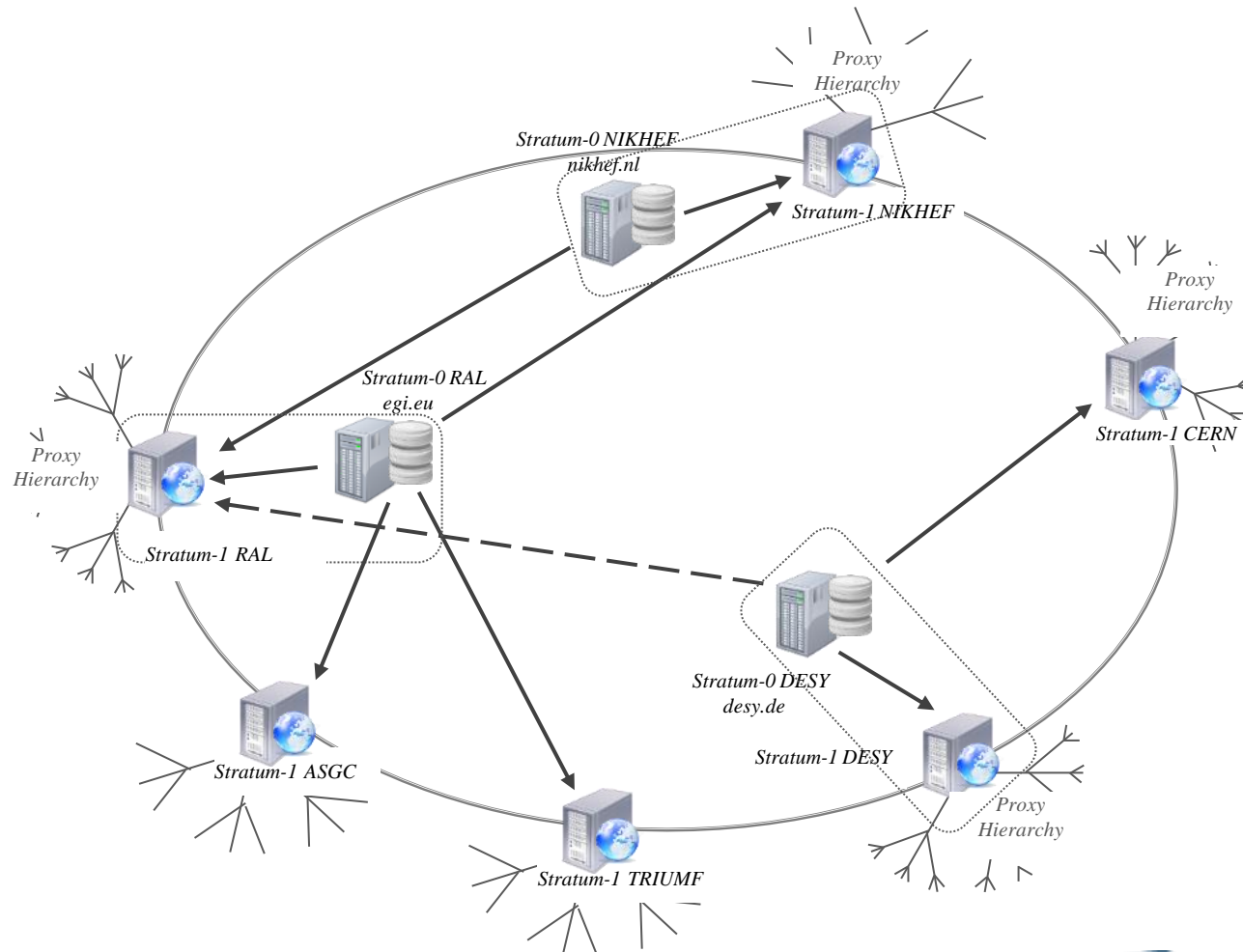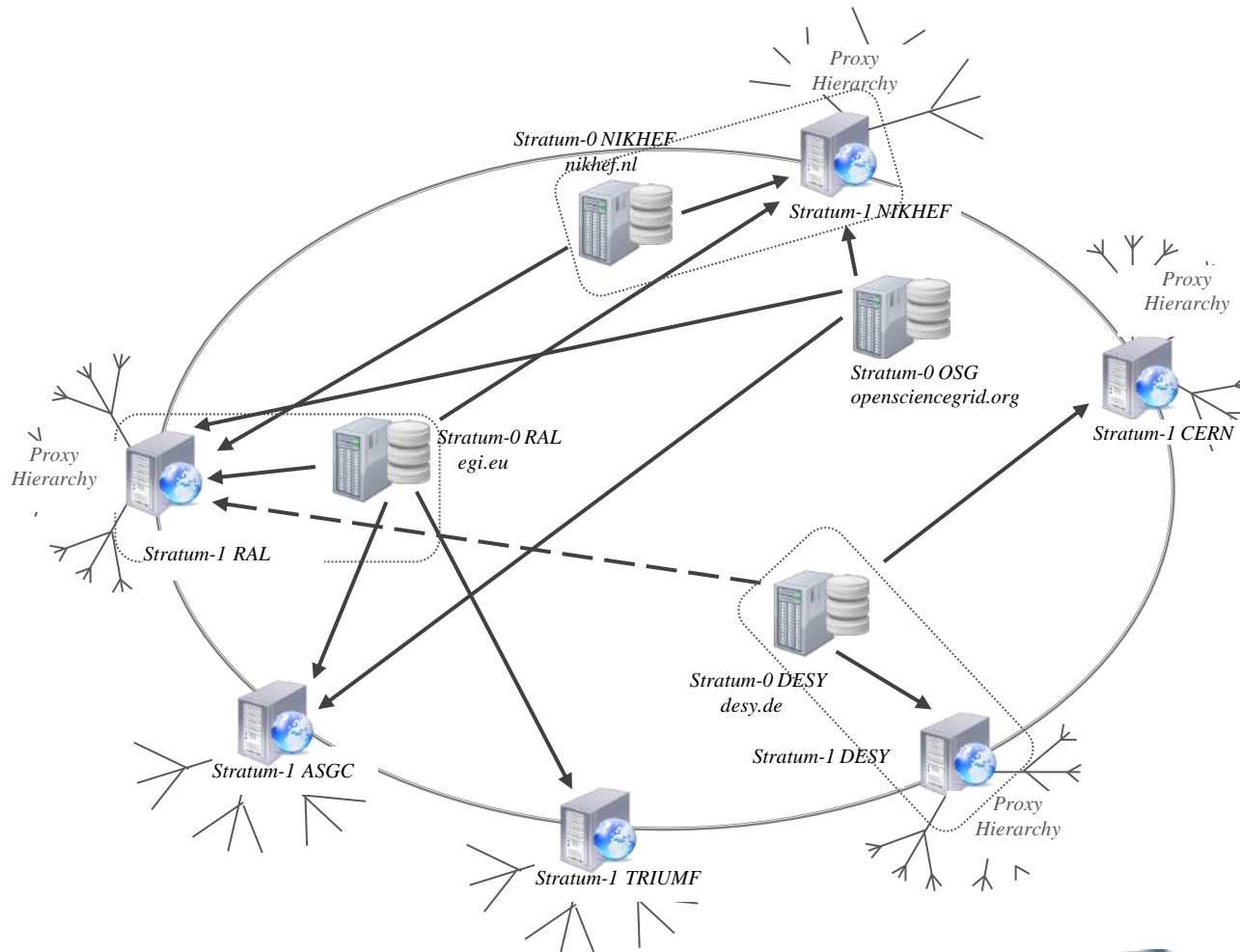
EGI CernVM-FS Infrastructure

- Domain '*gridpp.ac.uk*' less likely to continue - requests to create, host and publish repositories for various EGI VOs (not only UK)

- Following CernVM-FS WG meeting (March 2014) it was agreed that '*egi.eu*' would be the new CernVM-FS domain for new requests

- Work started to migrate '*gridpp.ac.uk*' -> '*egi.eu*'

- Existing '*gridpp.ac.uk*' repositories duplicated as '*egi.eu*'

EGI CernVM-FS Infrastructure - Present

EGI + OSG CernVM-FS Infrastructure - Present

EGI CernVM-FS Infrastructure

- 16 current repositories (only 5 at kick-off) hosted and published at RAL – ~500 GB

- RAL infrastructure
  - Stratum-0 – VM, 20GB RAM, 1.2TB HDD
  - CVMFS uploader – VM, 8GB RAM, 1TB HDD
  - Stratum-1 – HA 2-node cluster, metal boxes, 32GB, 12TB
    - *cvmfs-egi* - followed more or less Dave D. recipe – many thanks!
    - it replicates '*egi.eu*', '*opensciencegrid.org*', '*desy.de*', '*nikhef.nl*'
    - plan to integrate with WLCG Stratum-1 (*cernvmfs*)

EGI CernVM-FS Infrastructure

- Also squid machines at RAL are shared for Frontier and LHC/non-LHC CernVM-FS access

- Stratum-0 at DESY and NIKHEF (not as *egi.eu*)
  - 25 repos replicated at RAL (non-LHC, non-OSG)

- Other Stratum-1 replicas for *egi.eu* at NIKHEF, ASGC, TRIUMF

EGI CernVM-FS – Status and Plans

- Supporting sites asked to configure '*egi.eu*' (along with '*gridpp.ac.uk*') CernVM-FS domains
  - GGUS campaigns
    - i.e. IBEGRID sites for 'vo.phys.ibergrid.eu'
- Big help with cvmfs-keys v1.5
  - '*egi.eu*' configured by default
- '*gridpp.ac.uk*' to be phased out outside UK soon
  - will remain in use for few geographically UK VOs only

EGI CernVM-FS – Plans

- Still work to do with sites and VOs
  - but very good progress with 'biomed'
    - 44 out of 90 sites enabled CVMFS in Sep'14
    - 90 out of 169 CEs advertised the VO-biomed-CVMFS tag in Feb'15

- Ready to test at RAL on EGI infrastructure new features before deployment on WLCG infrastructure

EGI CernVM-FS – Plans

- ## EGI Operations
  - https://wiki.egi.eu/wiki/PROC20
    - the document describes the process of enabling the replication of CernVM-FS spaces across OSG and EGI CernVM-FS infrastructures
    - used for OSG repos replication
  - VO ID cards requirements
    - VO_<vo_name>_SW_DIR = /cvmfs/<repo_name>.egi.eu
    - VO-<voname>_CVMFS software tag

EGI CernVM-FS – Deployment Issues

- Virtual Organizations
  - are lacking manpower to progress more
  - technology and its advantages are understood, just they do not have time
  - subgroups within same VO are using subsets of software at specific sites
    - Coordination needed to consolidate SW under single CernVM-FS space
  - sometimes VOs wait too long for sites to help them with initial tests before rolling out more widely
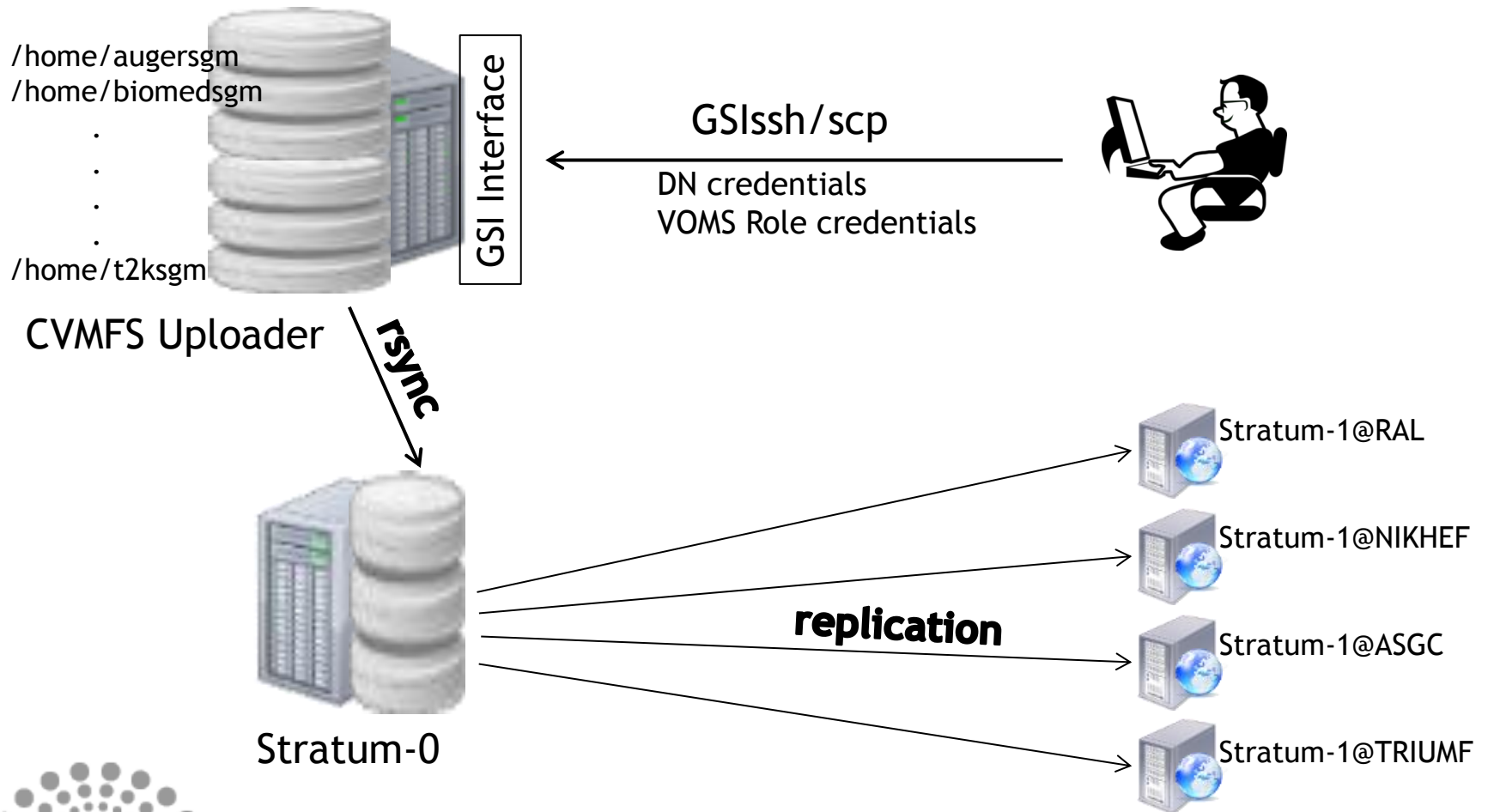
Software Installation Mechanism at RAL

- Initially local installation jobs
  - run by VO SGM accounts
  - upload SW into NFS area
  - rsync'ed into Stratum-0 then signed and published
- Replaced by web interface access
  - used by SGMs to upload, unpack tarballs into CVMFS uploader machine
  - based on DNs

Software Installation Mechanism at RAL

- Recently GSI interface added and configured on CVMFS uploader
  - faster and easier
  - allows a programmatic way to transfer and process the tarballs
  - initially based only on DN access, then VOMS Roles access also enabled – January 2015
  - software available at WN level after max 3 hours

CernVM
File system

Software Installation Mechanism at RAL

/home/augersgm
/home/biomedsgm
.
.
.
.
/home/t2ksgm

GSI Interface

CVMFS Uploader

rsync

GSIssh/scp

DN credentials
VOMS Role credentials

Stratum-0

replication

Stratum-1@RAL

Stratum-1@NIKHEF

Stratum-1@ASGC

Stratum-1@TRIUMF

Acknowledgments

- Many, many thanks to CernVM team (Dave, Rene, Jakob) for their invaluable support
- Also thanks to *cvmfs-*@cern.ch* members

Thank you!

(more) Questions?