



Baseline for WLCG Stratum 1 Operations

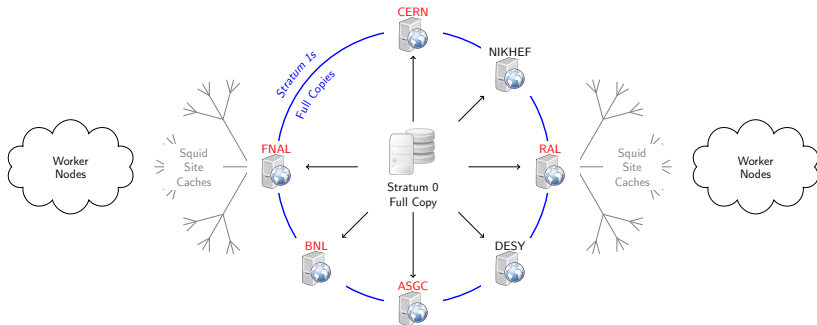
Jakob Blomer

GDB

February 8th, 2017



CernVM-FS Stratum 1 Network



- Backbone for CernVM-FS HTTP content distribution
- *Highly shared* infrastructure
- WLCG: not anymore *the* CernVM-FS infrastructure but a CernVM-FS infrastructure (others, e.g.: OSG, EGI)
- **Goal for today:** discuss baseline/guidelines wrt. WLCG (see list at the end)
 - What do experiments need/expect?
 - What would potential new stratum 1s need to provide?



- Loosely coupled set of web servers at CERN, RAL, BNL, FNAL, ASGC
- Example: `http://cvmfs-stratum-zero.cern.ch/cvmfs/alice.cern.ch`
`http://cvmfs-stratum-one.cern.ch/cvmfs/alice.cern.ch`
`http://cernvmfs.gridpp.rl.ac.uk/cvmfs/alice.cern.ch`
`http://cvmfs.racf.bnl.gov/cvmfs/alice.cern.ch`
`http://cvmfs.fnal.gov/cvmfs/alice.cern.ch`
`http://cvmfsrep.grid.sinica.edu.tw/cvmfs/alice.cern.ch`
- Requires maintenance of S0 – S1 replication, client configuration
- CDN key parameters
 - ~15 TB capacity per stratum 1 (RAID or replicated)
 - Current volume growth rate ~100 % per year
 - 1.2 B requests per month
 - 350 TB served per month
- Cost on Amazon per month:
 - S3 storage 15 TB: ~\$ 370
 - Request costs (0.5 cent / 10 000): ~\$ 600
 - Data egress (4 – 5 cent / GB): ~\$ 16,000 – \$ 20,000

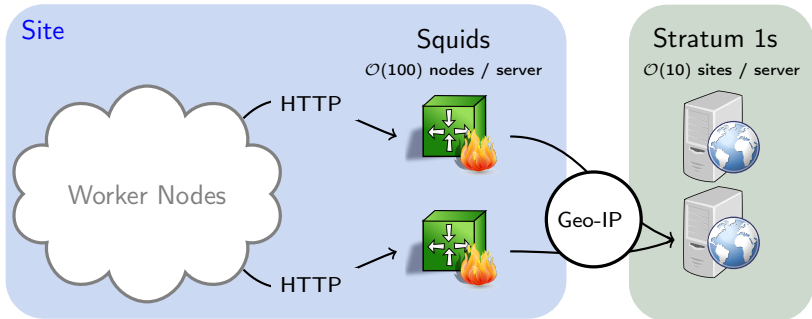


Typical Information Needs by Experiments

- 1 Which Stratum 1s are replicating my repository?
 - 2 Who is responsible for a particular Stratum 1?
 - 3 Do all the Stratum 1s serve the latest version of my repository?
 - 4 Which version of my repository is visible on a particular worker node?
 - 5 Do grid worker nodes have access to all the Stratum 1s (fail-over)?
 - 6 Is the distribution chain
Stratum 0 → Stratum 1 → Squid → worker node
healthy?
- Questions 1 – 4: retrievable through scripts (see backup slides)
Should we cover this through additional tooling / documentation?
 - Questions 5 – 6: needs help by conventions



Worker Node Configuration



Stratum 1 URLs burned into

- CernVM-FS client configuration through `cvmfms-config-...` RPMs
 - `cvmfms-config-default` provided by developers, to be replaced by `cvmfms-config-egi`
 - OSG covered by `cvmfms-config-osg`
- Squid configuration (ACLs), mostly covered by Frontier Squids



Suggested WLCG Stratum 1 Baseline / Guideline

(by and large covers current status)

- Stable list of Stratum 1 URLs
- Public list of WLCG repositories
(LHC production repos, sft.cern.ch, grid.cern.ch, cernvm-prod.cern.ch, geant4.cern.ch)
- Stratum 1 storage capacity: 20 TB, 50 % growth per year
- Stratum 0 – Stratum 1 synchronization frequency every 15 minutes
(the synchronization itself can take time for large change sets)
Do we need a guaranteed fast path? How fast?
- Running latest server software within a time window of, e.g., 2 months
- Connectivity: ports 80 and 8000 on the Internet, IPv4 and IPv6
- Replication (optionally) through LHCOPN
- Minimum capability for clients: 20 MB/s, 1000 req/s
- < 2 weeks downtime / year
- Set of public monitoring information, e.g. contact person, average throughput, list of provided repositories

Backup



How to Gather Repository Information I

- 1 Which Stratum 1s are replicating my repository?
- 2 Who is responsible for a particular stratum 1?
- 3 Do all the Stratum 1s serve the latest version of my repository?

Through CernVM-FS Monitor:

```
curl \
http://cernvm-monitor.cern.ch/cvmfs-monitor/alice.cern.ch/stratum0_details/
curl \
http://cernvm-monitor.cern.ch/cvmfs-monitor/alice.cern.ch/stratum1_details/4/
```

Self-maintained through `cvmfs_server update-repoinfo`:

```
target="http://cvmfs-stratum-zero.cern.ch/cvmfs/alice.cern.ch"
metainfo=$(curl $target/.cvmfspublished 2>/dev/null \
| grep "^M" | tr -d M)
{ printf "\x1f\x8b\x08\x00\x00\x00\x00"; \
curl $target/data/${metainfo:0:2}/${metainfo:2}M 2>/dev/null; } \
| gzip -dc 2>/dev/null
curl http://cvmfs-stratum-one.cern.ch/cvmfs/alice.cern.ch/.cvmfspublished \
2>/dev/null | grep ^S # Stratum 1 revision
curl http://cvmfs-stratum-one.cern.ch/cvmfs/info/v1/meta.json # admin info
```




- 4 Which version of my repository is visible on a particular worker node?

From the pilot job:

- Through extended attributes:
`attr -g revision /cvmfs/alice.cern.ch`
(Does not work on NFS-exported cvmfs clients)
- Through reading an application specific file that changes on publish, e.g.
`cat /cvmfs/alice.cern.ch/revision-sentinal`