WLCG Stratum 1 Operations Status and Plans

Dave Dykstra, dwd@fnal.gov

CernVM Workshop

12 September 2022

Fermilab

Stratum 1 configuration

- Stratum 1 setups vary, but in general
  - Any common CPU, memory
  - Disk usage currently in the 40T range for full replica, need extra for growth
    - RAID-10 recommended
      - For optimum write speed with **lots** of small files
      - Either by hardware or by software (zfs)
        - » Zfs can take advantage of SSD for caching metadata
- HA pair preferred
  - Some use primary/backup, switch by DNS
  - FNAL, RAL, and IHEP use cvmfs-hastratum1 package which enables quick, optionally automatic switchover sharing an IP address
    - Both machines need to update before exposing updates to clients
    - One download from upstream, then replicate from there to the other server
- Reverse-proxy frontier-squid on servers preferred
  - Helps for monitoring and for caching geo api lookups, the rest pass through to httpd

WLCG Stratum 1s map

WLCG Stratum 1s hosted distributions

Site	Distributions
ASGC	CERN, EGI, HSF, OSG
BNL	CERN, DESY (for osg), EGI (for osg), HSF, KEK, OSG
CERN	CERN, DESY, HSF, LSST, RAL
DESY	DESY, KEK
FNAL	CERN, DESY (for osg), EGI (for osg), HSF, KEK, OSG
IHEP	CERN, DESY, EGI, HSF, KEK, OSG
IN2P3	LSST
KEK	KEK
NIKHEF	EGI, OSG, RAL
RAL	CERN, DESY, EGI, HSF, KEK, OSG, RAL
Swinburne	CERN, DESY, EGI, HSF, KEK, OSG
TRIUMF	EGI
UNL	CERN (for replicas), DESY (for osg), EGI (for osg), OSG

- Not all repos replicated to all Stratum 1s
- EGI & OSG have a formal process
- Only one repo from KEK & DESY widely replicated

Configuration repositories

- Control over stratum 1 lists and all common cvmfs configuration is done through config repos
  - WLCG uses: default, egi, and osg
  - The one each site uses is based on who they trust
  - Central publication of the repository quickly updates the configuration on all cvmfs clients that use that config repo
- All 3 use different branches in one github repo, so there's a lot in common but they are not identical
- In some circumstances redirects requests through Content Delivery Network Cloudflare via openhtc.io domain

Replica management

- cvmfs-manage-replicas package enables automatically adding new repositories to replicate
  - Flexible configuration allows specifying another Stratum 1 to duplicate, matching wildcard subsets of repos on that Stratum 1
  - Many Stratum 1s replicate subsets that RAL has, and FNAL replicates what UNL has
  - Public keys for new domains automatically download from the config-repo github repository.
- Stratum 1s are configured to use `cvmfs_server snapshot –a` and `cvmfs_server gc –a` to operate on all repos
- UNL Stratum 1s are also doing `cvmfs_server check –a`

Subset of FNAL's manage-replicas.conf

```
keysource cvmfs-contrib/config-repo/master/etc/cvmfs/keys
replist http://cvmfs-egi.gridpp.rl.ac.uk:8000/cvmfs/info/v1/repositories.json

source  http://cvmfs-stratum-zero.cern.ch:8000
exclude alice-nightlies.cern.ch
repos   belle.cern.ch cms-ib.cern.ch cvmfs-config.cern.ch
repos   *.cern.ch

source http://oasis.opensciencegrid.org:8000
repos   config-osg.opensciencegrid.org oasis.opensciencegrid.org

replist http://oasis-replica.opensciencegrid.org:8002/cvmfs/info/v1/repositories.json
source  http://oasis-replica.opensciencegrid.org:8002
repos   *
```

Monitoring (1)

- Most of the monitoring is done via wlcg-squid-monitor.cern.ch
- Many monitors, but the most important one is based on the cvmfs-servermon package
  - Provides a single API point to probe if all repositories have had a successful snapshot recently, if it has not been a long time since a successful gc, and if the geo api works
    - Defaults to 8 hours and 10 days, respectively
  - Then wlcg-squid-monitor.cern.ch probes that API point and sends emails to cvmfs-stratum-alarm@cern.ch if status changes and also sends the status to the CERN XSLS Elasticsearch plot

Monitoring (2)

- The next most important monitor is frontier-awstats
  - Uploads data of where all requests come from to wlcg-squid-monitor
  - Has its own web page where data can be seen in table form, and the data is also used to look for “failovers”, where many connections are coming not through a site squid
    - Used in combination with squid registration
    - Clients at WLCG sites are configured (via the config repos) to fail to “backup proxies” at CERN and FNAL if their own squids fail, and that’s where most failovers are seen, but Stratum 1s are also watched directly

CVMFS Failover History

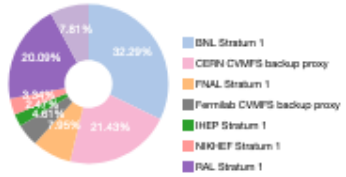
This is a visual representation of direct connections from non-squid machines to central groups of servers. The charts are both **viewers** and **controllers**: they allow for interactive exploration of the failover activity on display. The current filtering selections are noted at each chart's header.

Update

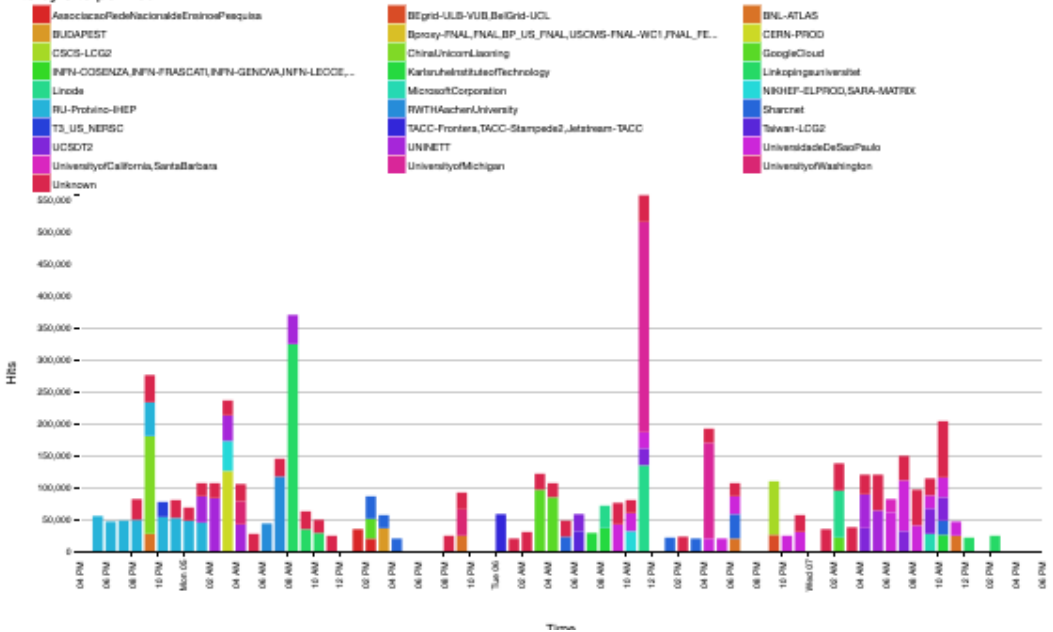
Time:
Sep 04, 2022 05:00 PM – Sep 07, 2022 05:00 PM

Time Zone:
Local

Machine Groups



Hits by site per hour



Monitoring (3)

- The distance monitor is also based on frontier-awstats, although supplemented with additional data beyond the standard awstats tool
  - IP addresses using every repo is collected
  - Plots the average distance in kilometers that clients are away from Stratum 1s per repo
  - The idea is that the plot can help identify when a repository should be replicated to another Stratum 1 because it is getting a lot of use over long distances

Monitoring (4)

- The CVMFS project monitors select repositories in more detail at https://cvmfs-monitor-frontend.web.cern.ch/
  - Not very helpful for alarms, just for overall status
  - Based on manually edited metadata
- Stratum 1s that have frontend squids also get MRTG plots of numbers of requests and data bytes sent each 5 minutes
  - Since not much caching is done, the in/out distinction isn't helpful
- Finally, all the repositories on the RAL and UNL Stratum 1s are probed daily to see if their root catalog is greater than 25MB
  - If too large, an email is sent to cvmfs-stratum-alarm@cern.ch

Links

- Stratum One admin guide:
  - https://twiki.cern.ch/twiki/bin/view/CvmFS/StratumOnes
- cvmfs-servermon:
  - https://github.com/cvmfs-contrib/cvmfs-servermon
- cvmfs-hastratum1:
  - https://github.com/cvmfs-contrib/cvmfs-hastratum1
- Configuration repository sources:
  - https://github.com/cvmfs-contrib/config-repo
- Links to WLCG squid monitors:
  - https://wlcg-squid-monitor.cern.ch