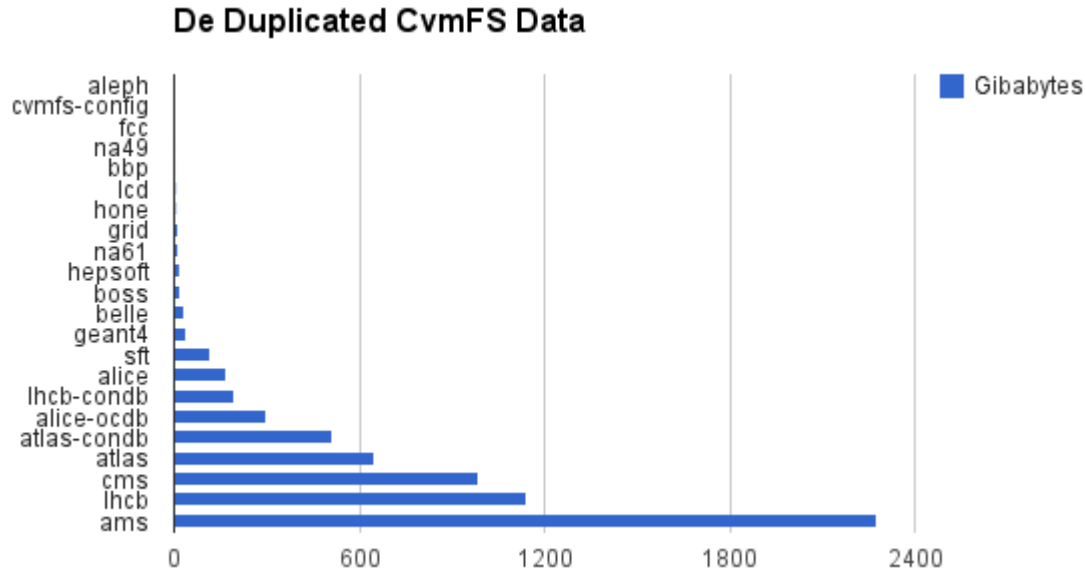


WLCG Stratum 0

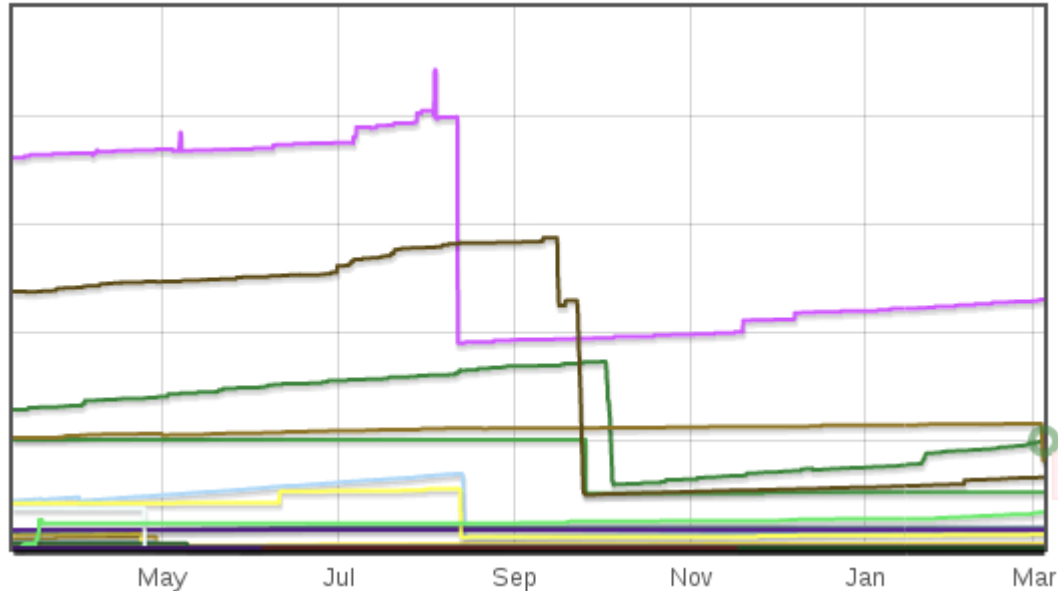
steve.traylen@cern.ch, CERN, IT-PES-PS
Thursday March 5th, CERN VM Workshop

Stratum 0 Repositories at CERN - Occupancy Now



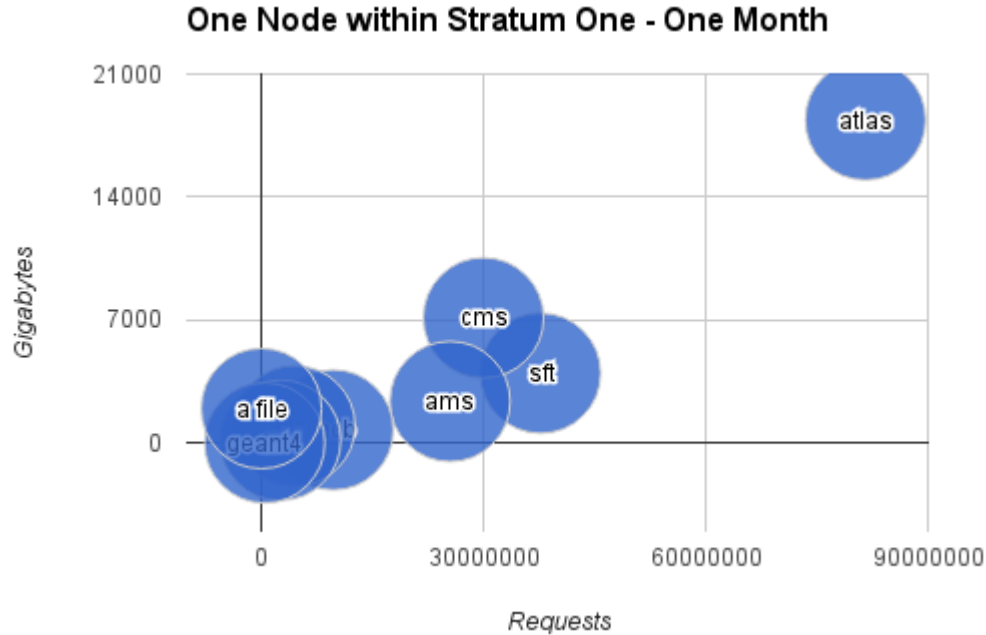
- AMS twice anyone else at 2.3 TB.
- Smallest is aleph with 34 KB.
 - AMS and Aleph use same dedicated resources, 1 CPU VM.
- LHCb is larger than it should be.
 - Not deduplicated.

Stratum 0 Repositories at CERN - Occupancy over last year



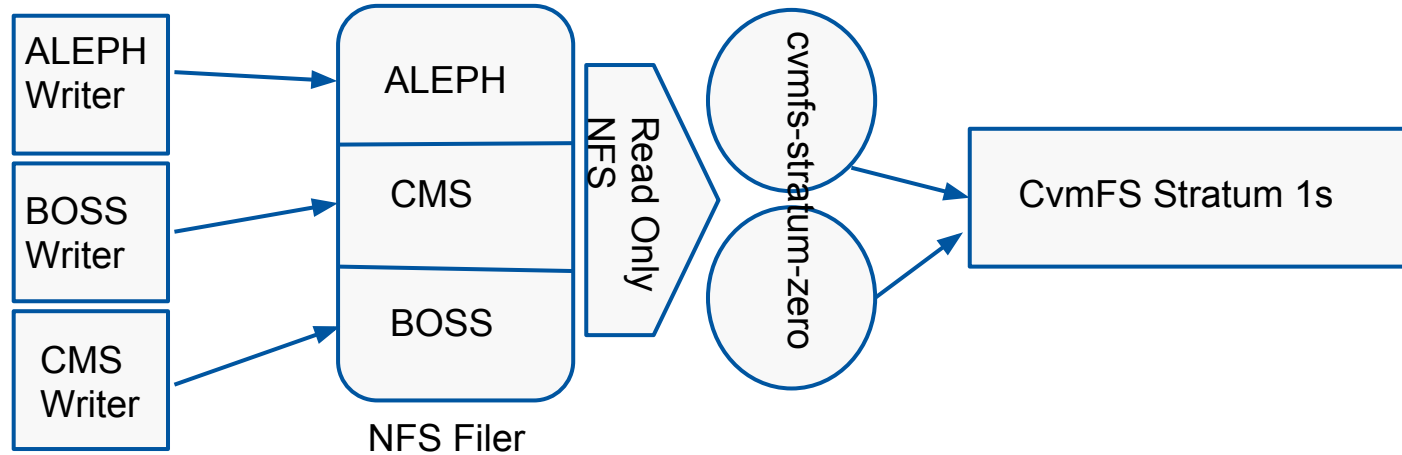
- Upgrade of to CvmFS 2.1 on Stratum 0 over summer.
- Allowed complete removal of full expanded file system.
 - LHCb (the flat line) only had it's redundant data deleted just now.

Downloads from CERN Stratum 1 at CERN during February.



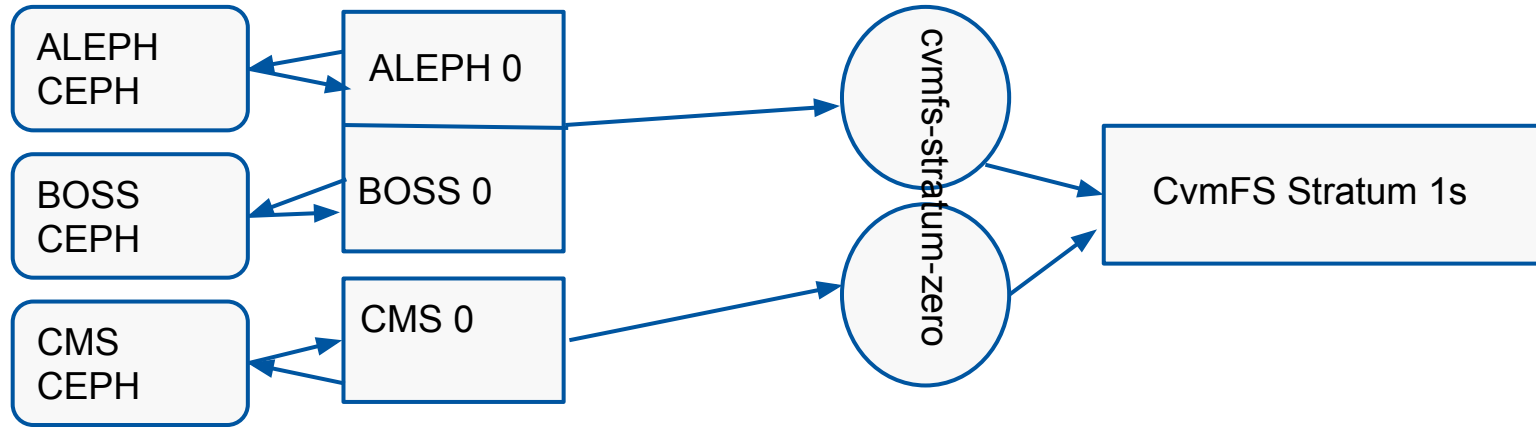
- Difficult to infer anything:
 - Atlas has the most nodes of fewest squids?
 - Maybe CMS does a better job of pinning files on batch workers?
- “a file” is just one file, accounts for 5% by size. Greater than LHCb and Alice combined.

Current stratum 0 setup, and problems



- NFS filer layer is vanishing:
 - other more flexible 'real' filesystems now available within CERN IT - CEPH.
- Writers boxes to ALEPH and BOSS are nearly never used, less than once per year.
- Writers have always been SPOF impacting writing only.

Future stratum 0 setup and comments.



- Data now on CEPH volumes.
- Stratum 0 downtime affects both writing and stratum 0->1 replication. :-(
 - CEPH volumes easier to mount on stand by nodes.
- Double up non-active customers on single nodes, e.g aleph and boss, not CMS and ATLAS.
- Migration can be done one repository at a time, downtime hours, remainder of this year.
 - Currently only theory.

Operations observations, wishlists - just one.

- CvmFS stratum 1s are vulnerable to certain deployments of CvmFS clients.
- Frequent examples:
 - One new node in CERN suddenly accounts 70% of all WLCG traffic.
 - 4 nodes in Sweden suddenly accounts for 50% of all WLCG traffic.
 - Development work in the UK unexpectedly is 50% of all requests.
 - Bursts of cloud machines at anywhere, large #s of nodes hurt the most.
- Current solution at Stratum 1s
 - Spare capacity - above cases were all individually absorbed.
 - Historical monitoring, email, tickets, typically one or two requests ongoing all the time.
 - This is probably doomed for the future, contacting owners of cloud machines...
- Some solutions.
 - geoip location of stratum 1s will probably equalize the situation across stratum 1s
 - White lists of squid servers - very inconvenient for all.
 - Throttling by network segment, domain , ... we accept laptops can be big as tier1s.
 - Automate discovery and use of close (or not) proxies.
 - CDNs or general DNS magic.
 - other....