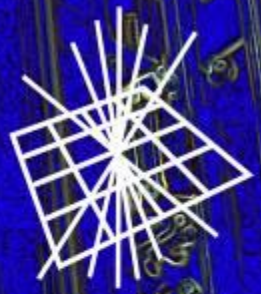


GridPP

UK Computing for Particle Physics

RAL Site Report

HEPiX Fall 2017 - KEK

16-20 October 2017

Martin Bly, STFC-RAL

Overview

- Hardware
- Networking
- Batch
- Storage
- Miscellaneous

Thanks to colleagues for contributions

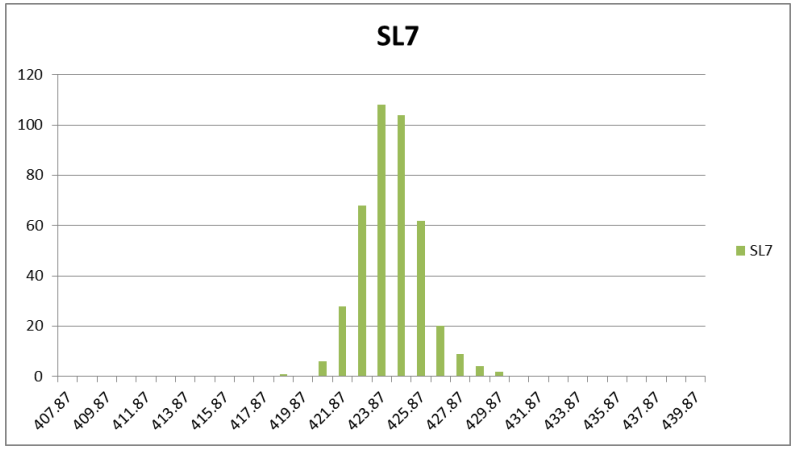
Tier1 Hardware

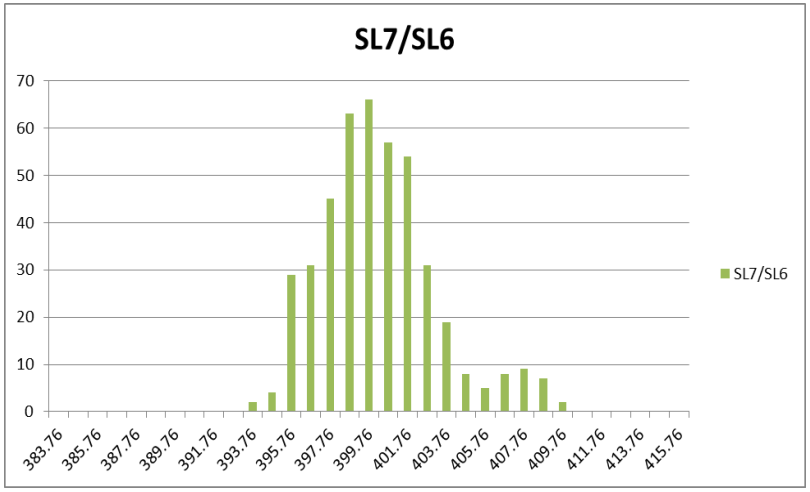
- CPU: ~240k HS06 (~24k cores)
  - FY16/17: ~20kHS06, 1920 cores (E5-2630-v4)
    - In production
  - FY17/18: procurement O(50k+) HS06
- Storage:
  - ~16.5 PB useable in Castor
  - ~13.3PB raw / for Ceph
  - FY 16/17: Additional 6720TB raw (~4.9TB configured) for Ceph
    - 35 x (Dell R630 + 2 x MD1400) units, SAS interconnect
      - Tested, awaiting transfer to production
  - FY 17/18: procurement O(20PB) raw for Ceph
- Tape: 10k slot SL8500 (one of two in system)
  - 80PB capacity (T10KD), ~30PB physics data

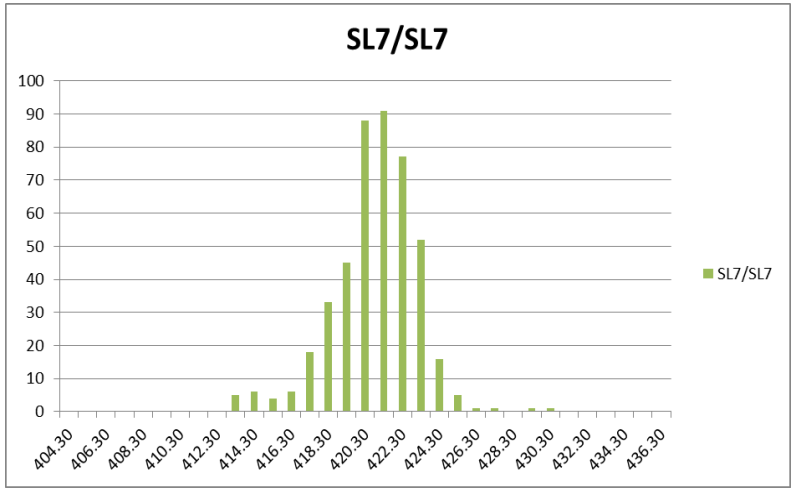
Benchmarking

Data from 2016 systems (E5-2630-v4)

Runs	412 2016 SL7 Bare Metal	440 2016 SL7 with SL6 Containers	450 2016 SL7 with SL7 Containers
Min	418.09	393.25	412.32
Max	429.73	409.66	429.32
Range	11.64	16.41	17.00
Geo Mean	423.87	399.76	420.30







Networking

- Tier1 WAN/LAN
  - No significant changes
  - Additional PerfSONAR nodes closer to the edge for diagnostics

- IPv6
  - IPv6 available on Tier1 network
  - Squid services dual stack since early August
  - CVMFS Stratum-1 dual stack since mid August
  - FTS due to be dual stack by end of 2017

Batch Farm

- WNs
  - ~26000 cores, ~22500 active
  - Migration to SL7 completed
    - Issues with SL7 installs on older generations (2010, 2011)
  - using the HTCondor Docker universe to run all jobs in Docker containers…
- … Containers
  - Queue for CentOS 7 configured on CEs
    - VOs can run Singularity containers inside the CentOS 7 Docker containers if they want
    - successfully tested by CMS but they're not yet doing this in production at RAL
- Xrootd Ceph gateway & proxies running in containers on each WN
  - Provides highly scalable access to ECHO
  - Used by ATLAS analysis jobs

Storage - Castor

- CASTOR version 2.1.16 deployed across the site
- ATLAS disk partially migrated to Echo, other VOs to follow
- Puppet-based configuration management infrastructure entirely decommissioned in favour of Quattor/SCDB

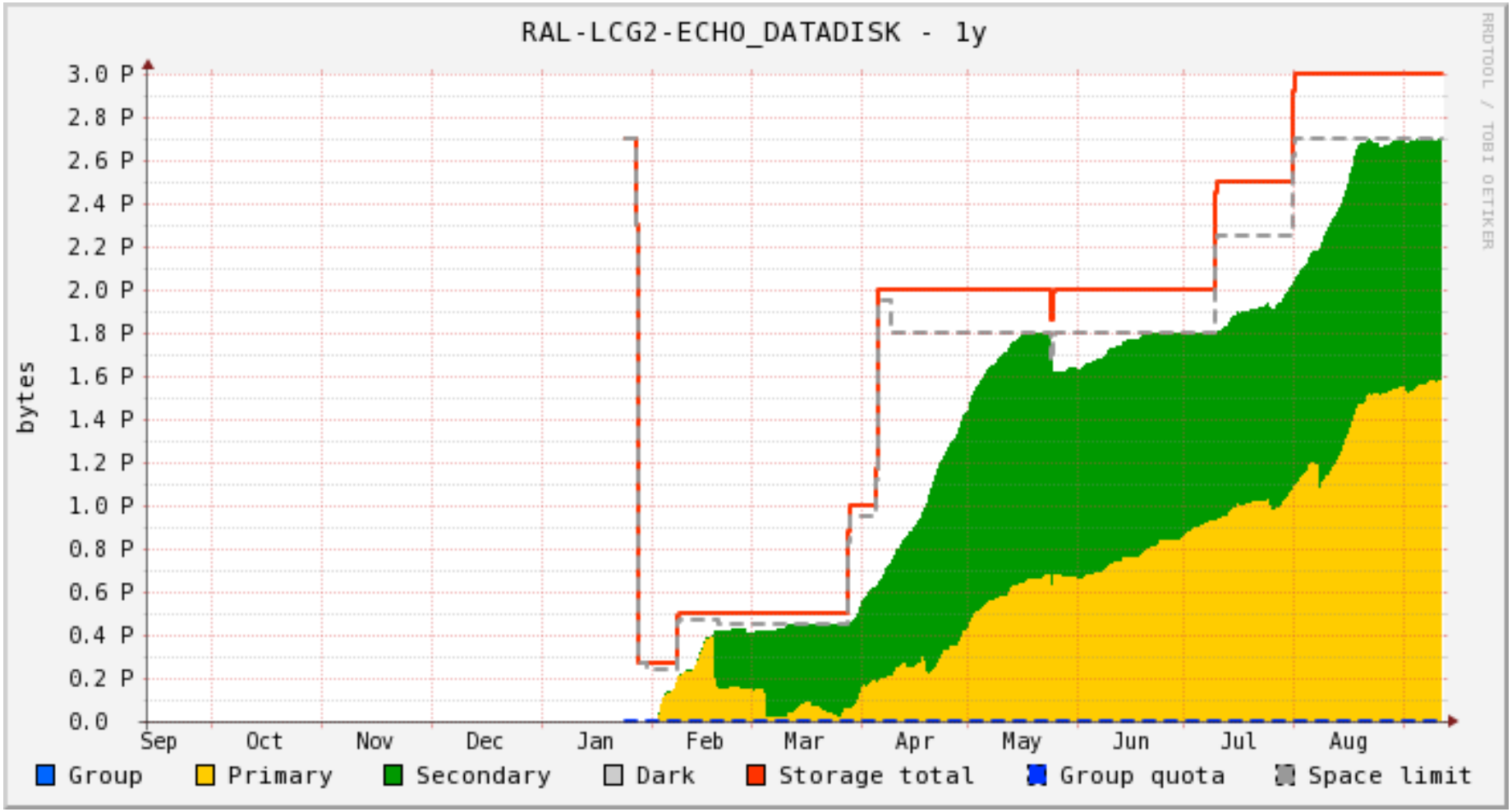
Storage - Echo (Ceph)

- Ceph cluster with Erasure Coded data pools (8+3)
- Disk only storage replacing Castor for LHC VOs
- 65 nodes x 36 disks @ 5TB
- 11.3PB raw space
- Probably the largest EC Ceph cluster in production
- GridFTP/XRootD service gateways in production

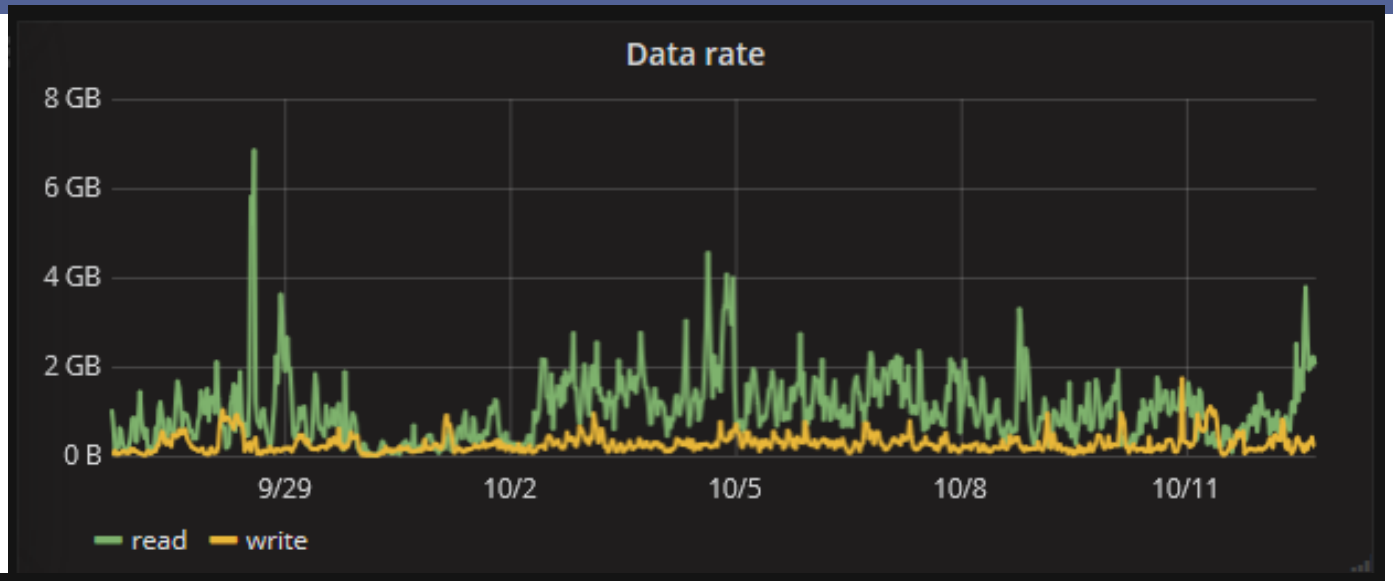
- Current numbers:
  - 3538TB raw used, => ~2.6PB data
  - Read load: 1-2GB/s typical, 5-6.5GB/s peak
  - Balancing traffic peak ~30GB/s (tuning limit)

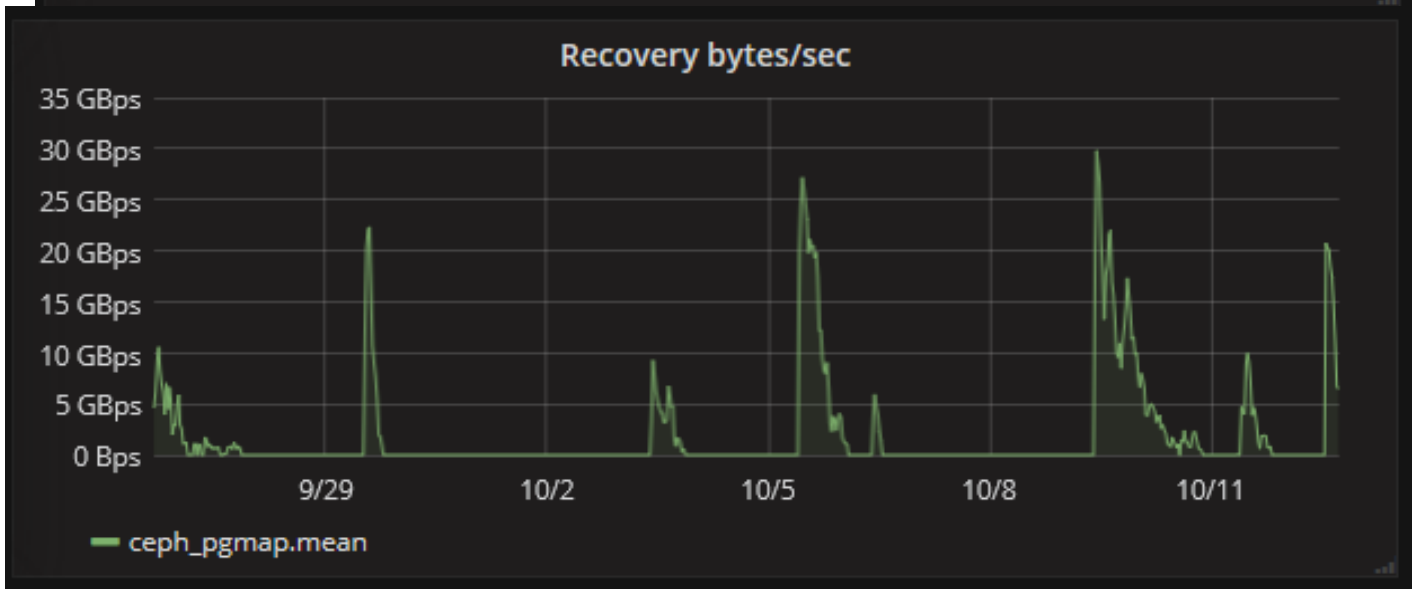
Echo II

- Atlas
  - production since April, intended pledge amount for year reached at end of July
- CMS
  - Extensive testing of all CMS activities
  - Uses direct i/o more than atlas, caching and prefetching at gateway layer
- LHCb –
  - Transfers into Echo via FTS
  - Redirection between Echo and CASTOR works
  - Gfal plugin for DIRAC being developed
- C.f talk by Ian Collier


RAL-LCG2-ECHO_DATADISK - 1y
RRDTOOL / TOBI OETIKER
bytes
3.0 P
2.8 P
2.6 P
2.4 P
2.2 P
2.0 P
1.8 P
1.6 P
1.4 P
1.2 P
1.0 P
0.8 P
0.6 P
0.4 P
0.2 P
0.0
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Group
Primary
Secondary
Dark
Storage total
Group quota
Space limit

Performance





Miscellaneous

- Move from Hyper-V to VMware
  - Cluster testing nearly complete
  - Migration planning in progress
- Databases
  - Plan to migrate to RH7 before 1 April 2018
- Patching for CVE-2017-1000253
- Continued purge of SL5 systems
- Adding support for Debian to Quattor
  - Support for switches etc

Questions?