RAL Site Report

HEPiX Spring 2016 – DESY Zeuthen

18-21 April 2016

Martin Bly, STFC-RAL

HEPiX Spring 2016 - RAL Site Report

- CPU: ~140k HS06 (~14.8k cores)
  - FY 15/16: additional ~106k HS06 in test
  - E5-2640v3 and E5-2630v3 (Supermicro, HPE)
    - 4 sleds in a 2U chassis
    - Back to 1G NICs
  - ~250 kHS06 in use in July

- Storage: ~16.5 PB disk
  - FY 15/16: additional ~13.3PB raw
  - CEPH spec
    - 2 x CPU, 64GB RAM, 36 x 6TB HDD, SAS HBA, 2 x 2 port 10G NICs

- Tape: 10k slot SL8500 (one of two in system)
  - 44PB (T10KC/D)
  - Migrations to D-only started
    - Atlas (~6PB) ongoing, LHCb (~3PB) to follow – estimated 1 month/PB

- # Tier1 LAN
  - Router plans
  - Phase 3: move the firewall bypass and OPN links to Tier1 routers
    - Scrapped – routers don't do what we need
  - Plan B:
    - Replace ancient UKLR router
      - Stacked S4810P
      - Will provide 40Gb/s pipe to border
      - Direct connection to Tier1 core
      - Landing point for OPN

- # Expansion capability of Tier1 network topology limited – need to look at moving from L2 to L3 mesh or other architectures

- # More frequent saturation of 10Gbit OPN link to CERN
  - Investigating options

- ## Batch system (HTCondor)
  - Developed a new method for draining worker nodes for multi-core jobs, enabling us to run pre-emptable jobs on the cores which would otherwise be idle
    - Running in production since late last year
    - Using this to run ATLAS event service jobs

- ## Mesos
  - Project to investigate management of services using Mesos
    - Container cluster manager
  - See talk by Andrew Lahiff (Ian Collier), today
    - https://indico.cern.ch/event/466991/contributions/1143587/

- Soon: larger storage for Stratum1 service (WLCG etc)

- CernVM-FS Workshop @ RAL, 6-8 June
  – Hosted by RAL with SCD and GridPP UK project
  – Invited speakers from Vmware, Linaro Enterprise Group, IBM HPC division, Pivotal, Mesosphere
- https://indico.cern.ch/event/469775
- Registration open…

- v2.1.4
  - stable running for the start of run 2
  - Major improvements in data throughput from disk thanks to scheduling optimisations
  - OS (SL6) and Oracle version upgrades for entire system
- v2.1.15 in test
  - investigating DB issues around memory use
    - Delayed slightly ☹ by broken HW now fixed ☺
- Starting integration work with the new Echo (Ceph) service

- # Cloud

  - Production service using OpenNebula

  - Department and wider use in STFC, including LOFAR

  - Infrastructure contribution from ISIS

  - See talk by Alex Dibbo

    - https://indico.cern.ch/event/466991/contributions/2136879/

- # CEPH

  - Production level service underpinning Cloud infrastructure

  - Working towards deployment for large scale science data storage

  - See talk by James Adams

    - https://indico.cern.ch/event/466991/contributions/2136880/

- Windows Hyper-V 2012 failover cluster is production ready
  - Beginning move to Windows Hyper-V 2012, but see VMWare as longer term solution
  - 5 hosts, 130 logical cores, 1280GB RAM total
  - Tiered EqualLogic arrays (82TB total)
    - Eql-varray2: 62TB, 7.2K SATA
    - Eql-varray3: 22TB, 10k SAS

- 10 systems testbed
- Dual stack
  - UI, Ceph gateway, squid, message broker, GocDB testing, …
- 10 more systems to be added
  - Frontier testing to begin
- OPN router replacement should enable better IPv6 connectivity options

- Tier1 funded by GridPP project
  - Grant renewed for FY16/17 to FY19/20 (GridPP5)

- Some internal reorganisation of staff grouping within Systems Division in Scientific Computing Department
  - No affect on Tier1 Service

- New mobile provider for STFC (UK Gov)
  - Phased migration in progress, no problems

- Old-style RAL addresses withdrawn
  - A.n.other@rl.ac.uk no longer works in most cases

- **Diamond Light Source (DLS)**
  - UK's national synchrotron science facility
- **SCD provides an archive for DLS data**
  - 1.1 Billion files / 4.7PB
  - Metadata is catalogued by ICAT (icatproject.org)
  - Data stored within CASTOR at RAL
  - Experimental data can be recalled from tape and downloaded at a later date from scientists' home institutions