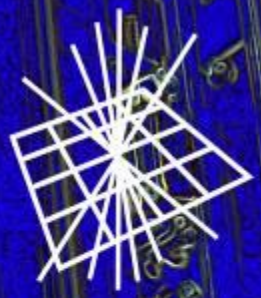

GridPP

UK Computing for Particle Physics

RAL Site Report

HEPiX Fall 2012, Beijing

15th – 19th October

Martin Bly, STFC-RAL

Science & Technology Facilities Council

e-Science

Overview

- News
- Developments
- Hardware
- Networking

RAL News

- Departmental Developments
  - Merger of e-Science Department and Computational Science and Engineering Department into Scientific Computing Department completed
  - “Science, Data and Solutions”
  - Dr Adrian Wander appointed Director of Scientific Computing and head of SCD
  - Four Divisions:
    - Systems (David Corney), Data Services (Juan Bicarregui)
      - Tier1 Group is in Systems Division
    - Technology (Mike Ashworth – acting), TBC (TBC)
  - Department includes services such as Tier1, BlueGene/Q (Blue Joule), iDataplex (Blue Wonder), Jasmine/CEMS and Emerald

Site Power

- Switch gear in power feeds to RAL site being replaced
  - Old paired system needed replacement
  - One half done, switches onto new switchgear during September
    - Significantly less risk of failure due to switchgear
    - But still risk due to outside influences (only a single feed)
  - Work progressing well on second set of switch gear
  - Enhanced risk period till end November
    - Estimate that can cut over to the other cable upstream of the switchgear in less that 48 hours

Machine Room

- In-row cooling working
  - Need to balance load in enclosed aisles to get the best possible cooling ability from both IRC and under floor system
- Enclosed aisles
  - Three cold aisles, some Tier1 kit in one of them
  - Work complete on system to drop enclosure roof panels when fire suppression system is triggered
  - Additional lights and warning sounders installed
  - Changes to machine room lighting (PIRs don't work)
  - Additional CCTV
- Floor 're-sealed'

Hardware changes

- Summary of April report:
  - 68 x 40TB SATA disk servers with 10GbE
  - 26 x 4-node double-twin WNs, X5650/E5645, 4GB/core, twin disk
  - 40GbE/10GbE and 10Gbe/1GbE switches, more tape servers

- New since April
  - Dell R610 database and tape servers
- To come
  - 2012 Capacity Procurements:
    - ITTs issued and bids received, in evaluation
  - Routers to connect to site core
    - Extreme x670V
  - More virtualisation hypervisors

Storage Issues

- None – fingers crossed!
- Seem to have entered a stable period with few disk servers causing problems other than the usual thrown disks
  - Probably due to program of testing and preventative upgrading of RAID controller firmware
- But…
  - … one of our suppliers went bust
  - Luckily only owed us maintenance and warranty on some batches of disk servers
  - Warranty on disks still available from OEMs
  - Not proving to be too much of a problem but does increase turnaround time for acquiring replacements – increased spare pool to compensate

Tape

- Tier1 has SL8500 tape library with 10K slots, 8 hand-bots
  - Second identical SL8500, with tape pass-though facility
  - 83 drives, most of which are T10KA/B/C
- Castor system has 18 x T10KA, 9 x T10kB and 14 x T10KC drives
  - Distributed between VO instances – some permanently assigned, some shared
  - Dedicated bulk repacking instance shares tape drives but has own fast cache servers
- Campaign to repack 1.5PB LHCb data on 2334 x T10KA tapes
  - 4 x T10KA reading, 2 x T10KC writing
  - Script kept read queue full, oldest tapes first
  - Enhancements in Castor 2.1.12 to use buffered tape marks allowed the T10KC to deliver about 220MB/sec on writes
  - 1.5PB repacked to 204 x T10KC tapes in about 4 weeks
  - 1 tape with 4 unrecoverable files – LHCB able to recover them from other sites
- ADS system now closed for science data
  - Data has been migrated off (Castor, …)
  - Tapes now being wiped - ~500TB

Networking

- Site
  - 'Early Summer' - boundary routers were replaced by two pairs of Extreme x670V units to provide resilient 40Gb/s connectivity for site and multiple 10GbE and 40GbE connections within site
  - Tier1 plan to use L3 capabilities of S4810s as routers between Tier1 and Boundary not possible – L3 capability does not (yet) include Policy Based Routing
  - New plan – use pair of Extreme x670V switches instead. Very similar to S4810p but can do PBR
  - Tier2 using S4810p with host based routing
- Asymmetric Data Transfer rates in/out of Tier1
  - Investigations still ongoing
- LAN
  - New Force10 routers and core switches to be deployed with mesh-type arrangement linked at multiple 40Gb/s with storage connectivity at 10Gb/s
  - PerfSonar monitoring for LHCOPN
    - Rollout across GridPP infrastructure almost finished

Questions?