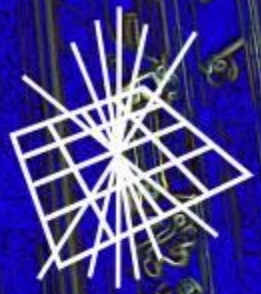


GridPP

UK Computing for Particle Physics

Tier-1 Data Storage Challenges Extreme Data Workshop

Andrew Sansum

20th April 2012

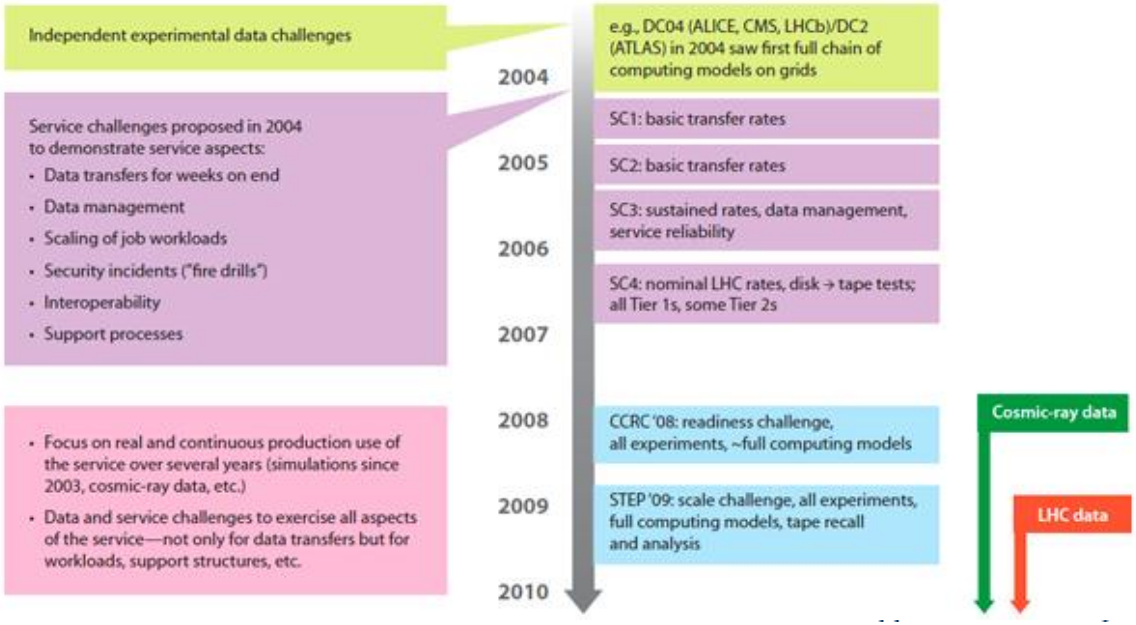
Science & Technology Facilities Council

Rutherford Appleton Laboratory

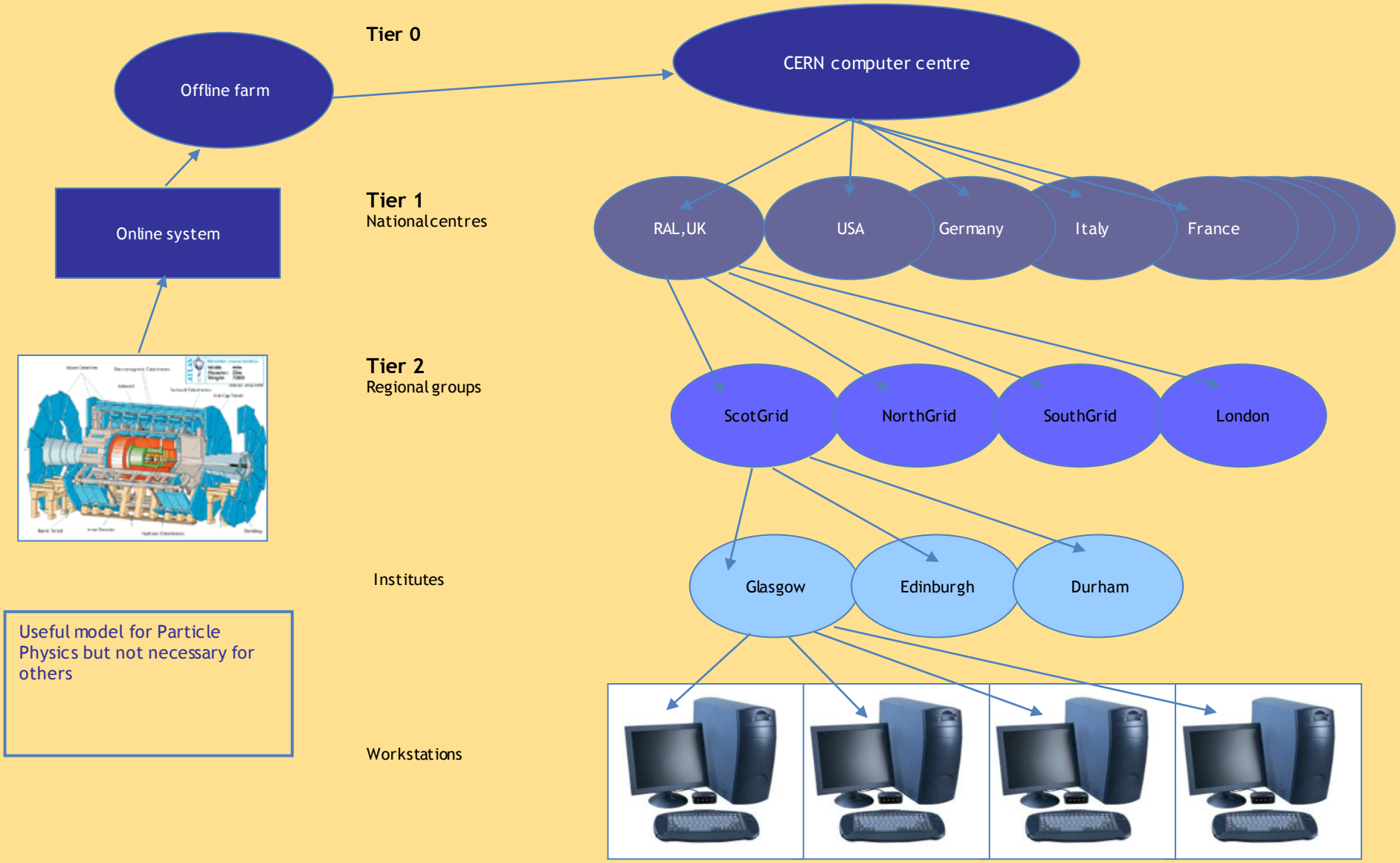
RAL Tier-1 in a Nutshell

- Primary role – compute resources for LHC. Find Higgs etc
- Part of the global WLCG Grid (globally 386M jobs and 164PB disk in 2011)
- STFC grant to UK GRIDPP collaboration.
- Memorandum of Understanding between CERN and STFC
- High availability (>98%), annual hardware procurement cycle

- RAL – 10PB disk, 10PB tape
- Making global storage grid work is hard!
- Started "production" in 2004.
- Evolved through series of service challenges.
- First physics in 2009

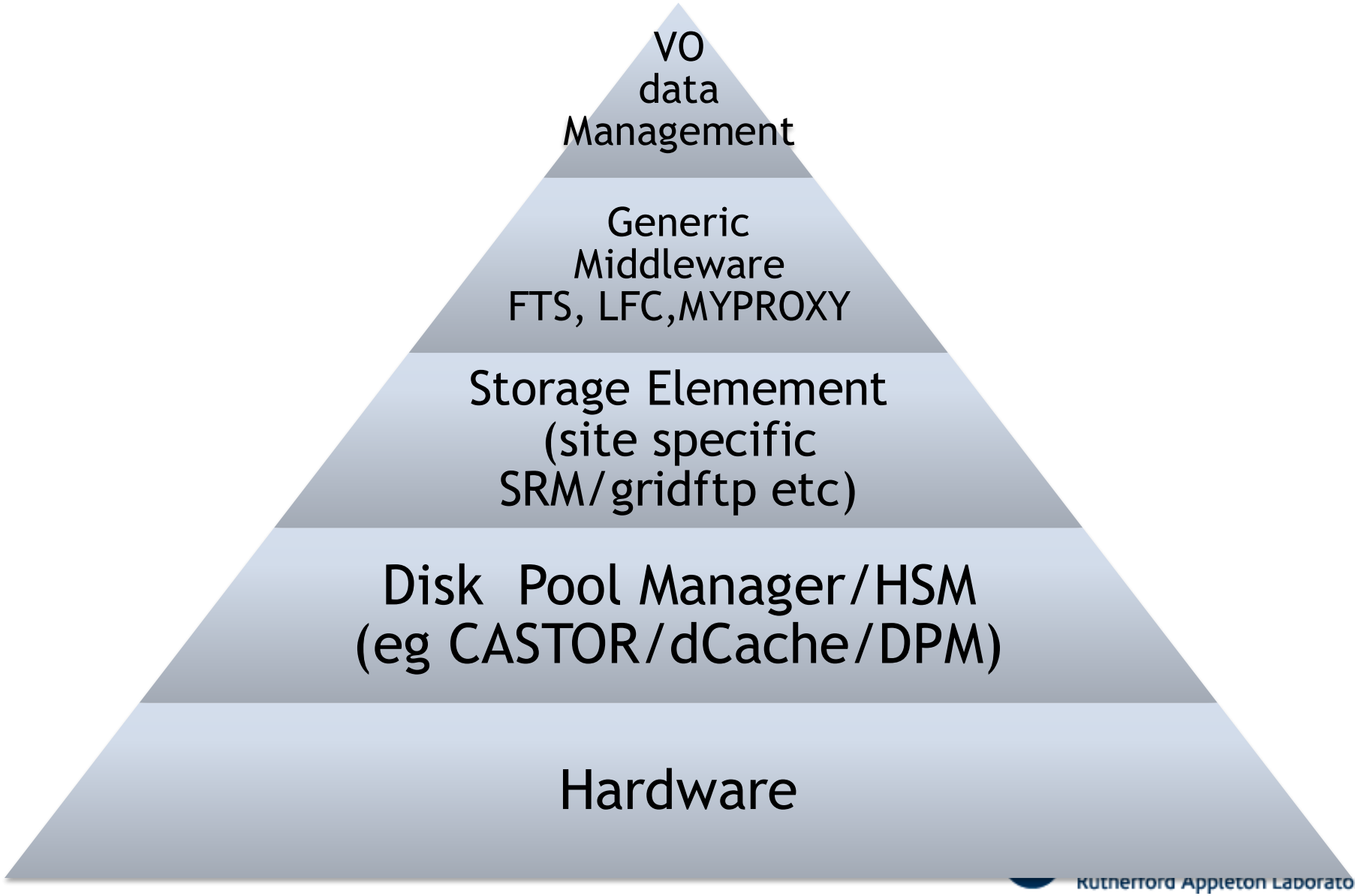


Context: WLCG architecture



Useful model for Particle Physics but not necessary for others

Storage Management Hierarchy



Challenge: Disk Hardware

- Experiments require:
  - High capacity (15PB by 2015)
  - High bandwidth (currently access disk at 50GBytes/second)
- Project requires low cost - £170 per TB
- CASTOR team prefer small allocation volumes (thin disk)

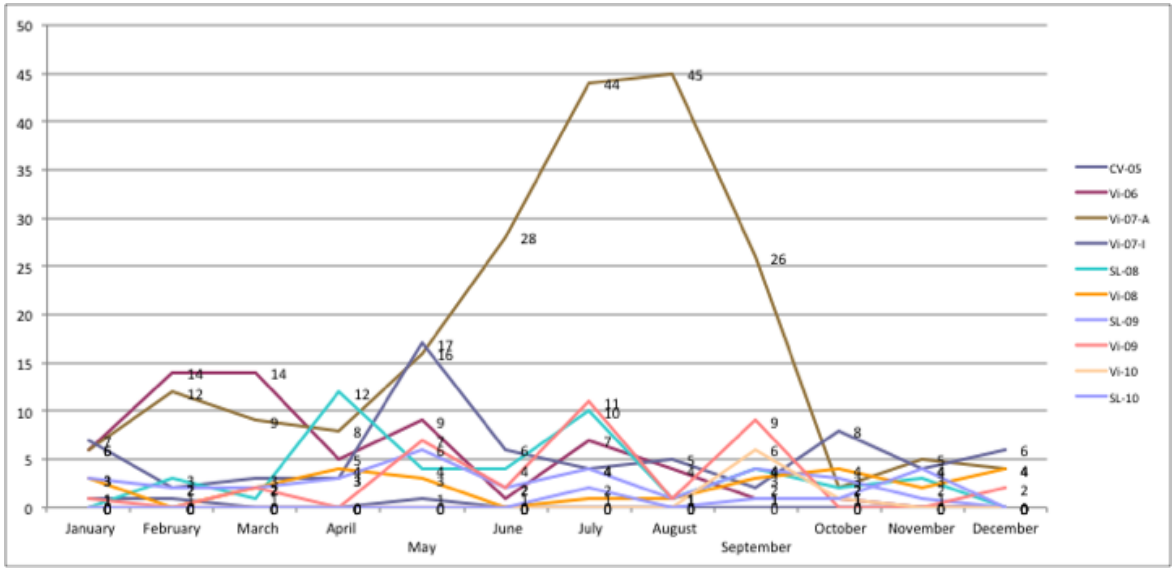
- Choose to buy many small disk servers
  - Currently 490 servers
  - 10PB disk capacity
- Only white box
  - SATA RAID 6
  - 16 or 24 disk drives

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Challenges: Disk Drives Fail

- Tier-1 has about 10,000 disk drives deployed
- Observed drive failure rate 3-6% pa normally
- Expect a drive failure every day in normal operation
- Sometimes generations fail (see Vi-07-A since April)

- Weed out early life failures with aggressive load/acceptance test
- Drive failure exception monitoring
- Dedicated "Hardware technician" who replaces drives, repairs hardware etc.
- Failure rate monitoring to detect long term trends/departure from norm
- Standby capacity to allow emergency phase-out of a whole generation when necessary

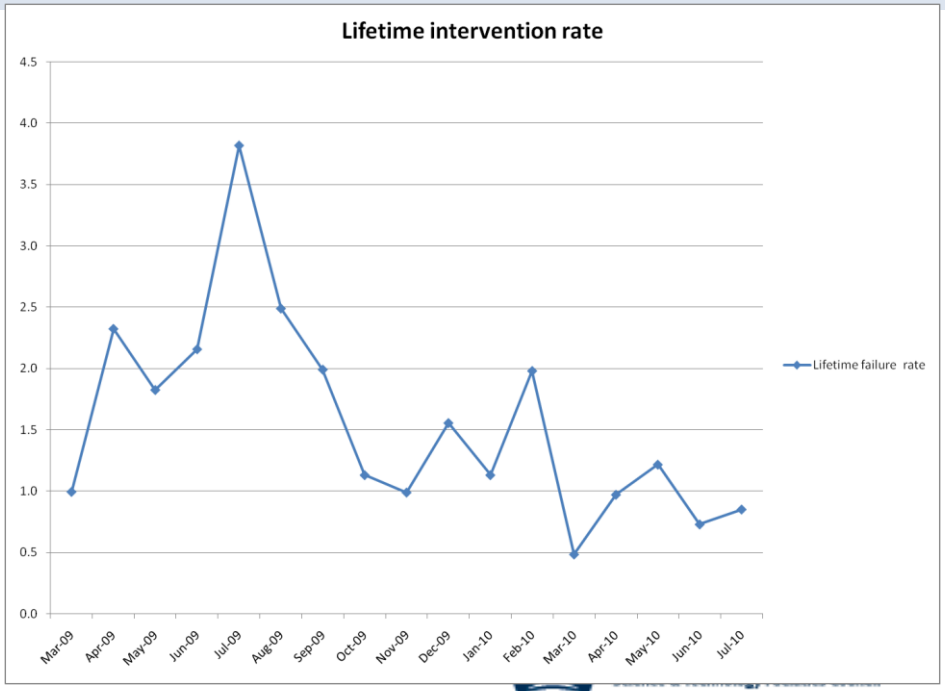


Science & Technology Facilities Council
Rutherford Appleton Laboratory

GridPP UK Computing for Particle Physic — Challenge: Disk Server Exceptions

- Experiments expect uninterrupted access to ALL their data, ALL of the time
  - Server failures can take small fractions of dataset offline for periods of hours to few days
- We observe a per server lifetime exception rate (excluding single drive failure) of about 1-2 (per server per 4 years)

- Leads to 10-20 server interventions per month (not easy)
- Disk server exception monitoring and integrated, ticketing and callout system
- Require VO input regarding speed required and retention policy.
- Detailed intervention process to ensure successful hand off between teams
- Daily inter-team review meetings to track current exceptions

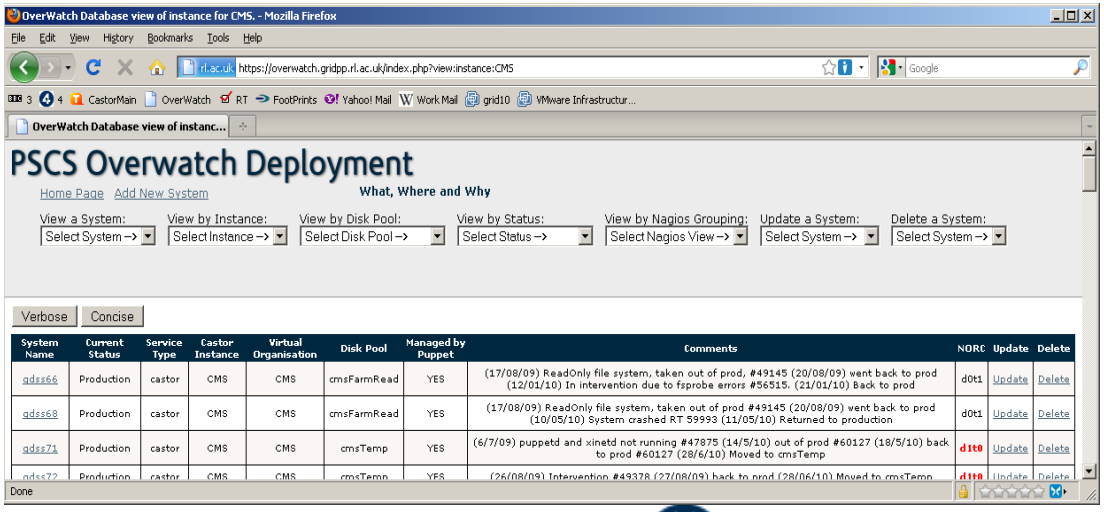


GridPP Challenge - Server Deployment

UK Computing for Particle Physics

- Tracking state of ~490 disk servers is not easy.
- Servers need to progress smoothly through deployment system (don't forget any, don't miss critical step ...)
  - Servers can be working, draining, waiting repair, repaired, certifying, ready to deploy, being deployed, deployed, etc etc
- Integrated management database to track state.

- Hardware tracking database integrated into deployment and monitoring systems.
- Detailed, documented, routine processes for deployment/removal
- Workflow managed through ticket system
- Regular deployment meetings to coordinate server moves



Science & Technology Facilities Council
Rutherford Appleton Laboratory

Challenge: Disk Procurement

- MoU commitments require timely hardware deployment
- Procurements search for disk solutions that are "High performance, high reliability, low cost".
  - Target price currently £170/TB
  - Currently buy 16 or 24 drive white box disk servers
- Accept that purchased hardware does not always work as described/required (or even work at all).

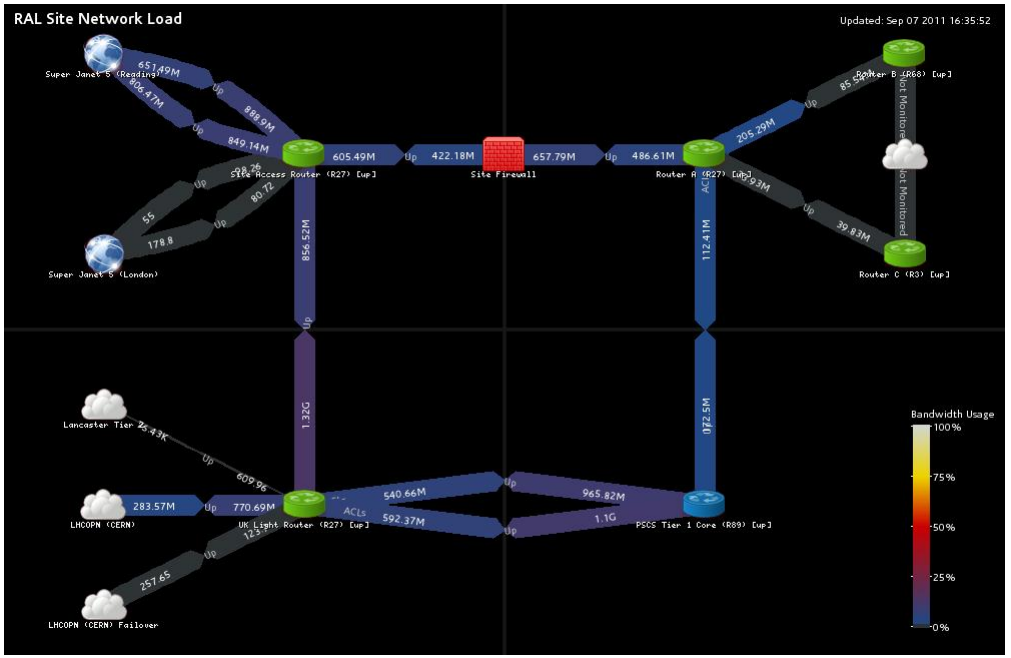
- Seek companies strong on project management with good internal processes and links to manufacturers
- Procure from 2 suppliers with 2 different technical solutions
- Aggressive 4 week acceptance test
- Maintain operational buffer of 50% of 1 year procurement against problems

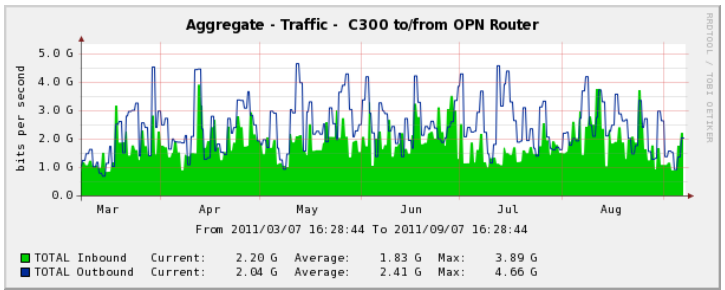
Science & Technology Facilities Council
Rutherford Appleton Laboratory

Challenge: Wide Area Network

- Require high bandwidth, low contention and resilience
- Team must be able to resolve complex performance problems on inter site transfers on deep middleware stack

- Access Tier-2s via production SJ5
  - Primary 2*10Gb/s to Reading
  - Secondary 2*10Gb/s to London
- LHC Optical Private Network (OPN) for access to Tier-0 and Tier-1s
  - Primary 10Gb/s to CERN
  - Secondary 10Gb/s to CERN by alternative physical path
- Typically 3Gb/s (with 10Gb/s peaks)



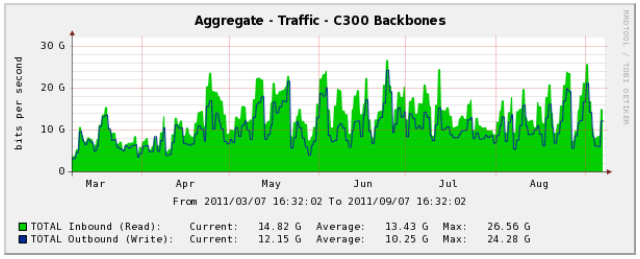


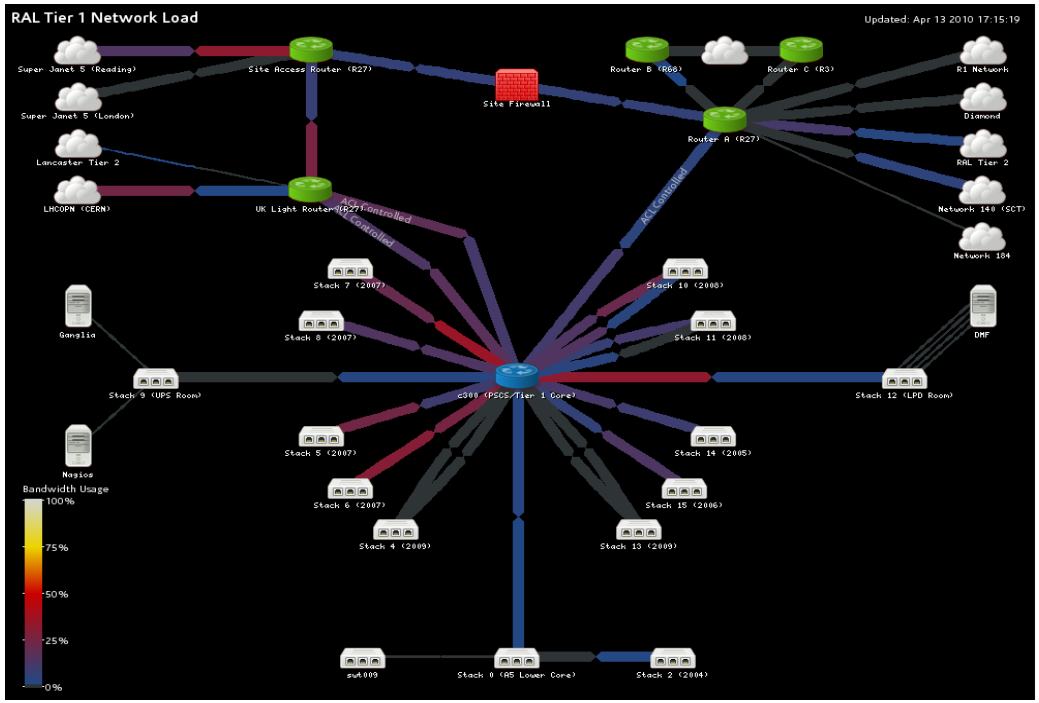
Science & Technology Facilities Council
Rutherford Appleton Laboratory

Challenge: Local Area Network

- Need to accommodate:
  - 500+ high bandwidth disk servers
  - 500 worker nodes
- Must be low cost (hence commodity)

- Star topology
  - 64 port Force10 C300 at centre
  - 12 Nortel stacks with multi-10Gb uplink
- Mix CPU nodes and worker nodes on same stacks to reduce inter stack traffic
- Routinely 20-40Gb/s with 100Gb/s peaks.





Science & Technology Facilities Council
Rutherford Appleton Laboratory

Challenge: Tape Robotics

- Experiments require 20PB long term "tape" storage by 2015
- Capacity requirements evolve according to LHC schedule
- Experiment Bandwidth estimates/access patterns uncertain
- Migrating between generations can be slow (0.5PB/month)

- Keep planning flexible – respond to changes
- Avoid buying too early (experiments don't always use what they say they will)
- "STK" SL8500 tape robot (2nd non-Tier-1 robot)
  - Currently using about 7500 tapes
  - 3 generations of tape drives:
    - 18 T10000A drives (LHCB+ALICE)
    - 9 T10000B drives (CMS)
    - 14 T10000C drives (ATLAS)

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Challenge: CPU Hardware

- No real challenge!
- Very few deployment / operational problems with CPUs!!
- Usually just work!!!

•Mainly buy white box

•Typically 4 motherboards in chassis solution most cost effective at the moment.

Some last thoughts

- Moving data reliably between multiple sites for multiple experiments at high speeds is not easy!
- A global project needs a global community working together. Responsiveness and good communications are essential. Automation is vital.
- High capacity storage operation is not easy! It is possible to make commodity hardware work well in high volume but you need mature processes.
- The operational cost of managing commodity storage hardware faults is < 30% of hardware price. Spending more is unlikely to be cost effective for us.
- A storage system that is resilient against any single hardware or service failure makes operational life less stressful – we have a system that is only partly so.