

# GridPP

UK Computing for Particle Physics

## Tier1 Site Report

HEPSysMan, RAL

10-11 June 2010

Martin Bly, STFC-RAL



Science & Technology Facilities Council  
e-Science

- RAL Stuff
- Building stuff
- Tier1 Stuff



- Plan to remove support for old-style [f.blogs@rl.ac.uk](mailto:f.blogs@rl.ac.uk) email addresses in favour of the cross-site standard [fred.bloggs@stfc.ac.uk](mailto:fred.bloggs@stfc.ac.uk) - **still under discussion**
  - significant resistance to this
  - no date as yet
- International Space Innovation Centre

*“A new environment where government, universities and industry can work together to better exploit scientific opportunities and create new technologies, applications and IP that exploits UK, European and Global Space Programmes”*
- Will focus on three key areas initially:
  - Understanding and countering **Climate Change**
  - Ensuring the **Security of space systems and services**
  - Exploiting the data generated by **Earth Observation satellites**

- UPS problems
  - Leading power factor due to switch-mode PSUs in systems
  - Causes 3KHz ringing on current, all phases
    - Load is small (80kW) compared to capacity of UPS (480kVA).
  - Most kit stable but EMC AX4-5 FC arrays unpredictably detect supply failure and shut down arrays
  - Longer feed cable from UPS to PDU has made 50% reduction in distortion
  - Proposed long-term solution: isolation transformer
- New chillers + pumps
  - Additional 1500kW cooling capacity: 4 chillers in 3+1 redundant configuration to provide at total of 2250kW
  - Additional and replacement higher capacity pumps fitted
- Dust in machine rooms



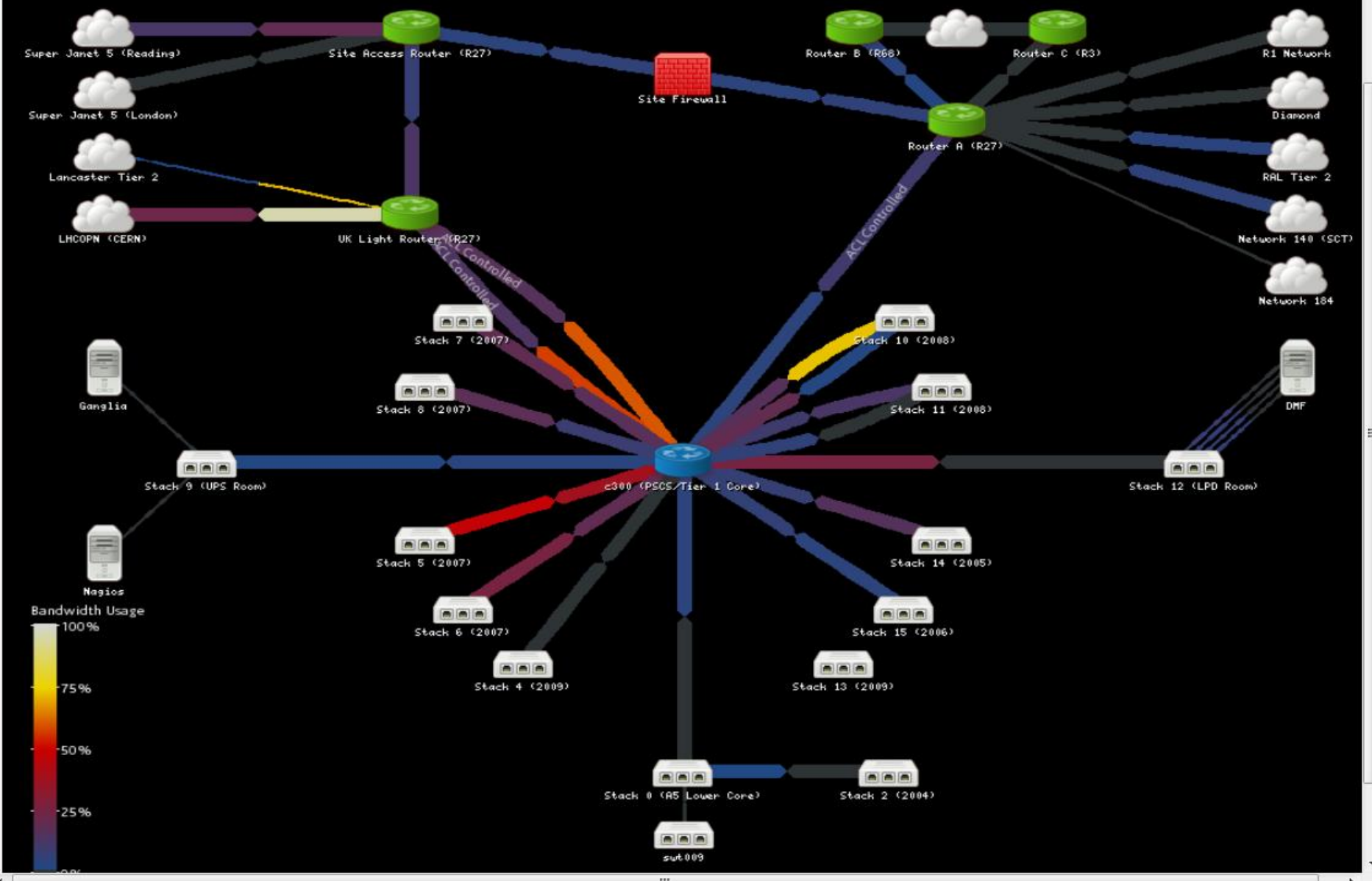
- 98 x SuperMicro 4U/24-bay chassis with 2TB SATA HDD
  - RAID6 (40TB) or RAID6+1 (38TB) on Adaptec or LSI controllers
  - ~3.8PB useable, Site => ~8PB
- 21 x SuperMicro Twin<sup>2</sup>, 2 x E5520, 3GB/core, 1T HDD
- 21 x HP SL6000 4-up chassis, 2 x E5520, 3GB/core, 1T HDD
  - ~15,500 HS06, Site => ~49,700 HS06, ~5260 cores
- FC storage arrays: 7 x Infotrend
- iSCSI arrays: 2 x Infotrend
- 40+ Dell R410 servers
  - Tape servers, virtualisation and services hosts
- 5 HP DL360 database servers with FC

- Additional Force10 C300
  - Twin for existing C300
  - ‘Spares’, testing
  - To be configured as a resilient routing pair
- WAN developments
  - Existing 10Gb/s site link to SuperJanet5 being doubled to 20Gb/s
    - Failover capacity maintained
  - LHCOPN failover link @ 10Gb/s
  - Both being commissioned
- LAN developments
  - RAL campus network-futures plan in development



RAL Tier 1 Network Load

Updated: Apr 07 2010 14:45:17



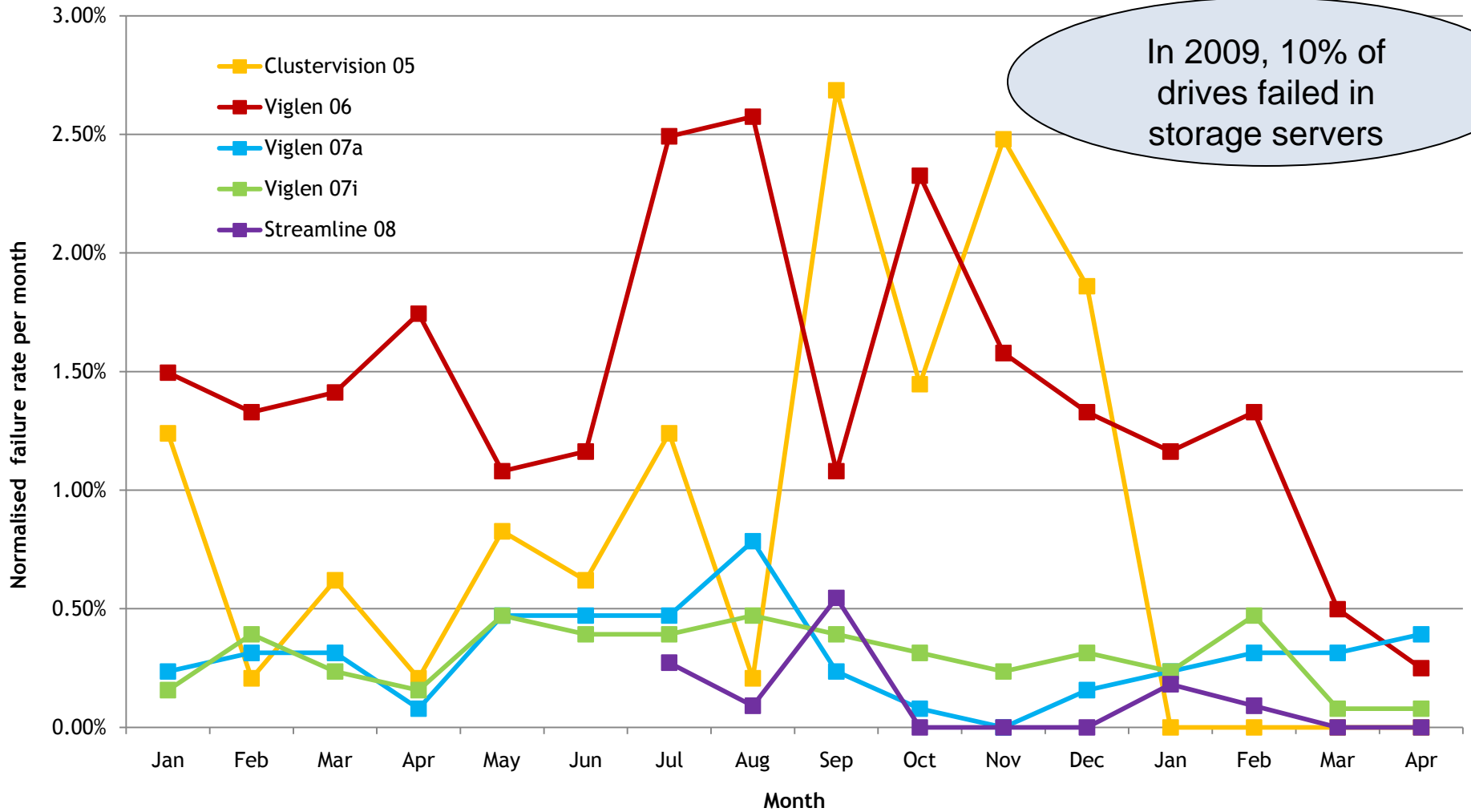


- One of two batches of the FY08/09 capacity storage failed acceptance testing: 60/110 servers (~1.2PB)
  - Rapid onset of failure mode
- Supplier involved
  - Significant testing with various ‘solutions’ unsuccessful
  - Escalated to HDD OEM
  - Additional testing program started
    - HDD FW variants
  - Unable to identify cause of problems
- Supplier/OEM agree swap out of disks
  - Combination clearly doesn’t work despite compatibility list!

- One of two batches of FY09/10 capacity storage delayed in proving tests: 60/98 servers (~2.3PB)
  - Multiple drive throws
  - Disks tested show now issues
- Supplier and OEMs involved
- New firmware for both controllers and drives
  - Flash in-situ
- Successful proving test
- Now in acceptance testing

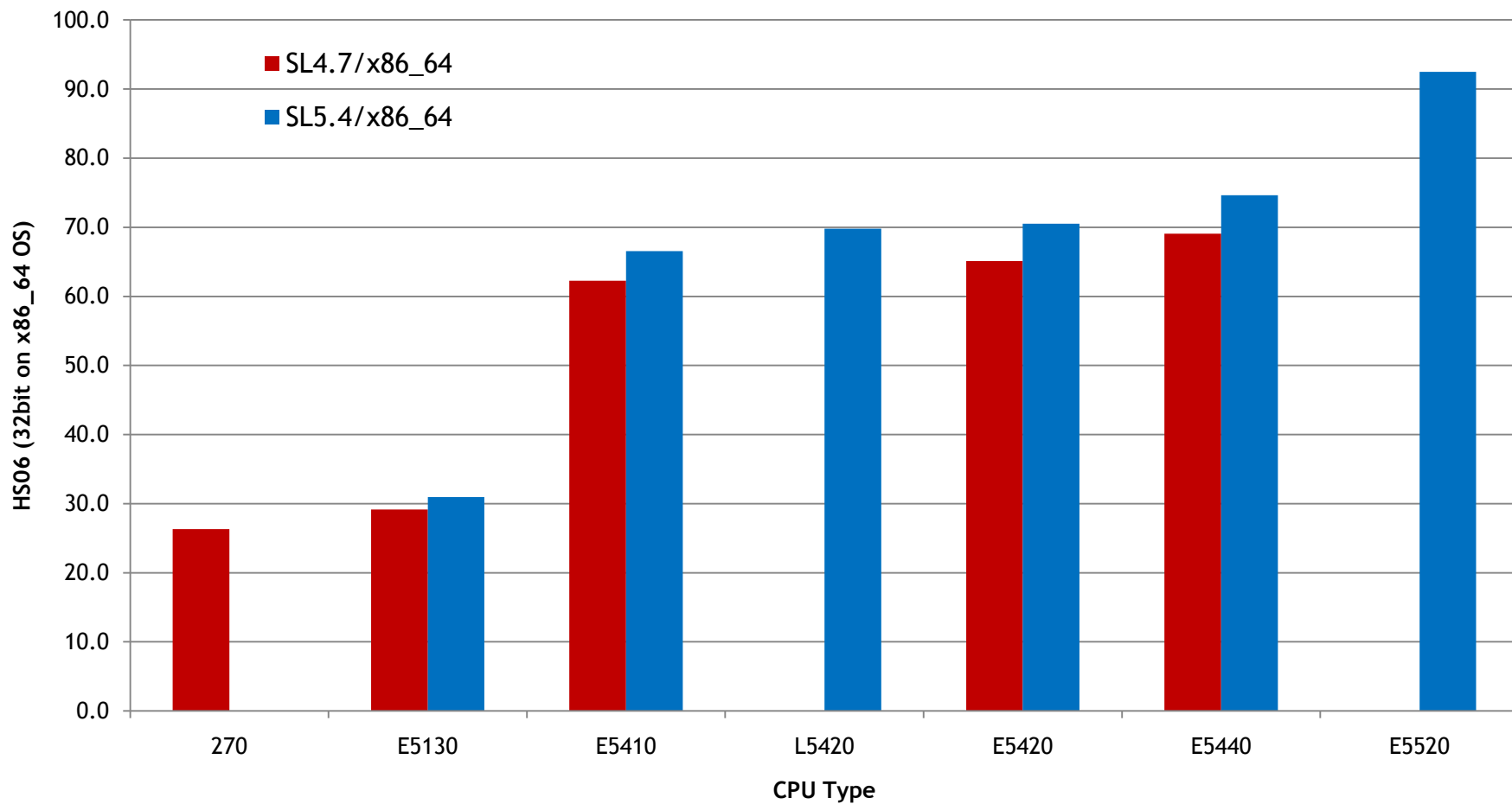


## Failure Rate Normalised by Installed Base 2009-10



In 2009, 10% of drives failed in storage servers

## HS06 Scores - RAL Worker Nodes



## Deployment with QUATTOR



### Already Deployed:

- Worker nodes, Batch server, UK-BDII, LHCb VO box

### In Progress now:

- Disk Servers, CEs, LFC, Local BDII

### Plan:

- All system types Quattorised by December 2010
- Final catch up rollout during end of year stop

- SL4 to SL5 migration for batch workers
- FroNTier deployment
- LFC cleanup
- FTS upgrade (to version 2.2.3)
- glxec deployment, SCAS/pilot roles in progress.
- Updates to hardware used for Top-BDII quintet.
- Other improvements in resilience e.g. enabling hot-swappable disks on WMS...

- **Done:**
  - Nameserver upgrade to 2.1.8-3 (September)
  - SRM update to 2.8-1 (October)
  - Database disk array problem and subsequent rewinding of database. (October)
  - Oracle database quarterly patching (November)
  - Move back to original disk arrays (January)
  - Update memory in database nodes (February)
- **To come:**
  - New database architecture and hardware
  - Updates to Castor (2.1.8/2.1.9)

