

# GridPP

UK Computing for Particle Physics

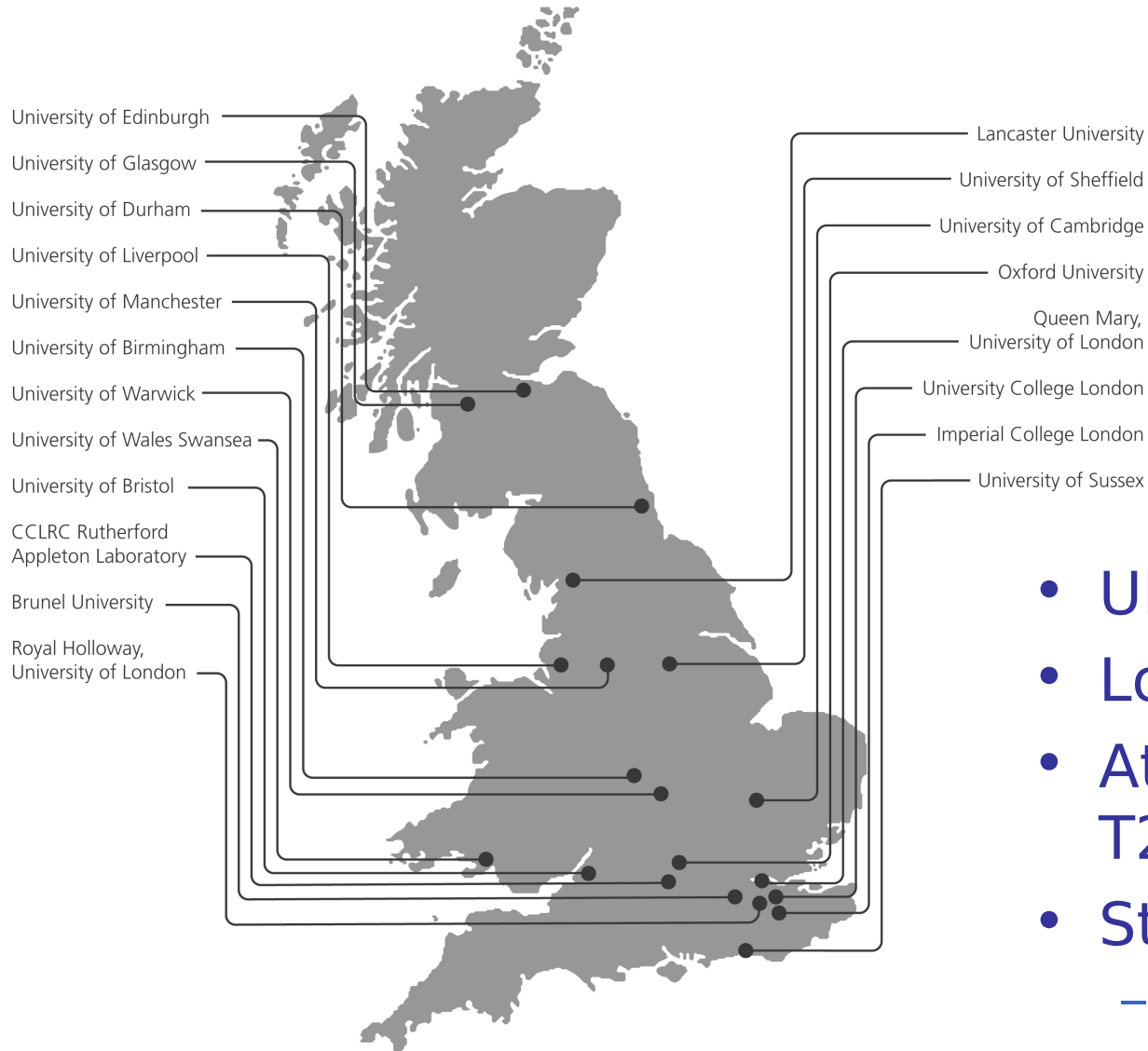
## Queen Mary, University of London (QMUL) Site report

Christopher J. Walker, Daniel Traynor



Queen Mary  
University of London

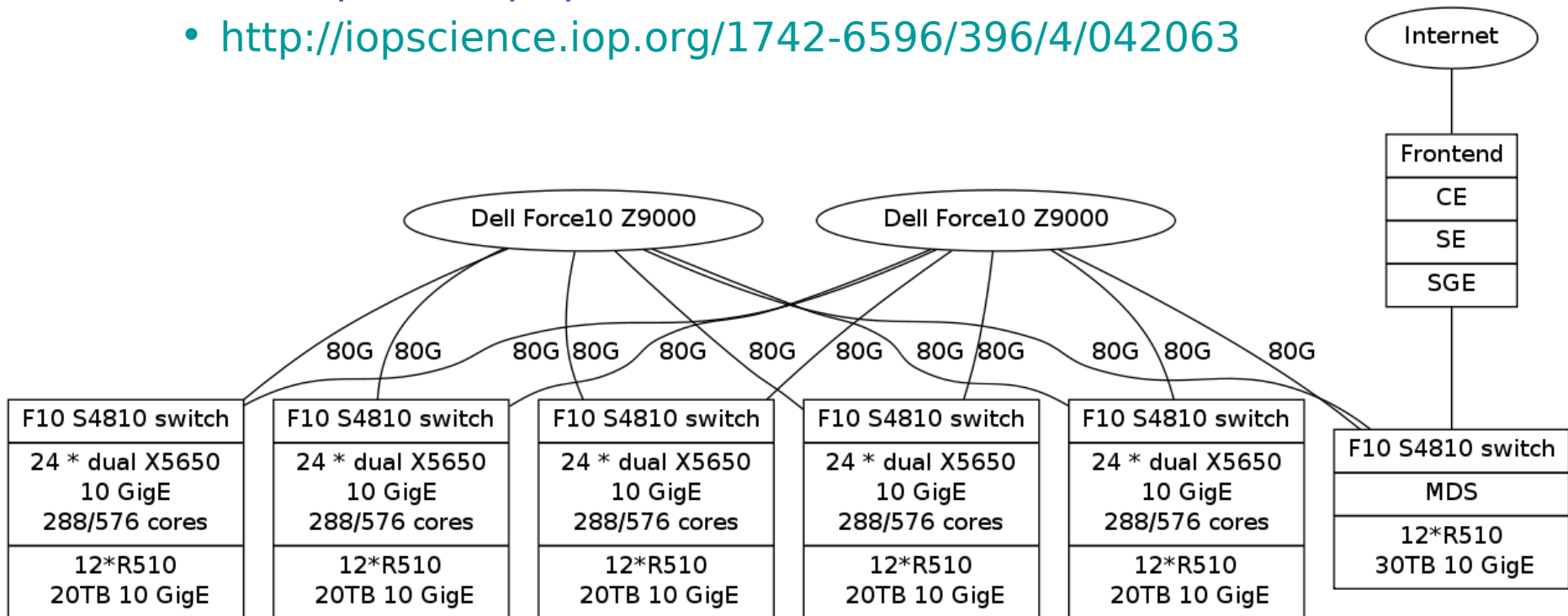
- QMUL's place in the world
  - London tier-2 / Atlas (T2K, SNO+, SuperB, CERN@School)
  - StoRM/Lustre
- Hardware
  - Computers and network topology
- Current issues
  - Vmem limits
  - Lustre >1.8 bugs preventing upgrade
- Conclusions



- UKI-LT2-QMUL
- London Tier-2
- Atlas, SNO+, T2K site
- StoRM
  - Lustre
  - X - Anney

- 3500 job slots
- 1.8 PB Disk
- Hardware (as last year – procurement this year)
  - CPU:
    - C6100: 100 \* Dual X5650 (24/48 Gig RAM, HT on) 10GigE
    - 64\* E5420 (2\*GigE)
  - Disk: (All 10GigE connected) – Lustre OSS
    - 60 \* Dell R510 (12\* 2 TB SATA)
    - 16 \* Dell R510 (12\*3TB NL-SAS)
    - 12 \* Dell PE1950 +2\* MD3000 (Legacy)
  - Network
    - LAN: 2\*z9000, 6\* S4810, 2\* Dell 8024 + Gbit switches
    - WAN: 10Gbit ( college resilient link)

- Mix CPU / Storage in rack
  - Balance power load
  - Balance network
  - See Chep 2012 paper
    - <http://iopscience.iop.org/1742-6596/396/4/042063>



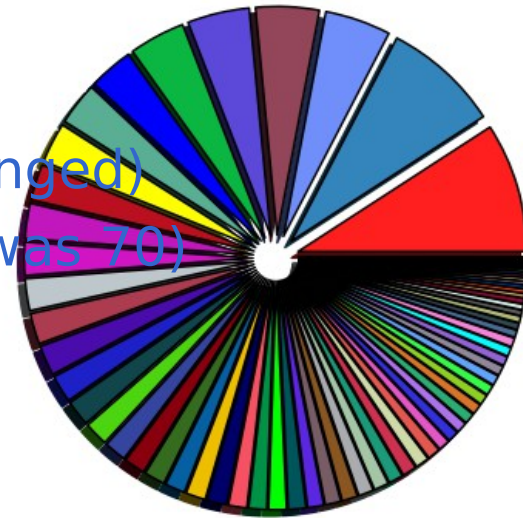
- All EMI or UMD-3
- 4 CEs
  - 2\* VM (Ce05, ce08),
  - 2\*E5460 (ce06,ce07)
  - Multicore (Gridengine)
- 4 SEs (Inc WebDAV)
  - SE03 10Gbit/s – production SE +GridFTP
  - SE04 10Gbit/s – Production GridFTP (Jumbo frames, rollout testing).
  - SE01 (VM)– Staged rollout
  - SE02 (VM)- IPv6 testing
- Argus (VM)
  - 2 Hz!!!
  - CEs + glexec
- Xrootd (VM)

- 2 Hz average !!!
- CEs (Cream)
  - OK
- SE (StoRM)
  - Not yet
    - Memory leak – now probably found (GGUS 104885)
    - Only StoRM, not GridFTP (yet) GGUS 96374
- Xrootd
  - Not yet
    - Requires packages from multiple potentially incompatible repositories

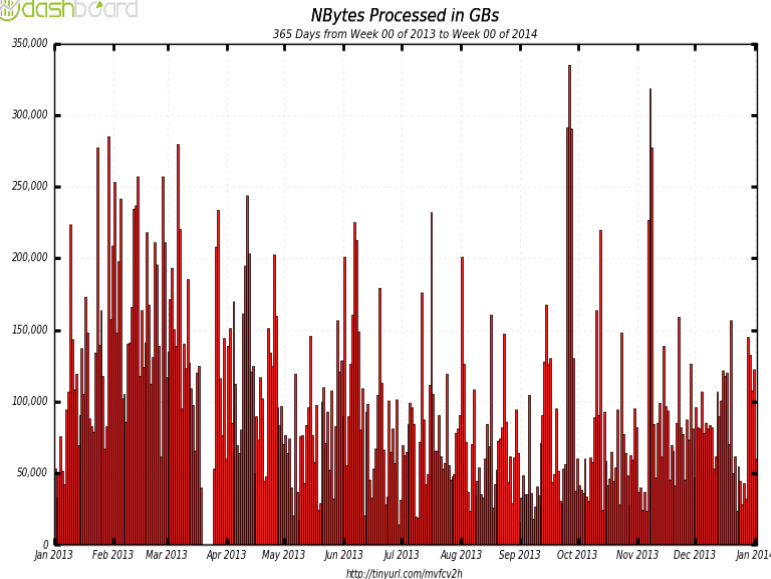
- **Processed (2013)**

- 35PB of data (was 25PB) (4.4% - previously 5.25% of Atlas data processed at Tier-2)
- 26 billion events (was 38)
- 30 million files (unchanged)
- 59 billion CPU seconds (unchanged)
- 64 billion wall clock seconds (was 70)

NBytes Processed in GBs (Sum: 808,320,782)



<http://tinyurl.com/mcw2nbs>

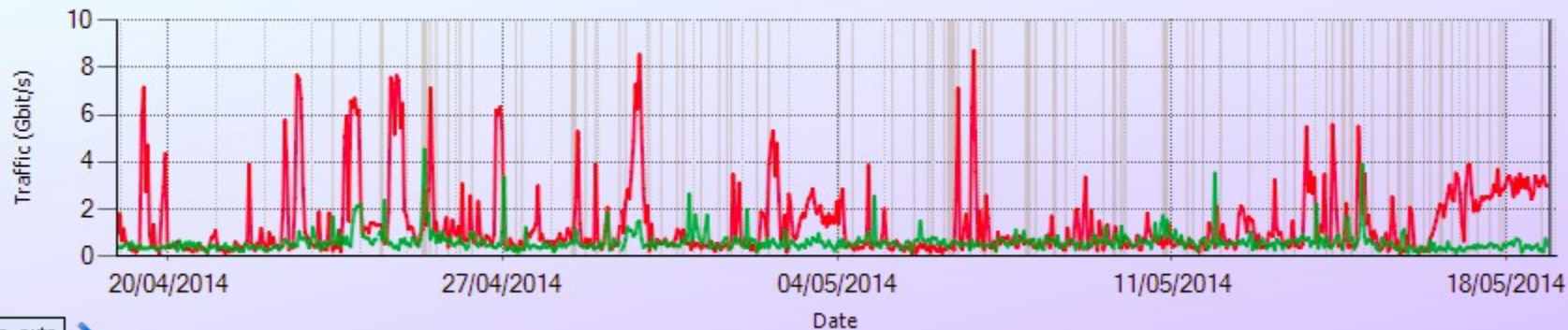


3)  
87,317  
393,198  
221)  
444)  
114,280)  
% (18,163,539)  
% (17,084,486)  
12,116)  
478)

■ MWT2 - 8.22% (66,477,818)  
■ WT2 - 4.47% (36,144,744)  
■ SWT2\_CPB - 3.95% (31,917,760)  
■ IN2P3-CC-T2 - 3.03% (24,517,419)  
■ LRZ-LMU - 2.55% (20,573,970)  
■ TW-FTT - 2.33% (18,819,608)  
■ GRIF-IRFU - 2.17% (17,510,396)  
■ UKI-NORTHGRID-MAN-HEP - 1.93% (15,614,477)  
■ PRAGUELCG2 - 1.70% (13,746,966)  
plus 65 more



Queen Mary and Westfield College, Mile End Road



Apr/May 2014

- 10Gbit/s capacity
  - Capable of reaching 10Gbit/s for short periods
  - Sufficient at the moment
    - Unless ATLAS plan significant increases next year...
- Janet link to college
  - Will be replaced with dark fibre
    - Upgrades should take 3 weeks, not 3 years

- Motivation
  - Local users “Interactive” analysis on Tier-2 disk
  - Non LHC VOs
- XrootD
  - Atlas only (test VM)
- WebDAV
  - EMI-3 StoRM production soon.
  - “Standard” clients (but X509 support limited)
  - RAL LFC now supports WebDAV
- GfalFS
  - Interesting – lots of bugs getting fixed
- Most other UK sites support/Can support these

- Job memory requirements
  - Passed to batch system as `s_vmem`
  - Good idea in principle
  - Problems in practice
- Jobs Killed with SIGXCPU
  - SGE uses this to indicate exceeding memory requirements
- GGUS #103028
  - Memory requirements 3500M/4000M
  - Use `vmPeakMax=1613708` `vmPeakMean=159099`
    - Still killed
  - Is this the job wrapper?

- Currently running 1.8.9
  - Patched to cope with OSS failure (see Jira)
- Anticipate – 2.5/2.6 with new procurement
  - How to copy 1.5 PB efficiently
  - LU-1482
    - Only owner can write extended attributes (checksum)
    - StoRM owns the file, VO write permission with ACLs

- Procurement
  - Storage
  - Compute
- New deployment system
  - Puppet +???
- Cloud
  - Cloudstack testing
  - Can we use Lustre as backend storage?
    - EG NFS export Lustre

- Network
  - 10Gig network performing well
    - No immediate need to upgrade
- StoRM
  - Performing well
- Lustre
  - Performing well
    - Challenges for upgrade