

Sheffield Computing

Matt Robinson

Elena Korolkova

Paul Hodgson

Tier-3

- ◆ Zero gridpp funding has gone into the Tier-3 cluster. It is entirely funded by the PPPA groups at Sheffield.
- ◆ Currently stands at 130 CPUs and ~50 TB of disk. Mostly built in-house.
- ◆ Running 32 bit SL4 until ATLAS SW stops supporting it.

Tier-3 primary server

- ◆ hep0:
 - ◆ 4 Opteron Cores, 8 GB RAM
 - ◆ Upgrade to 12 Cores & 32 GB at Xmas
 - ◆ ssh gateway (bastion-esque)
 - ◆ Account Server, Users' home areas

Tier-3 NFS Disk Servers

- ◆ Mostly in 3 (2 atlas, 1 general purpose) 13 TB (formatted) raid-5 arrays (mdadm, sata_mv). Also some older smaller servers.
- ◆ 3 13 TB machines each have 6 gigabit links split between groups of worker nodes.

Tier-3 desktops

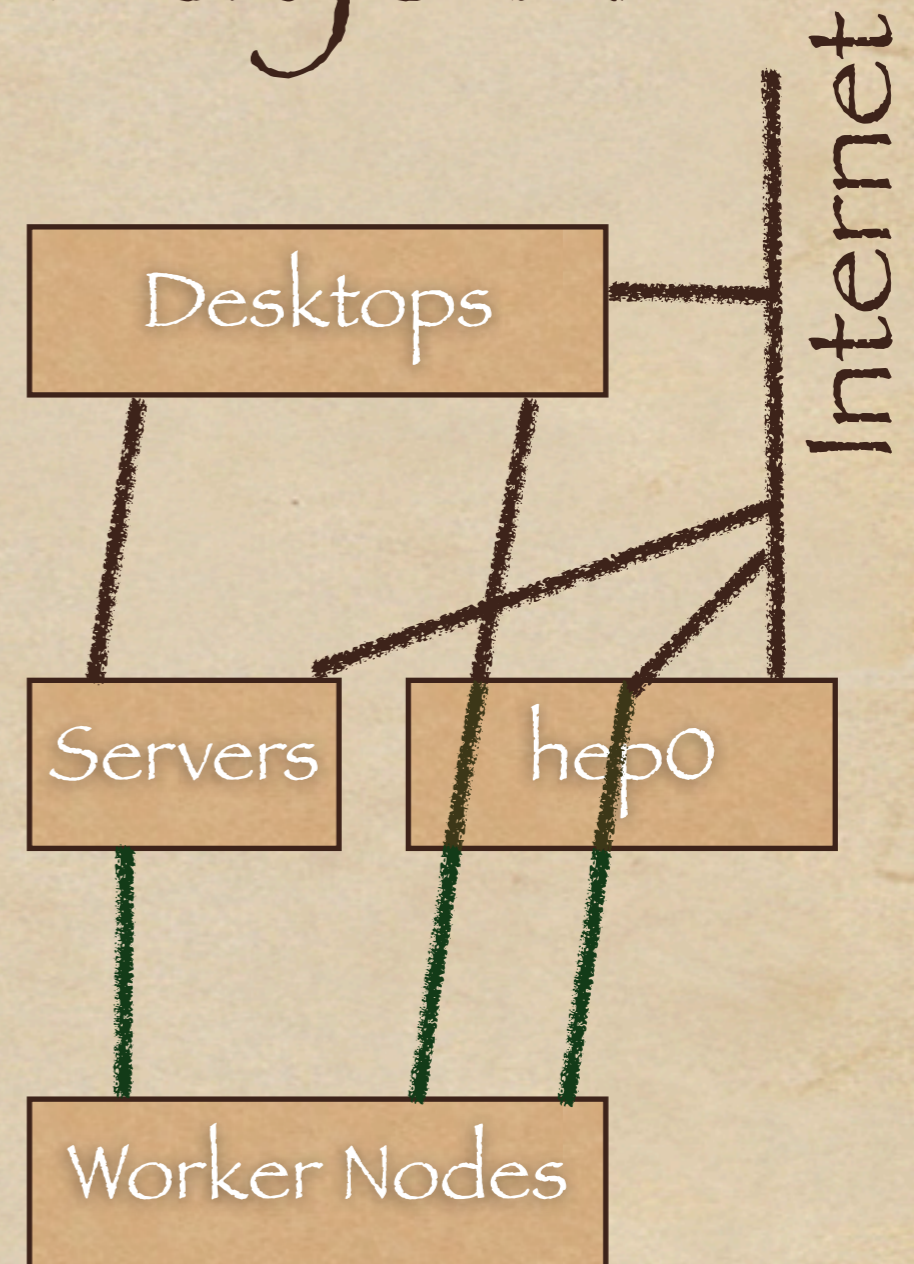
- ◆ About 40 machines, used by Academics RA's and PhD students.
- ◆ Same shared file-systems as servers and worker nodes
- ◆ Capable of direct job submission.
- ◆ Stuck with 100Mb University network.

Tier-3 worker nodes

- ◆ 50 machines. Networked in DMZ.
- ◆ Mixture of 32 bit and 64 bit. Gradually phasing out 32 bit machines in time for SL5 upgrade. 64 bit have 2 GB/core.
- ◆ Continuous HW upgrades as demand requires and funding allows.

Tier-3 Network Layout

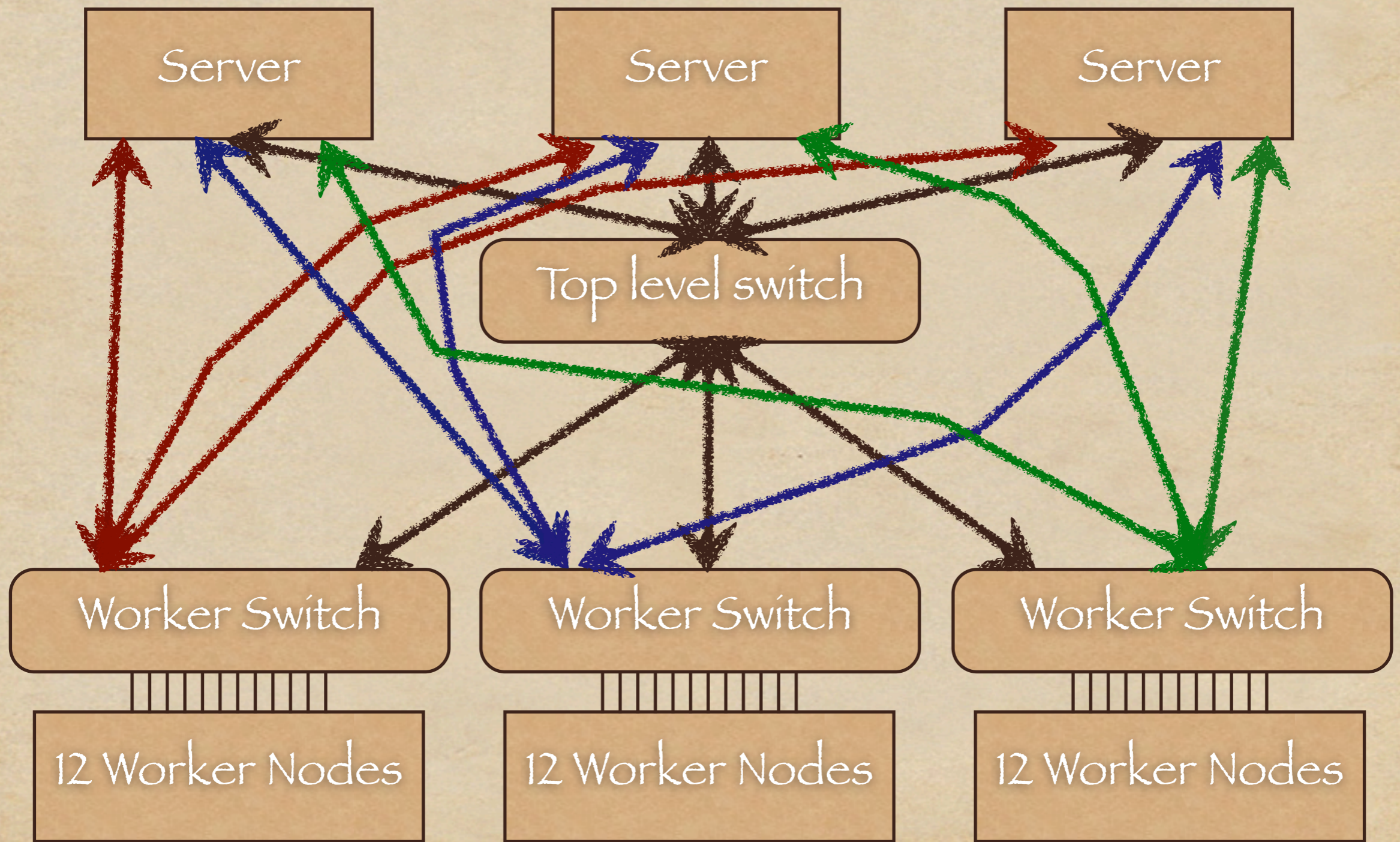
- ◆ Servers attached to DMZ and internet.
- ◆ Desktops attached to internet
- ◆ Workers attached to DMZ.



Tier-3 Network

- ◆ Workers and servers in dedicated machine room together, so many networking options.
- ◆ Multiple interfaces in each server so that data can bypass network bottlenecks
- ◆ Routing rules on disk servers enforce it.

Tier-3 Network Layout



Tier-3 Machine Room



Tier-3 Software

- ◆ Every machine configured as LCG UI.
- ◆ AFS on everything.
- ◆ About 10 ATLAS offline releases installed at any one time. Plus production caches.
- ◆ Any software the users want, they get.

Tier-3 P.S.

- ◆ NFS access to web server content from whole cluster.
- ◆ Supports many diverse experiments but choice of Linux version determined by ATLAS.
- ◆ No virtualisation.

Tier-3 Status

- ◆ Ganglia front page yesterday morning

<http://www.hep.shef.ac.uk/ganglia/>

Ganglia HEP Grid Report for Mon, 29 Jun 2009 12:26:56 +0100 [Get Fresh Data](#)

Last Sorted

HEP Grid > [Show Queues](#)

Monitoring

Running Jobs: **106**
 Pending Jobs: **0**
 Active CPUs: **106**
 Free CPUs: **24**

Temperature: **20.56** °C
 Alarms: **Enabled**

HEP Grid (3 sources) (tree view)

CPU's Total: **209**
 Hosts up: **90**
 Hosts down: **0**

Avg Load (15, 5, 1m): 52%, 56%, 57%
 Localtime: 2009-06-29 12:26

Workers (physical view)

CPU's Total: **130**
 Hosts up: **50**
 Hosts down: **0**

Avg Load (15, 5, 1m): 75%, 80%, 81%
 Localtime: 2009-06-29 12:26

Servers (physical view)

CPU's Total: **22**
 Hosts up: **7**
 Hosts down: **0**

Avg Load (15, 5, 1m): 29%, 30%, 32%
 Localtime: 2009-06-29 12:26

Desktops (physical view)

CPU's Total: **57**
 Hosts up: **33**
 Hosts down: **0**

Avg Load (15, 5, 1m): 9%, 10%, 12%
 Localtime: 2009-06-29 12:26

Snapshot of the HEP Grid | Legend

Workers Servers Desktops

Laptops

- ◆ Recommending Macs these days.
- ◆ No specific policy. No rules. About 50.
- ◆ New PC laptops are usually upgraded from Vista to XP and then dual booted with SL5 (custom kernel) by Muggins here. Takes about 1 day, but never had one yet that didn't work like this.

Tier-3 Upgrade plan

- ◆ Need to retire 32 bit desktops and workstations over the next 12 months, give or take.
- ◆ Building worker nodes mostly in ATX midi towers on shelves out of Phenoms.
- ◆ May put bigger disks in the disk servers.

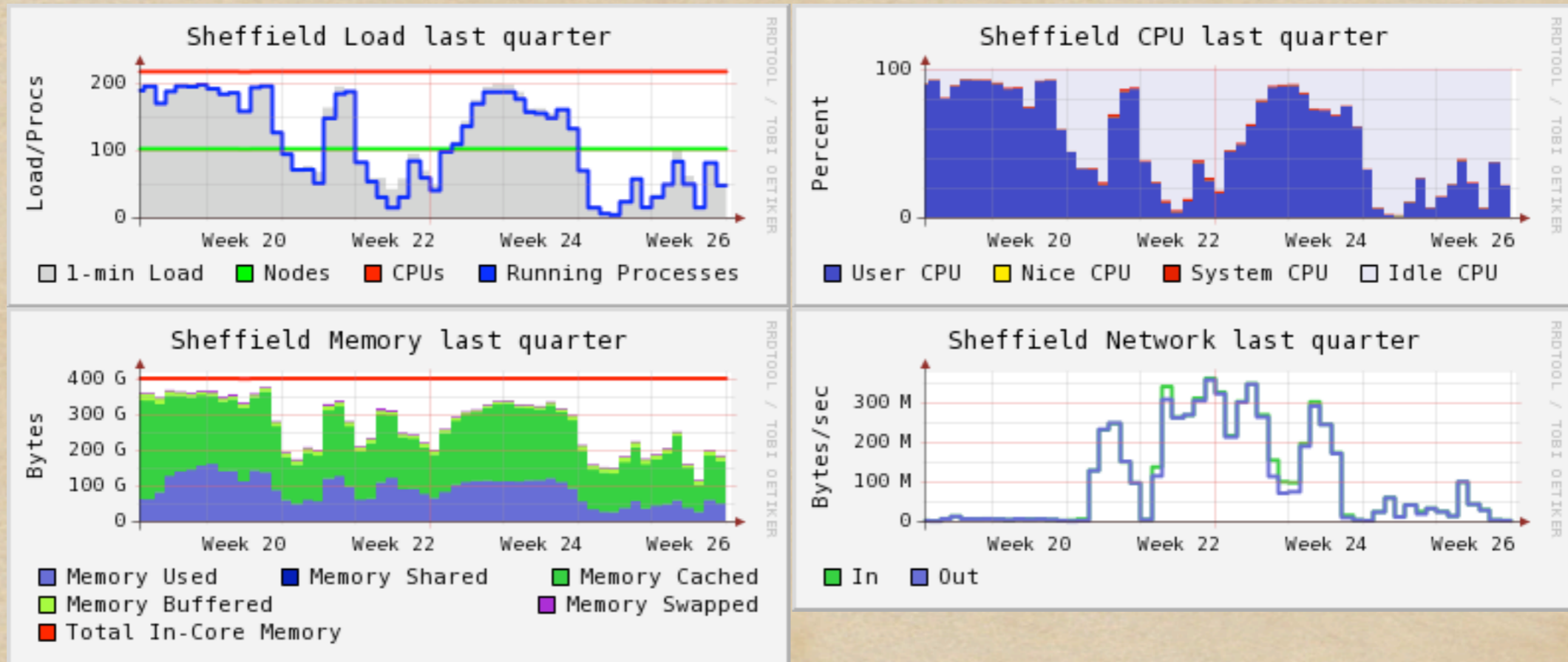
Tier-2 Machine Room



Tier-2

- ◆ 200 Single Core 2.4 GHz Opterons (2 GB/core) in 100 1u boxes. 64 bit SL4.
- ◆ PXE - Kickstart - Bash
- ◆ No resources shared with Tier-3
- ◆ 3 13 TB (formatted) raid-5 (mdadm) disk servers (DPM).
- ◆ Shared gigabit networking.

Tier-2 Monitoring



<http://lcg.shef.ac.uk/ganglia/>

Tier-2 Performance

- ◆ ~97% in Standard tests
- ◆ SI/Core still reasonable despite age
- ◆ Occasional “Thermal Events”
- ◆ Losing RAM, disks and PSU’s (irreplaceable), but spare machines coming available.

Obligatory: SAM Tests

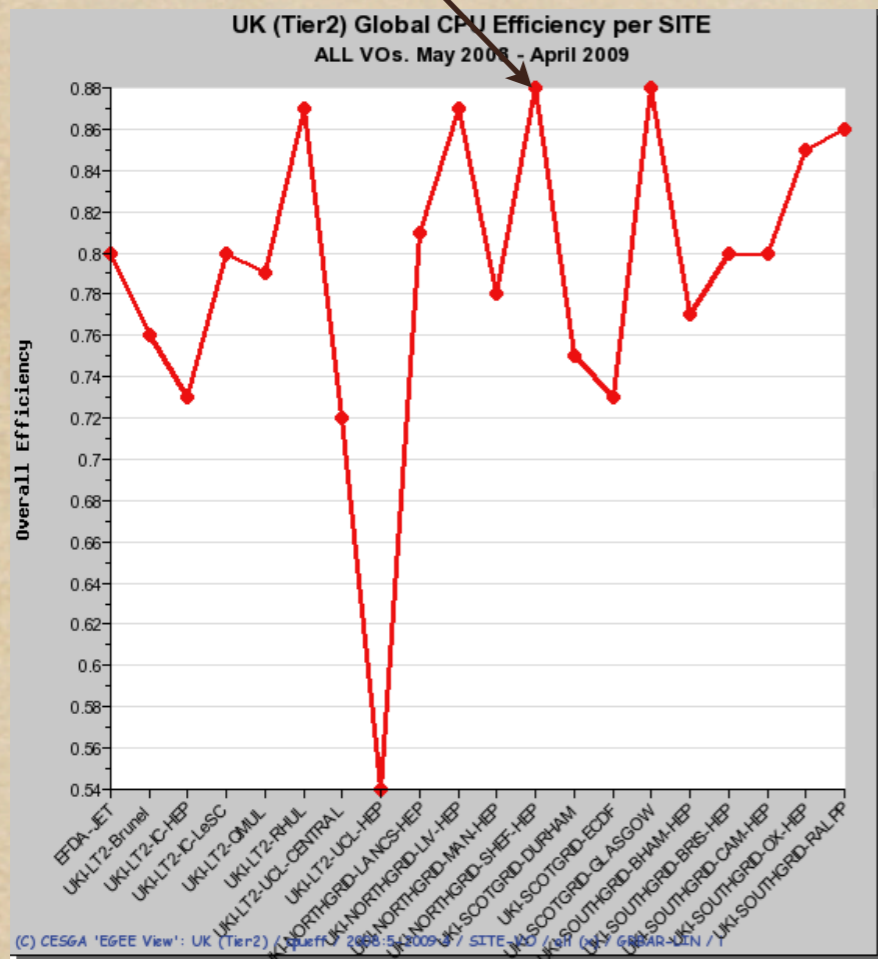
Site	1Q08	2Q08	3Q08	4Q08	1Q09	2Q09
EFDA-JET	92%	70%	91%	77%	96%	96%
RAL-LCG2_Tier-1	94%	97%	99%	98%	98%	92%
UKI-LT2-Brunel	84%	92%	84%	92%	100%	100%
UKI-LT2-IC-HEP	95%	92%	89%	95%	82%	88%
UKI-LT2-IC-LeSC	95%	91%	93%	85%	87%	85%
UKI-LT2-QMUL	41%	53%	94%	94%	96%	98%
UKI-LT2-RHUL	84%	96%	93%	90%	100%	99%
UKI-LT2-UCL-CENTRAL	50%	0%	7%	42%	68%	37%
UKI-LT2-UCL-HEP	77%	82%	77%	86%	68%	99%
UKI-NORTHGRID-LANCS-HEP	89%	94%	96%	83%	96%	100%
UKI-NORTHGRID-LIV-HEP	70%	97%	93%	98%	98%	99%
UKI-NORTHGRID-MAN-HEP	98%	87%	99%	92%	93%	96%
UKI-NORTHGRID-SHEF-HEP	98%	96%	96%	97%	97%	98%
UKI-SCOTGRID-DURHAM	94%	85%	95%	84%	91%	97%
UKI-SCOTGRID-ECDF	66%	72%	83%	99%	96%	93%
UKI-SCOTGRID-GLASGOW	85%	97%	96%	98%	97%	99%
UKI-SOUTHGRID-BHAM-HEP	87%	97%	97%	99%	98%	97%
UKI-SOUTHGRID-BRIS-HEP	99%	98%	98%	96%	98%	100%
UKI-SOUTHGRID-CAM-HEP	93%	85%	84%	90%	98%	95%
UKI-SOUTHGRID-OX-HEP	99%	90%	93%	88%	85%	99%
UKI-SOUTHGRID-RALPP	92%	91%	96%	95%	94%	92%
csTCDie	97%	92%	94%	89%	95%	89%
LondonGrid	75%	72%	79%	84%	86%	86%
NorthGrid	89%	94%	96%	93%	96%	98%
ScotGrid	82%	85%	92%	94%	97%	96%
SouthGrid	94%	92%	93%	94%	95%	96%

Hours ago:	Availability			
	24 Hrs	Week	Month	6 Mon
EFDA-JET	100%	96%	99%	96%
RAL-LCG2_Tier-1	0%	40%	86%	96%
UKI-LT2-Brunel	100%	100%	100%	100%
UKI-LT2-IC-HEP	0%	32%	83%	85%
UKI-LT2-IC-LeSC	0%	62%	91%	86%
UKI-LT2-QMUL	100%	100%	100%	97%
UKI-LT2-RHUL	100%	100%	100%	98%
UKI-LT2-UCL-CENTRAL	0%	0%	13%	53%
UKI-LT2-UCL-HEP	100%	99%	100%	83%
UKI-NORTHGRID-LANCS-HEP	100%	100%	100%	98%
UKI-NORTHGRID-LIV-HEP	100%	100%	99%	98%
UKI-NORTHGRID-MAN-HEP	100%	100%	99%	95%
UKI-NORTHGRID-SHEF-HEP	100%	100%	98%	98%
UKI-SCOTGRID-DURHAM	100%	100%	98%	97%
UKI-SCOTGRID-ECDF	100%	100%	99%	95%
UKI-SCOTGRID-GLASGOW	100%	100%	100%	98%
UKI-SOUTHGRID-BHAM-HEP	0%	63%	91%	97%
UKI-SOUTHGRID-BRIS-HEP	100%	100%	100%	99%
UKI-SOUTHGRID-CAM-HEP	100%	100%	95%	96%
UKI-SOUTHGRID-OX-HEP	100%	100%	100%	91%
UKI-SOUTHGRID-RALPP	100%	98%	93%	93%
csTCDie	100%	100%	87%	92%
Total	77%	86%	92%	93%

Tier-2 Productivity

Efficiency

Atlas Production



site	defined	assigned	waiting	activated	running	holding	transferring	success	failure	efficiency
RAL-LCG2	0	0	0	390	10	296	0	691428	103926	86.9%
UKI-SCOTGRID-GLASGOW	0	0	0	56	1	3	0	167941	19028	89.8%
UKI-LT2-QMUL	0	0	0	43	0	1	0	90153	13574	86.9%
UKI-NORTHGRID-MAN-HEP	0	0	0	42	0	3	0	32456	69114	32%
UKI-NORTHGRID-LANCS-HEP	0	0	0	41	0	5	0	67887	14324	82.6%
UKI-LT2-IC-HEP	0	0	0	60	0	10	0	65697	10964	85.7%
UKI-LT2-RHUL	0	0	0	9	1	7	0	60402	4542	93%
UKI-NORTHGRID-LIV-HEP	0	1	0	1124	1	5	2	47088	4207	91.8%
UKI-SCOTGRID-ECDF	0	0	0	29	1	2	0	32109	3399	90.4%
UKI-SOUTHGRID-CAM-HEP	0	1	0	48	1	0	2	31896	2962	91.5%
UKI-NORTHGRID-SHEF-HEP	0	0	0	120	0	3	0	30899	2151	93.5%
UKI-SOUTHGRID-RALPP	0	2	0	49	1	5	0	26694	2612	91.1%
UKI-SOUTHGRID-OX-HEP	0	0	0	4	0	0	0	19490	7279	72.8%
UKI-SCOTGRID-DURHAM	0	0	0	1	0	1	0	16107	1437	91.8%
UKI-LT2-Brunel	0	0	0	13	0	3	17	9987	691	93.5%
UKI-SOUTHGRID-BHAM-HEP	0	0	0	0	0	0	0	1299	65	95.2%
UKI-LT2-UCL-HEP	0	0	0	0	0	0	0	0	205	0%
total	0	4	0	2029	16	344	21	1391533	260480	84.2%

94%

Tier-2 upgrade plan

- ◆ CPU's to remain stable at 200. Enough spares around to achieve this.
- ◆ 3 13 TB disk servers to be upgraded to 26 TB by swapping out 1 TB disks for 2 TB. Probably in Autumn.
- ◆ Probably need to throw out Workers in 12-18 months.