Monitoring the Grid at local, national, and global levels

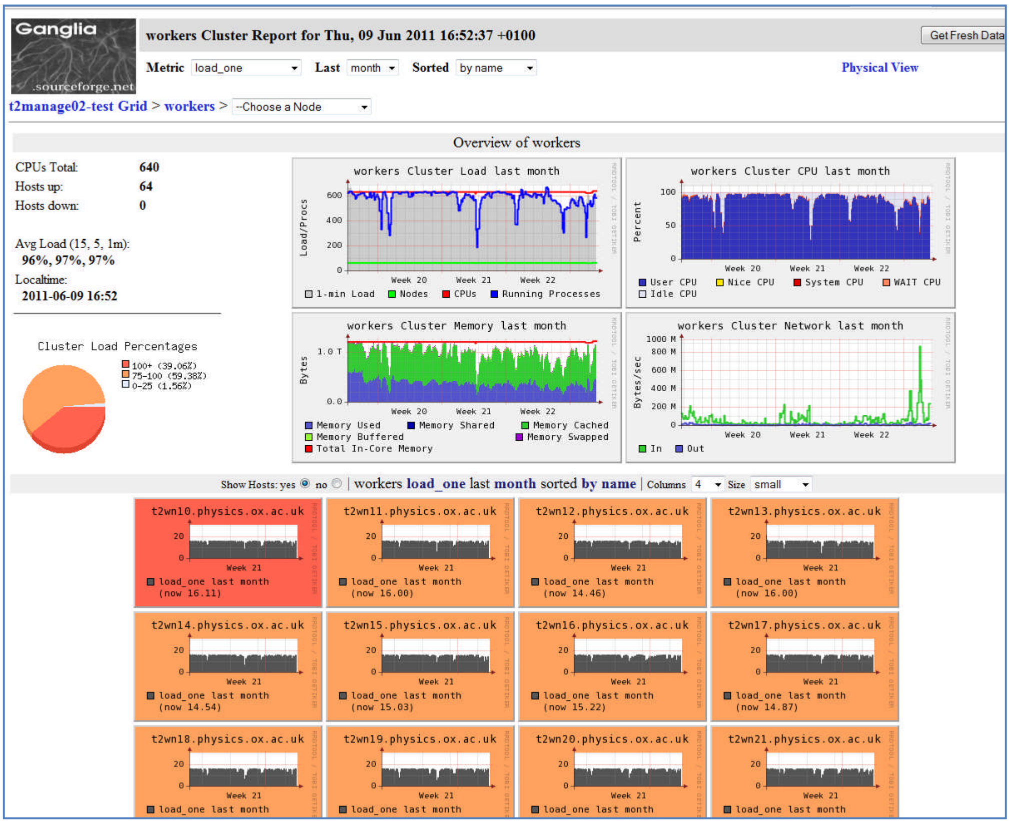
P D Gronbech [1]
1 Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

Email: p.gronbech1@physics.ox.ac.uk

Abstract. The World-wide LHC Computing Grid is a global infrastructure set up to process the experimental data from the experiments at the Large Hadron Collider located at CERN. The UK component is provided by the GridPP project across 19 sites at the universities and Rutherford Lab. To ensure that these large computational resources are available and reliable requires many different monitoring systems, ranging from local site monitoring of individual components, through UK-wide monitoring of Grid functionality, to the worldwide monitoring of resource provision and usage. In this paper we describe the monitoring systems used for the many different aspects of the system, and how some of them are being integrated together.

1. Local site monitoring

Local GridPP [1] site monitoring covers cluster load, batch system utilization, network bandwidth monitoring and fault condition monitoring. Ganglia [2] is the most common software used to monitor a cluster figure 1. It is easily installed on clients and allows data to be collected on a master node and displayed via a web server.



workers Cluster Report for Thu, 09 Jun 2011 16:52:37 +0100

Metric load_one Last month Sorted by name

t2manage02-test Grid > workers > --Choose a Node

Overview of workers

CPUs Total:	640
Hosts up:	64
Hosts down:	0

Avg Load (15, 5, 1m):
96%, 97%, 97%

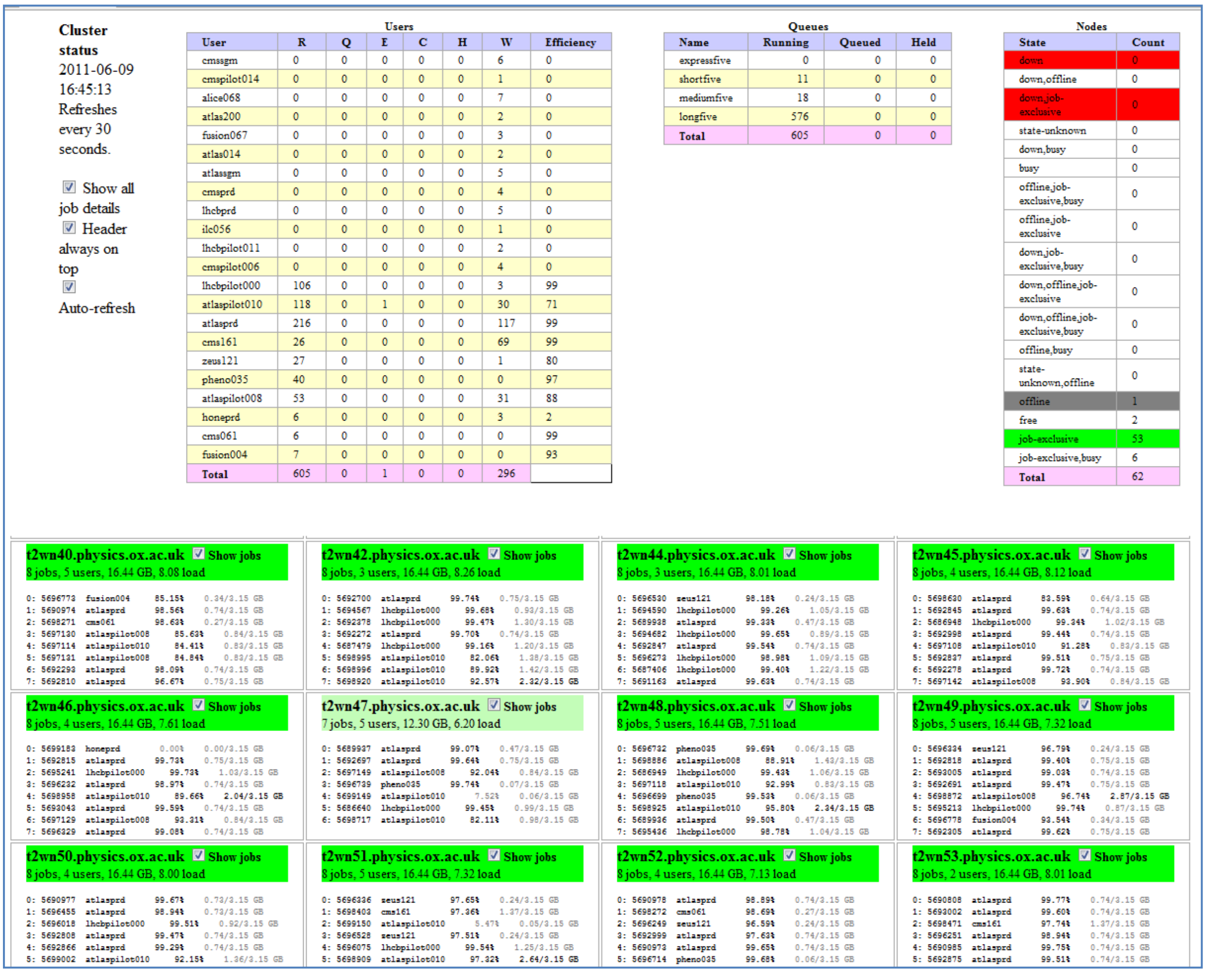
Localtime:
2011-06-09 16:52

Show Hosts: yes ● no ○ | workers load_one last month sorted by name | Columns 4 Size small

Figure 1. Ganglia monitoring showing a group of Worker Nodes.

Ganglia monitors and collects many parameters, such as load, memory used, network activity etc. The top part of the page shown in figure 1 shows an aggregate picture across the nodes in a particular logical group, such as worker nodes where the main priority is to ensuring they are busy. The web interface allows other aggregate parameters to be selected or the full details for any individual node can be displayed. The data are recorded, allowing time periods from an hour up to a year to be displayed. This can be extremely useful for spotting trends over time. Other parameters can be added to the Ganglia monitoring setup, but many sites are satisfied by the default setup.

Monitoring specific to the batch system used at a site is also typically used. Many GridPP sites use the torque batch system [3] (developed from PBS) monitored with pbswebmon [4], which provides a graphical way to monitor the occupancy of the cluster, and the job shares and efficiencies for each user figure 2. This package clearly shows which users jobs are running, something the ganglia programme does not do without modification. The efficiency of the job is also reported which can guide a systems administrator to investigate jobs which may be suffering through lack of I/O bandwidth to the disks or the network.



Cluster status
2011-06-09 16:45:13
Refreshes every 30 seconds.

☑ Show all job details
☑ Header always on top
☑ Auto-refresh

Users

User	R	Q	E	C	H	W	Efficiency
cmssgm	0	0	0	0	0	6	0
cmspilot014	0	0	0	0	0	1	0
alice068	0	0	0	0	0	7	0
atlas200	0	0	0	0	0	2	0
fusion067	0	0	0	0	0	3	0
atlas014	0	0	0	0	0	2	0
atlassgm	0	0	0	0	0	5	0
cmsprd	0	0	0	0	0	4	0
lhcbprd	0	0	0	0	0	5	0
ilc056	0	0	0	0	0	1	0
lhcbpilot011	0	0	0	0	0	2	0
cmspilot006	0	0	0	0	0	4	0
lhcbpilot000	106	0	0	0	0	3	99
atlaspilot010	118	0	1	0	0	30	71
atlasprd	216	0	0	0	0	117	99
cms161	26	0	0	0	0	69	99
zeus121	27	0	0	0	0	1	80
pheno035	40	0	0	0	0	0	97
atlaspilot008	53	0	0	0	0	31	88
honeprd	6	0	0	0	0	3	2
cms061	6	0	0	0	0	0	99
fusion004	7	0	0	0	0	0	93
Total	605	0	1	0	0	296	

Queues

Name	Running	Queued	Held
expressfive	0	0	0
shortfive	11	0	0
mediumfive	18	0	0
longfive	576	0	0
Total	605	0	0

Nodes

State	Count
down	0
down,offline	0
down,job-exclusive	0
state-unknown	0
down,busy	0
busy	0
offline,job-exclusive,busy	0
offline,job-exclusive	0
down,job-exclusive,busy	0
down,offline,job-exclusive	0
down,offline,job-exclusive,busy	0
offline,busy	0
state-unknown,offline	0
offline	1
free	2
job-exclusive	53
job-exclusive,busy	6
Total	62

t2wn40.physics.ox.ac.uk ☑ Show jobs
8 jobs, 5 users, 16.44 GB, 8.08 load

```
0: 5696773  fusion004      85.15%    0.34/3.15 GB
1: 5690974  atlasprd       98.56%    0.74/3.15 GB
2: 5698271  cms061         98.63%    0.27/3.15 GB
3: 5697130  atlaspilot008      85.63%    0.84/3.15 GB
4: 5697114  atlaspilot010      84.41%    0.83/3.15 GB
5: 5697131  atlaspilot008      84.84%    0.83/3.15 GB
6: 5692293  atlasprd       98.09%    0.74/3.15 GB
7: 5692810  atlasprd       96.67%    0.75/3.15 GB
```

t2wn42.physics.ox.ac.uk ☑ Show jobs
8 jobs, 3 users, 16.44 GB, 8.26 load

```
0: 5692700  atlasprd       99.74%    0.75/3.15 GB
1: 5694567  lhcbpilot000      99.68%    0.93/3.15 GB
2: 5692378  lhcbpilot000      99.47%    1.30/3.15 GB
3: 5692272  atlasprd       99.70%    0.74/3.15 GB
4: 5687479  lhcbpilot000      99.16%    1.20/3.15 GB
5: 5698995  atlaspilot010      82.06%    1.38/3.15 GB
6: 5698996  atlaspilot010      89.92%    1.42/3.15 GB
7: 5698920  atlaspilot010      92.57%    2.32/3.15 GB
```

t2wn44.physics.ox.ac.uk ☑ Show jobs
8 jobs, 3 users, 16.44 GB, 8.01 load

```
0: 5696530  zeus121        98.18%    0.24/3.15 GB
1: 5694590  lhcbpilot000      99.26%    1.05/3.15 GB
2: 5689938  atlasprd       99.33%    0.47/3.15 GB
3: 5694682  lhcbpilot000      99.65%    0.89/3.15 GB
4: 5692847  atlasprd       99.54%    0.74/3.15 GB
5: 5696273  lhcbpilot000      98.98%    1.09/3.15 GB
6: 5687406  lhcbpilot000      99.40%    1.22/3.15 GB
7: 5691163  atlasprd       99.63%    0.74/3.15 GB
```

t2wn45.physics.ox.ac.uk ☑ Show jobs
8 jobs, 4 users, 16.44 GB, 8.12 load

```
0: 5698630  atlasprd       83.59%    0.64/3.15 GB
1: 5692845  atlasprd       99.63%    0.74/3.15 GB
2: 5686948  lhcbpilot000      99.34%    1.02/3.15 GB
3: 5692998  atlasprd       99.44%    0.74/3.15 GB
4: 5697108  atlaspilot010      91.28%    0.83/3.15 GB
5: 5692837  atlasprd       99.51%    0.75/3.15 GB
6: 5692278  atlasprd       99.72%    0.74/3.15 GB
7: 5697142  atlaspilot008      93.90%    0.84/3.15 GB
```

t2wn46.physics.ox.ac.uk ☑ Show jobs
8 jobs, 4 users, 16.44 GB, 7.61 load

```
0: 5699183  honeprd         0.00%    0.00/3.15 GB
1: 5692815  atlasprd       99.73%    0.75/3.15 GB
2: 5695241  lhcbpilot000      99.73%    1.03/3.15 GB
3: 5696232  atlasprd       98.97%    0.74/3.15 GB
4: 5698958  atlaspilot010      89.66%    2.04/3.15 GB
5: 5693043  atlasprd       99.59%    0.74/3.15 GB
6: 5697129  atlaspilot008      93.31%    0.84/3.15 GB
7: 5696329  atlasprd       99.08%    0.74/3.15 GB
```

t2wn47.physics.ox.ac.uk ☑ Show jobs
7 jobs, 5 users, 12.30 GB, 6.20 load

```
0: 5689937  atlasprd       99.07%    0.47/3.15 GB
1: 5692697  atlasprd       99.64%    0.75/3.15 GB
2: 5697149  atlaspilot008      92.04%    0.84/3.15 GB
3: 5696739  pheno035       99.74%    0.07/3.15 GB
4: 5699149  atlaspilot010       7.52%    0.06/3.15 GB
5: 5686640  lhcbpilot000      99.45%    0.99/3.15 GB
6: 5698717  atlaspilot010      82.11%    0.98/3.15 GB
```

t2wn48.physics.ox.ac.uk ☑ Show jobs
8 jobs, 5 users, 16.44 GB, 7.51 load

```
0: 5696732  pheno035       99.69%    0.06/3.15 GB
1: 5698886  atlaspilot008      88.91%    1.43/3.15 GB
2: 5686949  lhcbpilot000      99.43%    1.06/3.15 GB
3: 5697118  atlaspilot010      92.99%    0.83/3.15 GB
4: 5696699  pheno035       99.53%    0.06/3.15 GB
5: 5698925  atlaspilot010      95.80%    2.34/3.15 GB
6: 5689936  atlasprd       99.50%    0.47/3.15 GB
7: 5695436  lhcbpilot000      98.78%    1.04/3.15 GB
```

t2wn49.physics.ox.ac.uk ☑ Show jobs
8 jobs, 5 users, 16.44 GB, 7.32 load

```
0: 5696334  zeus121        96.79%    0.24/3.15 GB
1: 5692818  atlasprd       99.40%    0.75/3.15 GB
2: 5693005  atlasprd       99.03%    0.74/3.15 GB
3: 5692691  atlasprd       99.47%    0.75/3.15 GB
4: 5698872  atlaspilot008      96.74%    2.87/3.15 GB
5: 5695213  lhcbpilot000      99.74%    0.87/3.15 GB
6: 5696778  fusion004      93.54%    0.34/3.15 GB
7: 5692305  atlasprd       99.62%    0.75/3.15 GB
```

t2wn50.physics.ox.ac.uk ☑ Show jobs
8 jobs, 4 users, 16.44 GB, 8.00 load

```
0: 5690977  atlasprd       99.67%    0.73/3.15 GB
1: 5696455  atlasprd       98.94%    0.73/3.15 GB
2: 5696018  lhcbpilot000      99.51%    0.92/3.15 GB
3: 5692808  atlasprd       99.47%    0.74/3.15 GB
4: 5692866  atlasprd       99.29%    0.74/3.15 GB
5: 5699002  atlaspilot010      92.15%    1.36/3.15 GB
```

t2wn51.physics.ox.ac.uk ☑ Show jobs
8 jobs, 5 users, 16.44 GB, 7.32 load

```
0: 5696336  zeus121        97.65%    0.24/3.15 GB
1: 5698403  cms161         97.36%    1.37/3.15 GB
2: 5699150  atlaspilot010       5.47%    0.05/3.15 GB
3: 5696528  zeus121        97.51%    0.24/3.15 GB
4: 5696075  lhcbpilot000      99.54%    1.25/3.15 GB
5: 5698909  atlaspilot010      97.32%    2.64/3.15 GB
```

t2wn52.physics.ox.ac.uk ☑ Show jobs
8 jobs, 4 users, 16.44 GB, 7.13 load

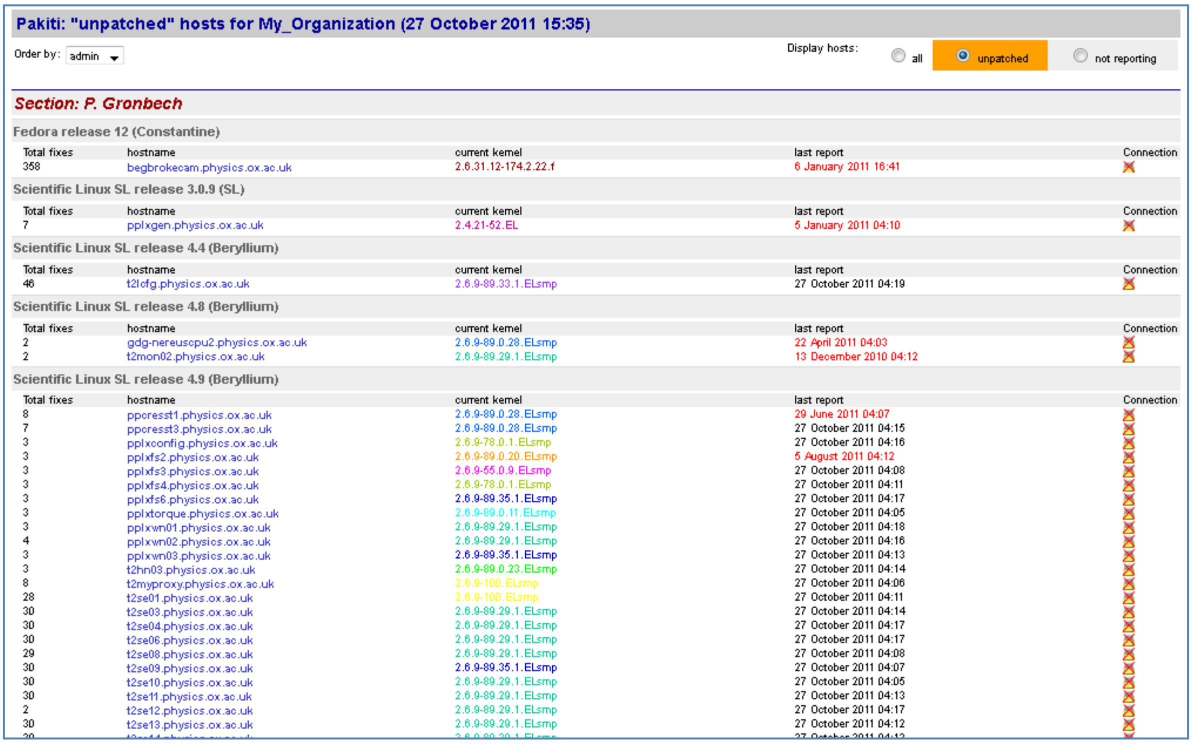
```
0: 5690978  atlasprd       98.89%    0.74/3.15 GB
1: 5698272  cms061         98.69%    0.27/3.15 GB
2: 5696249  zeus121        96.59%    0.24/3.15 GB
3: 5692999  atlasprd       97.63%    0.74/3.15 GB
4: 5690973  atlasprd       99.65%    0.74/3.15 GB
5: 5696714  pheno035       99.68%    0.06/3.15 GB
```

t2wn53.physics.ox.ac.uk ☑ Show jobs
8 jobs, 2 users, 16.44 GB, 8.01 load

```
0: 5690808  atlasprd       99.77%    0.74/3.15 GB
1: 5693002  atlasprd       99.60%    0.74/3.15 GB
2: 5698471  cms161         97.74%    1.37/3.15 GB
3: 5696251  atlasprd       98.94%    0.74/3.15 GB
4: 5690985  atlasprd       99.75%    0.74/3.15 GB
5: 5692875  atlasprd       99.51%    0.74/3.15 GB
```

Figure 2. pbswebmon.

For security reasons it is important for sites to keep their systems package rpms up to date. Switching on automatic “yum” updates can help this, but for some critical packages this could break something so it is often not done. Similarly Kernel updates require a reboot and this is often delayed to avoid downtime. A useful way of seeing how many outstanding updates are required on a particular system is to run Pakiti [5]. This package, originally known as yumit, was developed by GridPP staff at the Rutherford Appleton Lab. On each client it effectively runs a “yum check-update” command to list any packages that have new updates available. The results are stored on a database on the Pakiti server and displayed on a web page figure 3. The graphical display allows systems administrators to spot systems that have failed to pick up expected updates, or are behaving differently to that expected.



Pakiti: "unpatched" hosts for My_Organization (27 October 2011 15:35)

Order by: admin

Display hosts: ○ all ● unpatched ○ not reporting

Section: P. Gronbech

Fedora release 12 (Constantine)

Total fixes	hostname	current kernel	last report	Connection
358	begbrokecam.physics.ox.ac.uk	2.6.31.12-174.2.22.f	6 January 2011 16:41	☒

Scientific Linux SL release 3.0.9 (SL)

Total fixes	hostname	current kernel	last report	Connection
7	pplxgen.physics.ox.ac.uk	2.4.21-52.EL	5 January 2011 04:10	☒

Scientific Linux SL release 4.4 (Beryllium)

Total fixes	hostname	current kernel	last report	Connection
46	t2lcfg.physics.ox.ac.uk	2.6.9-89.33.1.ELsmp	27 October 2011 04:19	☒

Scientific Linux SL release 4.8 (Beryllium)

Total fixes	hostname	current kernel	last report	Connection
2	gdg-nereuscpu2.physics.ox.ac.uk	2.6.9-89.0.28.ELsmp	22 April 2011 04:03	☒
2	t2mon02.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	13 December 2010 04:12	☒

Scientific Linux SL release 4.9 (Beryllium)

Total fixes	hostname	current kernel	last report	Connection
8	ppcresst1.physics.ox.ac.uk	2.6.9-89.0.28.ELsmp	29 June 2011 04:07	☒
7	ppcresst3.physics.ox.ac.uk	2.6.9-89.0.28.ELsmp	27 October 2011 04:15	☒
3	pplxconfig.physics.ox.ac.uk	2.6.9-78.0.1.ELsmp	27 October 2011 04:16	☒
3	pplxfs2.physics.ox.ac.uk	2.6.9-89.0.20.ELsmp	5 August 2011 04:12	☒
3	pplxfs3.physics.ox.ac.uk	2.6.9-55.0.9.ELsmp	27 October 2011 04:08	☒
3	pplxfs4.physics.ox.ac.uk	2.6.9-78.0.1.ELsmp	27 October 2011 04:11	☒
3	pplxfs6.physics.ox.ac.uk	2.6.9-89.35.1.ELsmp	27 October 2011 04:17	☒
3	pplxtorque.physics.ox.ac.uk	2.6.9-89.0.11.ELsmp	27 October 2011 04:05	☒
3	pplxwn01.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:18	☒
4	pplxwn02.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:16	☒
3	pplxwn03.physics.ox.ac.uk	2.6.9-89.35.1.ELsmp	27 October 2011 04:13	☒
3	t2hn03.physics.ox.ac.uk	2.6.9-89.0.23.ELsmp	27 October 2011 04:14	☒
8	t2myproxy.physics.ox.ac.uk	2.6.9-100.ELsmp	27 October 2011 04:06	☒
28	t2se01.physics.ox.ac.uk	2.6.9-100.ELsmp	27 October 2011 04:11	☒
30	t2se03.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:14	☒
30	t2se04.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:17	☒
30	t2se06.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:17	☒
29	t2se08.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:08	☒
30	t2se09.physics.ox.ac.uk	2.6.9-89.35.1.ELsmp	27 October 2011 04:07	☒
30	t2se10.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:05	☒
30	t2se11.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:13	☒
2	t2se12.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:17	☒
30	t2se13.physics.ox.ac.uk	2.6.9-89.29.1.ELsmp	27 October 2011 04:12	☒

Figure 3. Pakiti.

The above three packages display results via a web interface that must be actively monitored. However, systems administrators want to be notified by an email or SMS message when systems start to fail for either hardware or software reasons. Nagios [6], which provides a very powerful framework, can be used to monitor the status of systems by running tests at specified intervals and then performing actions depending on the results. This could be emailing a warning message or running an event handler that takes remedial action to solve a problem. The advantage of Nagios is that it removes the need to monitor web pages when all is well but can provide notification via email, or SMS (as well as the web) when there is a problem. There are repositories of checking scripts that can be used to build up a customized array of tests for each system. If a specific check is not already available it is relatively simple to create it: For example to check the health of a particular hardware RAID array controller on a storage server or to check a particular piece of Grid middleware is running properly. Each time a new problem occurs the systems administrator can enhance the local Nagios monitoring, to watch for this new failure mode. Simple examples may be local disk partitions getting dangerously full, or a more sophisticated batch systems “black hole spotter”. A black hole, in this context, describes a batch worker (CPU node), which develops a fault such that it accepts new jobs from the queue, then immediately fails them and then takes another job. In this situation all the jobs on the queue can very quickly be drained via this node and no further jobs will get run by the cluster, as that worker node is always apparently free.

Network health, usage and bandwidth should also be monitored at sites. Many sites use Cacti [7] and/ or Network Weathermap [8] to monitor the switches in and around their clusters. The switch connecting a cluster to the outside world is of particular relevance to Grid sites as almost all the work comes from other WLCG [9] sites around the world, and only a minority from the sites own local or Tier 3 facilities. Monitoring the usage of this link helps the site to know if they have enough bandwidth to feed their cluster. If it is continuously saturated for many days then thought should be given to increasing the site connection to JANET [10]. (UK sites are mostly connected at 1Gbps with a few sites at 2-5Gbps and a couple at 10Gbps. Over the next couple of years most sites are expected to move to approximately 2-5Gbps). Cacti can also be setup to monitor other systems that support the Simple Network Management Protocol (SNMP), such as the Power Distribution Units in each rack figure 4.

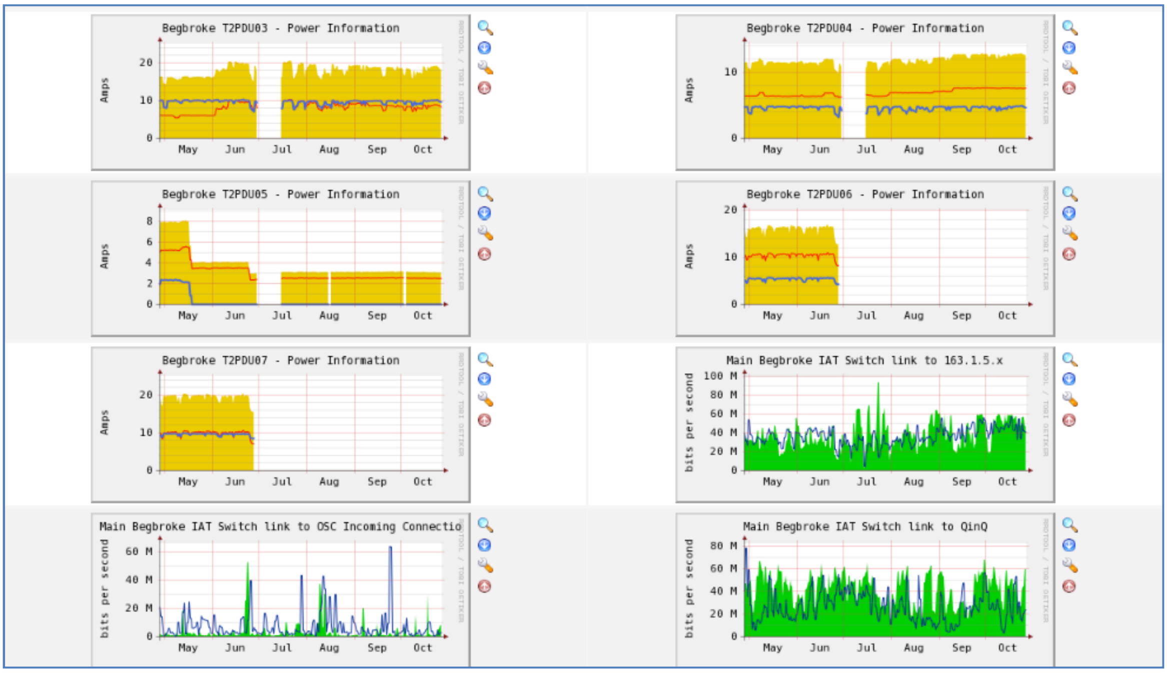
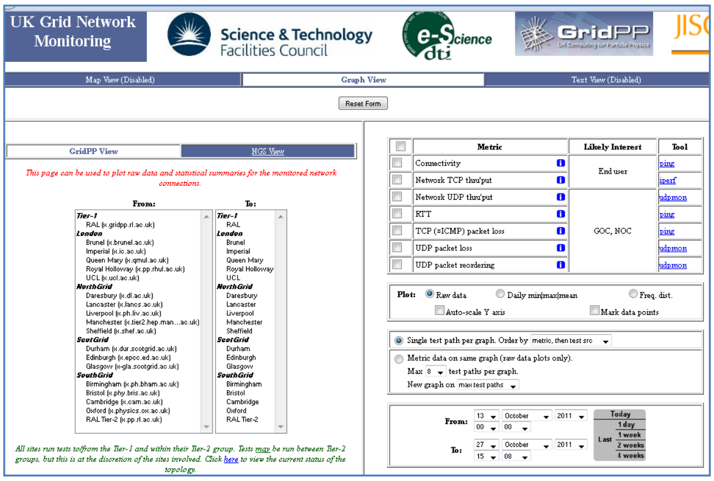


Figure 4. Cacti being used to monitor power and network switch traffic.

2. National testing

Although Cacti monitors the traffic actually flowing, it does not provide a measure of the available bandwidth between sites. The Gridmon [11] project was setup to address this problem and a dedicated ‘Gridmon’ test box is installed at each of the 19 GridPP sites to perform a matrix of iperf, udpmon and other network throughput tests. Each site performs tests against the other sites in its distributed Tier 2 and the Tier 1. The results are stored on a central database with a web frontend figure 5. This provides a useful history of the available bandwidth and can clearly show when the network gets degraded allowing problems to be diagnosed. A reduction in bandwidth can be easily seen which could be caused by a rate-cap that has been applied, or a faulty piece of networking equipment



UK Grid Network Monitoring

Map View (Disabled) | Graph View | Text View (Disabled)

Reset Form

GridPP View | NGS View

This page can be used to plot raw data and statistical summaries for the monitored network connections.

From:	To:
Tier-1	***Tier-1***
RAL (x.gridpp.rl.ac.uk)	RAL
London	***London***
Brunel (x.brunel.ac.uk)	Brunel
Imperial (x.ic.ac.uk)	Imperial
Queen Mary (x.qmul.ac.uk)	Queen Mary
Royal Holloway (x.pp.rhul.ac.uk)	Royal Holloway
UCL (x.ucl.ac.uk)	UCL
NorthGrid	***NorthGrid***
Daresbury (x.dl.ac.uk)	Daresbury
Lancaster (x.lancs.ac.uk)	Lancaster
Liverpool (x.ph.liv.ac.uk)	Liverpool
Manchester (x.tier2.hep.man...ac.uk)	Manchester
Sheffield (x.shef.ac.uk)	Sheffield
ScotGrid	***ScotGrid***
Durham (x.dur.scotgrid.ac.uk)	Durham
Edinburgh (x.epcc.ed.ac.uk)	Edinburgh
Glasgow (x-gla.scotgrid.ac.uk)	Glasgow
SouthGrid	***SouthGrid***
Birmingham (x.ph.bham.ac.uk)	Birmingham
Bristol (x.phy.bris.ac.uk)	Bristol
Cambridge (x.cam.ac.uk)	Cambridge
Oxford (x.physics.ox.ac.uk)	Oxford
RAL Tier-2 (x.pp.rl.ac.uk)	RAL Tier-2

All sites run tests to/from the Tier-1 and within their Tier-2 group. Tests may be run between Tier-2 groups, but this is at the discretion of the sites involved. Click here to view the current status of the topology.

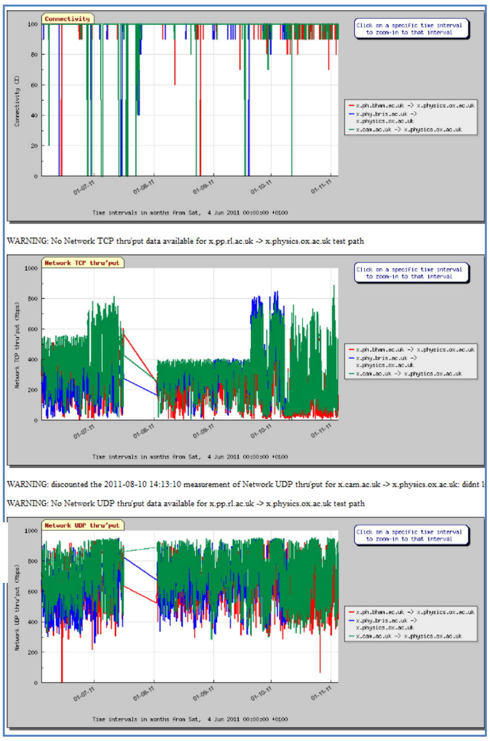
□	Metric	Likely Interest	Tool
□	Connectivity	End user	ping
□	Network TCP thru'put		iperf
□	Network UDP thru'put	GOC, NOC	udpmon
□	RTT		ping
□	TCP (≈ICMP) packet loss		ping
□	UDP packet loss		udpmon
□	UDP packet reordering		udpmon

Plot: ● Raw data ○ Daily min|max|mean ○ Freq. dist.
□ Auto-scale Y axis □ Mark data points

● Single test path per graph. Order by metric, then test src
○ Metric data on same graph (raw data plots only).
Max 8 test paths per graph.
New graph on max test paths

From: 13 October 2011 00 00
To: 27 October 2011 15 08

Today
Last 1 day
1 week
2 weeks
4 weeks



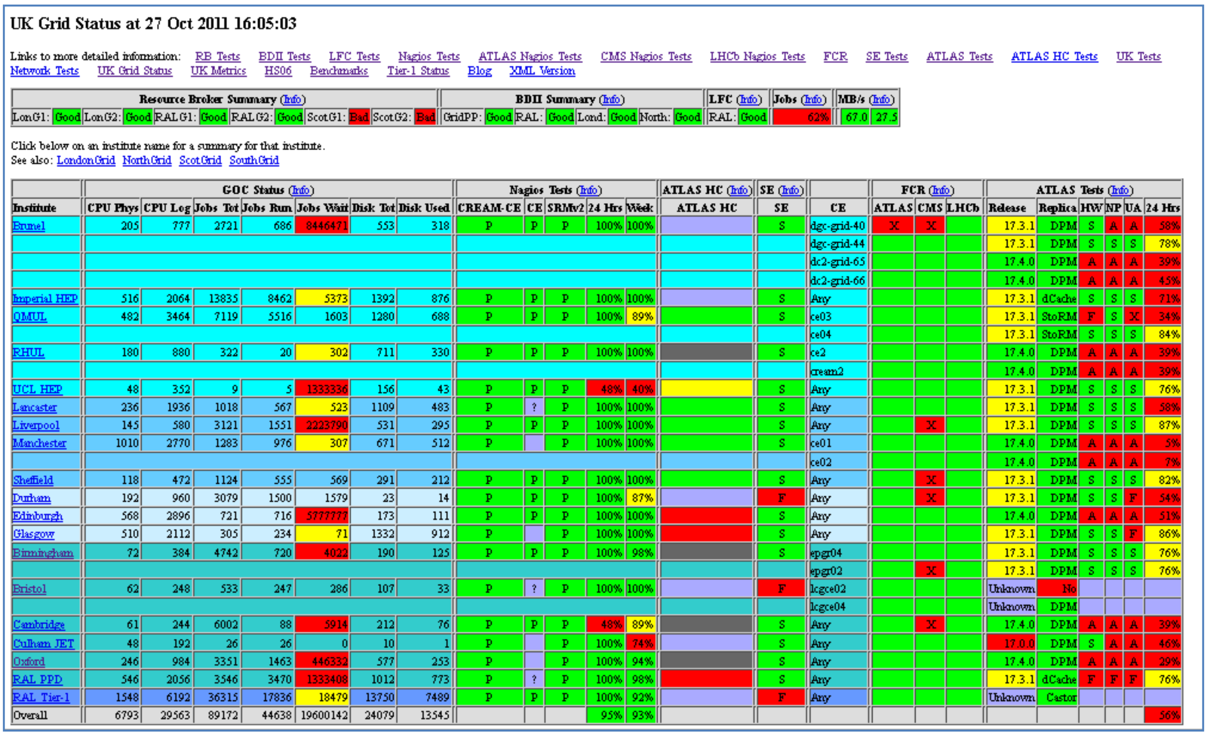
WARNING: No Network TCP thru'put data available for x.pp.rl.ac.uk -> x.physics.ox.ac.uk test path

WARNING: discounted the 2011-08-10 14:13:10 measurement of Network UDP thru'put for x.cam.ac.uk -> x.physics.ox.ac.uk: didnt

WARNING: No Network UDP thru'put data available for x.pp.rl.ac.uk -> x.physics.ox.ac.uk test path

Figure 5. Gridmon Network Bandwidth monitoring.

Other UK wide testing includes a GridPP-developed summation of relevant WLCG tests coupled with dedicated UK tests [12] developed by Prof. S. Lloyd at QMUL figure 6. This started out by just summarizing the Global Service Availability Metrics for the UK sites, but has expanded to include experimental tests and customized tests required by the UK.



UK Grid Status at 27 Oct 2011 16:05:03

Links to more detailed information: RB Tests BDII Tests LFC Tests Nagios Tests ATLAS Nagios Tests CMS Nagios Tests LHCb Nagios Tests FCR SE Tests ATLAS Tests ATLAS HC Tests UK Tests Network Tests UK Grid Status UK Metrics HS06 Benchmarks Tier-1 Status Blog XML Version

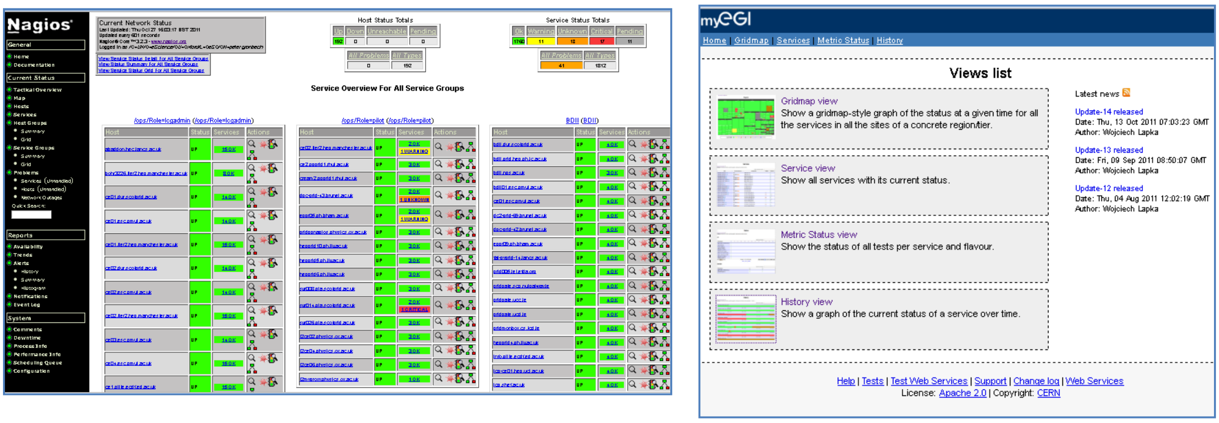
Resource Broker Summary (Info)												BDII Summary (Info)								LFC (Info)		Jobs (Info)	MB/s (Info)	
LonG1:	Good	LonG2:	Good	RAL G1:	Good	RAL G2:	Good	ScotG1:	Bad	ScotG2:	Bad	GridPP:	Good	RAL:	Good	Lond:	Good	North:	Good	RAL:	Good	62%	67.0	27.5

Click below on an institute name for a summary for that institute.
See also: LondonGrid NorthGrid ScotGrid SouthGrid

	GOC Status (Info)							Nagios Tests (Info)					ATLAS HC (Info)	SE (Info)		FCR (Info)			ATLAS Tests (Info)					
Institute	CPU Phys	CPU Log	Jobs Tot	Jobs Run	Jobs Wait	Disk Tot	Disk Used	CREAM-CE	CE	SRMv2	24 Hrs	Week	ATLAS HC	SE	CE	ATLAS	CMS	LHCb	Release	Replica	HW	NP	UA	24 Hrs
Brunel	205	777	2721	686	8446471	553	318	P	P	P	100%	100%		S	dgc-grid-40	X	X		17.3.1	DPM	S	A	A	58%
															dgc-grid-44				17.3.1	DPM	S	S	S	78%
															dc2-grid-65				17.4.0	DPM	A	A	A	39%
															dc2-grid-66				17.4.0	DPM	A	A	A	45%
Imperial HEP	516	2064	13835	8462	5373	1392	876	P	P	P	100%	100%		S	Any				17.3.1	dCache	S	S	S	71%
QMUL	482	3464	7119	5516	1603	1280	688	P	P	P	100%	89%		S	ce03				17.3.1	StoRM	F	S	X	34%
															ce04				17.3.1	StoRM	S	S	S	84%
RHUL	180	880	322	20	302	711	330	P	P	P	100%	100%		S	ce2				17.4.0	DPM	A	A	A	39%
															cream2				17.4.0	DPM	A	A	A	39%
UCL HEP	48	352	9	5	1333336	156	43	P	P	P	48%	40%		S	Any				17.3.1	DPM	S	S	S	76%
Lancaster	236	1936	1018	567	523	1109	483	P	?	P	100%	100%		S	Any				17.3.1	DPM	S	S	S	58%
Liverpool	145	580	3121	1551	2223790	531	295	P	P	P	100%	100%		S	Any		X		17.3.1	DPM	S	S	S	87%
Manchester	1010	2770	1283	976	307	671	512	P		P	100%	100%		S	ce01				17.4.0	DPM	A	A	A	5%
															ce02				17.4.0	DPM	A	A	A	7%
Sheffield	118	472	1124	555	569	291	212	P	P	P	100%	100%		S	Any		X		17.3.1	DPM	S	S	S	82%
Durham	192	960	3079	1500	1579	23	14	P	P	P	100%	87%		F	Any		X		17.3.1	DPM	S	S	F	54%
Edinburgh	568	2896	721	716	5777777	173	111	P	P	P	100%	100%		S	Any				17.4.0	DPM	A	A	A	51%
Glasgow	510	2112	305	234	71	1332	912	P		P	100%	100%		S	Any				17.3.1	DPM	S	S	F	86%
Birmingham	72	384	4742	720	4022	190	125	P	P	P	100%	98%		S	epgr04				17.3.1	DPM	S	S	S	76%
															epgr02		X		17.3.1	DPM	S	S	S	76%
Bristol	62	248	533	247	286	107	33	P	?	P	100%	100%		F	lcgce02				Unknown	No				
															lcgce04				Unknown	DPM				
Cambridge	61	244	6002	88	5914	212	76	P	P	P	48%	89%		S	Any		X		17.4.0	DPM	A	A	A	39%
Culham JET	48	192	26	26	0	10	1	P		P	100%	74%		S	Any				17.0.0	DPM	S	A	A	46%
Oxford	246	984	3351	1463	446332	577	253	P		P	100%	94%		S	Any				17.4.0	DPM	A	A	A	29%
RAL PPD	546	2056	3546	3470	1333408	1012	773	P	?	P	100%	98%		S	Any				17.3.1	dCache	F	F	F	76%
RAL Tier-1	1548	6192	36315	17836	18479	13750	7489	P	P	P	100%	92%		F	Any				Unknown	Castor				
Overall	6793	29563	89172	44638	19600142	24079	13545				95%	93%												56%

Figure 6. UK Grid Status.

The UK regional Nagios based Service Availability Monitoring (SAM) [13] is run by Oxford University. This service queries a central database (GOCDB) and Grid information services to create a list of sites and systems to be tested. The services offered are tested and the results of the tests are sent via an active MQ message bus to the EGI Regional Operations Dashboard [14] figure 8. Each region has an operator on duty that can raise alarm tickets against sites that have failed critical tests. All EGI [15] and WLCG sites have to meet agreed SLA’s, to respond to such tickets within a defined time depending on their tier status. The nagios web interface figure 7 can be used by Systems Administrators to investigate the error message resulting from the failed test. An alternative interface known as MyEGI is also available on the same server [16] that provides historical plots of availability.


Nagios
Current Network Status
Host Status Totals
Service Status Totals
Service Overview For All Service Groups
myEGI
Home | Gridmap | Services | Metric Status | History

Views list

Gridmap view
Show a gridmap-style graph of the status at a given time for all the services in all the sites of a concrete region/tier.

Service view
Show all services with its current status.

Metric Status view
Show the status of all tests per service and flavour.

History view
Show a graph of the current status of a service over time.

Latest news

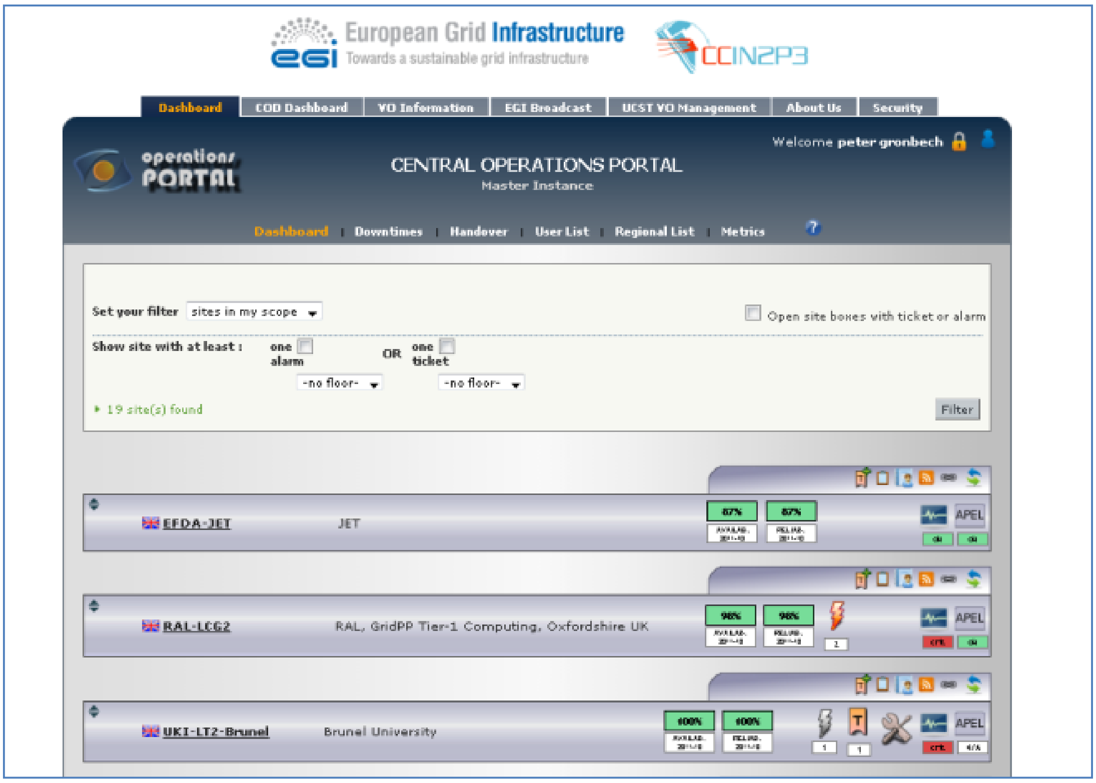
Update-14 released
Date: Thu, 13 Oct 2011 07:03:23 GMT
Author: Wojciech Lapka

Update-13 released
Date: Fri, 09 Sep 2011 08:50:07 GMT
Author: Wojciech Lapka

Update-12 released
Date: Thu, 04 Aug 2011 12:02:19 GMT
Author: Wojciech Lapka

Help | Tests | Test Web Services | Support | Change log | Web Services
License: Apache 2.0 | Copyright: CERN

Figure 7. Gridppnagios and the MyEGI portal.



Welcome peter gronbech

CENTRAL OPERATIONS PORTAL
Master Instance

Set your filter: sites in my scope

□ Open site boxes with ticket or alarm

Show site with at least : one alarm □ OR one ticket □

-no floor- -no floor-

19 site(s) found

EFDA-JET JET 87% 87%

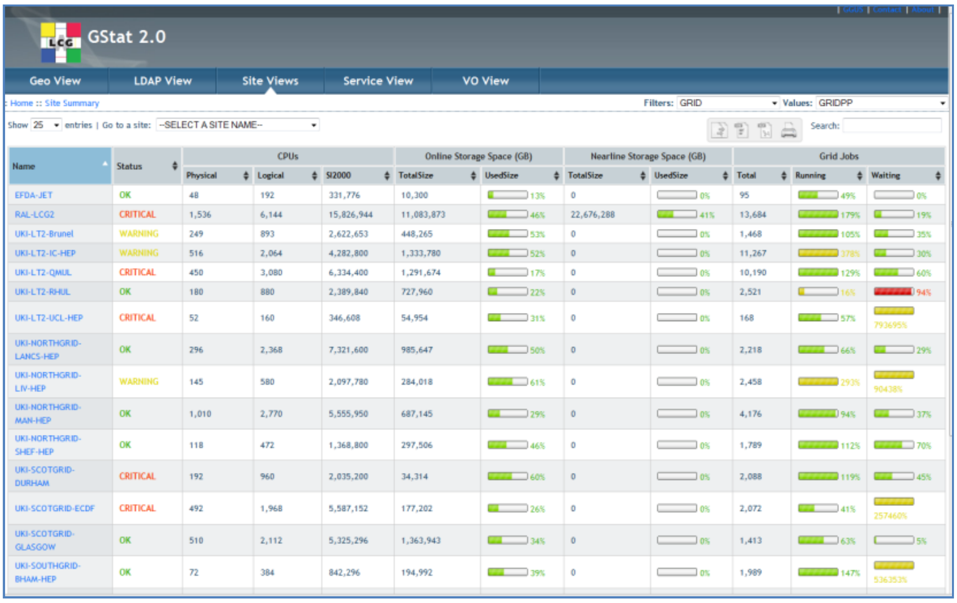
RAL-LCG2 RAL, GridPP Tier-1 Computing, Oxfordshire UK 98% 98% 1

UKI-LT2-Brunel Brunel University 100% 100% 1 1

Figure 8. Regional Operations Portal.

3. Global monitoring

There are many monitoring pages that look at all the European Grid Infrastructure (EGI) grid, or focus particularly on the WLCG grid.



Filters: GRID Values: GRIDPP

Show 25 entries | Go to a site: --SELECT A SITE NAME--

Name	Status	CPUs			Online Storage Space (GB)		Nearline Storage Space (GB)		Grid Jobs		
		Physical	Logical	SI2000	TotalSize	UsedSize	TotalSize	UsedSize	Total	Running	Waiting
EFDA-JET	OK	48	192	331,776	10,300	13%	0	0%	95	49%	0%
RAL-LCG2	CRITICAL	1,536	6,144	15,826,944	11,083,873	46%	22,676,288	41%	13,684	179%	19%
UKI-LT2-Brunel	WARNING	249	893	2,622,653	448,265	53%	0	0%	1,468	105%	35%
UKI-LT2-IC-HEP	WARNING	516	2,064	4,282,800	1,333,780	52%	0	0%	11,267	378%	30%
UKI-LT2-QMUL	CRITICAL	450	3,080	6,334,400	1,291,674	17%	0	0%	10,190	129%	60%
UKI-LT2-RHUL	OK	180	880	2,389,840	727,960	22%	0	0%	2,521	16%	94%
UKI-LT2-UCL-HEP	CRITICAL	52	160	346,608	54,954	31%	0	0%	168	57%	793695%
UKI-NORTHGRID-LANCS-HEP	OK	296	2,368	7,321,600	985,647	50%	0	0%	2,218	66%	29%
UKI-NORTHGRID-LIV-HEP	WARNING	145	580	2,097,780	284,018	61%	0	0%	2,458	293%	90438%
UKI-NORTHGRID-MAN-HEP	OK	1,010	2,770	5,555,950	687,145	29%	0	0%	4,176	94%	37%
UKI-NORTHGRID-SHEF-HEP	OK	118	472	1,368,800	297,506	46%	0	0%	1,789	112%	70%
UKI-SCOTGRID-DURHAM	CRITICAL	192	960	2,035,200	34,314	60%	0	0%	2,088	119%	45%
UKI-SCOTGRID-ECDF	CRITICAL	492	1,968	5,587,152	177,202	26%	0	0%	2,072	41%	257460%
UKI-SCOTGRID-GLASGOW	OK	510	2,112	5,325,296	1,363,943	34%	0	0%	1,413	63%	5%
UKI-SOUTHGRID-BHAM-HEP	OK	72	384	842,296	194,992	39%	0	0%	1,989	147%	536353%

Figure 9. GSTAT.

The GOCDB (Grid Operations Central Database) [17] registers details of all sites, their production and certification status and the contacts and services hosted there. Each site publishes the same information plus additional real time attributes via a site-BDII using LDAP. This can be seen via LDAP queries or by using the GSTAT [18] web pages figure 9. The GOCDB lists what is expected to be there and the GSTAT shows what is actually there, if there is a difference there may be a problem and this would be shown up by the Regional monitoring. An extension to GSTAT is the WLCG REBUS (Resource, Balance and Usage) [19] systems that makes use of published capacities and compares them against the pledged resources to the LHC [20] experiments, over time. This can highlight sites that have failed to provide what they promised.

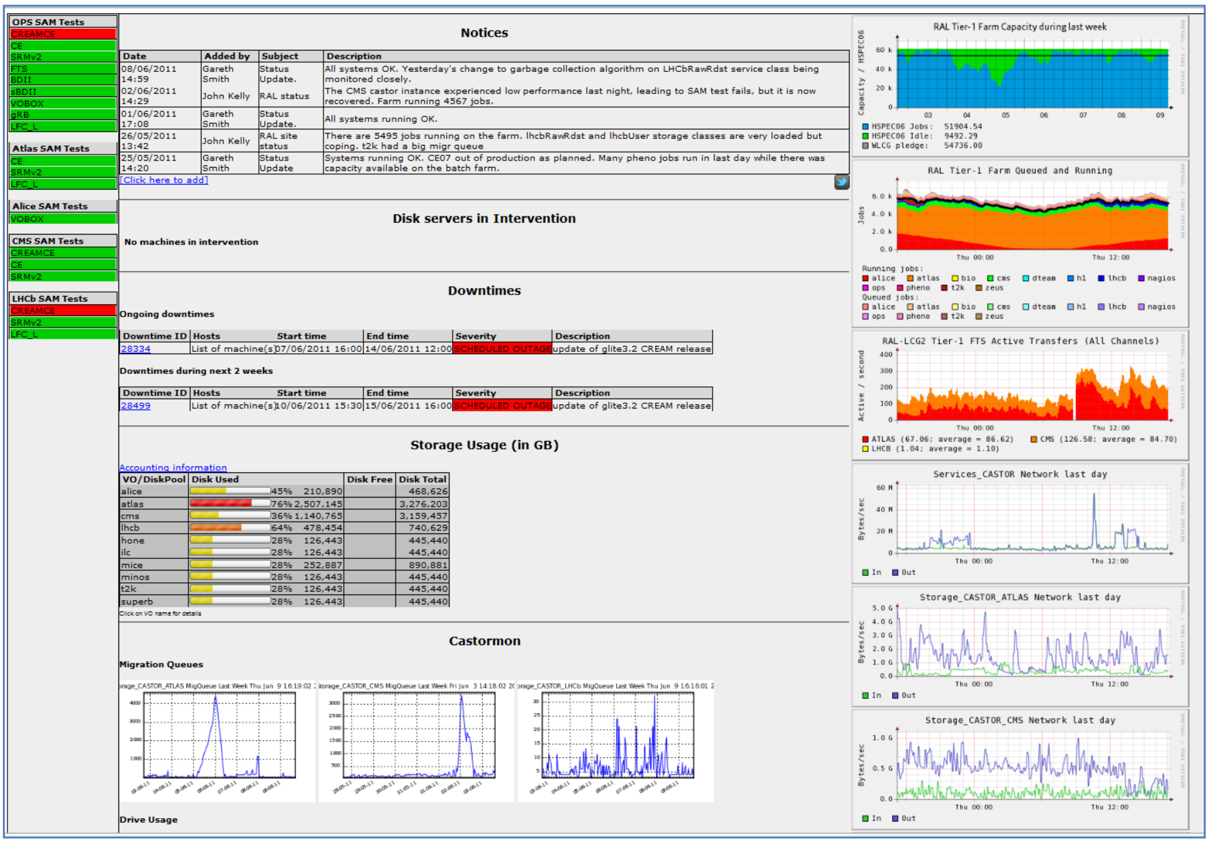
The actual work done at a site is measured by the Accounting system APEL [21], each night batch system logs are processed and correlated with the grid middleware logs to publish the number of jobs, and CPU time they took for each Virtual Organization. This is all collected together in a database and can be accessed by a very flexible web interface[22], which can show details of work carried out at individual sites, distributed Tier2s or by country, and can show the variations across time and users.

The main experiments at CERN submit jobs to the sites, with sophisticated job submission frameworks. The four LHC experiments dashboards now have a common frontend web page [23]. The dashboards track the jobs and report job submission, running, success or failure. Sites that have high failure rates are ticketed by the experiment team shifters, and can be automatically black listed from receiving new jobs.

4. Integrated dashboards

Systems Administrators are often overwhelmed by the number of different web sites and monitoring systems they should track. Attempts to integrate output from several systems into a site dashboard have been made at the Tier 1 figure 10 and some of the larger sites figure 11.

The RAL Tier1 Dashboard provides an overview of the Tier1 status. It was originally conceived as a dashboard to provide information externally (to the users). In practice it has proven an invaluable tool for internal monitoring as well. The dashboard pulls together a number of different sources of information. Several key ganglia plots provide trend information for batch and storage. Summaries of both the current SAM tests status and entries in the GOC DB are obtained via php scripts that request the information in XML and parse the result. A list of disk servers in intervention is provided by extract from the local database used to manage these systems. In addition the dashboard provides an area where Tier1 staff can display messages which are also copied to Twitter.



OPS SAM Tests
🟥 CREAMCE
🟩 CE
🟩 SRMv2
🟩 FTS
🟩 BDII
🟩 sBDII
🟩 VOBOX
🟩 gRB
🟩 LFC_L

Atlas SAM Tests
🟩 CE
🟩 SRMv2
🟩 LFC_L

Alice SAM Tests
🟩 VOBOX

CMS SAM Tests
🟩 CREAMCE
🟩 CE
🟩 SRMv2

LHCb SAM Tests
🟥 CREAMCE
🟩 SRMv2
🟩 LFC_L

Notices

Date	Added by	Subject	Description
08/06/2011 14:59	Gareth Smith	Status Update.	All systems OK. Yesterday's change to garbage collection algorithm on LHCbRawRdst service class being monitored closely.
02/06/2011 14:29	John Kelly	RAL status	The CMS castor instance experienced low performance last night, leading to SAM test fails, but it is now recovered. Farm running 4567 jobs.
01/06/2011 17:08	Gareth Smith	Status Update.	All systems running OK.
26/05/2011 13:42	John Kelly	RAL site status	There are 5495 jobs running on the farm. lhcbRawRdst and lhcbUser storage classes are very loaded but coping. t2k had a big migr queue
25/05/2011 14:20	Gareth Smith	Status Update	Systems running OK. CE07 out of production as planned. Many pheno jobs run in last day while there was capacity available on the batch farm.

[Click here to add]

Disk servers in Intervention

No machines in intervention

Downtimes

Ongoing downtimes

Downtime ID	Hosts	Start time	End time	Severity	Description
28334	List of machine(s)	07/06/2011 16:00	14/06/2011 12:00	SCHEDULED OUTAGE	update of glite3.2 CREAM release

Downtimes during next 2 weeks

Downtime ID	Hosts	Start time	End time	Severity	Description
28499	List of machine(s)	10/06/2011 15:30	15/06/2011 16:00	SCHEDULED OUTAGE	update of glite3.2 CREAM release

Storage Usage (in GB)

Accounting information

VO/DiskPool	Disk Used		Disk Free	Disk Total
alice	45%	210,890		468,626
atlas	76%	2,507,145		3,276,203
cms	36%	1,140,765		3,159,457
lhcb	64%	478,454		740,629
hone	28%	126,443		445,440
ilc	28%	126,443		445,440
mice	28%	252,887		890,881
minos	28%	126,443		445,440
t2k	28%	126,443		445,440
superb	28%	126,443		445,440

Click on VO name for details

Castormon

Migration Queues

Drive Usage

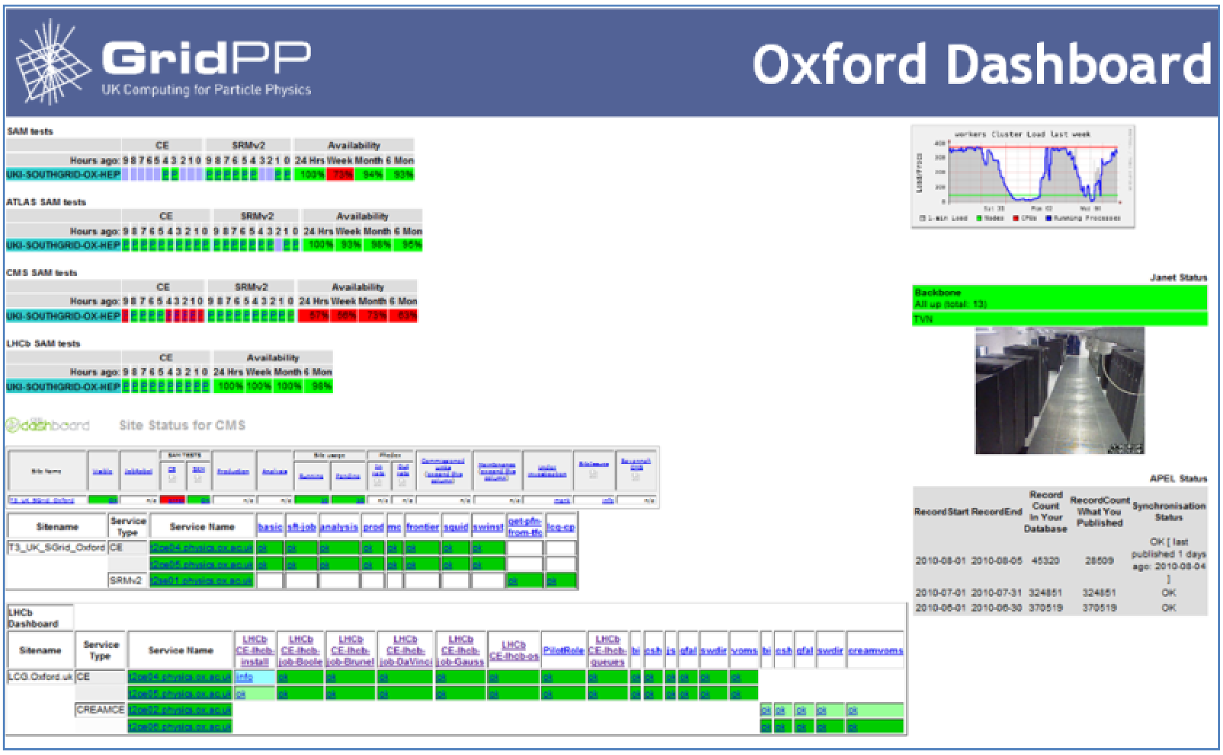
rage_CASTOR_ATLAS MigQueue Last Week Thu Jun 9 16:19:02 2
torage_CASTOR_CMS MigQueue Last Week Fri Jun 3 14:18:02 20
rage_CASTOR_LHCb MigQueue Last Week Thu Jun 9 16:16:01 2
RAL Tier-1 Farm Capacity during last week
Capacity / HSPEC06
60 k
40 k
20 k
0
03
04
05
06
07
08
09
HSPEC06 Jobs: 51904.54
HSPEC06 Idle: 9492.29
WLCG pledge: 54736.00
RAL Tier-1 Farm Queued and Running
Jobs
6.0 k
4.0 k
2.0 k
0.0
Thu 00:00
Thu 12:00
Running jobs:
alice atlas bio cms dteam h1 lhcb nagios
ops pheno t2k zeus
Queued jobs:
alice atlas bio cms dteam h1 lhcb nagios
ops pheno t2k zeus
RAL-LCG2 Tier-1 FTS Active Transfers (All Channels)
Active / second
400
300
200
100
0
ATLAS (67.06; average = 86.62)
CMS (126.58; average = 84.70)
LHCB (1.04; average = 1.10)
Services_CASTOR Network last day
Bytes/sec
60 M
40 M
20 M
In
Out
Storage_CASTOR_ATLAS Network last day
5.0 G
4.0 G
3.0 G
2.0 G
1.0 G
0.0
Storage_CASTOR_CMS Network last day
1.0 G
0.5 G
0.0

Figure 10. RAL Tier 1 Dashboard.

The Oxford dashboard is based on php code originally developed at Glasgow. It web scrapes the tables from Steve Lloyds SAM summary pages, the ganglia WN load graph, and some experiment dashboards. In addition as the computer room is in a remote location we find having a live camera feed useful. By showing the most relevant information on a single page that can be on display continuously in the support staff room, it can draw attention to faults at a glance.

5. Conclusions

Some of the many monitoring systems that are currently in use have been described. All areas; Fabric Monitoring, Grid Software; Experimental Job Submission and Accounting are continuously evolving. All providing views appropriate to each users particular interest. Sites running production resources have to take all of these into account. Systems administrators will use a combination of increasing the number of tests that local nagios monitoring systems check for and rectify automatically, coupled with use of site dashboards to reduce the time required to maintain an understanding of the health of their sites.



GridPP
UK Computing for Particle Physics

Oxford Dashboard

SAM tests

	CE										SRMv2										Availability			
Hours ago:	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	24 Hrs	Week	Month	6 Mon
UKI-SOUTHGRID-OX-HEP						P	P				P	P	P	P	P	P			P	P	100%	73%	94%	93%

ATLAS SAM tests

	CE										SRMv2										Availability			
Hours ago:	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	24 Hrs	Week	Month	6 Mon
UKI-SOUTHGRID-OX-HEP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P		P	P	100%	93%	98%	96%

CMS SAM tests

	CE										SRMv2										Availability			
Hours ago:	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	24 Hrs	Week	Month	6 Mon
UKI-SOUTHGRID-OX-HEP	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	57%	56%	73%	63%

LHCb SAM tests

	CE										Availability			
Hours ago:	9	8	7	6	5	4	3	2	1	0	24 Hrs	Week	Month	6 Mon
UKI-SOUTHGRID-OX-HEP	P	P	P	P	P	P	P	P	P	P	100%	100%	100%	98%

dashboard Site Status for CMS

Sitename	Service Type	Service Name	basic	sft-job	analysis	prod	mc	frontier	squid	swinst	get-pfn-from-tfc	lcg-cp
T3_UK_SGrid_Oxford	CE	t2ce04.physics.ox.ac.uk	ok	ok	ok	ok	ok	ok	ok	ok		
		t2ce05.physics.ox.ac.uk	ok	ok	ok	ok	ok	ok	ok	ok		
	SRMv2	t2se01.physics.ox.ac.uk									ok	ok

LHCb Dashboard

Sitename	Service Type	Service Name	LHCb CE-lhcb-install	LHCb CE-lhcb-job-Boole	LHCb CE-lhcb-job-Brunel	LHCb CE-lhcb-job-DaVinci	LHCb CE-lhcb-job-Gauss	LHCb CE-lhcb-os	PilotRole	LHCb CE-lhcb-queues	bi	csh	js	gfal	swdir	voms	bi	csh	gfal	swdir	creamvoms
LCG.Oxford.uk	CE	t2ce04.physics.ox.ac.uk	info	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok					
		t2ce05.physics.ox.ac.uk	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok	ok					
	CREAMCE	t2ce02.physics.ox.ac.uk															ok	ok	ok	ok	ok
		t2ce06.physics.ox.ac.uk															ok	ok	ok	ok	ok

Janet Status

Backbone
All up (total: 13)

TVN

APEL Status

RecordStart	RecordEnd	Record Count In Your Database	RecordCount What You Published	Synchronisation Status
2010-08-01	2010-08-05	45320	28509	OK [last published 1 days ago: 2010-08-04]
2010-07-01	2010-07-31	324851	324851	OK
2010-06-01	2010-06-30	370519	370519	OK

Figure 11. Oxford Site Dashboard**.**

Acknowledgements

This work is carried out on behalf of and with the support of the GridPP project [1].

References

[1] GridPP: the UK grid for particle physics D. Britton et al, UK e-Science All Hands Conference, Phil. Trans. R. Soc. A June 28, 2009 367:2447-2457; doi:10.1098/rsta.2009.0036
[2] Ganglia http://ganglia.sourceforge.net/
[3] Torque http://www.adaptivecomputing.com/products/torque.php
[4] Pbswebmon http://sourceforge.net/apps/trac/pbswebmon/wiki
[5] Pakiti http://pakiti.sourceforge.net/
[6] Nagios http://www.nagios.org/
[7] Cacti http://www.cacti.net/
[8] Network Weathermap http://www.network-weathermap.com/
[9] WLCG http://lcg.web.cern.ch/lcg/
[10] JANET http://www.ja.net/
[11] Gridmon http://gridmon.dl.ac.uk/gridmon/graph.html
[12] UKGrid Monitoring http://pprc.qmul.ac.uk/~lloyd/gridpp/ukgrid.html
[13] Distributed Nagios based SAM testing https://www.gridpp.ac.uk/wiki/UKI_WLCG_Regional_Nagios
[14] Regional Operations Dashboard https://operations-portal.in2p3.fr/dashboard
[15] EGI http://www.egi.eu/
[16] gridppnagios https://gridppnagios.physics.ox.ac.uk/nagios/ and https://gridppnagios.physics.ox.ac.uk/myegi
[17] GOCDB http://goc.egi.eu/
[18] GSTAT http://gstat-prod.cern.ch/gstat/summary/EGEE_ROC/UK/I/
[19] WLCG REBUS http://gstat-wlcg.cern.ch/apps/capacities/sites/
[20] The Large Hadron Collider http://lhc.web.cern.ch/lhc/

[21] APEL Accounting https://wiki.egi.eu/wiki/APEL
[22] Accounting Web Portal http://www4.egee.cesga.es/accounting/egee_view.html
[23] CERN Experimental Dashboards http://dashboard.cern.ch/