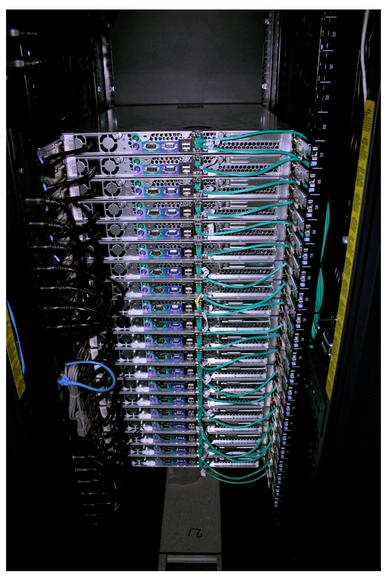
Large Cluster Management at a Tier2

The Nodes

- 2x2.8Ghz Xeons
- 2x250GB hdd
- 4GB RAM
- dual 1GgE



Each node has an onboard IPMI controller, that can be used to remotely manage and monitor the the chassis status. This, along with intelligent power distribution units, enables us to have a near 100% hands off management of the nodes.

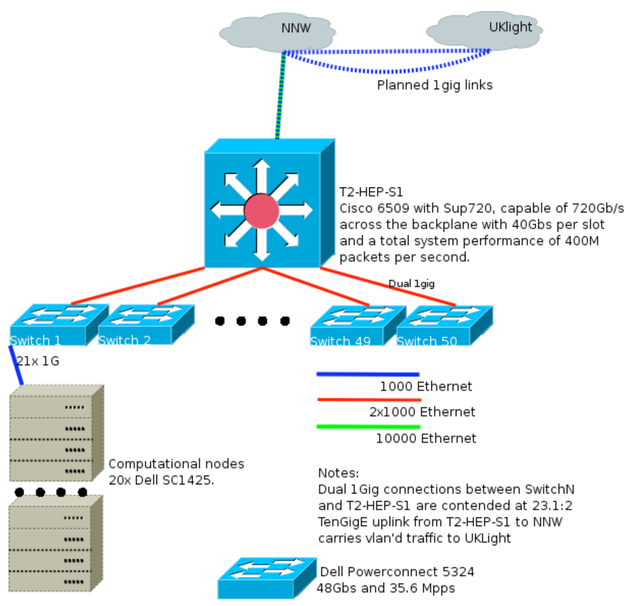
The Farm

- 1000 Nodes
- 2000x2.8Ghz Xeons
- 1/2 PB hdd
- 4TB RAM

The £1M farm is a sucessor to the Barbar Grid and LCG grid elements provided by the High Energy Physics group of the University of Manchester and one of the members of the regional tier2, Northgrid.

The Network

- 51 switches
- 50 Dell 5324
- 1 Cisco 6509
- 1250 Bit Links



Internal network connectivity is provided at 1Gbit/s to each node, with each of the 50 dcache doors having an extra 1Gbit/s connection on a seprate VLAN.

External Connectivity is currently at 10GBit/s and the process to upgrade to dual 10GBit/s should take place to coincide with the roll out of SJ5.

Author: Colin Morey
Presented by Dr A. Forti and Colin Morey

GridPP
UK Computing for Particle Physics

MANCHESTER 1824

Large Cluster Management at a Tier2

Kickstart

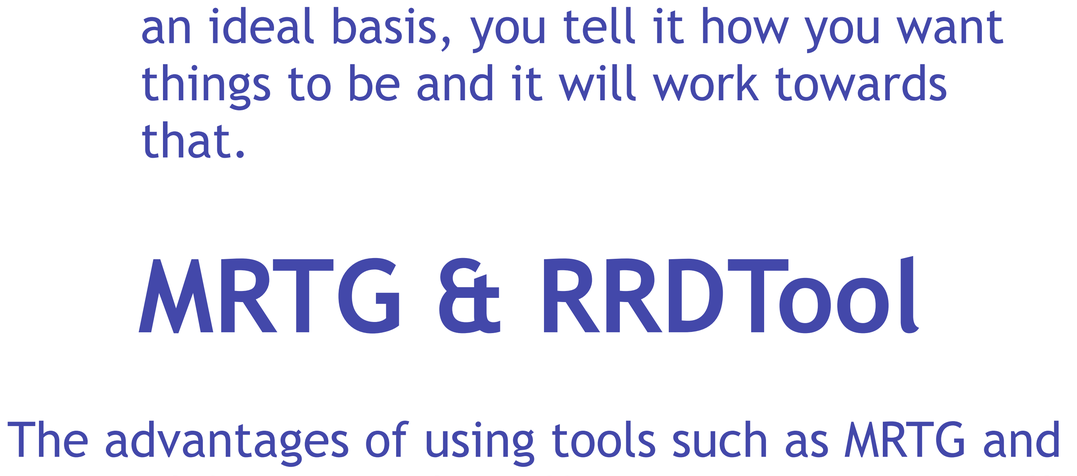
- 2 DHCP servers provide leases for the whole network, including the powerbars and the IPMI interfaces on the nodes.
- 4 Servers in total, connected via 2xGigE, All servers are identical and currently kept in sync via rsync from a designated ‘master’,
- With 3 servers, we were able to supply data at an aggregate rate of 600MB/s

Cfengine

- Cfengine is in the early stages of deployment at Manchester, with the aim of integrating it with the kickstart setup for easy synchronisation of all nodes.
- Cfengine operates on a convergence to an ideal basis, you tell it how you want things to be and it will work towards that.

MRTG & RRDTool

The advantages of using tools such as MRTG and RRDTool for network traffic monitoring are well known.
We have also started to monitor the power consumption of the individual nodes and plan to extend our monitoring to the physical environment of the servers.



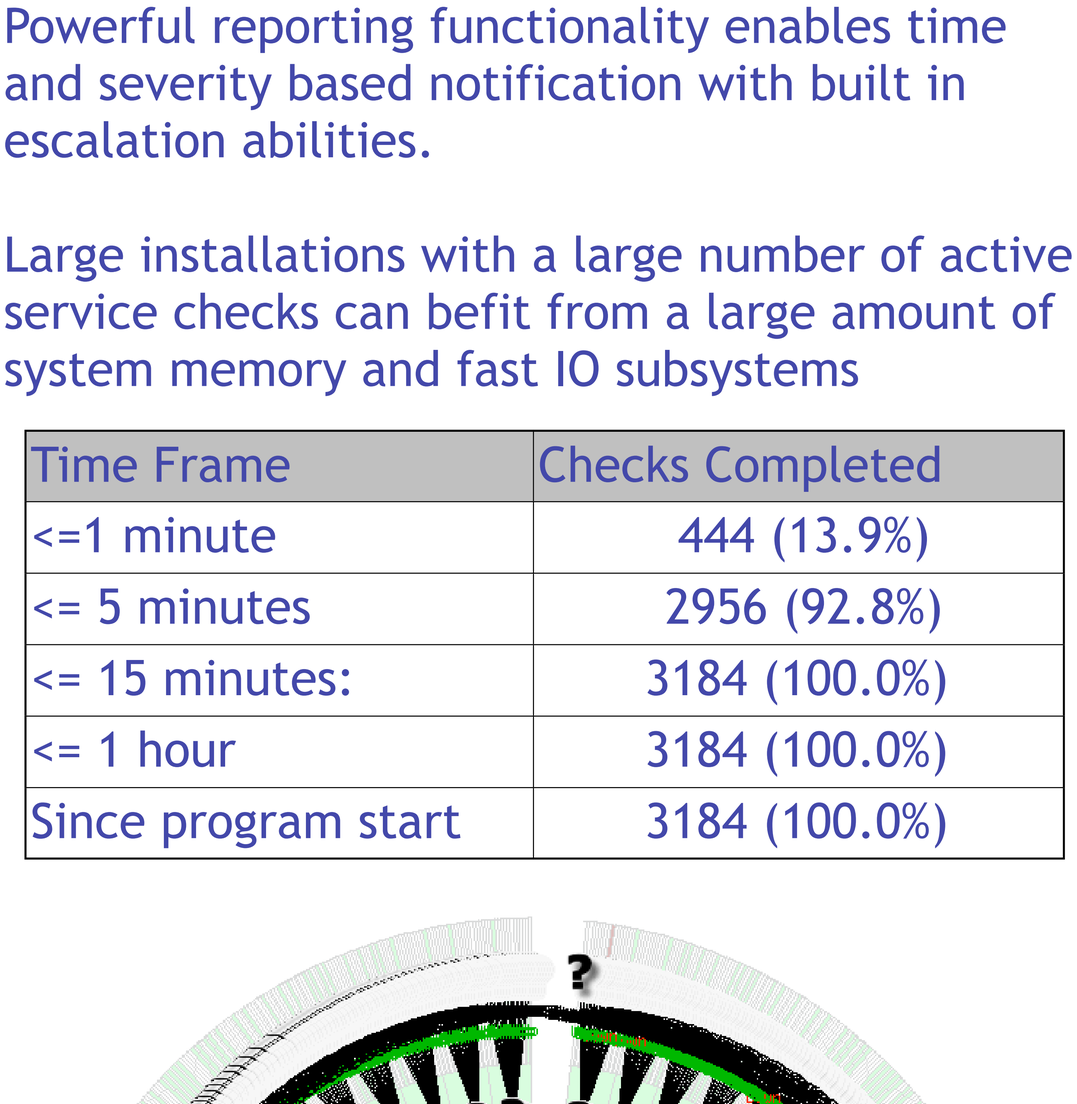
Nagios

Service based, open source monitoring tool, with an understanding of hierarchical relationships. Much of its power comes from its flexibility and extensibility with homegrown scripts written in a variety of languages.

Powerful reporting functionality enables time and severity based notification with built in escalation abilities.

Large installations with a large number of active service checks can befit from a large amount of system memory and fast IO subsystems

Time Frame	Checks Completed
<=1 minute	444 (13.9%)
<= 5 minutes	2956 (92.8%)
<= 15 minutes:	3184 (100.0%)
<= 1 hour	3184 (100.0%)
Since program start	3184 (100.0%)



Nagios Status Map

Author: Colin Morey
Presented by Dr A. Forti and Colin Morey