Large Cluster Management at a Tier2

C. Morey, A.Forti, Manchester Univerisity, United Kingdom

Abstract

The HEP department of the University of Manchester has purchased a 1000 node cluster. The cluster will be accessible to various VOs through EGEE/LCG grid middleware. In this paper and accompanying poster we will describe the systems behind the remote power control, centralised installation and updates, temperature, power and network performance monitoring.

INTRODUCTION

The Tier2 facility is a £1MP farm, the successor to the BaBar and LCG grid farms provided by the high energy physics group at the University of Manchester and is one of the members of the Northgrid regional tier2. The facility currently comprises of 1000 identically configured and maintained nodes, along with an accompanying networkr

FARM LAYOUT AND SPECIFICATIONS

Each node in the farm is a dual 2.8Ghz low voltage Xeon server with two 250Gb hard-drives with four Gigabytes of RAM and dual Gigabit Ethernet. Each node also has an onboard intelligent Platform Management Interface (IPMI)[1] controller that sits alongside the Gigabit Ethernet controllers and shares the same physical ports. This IPMI controller allows us to have ”hands off” monitoring of all aspects of the server including temperature and power status, and allows us to perform hard reboots without physical interaction. Internal connectivity within the farm is provided by a 24port, 10/100/1000bit/s Dell PowerConnect 5324 within each rack of 20 Machines. The extra four ports are spread between the power management, uplinks to the central switch and to the dcache door for the rack. The central switch is a Cisco 6509 with three 48 port 10/100/1000mbit/s blades, a four port 10Gigabit/s blade and a supervisor 720 management blade. The two links between the computational racks and the main switch are bonded together using 802.3ad[2] Link aggregation with 802.1q VLANs running over the bond, this enables us to make maximum use of the remaining capacity on the links, as the power management for the racks requires very little bandwidth.

INSTALLATION VIA KICKSTART

Initial installation of the nodes is done using the Redhat derived kickstart process. All nodes are configured to attempt to boot from the network as their primary method, only falling back to local disk if there is no boot file available. To configure this, a simple script is run which sets up links from various profiles on the primary kickstart server. The kickstart platform at Manchester currently consists of four dual PentiumIII rackmount servers with a single 200GB drive in each, connected directly into the Cisco switch via a bonded pair of Gigabit ethernet links. During initial testing, we sustained a constant draw from three of these servers at 600MB/s, installing the entire 1000 node farm in 40 mins, a vast majority of that time was spent with the nodes formatting local disks. Crucial to the process of the installation and indeed the running of the entire farm is a pair of DHCP servers which provide static leases for the entire farm. Currently the addition of a new machine, be node or server, into the farm requires manual addition in both DHCP and DNS, although there are plans to make this process more hands-off.

CFENGINE

Along side the kickstart process at Manchester, we install cfengine onto each node. Cfengine is an autonomous agent coupled with a high-level description language to form a ’convergence to ideal’ management tool, one which we have high hopes for. Currently we use it to provide notification of changes to binary files on the servers, with plans to expand its use across all aspects of node management. Supporting both cron-driven, and remote execution, the need to log onto a node to perform administration duties should be drastically reduced. Cfengine is aware of both OS and node types, and is capable of changing its behaviour and running process dependant upon those and many other factors. One of the key points of cfengine is an ability to take a system from whatever state and to bring it up to the current layout, even if the system has been down for many weeks and there have been many OS patches and toolkit upgrades that have happened since the last update. As cfengine handles this without operator intervention, it frees up the system administrators to continue on with their primary role of providing support to their users.

NETWORK AND ENVIROMENTAL MONITORING

At Manchester, we currently use a number of tools to monitor various aspects of the facility. For high level network usage and monitoring of the environment that the servers reside in, we use MRTG with RRDtool as a data storage back-end. MRTG is used for direct SNMP polling of both switches and the power distribution units in each of the racks. Widely available scripts create the network traffic plots we use for trend monitoring and some home-grown perl/bash hybrids are used to generate the power monitoring.

HOST AND SERVICE MONITORING

Service availability is monitored with the nagios monitoring tool. All hosts have a basic set of monitored services, ping and ssh/telnet. Initially we had problems with monitoring the 3000 plus targets that we had listed, investigation into the system usage showed contention on the disk subsystem as the OS paged memory in and out of swap whilst trying to write the nagios status log, doubling the amount of system memory removed this bottleneck and we have started to increase the number of targets we monitor. Our latest addition was the monitoring of the local PBS server and the state of the queues, this is done via check$_pbs.pl$

REFERENCES

[1] http://www.intel.com/design/servers/ipmi/

[2] http://standards.ieee.org/catalog/olis/lanman.html#csma

[3] http://standards.ieee.org/reading/ieee/std/lanman/restricted/802.1Q-2003.pdf

[4] http://en.wikipedia.org/wiki/VLAN

[5] http://www.cfengine.org

[6] http://www.mrtg.org

[7] http://www.rrdtool.org

[8] http://www.nagios.org

[9] http://www.hep.man.ac.uk/u/moreyc/tier2/nagios-plugins/