

## Cluster Development for Fall Semester 2008

Patrick Ford

This report will outline the results of my work developing the cluster over the Fall Semester 2008 and the Summer. Two documents are attached:

1. Installation and reference guide for the Frontend and NAS's.
2. Installation guide for OSG Production on the new Frontend.

Over the summer, David Pena and I tested out Rocks 4.3 on the old P3 frontend, but ran into difficulties configuring condor optimally for our uses. The process was simply trial and error which was time consuming. When the new hardware was ordered (Frontend, NAS, and sample compute node), we installed a beta version of Rocks 5.0 (anticipating the release of the full version) and configured most of the cluster components. By the time the Fall semester came, we had received the rest of the 20 new compute nodes and installed them into the newly configured Rocks 5.0 (release) cluster. The Frontend installation guide outlines the details of the cluster configuration.

The new cluster hardware is as follows:

### Frontend

8 Intel Xeon E5410 (233MHz, 2.33GHz)  
8GB 667MHz DDR2 (Buffered ECC Interleaved)  
500GB RAID5+Spare System Array  
250GB Squid Caching Drive

### NAS0

4 Intel Xeon E5410 (233MHz, 2.33GHz)  
4GB 667MHz DDR (Buffered ECC Interleaved)  
9.6TB RAID6 Array  
20GB Carved OS Partition (ext3)  
9.6TB Data Partition (XFS)

### Compute Nodes

160 Total Cores  
8 Intel Xeon E5410 (233MHz, 2.33GHz)  
16GB 667MHz DDR2 (Buffered ECC Interleaved)  
250GB System Drive

### Networking

Cisco Express 500G  
12 10/100/1000Base-T ports  
2 Linksys SR2016  
16 Gigabit ports

Over the fall semester, I made the steps to get us onto the Open Science Grid Production Grid and made us an official CMS Tier 3 site (details are outlined in the OSG installation guide). After receiving the CMS software from UF, I had a few more tweaks to perform before we started running successful CMS grid jobs. Since then we have received jobs from CMS and the Grid Laboratory of Wisconsin, as well as OSG swift workflows. Several professors are also using the cluster for their own simulations. The total resource hours are shown below.

Total Resource Hours on FLTECH							
	Total	g4hep	uscms01	mwood	zhang	fitzhang	glow
Oct 3, 2008	150507.1	74211	59956.96	16339.14	0	0	0
Oct 10, 2008	175437	96645	61119	17673	0	0	0
Nov 7, 2008	220460	114128	61967	18248	5803	0	20314
Nov 14, 2008	232592	114458	61967	18248	7410	0	30509
Dec 3, 2008	280330	114501	80071	24762	22232	19	38745
Dec 9, 2008	300535	114501	83396	26539	36028	1326	38745

I spent the rest of the semester optimizing the Frontend and compute node configuration, including disabling job pre-emption in Condor in favor of letting all jobs run to completion. I also created a Wiki on the cluster website, and more recently installed a gigabit port into the school's uplink router so that we have a full 1G connection to the Florida Lambda Rail.

At the end of the fall semester, we ordered additional UPS battery backup systems for the cluster, so that all machines would be covered in case of a power glitch. I installed the machines and configured the web-interface for external monitoring and automatic shutdown of critical servers (frontend and NAS).

The final cluster layout in the 50U rack is shown in an attached document (to scale).

I have performed bandwidth and filesystem performance tests over the Lustre cluster filesystem with Yujun Wu at UF to determine if we can use Lustre to transfer software or simulation data more quickly to Tier 3 grid sites. The results have been positive, especially after we installed the gigabit interface at the Florida Tech end. Tweaking the TCP settings of the frontend also improved iperf bandwidth and transfer performance with UF.

I attended the CMS Tier 3 meeting (October 23rd) at Fermilab where I viewed presentations from other Tier 3 sites and discussed our experiences running a grid site.

#### Future Work:

Continue network performance tests and test CMS software installation using Lustre.

Set up additional backup systems for final MT data.  
Ensure cluster stability and optimize.  
Get onto the CMS Tier 3 hypernews and collaborate with other T3 sites.