

cenpa-rocks Site Report

**KATRIN
TRIMS
MAJORANA
COHERENT
LEGEND
Project 8
Selena
Gravity**



CENPA

Center for Experimental Nuclear Physics and Astrophysics

Annual Report 2018

CENTER FOR EXPERIMENTAL NUCLEAR PHYSICS AND ASTROPHYSICS

University of Washington



Muon g-2
MuSun
ADMX
DAMIC
⁶He
⁶He-CRES
¹⁹⁹Hg

Annual Report 2018
University of Washington



Duncan Prindle, University of Washington

University of Washington Nuclear Physics Laboratory

1948 Construction of Cyclotron Building

1963 Construction of Van de Graaff Generator Building

1985 Construction of superconducting linear accelerator

1998 Name change to Center for Experimental Nuclear Physics and Astrophysics

2002 Mothball of superconducting linear accelerator

2008 Installation of Athena computer cluster

CENPA research emphasis is on fundamental symmetries and neutrinos

Currently we have roughly 8 interconnected research groups

ADMX, ^6He and ^6He -CRES are based at CENPA

Project 8, TRIMS and ^{199}Hg are based in the UW physics building

The muon group experiments (g-2, MuSun) are at Fermilab and PSI

The neutrino groups experiments (KATRIN, Majorana, Legend, Selena, Coherent) are at various facilities

Each project typically procures its own computing at computer centers such as NERSC, TACC, Fermilab, PNNL etc. Most of the analysis is embarrassingly parallel but with lots of data they need mass storage, fast I/O and many processors.

Not all computing needs at CENPA are covered under allocations at large computer centers so we are updating the Athena cluster.

Athena was installed in 2008

- 128 nodes (Dell PowerEdge 1950)
 - 2 CPUs, 8 cores, 8GB memory, 146 GB SAS drive
- CISCO Infiniband interconnect for all nodes
- 45 TB disk provided by Polyserve SAN available via NFS over GigE backbone
- Battery unit for controlled shutdown

CENPA had sufficient power and cooling for this cluster

Cluster was expected to run for about 4 years

Stopped paying for support, dropped Polyserve, battery unit stopped working but compute nodes still mostly worked and we maintained and used the cluster sporadically since then until..

Last year NERSC donated a previous generation of PDSF hardware to us (however all hard drives from compute nodes and storage had to be removed)

Renamed cluster: cenpa-rocks

Installed rocks 7 for cluster management with SGE as batch manager

Current compute node configuration

- 39 nodes DELL PowerEdge R410, 12 cores, 48GB memory, 146 GB disk
- 51 nodes DELL PowerEdge R410, 8 cores, 24 GB memory, 450 GB disk
- 47 original Athena nodes, 8 cores, 16 GB memory, 146 GB disk

Most nodes have infiniband connection

All nodes connected to switches via GigE copper

Using 1 switch per rack (4 racks), DELL PowerConnect 6248

Switches are stacked using bonded 10 Gbs links

With PDSF hardware we got storage hardware (but no disks)

4 servers, DELL PowerEdge R710

Each has 2 PERC H800 hardware raid cards, 2 10 Gbs fiber ports

12 DELL PowerVault MD1200, 12 slots each

We used one MD1200 to make raid6 array with 8 8TB drives

File system became corrupt quickly

Upgrade to latest PERC H800 firmware seems to have fixed this

Should be able to transfer data from raid to compute nodes at more than 6Gbs
In practice the best I see is 1.2 Gbs

I work part time on the cluster maintenance. I have no background in systems administration so its been a bit of a learning curve

Rocks is nice for managing the cluster

Cluster reinstalled itself after power failure (even SGE restarted running jobs)

Augmenting distribution easy if its a simple rpm. I couldn't figure out how to add pip3.6 for python 3.6. I also had trouble with conflicting rpms.

Documentation is very good for some aspects. I'm not sure where to look for existing rolls.

For CPU intensive jobs the cluster is working well. Monte Carlo simulations can use all the cores.

For I/O intensive jobs not so good

We are looking or tools to identify network bottle necks within the cluster

Typically see about 10% of data transfer rate I expect

We are trying to build expertise in diagnosing internal network issues

Actively looking to replace the older compute nodes (and find more hard drives)