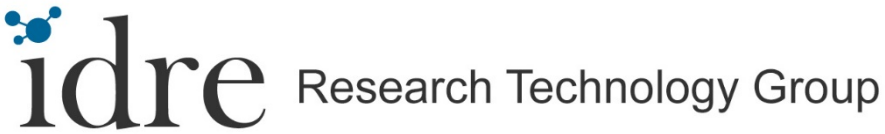

Hoffman2 Town Hall Meeting :: November 6, 2014

Agenda

- Introductions & state of the Cluster
- General Campus Users versus Contributors: What's the difference
- Storage and node service rates and what they cover
- Optimizing between HPC Storage and the Cloud Archival Storage System (CASS)
- Queuing system: Ten Top Tips
- Running jobs and HPC Storage information
- Q&A and Discussion – Let us hear from you

State of the Cluster

- 1,180 nodes / 12,220 cores
  - 824 contributed nodes / 8,952 cores
- 1,672 active users
- 290 research groups
  - 66 contributing research groups
- 90%+ average utilization
- Campus now funds all non-labor Cluster costs - $359K
- Scheduled Winter maintenance: December 22 from 0600-1800
- *CNSI Data Center decommission*
  - Will be returned to the campus within 1 year
  - Nodes in CNSI will be end of life and will not be moved

idre Research Technology Group

www.hoffman2.idre.ucla.edu

Hoffman2 Cluster User Guide

Search this site

Access | Computing | Data storage | File transfer | Software | FAQ | News

- Access
- Computing
- Data storage
- File transfer
- Software
- Cluster usage

User support

47
days until
Hoffman2 maintenance

2014-11-05: We are still updating... Please check back soon.

The Hoffman2 Cluster is a project of the Institute for Digital Research and Education (IDRE) Cluster Hosting Program. It opened to users on January 28, 2008. The Hoffman2 Cluster is managed and operated by the IDRE Research Technology Group under the direction of Bill Labate.

Serving the UCLA research community, the Hoffman2 Cluster has over 1,300 active users working with 250 faculty members from 100 UCLA departments, centers and institutes. Researches use more than 8,000,000 compute hours per month. Peak performance is 102.8 TFlops.

The Hoffman2 Cluster provides more than 11,000 processors in three data centers. Over 50 UCLA research groups have contributed nodes to the Hoffman2 Cluster. An additional 256 processors have been contributed by IDRE. Data centers, compute nodes and data storage are connected by Ethernet and InfiniBand networks.

Hoffman2 Cluster home directories are served from a NetApp storage system. All users have 20GB space for their home directories. Resource groups may purchased additional high-performance storage or archival storage. A state-of-the-art Panasas high performance parallel storage system serves the increasing active data storage demands of campus researchers.

The Hoffman2 Cluster is named after Paul Hoffman (1947-2003).

UCLA Institute for Digital Research and Education

Sign in | Report Abuse | Print Page | Powered By **Google Sites**

General Campus Users versus Contributors

- General Campus User
  - Access to only IDRE-provided surplus cores (~3,000 cores)
  - Can only run up to 24 hours
  - *Many other users in competition for resources*
  - 20GB of free storage per user
- Contributing User
  - Access to surplus cores on the entire cluster for 24 hours
  - Priority access over General Campus Users
  - Can run up to 14 days on contributed cores (longer runs by request)
  - Guaranteed start time
  - 20GB of free storage per user
- Either group can purchase additional storage in 1TB increments

Storage and Node Service Rates

- HPC Storage
  - Lower price for storage ($350/TB/Yr. versus $500)
  - Provides backup and no-backup options
- Compute nodes
  - Lower price per node ($5,916.88 versus $6,400)
  - Includes incremental IP network and Infiniband infrastructure
- Cloud Archival Storage Service (CASS)
  - Archival and backup storage
  - Multiple access methods (NFS, CIFS, iSCSI, Globus Online)
  - Multiple discounts up to 29.5%, federal and non-federal rates
- *All are fully self-sustaining rates*

CASS and Hoffman2

- HPC Storage is intended for active projects
- CASS is intended for archiving
  - Raw data
  - Inactive projects
  - Etc.
- Data transfers between CASS and Hoffman2
  - Use Globus
- Data transfers between CASS and your lab
  - NFS, SMB, iSCSI, Globus

Queuing System: Top Ten Tips

1. Memory size
2. Run-time limit
3. 1-CPU job should only use 1 CPU
   - Using more hurts other users sharing the same node
4. Use “job array” if submitting many jobs
5. Not every job can start immediately
   - Wait for resources to become available
   - High-priority vs. access to more resources

Queuing System: Top Ten Tips

6. h_data (memory size) is a per-slot value
7. (slots)*(h_data) must be smaller than RAM size
8. Make sure you have access to the nodes you request
9. Know when to submit high-priority jobs
10. Do not waste resources
    - Keep CPUs busy (doing useful computations)
    - Release resource as soon as you are done
    - Consider others in a shared environment

Running Jobs and Automatic Job Cleanup

- Jobs must be submitted through the queuing system
- Job monitoring based upon runtime, CPU consumption, and memory usage
- Completion of interactive processes (qrsh, abaqus, gaussview, etc.)

Hoffman2 HPC Storage

- Maximum of 10,000 files per directory
- /u/home vs. /u/project vs. /u/nobackup
- Migrating old /u/home sponsor's directories to /u/project or /u/nobackup
- Storage renewals

Q&A and Discussion