

The Discovery Cluster

<http://rc.dartmouth.edu>



Why use the cluster?

- Program that needs to run for a long time
- Large memory requirements
- Several jobs at a time
- Large data sets
- Files are backed up

Overview

- Discovery is comprised of 150+ compute nodes and has in total 3000+ cores available to the research community.
- The cluster uses the Linux operating system (Centos 6.8)
- Uses a batch system Torque resource manager along with Moab Workload Manager to Submit Jobs
- In aggregate the cluster has a total of 15.7TB of memory and more than 3PB of local/network disk space

Cluster Node Details



- The nodes of Discovery are broken out into 8 different groups (Cells) Each cell is comprised of different hardware.
- D) AMD 4284 3.0GHz (16 cores)
- E) AMD 4286 3.1GHz (16 cores)
- F) AMD 6348 2.8GHz (48-cores)
- G) Intel Xeon E5-2640 2.6 GHz (16 core)
 - 24 K80 GPU (119,000+ cuda cores)
- H) Intel Xeon E5-2470 2.xGHz (16 cores)
- J) Intel Xeon E5-2690 2.6GHz (24 core)
- K) Intel Xeon E5-2640 2.6GHz (16 core)
- M) Intel Xeon E5-2667 3.2GHz (16 core)

Applications

Discovery has an extensive software library, examples include:

- Python
- Datalad
- Cuda/8
- Tensorflow
- Singularity
- PyTorch
- Additional Applications are available upon request

What we will be going through

1. Logging into a head node (Discovery)
2. Environment modules
3. Upload data / program to cluster
4. Submitting batch / interactive job
5. Job monitoring

Logging on

- SSH (Secure Shell)
- Linux: `ssh -X username@discovery.dartmouth.edu`
- Mac: `ssh -Y username@discovery.dartmouth.edu`
- Windows
 - MobaXterm built in Xserver and sftp (free and recommended)
SSH secure shell or putty
- Changing your password:
 - Use the `passwd` command to make the change

Environment Modules

- To get a list of currently loaded modules:
 - **module list**
- To get a list of all available modules:
 - **module avail**
- To get a list of commands to use with modules:
 - **module help**

Moving files to / from the cluster

- Linux or Mac (CLI): sftp & rsync

CLI secure file transfer program – “sftp”

sftp username@discovery.dartmouth.edu

- put <file name> (mput filenames*)
- Get <file name> (mget filenames*)

- Another way to copy files to the cluster -- “rsync”

- rsync -av <files/folder to send> username@discovery.dartmouth.edu:.

Submitting basic job

You can easily submit a basic batch job simply by typing:

```
echo "hostname;sleep 120; date" | qsub
```

Once submitted you will see a job ID if successful.

Try these commands while the job is running:

```
showq -r -u username
```

```
checkjob jobid
```

Once complete, you should have two files .e & .o

Submitting Interactive Jobs

qsub -X -I -l nodes=1:ppn=1 -l walltime=20:00:00

- **qsub** (how to submit a job)
- **-X** (display)
- **-I** (interactive)
- **-l nodes=1:ppn=1** (One node one core)
- **-l walltime** (how long job will run)

Optional:

- **feature** (node set to submit to)

*Submitting an interactive job can take up to 2 minutes to start. *

Discovery Resources

- Research Computing:
research.computing@Dartmouth.edu
- Discovery
- <http://rc.dartmouth.edu>

Questions?