Introduction to Using HPC Clusters at UH

Hewlett Packard Enterprise Data Science Institute

University of Houston

Accounts and Allocations

- Before you can access any of the clusters you will need to request an account. The clusters are used for research, so you will need to be working on a research project with a P.I. This is verified with the P.I. The form can be found here.

  https://uh.edu/rcdc/getting-started/request-account.php

- Before you will be able to run any jobs, your P.I. will need to request an allocation. That form can be found here and is verified by an approval committee.

  https://uh.edu/rcdc/getting-started/request-allocation.php

- Because of the verification process, it can take up to a week before accounts or allocations are created.

Connecting to the Cluster

- Once your account has been approved and created, you will get an email with some general information.
- All access to the cluster is through secure shell. You will use your Cougarnet userid and Cougarnet password.

  If you are connecting from Linux or MacOS, you can use the ssh command.

  ```
  ssh userid@clustername.rcdc.uh.edu
  ```

  If you are connecting from Windows, you can use an application like PuTTY or MobaXterm.

  PuTTY - https://www.chiark.greenend.org.uk/~sgtatham/putty/

  MobaXterm - https://mobaxterm.mobatek.net/

- If you are not connecting from a UH wired connection or UH WiFi (UHSecure or eduroam) connection, you must use the UH VPN before trying to connect.

  https://uh.edu/infotech/services/computing/networks/vpn/

Hardware

- Carya
  - 246 48 core compute nodes with approximately 188Gb of usable memory
  - 12 52 core compute nodes with approximately 251Gb of usable memory
  - 8 48 core compute nodes with approximately 755Gb of usable memory
  - 16 48 core compute nodes with approximately 188Gb of usable memory and 2 NVIDIA V100 GPUs
  - 4 48 core compute nodes with approximately 377Gb of usable memory and 8 NVIDIA V100 GPUs
  - Mainly used for batch processing, not interactive

Hardware

- Sabine
  - 48 28 core compute nodes with approximately 125Gb of usable memory
  - 68 28 core compute nodes with approximately 251Gb of usable memory
  - 52 40 core compute nodes with approximately 188Gb of usable memory
  - 1 40 core compute nodes with approximately 755Gb of usable memory
  - 8 28 core compute nodes with approximately 251Gb of usable memory and 2 NVIDIA P100 GPUs
  - 4 28 core compute nodes with approximately 251Gb of usable memory and 8 NVIDIA V100 GPUs
  - Used for batch processing and interactive jobs

Directory/Folder Information

- You will have access to two different storage areas. The first is your home directory which will be 10Gb of space. The second will be the group project directory. The size of this directory will depend on the allocation requested by your P.I. Use the group project directory as your primary storage.

- Project directories are named after your P.I.
  Use the command `groups` to determine the group name. The first item in the list is the name you want. The directory will have a name like /project/davinci You can create your own directory there with the mkdir command.

```
$ groups
davinci appuser
```

- Your home directory is named with your cougarnet userid and will have a name like /home/ldavinci

  Note – the group project directory is not backed up. It is up to your group to determine their best practice of backing up.

Helpful Commands

- Quickly Navigate Between Filesystems – cd (change directory)

```
$ cd        # cd to home directory
$ cd -      # cd to previous directory

$ cd /project/groupname  # cd to your group project directory
```

- Show current directory

```
$ pwd
```

```
$ pwd

/home/ldavinci
```

Helpful Commands (cont)

- List the files in the current directory - ls

```
$ ls
$ ls -l    #long listing includes permissions and sizes
$ ls -a    #list all files include hidden files that name starts with a .
```

```
$ ls -al
total 64
drwxr-xr-x 14 ldavinci davinci  4096 Aug  7 09:58 .
drwxr-xr-x  4 root     root        0 Aug  7 13:44 ..
-rw-r--r--  1 ldavinci davinci   141 Jun 22 16:10 .bash_profile
-rw-r--r--  1 ldavinci davinci   313 Jun 30 13:12 .bashrc
-rwxr-xr-x  4 ldavinci davinci    33 Jun 24 15:44 job.sbatch
drwx------  2 ldavinci davinci   146 Aug  3 16:46 Downloads
```

Helpful Commands (cont)

- Make a copy of a file - cp

```
$ cp currentfile newfile
```

- Rename a file - mv (move)

```
$ mv currentfile newnamefile
```

- Make a new directory - mkdir

```
$ mkdir newdir
```

Helpful Commands (cont)

- **Show Disk Usage - df**

```
$ df -h directory-name
```

```
$ df -h /project/davinci
Filesystem                              Size  Used Avail Use% Mounted on
nfs-1-1-ib1:/export/project/davinci  2.0T  147G  1.9T   8% /project/davinci
```

```
$ df -h $HOME  # how much of your home directory is used

$ df -h .      # how much is used in the file system of your current
  directory (.)
```

Helpful Commands (cont)

- Check Your Allocation Balance For Running Jobs - sbalance

```
$ sbalance
```

```
$ sbalance
Account: davinci (DEFAULT)
Limit:      25000.00 SU
Unused:     24826.28 SU
Used:       173.72 SU (0.7 % of limit)
        User ldavinci used 26.2033 SU (15.1 % of total usage)
```

- sbalance --all will show the usage for all the users in your group
- 1 SU = 1 processor running for 1 hour

Helpful Commands (cont)

- Use installed software - module

```
$ module avail                  # list available software
```

```
$ module avail
--------------- /share/apps/modulefiles --------------------
  cudatoolkit/12.1    matlab/r2022b    matlab/r2023a (D)
--------------- /project/dsi/apps/modulefiles --------------
R/4.3.1-foss-2022a    python/3.10    PyTorch/2.0.1-foss-2022a
```

```
$ module load python/3.10     # load/add the python 3.10 module
```

Running Jobs

- Batch jobs are a type of non-interactive job that a user can submit to a queue. Once the requested resources are available, the batch job will automatically start running non-interactively on a compute node(s). Users do not need to be logged in or monitoring their batch jobs for them to automatically start.
- A batch job requires a short script with two sections. The first section (with lines beginning with #SBATCH) includes directives for the scheduler including what the job name should be, how many nodes are needed, what is the desired wall time, etc. The second section contains command-line instructions for applications to run. For example, that is where you would execute a python or matlab script. The job script can have any name.

Example Batch Job Script

- An example minimal batch job script is as follows:

```
#!/bin/bash
#SBATCH -J myjob                        # Job name
#SBATCH -o myjob.o%j                    # Name of stdout output file
#SBATCH -N 1                            # Total # of nodes
#SBATCH -n 28                           # Total # of processors
#SBATCH -t 01:30:00                     # Run time (hh:mm:ss)
#SBATCH --mail-type=END                 # When to get mail
#SBATCH --mail-user=your@email.address  # Address to get email (email lines are optional)
# Other commands must follow all #SBATCH directives...
module load needed_software
./mycode.exe
```

- Detailed information on the acceptable parameters as well as example job scripts can be found in the User Guide for the cluster you are using.

Additional sbatch directives

- #SBATCH --mail-type=XXX # You can have multiple mail-types separated by commas BEGIN,END or you can just use ALL
- #SBATCH --mem SIZE # include this for the memory size you need per node, default is 1G per processor
- #SBATCH --gpus=# # if you need gpus for your job, list the number needed
- #SBATCH --gpus=type:# # if you need a specific type and quantity (ex --gpus=volta:3 or --gpus=tesla:1)
- #SBATCH -A myproject # Allocation name (required if you have more than 1)

Submitting and Managing a Batch Job

- Once the job script is written, you can submit a batch job to the queue with:

```
$ sbatch myjobscript.sbatch
```

- Monitor the status of the job with `squeue` (pass the `--me` option to only show your jobs):

```
$ squeue --me
```

- Show estimated start times for only your jobs:

```
$ squeue --me --start
```

- Jobs can be deleted using the job ID:

```
$ scancel 123456
```

- Alternate for showing only your jobs (pass the `-u` flag to only show your jobs):

```
$ squeue -u $USER
```

Interactive Work

- Logging in to one of the clusters will place you on a shared login node. The login node is a shared resource and users should not run applications here (including Python, Matlab, R, or any other code downloaded from the internet). Instead, these applications should be run on a **compute node.**
- Use the command `hostname` to determine what type of node you are on. If the hostname command returns something like compute-X-Y, then you are on a compute node, otherwise you are on the login node.
- To run on a compute node interactively, you can use the `srun` command. The options are very similar to the directives to sbatch. You have to remain logged in for interactive jobs.

```
$ srun -N 1 -n 16 -t 2:00:00 --mem 32G --pty bash -l
```

- Make sure you are on a compute node and then you can run your commands. Type `exit` when you are done to log out of the compute node.

- If you are running interactively and need to display something graphical, you can add the option `--x11`
- The `srun` command still puts you in the queue, so you may not get an interactive session immediately.

squeue output

- When using the squeue command to check on jobs, you might see different messages
- When the job is running – ST (state) will be R and you will see which nodes are being used

```
JOBID   PARTITION  NAME     USER  ST  TIME       NODES  NODELIST(REASON)
952676  batch      NEB_14   user3 R   8-15:55:26     5  compute-0-[23,26],compute-2-[19,40,42]
```

- When the job is pending – ST (state) will be PD

```
JOBID   PARTITION  NAME     USER  ST   TIME      NODES   NODELIST(REASON)
945690  batch      new_mat  user2 PD   0:00         10   (Resources)

957072  batch      freq_o   user3 PD   0:00          4   (Priority)
```

- You will see the following reason for pending when your allocation does not have enough to run your job

```
JOBID   PARTITION  NAME     USER  ST   TIME      NODES   NODELIST(REASON)
941040  batch      mympij   user5 PD   0:00          2  (AssocGrpBillingMinutes)
```

- You will see the following reason for pending when you request more time than the maximum of 14 days

```
JOBID   PARTITION  NAME     USER  ST   TIME      NODES   NODELIST(REASON)

946027  batch      longjob  user7 PD   0:00          1   (PartitionTimeLimit)
```

Additional squeue information

- Pending reasons of Resources and Priority just mean the job is waiting to run. Use the `squeue --me --start` command to get an estimated start time
- If you see (Nodes required for job are DOWN, DRAINED or reserved for jobs in higher priority partitions) as a pending reason, it is similar to Resources and Priority
- If you see (ReqNodeNotAvail, Reserved for maintenance) as a pending reason, it is similar to Resources and Priority

General Information

- Do not run directly on the login node. Your jobs will be killed, since the job would impact the cluster, the job scheduler, and other users. Make sure to run either a batch job or an interactive job.

  If you are compiling and the instructions have make -j # as one of the steps, do not use a value larger than 4 for the number.

- The `sudo` command will not work for users of the cluster.

  If you are looking for information on the web and a suggestion has the sudo command as part of the suggestion, try without sudo.

- The `apt install` and `apt-get install` commands will not work on the cluster.

  These commands will sometimes show up as suggestions to fix a problem, but they are for a different operating system than what we run on the clusters.

Getting Help

- Additional training through the HPE Data Science Institute

  https://hpedsi.uh.edu/education/training/

- Online User Guide

  https://uh.edu/rcdc/support-services/user-guide/

- Need Help?

  https://support.hpedsi.uh.edu/

- Copy of the slides can be found in the Files section of this Microsoft Team