

# Compute Canada workshop

*Kim Dill-McFarland*

*2018-03-20*

## Outline

1. Systems available:
  - CC Cloud (Arbutus System Specs [https://docs.computeCanada.ca/wiki/Cloud\\_resources](https://docs.computeCanada.ca/wiki/Cloud_resources))
  - Cedar (System Specs <https://docs.computeCanada.ca/wiki/Cedar>)
  - Graham (System Specs <https://docs.computeCanada.ca/wiki/Graham>)
  - Niagara (Deployment later in the year System Specs <https://docs.computeCanada.ca/wiki/Niagara>)
2. List of available software ([https://docs.computeCanada.ca/wiki/Available\\_software](https://docs.computeCanada.ca/wiki/Available_software))
  - Most software is available by loading “modules”
  - Software can also be self-installed by user in their “home” directory
  - You can “bring your own license” for licensed software
3. Storage allocations:

| Storage Type | Who can request | Space Quota: Default | Space Quota:<br>Maximum available<br>via RAS by request           | # of Files<br>Quota |
|--------------|-----------------|----------------------|---|---------------------|
| /HOME        | NA              | 50GB per user        | NA  | 500K per<br>user    |
| /SCRATCH     | Any CC user     | 20TB per user        | 100TB per user  | 1M per user         |
| /PROJECT     | PI only         | 1TB per Group        | 10TB per group<br>(Cedar & Graham);<br>1TB per group<br>(Niagara) | 5M per<br>group     |
| /NEARLINE    | PI only         | NA                   | 5TB per group   | none                |

4. Logging in:
  - Can use any SSH client (MobaXterm and PuTTY for Windows; Terminal ssh for OS X and Linux)
  - Files can be copied to and from the systems using scp/sftp (supported by MobaXterm)
  - Standard Linux commands apply (ls, cat, less, etc...)
  - <https://docs.computeCanada.ca/wiki/SSH>
  - [https://docs.computeCanada.ca/wiki/Linux\\_introduction](https://docs.computeCanada.ca/wiki/Linux_introduction)
  - [https://docs.computeCanada.ca/wiki/Transferring\\_data](https://docs.computeCanada.ca/wiki/Transferring_data)
5. File systems:
  - Network filesystem - Your home directory is an example of a network filesystem. Data stored here is available throughout the HPC system and files stored here are backed up. Files stored here are typically slower to access, the data is actually stored on another computer and is being transmitted and made available over the network! /project and /home are both network filesystems.
  - Scratch - Some systems may offer “scratch” space. Scratch space is typically faster to use than your home directory or network filesystem, but is not usually backed up, and should not be used for long term storage. In the last example, /scratch is a scratch drive.
6. Submitting jobs:
  - Command to submit jobs is “sbatch”
  - Jobs are submitted via a job submission script
  - By default, the job output is placed in a file with the name “slurm-.out”
  - The job queue can be viewed using “squeue -u \$USER”

- To run interactive jobs (useful for testing):
- e.g. 1: `srun -pty -x11 -ntasks=2 -time=1:0:0 -account=def-venkmaha /bin/bash`
- e.g. 2: `salloc -time=1:0:0 -ntasks=2 -account=def-someuser`

## Example code

Example using user account **kadm** on resource **cedar**

### Log-in

```
ssh -Y kadm@cedar.computecanada.ca
```

-Y option allows use of graphical interfaces (GUI)

### Directories

- home: For backup files, programs
- Project: For larger files, data sets
- Scratch: Currently processing data, large allocation but not backed up
- Projects: **Ignore** Will be removed soon and does not exist on new accounts

### Usage

Check overall disk usage

```
diskusage_report
```

Or for a specific directory

```
lfs quota -u kadm /scratch
```

### Move files

Using **project1** folder created within the **scratch** directory.

Basic syntax is `scp file_you_want_to_move new_location`

### For your computer to Compute Canada

Terminal on your computer, not logged into CC

```
scp project3_test.gz kadm@cedar.computecanada.ca:~/scratch/project1/
```

### From Compute Canada to your computer

Terminal on your computer, not logged into CC

```
scp kadm@cedar.computecanada.ca:~/scratch/project1/test.txt ./
```

### Large files

For **very** large files, use [Globus](#)

## Software

Check what you have loaded  
`env`

Check what is available for loading  
`module avail`

Search for specific program like gcc  
`module avail | grep gcc`

Load default version (noted by D in module avail list)  
`module load gcc`

Load non-default version  
`module load gcc/7.3.0`

Check what version is loaded  
`gcc -v`

Unload a program  
`module unload gcc`

## Graphics programs

For Mac, need to download [XQuartz](#) to use anything that is GUI based (like xterm)

## Submit a job

Create .sh file

```
#!/bin/bash
#SBATCH --time=00:01:00
#SBATCH --account=def-shallam
echo 'Hello, world!'
sleep 30
```

- Time is as hr:min:sec
- Account is your username (if you have allocation) or your PIs account if you are working in their specified project

Submit the job  
`sbatch submit.sh`

See your queue  
`squeue -u kadm`

See the output  
`cat slurm-6278459.out`

## Help

For bioinformatics help, checkout the [helpdesk and forums](#)

For more tools and information, checkout [HPC Carpentry](#)