

INTRODUCTION TO CHEYENNE AND CASPER

BRIAN VANDERWENDE AND COLTON GRAINGER (SCRIBE)

1. CISL HELP DESK INFO

- ML 1B Suite 55
- 303-497-2400
- questions about this presentation? vanderwb@ucar.edu¹
- questions later? cislhelp@ucar.edu²

2. TUTORIAL

GLADE is the filesystem, and the CPUs are available on my particular node. Both CASPER and Cheyenne have the GLADE filesystem, but different login nodes. See the CISL course library³.

2.1. filespace.

```
/glade/u/home/$USER  
/glade/scratch/$USER  
/glade/work/$USER
```

Additionally, there are

- dedicated project spaces
- campaign storage for publication-scale storage lifespans
- dash process for collections in *long-term* storage
- HPSS tape archive for cold archival
 - e.g., to store satellite observations for 25 years

2.2. **transfers.** Globus services provides an NCAR GLADE endpoint. See Globus file transfers⁴.

- <https://www.globus.org/app/transfer>
- provides endpoints for transferring data 'tween endpoints (in parallel)

2.3. using modules. `module ...`

- “ (empty for help)
- `load <software>` and `unload <software>`
- `list` show all currently loadable modules
- `avail` or `av` show all currently loaded modules
- `purge` remove all loaded modules
- `save` create a saved set of software
- `restore` load a saved set of software
- `spider` to search for offerings of software available

What compilers are available?

Date: 2019-05-24.

url: `true`.

¹<mailto:vanderwb@ucar.edu>

²<mailto:cislhelp@ucar.edu>

³<https://www2.cisl.ucar.edu/user-support/training/library>

⁴<https://www2.cisl.ucar.edu/resources/storage-and-file-systems/globus-file-transfers>

- GNU
- Intel
- etc.

What software is available? All of R, MATLAB, Octave, Julia, Python, etc. Then `module av` will display default modules with the default version appended by (D), for example:

```
cmake/3.7.2
cmake/3.9.1      (D)
```

2.4. considerations when compiling software.

- Use `ncarc compilers` module along with library modules (e.g., `netcdf`) to simplify compilation, automatically adding flags for the make.
- Built software for either Cheyenne, or Casper, but usually not both.

2.5. run large take on compute nodes using batch jobs.

- many tasks are too large for the login nodes
- hence, schedule tasks to run on Cheyenne compute nodes using PBS or on Casper nodes using Slurm

```

                ssh cheyenne.ucar.edu
workstation -----> 6 login nodes -----> ~300 compute nodes
      \
      \
ssh casper.ucar.edu
  2 DAV login nodes -----> 28 data visualization nodes
```

2.6. example PBS and Slurm batch job scripts.

```
$ cat basic_mpi.pds
#!/bin/tcsh
#PBS -N hello_pbs
#PBS -A <project_code>
#PBS -j oe
#PBS -o pbsjob.log
#PBS -q regular
#PBS -l walltime=00:05:00 (don't under estimate)
#PBS -l select=2:ncpus=8:mpiprocs=8
```

```
#### set temp to scratch
setenv TMPR /clade/scratch/${USER}/temp
mkdir -p $TMPDIR
```

```
module load mpt/2.19
```

```
#### Run MPT MPI program
mpiexec ... TODO
```

Should you want to start with bash, don't forget `#!/bin/bash/ -l` for the login script.

Here're standard PBS commands for scheduling on Cheyenne compute nodes.

- `qsub <script>` submit batch job
- `qstat <jobid>` query job status
- `qdel <jobid>`
- `qinteractive -A <project>` run an interactive task on a Cheyenne compute node

2.7. PBS job submission queues. There's a wall clock of 12 hours for premium, regular, and economy submissions, with priority 1,2,3 respectively. Charges accrue on jobs in exclusive queues by $w \cdot n \cdot c \cdot q$ with

- w , wall-clock hours
- n , nodes
- c , 36 cores/node
- q , queue factor

In the shared queues (e.g., login shells), jobs accrue charges by $\frac{c}{3600}$ for c , wall-clock seconds.

2.8. Slurm. There are Slurm commands for scheduling on DAV (data analysis and visualization resources) in Casper. See “Starting jobs on Casper nodes | Computational Information Systems Laboratory”⁵.

```
$ cat gpu_job.slurm

#!/bin/tshc
#SBATCH -J sample_gpu
#SBATCH -n 6
... TODO
```

2.9. GUI applications. Use TurboVNC <https://www.turbovnc.org/> for MATLAB in IDE/development mode.

2.10. profiles for custom environments. Please don't load modules in the start up file! The *right way* to do this is to create, e.g., a default module set. The `default` module set is loaded at the start.

See <https://www2.cisl.ucar.edu/resources/computational-systems/cheyenne/user-environment/personalizing-start-files>.

```
module load ncl python nco mkl
module save default
```

2.11. examples.

- `ssh -l grainger -L <node>: ... TODO`
- need to provide a temporary login token (with Duo?)
- but how to load TurboVNC?

3. QUESTIONS

- Which python libraries are available?

See <https://www2.cisl.ucar.edu/resources/computational-systems/cheyenne/software/python>. Similar to using `conda`.

`ncar_pylib` loads a virtualenv with many packages. If the particular package is *not* listed, email cislhelp@ucar.edu⁶.

- Use `ncar_pylib` to creating your own clone of the NCAR package library.

From <https://www2.cisl.ucar.edu/book/export/html/5389>.

Creating a personal clone of the package library in your GLADE space is useful if you want to add or update packages, or develop Python code while using NCAR-provided packages. Simply specify the clone option (`-c`) to `ncar_pylib`, select a version of the package library, and choose a destination path for the clone.

```
ncar_pylib -c 20190118 /glade/work/$USER/personal_clone_name
```

⁵<https://www2.cisl.ucar.edu/resources/computational-systems/casper/starting-jobs-casper-nodes>

⁶<mailto:cislhelp@ucar.edu>

Then, use the standard virtual environment utility to activate your clone of the package library by sourcing it:

```
cd /glade/work/$USER/personal_clone_name/bin source activate.csh (tcsh) OR source activate (bash)
```

You can then install or upgrade packages from the Python Package Index using the pip utility. For example, to install your own copy of the seaborn plotting library, you could run:

```
pip install seaborn
```