

Introduction to batch computing on the Flux cluster: Torque PBS

28 Sep 2017

1. Components of the cluster management system
 - Login nodes
 - Compute nodes
 - Data transfer nodes
 - Home directory space
 - Scratch space
 - Batch job management software – Torque
 - Scheduling software – Moab
 - Billing system and account management
2. What you need to use the system
 - Login account, Duo authentication
 - Allocation: Could be provided by college (COE, LSA, or SPH) or Dept or PI or Rackham
3. A batch job runs a program when you're not around. Requires a working program.
4. Creating a first batch job and running it
 - create a batch file with a working script

```
$ mkdir IntroFlux
$ cd IntroFlux
$ nano hello.sh
----
echo "Hello, world"
----
```

- batch file structure: preamble and job portion
- The absolute, bare minimum that must be in a PBS script

```
#PBS -A training_flux
#PBS -q flux
./hello.sh
```

That will almost never be enough, but that's the minimum. Where's the money and which line do I stand in to pay.

5. Submit the job with `qsub hello.pbs`
6. When you run a program, output is printed to the screen. There are two kinds of output: regular, success output and error output.

7. By default, those will end up in separate files. That's not helpful for finding where an error may have occurred. So, we will often want to put both the errors and the output into one file.
8. You can specify all kinds of options in your PBS script
 - What are the available options?
 - We will go through them here starting with options that will typically not change much and ending with those that change a lot.

```
#####  
#####  
##### These are the most frequently changing  
options
```

```
##### Job name  
#PBS -N
```

```
##### Request resources here  
##### These are typically, number of processors, amount of memory,  
##### an the amount of time a job requires. May include processor  
##### type, too.
```

```
#PBS -l
```

```
##### Flux account and queue specification here  
##### These will change if you work on multiple projects, or need  
##### special hardware, like large memory nodes or GPUs or,  
##### or if you use software that is restricted to campus use.
```

```
#PBS -A  
#PBS -l
```

```
#####  
#####  
##### These are the least frequently changing options
```

```
##### Your e-mail address and when you want e-mail
```

```
#PBS -M  
#PBS -m
```

```
##### Join output and error; pass environment to job
```

```
#PBS -j
#PBS -V

# Add a note here to say what software modules should
be loaded.
# for this job to run successfully.
# It will be convenient if you give the actual load co
mmand(s), e.g.,
#
# module load intel/16.0.4
```

- To create a PBS script, add the preamble to any runnable script
- There are some other things are useful to have in the script. We provide a template you can use. To help you cement what the options are,

```
$ cp /scratch/data/workshops/IntroFlux/template.pbs he
llo.pbs
$ nano hello.pbs
Ctrl-R to read in hello.sh
```

- Complete the information in the preamble for `template.pbs`
- Note the if statements and the note about which modules are needed? (we'll come back to those in a bit).
- For now, let's run what we have

```
$ qsub template.pbs
```

- batch job manager commands

```
$ qsub <PBSscript.pbs>
$ qstat -u $USER
$ qstat <JobID>
$ qdel <JobID>
```

9. How to check on a job and on the line

```
$ checkjob -v <JobID>
```

10. Get everyone to submit a job at the same time here, then

```
$ showq -w acct=<AccountName>
```

11. Scheduler and how it is different from the batch manager

- scheduler determines the order in which things run and instructs the batch manager to start jobs.
- scheduler commands typically begin with an 'M', but not always

```
$ mdia -u $USER
```

shows which allocations can be used

```
$ mdia -a training_flux
```

shows procs and memory for an allocation

12. How to check on allocations

```
$ freealloc account_name
```

Maybe you are like me, submit a bunch of jobs, then realize that you didn't load the modules first. Aargh! You can use

```
$ cancel_my_jobs
```

to delete *all* your currently running or queued jobs. This wraps the `qdel` command with some options and error checking so you don't generate a ton of e-mail to us that you don't have permission to delete everyone else's jobs, too.

13. Running an interactive job

- What an interactive job is
- Run it when 1) you need more time or memory or threads than would be polite or allowed on a login node and/or 2) you need to run interactively with processors on more than one physical machine.
- To run an interactive job, you can put all the PBS options on the `qsub` command

```
$ qsub -I -V -l nodes=2:ppn=12,pmem=2gb,walltime=1:00:00 \
    -A training_flux -l qos=flux -q flux -j oe <pbs_script>
```

or you can add the `-I` option to `qsub` with a file

```
$ qsub -I <pbs_script>
```

Note: Resize your terminal window to the size you want *before* your submit an interactive job.

14. Copying data to and from Flux using the command line (Mac and Linux)

```
$ sftp flux-xfer.arc-ts.umich.edu
```

which will give you an interactive prompt, or

```
$ scp my_file flux-xfer.arc-ts.umich.edu:
$ scp my_data_file flux-xfer.arc-ts.umich.edu:data/
$ scp -r my_data_dir flux-xfer.arc-ts.umich.edu:my_study
```

- GUI tools, e.g., WinSCP, CyberDuck FileZilla