

Batch Systems

Running calculations on HPC resources

Dr Mark Bull, EPCC
markb@epcc.ed.ac.uk



Outline

- What is a batch system?
- How do I interact with the batch system
 - Job submission scripts
 - Interactive jobs
- Common batch systems
- Converting between different batch systems

Batch Systems

What are they and why are they used?

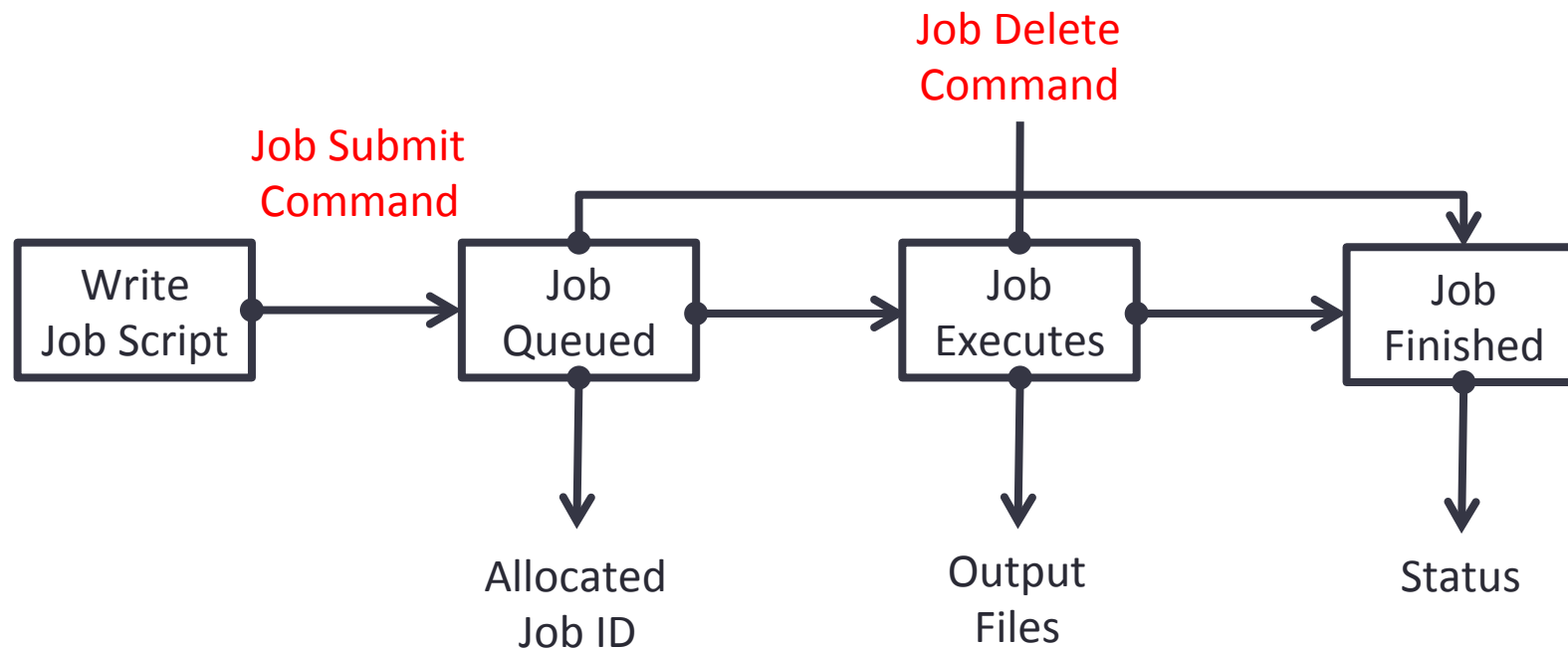
What is a batch system?

- A batch system controls access to the resources on a machine
- Used to ensure all users get a fair share of resources
 - As machine is usually oversubscribed
- Allows user to setup computational *job*, place it into batch queue and then log off machine
 - Job will be processed when there is space and time
 - Do not need to be continually logged-in for simulations to run
- Usually assumed that jobs are non-interactive
 - It runs for a time and produces results without intervention from the user
 - (Unlike interactive programs on a laptop.)

Reservation and Execution

- When you submit a job to a batch system you specify the resources you require:
 - Number of cores, job time,
- The batch system *reserves* a block of resources for you to use
- You can then use that block as you want, for example:
 - For a single job that spans all cores and full time
 - For multiple shorter jobs in sequence
 - For multiple smaller jobs running in parallel

Batch system flow



Running calculations

Interacting with the batch system

Batch and interactive jobs

- Most resources allow both batch and interactive jobs to be run through the batch system
- Batch jobs are non-interactive.
 - They run without user intervention and you collect the results at the end
 - Write a *job submission script* to run your job
- Interactive jobs allow you to use the resources interactively
 - For debugging/profiling
 - For visualisation and data analysis
- How you run these types of jobs differs with batch system and site

Job submission scripts

- Contain:
 - Batch system options
 - Commands to run
- Example:

```
#!/bin/bash -login
#PBS -N Weather1
#PBS -l mppwidth=4096
#PBS -l walltime=1:00:00
cd $PBS_O_WORKDIR
mpiexec -n 4096 ./weathersim
```

Common batch systems

Batch systems

- PBS, Torque
- Grid Engine
- SLURM
- LSF – IBM Systems
- LoadLeveller – IBM Systems

Migrating

Changing your scripts from one batch system to another

Conversion

- Usually need to change the batch system options
- Sometimes need to change the commands in the script
 - Particularly to different paths
 - Usually the order (logic) of the commands remains the same
- There are some utilities that can help
 - Bolt – from EPCC, generates job submission scripts for a variety of batch systems/HPC resources: <https://github.com/aturner-epcc/bolt>

Summary

Best practice

- Run short tests using interactive jobs if possible
- Once you are happy the setup works write a short test job script and run it
- Finally, produce scripts for full production runs
- Remember you have the full functionality of the Linux command line available in scripts
 - This allows for sophisticated scripts if you need them
 - Can automate a lot of tedious data analysis and transformation
 - ...be careful to test when moving, copying deleting important data – it is very easy to lose the results of a large simulation due to a typo (or unforeseen error) in a script