PBS to Slurm Transition

Transition from PBS to Slurm

- When the batch system changes from PBS to Slurm on June 4th 2019, the user interfaces, Auger and Swif will remain same.
- The existing jsub/jkill commands and jsub scripts, swif command and options used in Auger/PBS will continue to be used on Auger/Slurm.
- The nodes in Slurm are built to match the nodes in PBS. The same executable that works for PBS should work on the Slurm batch system.
- *There are few changes to emphasize.*

Job stdout and stderr

- In Auger/Slurm system, Auger will no longer manage user's stdout and stderr files, **Slurm will directly write the log message to the default or user specified location** so the log messages will be available to user during the run time of the job.
- The default job .out and .err files are located under **/farm_out/<user>**, name pattern will be **JOB_NAME-AUGER_JOB_ID-HOSTNAME.out** and **JOB_NAME-AUGER_JOB_ID-HOSTNAME.err.**
- If user changes the default logging location using the <Stdout> and <Stderr> tags, he/she **must create the log directory,** otherwise the job will fail with error **Job failed** due to invalid stdout/stderr.

Requesting memory

- In a change from PBS, Slurm uses *real* memory to schedule and kill the job, so user may be able to reduce their memory requested in the jsub script.

An example:

A 16 core job needs 60GB memory when using PBS, but 20GB may be enough for the same job running on Slurm.

The new features add to Auger-Slurm

- In Auger-Slurm, a user can request a whole node for a job. To do it just using **"CPU: 0** or **<CPU core="0" />** (in this case the memory tag will be ignored).

- A new optional attribute 'copyOption' (**copy or link)** is introduced to <Input> xml tag.

<Input src="/home/user/file1" dest="file" copyOption="link"/>

*<Input src="*mss:/mss/home/user/file.dat" dest="file.dat" *copyOption="copy"/>*

Other auger commands

- The **slurmjobs** command replaces the old jobinfo command. The notable change is that the output only contains jobs in the Slurm system.
- The **slurmHosts** command replaces the old farmHost command. It displays hosts and their static and dynamic resources/features.
- The **jobinfo** command is similar to the pbs jobinfo but outputs more useful information, such as MaxRss, MaxVMSize, AveDiskRead and AvgDiskWrite.

Slurm partition (queue)

- In slurm the production queue is split to two partitions (queues), general and production. The jobs submitted by production users (such as gxprojx, clas12 etc.) will route to production queue which has higher priority than general queue.
- These two partitions have identical default/max walltime and nodes.
- The map between user to partition is controlled by a database table, so a hall coordinators can request to change a user-partition map if needed.

How to use Auger-Slurm

During transition period (before June 4)

- Using swif, add **-slurm** option when call add-job to direct the jobs to slurm.
- Using auger, call all auger commands (jsub/jkill etc.) **/site/scicomp/auger-slurm/bin**

After June 4 (slurm will be the only batch system)

- All auger commands under /site/bin will be linked to /site/scicomp/auger-slurm/bin.
- Swif default will be same as “-slurm”.

Documentation

https://scicomp.jlab.org/docs/auger_slurm