Slurm Workload Manager Hands-On 03

In this hands-on, you are going to simulate a very large environment thanks to tuned Slurm setup. You will be able to setup different Job Preemption mechanism based on partition priority.

Estimated time : 25 minutes

ToDo

- Login to SuSE Studio Slurm Appliance
  https://susestudio.com/a/MnLYey/slurm
- Click on Test Drive, accept the EULA
- Enable the Networking
- Connect via SSH
- Stop the Slurm daemons
  ```
  service slurm stop
  ```
- Activate job preemption based on partition priority
  ```
  PreemptType=preempt/partition_prio
  PreemptMode=suspend,gang
  ```
- Setup the defatult preemtion mode to Suspend,gang
  ```
  PreemptMode=suspend,gang
  ```
- Extend the partition description with:
  ```
  low:     Shared=FORCE:1 Priority=25
           PreemptMode=suspend
  medium:  Shared=FORCE:1 Priority=20
           PreemptMode=off
  requeue: Shared=NO  Priority=50
           PreemptMode=requeue    GraceTime=120
  high:    Shared=FORCE:1 Priority=100
           PreemptMode=off
  ```
- Start the daemon with the following command line:
  ```
  service slurm startclean
  ```
- Submit large job as hpckp01 user through low priority queue.
- Submit another one with high priority queue requiring one of the nodes allocated for previous job. i.e.: `--nodelist=hsw0001` in order to force to suspend the low priority job.
- Submit a job using application-level checkpointing through `requeue` partition with user hpckp01. The first example will expect a Signal in order to generate the CHK. The second one will do it periodically.
  ```
  sbatch ~/src/Fortran/sigex_cr.sh
  sbatch ~/src/Fortran/periodic_cr.sh
  ```
- Once the job is running, submit another one requiring one of the nodes allocated for previous job in order to force preemption.

Requirements

- Desktop or Laptop with SSH client.

Quick Reference

Slurm commands

- sacct Extract accounting information from cluster.
- scancel Job deletion.
- sinfo Shows status information about cluster.
- squeue Status listing of jobs and queues.
- sview GUI to view job, node and partition information.
- smap CLI to view job, node and partition information.
- sbatch Command line interface to submit jobs.

Simulate Usage Activity

- You can use submit script based on examples folder.
- Or use the following long command line:
  ```
  sbatch -p high -n 512  \
         --wrap="env; srun -n 1 sleep 120"
  ```
- You can populate the virtual cluster with jobs using job arrays : `--array=1-1000`
- You can submit as a different user with **su** command line:
  ```
  su - hpckp01 -c "submit command"
  ```

Reference

- Slurm Training Slides
- Slurm Rosetta Stone
- Slurm Official Documentation