HPC Knowledge Portal

Slurm Training '15

Slurm Workload Manager Hands-On 01

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 hosts and partitions, and submit some jobs to this virtual cluster.

Estimated time : 30 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
- Add 512 new nodes (hsw[001-512]):
  - 2 sockets & 12 cores/socket
  - 256 GB
- Add 1024 new nodes (knl[0001-1024]):
  - 1 sockets & 72 cores/socket
  - 4 threads/core
  - 256 GB
- If you want, you can simulate a 9300 Intel Knights Landing cluster and call it Cori ;-)
- Setup a fat-tree topology in topology.conf
- Start the daemon with the following command line:
  ```
  service slurm startclean
  ```
- You can create a submit script based on example (right column) or in ~/src/examples
  ```
  sbatch -p high ~/src/examples/job_array.sh
  ```
- 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"
  ```

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.

Commonly used SLURM variables

- $SLURM_JOBID
- $SLURM_JOB_NODELIST : i.e. hsw[004,006]
- $SLURM_NNODES : Number of nodes
- $SLURM_SUBMIT_DIR

Submit Script Example

```
#SBATCH -J JobName
#SBATCH --time=00:10:00       # Walltime
#SBATCH --mem-per-cpu=1024    # memory/cpu (in MB)
#SBATCH -D /project/XXXX      # Working Directory
#SBATCH -o job-%j.%N.out      # OPTIONAL
#SBATCH -e job-%j.%N.err      # OPTIONAL
#SBATCH --mail-type=ALL       # email notification
#SBATCH --mail-user=youremail
#SBATCH -C hsw                # hsw=Haswell,knl=Knights Landing
#SBATCH --cpus-per-task=12    # 12 OpenMP Threads
#SBATCH --ntasks=72           # Number of MPI tasks
#SBATCH --nodes=36-72         # Number of nodes
##SBATCH --threads-per-core=1 # Only for HT and Intel Phi
#SBATCH --array=1-500         # Array definition
```

Reference

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