CS 4230: Parallel Programming

Lecture 4a: HPC Clusters

January 23, 2017
Outline

• Supercomputers
• HPC cluster architecture
• OpenMP+MPI hybrid model
• Job scheduling
• SLURM
Supercomputers

• Remember Top500 from a previous lecture?
• A supercomputer can be seen as a (large) collection of computing elements connected by a (often high-speed) network infrastructure (eg: Infiniband).
HPC Clusters

• You will be getting CHPC accounts soon (if not already)

   Available clusters,

   Ember, Kingspeak, Lonepeak, ...

• www.chpc.utah.edu
MPI+OpenMP hybrid model

https://computing.llnl.gov/tutorials/parallel_comp/images/hybrid_model.gif
Job Scheduling

• More users, less resources
• Job scheduling policy should ensure QoS, fairness, ...
• ssh-ing will land you on a ‘login node’
• Do **NOT** execute on ‘compute nodes’ directly
  – Exception: interactive nodes
• Always submit jobs to the job scheduler and it will run your jobs when resources are available
SLURM scripts

#!/bin/csh
#SBATCH --time=1:00:00 # walltime, abbreviated by -t
#SBATCH --nodes=2 # number of cluster nodes, abbreviated by -N
#SBATCH -o output.file # name of the stdout
#SBATCH --ntasks = 16 # number of MPI tasks, abbreviated by -n
#SBATCH --account=baggins # account - abbreviated by -A
#SBATCH --partition=kingspeak # partition, abbreviated by -p

# setenv, export, etc ...

# load appropriate modules
module load [list of modules]

# run the program
mpirun/aprun/srun [options] my_program [options]
SLURM commands

• `sbatch script`
• `squeue [-u username]`
• `scancel job_id`

```
tharindu@edison11:~$ squeue

 JOBID USER ACCOUNT NAME PARTITION QOS NODES TIME_LIMIT TIME  ST START_TIME
3513839 sstokes mp107 lbnoise debug debug  12 10:00  7:02  R  2017-01-18T09:14:34
3513796 zrsun m1394 2.55604 debug debug 10 30:00 25:46  R  2017-01-18T08:55:50
3487523 yuchung1 m2653 my_job regular normal  5 12:00:00 11:54:56  R  2017-01-17T21:26:40
3513846 lauch m77 my_job regular normal  1 10:00  4:01  R  2017-01-18T09:17:35
3513848 lauch m77 my_job regular normal  1 10:00  4:01  R  2017-01-18T09:17:35
3513860 lauch m77 my_job regular normal  1 10:00  4:01  R  2017-01-18T09:17:35
3513861 lauch m77 my_job regular normal  1 10:00  4:01  R  2017-01-18T09:17:35
3513570 desiproc desi DR4Tests shared normal  1 1:00:00 52:35  R  2017-01-18T08:29:01
3513872 lauch m77 my_job regular normal  1 10:00  1:58  R  2017-01-18T09:19:38
3513873 lauch m77 my_job regular normal  1 10:00  1:58  R  2017-01-18T09:19:38
3513572 desiproc desi DR4Tests shared normal  1 1:00:00 51:28  R  2017-01-18T08:30:08
3513575 desiproc desi DR4Tests shared normal  1 1:00:00 49:56  R  2017-01-18T08:31:40
3513681 mtnguyen m1924 aae debug debug  8 30:00 17:48  R  2017-01-18T09:03:48
3513682 mtnguyen m1924 ta2 debug debug  8 30:00 17:48  R  2017-01-18T09:03:48
3513744 hteng mp9 ces1.ctrl regular normal  90 30:00 16:13  R  2017-01-18T09:05:23
3513746 hteng mp9 ces1.ctrl regular normal  90 30:00 16:13  R  2017-01-18T09:05:23
3513748 hteng mp9 ces1.ctrl regular normal  90 30:00 16:13  R  2017-01-18T09:05:23
3513827 taojiang m1996 K15 debug debug  8 30:00 12:37  R  2017-01-18T08:58:59
3513339 gerbino planck ACT_lcdm regular normal  1 1:00:00 40:44  R  2017-01-18T08:40:52
3513774 gedong m808 gtcrun debug debug 32 30:00 10:21  R  2017-01-18T09:11:15
```

01/23/2017
CS4230
References

- SLURM tutorials,
  
  https://slurm.schedmd.com/tutorials.html