Resource Management

Addisu Gezahegn
University of Trieste
ICTP,Trieste
asemie@ictp.it

Resource Manager

- Is a centralized service that keeps track of the state of the resources availability and capability

- It accepts job submissions from users

- It delivers jobs to execution nodes

- Informs users about job status

- Maintain history of completed jobs

- Monitors progress of executing jobs

TORQUE Resource manager

- Terascale open source Resource and QUEue manager developed based PBS project

- Provides control over batch jobs and distributed computing resources

- It can be interfaced with schedulers such as maui, in which the scheduler(maui) decide which job to run during each scheduling cycle and directs torque based on policy settings

- Here we will use Maui scheduler instead of scheduler component of TORQUE

TORQUE daemons

- pbs – server

  ➢ Main controlling daemon

  ➢ Maintain state information in a database

- pbs – mom

  ➢ Runs in all nodes

  ➢ Communicate nodes status to pbs – server

  ➢ Agent that launches executable on nodes on behalf of pbs – server and monitors execution

- Trquathd

  ➢ User authentication daemon

  ➢ Runs on nodes that executes client commands
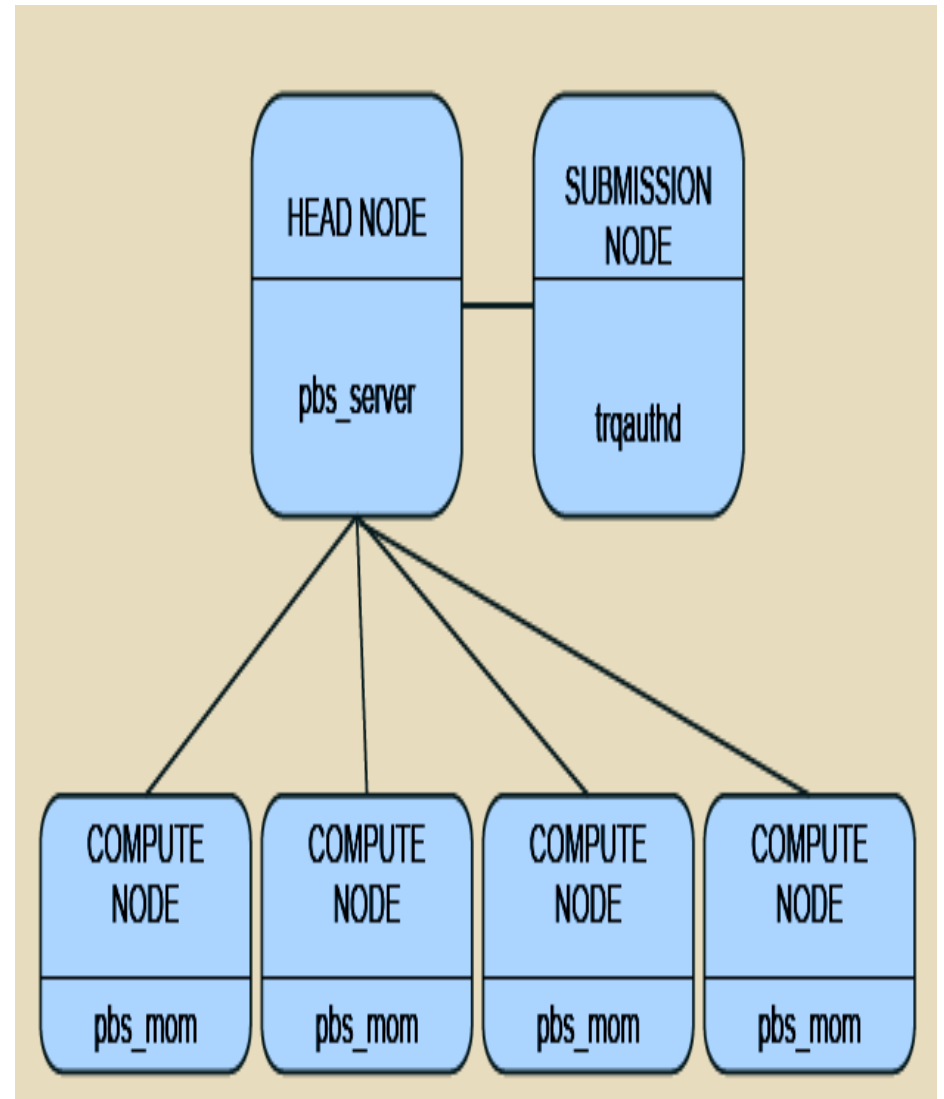
Submitting a job

- Creat a job script that contains information about the resource required and a command that you want to execute

- Use qsub command to submit your job to the resource manager

- Batch systems usually configured with multiple queue

```
#!/bin/tcsh

#PBS -N cyclone43

#PBS -l nodes=4:ppn=12

#PBS -l walltime=24:00:00

#PBS -q esp

#PBS -e cyclone_error

#PBS -o cyclone_output

echo \"Working directory is $PBS_O_WORKDIR\"

cd $PBS_O_WORKDIR

echo Running on host `hostname`

echo Time is `date`

echo Directory is `pwd`

module load openmpi/intel

module load netcdf/intel

cd $PBS_O_WORKDIR

mpirun -np 48 ./wrf.exe
```

Submitting a job

- To submit an interactive job use

 ➢ qsub -I -q serial -l walltime=1:00 -l nodes=1:ppn=2

 ➢ -I interactive

 ➢ -q qname submit a job to qname

 ➢ -l nodes=m:ppn=n, request m execution nodes with n processors for each

 ➢ -l walltime, request the time the jobs needs to use the resource

TORQUE commands

- To get information about queue status use qstat

  - qstat -r shows only running jobs

  - qstat -rn shows running jobs with allocated nodes

  - qstat -i shows idle jobs

  - qstat -u username shows jobs with the named user

- pbsnodes, examine nodes statistics

Maui job scheduler

- It is an open source job scheduler for clusters and supercomputers

- It is capable of supporting an array of scheduling policies

- It can decide job priorities dynamically based on job factors

- One can also configure Fairshare on maui that takes into account usage history for a user/group for priority calculation

Maui Scheduling Iteration

- Determine if job is eligible for considering in scheduling decision

- Priority assigned for each job based on several weighted factors such as time in queue, amount of resource requested, the owner...

- Begin with highest priority jobs

- Determine if the resource is available and place a reservation

- Communicate to resource manager to start jobs on reserved resources

Maui commands

- showstats, shows various accounting and resource usage statistics

- showq, display information about active, eligible, blocked or recently completed jobs

- showbf, shows what resources are available for immediate use

- checknode, shows detailed state information and statistics for nodes that run jobs

- mdiag, display information about various aspects of the cluster and the results of internal diagnostic tools

Thank you!
&
Let's go to Hands On