Resource Managers, Schedulers, and Grid Computing

James E. Prewett

October 8, 2008
Resource Managers

Practical: TORQUE Installation and Configuration

Schedulers

Practical: Maui Installation and Configuration

Grid Computing
What is a Resource Manager?

- Run jobs on [sets of] nodes
- Reports on resource utilization
- ... *that’s about it*
Popular Resource Managers

- TORQUE
- SLURM
- PBSPro
- OpenPBS\(^{1}\)
- Sun Grid Engine (SGE)
- LSF

\(^{1}\)DO NOT USE UNLESS YOU HAVE A DARN GOOD REASON!!!! USE TORQUE INSTEAD!
PBS Family Tree

NASA’s PBS

PBSPro

OpenPBS

TORQUE
**TORQUE Information**

<table>
<thead>
<tr>
<th>Vital Statistics:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Version:</td>
<td>2.3.3</td>
</tr>
<tr>
<td>Date:</td>
<td>August 15, 2008</td>
</tr>
<tr>
<td>Language:</td>
<td>C</td>
</tr>
<tr>
<td>Distribution Formats:</td>
<td>tar.gz</td>
</tr>
<tr>
<td>URL:</td>
<td></td>
</tr>
</tbody>
</table>
Building TORQUE

No surprises here...

- tar zxvf torque-2.3.3.tar.gz
- cd torque-2.3.3/
- ./configure
- make
- make install

and / or

- make packages²

- NOTE: No “init scripts” are installed by the make install step!
  Some examples exist in the contrib/init.d directory in the source package.

²This optional step will make self-extracting shell-script archives that you can unpack on your production machines.
Configuring TORQUE

An Execution Queue

Once the TORQUE server, pbs_server is running:

There /should/ be an execution queue named “workq” already defined, but just in case...

create queue workq
set queue workq queue_type = Execution
set queue workq enabled = True
set queue workq started = True
Configuring TORQUE (cont.)

A Debug Queue

A queue with a small wallclock and node limit for debugging purposes. Intended for quick turn-around times.

create queue debug
set queue debug queue_type = Execution
set queue debug resources_max.nodect = 1
set queue debug resources_max.nodes = 1
set queue debug resources_max.walltime = 00:30:00
set queue debug enabled = True
set queue debug started = True
A Routing Queue

A queue that decides which queue to route jobs into based upon their wallclock and node requirements.

```
create queue route
set queue route queue_type = Route
set queue route route_destinations = workq
set queue route route_destinations += debug
set queue route enabled = True
set queue route started = True
```
Configuring TORQUE (cont.)

Miscellaneous config options

Other TORQUE Server settings:

```
set server scheduling = True
set server managers = root@your.domain
set server managers += root@localhost
set server default_queue = route
set server resources_default.walltime = 00:30:00
set server node_ping_rate = 5
set server node_check_rate = 60
# for /fast/ job turn around
# a value like 60 or even 120 may be more reasonable for
# your system
set server job_stat_rate = 5
```
What is a Scheduler?

- Schedulers decide which jobs run in which order
- based on users’ importance
- based on job size
- based on job wallclock time
- based on 'fairshare'
- complex combinations of the above
Popular Schedulers

- MAUI
- MOAB
- LoadLeveler
- PBSPro
- LSF
- Sun Grid Engine (SGE)
### Maui Information

**Vital Statistics:**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Version:</strong></td>
<td>3.2.6 – Patch 19</td>
</tr>
<tr>
<td><strong>Language:</strong></td>
<td>C</td>
</tr>
<tr>
<td><strong>Distribution Formats:</strong></td>
<td>tar.gz</td>
</tr>
</tbody>
</table>
/Basic/ Maui Config (Part 1)

SERVERHOST your.server.edu
# primary admin must be first in list
ADMIN1 root

# Resource Manager Definition
RMCFG[YOUR.SERVER.EDU] TYPE=PBS@RMNMHOST@

# how often to query the resource manager (PBS) - 30 seconds
RMPOLLINTERVAL 00:00:30
RMPORT[15004]
SERVERPORT 42559
SERVERMODE NORMAL

# Admin: http://supercluster.org/mauidocs/a.esecurity.html
LOGFILE maui.log
LOGFILEMAXSIZE 10000000
LOGLEVEL 3

# Job Priority: http://supercluster.org/mauidocs/5.1jobprioritization.html
QUEUETIMEWEIGHT 1

# FairShare: http://supercluster.org/mauidocs/6.3fairshare.html
#FSPOLICY PSDEDICATED
FSDEPTH 7
FSINTERVAL 86400
FSDECAY 0.80
/Basic/ Maui Config (Part 2)

# Backfill: http://supercluster.org/mauidocs/8.2backfill.html
# BACKFILLPOLICY        BESTFIT
BACKFILLPOLICY        FIRSTFIT
RESERVATIONPOLICY     CUNANOENTHIGHEST

# Node Allocation: http://supercluster.org/mauidocs/5.2nodeallocation.html
NODEALLOCATIONPOLICY  MINRESOURCE

# QOS: http://supercluster.org/mauidocs/7.3qos.html
GROUPCFG[systaff]     PRIORITY=1000
GROUPCFG[DEFAULT]     PRIORITY=1
CREDWEIGHT 1
USERWEIGHT 1
USAGEWEIGHT 10
USERCFG[DEFAULT]      FSTARGET=25.0
USERCFG[download]     PRIORITY=100
USERCFG[special]      PRIORITY=10
USERCFG[DEFAULT]      PRIORITY=1
According to Ian Foster, a “Grid”:

- coordinates resources that are not subject to centralized control
- uses standard, open, general-purpose protocols and interfaces
- delivers nontrivial qualities of service
- ... most people just mean Globus ;)
## Globus Toolkit Information

<table>
<thead>
<tr>
<th>Vital Statistics:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Version:</td>
<td>4.2.0</td>
</tr>
<tr>
<td>Date:</td>
<td>September 2, 2008</td>
</tr>
<tr>
<td>Distribution Formats:</td>
<td>tar.gz</td>
</tr>
<tr>
<td>URL:</td>
<td><a href="http://www.globus.org/">http://www.globus.org/</a></td>
</tr>
</tbody>
</table>