Integrating SGE and Globus in a Heterogeneous HPC Environment

David McBride
<dwm99@doc.ic.ac.uk>

London e-Science Centre, Department of Computing, Imperial College
Presentation Outline

- Overview of Centre resources
- A brief description of Globus
- Integration of SGE
- Deployment issues
- Further information
LeSC Hardware

Saturn E6800
Codon V880
Sunshine V880
Rhea V880

Viking Linux SMP P4

Pioneer Linux Athlon

Atlas Tru64 SMP Alpha

Mariner Linux SMP P4

SGEEE
SGE
RMS
LeSC Firewall
Imperial Firewall
Sun Grid Engine Deployment

- SGEE v5.3 deployed on each cluster
  - Uses reserved port-based authentication
  - Single scheduler for all production resources
  - Multiple queues exist with different constraints
    - A subset of nodes have restrictions on maximum runtime to allow a quick response time for short tasks
    - Multiple environments are available which specify how many nodes a job may use. Some are restricted.
    - Myrinet-equipped nodes share a primary queue for parallel jobs and a secondary, restricted queue for sequential tasks
The Globus Toolkit

- Open-source project managed by the Globus Alliance (formerly Globus Project)
- Provides a hosting environment for OGSA Grid Services
- Also provides an implementation of core services and client-side tools
Major Globus version differences

- Globus 2.x implements bespoke standards
  - Job execution framework uses a separate gatekeeper process
- Globus 3.x implements OGSI standards
  - Job execution framework uses a grid service to provide gatekeeper functions.
- Both use a modular Perl “job-manager” to handle scheduler interaction
Globus 3.x Job Execution Architecture

See http://www-unix.globus.org/developer/gram-architecture.html for more details
Globus and SGE integration

- Globus jobs are specified using RSL, the Resource Specification Language
- Used to construct the local execution environment and specifies the program to be executed.
- job-manager executes parsed RSL instructions and monitors running jobs
- LeSC Perl module provides SGE-specific backend within the Globus framework
Globus 3.x and SGE integration

Globus Server

Master Hosting Environment (MHE)
- Master Managed Job Factory Service (MMJFS)
- Suid-root wrapper

User Hosting Environment (UHE)
- Managed Job Factory Service
- Managed Job Service

User
- jbloggs

Proxy

Scheduler

Sun Grid Engine
- qsub
- qstat

Job Manager
- Fork
- SGE

Fork
Deployment Issues

- Globus Toolkit is an evolving codebase
- OGSI-capable 3.x series is relatively new
- Not quite production quality
  - UHE's are intended to be short lived but are sometimes not reaped
  - Components initialised on demand sometimes exceed timeouts set for their response
• Globus Alliance:
  http://www.globus.org/

• SGE and Globus integration:
  http://www.lesc.ic.ac.uk/projects/epic-gt-sge.html
  – GPT packages and documentation freely available