Overview

• The ReSS Project (collaboration, architecture, …)
• ReSS Validation and Testing
• Project Status and Plan
• ReSS Deployment
The OSG Resource Selection Service (ReSS)

The ReSS Project

- The Resource Selection Service implements cluster-level Workload Management on OSG.
- The project started in Sep 2005
- Sponsors
  - DZero contribution to the PPDG Common Project
  - FNAL-CD
- Collaboration of the Sponsors with
  - OSG (TG-MIG, ITB, VDT, USCMS)
  - CEMon gLite Project (PD-INFN)
  - FermiGrid
  - Glue Schema Group
Motivations

- Implement a light-weight cluster selector for push-based job handling services
- Enable users to express requirements on the resources in the job description
- Enable users to refer to *abstract* characteristics of the resources in the job description
- Provide soft-registration for clusters
- Use the standard characterizations of the resources via the Glue Schema
The OSG Resource Selection Service (ReSS)

Technology

- ReSS basis its central services on the Condor Match-making service
  - Users of Condor-G naturally integrate their scheduler servers with ReSS
  - Condor information collector manages resource soft registration
- Resource characteristics is handled at sites by the gLite CE Monitor Service (CEMon)
  - CEMon registers with the central ReSS services at startup
  - Info is gathered by CEMon at sites running Generic Information Providers (GIP)
  - GIP expresses resource information via the Glue Schema model
  - CEMon converts the information from GIP into old classad format. Other supported formats: XML, LDIF, new classad
  - CEMon publishes information using web services interfaces
The OSG Resource Selection Service (ReSS)

Architecture

- Info Gatherer is the Interface Adapter between CEMon and Condor
- Condor Scheduler is maintained by the user (not part of ReSS)
The OSG Resource Selection Service (ReSS)

Resource Selection Example

universe = globus
globus scheduler = $$(GlueCEInfoContactString)
requirements = TARGET.GlueCEAccessControlBaseRule == "VO:DZero"
executable = /bin/hostname
arguments = -f
queue

MyType = "Machine"
Name = "antaeus.hpcc.ttu.edu:2119/jobmanager-lsf-dzero.-1194963282"
Requirements = (CurMatches < 10)
ReSSVersion = "1.0.8"
TargetType = "Job"
GlueSiteName = "TTU-ANTAEUS"
GlueSiteUniqueID = "antaeus.hpcc.ttu.edu"
GlueCEName = "dzero"
GlueCEUniqueID = "antaeus.hpcc.ttu.edu:2119/jobmanager-lsf-dzero"
GlueCEInfoContactString = "antaeus.hpcc.ttu.edu:2119/jobmanager-lsf"
GlueCEAccessControlBaseRule = "VO:dzero"
GlueCEHostingCluster = "antaeus.hpcc.ttu.edu"
GlueCEInfoApplicationDir = "/mnt/lustre/antaeus/apps"
GlueCEInfoDataDir = "/mnt/hep/osg"
GlueCEInfoDefaultSE = "simgorgh.hpcc.ttu.edu"
GlueCEInfoLRMSType = "Istf"
GlueCEPolicyMaxCPUTime = 6000
GlueCEStateStatus = "Production"
GlueCEStateFreeCPUs = 0
GlueCEStateRunningJobs = 0
GlueCEStateTotalJobs = 0
GlueCEStateWaitingJobs = 0
GlueClusterName = "antaeus.hpcc.ttu.edu"
GlueSubClusterWNTmpDir = "/tmp"
GlueHostApplicationSoftwareRunTimeEnvironment = "MountPoints,VO-cms-CMSSW_1_2_3"
GlueHostMainMemoryRAMSize = 512
GlueHostNetworkAdapterInboundIP = FALSE
GlueHostNetworkAdapterOutboundIP = TRUE
GlueHostOperatingSystemName = "CentOS"
GlueHostProcessorClockSpeed = 1000
GlueSchemaVersionMajor = 1
...
The OSG Resource Selection Service (ReSS)

Glue Schema to old classad

Mapping

Mapping the Glue Schema “tree” into a set of “flat” classads:
all possible combination of
(Cluster, Subcluster, CE, VO)

...
Mapping the Glue Schema “tree” into a set of “flat” classads:
all possible combination of
(Cluster, Subcluster, CE, VO)
Mapping the Glue Schema “tree” into a set of “flat” classads:
All possible combination of (Cluster, Subcluster, CE, VO)
Mapping the Glue Schema “tree” into a set of “flat” classads:
All possible combination of (Cluster, Subcluster, CE, VO)
Mapping the Glue Schema “tree” into a set of “flat” classads:
All possible combination of (Cluster, Subcluster, CE, VO)
Impact of CEMon on the OSG CE

- We studied CEMon resource requirements (load, mem, ...) at a typical OSG CEs
  - CEMon pushes information periodically
- We compared CEMon resource requirements with MDS-2 by running
  - CEMon alone (invokes GIP)
  - GRIS alone (Invokes GIP) queried at high-rate (many LCG Brokers scenario)
  - GIP manually
  - CEMon AND GRIS together

- Conclusions
  - running CEMon alone does not generate more load than running GRIS alone or running CEMon and GRIS
  - CEMon uses less %CPU than a GRIS that is queried continuously (0.8% vs. 24%). On the other hand, CEMon uses more memory (%4.7 vs. %0.5).

US CMS evaluates WMS’s

- Condor-G test with manual res. selection (NO ReSS)
  - Submit 10k sleep jobs to 4 schedulers
  - Jobs last 0.5 – 6 hours
  - Jobs can run at 4 Grid sites w/ ~2000 slots
- When Grid sites are stable, Condor-G is scalable and reliable

Study by Igor Sfiligoi & Burt Holzman, US CMS / FNAL, 03/07
https://twiki.grid.iu.edu/twiki/bin/view/ResourceSelection/ReSSEvaluationByUSCMS
The OSG Resource Selection Service (ReSS)

ReSS Scalability

- Condor-G + ReSS Scalability Test
  - Submit 10k sleep jobs to 4 schedulers
  - 1 Grid site with ~2000 slots; multiple classad from VOs for the site

- Result: same scalability as Condor-G
  - Condor Match Maker scales up to 6k classads
ReSS Reliability

- Same reliability as Condor-G, when grid sites are stable
- Failures mainly due to Condor-G / GRAM communication problems.
- Failures can be automatically resubmitted / re-matched (not tested here)

Note: plotting artifact

20K jobs

130 jobs
The OSG Resource Selection Service (ReSS)

Project Status and Plans

• Development is mostly done
  – We may still add SE to the resource selection process
• ReSS is now the resource selector of Fermigrid
• Assisting Deployment of ReSS (CEMon) on Production OSG sites
• Using ReSS on SAM-Grid / OSG for DZero data reprocessing for the available sites
• Working with OSG VOs to facilitate ReSS usage
• Integrate ReSS with GlideIn Factory
• Move the project to maintenance
The OSG Resource Selection Service (ReSS)

ReSS Deployment on OSG

<table>
<thead>
<tr>
<th>Site</th>
<th>Gatekeeper</th>
<th>2007-03-05-18-00-08</th>
<th>2007-03-06-00-00-07</th>
<th>2007-03-06-12-00-09</th>
<th>2007-03-07-00-00-08</th>
<th>2007-03-07-18-00-09</th>
<th>2007-03-07-12-00-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cornell/LEPP</td>
<td>linux6211.lns.cornell.edu:2119/jobmanager-sge</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>NERSC-VM-VTB0</td>
<td>osp-vtb00.nersc.gov:2119/jobmanager-sge</td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>TTU-ANTAEUS</td>
<td>antaeus.hpcc.ttu.edu:2119/jobmanager-lsf</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>UCRHEP</td>
<td>top.ucr.edu:2119/jobmanager-condor</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>UCTier3</td>
<td>uct3-edge6.uchicago.edu:2119/jobmanager-pbs</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>UVaHEP-T3</td>
<td>osg.hep.phys.virginia.edu:2119/jobmanager-pbs</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>Vanderbilt</td>
<td>vmpg01.vampire.2119/jobmanager-pbs</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
<td>Down</td>
</tr>
<tr>
<td>red-uml.edu:2119/jobmanager-pbs</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
</tr>
<tr>
<td>stitch.esc.rutgers.edu:2119/jobmanager-condor</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
<td>Up</td>
</tr>
</tbody>
</table>

Click [here](#) for live URL
Conclusions

• ReSS is a lightweight Resource Selection Service for push-based job handling systems
• ReSS is deployed on OSG 0.6.0 and used by FermiGrid
• More info at http://osg.ivdgl.org/twiki/bin/view/Resource Selection/