FermiGrid—A Condor-based Heterogeneous Campus Grid

Steven Timm
Condor Week, May 2, 2007
Introduction

- “Let all the dirt hang out” M. Livny, 3/5/2007
- FermiGrid and the Open Science Grid
- Why have a campus grid?
- Components of FermiGrid
  - Common Site Authorization
  - Computing Clusters
  - Information System
  - Scheduling Algorithm
  - Storage
- Experiences managing large condor pools
- High Availability Services
- Future Plans

Steven Timm  http://fermigrid.fnal.gov/
Condor Week 2007
FermiGrid and Open Science Grid

- Most OSG jobs are submitted via Condor-G
  - Grid Universe / type gt2
- OSG made of 31 Virtual Organizations
  - Seven of them connected to Fermilab projects
  - Four more that we run services for.
- Typically we have three of the top five busiest sites on OSG.
- Currently 3281 VM's on the grid, 860 more nodes soon to come.
- Typical Fermilab application—load a chunk of data, chew on it for hours, send back the result. (Embarassingly parallel)
- Jobs not checkpointable, our users don't like to get pre-empted.
- Management wanted a single point of entry for outside grid jobs to Fermilab, thus FermiGrid was born.
Why Have a Campus Grid?

- Unified point of entry and management
- Present unified interface to the outside grid, masking some of the complicated cluster structure.
- Support more features, i.e.
  - Mounting of experiment-specific mass storage (applications and data)
  - Alternate methods of job submission
- Certain I/O intensive jobs run best close to home.
- Make use of [bleeding-edge pre-alpha] software not yet part of Open Science Grid.
- Steer outside jobs to where they are most likely to run.
Common Site Authorization

### Step 1
User issues `voms-proxy-init` command and receives VOMS signed credentials.

### Step 2
User submits their grid job via `globus-job-run`, `globus-job-submit`, or `condor-g`.

### Step 3
Gateway requests GUMS mapping based on VO & Role.

### Step 4
Gateway checks against Site Authorization Service clusters, sends ClassAds via CEMan to the site-wide gateway.

### Step 5
Grid job is forwarded to the target cluster.

Periodic synchronization

BlueArc

CMS WC1, CDF OSG1, CDF OSG2, D0 CAB2, GP Farm

Site Wide Gateway

VOMS Server

GUMS Server

SAZ Server
Computing Clusters

- Four major clusters
  - CDF: 120 nodes/600 VM's, soon to grow to 520 nodes/2600VM's
  - D0: currently 200 nodes/400VM's on the grid, soon to grow to 800 nodes/2000VM's.
  - CMS: currently >1900 VM's on the grid.
  - General Purpose: currently 388 VM's on the grid, 240 more coming soon.
Information System

- Generic Information Providers make information about the cluster in LDIF form, suitable for serving with an LDAP server.
- Web service called CEMon bundles this into “Old Classad” dialect and sends it to a central collector
- We end up with a condor pool which is a cluster of clusters
- One classad per cluster per VO.
- [http://fermigrid.fnal.gov/classads.html](http://fermigrid.fnal.gov/classads.html)
- Can use in two ways—>
  - Condor schedd on anyone's desktop using this as master
  - Globus gatekeeper in front of the condor_schedd
Scheduling Algorithm

- We are emulating Globus toolkit's jobmanager-condor
- Take incoming job from the grid and resubmit it as a 2nd grid job to one of our four subclusters
- Use Condor-G matchmaking to pick which one.
- Match on requirements including support of VO
- Rank on number of free CPU's
- If job is matched and doesn't start for two hours, hold it and release/rematch it to another cluster.
- Add extra field GlueRequirements to globus RSL
  - GlueRequirements are used to do the condor-G match
  - CondorSubmit requirements passed through to final cluster
Scheduling Algorithm

[root@fermigrid1 tmp]# more condor.engage.17171.1177960558.jdl
#
# description file for condor submission
#
Universe = grid
Notification = Never
WhenToTransferOutput = ON_EXIT_OR_EVICT
x509userproxy = /grid/home/engage/.globus/job/fermigrid1.fnal.gov/17171.1177960558/x509_up
Executable = /grid/home/engage/.globus/.gass_cache/local/md5/70/a2/73/ced2f900abd5a19b63da86a93b/md5/b8/cb/3d/46a47539a6e551a181bfe45022/data
globusscheduler=$$(GlueCEInfoContactString)
globusrsl = (maxWallTime=1200)(jobtype=multiple)(count=1)(condorsubmit=('+ForwardHost' 'fermigrid1.fnal.gov')('+ForwardJobID' '$(Cluster)'))
machine_list_length = 5
Rank=GlueCEStateFreeJobSlots
Arguments = '­­run­id=tyr2nd­­job­id=0000498'
InitialDir = /grid/home/engage//gram_scratch_DDMkyQAFZI
Input = /dev/null
Log = /usr/local/vdt-1.6.1/globus/tmp/gram_job_state/gram_condor_log.17171.1177960558
log_xml = True
globus_resubmit = (NumGlobusSubmits <= NumSystemHolds) & (NumSystemHolds < 11)
globus_rematch = (EnteredCurrentStatus-LastMatchTime) > 30
Output = /grid/home/engage/.globus/job/fermigrid1.fnal.gov/17171.1177960558/stdout
Error = /grid/home/engage/.globus/job/fermigrid1.fnal.gov/17171.1177960558/stderr
queue 1

Steven Timm  http://fermigrid.fnal.gov/
Condor Week 2007
Storage

- OSG allows
  - NFS-mounted DATA area or
  - Grid-accessible area served by Storage Resource Manager.
- We have both.
- DATA area served by BlueArc NAS appliance (7TB)
- Volatile SRM-dCache Grid accessible area, for temporary files while jobs are running (5TB).
Experiences Managing Large Condor Pools

- Avoid NFS unless you have an enterprise NAS server like we do.
- NFS and condor_collectors don't mix.
- Even now condor config files and executables local to each machine.
- TCP is good, use it where you can.
- Tune network parameters esp. on collector/negotiator machine.
- Can't have too much RAM or swap.
- Multiple schedd's on same machine can help distribute the load. (If you configure them right).
- Train your users not to submit thousands of 30-second jobs simultaneously.
- Quill, friend or foe?
  - Keep your DB vacuumed, make friends with Condor-DB team.
FermiGrid HA Services

![Diagram of BlueArc with heartbeats between XEN1, XEN2, and XEN3 nodes.]

Steven Timm  http://fermigrid.fnal.gov/
Condor Week 2007
FermiGrid HA Gatekeeper

BlueArc

XEN1

XEN2

XEN3

Active

Standby

Active

Standby

Active

Standby

Active

Heartbeat

FermiGrid

Steven Timm  http://fermigrid.fnal.gov/
Condor Week 2007
Future Plans

- We need to make the equivalent forwarder for Globus Toolkit 4 (Web Services)
- Need uniform convention across OSG for MPI parallel jobs
- Investigating kerberos-authenticated Condor dedicated scheduler for this purpose, and MPICH-G2
- Enhance information system
  - How much do we tell the OSG? CPU type? Architecture? How many of each we have?
  - Give enough info to users but not too much for hackers
  - Also make sure the information provider can finish before the new one starts 5 minutes later.
Why not use Condor-C?

- Condor-C now has features to submit from Condor to Condor and from Condor to PBS
- 2 ½ years ago when we started, those features didn't exist.
- OSG and FermiGrid use extended globus callouts not currently supported by Condor GSI authentication (but we hear they will be soon.)
- Schedd-on-the-side is interesting but many of our jobs currently not easily sandboxable.
- Nevertheless we are watching and learning
Result: Happy Puppies

Ziggy is happy with any system that lets his humans get home to feed him supper on time!