

OSG Lessons Learned and Best Practices

Steven Timm, Fermilab

OSG Consortium

August 21, 2006

Site and Fabric Parallel Session

Introduction



- Ziggy wants his supper at 5:30 PM
- Users submit most jobs at 4:59 PM (on Friday).
- What to do so the system will stay up and Ziggy will get his supper on time?

Outline

- Pre-Installation Planning
 - Hardware Selection
 - File Systems
- Post-install Configuration
- Troubleshooting
 - How to track a job

Hardware selection

- Compute Element head node
 - Should have at least 4GB of RAM. We have 6GB and we still swap.
 - At least 2 cpu's, as fast as you can get.
 - Enough disk to hold the VDT software and logs (250GB or so).
 - For small-medium size cluster, can be NFS server for home areas too.
- If you are supporting a VO, use a 2nd machine to be your VOMS server, of equal size to above.
- GUMS (Grid User Management System) is good way to manage users, if you're running it, use a 3rd server of equal size to above.

File systems

- Some documents have already been written—see DOCDB
- OSG Compute Element Best Practices
 - <http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=379>
 - (OSG Compute Element)
- NFS-Lite
 - <http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=382>

Four File Systems Needed

- Home areas for each user—quota 50MB per user
 - Nothing permanent to be stored here
- APP area for each VO's applications—quota 10GB per VO
 - Any way to do garbage collection here, or make VO's do it?
- DATA area for each VO's data, 2GB per job slot
 - Again, any way to make VO's clean up after themselves? right now using quotas to keep any one VO from overrunning the area
 - Even if you have a SITE_READ, SITE_WRITE storage element should have DATA too..
- WN_TMP area on each worker node
 - In past VDS has require this be a fixed path—will change for OSG-0.5.0 to be able to be assigned at execute time

To NFS or not to NFS

- Pre-web-services GRAM makes pig-headed use of NFS, tolerates NFS failures badly. (WS GRAM doesn't need so much NFS).
- Small to medium cluster (120 machines)—Get user home areas on globus gatekeeper machine itself, make second server with APP, DATA.
- Condor negotiator/collector and NFS servers don't mix.
- Big cluster--Use NFS-lite configuration or some global file system for home areas.
 - Bluearc NAS head has worked well for us thus far.
 - Others out there—Ibrix, Panasas, etc.

Pre-inst planning, Software

- Condor
 - Install first and separately out of VDT area.
 - Condor needs upgraded more than rest of VDT
 - Some versions may not be available in VDT—use RPM
 - We install rpm on every WN—helps with network trouble
 - Needed for ManagedFork even if main batch system isn't condor.
- Accounts—create them in advance. Suggest /sbin/nologin shell
- Ganglia—have your gmetad elsewhere than your head node, big network traffic generator and cpu load too.

Software, during install

- I install worker node client on head node and on each worker
- Big load on VDT server and on CRL servers—in 0.5.0 plan to use Squid to mitigate this.
- ManagedFork—This is a lifesaver. Only way to log and control what users are doing with the jobmanager-fork.
- `$VDT_LOCATION/globus/TRUSTED_CA`
 - Is the default location for certificates now
 - But `/etc/grid-security/certificates` must also be symlinked to this, otherwise Condor `>=6.7.18` doesn't know where to get certificates from.
 - Also make sure worker node client `TRUSTED_CA` points here too

Monitoring

- 4 Major monitoring packages
 - MonALISA
 - MIS-CI (feeds GridCat, did feed ACDC dashboard)
 - CEMon->Generic Information Provider
 - GRIS->Generic Information provider
- Frequencies of all monitoring software in the OSG are adjustable somehow— not enough time to go into detail here.
- Going slower than the default frequency essential to smooth operation
- More things calling condor_q, more room for trouble
- Recent report—grep can be faster if LOCALE=C rather thanUS-ENGLISH

Monitoring

- If vo is in grid3-user-vo-map.txt, need to support it
- grid3-user-vo-map.txt has to be right before GIP is configured
- generic info provider—be sure it advertises the right number of job slots, otherwise it's a “kick me” sign to the LCG.
- Stuff breaks—look at daily MonALISA email, check GRIS is up
- CEMon is new, we have been running it at FNAL under OSG 0.4.1. Can be memory-hungry.
- Do we have to split some monitoring off to another node? maybe eventually, especially with advent of web services.
- Monitoring should use condor_q -format and condor_status -format

Importance of a central submit node

- Have a central submit node!!
 - Discourage people from submitting from their desktops
 - Can't control the configuration of desktops, lots of ways to make it insecure
 - With central submit node, can kill rogue condor-G processes if necessary
 - Allows you to see the same condor errors they are seeing
 - Most serious grid work needs bigger staging disk than a desktop will provide anyway
 - Laptops turn off and jobs get held that would otherwise succeed

Tracking jobs

- From submission job, get GridJobId—`condor_q -l | grep GridJobId`
- GridJobId = "gt2 fngp-osg.fnal.gov/jobmanager-condor https://fngp-osg.fnal.gov:18764/10515/1156102663/"
- in `/var/log/messages`
- Aug 20 14:37:48 fngp-osg gridinfo[10515]: JMA 2006/08/20 14:37:48 GATEKEEPER_JM_ID 2006-08-20.14:37:43.0000010457.0000000000 for /DC=gov/DC=fnal/O=Fermilab/OU=People/CN=Charles C. Polly/UID=polly on 131.225.167.42
- Aug 20 14:37:48 fngp-osg gridinfo[10515]: JMA 2006/08/20 14:37:48 GATEKEEPER_JM_ID 2006-08-20.14:37:43.0000010457.0000000000 mapped to minboone (12755, 5468)
- Aug 20 14:37:48 fngp-osg gridinfo[10515]: JMA 2006/08/20 14:37:48 GATEKEEPER_JM_ID 2006-08-20.14:37:43.0000010457.0000000000 has GRAM_SCRIPT_JOB_ID 507209 manager type condor

Wish List

- Better audit trail in Condor_G—source job ID available to remote and vice versa
- globus-job-manager reaper, kill hung processes
- Better home area cleanup of /home/.globus/.gass_cache/*
- mysql root is open, why?
- Policies for VO's to clean up APP and DATA
- Watchdog to look for broken services.

Conclusion

- Ziggy is getting his supper!



Condor Configuration—what we do

- Main condor_schedd on same node as OSG compute element
- Second dedicated machine to be condor collector/negotiator
 - should not be an NFS server of any sort.
- Third machine for local submits to the cluster and run condorview.
- Requested 3 more machines
 - Second globus gatekeeper/condor_schedd machine for internal jobs
 - Separate server for postgres
 - Login node for submitting outbound grid jobs.

Condor continued

- From previous batch system, users are used to no preemption, we try to keep it that way.
- Group Quotas keep any one group from taking over the whole cluster with week-long jobs
- Each job auto-assigned to a group via wrapper on condor_submit
- Priority factors ensure local user jobs start ahead of OSG jobs.
- Use condor_quill on our most busy schedd. (database frontend makes condor_q, condor_history go faster).
- Use TCP to talk to the collector.
- Condor config file and condor software installed independently on each node, no nfs.
- **IMPORTANT—change HOSTALLOW_WRITE to be something other than “*”**