Batch Systems
WLCG GDB.
Oct 10th 2008

Steve Traylen, CERN, steve.traylen@cern.ch

Information Society

Enabling Grids for E-sciencE

– Review of batch systems in use on WLCG/EGEE.

– WLCG + community support for batch systems.

– Existing Documents

– Recent Changes in the EGEE Deployment

– Missing Documents and Instructions.

Batch Type	Versions	Number	Total
SGE and GE(?)	6, 6.0u8, u9 6.1u3, u4, u5	1, 4, 1 3, 5, 1	15
LoadLeveler	3.4.2.2	1	1
PBSPro	PBSPro_7.0.0, 8.0.0 PBSPro_9.1.0, 9.2.0	1, 1 3, 1	6
BQS	?	1	1
Condor	6.8.2, 3, 4, 5, 6, 8, 6.9.5 7.0.1, 2, 3, 4, 5, 7.1.0	1, 1, 1, 2, 1, 1, 3 2, 1, 4, 8, 2, 1	28
LSF	6.0, 6.1, 6.2 7, 7.0.1, 7.0.2	3, 2, 6 2, 1, 3	17
PBS and Torque	1.0.1p5, 1.1.0p0, 1.2.0p2 2.1.6, .7, .8, .9, .10 2.3.0 , .3, .4	3, 1, 1 12, 1, 6, 32, 1 95, 6, 1	159
			226

- **Numbers from** *GlueCEInfoLRMSType* **and** *LRMSVersion*
  - Could be cleaned up but they have never been used.
- **Batch systems are very diverse.**
  - 7 Batch Systems but many versions within that.
    - § Torque with Maui or pbs_sched
    - § SGE or SGE and Maui.
    - § OpenPBS , Torque and PBSPro
    - § 13 versions of Condor.
- **Torque is dominant:**
  - 70% of batch systems.
  - Only "blessed" gLite batch system.
    - § 42% of batch systems run current gLite versions, torque 3.3.0.
    - § New one will go to certification shortly...

- **EGEE Batch Systems Support Groups.**
  - https://twiki.cern.ch/twiki/bin/view/EGEE/BatchSystems
  - Torque -
    - § SA3 Core Certification Platform.
    - § General Torque integration NIKHEF and BLAH particulars INFN.
  - SGE
    - § CESGA - https://twiki.cern.ch/twiki/bin/view/LCG/ImplementationOfSGE
    - § Well supported, an active community in EGEE.
  - Condor
    - § PIC Coordinates(they use SGE?).
    - § Mailing list: project-eu-egee-batchsystem-condor@cern.ch
    - § Condor choice often motivated outside WLCG.
      - *Integration to campus grid.*

- **PBSPro**
  - Altair have offered to EGEE to help integrate to gLite better.
    - § Ask them if you have their commercial support.

Enabling Grids for E-sciencE

- **Torque and Maui User Guides**
  - Hosted by ClusterResources
  - Torque guide is good, maui is poor and plain wrong in places.
  - http://www.clusterresources.com/pages/resources/documentation.php
  - Community mailing lists are also hosted here.
- **Torque-Maui Tutorial for WLCG Ops WorkShop.**
  - Explanation of how the Maui Schedular Makes Decisions.
  - Fairshare, Account, Groups, Users, ...
  - https://twiki.cern.ch/twiki/pub/LCG/WhiteAreas/TorqueMauiTutorial-Traylen-White-April2008.pdf
- **Job Priority Working Group.**
  - Contains details on supporting, e.g high priority jobs within WLCG.
  - More generally restricting a batch queue to a particular VOMS FQAN.
    - § http://egee-intranet.web.cern.ch/egee-intranet/NA1/TCG/wgs/priority.htm
  - Simpler Document Concerning the Publishing Only
    - § http://goc.grid.sinica.edu.tw/gocwiki/How_to_publish_queues_with_access_restricted_to_a_FQAN

Enabling Grids for E-sciencE

- **MAUI Cookbook for LCG.**
  - Old, but still a good introduction.
    - § Everything is still correct.
  - These were the manual instructions for what YAIM now does.
  - http://grid-deployment.web.cern.ch/grid-deployment/documentation/Maui-Cookbook.pdf

- **Short Dead Line Jobs.**
  - Describes Torque and grid setup to run jobs "instantly".
    - § http://egee-intranet.web.cern.ch/egee-intranet/NA1/TCG/wgs/sdj.htm

- **Monitoring Jobs.**
  - Alternate simpler solution for SDJ.
    - § http://goc.grid.sinica.edu.tw/gocwiki/How_to_reserve_CPU_for_monitoring_tools

- **LSF, YAIM and Grid.**
    - § https://twiki.cern.ch/twiki/pub/LCG/WhiteAreas/lsf.pdf

Enabling Grids for E-sciencE

- **Individual Queues for VOs.**
  - VO queues not needed since VOViews are now available.
    - § Supported by YAIM and the Middleware (WMS)
  - Many sites (inc large ones) have removed their VO queues.

- **YAIM and Middleware now support.**
  - Mapping a FQAN to a particular uid pool/gid pair.
  - Submitting a FQAN to particular queue.
    - § Due to the publication of that queue.
  - Publishing a VOView for that FQAN.

Enabling Grids for E-sciencE

- **Killing Jobs at Limits.**
  - Handling memory limits in particular.
  - Some sites do this but best practices are not well established.
    - § It's a confusing topic:
      - *vmem vs pmem,*
      - *kernel version dependent*
      - *linux calculates it incorrectly anyway.*

- **Handling Inefficient Jobs**
  - Tracking down stuck jobs.
  - Understanding when they can be terminated.
    - § Who if anyone should be notified....

- **Tracing and Profiling Jobs**
  - Catching (in)frequent errors in jobs.
  - Extends beyond the batch system, security .....

Enabling Grids for E-sciencE

- **Implementing FQAN support at Batch.**
  - Currently quite distributed. YAIM, SDJ WG, JP WG.
  - The cookbook must be extended for these.
  - Testing the implementatons with voms-proxy-fake.

- **Introduction of CREAM.**
  - Initially zero changes in batch configuration.
  - CREAM *should* be a drop in with respect to batch.
  - Much of the current testing in PPS is for this model.

- **Benefiting from new CREAM Features**
  - CREAM will pass submission parameters to the BATCH system.
  - Some documentation, slides exist for the pass to batch.
    - § Previous tests of BLAH in gLiteCE were done in isolation.
      - *Lots of other changes since then, e.g VOViews.*
    - § It must be reviewed, developed and rewritten.

Enabling Grids for E-sciencE

- **Batch systems are very diverse.**
  - Type and version is decoupled from middleware.
- **There is a lot of documentation.**
  - EGEE / WLCG supplies enough for the common cases.
- **CookBook**
  - Was good but needs an update for the newer special cases.
    - § A quality reference document.
  - Some examples in cookbook reported not to work.
- **Configuring batch to benefit from CREAM.**
  - There is development and documents to write here.