



Job management: can we deliver the workload?

Maarten Litmaath
CERN



Yes, but ...

- LCG-CE
- CREAM
- WMS
- Pilots
- gLExec
- SCAS
- Argus
- SW area
- MyProxy
- WN properties
- BDII
- ...



- lcg-CE version 3.1.35 (Aug 31) brings important fixes
 - globus-gma-1.0.13 protects CE better against overload
 - cleanup-jobdirs-1.0.0 removes stale job directories
 - Fixes bug <https://savannah.cern.ch/bugs/index.php?53358>
 - Important for production and pilot jobs
 - YAIM will start the gatekeeper with GLOBUS_GMA=true
 - APEL more robust, various bugfixes



Mapping users

- Many more pool accounts are needed for ATLAS, CMS, ...
 - 200 accounts in YAIM example users.conf insufficient for experiments with many more active users
 - Only idle accounts can be recycled
 - Recycling is not desirable, should rarely happen
- Various sites have reported mapping problems
 - User proxies may have primary FQANs that are not explicitly supported by the CE
 - Typically national groups within experiment VO
 - Fallback on classic grid-mapfile may give unwanted mapping
 - User jobs may unexpectedly run e.g. under "sgm" account
 - Fix 1: add wildcard catch-all rules to YAIM groups.conf
 - Fix 2: use VO_<vo>_UNPRIVILEGED_MKGRIDMAP



- 30 instances
 - T0, 5 T1, 19 T2
- Still published with GlueCEStateStatus Special
 - To prevent accidental matching by older WMS versions
- Not a lot of experience yet
 - ALICE direct job submission OK
 - CMS have just started dedicated tests via WMS 3.2
 - Various instances found not in good shape
 - Patch #3259 will bring important improvements
- Condor-G submission not yet OK
 - Failure rates too high
 - Timeline weeks/months
 - Fixes needed for ATLAS PanDA



- Version 3.1.20 (Aug 24) has 3.2 code backported to gLite 3.1
 - Major internal redesign, many bugfixes
 - Best version ever, but check the known issues page!!!
- Can once again submit to CREAM
 - All WMS instances should be upgraded before end October
 - To allow CREAM CEs to have GlueCEStateStatus Production
 - Patch #2862 (PPS) has important improvements for ICE
- Can submit to NorduGrid
 - Used by CMS
- Advanced monitoring system developed at CNAF
 - <https://cert-wms-01.cnaf.infn.it:8443/wmsmon/main/main.php>
- Simple monitoring system used at CERN not ready for release
 - <http://wmsmon.cern.ch/monitoring/monitoring.html>



Pilot jobs

- Each experiment has a pilot job framework
 - ALICE: AliEn
 - ATLAS: PanDA
 - CMS: glideinWMS, only used on OSG
 - LHCb: DIRAC
- All examined by GDB Pilot Job Frameworks Review group
- Current usage
 - Production managers run VO workload for many/all users
 - Individual users may be able to run their own jobs
- Foreseen usage
 - Pilot jobs use glEXEC to run payload under user account
 - GlEXEC on WN maps user proxy like CE would have done
 - Shield user jobs & proxies, protect pilot proxy



- Released to production early July after long testing on PPS
 - ATLAS and LHCb adapted their frameworks to glexec
 - Can use glexec optionally (and if it works)
 - ALICE are adapting their code
- Not yet available on SL5
 - patch #3192 in certification
- So far only the PPS testing sites deployed it
- Comes with a new service → SCAS
 - Remembers the mappings
 - Can share the gridmapdir with LCG-CE, CREAM, ...
- An alternative new service is being certified → Argus
 - Seen as longer-term strategy
 - Can share the gridmapdir with SCAS etc.



Other concerns

- Shared SW area can become a bottleneck
 - Problems reported by various big sites
 - ALICE suggest BitTorrent!
- MyProxy
 - Single point of failure for many scenarios
- WN
 - Efficiency of matching job requirements to WN properties
 - <https://twiki.cern.ch/twiki/bin/view/EGEE/WNWorkingGroup>
- BDII
 - Version 5 should be more robust
 - GFAL failover will skip “empty” BDIIs (“No entries for host”)
 - LCG_GFAL_INFOSYS should have ordered list of BDII
 - Some tools cannot handle that yet