NBS project :
Batch migration to
Grid Engine

Philippe.Olivero @cc.in2p3.fr

Rencontre LCG-France 1 décembre 2011

Overview

o **Reminder**

o **Steps of the project**

o **Migration to GE : Method**

o **Deployment : fixed technical problems and Oracle Support**

o **Current situation**

o **Conclusion : benefit and loss**

Reminder

o **ESC recommandations:**

- o **2008 Local batch system review submitted to foreign external experts**
- o **2009 Decision to change the bactch system : Local study of batch systems**
- o **2010 Decision to start a migration to Grid Engine (versus LSF)**

o **Main reasons :**

- o **To scale face to predicted CPU load increase**
- o **To decrease FTE's (3) necessary to manage BQS**
- o **To break CC isolation in the HEP community**
- o **To minimize efforts to adapt the Batch system to Grid Middleware changes**

o **GE in the Hep communnity**

- o **Oracle GE : only CC**
- o **Oracle Univa : Desy, Kit, PDSF, GSI, Cesga + some small sites**
- o **Son Of Oracle : 3, 4 sites**

Steps

o **2009 June**   **Decision to change the batch system**

o **2010 February**   **Decision to select Grid Engine (vs LSF)**

o **2010 May**   **Pre-production GE cluster for CC-experts to work**

o **2010 October**   **Pre-production cluster open to CC : usage and work**

o **2010 December**   **Pre-production cluster open to selected guinean pigs**

o **2011 April**   **Pre-production cluster open to users, calls for migration**

o **2011 June**   **Official Production cluster [29%]   with Oracle Support**

o **2011 November**   **Production cluster [~ 90%]   -- Wlcg jobs on GE**

o **2011 December**   **6th : End of migration to GE – BQS decommissioned**
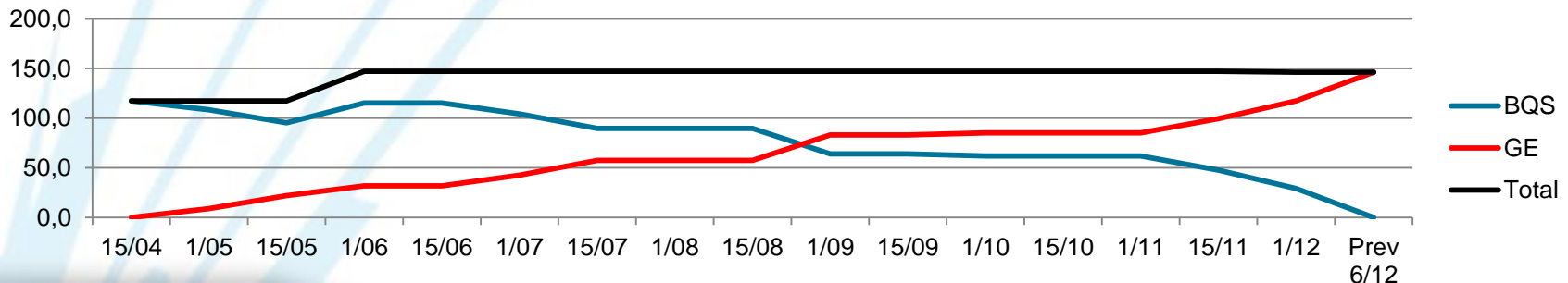
Migration to GE - Method

o **A project manager appointed by CC Steering Commitee**

o **Constitution of a first group to cope with NBS project (New Batch System)**

    o **the Operation Group leader, 1 batchMaster, 1 BQS developer,**

        1 SysAdmin, 2 help-desk, 1 Decisional (accounting) and the Quality manager

    o **Episodic collaborators from other teams :**

        Accounting, SysGrid, Dev, Storage, sysAdmins

o **Definition of milestones and technical problems to solve**

  A responsible appointed by the group per each technical problem to follow

o **~2 weekly basis coordination meeting**

o **Intermediate specific technical meetings with CC experts**

o **Tests clusters to investigate, to try solutions, to set configurations,**

o **Tests by first guinea pigs**

Deployment : fixed technical points

o **AFS tokens renewal**

o **Force local disk usage for stdout and stderr (and not in /afs::$HOME)**

o **Disk space limits per job**

o **Accounting compatibility, local and Grid reports**

o **Jobs flow regulation to control storage systems accesses (using GE facilities)**

o **GPFS secured access control**

o **Improvements of Jobs information (especially for Grid Jobs)**

o **gLite Cream-CE adaptation (on going process to include this work in official release)**

o **(quite) light developments (in prlog/epilogs)**

o **Efforts on internal and end-users documentation**

Deployment : Oracle Support tickets

o **Available since June 2011**

o **Loss of AFS tokens when masters switch**   **[patched]**

o **Loss of worker spool files at end of Job**   **[patched]**

o **Episodic and unexplained long scheduling execution**   **[upgrade recommanded]**

o **Bug in midnight jobs accounting**   **[workaround]**

o **Statistics since June 2011:**

    o **9 open tickets**

    o **3 correctly closed (patches and workaround)**

    o **6 pending answers for now (5 waiting state, 1 upgrade)**

o **A globally good satisfaction for Oracle reactivity and answers**

o **Planned upgrades (current is 2.6 Update 5 with corrective patches):**

    o **Update 7 as soon as patches are available**

    o **Update 8 when release is available (2012Q1)**

Current situation

o **~90% really migrated today**

o **~ 5 % minimum will be migrated by December 6th**

o **LHC grid jobs totally out of BQS since November 24th**

o **Last machines will be migrated during outage of December 6th**

o **Use of Oracle support with a (pretty) good reactivity**

o **Ready to scale for the future needs**

o **BQS decommissionning on time**

o **Planned actions to make CC developed code for Cream-CE in the official release**

o **First actions to initiate a GE community [mainly in HEP]**

   o **Hepix Reports (Batch survey and study, Tests robustness, NBS project)**

   o **First list of contacts (Kit, Desy and smaller sites) to organise a GE-workshop**

Benefit and Loss

o **Pros**

- o **Batch System FTEs decreased from 3 (inc. 2.2 Dev) to 1.5 (will still decrease)**

- o **One single farm for all types of jobs**

- o **Ease of administration and configuration per user/group/machine, …**

o **Cons**

- o **Lack of jobs information**

- o **No smooth spawn of jobs (« Distribution rate »)**

Questions ?