



# Migration BQS vers GE

Suzanne Poulat – groupe Exploitation





# Plan



- Généralités
- Mutualisation du cluster
- Régulation de la production
- Limites de GE
- Nouvelles fonctionnalités
- Plateforme actuelle
- Documentation



# Généralités



- Queues d'exécution standard
  - Short, medium, long
- Queues spécialisées associées à des regroupements de machines, et réservées à des groupes d'utilisateurs
  - Parallèles, Interactives, « demon », beaucoup de mémoire ou CPU
- Requêtes possibles : cpu, mémoire, disque, durée de résidence en machine, architecture, os, complexes, projet...
- Dépassements contrôlés
  - de façon native : cpu, mémoire, temps de résidence.
  - Via les quotas pour le disque



## Mutualisation du cluster (1)



- « Fair share » à 2 niveaux et par projets
- Entre groupes. A chaque groupe est associé :
    - Un projet
    - Un nombre de tickets, proportionnel à l'objectif
  - Intragroupes (différents rôles au sein du groupe).  
A chaque rôle est associé :
    - Un projet
    - Un pourcentage
    - Un ensemble d'utilisateurs



## Mutualisation du cluster (2)



### ■ Exemple

- Atlas : 2400
  - T1 : 50%
  - T2 : 50%
    - » Production 10%
    - » Analyse 90%
- Lhcb : 1500
  - Pilot : 90%
  - Production : 10%
- Hess : 140

### ■ Projets par défaut

- au niveau cluster (pour ceux qui omettent de préciser le projet)
- Au niveau du groupe (pour ceux qui ont plusieurs projets)

### ■ Possibilité de répondre aux urgences

- Priorité associée au job
- Supplément de « faire-share » au niveau du groupe ou du projet



## Régulation de la production (1)



- Régulation des accès aux services de stockage
  - Un complexe par service (hpss, sps, xrootd, dcache, etc) avec une limite globale
  - Une limite par groupe, projet ou utilisateur
  - Ne peut pas être à usage exclusif d'un groupe
  - ~20 complexes pour ~200 ressources dans BQS
- Limite générale par groupe et/ou projet et/ou utilisateur
- Possibilité d'interdire soumission à un utilisateur ou un groupe ou à tous



## Régulation de la production (2)



- Régulation de la charge des machines
  - Classement des machines en fonction de la charge (cpu, mémoire, disque)
  - Seuils limites de charge qui bloquent l'entrée de nouveaux jobs
  - Plus un nombre limités de jobs par machines



## Limites de GE



- Informations sur les jobs (pas de « qselect »)
  - qstat (jobs en queue et en machine)
  - qacct (jobs terminés, infos dans fichier plat)
- Pas de normalisation du cpu
  - qsub avec valeurs réelles (évaluation du temps sur les machines les moins puissantes)
  - qacct somme les cpu réels
  - Mais coefficient sur chaque machine, utilisé pour mise en exécution des jobs (fair share), qstat et l'accounting « maison »
- Complexes :
  - pas de « distribution rate »
  - pas de notion de propriété, ni de statut
- qsub est permissif : les droits d'accès aux queues, et les limites sont vérifiés lors du passage en exécution
- Les « profiles » ne sont pas « sourcées » au démarrage



## Nouvelles fonctionnalités



- Limite sur temps de résidence en machine -> pas de contrôle « maison » sur les jobs
- Jobs paramétriques
- Jobs interactifs (sur un pool de machines)
- Jobs parallèles (sur un pool de machines) et jobs monoprocesseurs dans le même cluster GE
- Plus de fonctionnalités coté jobs parallèles
- Possibilité de réservation
- Accès exclusif



## Plateforme de Pré-production



- Configuration :
  - 38 workers : ccwge0001-38
  - une interactive ccige
  - un CE avec CreamCE : cccreamceli05
- Début mars : ouverture de la ferme de pré-production
- Début avril : ouverture du cluster de production



- GE : soumettre un job :  
<http://cc.in2p3.fr/docenligne/969>
- GE : liste des commandes :  
<http://cc.in2p3.fr/docenligne/970>
- Correspondance des commandes BQS et GE :  
<http://cc.in2p3.fr/docenligne/972>
- Sources :  
<http://sourceforge.net/projects/gridscheduler/>