

Point de vue (personnel) d'un utilisateur ATLAS de la grille

Olivier Arnaez
Laboratoire d'Annecy-le-vieux de Physique des Particules
CNRS/IN2P3

Réunion LCG-France
19/05/2009

Sommaire

- Les différents types de « jobs »
- Cas d'utilisation de la grille
- Outils utilisés
- Choix de backend/site
- Difficultés rencontrées

Les différents « jobs »

- Analyse de données (cosmiques, MC)
- De façon routinière, validation du software ATLAS
- Utilisation de code produisant des données intermédiaires ayant vocation à être partagées à travers la communauté ATLAS

Cas d'utilisation

Sur grille

- Input = datasets officiels
 - Traitement de données (événements cosmiques seulement pour l'instant...)
 - Traitement de “productions MC officielles”
- Traitement = job Athena (framework ATLAS)
- Output = dataset d'histogrammes/NTuples

Cas d'utilisation

Hors grille

- Généralement :
 - Récupération en local (dq2-get → stockage LAPP non-grille) des histogrammes/NTuples précédents
 - Traitement = code C++/ROOT
 - Output = histogrammes
- Mais aussi parfois :
 - Traitement de productions officieuses (batch, castor/dcache)

Outils utilisés

- pAthena : outil de soumission de jobs PANDA
- GANGA : outil de soumission/monitoring de jobs PANDA/LCG/NdG/LSF
- PANDA Monitor
- Commandes dq2
- Commandes glite, lcg-*

Choix de « backend »

- Méthode empirique basée sur l'expérience...
- Utilisation de GANGA avec LCG principalement
- ...mais les jobs à lancer régulièrement (validation du soft) « passent » plus rapidement à BNL qu'au CERN —> PANDA pour ceux-la
- Lorsque j'ai défini un job avec GANGA ou pAthena, changement rare d'outil

Choix de « sites »

- Évolution de la politique de réplication des données avec la liste des formats de données dans ATLAS... —→ difficile à suivre !
- Jobs utilisant les données brutes —→ CERN
- Autres jobs, je laisse GANGA choisir le site dans le nuage désiré
- Généralement j'essaye FR en premier, CA en second puis FZK si nécessaire

Difficultés

Fréquence



- Problèmes de code (seg. fault, ...)
- Problèmes de « conditions », accès aux bases de données, principalement sur les T2/T3 mais aussi sur les T1
- Problèmes d'accès aux données (catalogue, XROOTD,...)
- Problèmes de RB (job soumis mais aucun site pour le tourner...)

Difficultés (2)

- Si le problème ne vient pas de moi
 - Demande d'aide au « support applicatif » du LAPP (Sabine, Stephane, Eric, Gael)
 - demande d'aide sur le e-group ATLAS (60 messages en 2009...)
- Si le problème persiste :
 - changement de nuage
 - dq2-get et job local si la taille du dataset le permet
 - batch si dataset important
 - Sur demande des experts de l'e-group, bug report (trop compliqué à mon goût...)

Conclusion

- Utilisation intensive de la grille depuis des mois
- Sensation que la qualité de service augmente
- Mais :
 - Problèmes d'accès aux bases de données très fréquents et gênants (ATLAS seulement ?)
 - Support (choix du contact, interface) parfois complexe