

mardi 25 novembre 2008

Services centraux de stockage grille au CC

Optimisations et évolutions

David Bouvet / Jonathan Schaeffer / Lionel Schwarz

dapnia

cea

saclay

- Service de transfert: FTS
- Service de stockage: dCache
- Service de catalogue: LFC

- Version 2.0 sur SL3.0.5
- Hardware: 2 noeuds V20Z (bi-pro 2GB RAM)
- DB back-end: cluster Oracle10g du CC
- 74 canaux définis en mode '*urlcopy*'
- ~4000 jobs soumis/jour

- Répartition des agents de canaux sur les 2 noeuds
 - Réalisable sans interruption de service en qq minutes
- Problématique du site GRIF (5 SE)
 - Mise en place d'1 canal par SE
- Détection des pannes et redémarrage automatique
- Service stable

- FTS 2.1 sur SLC4 prochainement en tests
 - twiki.cern.ch/twiki/bin/view/LCG/FtsChangesFrom20To21
 - Meilleure classification des erreurs
 - Source space token
 - Optimisation des performances dans la préparation du transfert

- Améliorer le niveau de fiabilité
 - Redondance des webservices
 - Machine de rechange
- Capture/remontée des erreurs
- Configuration partagée entre les noeuds
- Ajout de noeuds supplémentaires (meilleure répartition des canaux et load balancing Tomcat)

- Version: dcache-server-1.9.0-4 (13 nov)
- Hardware: 3 core servers en SL4/64, ~85 serveurs de disque X4500 (Thumpers) en Solaris10
- Volumétrie: 20M fichiers et répertoires, 1.2Po disque
- Débit moyen: 150Mo/s en entrée, 500Mo/s en sortie

- Software: Java6, postgresql-8.3
- Archivage de la base de comptabilité dCache
- fastPNFS (13 nov)
- Aggrégation des interfaces réseau: 2*1Gbs

- Redondance des portes d'cap
 - Répartition de charge et redondance par 'balance'
- Mise en place du service minimum import T0
 - Isolation des serveurs de transfert
- Détection des pannes et redémarrage automatique
- Double-alimentation des core-servers

- Exploitation du service
 - Amélioration des détections/diagnostic d'erreurs
 - Amélioration des procédures et outils d'installation
 - Automatisation des tâches routinières
- Fiabilité/disponibilité
 - Machine de rechange
 - Etude serveur srm/VO
- Transferts
 - GFTP2 (tests OK)

■ Métadonnées

- PNFS -> Chimera (1 base /VO ?)
- Réplication des bases PostgreSQL/amélioration du processus de sauvegarde
- Log des suppressions de fichiers

■ Interfaçage HPSS

- Amélioration du prestaging massif

- LFC local ATLAS (`lfc-prod.in2p3.fr`)
 - Version 1.6.11-3, SL4/64bits
 - Hardware : 2 noeuds V20Z
 - DB back-end : cluster Oracle10g du CC
- Replica read-only du LFC central LHCb (`lfc-lhcb-ro`)
 - Version 1.6.11-3, SL4/64bits
 - Hardware : 3550
 - DB back-end : cluster Oracle10g du CC -> stream Oracle (via LCG 3D) pour la réplication de la base du CERN

■ Fiabilité

- Service très stable
- DNS load balancing sur `lfc-prod.in2p3.fr`

■ Améliorations/évolutions

- Détections des pannes avec corrections automatiques ou remontée d'alarmes (rajout de sondes)