



Gestion de grands volumes de données géographiquement distribués : DPM, XRootD et iRODS

Jérôme PANSANEL

Storage Day 2018

Contexte

- IPHC : un institut pluridisciplinaire composé de quatre départements
- Impliqué dans les expériences CMS et ALICE (LHC)
- Centre de calcul créé fin 2006
- Partenaire de WLCG et LCG-France en 2007
- Tier-2 CMS et ALICE depuis juin 2008
- Ouverture aux autres disciplines scientifiques en 2010 (Biomed, Renabi, Egeode, ILC, ...)
- Partenaire de France Grilles depuis 2010
- Déploiement de la solution iRODS en 2011
- Début du Cloud en 2013
- Implication dans l'IFB (BISTRO)
- Fin 2017 : lancement de la plate-forme labellisée SCIGNE

Le calcul et les données à l'IPHC

- 150 serveurs de calcul (3000 cœurs)
- 30 serveurs de stockage (2,5 Po de stockage utile)
- Backbone à 80 Gb/s
- Accessible en 10 Gb/s
- Robotique de sauvegarde
- Une équipe de 7 personnes (3,1 ETP)
- Plusieurs technologies en production :
 - Grille : Middleware UMD et DIRAC
 - Cloud : OpenStack (+ conteneurs)
 - Stockage : DPM, XRootD, iRODS et CEPH

DPM : contexte

- Expériences du LHC organisent les données en tiers
- CERN → Tier 0, 14 T1 (12 pays) et ~ 70 T2
- 1 type de catalogue / expérience
- Distribution des données avec l'outil File Transfer Service (FTS)
→ <http://fts3-service.web.cern.ch/>
- Support des protocoles GridFTP et SRM (HTTPS et WebDav plus récemment)
- Authentification forte basée sur les certificats x509
- Solutions communément déployées :
 - dCache - <https://www.dcache.org/>
 - DPM - <http://lcgdm.web.cern.ch/dpm>



WLCG
Worldwide LHC Computing Grid

DPM

- Logiciel libre
- Gestion de plusieurs péta-octets de données
- Serveurs de stockage derrière un frontal (*head node*)
- Gestion d'un espace de nommage hiérarchique (arborescence)
- Catalogue DPNS comprenant les fichiers et les répertoires
- Catalogue DPM pour les requêtes et les interactions avec les serveurs sur chaque site pour accéder aux serveurs
- Accès CLI, API C et Python
- Catalogue logique externe LFC pour accéder aux différents sites hébergeant un service DPM

Avantages

- Déploiement facile avec Puppet (également Quattor)
- Authentification basée sur les certificats x509 (GSI)
- Chiffrement des flux
- Excellent passage à l'échelle
- Stabilité de la solution
- Réplication, réservation d'espace, ...
- Possibilité d'être un frontal S3

Inconvénients

- Pas d'interface graphique (hormis https et WebDav ...)
- Authentification basée sur les certificats x509 (GSI)
- Gestion des petits fichiers (~ Ko)
- Option de réplication limitée
- Pas de mécanisme de *fail-over*
- Pas de *tiering*

XrootD : contexte

- Solution développée à l'origine pour le framework d'analyse ROOT et utilisé par l'expérience ALICE
- Logiciel libre
- Accès rapide et avec une grande bande passante aux données
- Système de redirection pour assurer la haute disponibilité des données
- Authentification GSI, Kerberos et identifiant / mot de passe

→ <http://xrootd.org/>



XRootD

Avantages

- Déploiement facile avec Puppet (également Quattor)
- Plugin pour DPM
- Excellente efficacité des transferts
- Redirection (haute-disponibilité des données)

Inconvénients

- Pas d'interface graphique (hormis https et WebDav ...)
- Configuration complexe
- Documentation succincte
- Pas de méta-données
- Pas de *tiering*

IRODS : contexte

- 2011 : début du Cloud dans les infrastructures distribuées
- Un système utilisable depuis le Cloud et la grille
- Facile à déployer (au moins la partie cliente)
- Authentification mutuelle (identifiant / mot de passe et GSI)
- Système de fichiers accessible depuis n'importe quelle machine cliente (ordinateurs, smartphones et tablettes)
- Possibilité de réplication
- Gestion des méta-données
- Déploiement et maintenance aisés du service

iRODS

- Logiciel libre sous licence BSD
- Virtualisation de différents type de stockage dans un seul espace de nom (zone)
- Passage à l'échelle
- Méta-données
- Automatisation de la gestion des données grâce au moteur de règles (vérification des types, extraction des méta-données, ...)
- Transfert parallèle pour les données volumineuses
- Sécurisation des données grâce à la réplication et la gestion des accès

→ <https://irods.org/>

The logo for iRODS, featuring a lowercase 'i' in teal and the letters 'RODS' in a bold, grey, sans-serif font.

Avantages

- Solution facile à déployer
- Service utilisé par plusieurs communautés scientifiques différentes
- Support commercial
- Interface Web
- Montage Fuse
- Flexibilité
- Groupe d'utilisateurs dynamique
- Méta-données

Inconvénients

- Configuration complexe (règle et micro-service)
- Modifications importantes entre les versions

Prospectives

CEPH – <http://ceph.org>

- Plusieurs services (DPM, XRootD, iRODS)
- Possibilité d'utiliser le même *backend*
- Qualification en cours d'une solution de 500 TB
- Projet de réplication d'une partie du stockage au LAL (reprise d'activité OpenStack)

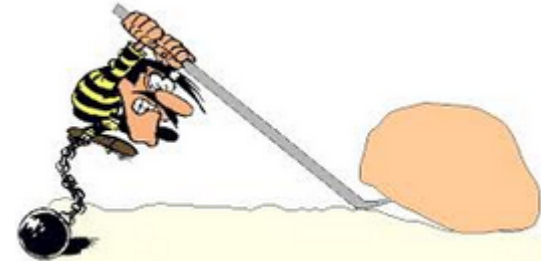
OpenIO – <http://openio.io>

- Solution pour le stockage géographiquement distribué
- Efficace avec les petits fichiers
- Réplication, facilité de placement des données près du calcul, ...
- Test en cours entre l'IPHC et l'IHES / école polytechnique

Freins et leviers



- Multiplicité des systèmes
- Configuration complexe avec un impact potentiel sur les performances
- Différentes interfaces pour les utilisateurs
- Différents systèmes d'authentification / autorisation



- Solutions reconnues pour leur possibilité de passage à l'échelle
- Solutions stables et bien maintenues
- Fonctionnement avec du matériel standard / pas de blocage vendeur
- Interopérabilité avec les partenaires français et européens
- Flexibilité des solutions proposées aux utilisateurs



Questions ?