



Jeudi 21 juin 2012

WLCG Technology Evolution Groups (TEG) Data and Storage Management

Yvan Calas

yvan.calas@cc.in2p3.fr

dapnia
cea
saclay

 CNRS
CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE



TEG Data and Storage Management



- Création en novembre 2011 de plusieurs groupes de travail sur les thèmes suivants:
 - *Data management* (30 membres).
 - *Storage Management* (~40 membres).
 - *Workload Management*.
 - *Databases*.
 - *Security*.
- Mission: « Réexaminer l'implémentation des infrastructures utilisées à la lumière de l'expérience acquise avec les données du LHC, et prendre en considération l'évolution des technologies. »

TEG Data and Storage Management



- Fusion des groupes « *Data Management* » et « *Storage Management* »
 - thèmes abordés très liés.
- ~7 meetings téléphoniques (EVO) + workshop « *face to face* » Amsterdam fin janvier 2012.
- Recommandations du TEG rendues le 13 avril 2012 (document de 40 pages...):

<https://twiki.cern.ch/twiki/pub/LCG/ReportDataStorageTEG/DataandStorageTEG-SummaryofRecommendationsv4.0.pdf>

■ Actuellement 2 couches:

– *Archive layer*:

- copie sur tapes des fichiers à stocker de manière permanente.
- coût faible, plus grande fiabilité, latence élevée.
- remplacée par une solution *disk-only* sur le moyen ou long terme?
 - diminution latence, augmentation du coût.
 - réplication d'un même fichier sur plusieurs sites.



Architecture WLCG pour la gestion des données



– *Placement (access) layer:*

- couche responsable de l'accès aux datasets par les clients.
- utilisation de FTS pour le transfert des données au sein de cette couche (transferts inter-sites).
- utilisation de SRM pour ramener les fichiers sur disque au niveau d'un site (*srm-bring-online*).

Architecture WLCG pour la gestion des données



- Introduction d'une nouvelle couche: *Federation Layer*.
 - Fédération de cache disque entre les différents sites:
 - But: augmenter la disponibilité des données au delà d'un unique site (utile en cas d'indisponibilité temporaire ou persistante d'un site par exemple).
 - Principe: si des données (un fichier) demandées par une application ne peuvent pas être trouvées à l'endroit où tourne l'application, le protocole redirige la requête à la couche Federation pour trouver un autre site où se trouve ce fichier (*redirection capabilities* du protocole d'accès client).
 - Plusieurs implications:
 - un protocole d'accès commun entre les sites (*federation protocol*).
 - un espace de nommage global.
 - Les données fédérées sont en lecture seule.
 - Solution testée par Alice, Atlas et CMS: protocole xrootd (mécanisme de redirection déjà implémenté).
 - Autre solution (non testées): HTTP/WebDAV.



Recommandations



- Développement d'un plugin HTTP pour Xrootd.
- Mise en place d'un outil de surveillance de la bande passante utilisée dans les mécanismes de fédération pour étudier/évaluer les accès/transferts opportunistes.
- Création de plusieurs groupes de travail, eg.:
 - Publication des nouvelles données dans la *read-only placement layer*.
 - Etudier une plus stricte séparation entre les données *read-only* et *read-write*.



Protocoles d'accès point-à-point



- Actuellement aucun protocole standard pouvant remplacer GridFTP mais amélioration de l'efficacité par:
 - changement de version.
 - réutilisation de connexions GFTP déjà ouvertes.
- Vérifier que les systèmes de stockage supportent bien Xrootd, notamment en terme de latence et de stabilité sous forte charge.
- Tester HTTP intensivement (protocole bien adapté pour la *Federation layer*)



Recommandations pour SRM



- SRM doit être conservé dans le court/moyen terme.
 - toujours indispensable sur les sites archivant les données (*custodial data*). Pas de solution de remplacement pour le moment: SRM doit restée l'interface par défaut.
- Cependant:
 - Seulement une fraction des fonctionnalités est utilisée.
 - SRM utilisé uniquement dans la communauté HEP.
 - Recherche de solutions alternatives supportant FTS et lcg_utils pour permettre l'interopérabilité entre sites.
 - protocoles propriétaires: Amazon S3, DropBox.
 - standards existants: CDMI, WebDAV.
 - Groupe de travail à créer pour évaluation de ces nouvelles solutions.



Conclusions « Post-TEG »



- Lors du GDB du 13 mai 2012, Ian Bird a proposé de créer plusieurs groupes chargés de concentrer leurs efforts sur:
 - La fédération.
 - Le *benchmarking* (étude des outils d'analyse de performance existants pour le stockage).
 - Le remplacement (?) de SRM (*Storage Interface Group*).