



Nouveaux services MUST

Service de gestion de données iRODS

M. GOUGEROT, C. BARBIER

12 décembre 2017

Plan

- Rappel du besoin
- iRODS
 - Présentation générale
 - Architecture
 - Fonctionnalités
- iRODS sur MUST
 - Historique
 - Exemples d'utilisation

Rappel du besoin

Besoin exprimé par le laboratoire EDYTEM :

- Augmentation importante de la volumétrie des données scientifiques issues de leurs instruments/capteurs
- Gestion de métadonnées
- Automatisation des transferts
- Performance réseau
- Interface de transfert/consultation graphique conviviale (Windows-like)

iRODS : Présentation générale

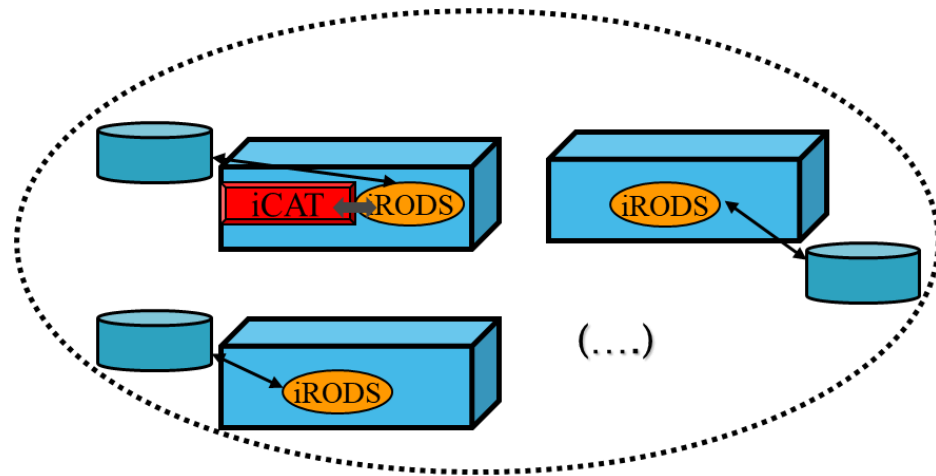
- iRODS = “integrated Rule Oriented Data Systems”
→ Système de virtualisation de stockage sur une architecture distribuée (hétérogène)
- Origine : 2006, RENC (institut de l’“University of North Carolina”)
- Consortium pour le développement et la gouvernance
- Réseau de partenaires pour le support commercial
- Open source (CC-IN2P3 contributeur)
- <https://irods.org>

iRODS : Architecture

Des serveurs :

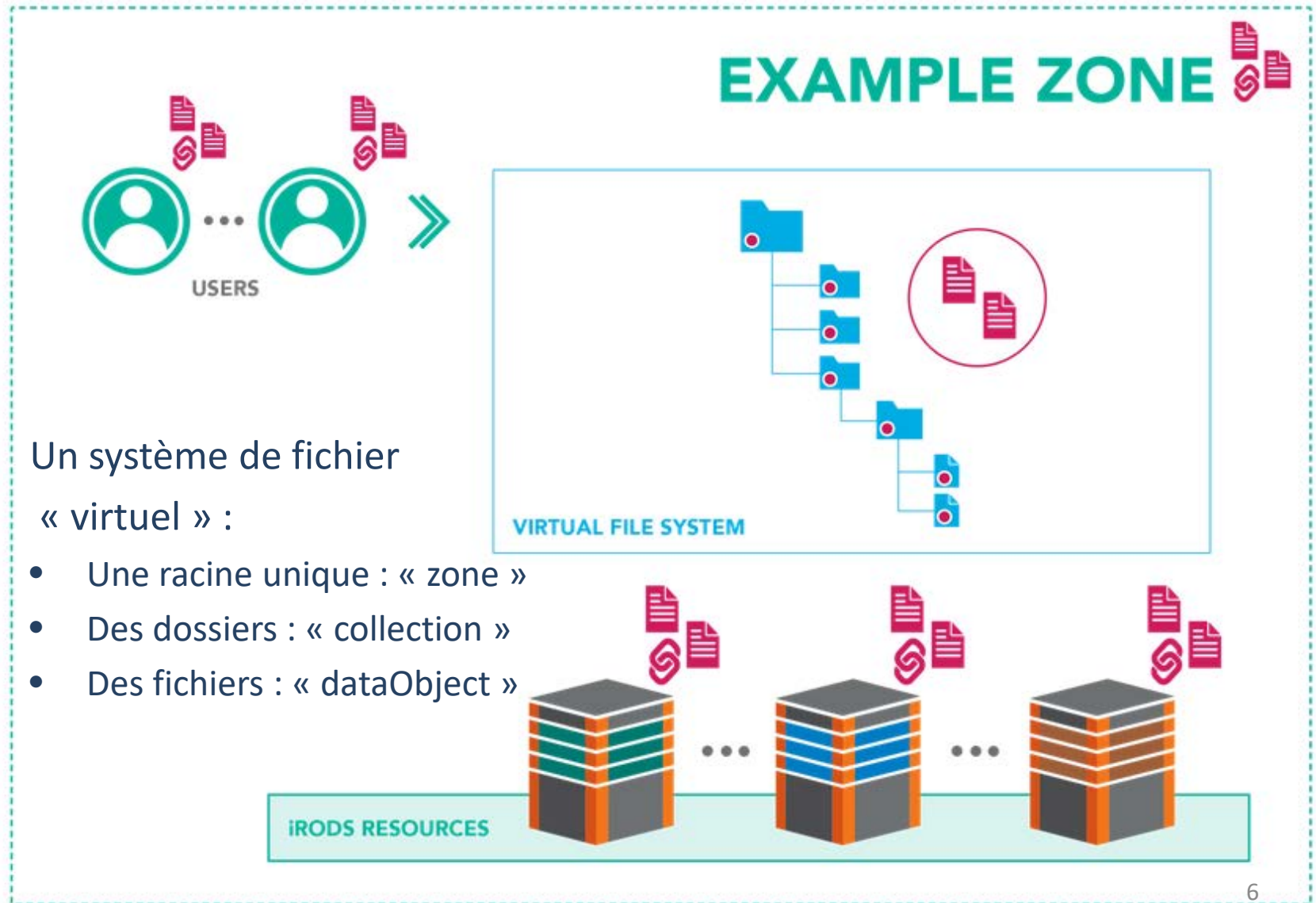
– Un service de catalogue (« provider ») pour stocker les infos (= base de données):

- Métadonnées
- Utilisateurs
- Ressources



– Des serveurs de ressources distribués (« consumer »)

iRODS : Architecture



iRODS : Architecture

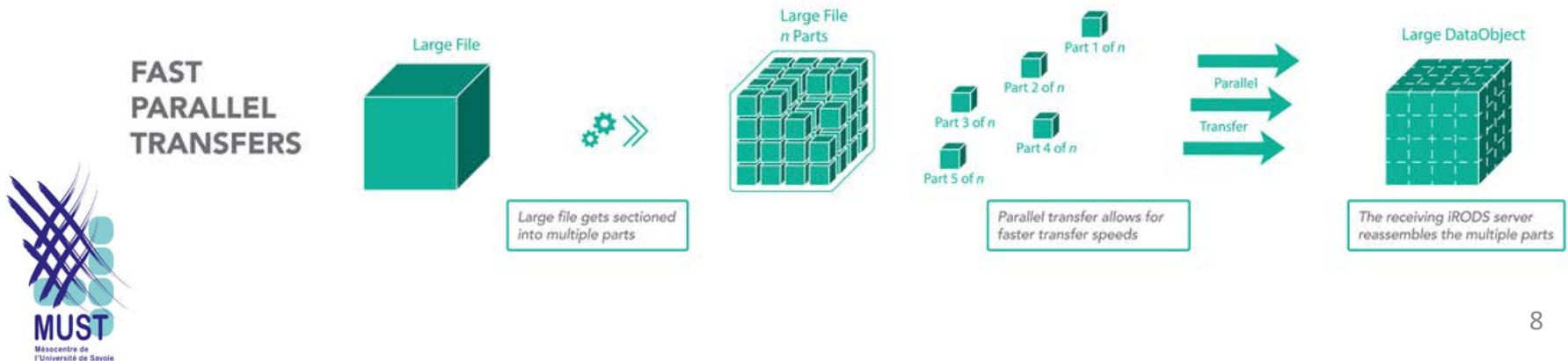
Des clients :

- Commandes client de base style Unix : iCommands
 - **iinit** : pour initialiser votre environnement et saisir votre mot de passe de session
 - **ils** : pour voir la liste des objets et collections
 - **iput** : pour déposer un objet (avec l'option "-P" pour suivre l'avancement)
 - **iget** : pour récupérer un objet
 - **icd** : pour changer de dossier (collection)
 - **ipwd** : pour voir le dossier en cours
- Possibilité de portail web, d'API REST
- Clients graphiques pour Windows, Mac, ...

iRODS : Fonctionnalités de base

Les fonctions de base :

- Un compte par utilisateur, notion de groupe
- Droits d'accès
- Quotas (par utilisateur, par groupe)
- Performance et sécurité :
 - Réplication
 - Distribution
 - Transferts parallélisés



iRODS : Fonctionnalités avancées

Les règles : actions automatisées

- Ecrites dans un langage libre si plugin disponible , exemple python
- Les règles s'appuient sur des microservices intégrés au serveur iRODS
- Usages :
 - Placement de fichiers (tiering, réplication)
 - Calcul et enregistrement de métadonnées
 - Conversions
 - Contrôle d'accès, audit

iRODS : Fonctionnalités avancées

Les métadonnées:

– triplet AVU = Attribut Valeur Unité

Exemples :

- "taille" "160" "cm"
- "poids" "60" "kg"

– Syntaxe permettant de faire des requêtes

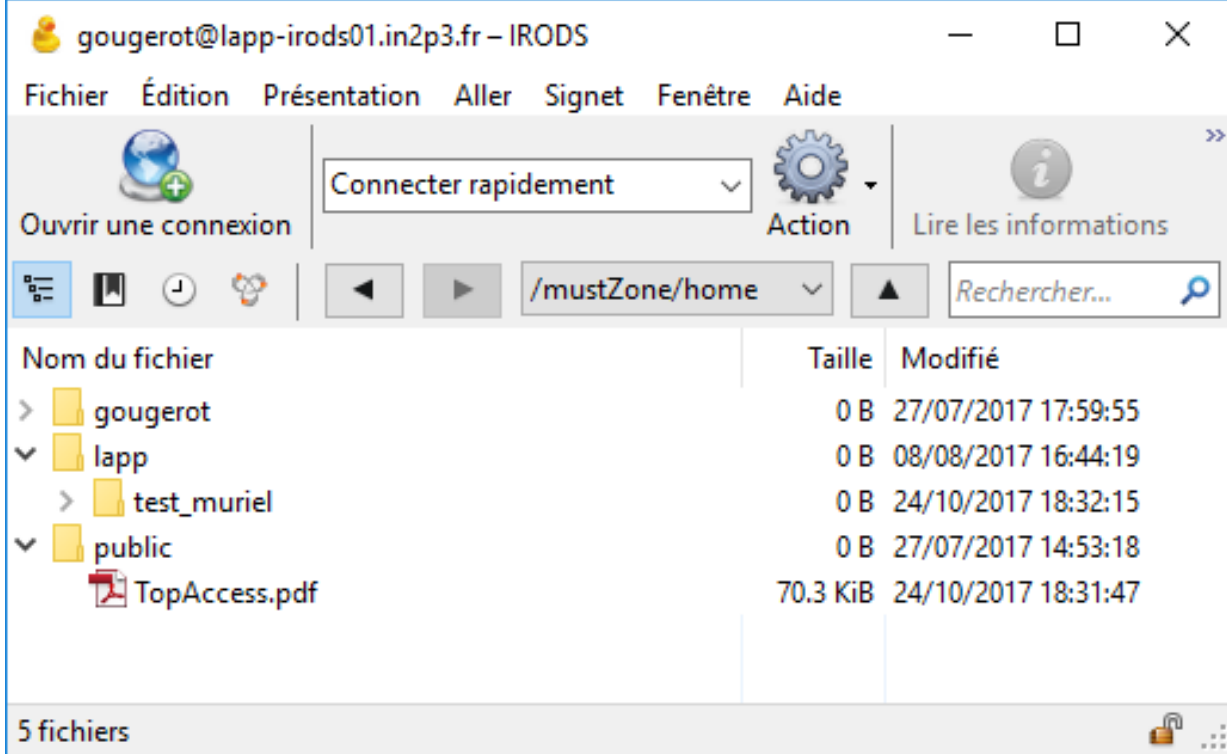
iRODS sur MUST

iRods sur MUST : Historique

- Mi 2015 : discussions avec EDYTEM sur solutions de stockages possibles dont iRODS
- Début 2016 : expression du besoin et demande de tests
- Mi 2016 : tests par EDYTEM sur l'instance iRODS du CC-IN2P3 et confirmation du besoin
- Fin 2016 : approbation en comité de pilotage MUST pour nouveau service iRODS
- Mi 2017 : création d'une maquette pour MUST
- Octobre 2017 :
 - Mise en production du service iRODS MUST avec 1To de stockage pour EDYTEM (6 utilisateurs)
 - Documentation utilisateur sur le wiki de MUST :
<http://lappwiki01.in2p3.fr/Support-Applicatif/index.php/IRODS>

iRODS sur MUST : Démonstration

```
#
#ils /mustZone
/mustZone:
  C- /mustZone/AUTRE-LABO
  C- /mustZone/EDYTEM
  C- /mustZone/home
  C- /mustZone/LAPP
  C- /mustZone/trash
#
#icd /mustZone/LAPP
#
#ipwd
/mustZone/LAPP
#
#ils -l /mustZone/home/public
/mustZone/home/public:
  gougerot                0 must1Resc          72018 2017-10-24.18:31 & TopAccess.pdf
#
..
```



The screenshot shows the iRODS web interface for the user 'gougerot@lapp-irods01.in2p3.fr'. The interface includes a menu bar (Fichier, Édition, Présentation, Aller, Signet, Fenêtre, Aide) and a toolbar with options like 'Ouvrir une connexion' and 'Connecter rapidement'. The main area displays a file browser view of the directory '/mustZone/home'. The file list is as follows:

| Nom du fichier | Taille | Modifié |
|----------------|----------|---------------------|
| > gougerot | 0 B | 27/07/2017 17:59:55 |
| ▼ lapp | 0 B | 08/08/2017 16:44:19 |
| > test_muriel | 0 B | 24/10/2017 18:32:15 |
| ▼ public | 0 B | 27/07/2017 14:53:18 |
| TopAccess.pdf | 70.3 KiB | 24/10/2017 18:31:47 |

The status bar at the bottom indicates '5 fichiers'.

iRODS sur MUST : Démonstration

```
#
#imeta add -d individu_002 "taille" "160" "cm"
#
#imeta add -d individu_002 "poids" "60" "kg"
#
#imeta ls -d individu_002
AVUs defined for dataObj individu_002:
attribute: poids
value: 60
units: kg
----
attribute: taille
value: 160
units: cm
```

```
#imeta qu -d taille '>' 180
XXXXX cmdToken[cmdIx] =
collection: /mustZone/LAPP/Population
dataObj: individu_004
----
collection: /mustZone/LAPP/Population
dataObj: individu_005
#
```



```
▼ LAPP
  > barbier
  gros_fichier
  ▼ Population
    individu_001
    individu_002
    individu_003
    individu_004
    individu_005
  test7
  test8
```

```
#imeta qu -d poids '<=' 80
XXXXX cmdToken[cmdIx] =
collection: /mustZone/LAPP/Population
dataObj: individu_001
----
collection: /mustZone/LAPP/Population
dataObj: individu_002
----
collection: /mustZone/LAPP/Population
dataObj: individu_003
#
```

Conclusion

- Solution répondant bien au besoin exprimé
- Perspectives
 - Déploiement d'un portail web et/ou de clients web
 - Ouverture aux autres laboratoires/partenaires de MUST
- Pour toute demande (accès, groupe, quota ...), nous contacter via support-usmb@lapp.in2p3.fr

Questions ?

iRODS : quelques iCommands

- iinit : pour initialiser votre environnement et saisir votre mot de passe de session
- ils : pour voir la liste des objets et collections
- iput : pour déposer un objet (avec l'option "-P" pour suivre l'avancement)
- iget : pour récupérer un objet
- icd : pour changer de dossier (collection)
- ipwd : pour voir le dossier en cours

iRods sur MUST : Détails

- Version : 4.2.1
- iCommands disponibles sur le portail interactif lappuds
- Zone : mustZone
- Service : lapp-irods.in2p3.fr
 - 1 serveur pour le catalogue (base postgresql)
 - 1 seul serveur de ressource pour l'instant
 - Utilisation du stockage SAN (GPFS)