Le service « iRODS » de France-Grilles

Offre aux utilisateurs

Le groupe iRODS de France-Grilles

13 février 2014, Clermont-Fd

Le logiciel

Integrated Rule-Oriented Data System

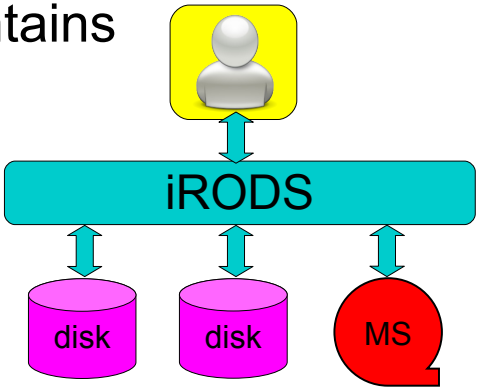
iRule Oriented Data Systems

- Projet né en 2006, successeur de *SRB*
- Open Source, licence BSD
- Développé par le groupe DICE et quelques collaborateurs (dont le CC-IN2P3)

Data Intensive Cyber Environments Group

Gestionnaire de données de type « grille »

- Gestion de collections de données distribuées sur des sites lointains
  - sites équipés d'infrastructures (disque, bande) hétérogènes
- Mécanisme uniforme pour accéder aux données



Plus en détails

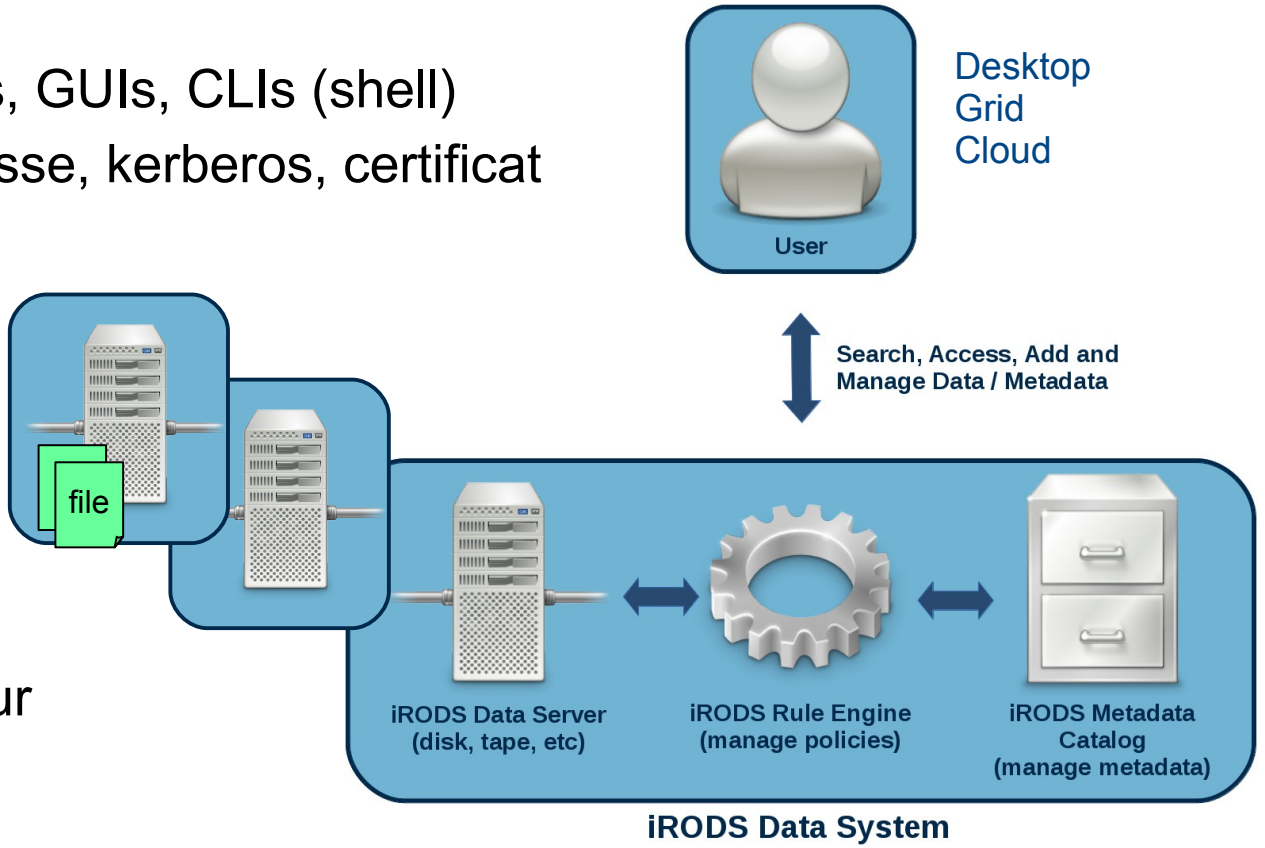
- Système de métadonnées pour manipuler les données
- Organisation logique des fichiers indépendante de leur organisation physique
- Cohérence et homogénéité dans la gestion des données renforcées
  - intègre la politique de gestion des données (décidée par leur propriétaire)

- iRODS ne remplace pas une solution locale pour des accès parallèles massifs

iRODS techniquement

– Trois éléments principaux définissent un système iRODS
  • Base de données + « moteur de règles » + ressource
  • Plusieurs ressources distantes peuvent co-exister

– Interfaces utilisateurs : APIs, GUIs, CLIs (shell)
– Authentification : mot de passe, kerberos, certificat



Gestion des données

– Flot des données prédéfini
et transparent pour l'utilisateur
– A la demande

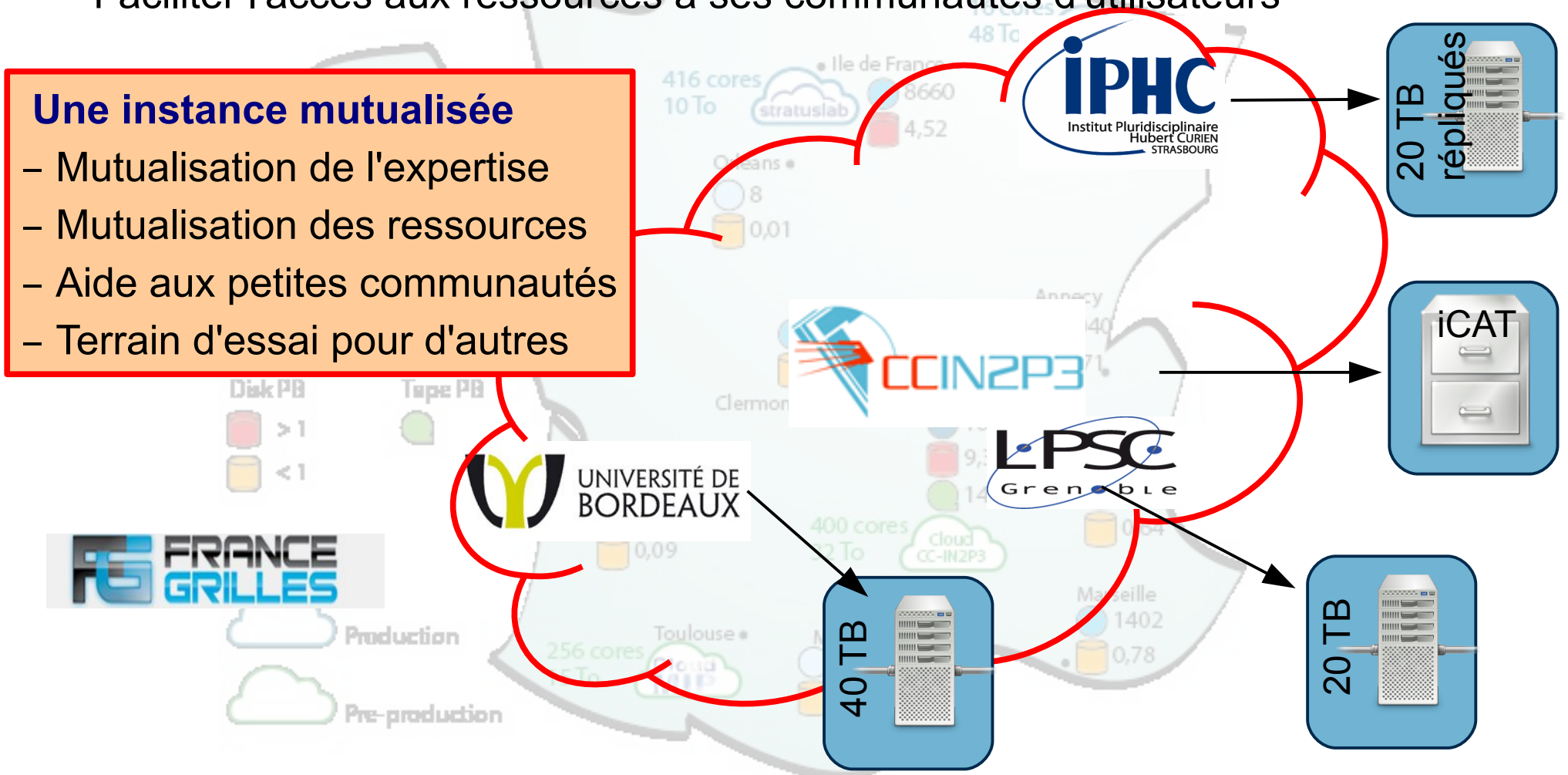
Une instance iRODS nationale

Instance iRODS France Grilles

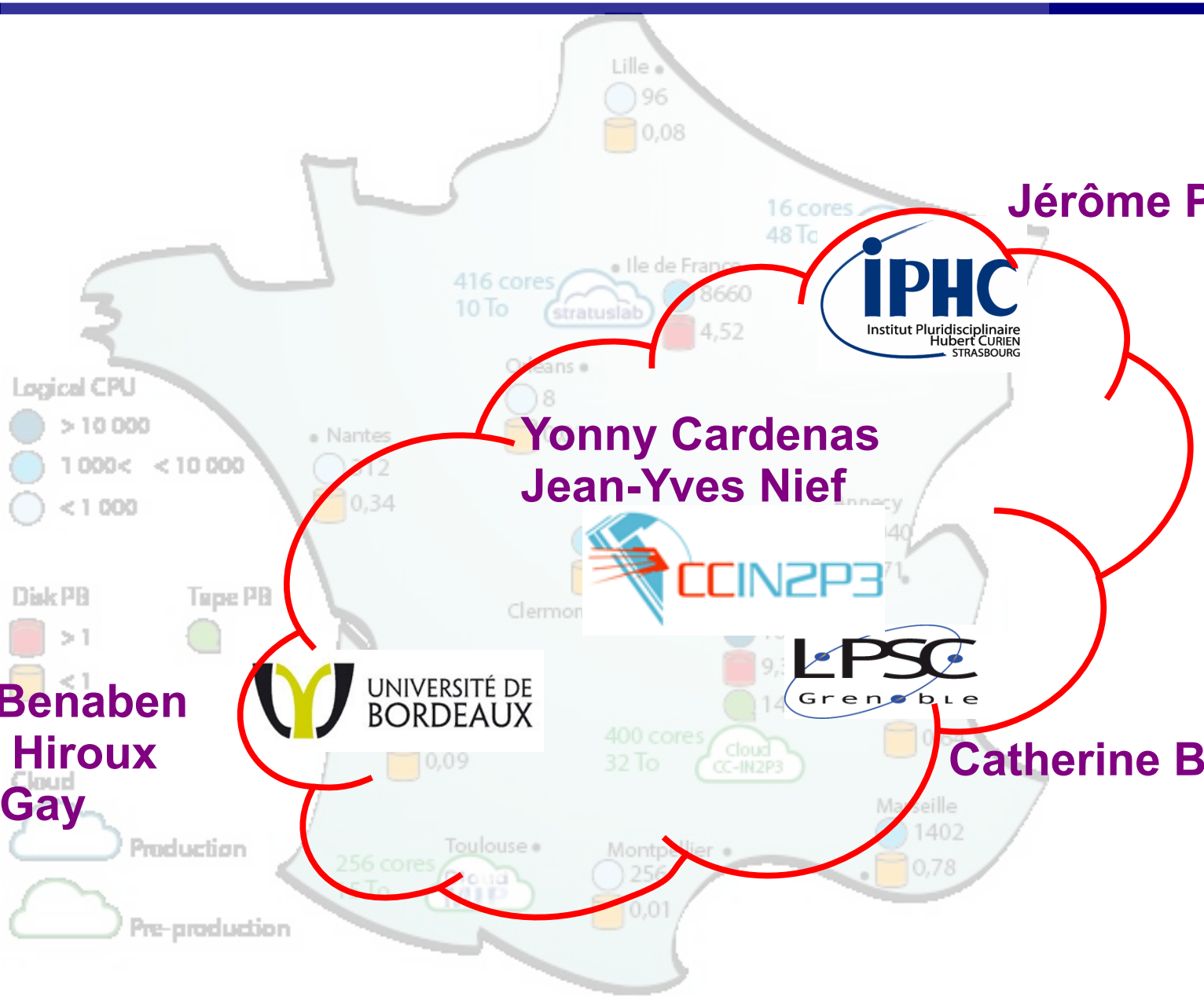
- Coordonnée par France Grilles
  - Faciliter l'accès aux ressources à ses communautés d'utilisateurs

Une instance mutualisée

- Mutualisation de l'expertise
- Mutualisation des ressources
- Aide aux petites communautés
- Terrain d'essai pour d'autres



L'équipe des administrateurs



Les utilisateurs

Profil

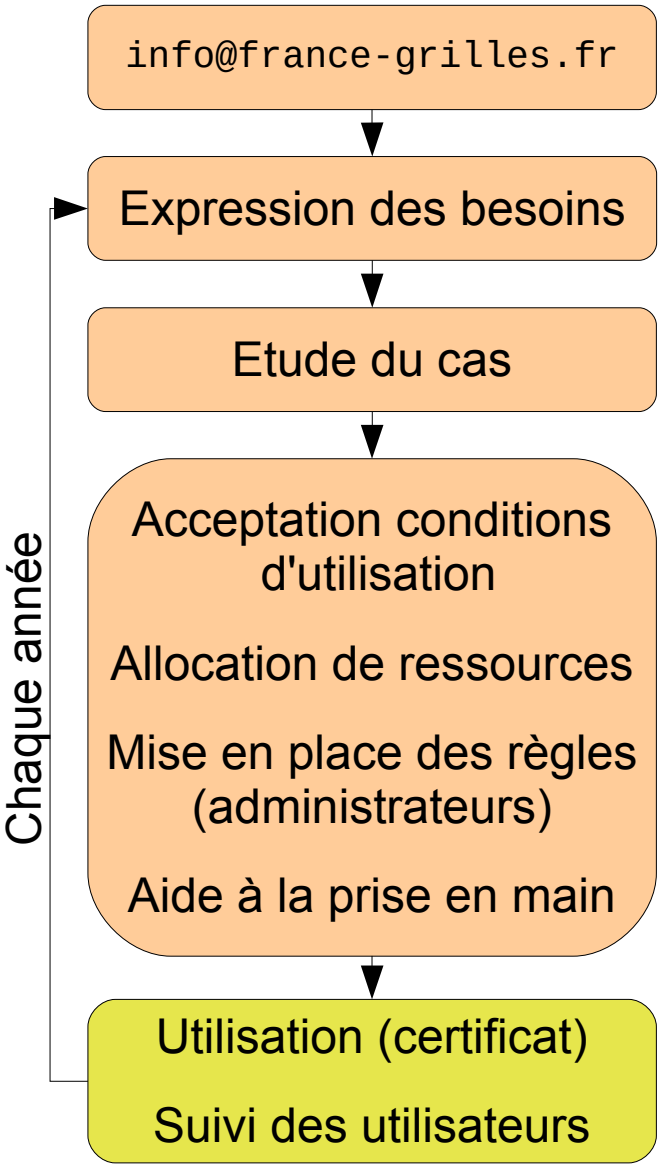
– Données scientifiques
– Tout domaine
– Utilisateur seul ou un groupe
  • L'accès est nominatif

Besoins

– Stockage et distribution de données
– Gestion de flots de données plus complexes avec des actions implicites (réplication, création de métadonnées, déclenchement de traitement, ...)
– Ou test du produit iRODS

Utilisation

– Copie, réplication, création de métadonnées
– Contrôle des accès
– Accès depuis les nœuds de calcul FR-Grilles



Gestion implicite

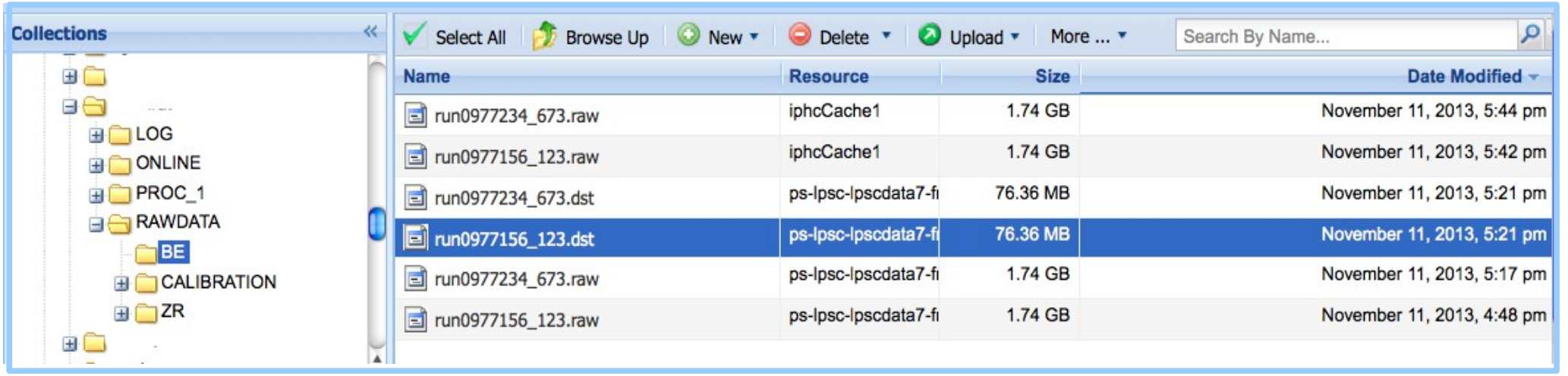
Vue administrateur

- Par la définition de règles (actions)
- Basées sur d'autres règles et les (~250) micro-services proposés par iRODS
- Interprétées automatiquement par le moteur de règles unique
  - Transparent pour les utilisateurs

```
AcPostProcForPut {   # executed with each new file loaded in iRODS
    on ($objPath like "/tempZone/home/cines-test/archivage/*.tar") {
        acAdonis1($objPath,"test");  # untar the file and register all the created files in iRODS
    }
    msiSysReplDataObj("SiteAResc1","null"); # replication on an other disk space called “SiteAResc1”
    msiSysReplDataObj("diskcache“,"null");  # replication on an other disk space called “diskcache”
    msiSysMetaModify("comment++++numRepl=0",“Original file");  # modify the file metadata
    called comment for replica number 0
}
```

Fichiers logiques

Vue utilisateur



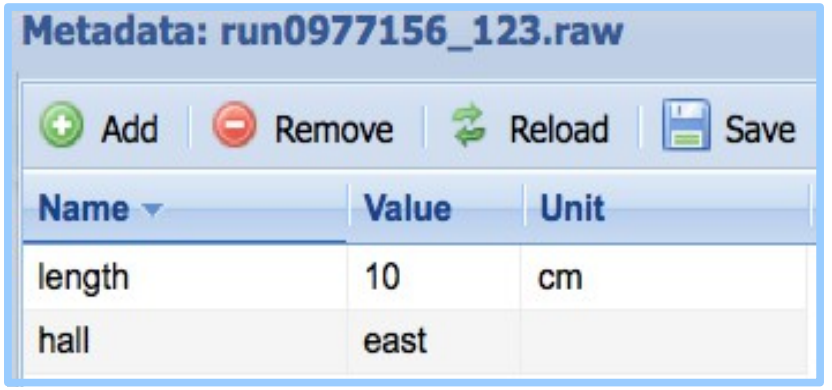

```
[user ~]$ ils
/frgrid/home/UNECOLLAB/RAWDATA:
 C- /frgrid/home/UNECOLLAB/RAWDATA/CALIBRATION
 C- /frgrid/home/UNECOLLAB/RAWDATA/BE
 C- /frgrid/home/UNECOLLAB/RAWDATA/ZR
[user ~]$ ils -l BE/
/frgrid/home/UNECOLLAB/RAWDATA/BE:
  owner    0 ps-lpsc-lpscdata7-fr   80072192 2013-11-11.16:21 & run0977156_123.dst
  owner    0 ps-lpsc-lpscdata7-fr 1748189011 2013-11-11.15:48 & run0977156_123.raw
  owner    1 iphcCache1           1748189011 2013-11-11.16:42 & run0977156_123.raw
  owner    0 ps-lpsc-lpscdata7-fr   80072192 2013-11-11.16:21 & run0977234_673.dst
...
```

Métadonnées

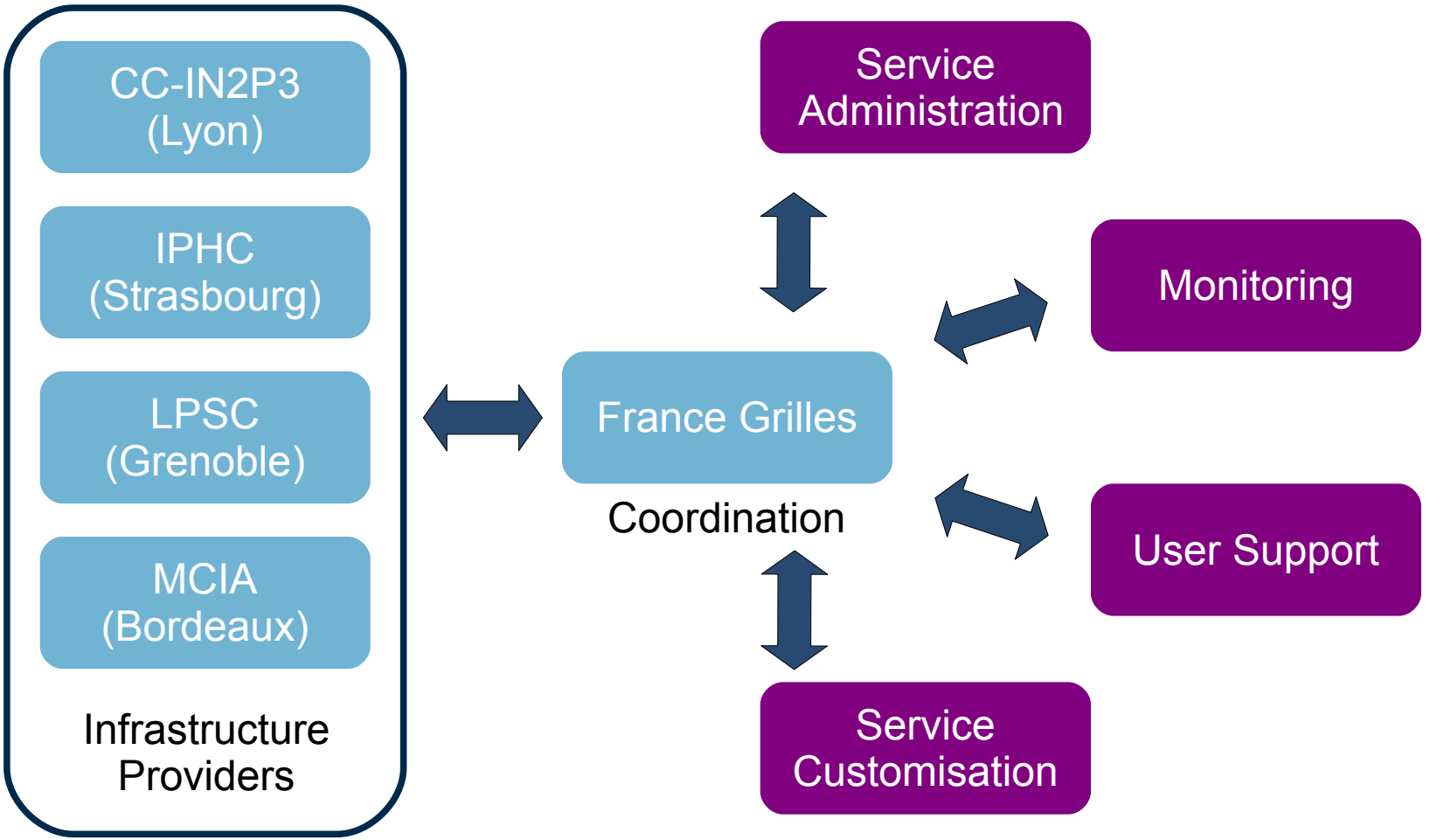
Vue utilisateur

- Attachées à : une collection, un fichier, un utilisateur
- Triplet : nom attribut unité

```
[user ~]$ imeta add -d run0977156_123.raw length 10 cm
[user ~]$ imeta add -d run0977156_123.raw hall east
[user ~]$ imeta ls -d run0977156_123.raw
AVUs defined for dataObj run0977156_123.raw:
attribute: length
value: 10
units: cm
----
attribute: hall
value: east
units:
[user ~]$ imeta -d qu hall east
collection: /frgrid/home/UNECOLLAB/RAWDATA/ZR
dataObj: run0977156_123.raw
----
collection: /frgrid/home/UNECOLLAB/RAWDATA/ZR
dataObj: run0817773_556.raw
```



Les services



Les services

Service Administration

Monitoring

L'équipe iRODS assure :

L'administration du service

- Serveurs de données et iCAT
- Préparation d'outils spécifiques et leur diffusion (RPM, VM, etc)
  - Documentation utilisateurs de France-Grilles
    - https://forge.in2p3.fr/projects/france-grilles-documentation/wiki/Utiliser_le_service_iRODS

La surveillance du système

- Qualité du service rendu
- Service 24/24 et 7/7
  - sauf arrêts planifiés annoncés en avance : irods-user-l@FRANCE-GRILLES.FR

Les services

User Support

Service Customisation

Accueil des utilisateurs

– Expression des besoins : info@france-grilles.fr

– Aide à la prise en mains du service

- Tutoriels planifiés : http://www.france-grilles.fr/-Formation-
- Documentation pour les utilisateurs (dans le wiki France-Grilles)

  https://forge.in2p3.fr/projects/france-grilles-documentation/wiki/Utiliser_le_service_iRODS
- Aide personnalisée pour adapter le service aux besoins
  - Mise en place de règles pour la gestion implicite des données

Suivi continu des utilisateurs

– Point de contact pour des conseils/incidents : irods-support-l@FRANCE-GRILLES.FR

– Forum de discussion et mail : https://forge.in2p3.fr/projects/france-grilles-documentation/boards

irods-user-l@FRANCE-GRILLES.FR

Revue de l'usage du service chaque automne

– Processus d'amélioration continue du service

– Vision globale des volumétries accordées

Pré-requis et conditions d'accès

Prérequis :

(Après votre expression d'intérêt et l'acceptation du projet)

- Avoir une certificat électronique d'une autorité académique
- Appartenance à la VO vo.france-grilles.fr (acceptations des conditions)
- Signature des conditions d'accès du service iRODS (« Acceptable Use Policy »)

Extrait des conditions d'accès au service :

- L'accès doit être demandé explicitement et individuellement.
- Pour un groupe travaillant sur un projet commun, une personne « contact » sera demandée.
- Les projets acceptés le sont jusqu'à l'automne qui suit.
- Pour tout accès autorisé, l'utilisateur sera inscrit à la liste des utilisateurs.
- Les données dans iRODS doivent être à caractère scientifique et liées au projet initial.
- France-Grilles ne peut être tenu responsable de la perte ou du vol des données.
- A la fin de la période agréée d'utilisation du service, l'utilisateur doit rapatrier ses données si elles lui sont utiles.

Avancement et futur du projet

- ICAT et ressources en production (80 To nets disponibles)
- Clients iRODS déployés sur les sites FR-Grilles
- Accueil : procédure et conditions d'accès au service définies
- Canaux de communication établis
- Documentation et forum utilisateurs prêts

A ce jour

L'infrastructure et la partie « utilisateurs » sont en place pour accueillir les premiers utilisateurs.

- Gain en stabilité de service ↔ retour d'utilisateurs
- Mise en place d'un monitoring centralisé
- Mise en place du support interne

Cet automne

Passage en production (service stable et testé) en septembre 2014.

- Créer une synergie en France avec un partage d'expertise
  - Supports (forum, wiki, mailing liste)
- Pérennisation de l'infrastructure de France-Grilles

Pour conclure

Gestionnaire de données mutualisée

- Mutualiser l'administration
- Coordonner le support et la formation
- Soutien des petites communautés

Infrastructure en place et prête à être utiliser

- Aujourd'hui :
  - comment les utilisateurs peuvent utiliser à l'infrastructure
  - la gestion implicite des données ne sera pas traitée (étude au cas-par-cas)
- En fin de journée :
  - possibilité pour vous de devenir « early adopter », c'est-à-dire d'utiliser iRODS dès maintenant (sans forte garantie de stabilité)