



mardi 3 avril 2007

# SRB et iRODS au CC-IN2P3

Etat des lieux et perspectives

dapnia  
cead  
saclay

CNRS  
CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE



## Vue d'ensemble de la présentation



- Introduction à SRB:
  - Fonctionnalités.
  - Utilisation de SRB dans le monde.
- Utilisation de SRB au CC-IN2P3:
  - Exemples d'expériences/projets utilisant SRB.
  - Développements spécifiques.
- Bilan et perspectives d'utilisation de SRB.
- Introduction à iRODS.
- Participation du CC-IN2P3 à iRODS.



# Introduction à SRB



- Collaborations internationales:
  - Données (publiques, privées) distribuées sur de nombreux sites.
  - Stockage hétérogène (fichiers/bases de données/flux de données, URL, cartouches/disques).
  - OS hétérogènes.
  - Authentification / droits d'accès aux fichiers dans l'organisation virtuelle.
  - Authentification / droits d'accès à travers différentes organisations virtuelles.
- Virtualisation du stockage nécessaire:
  - Vue hiérarchique (organisation logique) de l'ensemble des données.
  - Applicatif qui gère les données indépendant:
    1. des évolutions technologiques des systèmes de stockage.
    2. des modifications de l'organisation locale des fichiers (ex: serveurs, points de montage, etc...).

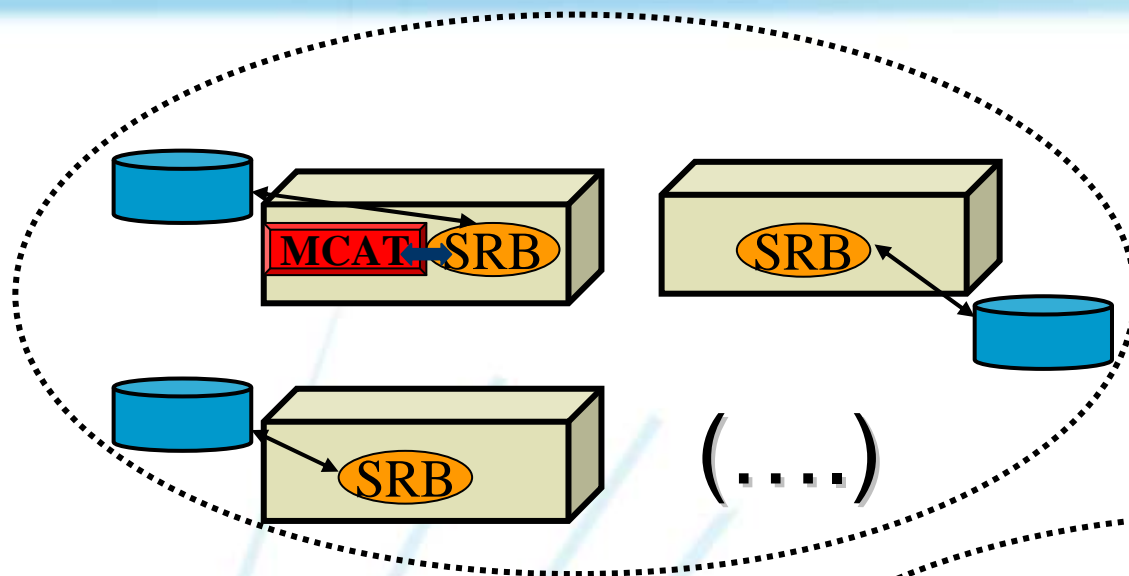


## Introduction à SRB

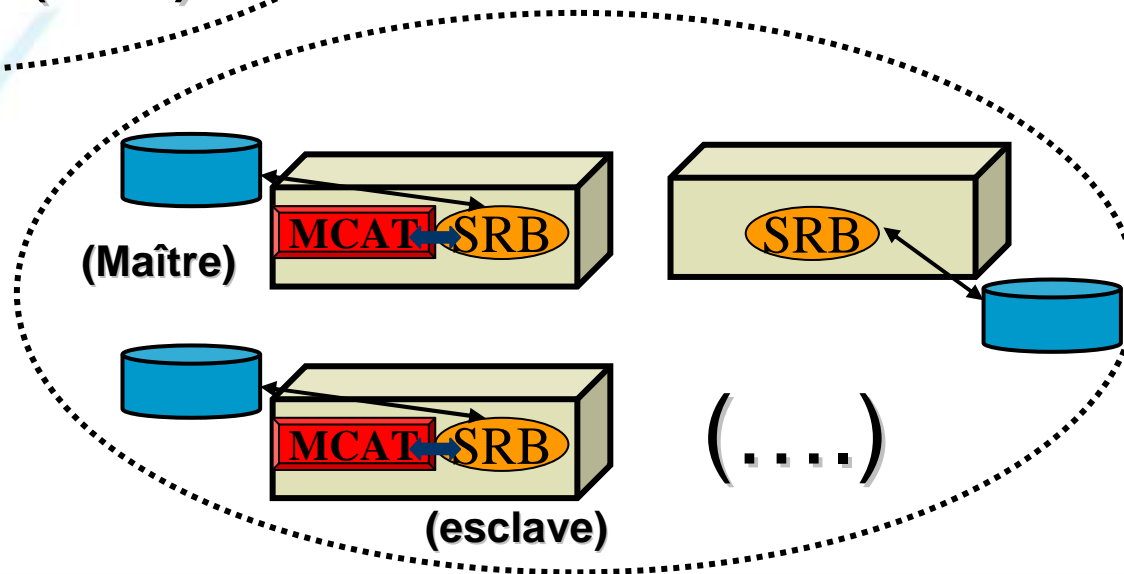


- SRB offre une réponse à ces aspects et à bien d'autres...
- Développé au SDSC: démarrage en 1998.
- Sous licence de l'Université de Californie / General Atomics.
- Développeurs au contact permanent des utilisateurs:
  - Ajout de fonctionnalités demandées par la communauté.
- Portage sur un grand nombre d'OS et de plateformes.
- Support de nombreux systèmes de stockage, non limitatif.

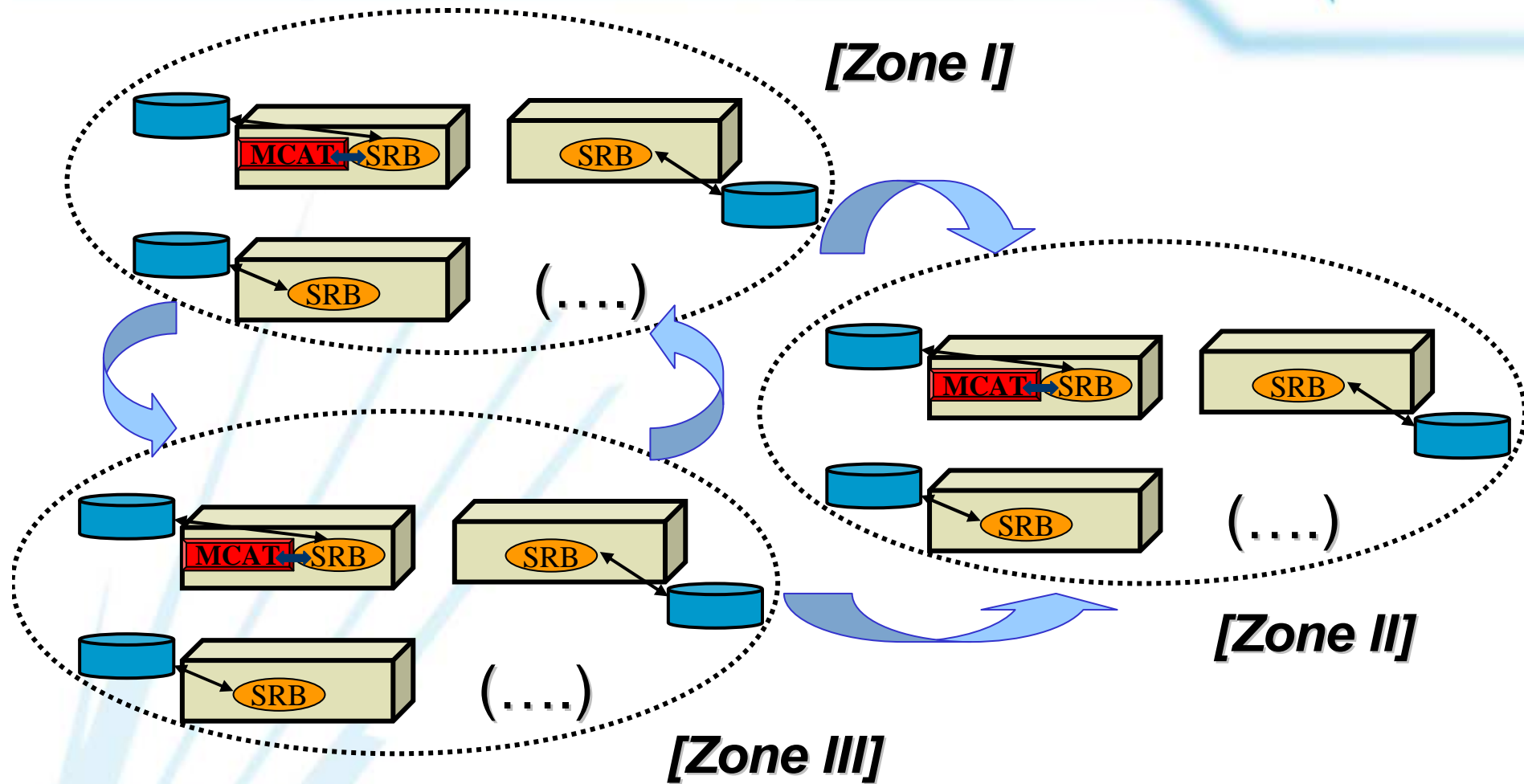
# Architecture de SRB (1 fédération)



OU



# Architecture de SRB (n fédérations)








# Fonctionnalités de SRB



- Organisation logique des données découplée de l'organisation physique:

Collection:	<b>prod02</b>
Collection Parent:	<b>/in2p3/mc/neutrino/nc</b>
Propriétaire:	<b>ant_write@ccin2p3</b>

Nom Donnée	Date Creation	Propriétaire	Num. Replica	Num. Version	Taille	Type Donnée
 <a href="#">nul01_evt.gz</a>	2006-10-24-14.15.30	ant_write@ccin2p3	0	0	27808157	generic
 <a href="#">nul02_evt.gz</a>	2006-10-24-14.15.30	ant_write@ccin2p3	0	0	27778433	generic
 <a href="#">nul03_evt.gz</a>	2006-10-24-14.15.28	ant_write@ccin2p3	0	0	27209758	generic

- Différents outils de navigation: GUI, Web, APIs, Scommands (Scd, Smkdir, Sput ...).
- Authentification: pwd, certificat X509.
- Organisation des utilisateurs par:
  - Type (sysadmin, domainadmin, simple utilisateur...).
  - Zones, domaines, groupes.
- ACL sur les objets et les répertoires.
- Notion de tickets: droits temporaires sur un fichier.



## Fonctionnalités de SRB



- Gestion des réplicats et des versions.
- Accès aux données à partir de leurs attributs au lieu de leur nom et emplacement physique.
- Recherche par métadonnées associées aux fichiers.
- Annotations possibles sur les fichiers.
- Audit: enregistrement de toutes les actions effectuées.
- Hiérarchie des ressources de stockage:
  - Ressources logiques: ensemble de ressources physiques.
  - Ressources composites: ressources caches (temporaires, ex: disque) + 1 ressource archive (ex: HPSS).





## Fonctionnalités de SRB



- Bases de données:
  - Extension du MCAT: interface de votre schéma relationnelle avec celui de SRB.
  - Objets cachés: accès à des bases de données via SRB (serveur proxy dans ce cas).
- Interface possible avec:
  - N'importe quel système offrant des APIs POSIX.
  - Castor, gridFTP, dCache....



## Utilisation de SRB dans le monde



- Utilisé en: Amérique, Asie, Australie, Europe.
- Disciplines:
  - Biologie (ex: FACETS), applications biomédicales (ex: BIRN, BBSRC, IB ...).
  - Astrophysique (NVO), LSST, Sciences de la Terre (ex: NASA, VORB).
  - Matière condensée (rayonnement synchrotron: UK, Australie).
  - Bibliothèques numériques (ex: NARA, bibliothèque du Congrès aux USA).
  - Sciences humaines.
  - HEP (BaBar, Belle).
- Projets de grilles:
  - TeraGrid (USA).
  - IVOA (International Virtual Observatory Alliance).
  - Grilles nationales en Australie, Hollande, UK, Allemagne.
  - Un des 9 projets de bases dans OGF (Open Grid Forum).
- Secteur privé:
  - Rolls Royce (suivi des anomalies sur les réacteurs d'avion).
  - Version de SRB distribuée par Nirvana (General Atomics): diffère du SRB SDSC !!!

# Utilisation de SRB au CC-IN2P3



<b>HEP</b>	<i>BaBar</i>	Site « miroir » de SLAC
	Calice	Archivage de données
	CMOS	Archivage de données
	QCD	~dizaines To / an
<b>Astroparticule</b>	Antares	Tier 0: ~200 To / an
	Auger	Tier 0: ~dizaines To / an
	Edelweiss	Tier 0: ~dizaines To / an
	<i>SN Factory</i>	Au cœur du online: ~Go / j
<b>Biomédical</b>	<i>BioEmergence</i>	Tier 0: Projet européen
	Mammographie	Projet avec le LIRIS (1 thèse)
	<i>Neuroscience</i>	ISC, CHU de Lyon, Strasbourg



## SRB au CC-IN2P3: côté serveur



- Manpower: Pascal + JY (??% d'un FTE)
- 3 serveurs:
  - 1 V440 et 1 V480 (Sparc III): Solaris 9.
  - 1 V40z (AMD Opteron): Solaris 10.
  - Total de 8 To de disque.
  - 1 serveur x4500 avec 20 To à venir + 1 serveur type??
- Bases de données: Oracle 10g
- HPSS: interface avec HPSS v5
- TSM: backup des données résidentes sur disque.



## SRB au CC-IN2P3: côté serveur



- Surveillance et redémarrage automatique des services (*à faire: inclure dans NGOP*).
- Gestion automatisé des ressources composites (disques SRB + HPSS):
  - Migrations/Purges régulées (fréquence, nbre de tâches en // etc...).
  - Notification via email en cas d'anomalie.
  - Base pour iRODS.

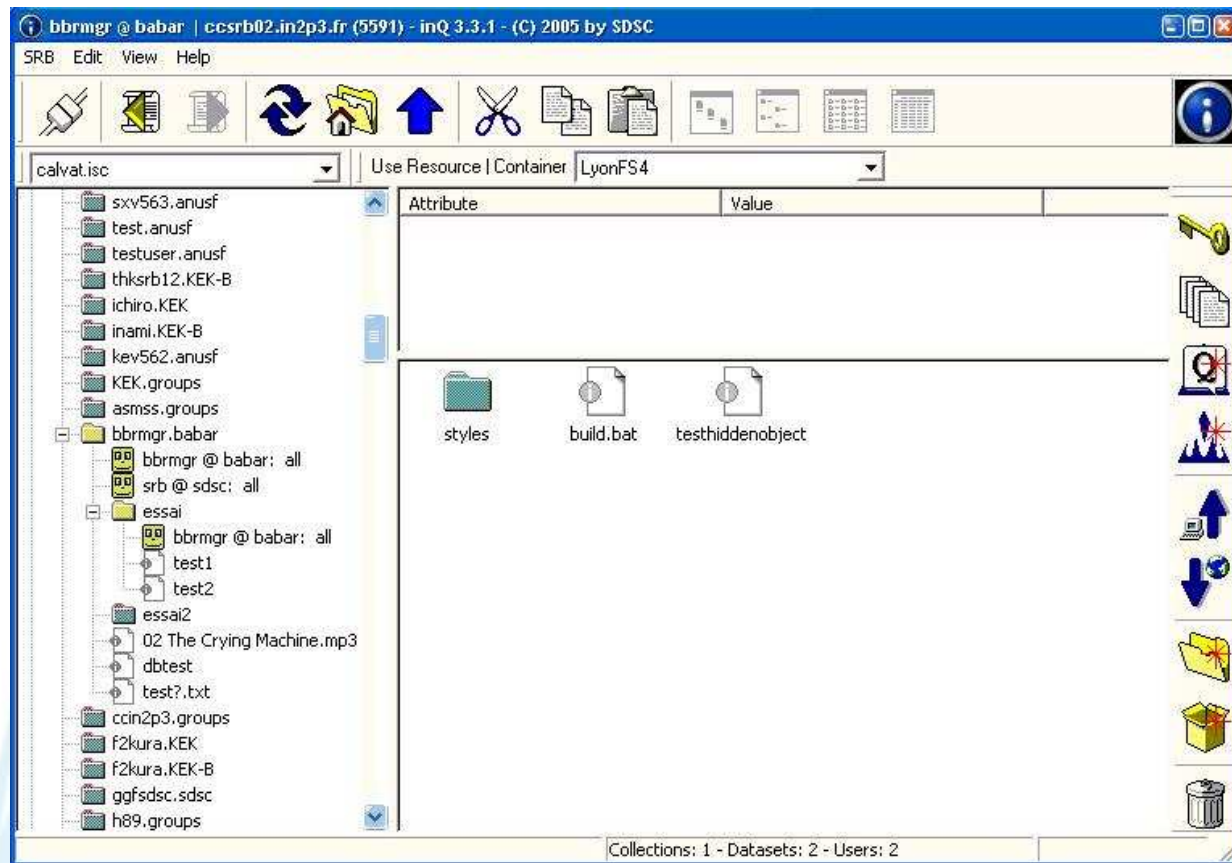


## SRB au CC-IN2P3: côté client



- Qu'utilisent-ils ?
- Code utilisant les:
  - Scommands:
    - Linux: Debian, Red Hat, Suse.
    - Solaris, Mac OSX, Windows.
  - APIs: Java, C++, Perl.
- Clients graphiques:
  - inQ (Windows).
  - mySRB (Web).

# SRB au CC-IN2P3: côté client



**inQ pour  
Windows**



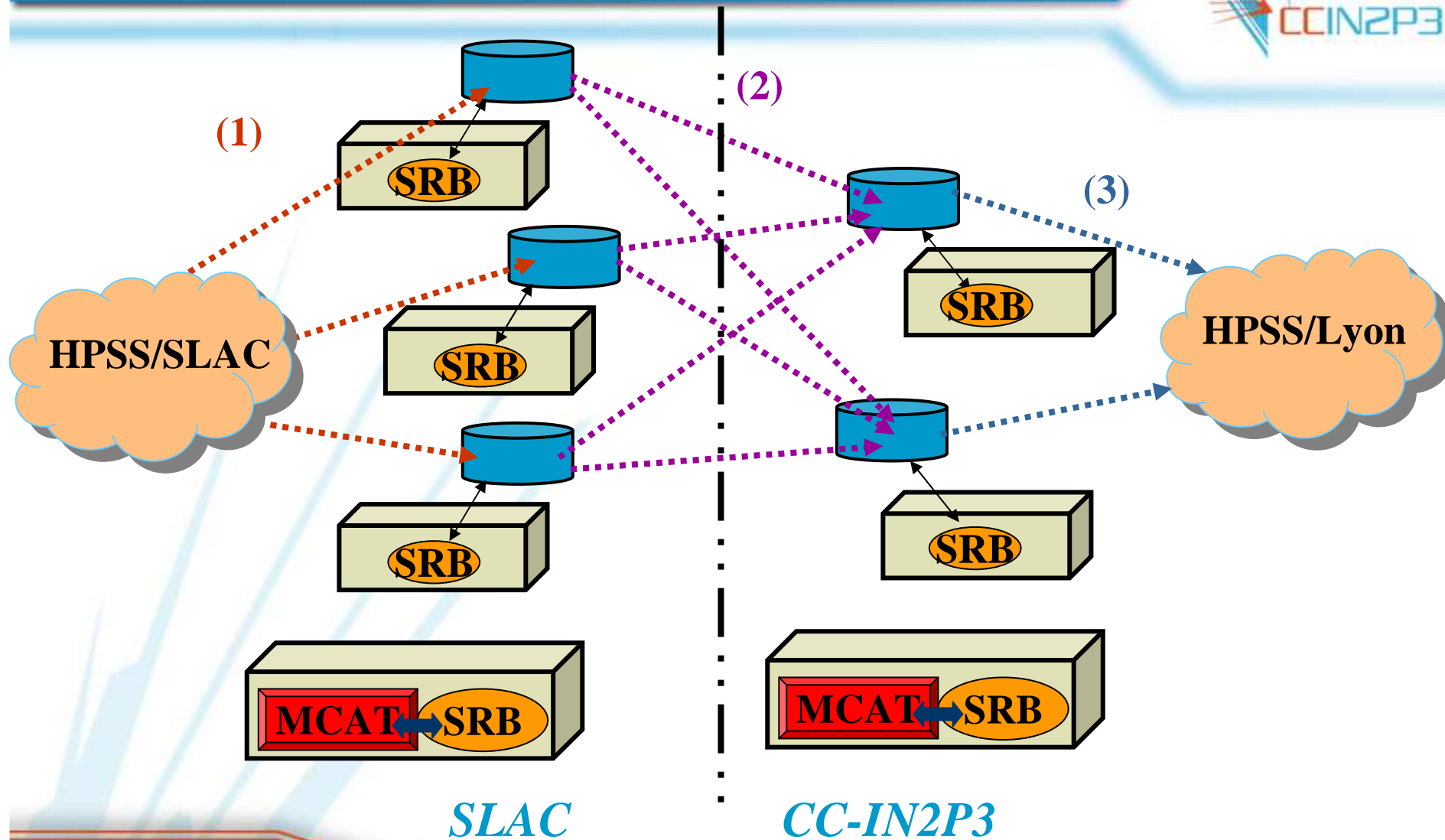
## Ex. d'utilisation au CC: BaBar



- CC-IN2P3: Tier A de BaBar.
- Sous-ensemble des données produites au CC-IN2P3 pour analyse.
- SRB utilisé depuis 2004 en prod:
  - Développement rapide des applicatifs.
  - Fiabilité.
  - Extensibilité, souplesse d'administration.



# Ex. d'utilisation au CC: BaBar





## Ex. d'utilisation au CC: BaBar

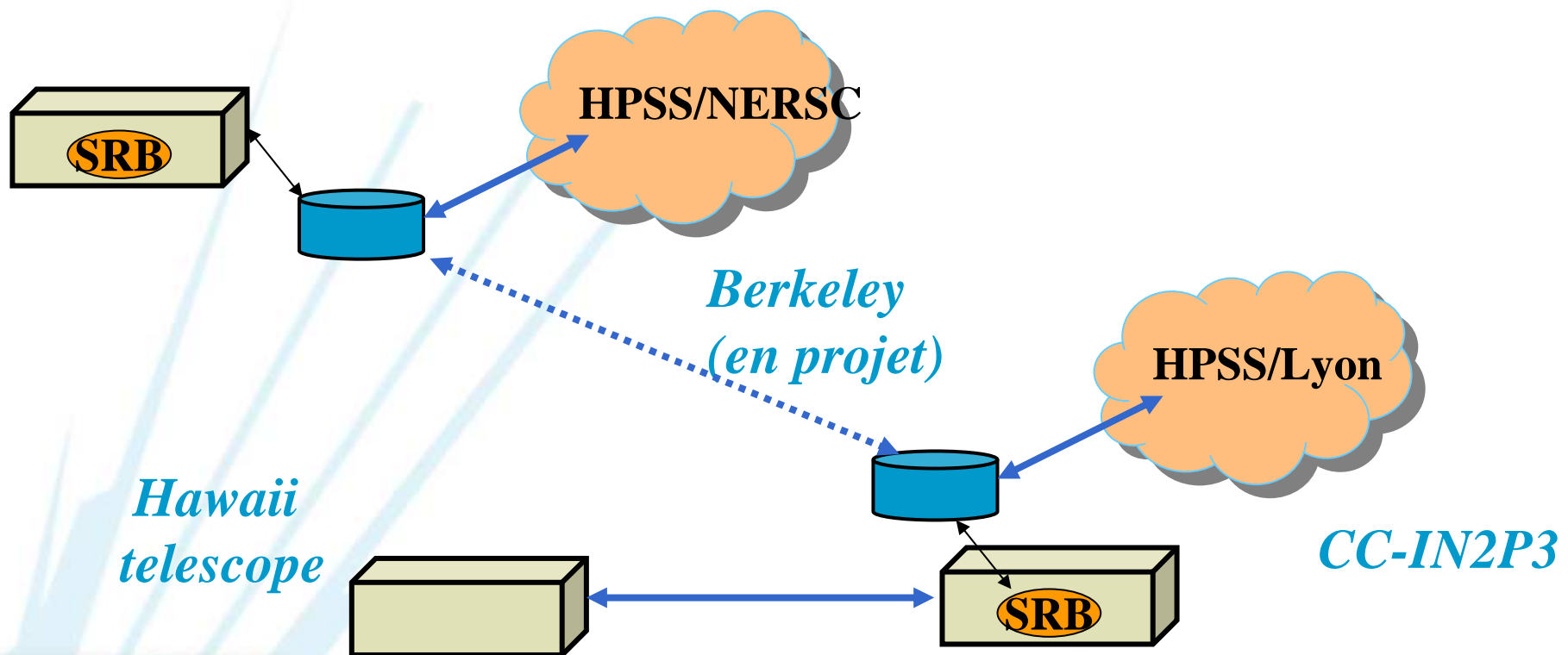


- Procédure d'import multithreads.
- **Entièrement automatisée:**
  - enregistrement à SLAC.
  - détection de fichiers absents à Lyon.
  - Récupération des erreurs.
- Jusqu'à 5 To / jour (pas de limitation).
- 330 To importé, 630000 fichiers.
- Jusqu'à fin 2008, volume x2 au minimum.

## Ex. d'utilisation au CC: SN Factory



- SRB: partie intégrante du online!
- Utilisation comme outil collaboratif (mise en commun de fichiers privés).



# Ex. d'utilisation au CC: Neuroscience



**IRM**

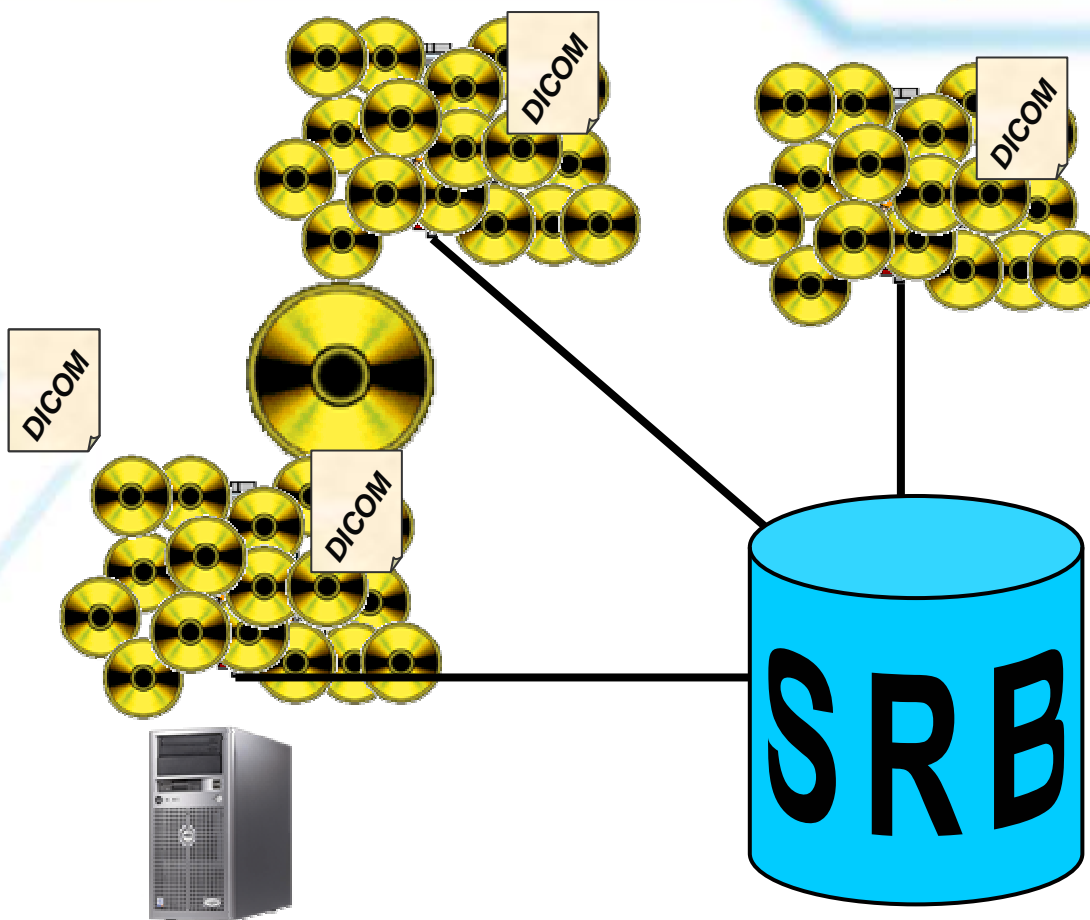
Siemens MAGNETOM  
Sonata Maestro Class 1.5 T

Acquisition



**Consol**

• push DICOM



**Export PC (serveur DICOM,  
clients SRB)** Dell PowerEdge 800

Siemens Celsius Xeon  
(Window NT)  
SRB et iRODS au CC-IN2P3

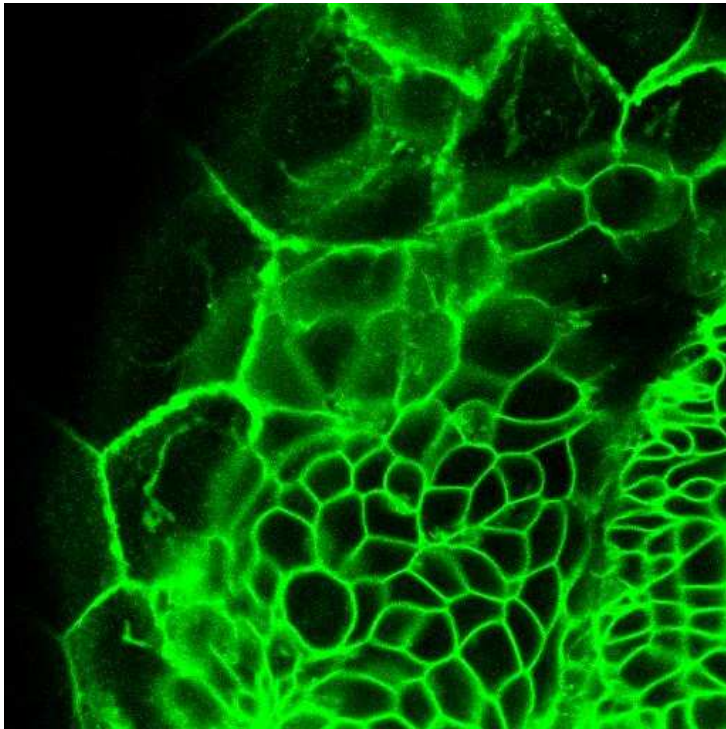


## Ex. d'utilisation au CC: Neuroscience



- **Développements:**
  - Enregistrement automatisé des fichiers dans SRB à partir des hôpitaux (Lyon, Strasbourg).
  - Gestion de l'anonymisation (prise en compte aussi au moment de l'acquisition).
  - Interface graphique SRB dans MATLAB.
- **Chaîne totalement opérationnelle.**
- **A faire:**
  - Interface serveur DICOM / serveur SRB au CHU.
  - Officialisation auprès de la CNIL: document rédigé avec un juriste de la DSI.
  - Projet DatScan (participation de General Electrics).
- **A terme: participation à BIRN ?**

## Ex. d'utilisation au CC: BioEmergence



- Embryogénèse: poisson zèbre.
- Plusieurs microscopes à travers l'Europe.
- Données poussées des microscopes dans SRB.
- SRB intégré dans le workflow (calcul distribué).
- CC-IN2P3: référentiel de la collaboration.
- Serveurs SRB dans d'autres labos ?



## Bilan d'utilisation au CC-IN2P3



- Beaucoup de fonctionnalités utilisées...
- ... mais pas toutes ☹, par exemple:
  - Extension du schéma du MCAT.
- Développements:
  - Côté serveur (surveillance, gestion des ressources composites, ...).
  - Côté client (applicatifs pour BaBar, neuroscience etc...).
- Documentation (FAQ).
- Applications graphiques (ex: inQ) alléchantes mais dangereuses:
  - Trop facile → utilisées en toute inconscience.
- Aussi vrai pour les APIs, Scommands...



## Perspectives globales pour SRB



- Outil à maturité, plus de développements.
- Support assuré pour au moins 3 ans (corrections de bugs).
- Bonne adéquation avec un grand nombre de besoins, mais...





## De SRB à iRODS



- ...SRB n'est pas open source:
  - Problème pour projets avec partenaires privés.
  - Distribution libre du code...
  - **Pas un problème dans le secteur académique!!!**  
*(accès et modification du code source tout à fait possible).*
- Code et schéma de la base de données à retravailler: simplifier (rappel: démarrage en 98).



## De SRB à iRODS



- Virtualisation du stockage pas suffisante.
- Applications clientes s'appuyant sur ces middlewares:
  - Pas de garde-fous.
  - Pas de garanties d'application stricte d'une politique de préservation des données.
- Nécessité pour un projet de distribution de données de définir une politique homogène et cohérente de:
  - gestion des données.
  - gestion des ressources de stockage.
- Essentiel dans les projets d'archivage massif de fichiers (bibliothèques numériques...).
- Aucun outil de grille actuellement ne permet cela.



## De SRB à iRODS



- Cas typiques de chausse-trappes:
  - Non respect de certaines règles préétablies.
  - Plusieurs applicatifs coexistent pour gérer les données et le workflow → *incohérences potentielles*.
  - Plusieurs versions du même applicatif utilisé au sein du projet → *incohérences potentielles*.
- Délester l'application cliente des contraintes locales à chaque système.
- Solution: virtualisation de la politique de gestion des données, exprimée sous forme de règles.



## Quelques exemples...



- ACL personnalisés dans le système:
  - Interdire effacement de fichiers dans une arborescence particulière même par le propriétaire.
- Intégrité et sécurité des données:
  - Checksum automatique lancé en arrière plan.
  - Anonymisation à la volée même si non effectuée par le client.
- Enrichissement des métadonnées:
  - Enrichissement automatisée des métadonnées associées aux objets (dans ou hors de la db du système).
- Personnalisation des paramètres de transfert:
  - Nombre de flux, taille des paquets, fenêtre TCP variant en fct de l'origine de la connexion à un serveur.
- ... à vos plumes ...



- **iR**ule **O**riented **D**ata **S**ystems.
- Projet démarré en Janvier 2006.
- Première version officielle en Décembre 2006 (v 0.5).
- Financé par: NSF, NARA (National Archives and Records Administration).
- CC-IN2P3, e-science (UK): collaborateurs.



- Reprend les principes de SRB.
- iCAT de iRODS  $\Leftrightarrow$  MCAT de SRB.
- Mais va beaucoup plus loin:
  - Gestion des données basée sur des règles paramétrées côté serveur.
  - Personnalisation complète du système sans avoir à modifier le code iRODS.
  - Écriture de ses propres services en ajoutant ses propres modules.
  - Virtualisation de la politique de gestion des données.
  - Espace logique pour les règles:
    - Regroupement en ensembles de règles.
    - Gestion des versions.



# iRODS: les règles



- Une règle (préfixe *ac*) est composée de:
  1. **Nom.**
  2. **Condition.**
  3. Appel de fonction(s): autre(s) règle(s) ou micro-services.
  4. **Récupération en cas d'erreur.**
- Un micro-service (préfixe *msi*):
  - Effectue une tâche précise, peuvent faire appel à des fonctionnalités internes de iRODS.
  - Interface standard fournie avec les micro-services
- Exemple de règle (appelée lors de l'effacement d'un fichier):  
***acDataDeletePolicy//nop/nop***



# iRODS: les règles



- Modification de la règle précédente:
  - Empêcher l'effacement des données dans /in2p3/RealData:  
*acDataDeletePolicy/\$objPath like /in2p3/RealData\*/msiDeleteDisallowed/nop*
- Enchaînement de micro-services et de règles (exemple: création d'un utilisateur):

*acCreateUser//msiCreateUser##acCreateDefaultCollections##msiCommit/  
msiRollback##msiRollback##nop*

*acCreateDefaultCollections//acCreateUserZoneCollections/nop*

*acCreateUserZoneCollections//acCreateCollByAdmin(/\$rodsZoneProxy/home,\$ot  
herUserName)##acCreateCollByAdmin(/\$rodsZoneProxy/trash/home,\$otherUserN  
ame)/nop##nop*

*acCreateCollByAdmin(\*parColl,\*childColl)//msiCreateCollByAdmin(\*parColl,\*child  
Coll)/nop*

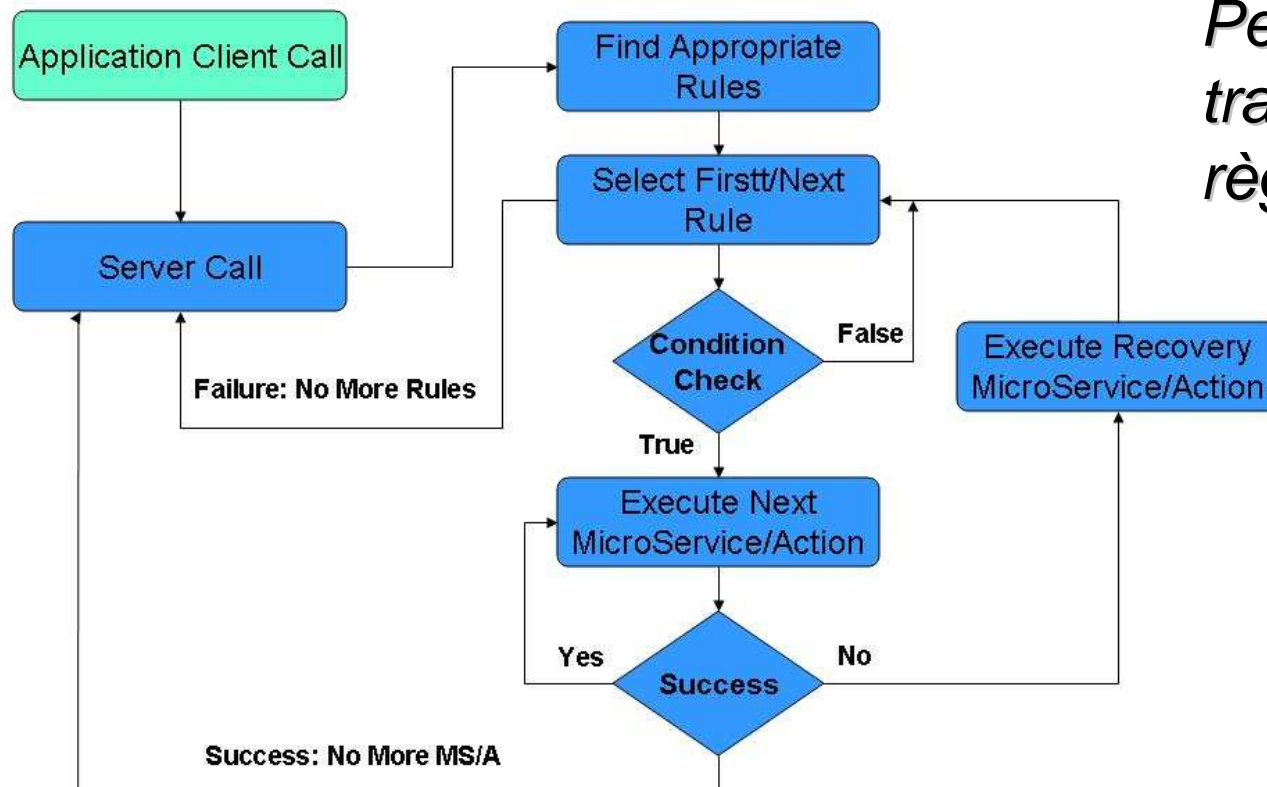




# iRODS: Rule Engine



## Rules Flow



*Permet le traitement des règles.*

1



## iRODS: Rule Engine



- Possibilité d'avoir des actions en tâche de fond au sein d'une règle (liste gérée par la base de données).
- Possibilité pour un utilisateur de:
  - Surveiller l'état des tâches différées.
  - Modifier des paramètres pour des tâches différées.
  - Annuler des tâches différées.
  - Soumettre ses propres règles à un serveur.



## iRODS: à faire



- Support de GSI.
- Sessions à durée limitée.
- Clients Python, Java.
- Interface graphique (Java, Python ou autre).
- Driver pour HPSS, SAM-QFS, autres...
- Portage + complet de Unix/Linux (déjà RH, Mac OS, Solaris!), sur Windows.
- Support de Oracle, MySQL (PostGres pour l'instant).
- Méthode pour les utilisateurs d'influencer les règles (mais pas d'en écrire).
- Scripts d'installation et de tests plus complets.
- Agrégation des petits fichiers.
- **Outils de migration de SRB vers iRODS.**
- Extension du schéma de iRODS.
- Zones, fédérations.
- Auditing.
- API en PHP utilisant XML pour communiquer avec iRODS. Ajax ?



## Participation du CC-IN2P3 à iRODS



- Serveur au CC:
  - Utilisation des iCommands ( $\Leftrightarrow$  Scommands).
- Spécifications, tests, debugging.
- Programme de tests.
- A faire, écriture de micro-services:
  - contrôle d'accès personnalisé aux serveurs, gestion des ressources composites (caches + archives), anonymisation des données au format DICOM...
  - interface HPSS.



## Conclusion



- SRB:
  - Middleware stable et riche en fonctionnalités.
  - Communauté très diversifiée.
- ~1 Po enregistré dans SRB vers fin 2008 au CC-IN2P3.
- iRODS ouvre des perspectives très riches:
  - Personnalisation poussée du paramétrage des services.
  - Côté client: virtualisation des règles de gestion des données.
- Excellente collaboration avec le SDSC.
- **... mais un bon middleware ne fait pas tout:**
  - Définition et élaboration de la politique de préservation des données.
  - Définition et élaboration des workflows: évident mais trop souvent négligé.



## ■ SRB:

- [http://www.sdsc.edu/srb/index.php/Main\\_Page](http://www.sdsc.edu/srb/index.php/Main_Page)
- <http://cc.in2p3.fr/rubrique339.html>

## ■ iRODS:

- [http://irods.sdsc.edu/index.php/Main\\_Page](http://irods.sdsc.edu/index.php/Main_Page)