



# Installation de logiciels dans AFS

## Installation via la Grille

[afsmaster@cc.in2p3.fr](mailto:afsmaster@cc.in2p3.fr)

Contraintes dues à AFS

Contraintes des Expériences

Systemes utilisés et softgrp

# Partage, lecture de petits fichiers !



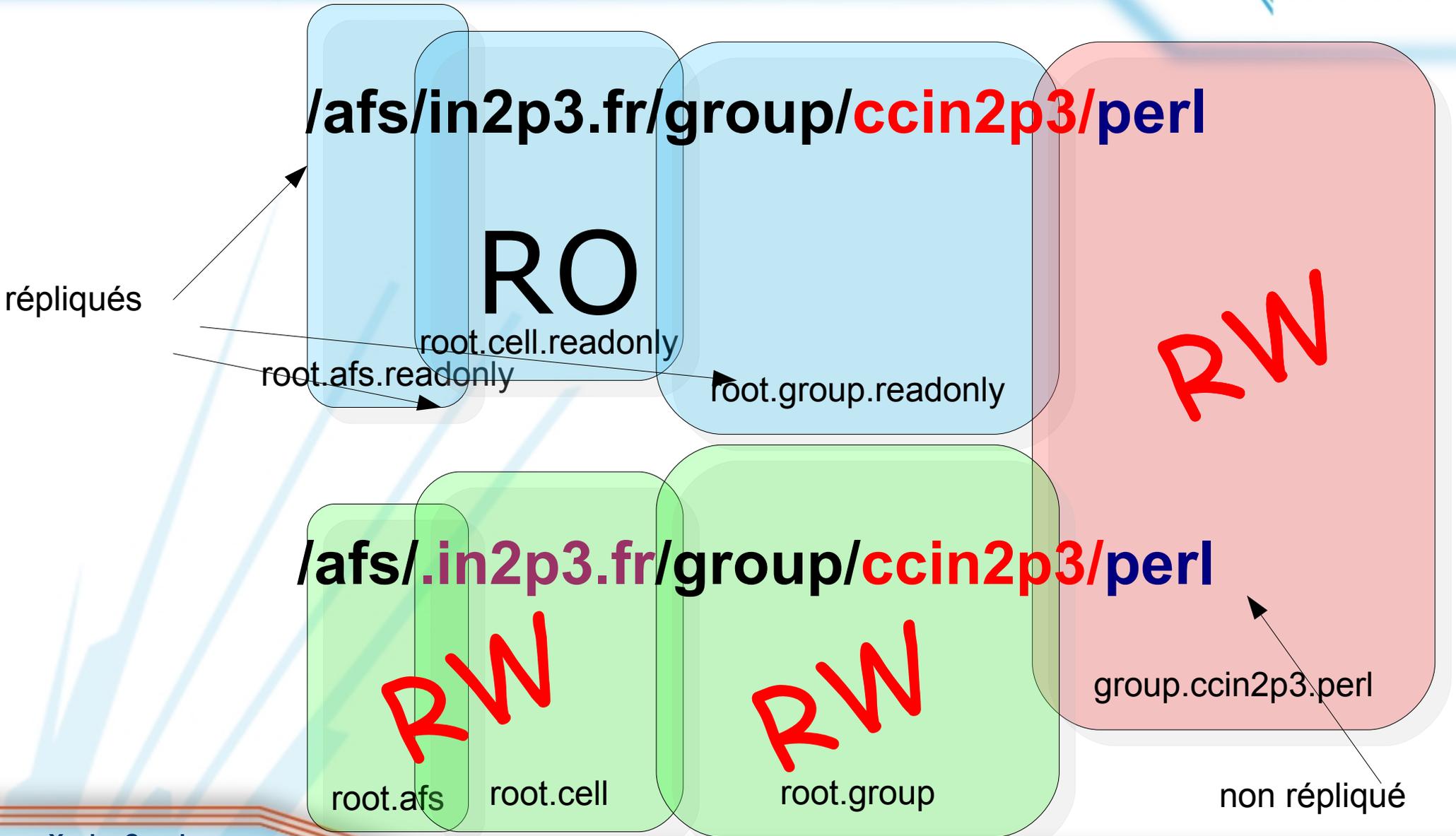
- Système de fichiers distribué à racine unique  
Partage de données aisé  
Cellule `/afs/in2p3.fr`
- Architecture client/serveur
- Scalable par réplication en lecture seule  
Distinction des accès en écriture ou lecture seule  
**`/afs/in2p3.fr` vs `/afs/.in2p3.fr`**
- Migrations transparentes entre serveurs, sans arrêts de production

## AFS n'est pas :



- Un système de distribution de données en masse, de hautes performances
  - Pas de gros volumes
  - Pas de gros fichiers
- Un système de distribution de données à accès intensif
  - Pas les performances de GPFS
  - Sectarisation des volumes sur les serveurs selon le type d'utilisation

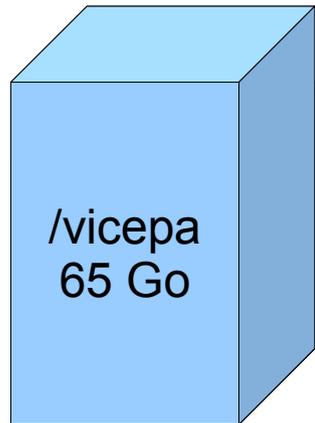
# Path RO, Path RW, volumes



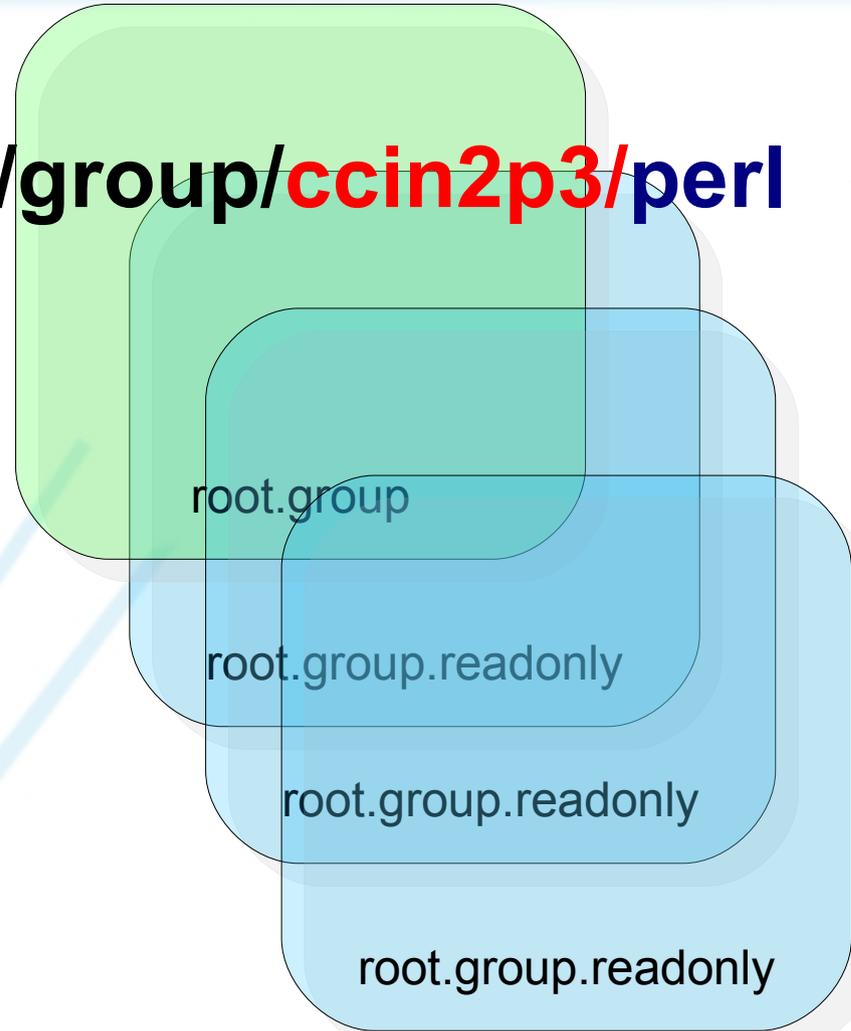
# Volumes, partitions et serveurs



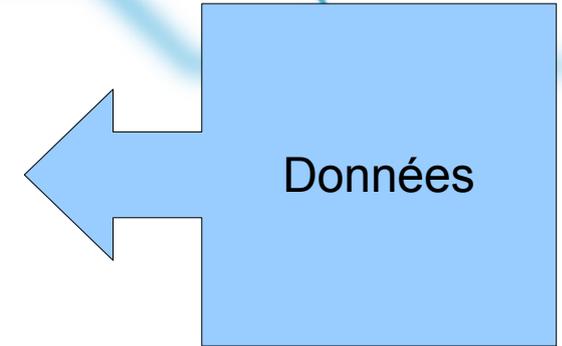
**/afs/.in2p3.fr/group/ccin2p3/perl**



Partitions gérées par les  
afsmasters/sysadmins



Réplication gérée par les afsmasters



Contenu géré par  
l'utilisateur d'un volume

- La gestion de volumes nécessite des privilèges d'administration AFS
- La réplication n'est **pas** automatique
- Le cachemanager gère automatiquement le type de volume accédé
  - Mais si l'on quitte le path RO, on peut pas y revenir automatiquement

- Plusieurs espaces utilisés
  - \$GROUP\_DIR
  - \$THRONG\_DIR
  - \$GRID\_TOOLKIT
  - \$SYSTEM
- Libre choix des expériences pour les 3 premiers
- \$GROUP\_DIR propose l'espace le plus vaste

## Contraintes dues à LHC

- Nombre d'accès par job
- Nombreux jobs
- Grosse volumétrie
- Lecture principalement
- Paths écrits en dur
- Installation automatique sans contrôle fin du CC
- Installation  $\neq$  Mise à jour

## Mécanisme AFS

- Configuration cache
- Réplication RO
- Volumes multiples
- Volumes RO
- Path RW vs RO
- Outil de réplication ?

- Demande initiale d'un seul gros volume
- Installations de releases dans  
`$VO_ATLAS_SW_DIR` de  
`{software,dev,prod,nightlies,ddm}/releases/<rel#>`
- Choix du `GROUP_DIR` à cause
  - Des outils de gestion des volumes
  - Du type d'accès

**`VO_ATLAS_SW_DIR=/afs/in2p3.fr/group/atlas/sw/`**

- Ajout d'une hiérarchie de répertoires/volumes qui impose des interventions privilégiées AFS
- Une release s'installait via :
  - mkdir \$GROUP\_DIR/sw/prod/releases/<rel#>  
**=> 1 gros volume READ-WRITE**
  - Demande de réplication et lancement de production  
**=> synchronisation nécessaire**
- Paths en dur, donc impossibilité d'utiliser /afs/.in2p3.fr vs /afs/in2p3.fr  
**=> manipulations des points de montage**

- Volumes RW montés sur \$GROUP\_DIR/sw/prod/releases/<rel#>
  - Ok pour les patchs manuels, réplication RO facilitée
    - => **fonctionnement normal du « vos release »**
- Gros patchs via l'outil de déploiement
  - => **manipulation RO → RW qui ne peut se faire sous charge de prod importante**

- Création simplifiée de volume
- Réplication simple
- Aucune manipulation sur les points de montage
- Pas de gestion utilisateur du quota
- Pas de gestion utilisateur des pools de serveurs
  
- Reste à :
  - Gérer les codes retour
  - Gérer les PATHs « en dur »

## Deux commandes mise en place



- `/usr/local/shared/bin/afs_create`
- `/usr/local/shared/bin/afs_release`
- Localisé dans le PATH issu des scripts de profile AFS

**`afs_create <path>`**  
**`afs_release <path>`**

- Installation par un utilisateur AFS (token)
- Utilisation de « adminexec »
- Privilèges positionnés via sprivc
  - afsad.rocreate
  - afsad.rorelease
- Scripts shell
- Fichier de configuration \$USS/group/ro\_software
- La contrainte majeure sur cette infrastructure :  
Adminexec => ccsv99 => AIX 4.3.2 => ksh

## ■ AVANT

```
mkdir $PATH  
<installation>
```

## ■ APRÈS

```
afs_create $PATH_RW  
<installation>  
afs_release $PATH
```

## ■ AVANT

```
cd $PATH  
<patch>
```

## ■ APRÈS

```
cd $PATH_RW  
<patch>  
s|.in2p3.fr|in2p3.fr|  
afs_release $PATH
```

- Locks d'importance différente imbriqués
  - Afsmaster
  - Czar
  - Utilisateur afs\_{create,release}
- Plusieurs paths possibles pour une expérience
- Plusieurs pools de serveurs possibles pour une expérience
- Pools de serveurs potentiellement distincts par expérience

# Pourquoi ne pas avoir utilisé les outils en place ?



- Ce n'est pas le czar qui fait ces installations
- Logiciels inadaptés
- Pas de réplication autorisée aux czars

**\$GROUP\_DIR/<volume> est RW !**

**/afs/in2p3.fr/group/ccin2p3/perl**

**Utilisation de \$SFTGRP :**

**/afs/in2p3.fr/sftgroup/ccin2p3/perl**

- ccsvsn14 remplace ccsv99
  - Solaris 10
  - OpenAFS 1.4.10
  - Perl
- Kerberos 5
  - Autres outils de délégation de privilège

# La contrainte majeure pour les expériences



- Déplacer les softs de \$WTF vers \$SFTGRP
- Gestion des liens inscrits « en dur »
- Les liens dans AFS ne sont pas une bonne idée si les PATHs sont inscrits en dur
  - /afs/.in2p3.fr/sftgroup/ccin2p3/perl contient le path RW
  - un vos release pousse en RO ce path
  - L'appel à /afs/in2p3.fr/sftgroup/ccin2p3/perl attaque toujours le path RW, même si le binaire est appelé depuis un volume RO