



Accès à la Grille en Rhône-Alpes

Yonny CARDENAS
Journées Grilles France,
Lyon, 14 - 16 Octobre 2009

Contenu



- Introduction
- Grille Rhône-Alpes
- Portage des Applications
- Conclusions

Introduction



Pourquoi faciliter l'accès à la grille ?

Sondage deuxième semestre 2008, prospection Institut de Grilles 3000 chercheurs

La très grande majorité (environ **80%**) des chercheurs encore peu impliqués dans les grilles mais ils jugent qu'il existe un très fort potentiel dans cet outil qu'ils souhaiteraient mettre à profit.

Problématique



Point de blocage principal l'**adaptation** entre les connaissances domaine scientifique spécifique et les techniques des grilles.

Recommandation



Groupe transverse: Accès à la grille

Structure pour le portage des applications

équipe d'ingénieurs -> interface entre l'infrastructure grille et les utilisateurs de différentes disciplines scientifiques

Mise en œuvre dans la Grille Rhône-Alpes.



Grille Rhône-Alpes

Grille Rhône-Alpes



Structure qui mutualise les ressources et services informatiques de plusieurs unités situées sur la région Rhône-Alpes en facilitant l'accès auprès de plusieurs disciplines scientifiques.

NGI Française



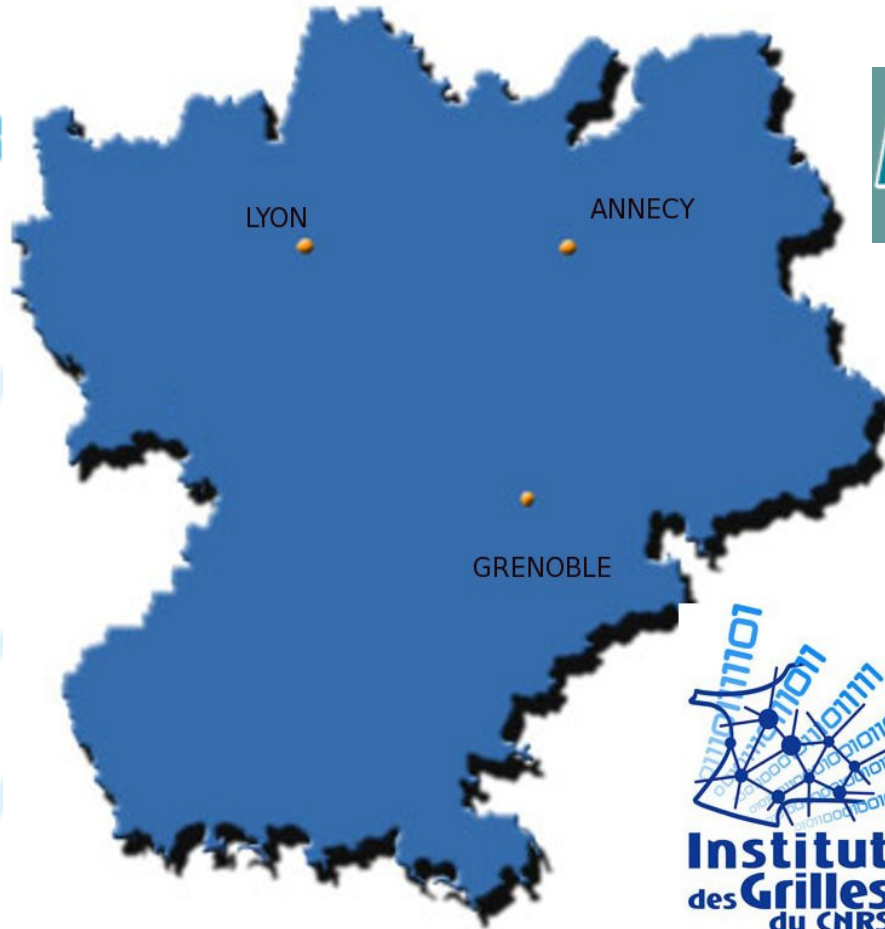
NGI



NGI ...

Grilles Régionales

▶ Composition



▶ Ressources



Ressources sur les 5 sites (potentiellement) :

Calcul: 7000 Cœurs (*)

Stockage: 51 Téraoctets (*)

Middleware:

Job management :

gLite, LRMS (spécifiques)

Autres (Cigri, OAR, DIET)

Stockage:

SRM, IRODS

Utilisateur:

JSAGA implementation SAGA

Organisation virtuelle:

vo.rhone-alpes.idgrilles.fr

contact:

contact@rhone-alpes.idgrilles.fr

Grille Rhône-Alpes



Objectifs:

Permettre aux scientifiques de faire d'une meilleure façon la recherche.

Faciliter l'accès aux infrastructures grilles à plusieurs communautés scientifiques.

Adapter différentes technologies grilles pour mieux répondre aux divers besoins.

Caractéristiques



Proposer une assistance et conseil technique aux utilisateurs.

Répondre aux besoins spécifiques des utilisateurs/ communautés locales:

- Exemple: Biomédical (Lyon)

Approche: Accompagnement intégral pour le portage des applications.

Portage des Applications

Grille Rhône-Alpes

Processus de portage



- Prise de contact
- Évaluation des besoins
- Analyse
- Plan de Portage
- Plan de mise en production
- Production

Processus de portage



- **Prise de contact**
 - Initiative de l'utilisateur
 - Travail prospectif
- **Évaluation des besoins**
 - Établir bénéfices à attendre
- **Analyse**
 - Analyse de fiabilité
 - Évaluation des alternatives techniques
 - Estimation initiale de ressources

Processus de portage



- Plan de Portage (1/2)
- Identification de techniques et procédures grilles à utiliser
 - contexte distribué: calcul - données - accès
- Adaptation de l'application (réingénierie)
 - découpage de calcul global (tâches, jobs, concurrence, robustesse)
 - gestion des données (distribution, accès)
 - bibliothèques, licences

Processus de portage



- Plan de Portage (2/2)
- Adaptation pour la production
 - adaptation des outils ou services grilles (environnement)
 - optimisation et bonnes pratiques
 - post production (récupération de résultats, validation)
- Prototype/testing/optimisation
 - Local -> système de batch -> un site -> plusieurs sites

Processus de portage



- Plan de mise en production
 - Allocation et utilisation des ressources
 - Vérification et/ou configuration de services et environnement
 - Répartition de la charge (temps, sites, ressources)

- Production
 - Validation de résultats

HOVERGEN - LBBE



LBBE - Laboratoire Biométrie et Biologie Évolutive CNRS
U. Lyon1

Mise à jour de bases de données biologiques dédiées à la phylogénie (étude de l'évolution des organismes vivants afin d'établir leur parenté).

Calcul intensif (Blast): Comparaison entre elles de centaines de milliers de séquences biologiques et d'alignements de séquences et d'arbres phylogénétiques.

HOVERGEN (1)



HOVERGEN base de familles de gènes de vertébrés.

Calcul: 35 000 jobs exécutés, 1455 millions heures SI2K

Jobs : 40 min

Stockage: SE/SRM 26 000 fichiers

Base de données HORVERGEN distribuée (sur 4 sites)

Calcul en 1 semaine au lieu 2 ans (avril 2009)

HOVERGEN (2)



HOVERGEN base de familles de gènes de **toutes les espèces.**

Calcul: 1200 jobs, 5280 millions heures SI2K

Jobs: 20 heures

Optimisation du calcul – jobs longs

Stockage: IRODS (future utilisation de metadonnées)

Calcul en 1 semaine au lieu 8 ans (en cours)

Gestion de données



Centre de recherche en imagerie médicale
CREATIS CNRS UMR5220–INSERM U630

Trois projets de recherche biomédicale
(Poumon 4D, Renata, Cœur).

Gestion de données



Stockage et partage de manière sécurisée de larges volumes de données (images médicales).

Middleware: IRODS (SRB)

Gestion d'accès basée sur metadonnées DICOM.

Base d'images médicales partagée:

400 000 fichiers en production

Conclusions



- Processus de portage devient un processus de réingénierie de l'application
- Assistance et conseil technique continus (pris en charge par un groupe dédié)
- Comprend de l'expression de besoins jusqu'à mise en production
- Pas de cas standard, chaque cas est particulier
- Adaptation des outils grilles et/ou intégration différentes technologies grilles qui répondent mieux aux besoins
- Demande un effort considérable (utilisateur, médiateur) difficile à estimer au départ



Accès à la Grille en Rhône-Alpes

Yonny CARDENAS
Journées Grilles France,
Lyon, 14 - 16 Octobre 2009

- Proposition

Workshop Grille Rhône-Alpes

2ème ou 3ème semaine de décembre

CC-IN2P3 Lyon