



Enabling Grids for  
E-science in Europe

[www.eu-egee.org](http://www.eu-egee.org)

*Tutorial Grille (LCG/EGEE)*

# Middleware EGEE



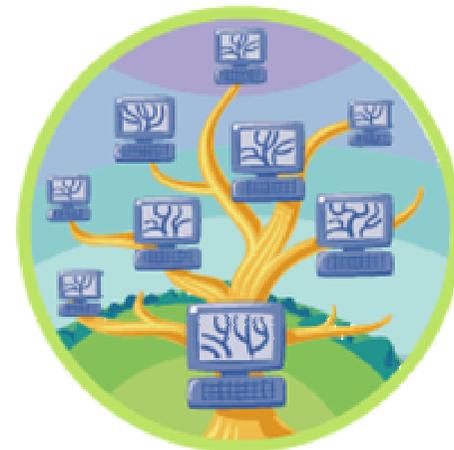
E.Fede (CNRS/LAPP)

Tutorial Utilisateurs EGEE

26 Septembre 2007

EGEE is a project funded by the European Union under contract IST-2003-508833

- Pourquoi la grille
- Le Middleware de grille
  - Pourquoi un middleware ?
  - Les challenges à relever.
- Historique
- Le Middleware d'EGEE



# Pourquoi la grille ?

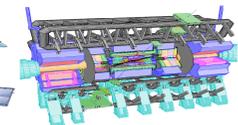
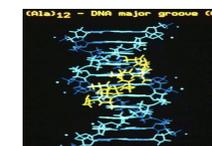
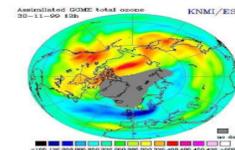
- Pour en avoir plus
  - A travers la mutualisation
    - Des ressources :
      - Plus de sites, plus d'applications, plus de services
    - Mise en commun des expertises
    - Modèle :  $X \text{ mutualisé } Y > X+Y$
- Pour en avoir moins
  - A travers la mutualisation
    - Des coûts :
      - Optimisation des infrastructures
    - Modèle :  $\text{Efficacité}(X \text{ mutualisé } Y) > \text{Eff}(X) + \text{Eff}(Y)$
- Pour simplifier/standardiser
  - Uniformisation de l'accès à l'outil informatique

# Pourquoi un middleware de grille

- Les ressources
  - De calcul, de stockage, de communication,...
  - Des données
- Middleware (intergiciel) : services de base permettant l'accès transparent aux ressources
  - Soumission de jobs
  - Accès aux données
  - ....
- Les applications
  - Consommateurs des ressources
  - Accèdent aux ressources a travers le middleware
  - De nature très diverses, avec des modèles d'organisations disparates



## Middleware de la grille



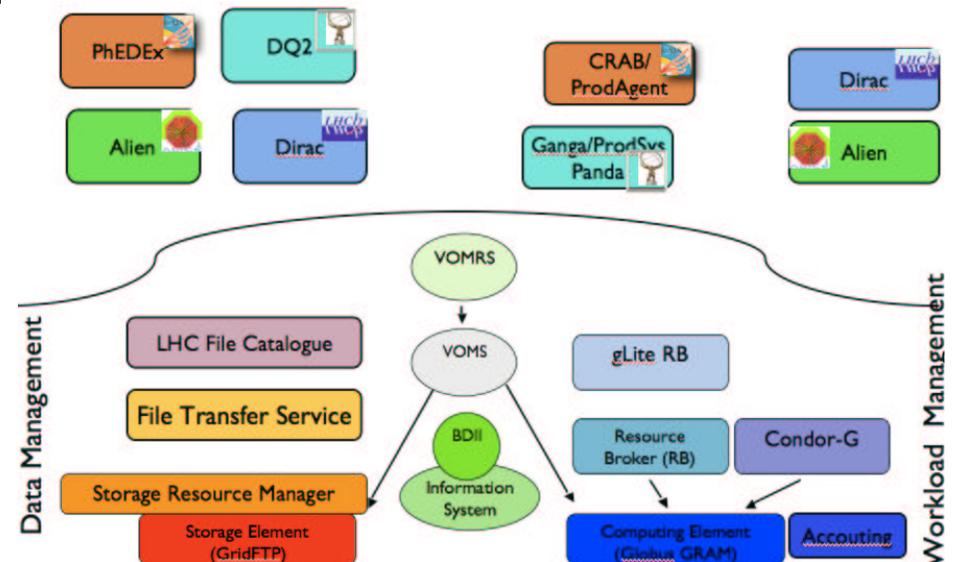
# Les principaux challenges du middleware

- Fournir des interfaces standard vers les ressources de calcul et aux données dans un contexte de sécurité intégrée
  - Ne doit pas être spécialisée pour une application
  - Définir des APIs pour l'ensemble des services
- Ressources et utilisateurs appartenant à des domaines « administratifs » différents
  - Pas d'accords bilatéraux directs
- Gestion de ressources dynamiques et hétérogènes
  - Forte distribution qui rend impossible des choix homogènes
- Accounting des ressources mises à disposition
  - Doit être possible d'exécuter un contrôle à posteriori
  - Valorisation des contributions
- Grands volumes de données distribuées
- Contrôle de l'accès aux ressources et aux données
  - Encryptage nécessaire pour certains types de données

- Plusieurs générations de middleware/grille
  - Fin des années 90
    - Seti @ home
    - Pas vraiment une grille mais plutôt du calcul distribué, fruit de recherches sur le meta scheduling
    - Sont définies les bases du grid computing : The Grid, I. Foster, C. Kesselman, 1998
  - 99-2000
    - Globus, Legion
    - Apparition des services d'informations, et des aspects liés à la sécurité
  - Depuis 2001
    - EDG, EGEE, Nordugrid, OSG, GridEngin
    - Gestion des données, des autorisations, de l'accounting,.....

# Middleware d'EGEE

- gLite 3.0 /3.1
  - Tourne sous OS SL3/SL4 (Redhat Enterprise)
- Constituants
  - Produits/services développés par le projet EGEE
  - Produits/services développés par d'autres projets
    - Anciens projets tel que EDG
    - Projets actuels tel que OSG
  - Produits/services venant de « l'industrie »
    - Condor
    - LDAP



# Services de base du MW

- Sécurité (authentification) : GSI
  - Nom usuel : « les certifs »
  - Basé sur des certificats
  - Institutionnalisé à travers des groupes de coordination européen, mondial
  - Permet l'identification des individus, des machines et des services
- Sécurité (autorisation): VOMS (Virtual Organisation Membership Service)
  - Nom usuel :VOMS
  - Service transversal à de nombreux services
  - Gestion des autorisations
    - Droits accès aux ressources (calcul, stockage)
    - ACL sur les data
    - Utilisations des services
    - ...
  - Service utilisé par toutes les grilles (OSG, Nordugrid,...)
    - Permet une vue cohérente des VOs dans tous les grilles

# Services de base du MW

- **Système d'information: Publication des ressources et informations diverses (statiques et dynamiques)**
  - Nom usuel : BDII
  - Mise à jour en permanence, fréquence de l'ordre de 5 mn
  - Basé sur LDAP et un schéma d'information (GlueSchema)
  - Utilisé principalement par les services eux même
    - Interrogeable par l'utilisateur
  - Organisé de façons hiérarchiques
- **Interface utilisateur**
  - Nom usuel : UI
  - C'est le client de la grille, c'est pas une GUI
  - Ensemble d'API (bibliothèques, modules Perl/Python...)
    - Permet l'interfaçage des services de grille au niveau des applications
  - Outils « ligne de commande » pour les fonctions les plus courantes

# Services de computing du MW

- 2 composants principaux , 2 générations : La seconde génération arrive en ce moment.
- Resource Broker : meta-scheduler faissant la correspondance entre requirements et ressources
  - Nom usuel : RB (génération 1) , gLite WMS (génération 2)
  - Sélectionne un site (CE) en s'appuyant sur le système d'Information qui connaît l'état réel des ressources
  - Intègre des contraintes sur la localisation des données
  - Fonctionnalités avancées : bulk submission, Job paramétrique, DAG... (génération 2)
  - Basé sur Condor-G

# Services de computing du MW

- Computing Element : interface avec les jobmanager locaux des sites
  - Nom Usuel : CE
  - S'interface avec Condor, LSF, Torque/MAUI, SGE...
- Autres services
  - RB Light: permettant de faibles latences de soumission
  - Moteurs de workflow
  - Outils de soumission
    - Développés et utilisés par des expériences
    - S'interfacent principalement avec le RB
    - GANGA

# Services de data management du MW

- 2 composants principaux
- File Catalog : Catalogue de nommage des data
  - Nom usuel : LFC,.....
  - Non distribué, éventuellement répliquable (Oracle)
  - Peut associer plusieurs replicas à un nom logique
    - Permet le déplacement transparent des données physiques
- Storage Element :
  - Nom usuel : SE
  - Unité de stockage de la grille
  - Basé sur SRM (Storage Resource Manager)
  - Permet de répartir localement les fichiers sur plusieurs serveurs
  - 2 principaux protocoles d'accès : gsiftp et rfio
    - Gsiftp : transfert intégral du fichier ( c'est du ftp sécurisé)
    - Rfio : accès direct au contenu du fichier, API Posix like

# Services de data management du MW

- Autres services
  - FTS : File Transfer Service
    - Service permettant de gérer en batch les transferts
    - Basé sur SRM
    - Notion de canaux (channel) entre sites/SE
    - Optimise l'utilisation des ressources :
      - Nombre de transferts concurrents...
      - Gestion des reprises en cas d'erreur
  - Service de management de meta data
  - Service d'encryptions au vol des data
    - Utilisé principalement par Biomed

# Autres services du MW

- Accounting
  - Collecte d'informations auprès de Computing Element
  - Publications des consommations par nœud de grille
  - Très utile pour tuner les politiques d'ordonnancement des tâches sur la grille
- RGMA
  - Service d'information parallèle au BDII

# Question