



Initiation à l'infrastructure

David Bouvet, David Weissenbach
Université Lille 1, 18-20/05/2011

- **Le Middleware/Intergiciel d'EGI (European Grid Infrastructure)**
- **Infrastructure opérationnelle de EGI**
- **Infrastructure pour la formation**

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

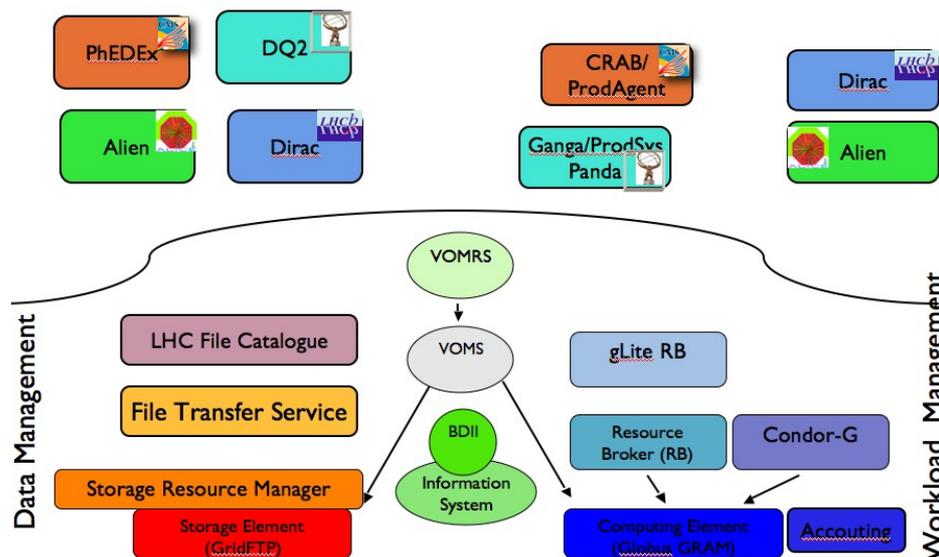
➤ gLite 3.1/3.2

- Tourne sous OS SL4/SL5 (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



- **Sécurité (authentification) : GSI**
 - 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 d'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 VO dans tous les grilles

➤ **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 min
- 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, ce n'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

- **Workload Management System (WMS) : *meta-scheduler* faisant la correspondance entre besoins et ressources**
 - Nom usuel : 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 : soumission par lot, job paramétrique, DAG... (génération 2)
 - Basé sur Condor-G

- **Computing Element : interface avec les *jobmanagers* locaux des sites**
 - Nom usuel : CE ou CREAMCE
 - S'interface avec Condor, LSF, Torque/MAUI, SGE...

- **Autres services**
 - Moteurs de *workflow*
 - Outils de soumission
 - Développés et utilisés par des expériences
 - S'interfacent principalement avec le WMS

Services de gestion des données du MW

➤ File Catalog : catalogue de nommage des données

- 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 façon Posix

➤ 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 metadonnées
- Service d'encryptage au vol des données
 - Utilisé principalement par Biomed

➤ *Accounting*

- Collecte d'informations auprès des CE
- Publications des consommations par nœud de grille
- Très utile pour régler les politiques d'ordonnancement des tâches sur la grille

➤ **RGMA**

- Service d'information parallèle au BDII



Infrastructure Opérationnelle de EGI

➤ Infrastructure opérationnelle

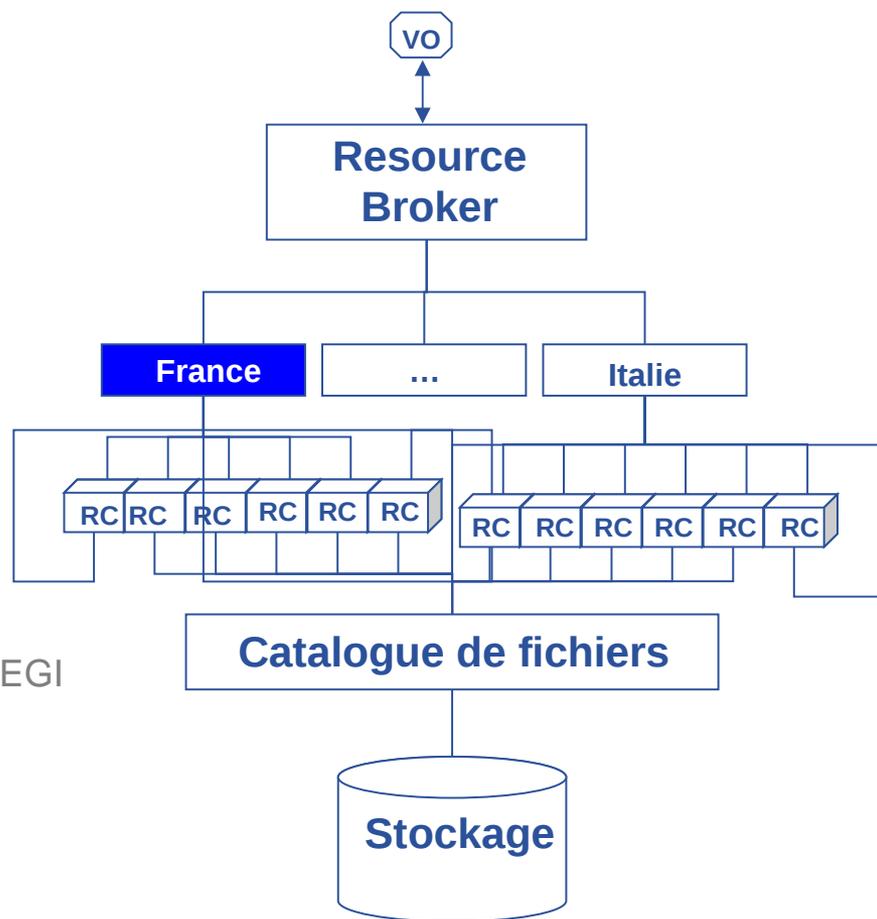
- Activité SA1 d'EGI: déploiement, exploitation et support de la grille
- Basée sur les NGI (National Grid Initiative)
 - France Grilles
 - IGI (Italie)
 - IBERGRID
 - Allemagne
 - NGF (Europe du Nord)
 - Russie
 - Pologne
 - ...

➤ Activités globales

- Définition des procédures d'exploitation
- Administration globale
 - Serveurs VOMS
 - Catalogues de fichiers
 - RB/WMS
 - Système d'information de la grille
- Monitoring
- User Support / Astreinte
- Accounting

➤ Activités régionales

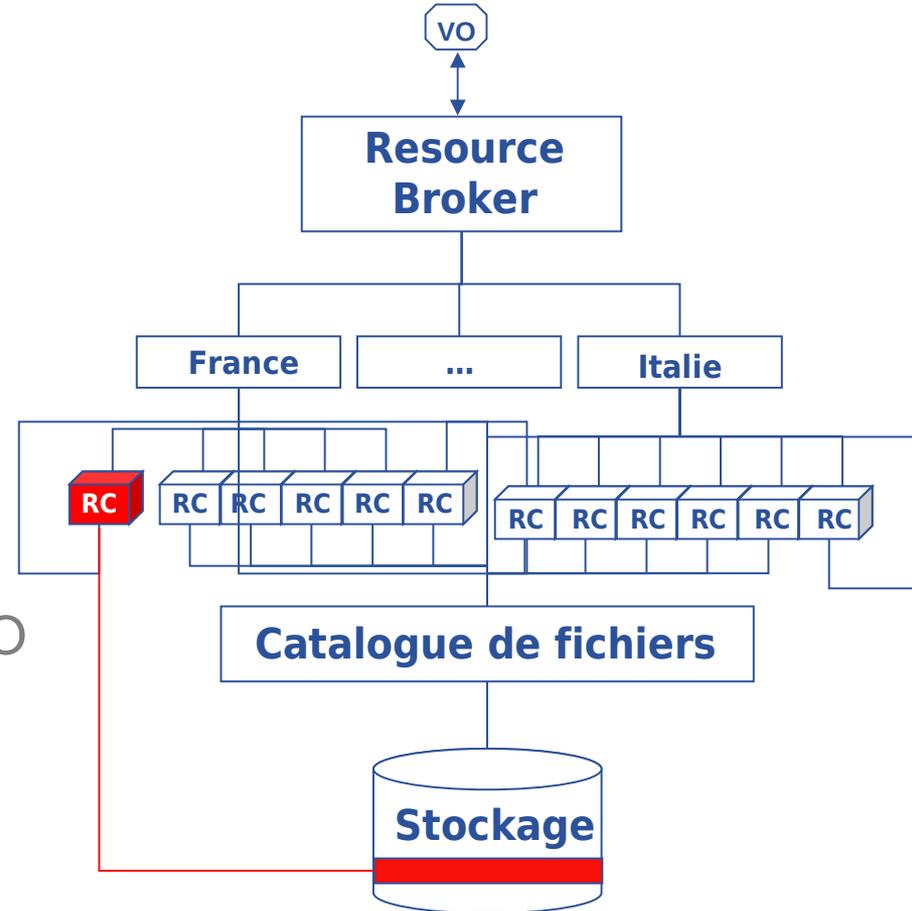
- Représentant des sites d'une région auprès d'EGI
 - Assistance aux sites de la région
 - Canal de communication entre EGI et les sites
- Gestion régionale de la sécurité
- Monitoring régional
- Certification des sites



➤ Resource Centre / Site

fournit à la grille

- Des ressources de calcul
 - Computing Element
- Des ressources de stockage
 - Storage Element
- Eventuellement des services/VO



- **Configuration minimum nécessaire pour une VO**
 - 1 nom de VO unique pour la grille + 1 VO manager
 - 1 serveur de gestion des membres/rôles (VOMS)
 - 1 (au moins) WMS
 - 1 (au moins) User Interface (UI)
 - 1 (au moins) Centre de Ressource (RC)
 - 1 système d'information indexant les ressources (BDII)

- **Options**
 - 1 service de catalogue de fichiers (LFC)
 - 1 service de transfert de fichiers (FTS)

- **Monitoring des sites**
- **Déploiement du *middleware***
- **Remontée des incidents**
- **Récolte des doléances**
 - des sites
 - des VO
- **Diffusion d'informations**
 - mises à jour
 - sécurité
- **Génération de statistiques**
- **Maintient d'une grille fonctionnelle**

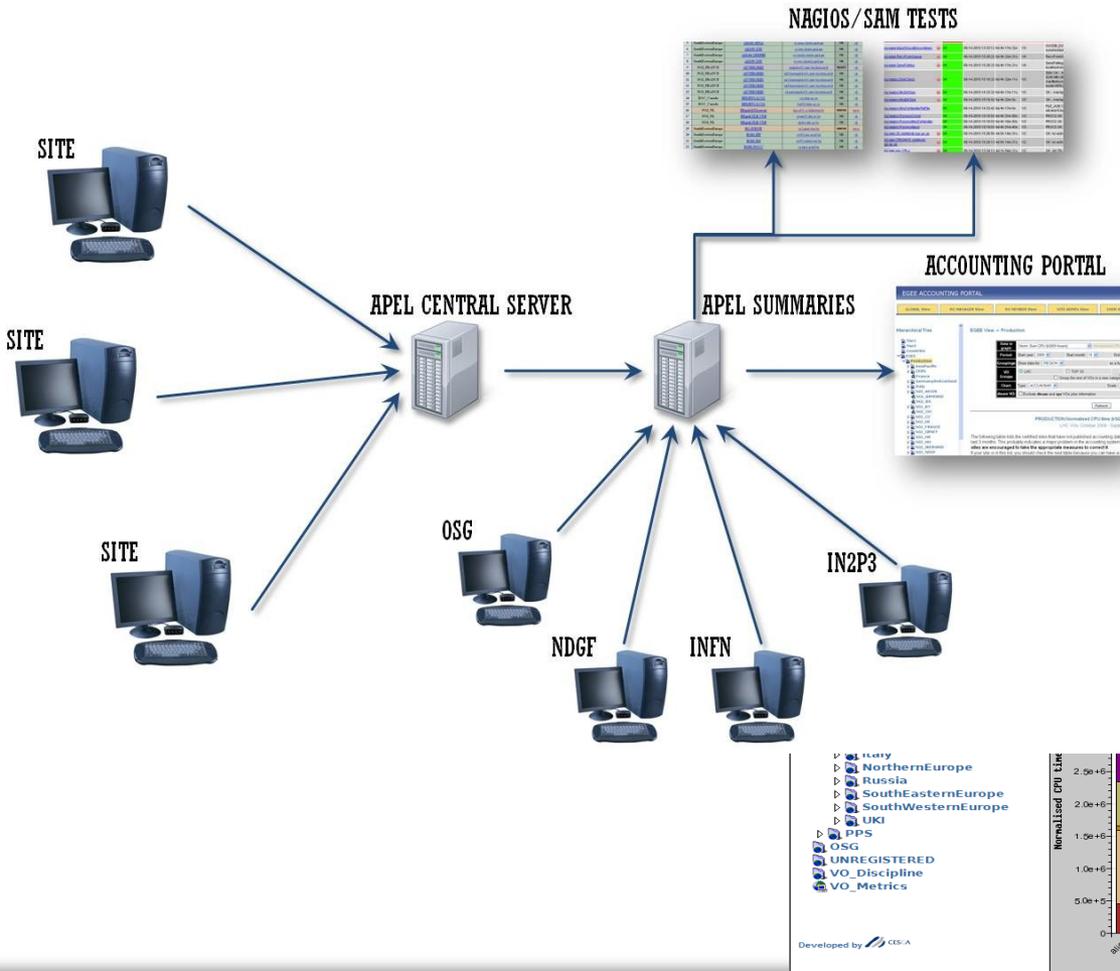
- **Grid Operations Centre core Database (GOCDB)**
 - <https://goc.gridops.org/>
- **GOC monitoring tools**
 - <http://goc.grid-support.ac.uk/gridsite/monitoring>
- **GOC job monitor**
 - <http://gridportal.hep.ph.ic.ac.uk/rtm/>

➤ 1 équipe par NGI d'opérateurs

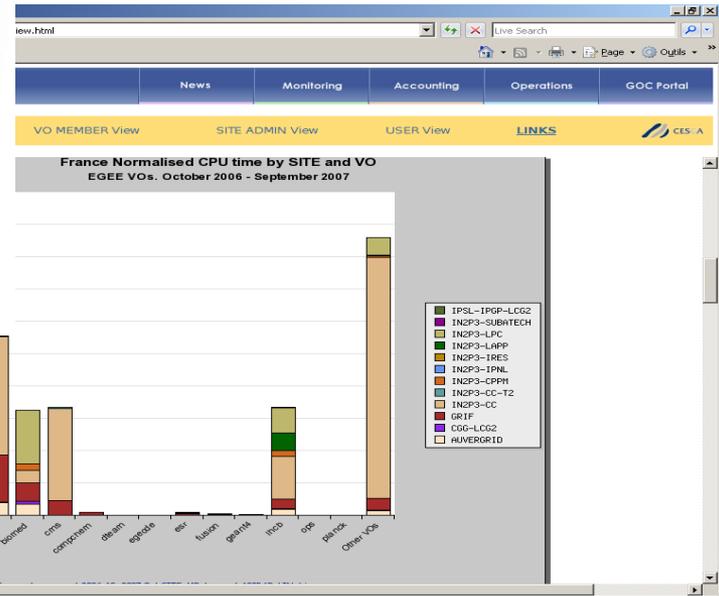
- Surveillance des sites de la NGI

➤ Outils

- Operation Portal
 - <http://cic.gridops.org>
 - Point d'entrée sur les autres outils
 - Vue intégrée des autres outils à l'usage des opérateurs
- Nagios
 - <https://ccnagboxli01.in2p3.fr/nagios/>
- GSTAT
 - <http://www.gstat.org/>



- Basé sur APEL et ActiveMQ
- Déployé sur tous les sites



- Global Grid User Support
- Création de tickets par mail ou formulaire web
- Affectation à des *Support Units*
- Supervision possible de la NGI
- Interfaçage possible au système de tickets de la NGI
- Notifications

<http://www.ggus.org>



Welcome to Global Grid User Support

Tickets @ GGUS

- ▶ [Information](#) on your GGUS account
- ▶ [Submit a new ticket](#) via browser
- ▶ [Submit a new ticket](#) via email

GGUS tickets for Frederic Schaefer

ID	Status	Date	Info
▶ 27071	solved	2007-09-20	RGMA Schema/Registry down ?
▶ 27030	solved	2007-09-19	RGMA data transmission is inefficient
▶ 27017	solved	2007-09-19	RGMA buffers seem to cause wrong RGMA (a...
▶ 23898	solved	2007-06-28	accounting records available in R-GMA, u...
▶ 21217	solved	2007-04-25	accounting Specint history displayed is ...

6 more tickets. ▶ [Show all](#)

Latest open tickets

ID	VO	Info
▶ 27714	cms	cms jobs abort at UKI-LT2-IC-HEP
▶ 27711	other	Yaim guide doesn't document GLITE_USER* variables
▶ 27710	cms	cmssgm jobs aborting at UKI-LT2-QMUL
▶ 27708	none	MyProxy failure on skurut3.cesnet.cz (prague_cesne...
▶ 27707	none	SE failure on se.bfg.uni-freiburg.de (UNI-FREIBURG...
▶ 27706	none	SE failure on castorsc.grid.sinica.edu.tw (Taiwan-...
▶ 27705	dteam	no permission to write into path
▶ 27704	other	glite-WMS requires boost and boost-devel
▶ 27703	none	MyProxy failure on dg15.cc.kek.jp (JP-KEK-CRC-01)
▶ 27701	atlas	ATLAS production job are failing in lcgce.ijs.si -...
▶ 27700	cms	CMS production jobs failing at TW-FTT
▶ 27695	none	CERN APEL accounting descrapencies
▶ 27694	none	RGMA failure on lcg-se.rcf.uvic.ca (VICTORIA-LCG2)
▶ 27692	none	SRM failure on lcg-se.rcf.uvic.ca (VICTORIA-LCG2)
▶ 27688	none	CE failure on cmsrm-ce01.roma1.infn.it (INFN-ROMA1...

▶ [Show all open tickets](#)

▶ [Show open ENOC tickets](#)

▶ [Search ticket database](#)



Infrastructure pour la formation

- Formation gérée par les NGI
- En France par France Grilles
 - <http://formation.france-grilles.fr/>
 - Création d'une VO formation : vo.formation.idgrilles.fr

- **Projet EGEE : <http://www.eu-egee.org/>**
- **LCG**
 - GOC : <http://goc.grid-support.ac.uk/>
 - Projet LCG : <http://lcg.web.cern.ch/LCG/>
- **Site de vulgarisation de la grille :**
<http://gridcafe.web.cern.ch/gridcafe/>