

# Le projet EGEE

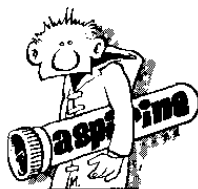
(*Enabling Grids for E-science in Europe*)

Août 2004

---

|            |                                                              |           |
|------------|--------------------------------------------------------------|-----------|
| <b>I</b>   | <b><u>Présentation générale d'EGEE</u></b>                   | <b>2</b>  |
| 1          | Les Grilles                                                  | 2         |
| 2          | Organisation du projet EGEE                                  | 3         |
| 3          | Accès sécurisé aux ressources                                | 8         |
| 4          | Composants de la Grille                                      | 9         |
| 5          | Le système d'information : R-GMA                             | 10        |
| 6          | Gestion des données                                          | 11        |
| 7          | Soumission d'un job                                          | 13        |
| <br>       |                                                              |           |
| <b>II</b>  | <b><u>Manipulations sur la Grille</u></b>                    | <b>16</b> |
| 8          | Obtention d'un certificat                                    | 16        |
| 9          | Soumission d'un job simple                                   | 16        |
| 10         | Interrogation du système d'information R-GMA                 | 18        |
| 11         | Manipulation des fichiers : <code>edg-replica-manager</code> | 19        |
| <br>       |                                                              |           |
| <b>III</b> | <b><u>Appendices</u></b>                                     | <b>22</b> |
| 12         | Sigles et liens                                              | 22        |
| 13         | Autres liens utiles                                          | 25        |

---



La plupart des sigles utilisés sont listés dans la section **12**. Par ailleurs, un **dictionnaire** des termes utilisés dans le *Grid Computing* peut-être consulté sur le site du CNAF.

Les termes apparaissant en bleu sont des liens hypertextes. Pour pouvoir les suivre, récupérer la version PDF sur <http://www.ipgp.jussieu.fr/~moguilyn/GRILLE/EGEE.pdf>, ou regarder les adresses dans les sections **12** et **13**. Les termes apparaissant en rouge sont des liens internes au document.

## Première partie

# Présentation générale d'EGEE

---

## 1 Les Grilles

Étant donné la demande croissante des scientifiques en ressources informatiques (calcul et stockage), l'idée de Grille de Calcul et de Grille de Données est apparue, de manière à mutualiser des ressources réparties, afin de pouvoir :

- accéder à de la puissance de calcul,
- accéder à des données géographiquement dispersées,
- bénéficier d'une continuité de service (si un nœud de calcul est en panne, le travail peut s'exécuter sur un autre),

et cela, de n'importe où.

La situation idéale étant de pouvoir utiliser ces ressources sans se soucier de l'endroit où elles se trouvent, de nombreux projets ont vu le jour afin de développer du *middleware*, c'est-à-dire le logiciel nécessaire pour interfacer les ressources disponibles d'un côté, et les applications de l'autre, de la manière la plus transparente possible pour les utilisateurs.

Il s'agit donc de mettre en place une infrastructure matérielle (ou d'utiliser celle existante) et logicielle permettant de partager, de manière sûre, des ressources hétérogènes, géographiquement distribuées, et gérées de façon décentralisée, tout en prenant en compte la diversité des besoins (puissance de calcul, réseaux rapides, stockage, logiciels...).

Pour mettre en application ce concept de Grille, de nombreux projets ont vu le jour, et en Europe, beaucoup sont partiellement financés par la Commission Européenne.

**GridStart** est une initiative de la Commission Européenne, créée pour consolider les avancées techniques en Europe, faciliter les interactions entre des projets similaires en Europe et dans le reste du monde, et encourager la participation de l'Industrie et la recherche des applications portables sur les Grilles. La page Web de GridStart comporte, entre autres, un **portail** sur de nombreux projets de grille. Ceux participant directement à GridStart sont :

- **CrossGrid** pour la construction d'une grille à grande échelle pour les visualisations et les simulations nécessitant du temps réel ;
- **DAMIEN** qui développe une architecture de *middleware* pour une supervision générique des systèmes ;
- **DataTAG** pour la mise en place d'une interconnexion haut débit entre des grilles européennes et américaines ;
- **EGSO** pour la création d'un observatoire solaire virtuel ;
- **GRIA** (*Grid Resources for Industrial Applications*) pour la mise à disposition de ressources pour l'Industrie ;
- **GridLab** qui développe du logiciel pour permettre aux applications génériques de s'adapter aux changements d'environnement Grille et d'exploiter pleinement des ressources dynamiques ;

- **GRIP** (*GRid Interoperability Project*) dont le but est de permettre l'interopérabilité entre **GLOBUS** et **UNICORE**, qui sont à la base de la plupart des *middlewares* Grille ;
- **DataGrid** ou EDG (*European DataGrid*) pour le développement d'une grille pour les scientifiques nécessitant du calcul intensif et le partage de grandes bases de données réparties. L'infrastructure utilisée par le projet EDG qui s'est terminé en mars 2004, et le *middleware* développé (EDG 2.x) serviront de base au projet EGEE.

Sont aussi impliqués dans GridStart, les projets **bioGRID**, **COG** (*Corporate Ontology Grid*), Flow-Grid, **GEMSS** (*Grid-Enabled Medical Simulation Services*), **GRACE** (*GRid seArch & Categorization Engine*), **GRASP** (*Grid Based Application Service Provision*), **MammoGrid**, **MOSES** (*MODular and Scalable Environment for the Semantic web*), **OpenModGRID** (*Open Computing GRID for Molecular Science and Engineering*), et **SeLeNe** (*Self e-Learning Networks*).

Par ailleurs, l'**IDRIS**, après avoir participé au projet **EuroGrid** qui s'est terminé en janvier dernier, s'implique maintenant dans deux nouveaux projets :

- **DEISA** (*Distributed European Infrastructure for Supercomputing Applications*), qui est un consortium de centres nationaux de calcul en Europe qui se propose de déployer et d'exploiter une infrastructure de calcul distribué de très haute performance (pour lequel l'**IDRIS** est le coordinateur),
- **HPC-Europa** dont l'objectif est la mise en œuvre d'une activité d'accès trans-national au calcul de haute performance autour de centres nationaux d'informatique scientifique.

## 2 Organisation du projet EGEE

Le projet **EGEE** (*Enabling Grids for E-science in Europe*) vise à intégrer les efforts nationaux, régionaux et thématiques dans une infrastructure pour la Recherche Européenne. Cette infrastructure sera construite sur le réseau **GEANT** (voir figure 1 page 4), qui relie les réseaux académiques nationaux ou NRENs (*National Research and Educational Network*). Actuellement, 33 NRENs sont impliqués, avec des débits variant de 34 Mo/s à 10 Gb/s. EGEE exploitera l'expertise acquise par d'autres projets Grille comme **EDG** et devra assurer l'interopérabilité avec d'autres Grilles comme **Cyberinfrastructure**, de la NSF américaine, afin d'établir une Grille mondiale.

### 2.1 Partenaires

L'union européenne contribue pour 32 M€ au projet EGEE qui a démarré le 1<sup>er</sup> avril 2004, et qui est piloté par le **CERN** (Centre Européen de la Recherche Nucléaire). 160 chercheurs sont déjà impliqués et plusieurs postes en CDD ont été ouverts.

EGEE est un projet de 2 ans dans le cadre d'un programme européen de 4 ans. Des applications pilotes ont été choisies pour guider l'implémentation et certifier la performance et la fonctionnalité de cette Grille :

- le **LCG** (*Large Hadron Collider Computing Grid project*) pour la physique des particules, qui est une collaboration internationale regroupant les expériences du LHC, des centres de calcul régionaux et des Instituts de physique. Le collisionneur LHC, dont la mise en service est prévu au CERN en 2007, devrait produire 5 PetaOctets<sup>1</sup> de données par an ;

---

<sup>1</sup> 1 Peta = 10<sup>15</sup>

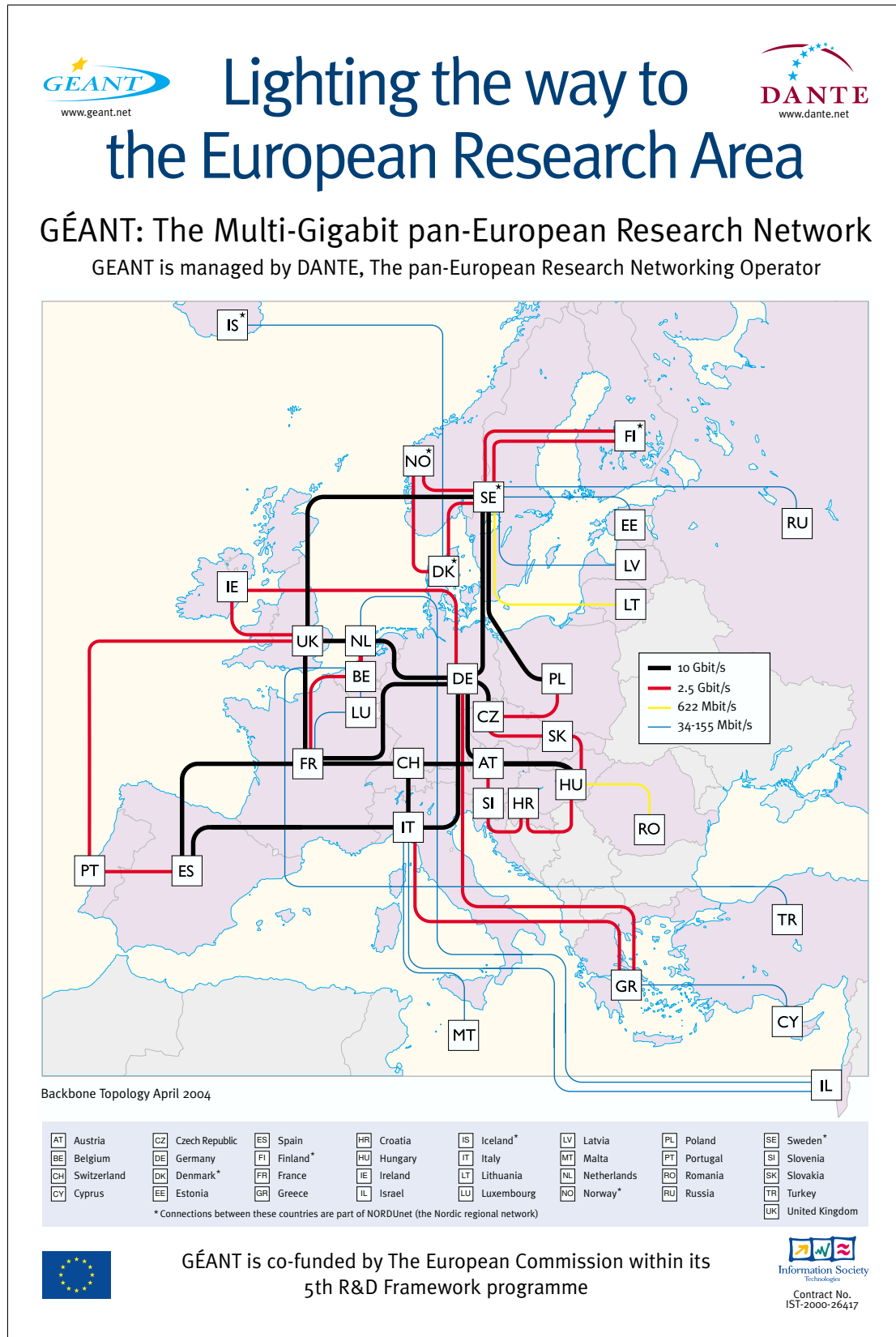


FIG. 1 – Topologie du réseau **GEANT** en avril 2004 (source : DANTE).

- le biomédical (traitement d'images médicales, bioinformatique, base de données mammographiques, aide à la décision clinique...);
  - les applications « génériques », coordonnées par Roberto BARBERA, parmi lesquelles
    - celles développant des interfaces à la Grille comme
      - **GENIUS** (*Grid Enabled web eNvironment for site Independent User job Submission*) : portail Web,
      - **GILDA** (*Grid INFN Laboratory for Dissemination Activities*) : outil de dissémination permettant de tester la Grille ;
    - d'autres spécifiques, sélectionnées par l'EGAAP (*Generic Application Advisory Panel*), qui alloue des forces de travail pour une période donnée. À la dernière réunion de l'EGAAP qui a eu lieu le 14 juin 2004 au CERN, quatre communautés ont été sélectionnées :
      - Astrophysique (traitement des données du télescope MAGIC),
      - Chimie statistique,
      - SIMDAT (applications industrielles),
      -  **Sciences de la Terre**, coordonnées par Monique PETITDIDIER de l'IPSL (Institut Pierre Simon Laplace). L'IPSL assure une mise en commun de compétences concernant les sciences de l'environnement terrestre et planétaire.
- Les activités du groupe **Sciences de la Terre** sont structurées en thématiques :
- observation de la terre par satellite (**ESA**, **KNMI**, IPSL, **UTV**),
  - climat (**DKRZ**, IPSL),
  - **physique du globe (IPGP)**,
  - hydrologie (**Université de Neuchâtel**),
  - géophysique (**CGG**).

La première phase du projet EGEE, dont la fin est prévue en 2006, devrait fournir une plate-forme de production pour la Recherche Européenne, la 2<sup>e</sup> phase devrait permettre d'offrir une couverture géographique plus large et impliquer plus de communautés d'utilisateurs. Le projet prévoit 5000 utilisateurs (dont 3000 après 2 ans), dans au moins 5 disciplines. Avant fin 2004, EGEE devrait regrouper 4000 CPU sur une trentaine de sites.

Les partenaires européens EGEE (environ 70 institutions dans 27 pays) sont répartis dans 10 « régions » ou **fédérations** : CERN, France, Italie, Irlande + Grande Bretagne, Allemagne + Suisse, Europe du Nord (Belgique + Scandinavie + Pays-Bas), Europe du Sud-Ouest (Espagne + Portugal), Europe du Sud-Est (Bulgarie, Chypre, Grèce, Israël, Roumanie), Europe Centrale (Autriche, Hongrie, Pologne, République Tchèque, Slovaquie), et Russie.

Au niveau français, sont impliqués :

- le **CEA/DSM** (Commissariat à l'Énergie Atomique, Direction des Sciences de la Matière),
- la **CGG**,
- le CNRS, et en particulier l'**IN2P3** (Institut National de Physique Nucléaire et de Physique des Particules), qui est un institut du CNRS dont la mission est de promouvoir et de fédérer les activités de recherche dans ces domaines de la physique,
- la **CS SI** (Compagnie des Signaux, Systèmes d'Information),
- la **CRSA** (Centrale Recherche S.A.), émanation de l'**École Centrale de Paris**.

Les États-Unis, le Japon et l'Asie Pacifique devraient aussi rejoindre le projet. La liste complète des **partenaires** peut être visualisée sur le site d'EGEE. Les autres communautés d'utilisateurs potentielles entreront en contact avec EGEE grâce aux événements organisés par le groupe de dissémination.

EGEE se servira de ressources nationales et de celles de centres de calcul, qui, en contre partie, pourront exécuter des applications sur de grosses configurations, avoir accès à la compétence de spécialistes, et entrer en contact avec d'autres utilisateurs, avec qui pourront se créer de nouvelles collaborations.

Il est prévu que l'Industrie participe à EGEE de plusieurs manières : en tant que fournisseur de main d'œuvre, en tant qu'utilisateur, et pour assurer la maintenance à long terme des services Grille. Déjà, HP, IBM, GridXpert, DATAMAT et Microsoft font partie de ce forum de l'Industrie d'EGEE.

## 2.2 Structuration du projet EGEE

EGEE est structuré en 3 grandes zones d'activité : les services, la re-ingénierie du *middleware*, le réseau au sens large du terme.

### ❶ Les services

L'activité « services » est en charge de la mise en place et de l'administration d'une plate-forme de production. Elle est organisée en 4 entités de la manière suivante :

- La direction centrale ou OMC (*Operations Management Center*), au CERN, qui assure, entre autres, la coordination centrale, les relations avec la Commission Européenne et les autres Grilles. Ce groupe reçoit les recommandations d'un conseil d'exploitation ou OAG (*Operations Advisory Group*), regroupant des représentants des autres entités de l'activité « services ».
- Les pivots d'infrastructure ou CIC (*Core Infrastructure Center*), qui fournissent l'infrastructure de base pour la Grille et assurent son bon fonctionnement. En particulier, ils doivent :
  - analyser et optimiser les performances,
  - assurer le bon fonctionnement des services de base, comme la répartition des jobs sur la Grille, la gestion des catalogues de données, les services d'autorisation, d'information, la collecte et la publication des informations sur l'utilisation des ressources,
  - fournir du support aux centres régionaux (ou ROC).

Les CIC sont basés principalement en Grande-Bretagne, France, Italie et au CERN.

- Les centres d'exploitation régionaux ou ROC (*Regional Operation Center*), qui coordonnent les ressources des sites et le support :
  - validation, déploiement et installation du *middleware*,
  - coordination des mises à jour,
  - mise à disposition de documentation et de *software*,
  - surveillance de la Grille et remontée des problèmes, maintenance d'une base de données des problèmes rencontrés,
  - gestion des organisations virtuelles (VO) qui représentent les communautés d'utilisateurs.

- Les centres de ressources ou RC (*Regional Center*), qui sont les sites disposant des ressources matérielles proprement dites : puissance de calcul, stockage, connectivité...

Les RC français sont à l'**École Centrale Paris**, à la CGG, et, dans l'IN2P3, au **LAL Orsay**, au **LPC** Clermont- Ferrand, au **CPPM** de Marseille et au **Centre de Calcul** de Lyon (qui est aussi ROC et CIC).

## ② La re-ingénierie du *middleware*

Le *middleware* actuel, ou logiciel d'interfaçage entre les utilisateurs et les ressources de la Grille, ne suit pas forcément les recommandations de l'**OGSA** (*Open Grid Services Architecture*), qui préconise des standards pour la construction de Grilles. Il faut donc modifier les composants actuels du *middleware* pour qu'ils respectent ces standards, et puissent évoluer tout en assurant une interopérabilité avec les autres Grilles. La France est principalement chargée, dans cette activité, de la partie « Assurance Qualité ».

La 1<sup>re</sup> version du *middleware* utilisée sur EGEE sera LCG-2 (logiciel utilisé par les expériences du LHC), basée sur :

- **EDG** : *middleware* écrit par et pour le projet *European DataGrid* qui s'est terminé en mars 2004,
- **VDT** (*Virtual Data Toolkit*) : boîte à outils dont le but est de permettre une installation et une configuration simple du logiciel Grille,
- **AliEn** : ensemble d'outils développé pour l'expérience ALICE (LHC) pour interfacer la Grille.

Par la suite, LCG-2 sera remplacé par EGEE-1, cette dernière étant basée non plus sur Globus 2, mais sur les services Web.

La 1<sup>re</sup> plate-forme supportée est RedHat 7.3.

## ③ Le réseau

Les activités « réseau » ont pour but de faciliter l'intégration de nouveaux utilisateurs, de nouvelles communautés scientifiques et organisations virtuelles (VO), ce qui implique de prendre en charge la dissémination, la formation, l'identification des applications et le support, et la coopération internationale. Afin de permettre aux utilisateurs d'accéder aux réseaux académiques connectés à la Grille, différents modèles de documents sont utilisés :

- les SLR (*Service Level Requirements*)  
où les utilisateurs expriment leurs besoins en ressources réseaux (bande passante, QoS...);
- les SLS (*Service Level Specifications*)  
qui sont une spécification technique des ressources réseaux disponibles ;
- les SLA (*Service Level Agreement*)  
qui définissent les procédures opérationnelles, administratives et légales, et les structures à mettre en place pour permettre aux utilisateurs d'une VO d'accéder à un réseau académique.

## 2.3 Organisation des activités (ou groupes de travail)

**Le réseau** (*Networking Activities*), 28 %

- NA1** Direction du projet (responsable : Fabrizio GAGLIARDI).
- NA2** Dissémination (resp. John DYER) :  
basée sur le Web, organisation de conférences, publications.
- NA3** Formation des utilisateurs (resp. Malcolm ATKINSON) :  
production de matériel et organisation de cours.



- NA4** Identification des applications et support (resp. Vincent BRETON) : support pour les applications de la Physique des Hautes Énergies et bio, définition d'interfaces communes, création d'un forum de l'Industrie, intégration de nouvelles communautés. Ces nouvelles communautés candidates sont évaluées par l'EGAAP, une espèce de conseil scientifique.
- NA5** Coopération internationale (resp. Fabrizio GAGLIARDI) : en particulier avec la grille américaine **Cyberinfrastructure**, et d'autres projets européens (voir **GridStart**).

#### Les services (*Specific Service Activities*), 48 %

- SA1** Support européen de la Grille (resp. Ian BIRD) : administration et exploitation : cœur des services d'infrastructure, contrôle et gestion de la Grille, déploiement du *middleware*, support utilisateur.
- SA2** Gestion des ressources réseau (resp. Jean-Paul GAUTIER) : définition des besoins et spécification des services pour la détermination des schémas de réseau, en collaboration avec **GEANT** (et bientôt son successeur **GN2**) et les réseaux académiques nationaux. Ce groupe de travail, ainsi que le JRA4 (voir plus loin), est sous la responsabilité de l'Unité Réseaux du CNRS (**UREC**).

#### Activités de recherche (*Research Activities*), 24 %

- JRA1** Re-ingénierie du *middleware* et intégration des services Web (resp. Frederic HEMMER) : mise à jour du *middleware* existant, ajout des fonctionnalités manquantes, test et validation.
- JRA2** Assurance qualité (resp. Gabriel ZAQUINE) : vérification de la conformité aux besoins d'EGEE.
- JRA3** Sécurité (resp. Åke EDLUND) : développement d'une architecture sécurisée.
- JRA4** Développement des services réseau (resp. Javier ORELLANA) : cette activité devrait fournir les interfaces entre le *middleware* Grille et les utilisateurs pour effectuer des réservations de bande passante, fournir des outils de diagnostics et de surveillance des performances réseau, et étudier l'impact d'IPv6.

### 3 Accès sécurisé aux ressources

L'accès à l'infrastructure EGEE est assuré par l'utilisation de certificats numériques, associés à des mots de passe. Les détails peuvent être trouvés sur le Web, à l'adresse <http://igc.services.cnrs.fr/Datagrid-fr/>.

Parmi les autorités de certification (CA), *Datagrid-fr* délivre les certificats pour le CNRS, et une vingtaine de laboratoires y sont inscrits.

Par ailleurs, l'**EUGridPMA** (*European Policy Management Authority for Grid Authentication e-Science*) a été créée pour coordonner l'activité des différents CA pour les projets européens, les projets nationaux des pays européens et autres projets associés comme le LCG.

Enfin, chaque laboratoire doit être rattaché à une des VO (*Virtual Organisation*) existantes. Une VO regroupe des utilisateurs partageant des besoins similaires et des ressources communes dans leurs applications, et chaque VO maintient un annuaire **LDAP** des utilisateurs qui y sont rattachés.



Avec la version 2 du logiciel EDG, il est possible de faire partie de plus d'une VO.

Les droits d'accès sont liés au sujet du certificat. L'accès est contrôlé par un *proxy*<sup>2</sup>, de validité limitée dans le temps, signé par la clé privée de l'utilisateur.

Deux types de *proxy* peuvent être générés :

1. un *proxy* Globus qui vérifie seulement l'identité de l'utilisateur,
2. un *proxy* **VOMS** (*VO Membership Service*) qui contient de l'info sur des autorisations supplémentaires au sein des VO, de groupes, ou de rôles dans ces organisations. La méthode VOMS ne peut être utilisée que si la VO de l'utilisateur possède un serveur VOMS.

Pour pouvoir lancer des applications sur la Grille, un utilisateur devra tout d'abord obtenir un certificat, s'inscrire sur une VO et avoir un compte sur une machine ayant une interface utilisateur (UI) sur la Grille.

Quelques indications pour l'obtention d'un certificat sont données dans la section 8.

## 4 Composants de la Grille

L'utilisateur interagit avec la Grille via une interface utilisateur (UI) lui permettant :

- d'interroger le système d'information pour connaître les ressources disponibles (section 5) ;
- de manipuler des fichiers : les déposer sur des unités de stockage (SE), les enregistrer sur des catalogues, les dupliquer... (voir section 6) ;
- de soumettre des jobs (voir section 7).

Les sites mettant à disposition des ressources possèdent un ou plusieurs des éléments suivants :

- **UI** (*User Interface*) : Interface utilisateur  
Machine depuis laquelle un utilisateur soumet un job.
- **CE** (*Computing Element*) : Élément de Calcul comprenant :
  - un **GK** (*GateKeeper*) : frontal qui assure la correspondance entre un utilisateur de la Grille et un compte disponible sur le CE pour la VO de l'utilisateur ;
  - un ou plusieurs **WN** (*Worker Nodes*) : nœuds effectifs de calcul, derrière un système de batch (BQS, PBS, LSF).
- **SE** (*Storage Element*) : Élément de stockage  
Il s'agit d'un serveur de disques (éventuellement de cassettes, comme celles gérées par le système de stockage hiérarchique Castor du CERN).
- **MON** (*Monitoring Node*) : Nœud de contrôle  
Machine chargée de faire tourner les *servlets*<sup>3</sup> du système d'information R-GMA (voir section 5) pour le site, et le service d'optimisation des duplications de fichiers (ROS : *Replica Optimization Service*).

<sup>2</sup>ici, enregistrement "caché" servant de passerelle sécurisée entre l'utilisateur et la Grille

<sup>3</sup>*servlet* : programme Java qui réside et s'exécute sur un serveur

Une partie de ces éléments peut-être visualisée sur la figure 3 page 14. D'autres services globaux, non spécifiques à un site, sont nécessaires au fonctionnement de la Grille, parmi lesquels :

- **RB** (*Resource Broker*) : Courtier de ressources, machine qui recherche sur la Grille l'élément de calcul le plus adéquat pour l'exécution du job, en prenant en compte la proximité des données éventuellement requises.
- **RLS** (*Replica Location Service*) : Service assurant la maintenance des catalogues contenant la localisation des fichiers originaux et de leurs copies, sur un ou plusieurs éléments de stockage (SE), pour une VO donnée (voir section 6).
- **IC** (*Information Catalog*) : Service d'information (R-GMA) qui délivre des informations sur l'état de la Grille : infrastructure et applications (voir section 5), principalement utilisé par le RB.
- **LB** (*Logging & Bookkeeping*) : Service de journalisation, construit sur une base de données MySQL, qui récupère les événements relatifs à la vie d'un job (soumission, CE d'exécution...) et les rend accessibles aux utilisateurs.

## 5 Le système d'information : R-GMA

Le but du système d'information de la Grille est de permettre la gestion et la surveillance des ressources distribuées. Il est utilisé pour trouver les services disponibles à un moment donné, et pour contrôler les applications. Sur la dernière version du logiciel EDG repris par EGEE, son architecture est basée sur le GMA (*Grid Monitoring Architecture*), défini par le **GGF** (*Global Grid Forum*) et il utilise les *servlets*.

Les composants de GMA sont :

- les consommateurs,
- les producteurs,
- le service d'enregistrement (actuellement un par VO).

Les producteurs d'info s'enregistrent de manière à ce que le service d'enregistrement puisse publier le type et la structure des données qu'ils veulent rendre disponible sur la Grille. L'information fournie suit le **schéma GLUE** (*Grid Laboratory Uniform Environment*). Ce modèle d'information, qui décrit les ressources en calcul et en stockage, a déjà été adopté par plusieurs autres projets Grille et permet donc l'interopérabilité entre ces différentes Grilles.

Les consommateurs peuvent interroger le service d'enregistrement pour savoir quel type d'information est disponible et quels sont les producteurs qui peuvent la fournir. Ensuite, le consommateur peut interroger directement le producteur.

L'architecture supporte aussi de pouvoir rassembler l'information fournie par divers producteurs pour rendre une info "digérée" aux consommateurs.

Le R de **R-GMA** signifie qu'il s'agit d'un GMA relationnel, qui donne l'impression à l'utilisateur que toute l'info est contenue dans une base de données unique. Ce modèle, supportant **SQL**, fournit aussi bien de l'info statique que dynamique, sur les ressources de la Grille et les applications.

Les utilisateurs de R-GMA peuvent être de 3 types :

1. les indirects, qui obtiennent l'info via d'autres outils interfacés avec R-GMA (comme PULSE qui est un outil graphique de visualisation),

2. ceux qui utilisent les outils R-GMA pour examiner directement les données (commande `edg-rgma`),
3. les programmeurs qui écrivent du code (en Java, C, C++, Perl ou Python) pour interfacer directement R-GMA.

Dans la version précédente du logiciel EDG, le système d'information était basé sur le **MDS** (*Monitoring and Discovery Services*) de **GLOBUS**, qui rassemblait et publiait les informations via des annuaires LDAP. Afin de garder la compatibilité avec MDS, l'outil **GIN** permet d'interroger les producteurs d'info à la manière de MDS et de les publier sous R-GMA, et l'outil **GOUT** permet, lui, de republier l'info de R-GMA à un serveur LDAP.

Les informations disponibles par R-GMA peuvent être consultées par un navigateur, par exemple, à l'adresse : <http://gppic05.gridpp.rl.ac.uk:8080/R-GMA/index.html>.

Des exemples de requêtes à R-GMA sont présentés dans la section **10**.

## 6 Gestion des données

### 6.1 Gestion des fichiers

Lors de la soumission de jobs utilisant peu de données, les fichiers d'entrée nécessaires peuvent être envoyés avec l'exécutable, via l'attribut `InputSandbox` d'un script JDL (voir section **7**). En revanche, lorsque ces fichiers sont volumineux (plusieurs MégaOctets) et répartis sur la Grille, il est préférable que l'application s'exécute sur un nœud d'exécution (CE) proche d'un élément de stockage (SE) où se trouve une copie des données.

Pour assurer cette proximité CPU/données, il existe des outils permettant de déposer des fichiers sur la Grille, de les dupliquer sur différents SE, de connaître leurs localisations, de les enregistrer sur des catalogues, de publier ces catalogues pour qu'ils soient connus de tous les utilisateurs appartenant à une même VO.

Ces opérations de manipulation des données et des catalogues sur la Grille sont assurées par le **Replica Manager** qui intègre tous les outils permettant de gérer les données sur la Grille, et assure un point d'entrée unique pour les autres utilisateurs et services de la Grille.

Les fichiers déposés peuvent être identifiés de différentes manières :

- **SFN** (*Storage File Name*) : identifiant de fichier physique. Exemple :  
`host1.cern.ch/directory1/directory2/filename`
- **SURL** (*Storage URL*), appelé aussi **PFN** (*Physical File Name*) : SFN reconnu par un *Storage Resource Manager* (SRM). Exemple :  
`srm://host1.cern.ch/directory1/directory2/filename`
- **GUID** (*Grid Unique Identifier*) : nombre de 128 bits, unique et immuable, généré par le RMS pour un SURL. Exemple :  
`guid:73e16e74-26b0-11d7-b1e0-c5c68d88236a`
- **LFN** (*Logical File Name*) : alias d'un GUID, défini par l'utilisateur. Exemple :  
`lfn:mydata`

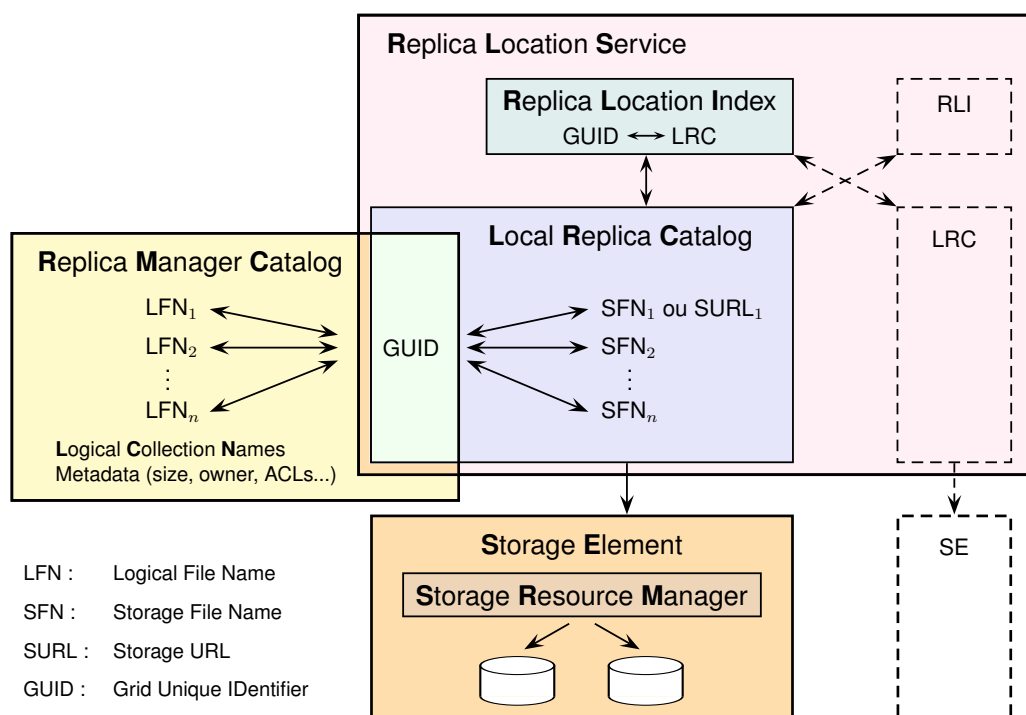


FIG. 2 – EDG Replica Manager.

Le *Replica Manager*, dont les composants et leurs interactions peuvent être visualisés sur la figure 2, est donc composé des éléments suivants :

- le **RLS** (*Replica Location Service*)  
composé lui-même du :
  - **RLI** (*Replica Location Index*)  
index qui assure la correspondance entre les GUID et les LRC ;
  - **LRC** (*Local Replica Catalog*)  
qui stocke la correspondance entre le GUID et les SURL, ainsi que les metadonnées concernant les copies physiques et leur replicas pour une VO donnée.
- le **RMC** (*Replica Metadata Catalog*)  
Stocke la correspondance entre les GUID et les LFN (alias) ainsi que les metadonnées sur les GUID (taille, propriétaire, ACL...) pour tous les fichiers d'une VO donnée. Le RMC stocke aussi les infos sur les LCN (*Logical Collection Name*), un LCN étant un label qui identifie un ensemble de fichiers.

Les fichiers conservés sur une unité de stockage (SE) peuvent être accédés depuis les applications avec différents protocoles :

**gridftp** : protocole de transfert élémentaire de **Globus** contenant une authentification par certificat sur le FTP standard, et quelques possibilités d'optimisation des transferts. `gridftp` est disponible sur tous les sites disposant d'une connexion IP, mais requiert la copie explicite des fichiers (via la commande `globus-url-copy`) sur le nœud effectif de calcul (WN) avant utilisation.

**file** : pour l'accès aux fichiers "proches" de l'élément de calcul (`CloseCE`), lorsque le système de fichiers du SE est distribué (NFS ou AFS). L'accès `file` est transparent et ne demande pas de

copie, mais il ne peut être utilisé que si les fichiers sont montés sur le CE (point d'entrée du *job manager*), et sur le nœud effectif de calcul (WN) où le job s'exécutera (voir section 4).

**rftio** (*Remote File I/O*) : protocole implémenté dans une librairie du CERN utilisée principalement pour accéder aux fichiers stockés sur cassette (système de stockage hiérarchique Castor).

Lorsqu'une application requiert un fichier (LFN ou GUID...), le RLS est contacté pour connaître le SURL d'une de ces instances physiques. Le SRM retourne alors un TURL (*Transport URL*) contenant l'info nécessaire pour que l'application puisse accéder au fichier physique (protocole, *host*, port, chemin d'accès). Par exemple :

```
gsiftp://host1.cern.ch/directory1/directory2/filename
```

L'optimisation de l'accès aux données est faite par le ROS (*Replica Optimization Service*) dont la fonction principale est d'évaluer le coût de l'accès à un ensemble de fichiers sur la Grille. Ce service est utilisé entre autres par le *Resource Broker* lors de son choix du CE d'exécution.

Par ailleurs, les utilisateurs peuvent s'abonner au RSH (*Replica Storage Handler*), par l'intermédiaire du service RSS (*Replica Subscription Service*) pour automatiser la duplication de leurs données sur différents SRM.

Des exemples de manipulation de fichiers sont donnés dans la section 11.

## 6.2 Gestion des méta-données

Certains services doivent maintenir des méta-données persistantes pour des bases de données relationnelles (SGBDR ou *RDBMS*) déportées.

Comme les SGBDR existants ne sont pas accessibles directement par la Grille, et ne peuvent interagir entre eux directement, **Spitfire** est utilisé pour stocker et récupérer les méta-données.

Spitfire est donc un *middleware* qui fournit une interface uniforme, un modèle de données, et un protocole d'accès sécurisé. Il est composé d'un serveur primaire, et de bibliothèques pour les clients.

Les applications dont l'édition de liens a été faite avec la bibliothèque de Spitfire client, communiquent avec une instance du serveur. Ce serveur est mis devant un SGBDR (MySQL ou Oracle) et assure un accès sécurisé à la base de données en question.

Le serveur est implémenté en Java (avec Apache Axis sous Tomcat), et la bibliothèque cliente permet d'appeler les services Web au travers de fichiers WSDL (*Web Services Description Language*) à partir de Java, C et C++.

Il existe également une interface Web indépendante qui :

- coté client permet d'accéder à la base de données,
- coté serveur, permet d'interroger et de manipuler la base. Une interface d'administration permet de définir des règles d'accès pour les utilisateurs et les groupes.

## 7 Soumission d'un job

Les utilisateurs désirant utiliser la Grille, doivent, après avoir récupéré un certificat, initier un *proxy*.

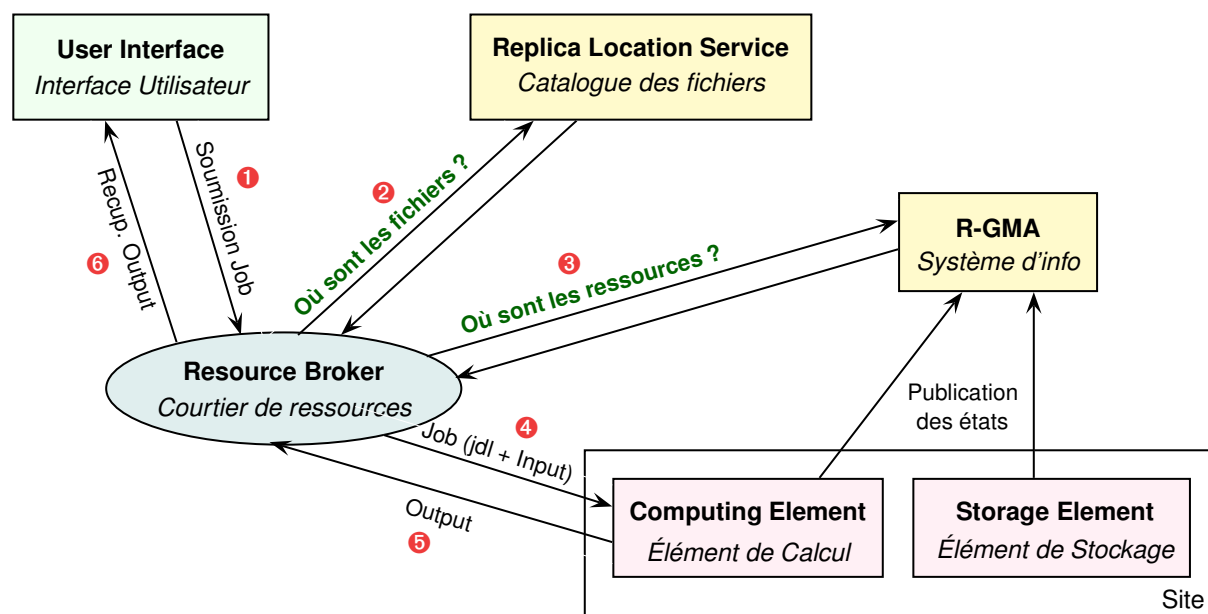


FIG. 3 – Vie d'un job.

La soumission d'un job se fait depuis une interface utilisateur (UI), en envoyant au courtier de ressources (RB) un script JDL (*Job Description Language*) décrivant les caractéristiques du job et ses pré-requis sous forme d'attributs. Ces attributs décrivent les ressources requises (CPU, mémoire, logiciels..), et les fichiers nécessaires. La [liste complète de ces attributs](#) est disponible sur le Web.

Avec ces informations, le RB déterminera l'élément de calcul le plus optimal pour l'exécution.

Le choix du CE d'exécution se fera de manière différente suivant le cas :

**Soumission directe :** Il est possible de spécifier, dans le script JDL, sur quel CE doit s'exécuter le job, auquel cas le RB sert seulement à vérifier la syntaxe du JDL soumis.

**Soumission sans demande d'accès à des fichiers de la Grille :** Dans ce cas, le RB contacte le système d'information (R-GMA) et récupère une liste de CE disposant des ressources de calcul demandées (CPU, mémoire, logiciels), et sur lesquels l'utilisateur a le droit de tourner (informations statiques). Ensuite, il va interroger chacun des CE de cette liste pour déterminer le meilleur candidat, d'après, par exemple, le nombre de CPU ou la mémoire libre (informations dynamiques).

**Soumission avec demande d'accès à des fichiers préalablement stockés sur la Grille :** Dans ce cas, le RB doit d'abord interroger le catalogue de la VO de l'utilisateur (via le RLS) pour savoir où est stockée une copie des fichiers requis par le job, et établir une liste des CE éligibles, c'est-à-dire proches des données requises et sur lesquels l'utilisateur est autorisé à tourner.

Une fois le CE d'exécution choisi, le RB lui soumet le job avec les infos récupérées et prévient le service de *Logging & Bookkeeping* (LB). Ce service peut être interrogé à tout moment par l'utilisateur (commande `edg-job-status`), pour savoir où en est le job soumis.

Lorsque l'exécution commence, les fichiers "locaux" sont envoyés avec l'exécutable, et lorsque le LB annonce que le job est terminé, le CE renvoie les fichiers de sorties éventuels au RB, qui peuvent être récupérés par l'utilisateur (`edg-job-get-output`).

Les différentes étapes de la vie d'un job peuvent être visualisées sur la figure 3, et un exemple de soumission d'un job simple est donné dans la section 9.

Avec la nouvelle version du logiciel EDG, il devrait être possible de soumettre des jobs MPI, avec de nombreuses restrictions, des jobs interactifs, et de faire du *check point / restart*.



## Deuxième partie

# Manipulations sur la Grille

## 8 Obtention d'un certificat

Comme indiqué dans la section 3, l'utilisateur voulant travailler sur la Grille devra d'abord obtenir un certificat électronique.

Attention, sous Unix ou Linux, la demande et la récupération de certificat ne peuvent se faire qu'avec le navigateur NETSCAPE 4.75 ou 4.78. À l'IPGP, NETSCAPE 4.78 est installé sur *falbala*, dans `/usr/dt/bin/`.

Le certificat (valable 1 an) peut être demandé à l'adresse

<http://igc.services.cnrs.fr/Datagrid-fr/index.html>,

ensuite, il faudra :

- le récupérer et le sauvegarder sous le nom `usercertkey.p12`,
- lui mettre les bons droits d'accès (`chmod 600`),
- le transformer du format PKCS12 au format PEM avec la commande `pkcs12-extract`,
- mettre les 2 fichiers créés, `usercert.pem` et `userkey.pem` dans le répertoire `$HOME/.globus`.

### Avertissement :

*Le middleware utilisé par EGEE est pour le moment basé sur celui produit par EDG, mais cela devrait bientôt changer. Aussi, les exemples donnés ici fonctionnent en juin 2004, mais probablement pour une période limitée. Toutes les commandes utilisées ici, et les réponses obtenues, ont été passées depuis les interfaces utilisateur (UI) `albanov.datagrid.jussieu.fr` et `ccedgui.in2p3.fr`, depuis un compte utilisateur appartenant à l'organisation virtuelle (VO) `earthobs` (ou `eo`).*

## 9 Soumission d'un job simple

La soumission des jobs (expliquée dans la section 7) se fait depuis une interface utilisateur (UI), où il faudra créer un *proxy*, avec la commande `grid-proxy-init`, valable par défaut pendant 12 h. L'utilisateur soumet ensuite un script JDL décrivant les caractéristiques de son job et les pré-requis.

### Parmi les attributs d'un script JDL, on trouve :

|                                  |                                                                               |
|----------------------------------|-------------------------------------------------------------------------------|
| <code>Executable</code>          | nom de l'exécutable ou de la commande à exécuter                              |
| <code>Arguments</code>           | arguments de l'exécutable                                                     |
| <code>VirtualOrganisation</code> | VO de l'utilisateur                                                           |
| <code>InputData</code>           | liste des fichiers sur SE nécessaires à l'exécutable sous forme LFN ou PFN    |
| <code>ReplicaCatalog</code>      | identifiant du RC à consulter en cas d'utilisation de LFNs                    |
| <code>StdInput</code>            | nom du fichier standard d'entrée                                              |
| <code>StdOutput</code>           | nom du fichier standard de sortie                                             |
| <code>StdError</code>            | nom du fichier standard d'erreur                                              |
| <code>InputSandbox</code>        | liste des fichiers à envoyer avec l'exécutable                                |
| <code>OutputSandbox</code>       | liste des fichiers à récupérer à la fin du job                                |
| <code>Requirements</code>        | requis pour le job : architecture du nœud de calcul, mémoire, logiciels, etc. |

Les Requirements sont exprimés sous forme d'expressions suivant le schéma **GLUE**, qui est un modèle d'information décrivant les ressources en calcul et en stockage, déjà adopté par plusieurs autres projets Grille. Les attributs qui décrivent les nœuds de calcul sont préfixés par `.other`, par exemple,

```
Requirements = (other.GlueHostMainMemoryRAMSize >= 512) &&
               (other.GlueHostOperatingSystemName == "RedHat") &&
               (other.GlueHostOperatingSystemRelease == 7.3 ) ;
```

demande que le nœud de calcul soit en RedHat 7.3 et possède une mémoire vive d'au moins 512 Mo.

### Exemple simple de script JDL :

```
# hello.jdl
Executable      = "/bin/echo";
Arguments       = " Hello world";
VirtualOrganisation = "eo";
StdOutput       = "hello.out";
StdError        = "hello.err";
OutputSandbox   = {"hello.out", "hello.err"};
```

Ce script demande l'exécution de la commande `"/bin/echo Hello world"`, et la récupération des sortie et erreur standard dans les fichiers `hello.out` et `hello.err`.

Soumission du job :

```
edg-job-submit hello.jdl
```

la commande rend alors un JID (*Job Identifier*) permettant de suivre le job. Ce JID contient entre autres, l'adresse du *Resource Broker*, ou RB, traitant le job. Par exemple :

```
https://boszwi.jn.nikhef.nl:9000/QM1QBA3MNWgW42Z3JHrS4Q
```

Pour éviter de manipuler cet identificateur, la soumission peut se faire en le redirigeant dans un fichier grâce à l'option `-o` (comme `output`) :

```
edg-job-submit -o hello.jid hello.jdl
```

Pour savoir où en est le job :

```
edg-job-status -i hello.jid
```

le `-i` (comme `input`) permet de stipuler le fichier contenant le JID du job suivi.

Quand le "Current Status" du job sera "Done (Success)" (après être passé par les états Submitted, Waiting, Ready, Scheduled, Running), cela voudra dire que le job est terminé, et que les sorties peuvent être récupérées par la commande :

```
edg-job-get-output -i hello.jid
```

les fichiers spécifiés par l'attribut `OutputSandbox` pourront alors être récupérés dans un répertoire du UI (généralement `/tmp/xxx`) dont le nom sera affiché lorsque la commande `edg-job-get-output` aura rendu la main.

## 10 Interrogation du système d'information R-GMA

Le système d'information de la Grille R-GMA (voir section 5) permet la gestion et la surveillance des ressources disponibles. Une interface ligne-à-ligne, `edg-rgma`, permet de consulter les informations publiées.

Les pages suivantes donnent quelques exemples de requêtes de type SQL.

### Liste des tables publiées par R-GMA

```
edg-rgma -c "show tables"
```

le `-c` (comme `commande`) demande l'exécution de la commande (ou requête) qui suit, ici la liste des tables publiées. Le résultat obtenu (et tronqué) est :

| tableId | name                                     |
|---------|------------------------------------------|
| 1       | GlueCE                                   |
| 2       | GlueCEAccessControlBaseRule              |
| 3       | GlueCluster                              |
| 4       | GlueSubCluster                           |
| 5       | GlueSubClusterSoftwareRunTimeEnvironment |
| 6       | GlueHostRemoteFileSystem                 |
| 7       | GlueCESEBind                             |
| 8       | GlueSE                                   |
| ...     |                                          |
| 52      | ETT_tst                                  |
| 53      | ComputingLoadExample                     |

### Liste des caractéristiques des éléments de calcul (CE) publiées

```
edg-rgma -c "describe GlueCE"
```

Table: GlueCE

|                       |                  |                       |                   |              |
|-----------------------|------------------|-----------------------|-------------------|--------------|
| UniqueID              | Name             | GlueClusterUniqueID   | TotalCPUs         | LRMSType     |
| LRMSVersion           | GRAMVersion      | HostName              | GatekeeperPort    | RunningJobs  |
| WaitingJobs           | TotalJobs        | Status                | WorstResponseTime |              |
| EstimatedResponseTime | FreeCpus         | Priority              | MaxRunningJobs    | MaxTotalJobs |
| MaxCPUTime            | MaxWallClockTime | InformationServiceURL | MeasurementDate   |              |
| MeasurementTime       |                  |                       |                   |              |
| ...                   |                  |                       |                   |              |
| VARCHAR(128)          | VARCHAR(255)     | VARCHAR(100)          | INT               | VARCHAR(255) |
| VARCHAR(255)          | VARCHAR(255)     | VARCHAR(128)          | VARCHAR(128)      | INT          |
| ...                   |                  |                       |                   |              |
| 1 Rows in set         |                  |                       |                   |              |

## Nombre de processeurs disponibles pour chaque queue des CE enregistrés

```
edg-rdma -c "latest select UniqueID,TotalCPUs from GlueCE"
```

demande la liste des identifiants et le nombre de processeurs des queues des éléments de calcul (CE) enregistrés, *latest* limitant l’affichage aux dernières (plus récentes) données publiées.

```
+-----+-----+
| UniqueID                                     | TotalCPUs |
+-----+-----+
| farm003.hep.phy.cam.ac.uk:2119/jobmanager-pbs-workq | 1          |
| dgrid-2.srce.hr:2119/jobmanager-pbs-medium          | 2          |
| hudson.datagrid.jussieu.fr:2119/jobmanager-pbs-short | 1          |
| dgrid-2.srce.hr:2119/jobmanager-pbs-long            | 2          |
...
| ce001.fzk.de:2119/jobmanager-pbs-medium             | 2          |
| grid0007.esrin.esa.int:2119/jobmanager-pbs-short    | 56         |
| ccgridli05.in2p3.fr:2119/jobmanager-pbs-short       | 6          |
| ccgridli05.in2p3.fr:2119/jobmanager-pbs-medium     | 6          |
| marseille-ce01.datagrid.cnrs.fr:2119/jobmanager-pbs-short | 3          |
+-----+-----+
18 Rows in set
```

## 11 Manipulation des fichiers : **edg-replica-manager**

L’**edg-replica-manager** (voir section 6) permet de déposer, dupliquer, manipuler des fichiers sur la Grille et de les enregistrer sur des catalogues. La commande **edg-rm** (ou **edg-replica-manager**) permet d’effectuer ces opérations.

### Ressources disponibles

Pour savoir quelles ressources sont disponibles (en particulier les unités de stockage ou SE) pour la VO *earthobs*, on peut passer la commande :

```
edg-rm -vo=eo printInfo
```

et l’on obtiendra la réponse :

```
VO used           : eo
default SE       : ccgridli02.in2p3.fr
default CE       : ccgridli05.in2p3.fr
Info Service     : RGMA

RMC endpoint     : http://gppls05.gridpp.rl.ac.uk:8080/eo/edg-replica-metadata-catalog/
                  services/edg-replica-metadata-catalog
LRC endpoint     : http://gppls05.gridpp.rl.ac.uk:8080/eo/edg-local-replica-catalog/
                  services/edg-local-replica-catalog
ROS endpoint     : http://ccgridli08.in2p3.fr:8080/eo/edg-replica-optimization/
                  services/edg-replica-optimization

List of CE ID's  : ccgridli05.in2p3.fr:2119/jobmanager-pbs-long
                  ccgridli05.in2p3.fr:2119/jobmanager-pbs-medium
                  ccgridli05.in2p3.fr:2119/jobmanager-pbs-short
```

```

ce001.fzk.de:2119/jobmanager-pbs-long
ce001.fzk.de:2119/jobmanager-pbs-medium
ce001.fzk.de:2119/jobmanager-pbs-short
dgrid-2.srce.hr:2119/jobmanager-pbs-infinite
...
List of SE ID's      : ccgridli02.in2p3.fr
                     dgrid-3.srce.hr
                     gppse05.gridpp.rl.ac.uk
                     grid-s2.ifae.es
                     grid0010.esrin.esa.int
                     marseille-se01.datagrid.cnrs.fr
...
SE at datagrid.cnrs.fr :
    name : datagrid.cnrs.fr
    host : marseille-se01.datagrid.cnrs.fr
    type : edg-se
    accesspoint : /flatfiles/SE00
    endpoint : http://marseille-se01.datagrid.cnrs.fr:8080/
              edg-se-webservice/services/edg-se-webservice
    VOs : dzero,babar,iteam,tutor,wpsix,eo,biome,lhcb,cms,atlas,alice
    VO directories : dzero:/dzero,babar:/babar,iteam:/iteam,tutor:/tutor,
                    wpsix:/wpsix,eo:/eo,biome:/biome,lhcb:/lhcb,cms:/cms,
                    atlas:/atlas,alice:/alice
    protocols : file,rfio,gsiftp
...

```

La sortie affichée montre que l'on peut écrire sur `marseille-se01.datagrid.cnrs.fr` où l'on va copier et enregistrer un fichier `GM1.txt` sous le nom logique `GMse1.txt` :

```

edg-rm --vo eo --insecure copyAndRegisterFile \
    file:`pwd`/GM1.txt \
    -l lfn:GMse1.txt \
    -d marseille-se01.datagrid.cnrs.fr

```

où `--insecure` stipule de ne pas utiliser le protocole https,  
`-l` nom logique du fichier,  
`-d` destination ou SE cible.

La réponse obtenue :

```
guid:c3033943-ae4e-11d8-840b-d4ce59be34b5
```

donne le *Grid Unique Identifier* ou GUID du fichier.

On peut vouloir dupliquer ce fichier (et enregistrer sa copie dans le catalogue de la VO) sur un autre SE, sur `ccgridli02.in2p3.fr` par exemple :

```

edg-rm --vo eo --insecure replicateFile \
    guid:c3033943-ae4e-11d8-840b-d4ce59be34b5 \
    -d ccgridli02.in2p3.fr

```

L'affichage de la liste des replicas déposés pour un GUID donné peut se faire avec l'option `listReplicas` de `edg-rm` :

```

edg-rm --vo eo --insecure listReplicas \
    guid:c3033943-ae4e-11d8-840b-d4ce59be34b5

```

la réponse obtenue :

```
srm://marseille-se01.datagrid.cnrs.fr/eo/generated/2004-05-25/  
    filead30fb22-ae4e-11d8-840b-d4ce59be34b5  
srm://ccgridli02.in2p3.fr/eo/generated/2004-06-01/  
    file21d31cfd-b3d1-11d8-9857-8dc8a61e8f8a
```

donne cette fois les identificateurs SRM (*Storage Resource Manager*) des replicas, contenant le nom du SE où ils sont stockés.

Ici, le 1<sup>er</sup> SRM affiché est celui du fichier déposé sur la Grille par l'opération `copyAndRegisterFile`, et le 2<sup>e</sup>, celui du fichier déposé par `replicateFile`.

## Troisième partie

# Appendices

## 12 Sigles et liens

|           |                                                                                                                                                                                                                                                                                      |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AliEn     | <i>Alice Environment</i><br><a href="http://www.slac.stanford.edu/BFROOT/www/Computing/Offline/BaBarGrid/meetings/15Dec2003/AliEnforBaBar.pdf">http://www.slac.stanford.edu/BFROOT/www/Computing/Offline/BaBarGrid/meetings/15Dec2003/AliEnforBaBar.pdf</a>                          |
| CA        | <i>Cerification Authority</i><br>CA du CNRS : <a href="http://igc.services.cnrs.fr/Datagrid-fr/">http://igc.services.cnrs.fr/Datagrid-fr/</a>                                                                                                                                        |
| CE        | <i>Computing Element</i>                                                                                                                                                                                                                                                             |
| CERN      | Centre Européen de la Recherche Nucléaire<br><a href="http://public.web.cern.ch/public/">http://public.web.cern.ch/public/</a>                                                                                                                                                       |
| CIC       | <i>Core Infrastructure Centre</i>                                                                                                                                                                                                                                                    |
| CNAF      | <i>Centro Nazionale per la Ricerca e Sviluppo nelle Tecnologie Informatiche e Telematiche dell'INFN</i>                                                                                                                                                                              |
| CPPML     | Centre de Physique des Particules de Marseille<br><a href="http://marwww.in2p3.fr/">http://marwww.in2p3.fr/</a>                                                                                                                                                                      |
| DEISA     | <i>Distribued European Infrastructure for Supercomputing Applications</i><br><a href="http://www.deisa.org">www.deisa.org</a>                                                                                                                                                        |
| EDG       | <i>European DataGRID</i><br><a href="http://eu-datagrid.web.cern.ch/eu-datagrid">http://eu-datagrid.web.cern.ch/eu-datagrid</a><br><a href="http://www.ipgp.jussieu.fr/~moguilly/GRILLE/EDG.pdf">http://www.ipgp.jussieu.fr/~moguilly/GRILLE/EDG.pdf</a>                             |
| EGAAP     | <i>Generic Application Advisory Panel</i>                                                                                                                                                                                                                                            |
| EGEE      | <i>Enabling Grids for E-science in Europe</i><br><a href="http://public.eu-egee.org/">http://public.eu-egee.org/</a>                                                                                                                                                                 |
| EUGridPMA | <i>European Policy Management Authority for Grid Authentication e-Science</i><br><a href="http://www.eugridpma.org">http://www.eugridpma.org</a>                                                                                                                                     |
| GEANT     | <i>Gigabit European Academic Network</i><br><a href="http://msi.lms.lt/5FP/IST/ist_geant.html">http://msi.lms.lt/5FP/IST/ist_geant.html</a><br>GN2 : <a href="http://public.eu-egee.org/kickoff/programme/slot-dv.html">http://public.eu-egee.org/kickoff/programme/slot-dv.html</a> |
| GENIUS    | <i>Grid Enabled web eNvironment for site Independent User job Submission</i><br><a href="https://genius.ct.infn.it/">https://genius.ct.infn.it/</a><br><a href="https://grid-tutor.ct.infn.it/">https://grid-tutor.ct.infn.it/</a>                                                   |
| GGF       | <i>Global Grid Forum</i><br><a href="http://www.gridforum.org/">http://www.gridforum.org/</a>                                                                                                                                                                                        |
| GILDA     | <i>Grid INFN Laboratory for Dissemination Activities</i><br><a href="http://gilda.ct.infn.it/">http://gilda.ct.infn.it/</a><br><a href="https://grid-demo.ct.infn.it/">https://grid-demo.ct.infn.it/</a>                                                                             |
| GK        | <i>GateKeeper</i>                                                                                                                                                                                                                                                                    |
| GLUE      | <i>Grid Laboratory Uniform Environment</i>                                                                                                                                                                                                                                           |

|       |                                                                                                                                                                                                                         |
|-------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|       | <a href="http://pcct108.ct.infn.it/workshop-infncgrid/slides/andreozi.ppt">http://pcct108.ct.infn.it/workshop-infncgrid/slides/andreozi.ppt</a>                                                                         |
| GMA   | <i>Grid Monitoring Architecture</i>                                                                                                                                                                                     |
| GOC   | <i>Grid Operation Center</i>                                                                                                                                                                                            |
| GUID  | <i>Grid Unique Identifier</i>                                                                                                                                                                                           |
| IC    | <i>Information Catalog</i>                                                                                                                                                                                              |
| IDRIS | Institut du Développement et des Ressources en Informatique Scientifique<br><a href="http://www.idris.fr/">http://www.idris.fr/</a>                                                                                     |
| IN2P3 | Institut National de Physique Nucléaire et de Physique des Particules<br><a href="http://www.in2p3.fr/">http://www.in2p3.fr/</a><br>Centre de Calcul de l'IN2P3 : <a href="http://cc.in2p3.fr/">http://cc.in2p3.fr/</a> |
| JDL   | <i>Job Description Language</i><br><a href="http://server11.infn.it/workload-grid/docs/DataGrid-01-TEN-0142-0_2.pdf">http://server11.infn.it/workload-grid/docs/DataGrid-01-TEN-0142-0_2.pdf</a>                        |
| LAL   | Laboratoire de l'Accélérateur Linéaire<br><a href="http://www.lal.in2p3.fr/">http://www.lal.in2p3.fr/</a>                                                                                                               |
| LB    | <i>Logging &amp; Bookkeeping</i>                                                                                                                                                                                        |
| LCAS  | <i>Local Centre Authorization Server</i>                                                                                                                                                                                |
| LCG   | <i>LHC Computing Grid project</i>                                                                                                                                                                                       |
| LCN   | <i>Logical Collection Name</i>                                                                                                                                                                                          |
| LDAP  | <i>Lightweight Directory Access Protocol</i><br><a href="http://hepunix.rl.ac.uk/edg/wp3/documentation/wp3-ldap_user_guide.html">http://hepunix.rl.ac.uk/edg/wp3/documentation/wp3-ldap_user_guide.html</a>             |
| LFN   | <i>Logical File Name</i>                                                                                                                                                                                                |
| LHC   | <i>Large Hadron Collider</i><br><a href="http://www.cern.ch/lcg/">http://www.cern.ch/lcg/</a>                                                                                                                           |
| LPC   | Laboratoire de Physique Corpusculaire<br><a href="http://clrwww.in2p3.fr/">http://clrwww.in2p3.fr/</a>                                                                                                                  |
| LRC   | <i>Local Replica Catalog</i>                                                                                                                                                                                            |
| LRMS  | <i>Local Resource Management System</i>                                                                                                                                                                                 |
| MDS   | <i>Monitoring and Discovery Services</i><br><a href="http://www.globus.org/mds/">http://www.globus.org/mds/</a>                                                                                                         |
| MON   | <i>Monitoring Node</i>                                                                                                                                                                                                  |
| NOC   | <i>Network Operation Centre</i>                                                                                                                                                                                         |
| NREN  | <i>National Research and Educational Network</i>                                                                                                                                                                        |
| NSF   | <i>National Science Foundation</i>                                                                                                                                                                                      |
| OAG   | <i>Operations Advisory Group</i>                                                                                                                                                                                        |
| OGSA  | <i>Open Grid Services Architecture</i><br><a href="http://www.globus.org/ogsa/">http://www.globus.org/ogsa/</a>                                                                                                         |
| OMC   | <i>Operations Management Centre</i>                                                                                                                                                                                     |
| PFN   | <i>Physical File Name</i>                                                                                                                                                                                               |



|       |                                                                                                                                                                                                                                                                                                                    |
|-------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| R-GMA | <i>Relationnal Grid Monitoring Architecture</i><br><a href="http://gppic05.gridpp.rl.ac.uk:8080/R-GMA/index.html">http://gppic05.gridpp.rl.ac.uk:8080/R-GMA/index.html</a><br>Navigateur : <a href="http://gppic05.gridpp.rl.ac.uk:8080/R-GMA/index.html">http://gppic05.gridpp.rl.ac.uk:8080/R-GMA/index.html</a> |
| RB    | <i>Resource Broker</i>                                                                                                                                                                                                                                                                                             |
| RC    | <i>Regional Center</i>                                                                                                                                                                                                                                                                                             |
| RDBMS | <i>Relational Database Management System</i> (SGBDR en français)                                                                                                                                                                                                                                                   |
| RLI   | <i>Replica Location Index</i>                                                                                                                                                                                                                                                                                      |
| RLS   | <i>Replica Location Service</i>                                                                                                                                                                                                                                                                                    |
| RMC   | <i>Replica Metadata Catalog</i>                                                                                                                                                                                                                                                                                    |
| ROC   | <i>Regional Operations Centre</i>                                                                                                                                                                                                                                                                                  |
| ROS   | <i>Replica Optmization Service</i>                                                                                                                                                                                                                                                                                 |
| RSH   | <i>Replica Storage Handler</i>                                                                                                                                                                                                                                                                                     |
| RSS   | <i>Replica Subscription Service</i>                                                                                                                                                                                                                                                                                |
| SE    | <i>Storage Element</i>                                                                                                                                                                                                                                                                                             |
| SFN   | <i>Storage File Name</i>                                                                                                                                                                                                                                                                                           |
| SGBDR | Système de Gestion de Bases de Données Relationnelles ( <i>RDBMS</i> en anglais)                                                                                                                                                                                                                                   |
| SLA   | <i>Service Level Agreement</i>                                                                                                                                                                                                                                                                                     |
| SLR   | <i>Service Level Requirements</i>                                                                                                                                                                                                                                                                                  |
| SLS   | <i>Service Level Specifications</i>                                                                                                                                                                                                                                                                                |
| SQL   | <i>Structured Query Language</i> (langage de requêtes structurées)<br><a href="http://www.sql.org/">http://www.sql.org/</a>                                                                                                                                                                                        |
| SRM   | <i>Storage Resource Manager</i>                                                                                                                                                                                                                                                                                    |
| SURL  | <i>Storage URL</i>                                                                                                                                                                                                                                                                                                 |
| TURL  | <i>Transport URL</i>                                                                                                                                                                                                                                                                                               |
| UI    | <i>User Interface</i>                                                                                                                                                                                                                                                                                              |
| UREC  | Unité Réseaux du CNRS<br><a href="http://www.urec.cnrs.fr/">http://www.urec.cnrs.fr/</a>                                                                                                                                                                                                                           |
| URL   | <i>Unified Resource Locator</i>                                                                                                                                                                                                                                                                                    |
| VDT   | <i>Virtual Data Toolkit</i><br><a href="http://www.lsc-group.phys.uwm.edu/vdt/">http://www.lsc-group.phys.uwm.edu/vdt/</a>                                                                                                                                                                                         |
| VO    | <i>Virtual Organisation</i>                                                                                                                                                                                                                                                                                        |
| VOMS  | <i>Virtual Organisation Membership Service</i><br><a href="http://hep-project-grid-scg.web.cern.ch/hep-project-grid-scg/voms.html">http://hep-project-grid-scg.web.cern.ch/hep-project-grid-scg/voms.html</a>                                                                                                      |
| WMS   | <i>Workload Management Service</i>                                                                                                                                                                                                                                                                                 |
| WN    | <i>Worker Node</i>                                                                                                                                                                                                                                                                                                 |
| WSDL  | <i>Web Services Description Language</i>                                                                                                                                                                                                                                                                           |

## 13 Autres liens utiles

➤ **Le projet EGEE :** <http://public.eu-egee.org/intro>

- Fédérations participant à EGEE :

[egee-intranet.web.cern.ch/egee-intranet/dissemination/images/TA/figure3.gif](http://egee-intranet.web.cern.ch/egee-intranet/dissemination/images/TA/figure3.gif)

- Partenaires d'EGEE :

<http://public.eu-egee.org/partners>

- Fédération française d'EGEE :

<http://egee.in2p3.fr/>

Partenaires français :

- CEA/DSM (Commissariat à l'Énergie Atomique, Direction des Sciences de la Matière) :

<http://www-dsm.cea.fr/>

- CGG (Compagnie Générale de Géophysique) :

<http://www.cgg.com/>

- CS (Compagnie des Signaux, Systèmes d'Information) :

<http://www.c-s.fr/>

- Centrale Recherche S.A. :

<http://www.ecp.fr/fr/F-recherche/F4-crsa/crsa.html>

émanation de l'École Centrale de Paris : <http://www.ecp.fr/>

- Partenaires des applications **Sciences de la Terre** :

- ESA (*European Space Agency*) : <http://www.esa.int/>

- KNMI (*Koninklijk Nederlands Meteorologisch Instituut*) : <http://www.knmi.nl/>

- IPSL (*Institut Pierre Simon Laplace*) : <http://www.ipsl.jussieu.fr/>

- DKRZ (*Deutsches Klimarechenzentrum*) : <http://www.dkrz.de/>

- UTV (*Università Tor Vergata*) : <http://www.disp.uniroma2.it/>

- IPGP (*Institut de Physique du Globe de Paris*) : <http://www.ipgp.jussieu.fr/>

- Université de Neuchâtel : <http://www.unine.ch/>

- CGG (Compagnie Générale de Géophysique) : <http://www.cgg.com/>

➤ D'autres projets Grille dont ceux affiliés à GridStart (<http://www.gridstart.org/>) :

- CrossGrid : <http://www.eu-crossgrid.org/>

- DAMIEN : <http://www.hlrs.de/organization/pds/projects/damien/>

- DataTAG : <http://datatag.web.cern.ch/datatag/>

- EGSO : <http://www.egso.org>

- EuroGrid : <http://www.eurogrid.org/>

- GRIA : <http://www.gria.org/>

- GridLab : <http://www.gridlab.org/>

- GRIP : <http://www.grid-interoperability.org/>

- bioGRID : <http://www.bio-grid.net/index.jsp>

- GEMSS : <http://www.ccrl-nece.de/gemss/>

- GRACE : <http://www.grace-ist.org/>

- GRASP : <http://eu-grasp.net/english/default.htm>

- MammoGrid : <http://mammogrid.vitamib.com/>
  - MOSES : <http://www.hum.ku.dk/moses/>
  - OpenModGRID : <http://www.openmolgrid.org/>
  - SeLeNe : <http://www.dcs.bbk.ac.uk/selene/>
  - DEISA : <http://www.deisa.org/>
  - HPC-Europa : <http://www.hpc-europa.org/>
  - Cyberinfrastructure : <http://www.adec.edu/nsf/nsfcyberinfrastructure.html>
- Les certificats électroniques :
- Au CNRS : [http://www.urec.cnrs.fr/igc/Certifs\\_CNRS.html](http://www.urec.cnrs.fr/igc/Certifs_CNRS.html)
  - Demande de certificat CNRS : <http://igc.services.cnrs.fr/Datagrid-fr/index.html>
- Bases de *middleware* Grille :
- GLOBUS : <http://www.globus.org/>
  - UNICORE : <http://www.unicore.org/>
- Dictionnaire des termes du *Grid Computing* :  
<http://grid-it.cnaf.infn.it/fileadmin/users/dictionary/dictionary.html>
- Image aspirine : <http://home.nordnet.fr/~ecarpentier/maths/aspirine.gif>