

Introduction à la Grille et à EGEE

Laboratoire de l'Accélérateur Linéaire

2 février 2007

Machines utilisées pour les exercices

grid01.lal.in2p3.fr	BDII (système d'information)
grid08.lal.in2p3.fr	Serveur R-GMA pour le LAL
grid09.lal.in2p3.fr	Resource Broker (RB) au LAL
grid10.lal.in2p3.fr	Computing Element (CE) au LAL
grid11.lal.in2p3.fr	"User Interface" (UI)
grid12.lal.in2p3.fr	Serveur VOMS pour la VO vo.u-psud.fr
grid14.lal.in2p3.fr	Serveur LFC pour la VO vo.u-psud.fr
bdii.grif.fr	BDII pour le site GRIF

1 Etapes Préliminaires

1.1 Postes de travail

1.1.1 Les iMacs

Les iMacs sont disponibles pour les étudiants. Il est préférable de travailler en binôme. L'application "Terminal" permet d'accéder aux autres machines (principalement le "User Interface") par **ssh**. Codes d'accès pour les iMacs :

- utilisateur : ens<numéro de poste¹>
- mot de passe : etud0202

Le tableau suivant a des touches utiles pour les commandes dans ce document.

Touches pour les caractères spéciaux

{	alt + (
}	alt +)
[alt + maj + (
]	alt + maj +)
\	alt + maj + /
	alt + L
~	alt + n

1.1.2 Votre propre ordinateur

Si vous avez apporté votre propre ordinateur, vous pouvez le brancher sur le réseau sans fil du LAL. Vous avez besoin d'un client ssh et d'un navigateur (pas safari) installés dans votre ordinateur.

¹le numéro de poste est un chiffre que vous trouverez sur le pied de votre ordinateur

Utilisez le réseau "LALpublic". Ce réseau est disponible près de la Salle Bleue. Il n'y a pas de clé d'accès.

Si vous êtes du LAL, vous pouvez également utiliser le réseau sans fil "LAL-WaveLAN" si vous préférez.

1.2 Ouvrez une session sur l'ordinateur "User Interface"

Vous allez utiliser la machine `grid11.lal.in2p3.fr` pour accéder à la grille. Des comptes y ont été créés : `etud01`, `etud02`, etc. Le mot de passe associé à chaque compte est "etud0202".

Ouvrez une session sur cette machine avec le nom d'utilisateur donné en utilisant la commande `ssh`. Si vous voulez utiliser X11 (`xemacs` par exemple), vérifiez que vous avez démarré un serveur X sur votre ordinateur et que vous avez défini la variable d'environnement `DISPLAY`. (Le serveur X démarre avec l'application X11 dans les iMacs.) Dans ce cas, utilisez l'option "-Y" avec la commande `ssh`.

```
export DISPLAY=:0.0
ssh -Y etud01@grid11.lal.in2p3.fr
```

1.3 Récupération des fichiers

Pour faciliter les exercices, il existe un dossier d'archives contenant tous les fichiers nécessaires.

1. L'archive est disponible sur la page de l'agenda du tutorial et aussi dans le "User Interface", `grid11.lal.in2p3.fr`. Depuis `grid11.lal.in2p3.fr`, récupérez l'archive contenant ces fichiers :

```
cp /usr/share/tutorial/20070202-tutorial.tar.gz .
```

2. Décompressez l'archive dans votre "home directory" :

```
tar xzf 20070202-tutorial.tar.gz
```

Les fichiers sont dans le répertoire `tutorial_material`.

2 Authentification et Autorisation

2.1 Demande de certificat

On va utiliser la procédure normale pour obtenir un certificat. Les certificats pour cette formation ont une limite de validité de quelques jours. Normalement un certificat a une limite de validité d'un an et peut être renouvelé.

1. Connectez-vous sur l'autorité de certification CNRS² pour la France. *N'utilisez pas le navigateur safari.*

²<http://igc.services.cnrs.fr/GRID-FR/certificats.html>

2. Cliquez sur “Certificat personnel” à gauche de la page.
3. Remplissez les données du premier formulaire. Pour le nom utilisez “EtudiantUn”, “EtudiantDeux”, etc. Utilisez votre propre prenom. Cliquez sur “[suite]”. La procédure vérifie votre adresse électronique. Il est impératif que vous fournissiez une adresse valide et que vous puissiez accéder à votre boîte aux lettres à distance.
4. Choisissez votre organisme, puis cliquez sur “[suite]”. Indiquez CNRS.
5. Choisissez votre laboratoire en consultant la liste, puis cliquez sur “[suite]”. Indiquez le LAL.
6. Un récapitulatif des informations apparaît. Vérifiez que les informations sont correctes. Si vous cliquez sur “[suite]”, la demande est envoyée.
7. L’autorité de certification envoie un message électronique à l’adresse que vous avez indiquée. Dans ce message il y a un lien que vous devez visiter pour valider votre demande. Si vous utilisez les filtres SPAM, vérifiez que ce message de confirmation n’est pas marqué comme SPAM.
8. Une fois la demande acceptée, vous allez recevoir un deuxième message électronique contenant un lien à cliquer. A l’ouverture de la page correspondante, votre certificat est importé dans votre navigateur. *Vous devez ensuite utiliser le même navigateur que vous avez utilisé pour faire la demande.*

2.2 L’enregistrement dans une VO

L’accès aux ressources de la grille est contrôlé par des organisations virtuelles (VO). Il est nécessaire d’être membre d’au moins une VO. Une fois que vous avez votre certificat, il est nécessaire de contacter le “VO Manager” d’une VO pour faire l’enregistrement.

Le site du “CIC Operations Portal”³ contient les informations pour les VO principales et le lien pour s’enregistrer dans ces VO. Cette page ne contient pas toutes les VO. Il peut être nécessaire de contacter quelqu’un de votre organisation pour connaître le serveur correct.

Aujourd’hui, vous allez utiliser la VO “vo.u-psud.fr”.

1. Allez au serveur VOMS⁴ pour la VO vo.u-psud.fr avec un navigateur sur lequel votre certificat est installé.
2. Cliquez sur le lien “New user registration” dans la liste à gauche.
3. Remplissez le formulaire et cliquez sur le bouton “I have read and agree to the VO’s Usage Rules”.
4. Vous allez recevoir un message électronique pour confirmer votre adresse électronique. (Vérifiez encore que le message n’est pas marqué comme SPAM.) Cliquez sur l’URL de confirmation indiquée.
5. Le “VO Manager” doit valider votre demande. Une fois la demande validée, vous allez recevoir un message électronique.

³<https://cic.gridops.org/>

⁴<https://grid12.lal.in2p3.fr:8443/voms/vo.u-psud.fr>

2.3 Exportation, installation et utilisation du certificat

Votre certificat est installé dans votre browser. Pour utiliser la grille, on doit extraire ce certificat et l'installer dans le "User Interface" grid11.lal.in2p3.fr.

1. Utilisez la fonction "export" de votre navigateur pour faire une sauvegarde de votre certificat au format pkcs12 (fichier .p12).

Firefox Sélectionnez le menu FireFox/Preferences. Cliquez sur "Avancer" à droite et l'onglet "Chiffrement". Cliquez sur le bouton "Afficher les résultats...", sélectionnez votre certificat et utilisez la fonction "importer".

Mozilla Dans le menu Edit/Preferences allez dans l'onglet "Privacy & Security" puis dans "certificates". Cliquez sur "Manage Certificates", et utilisez la fonction "backup".

Netscape Cliquez sur l'icône "sécurité" (ou passez par le menu Communicator/outils/Information sur la sécurité), puis cliquez sur le choix "vos certificats".

Internet Explorer Dans le menu Outils/Options Internet, allez dans l'onglet "contenu" et cliquez sur "certificats". Cliquez ensuite sur "exporter".

La procédure va demander un mot de passe pour sauvegarder votre certificat. Ne l'oubliez pas.

2. Copiez votre certificat dans la machine grid11.lal.in2p3.fr. Remplacez *cert.p12* par le nom exact de votre certificat dans toutes les commandes suivantes. Il a peut-être une extension ".pfx".

```
scp cert.p12 etud01@grid11.lal.in2p3.fr:~/cert.p12
```

3. Connectez-vous via **ssh** sur grid11.lal.in2p3.fr et créez, dans votre "home directory", le répertoire *.globus*. Descendez dans ce répertoire.

```
mkdir .globus
cd .globus
mv ~/cert.p12 .
```

4. Convertissez votre certificat au bon format et donnez-lui les bons droits. Pour utiliser la grille, vous avez besoin de deux fichiers *usercert.pem* et *userkey.pem* qui contiennent votre clé publique et votre clé privée.

```
openssl pkcs12 -in cert.p12 -clcerts -nokeys -out usercert.pem
openssl pkcs12 -in cert.p12 -nocerts -out userkey.pem
```

Ces commandes vont vous demander d'entrer le mot de passe que vous avez utilisé pour créer le fichier *cert.p12*. Pour créer le fichier *userkey.pem*, la commande vous demande aussi de donner un nouveau mot de passe associé pour protéger le fichier. Normalement on utilise le même mot de passe que pour le fichier *cert.p12*.

5. Vérifiez que les fichiers disposent des bons droits. Le fichier *userkey.pem* doit avoir les droits 0400 ; le fichier *usercert.pem* doit avoir les droits 0444. Changez les droits avec **chmod** si nécessaire.

```
ls -l ~/.globus
chmod 0400 ~/.globus/userkey.pem
chmod 0444 ~/.globus/usercert.pem
```

6. Repérez les différents champs du certificat (sujet, validité...) en lisant son contenu avec la commande **openssl**.

```
openssl x509 -text -noout -in ~/.globus/usercert.pem
```

Quel est votre "Subject" ? Quelle est la date de fin de validité de votre certificat ?

2.4 Manipulation d'un proxy

Pour permettre à vos jobs d'accéder aux services de la grille, vous devez créer un proxy. Un proxy est un fichier signé avec votre certificat qui est envoyé avec vos jobs. Les proxies sont valables pour une durée limitée, normalement 12 à 24 heures.

1. Depuis la machine grid11.lal.in2p3.fr, demandez un proxy.⁵

```
voms-proxy-init --voms vo.u-psud.fr
```

Cette commande crée un fichier signé avec votre certificat que vos jobs peuvent utiliser pour accéder aux services de la grille. Il vous est demandé de donner votre mot de passe.

2. Affichez les informations du proxy créé.

```
voms-proxy-info
```

Les informations intéressantes sont le "subject", l'équivalent de votre "username" dans la grille, le "timeleft", le temps durant lequel le certificat est encore valable, et le nom de fichier pour le proxy.

3. Supprimez explicitement le proxy.

```
voms-proxy-destroy
```

Cette commande supprime le proxy dans l'ordinateur local. Les copies éventuellement envoyées avec vos jobs ne sont pas affectées. Vérifiez que le proxy a bien été supprimé avec la commande **voms-proxy-info**.

4. Regardez les autres options disponibles avec la commande **voms-proxy-init**. (Utilisez **man** ou l'option "-help".) On peut changer, par exemple,

⁵L'erreur "Cannot find file or dir : ... /glite/vomses" est normale.

la durée de vie et la taille (nombre de “bits”) du proxy. Normalement les valeurs par défauts sont correctes.

Les anciennes commandes **grid-proxy-*** pour manipuler les proxies existent toujours. Mais ces commandes ne supportent pas de mettre les informations de votre VO dans le proxy. Les proxies VOMS sont 100% compatibles avec l’ancien format et il est préférable d’utiliser les commandes **voms-proxy-***.

Il existe aussi des commandes **myproxy-***. Elles permettent le renouvellement automatique des proxies pour des jobs de très longue durée. Ces commandes ne seront pas utilisées dans ce tutorial.

2.5 Utilisation des groupes et rôles dans une VO

Les organisations virtuelles qui utilisent VOMS (Virtual Organisation Membership Server) peuvent définir des groupes et des rôles. Tous les groupes que vous avez le droit d’utiliser sont décrits dans votre proxy VOMS. Quand on crée un proxy, il est nécessaire de demander d’ajouter les rôles explicitement.

Actuellement les groupes et les rôles ne sont pas respectés par tous les services grille. Dans le futur on pourra les utiliser pour le contrôle d’accès aux fichiers et aux services.

1. Pour lister les groupes de votre proxy utilisez la commande **voms-proxy-info** avec l’option “-all” (peut-être vous devrez créer un nouveau proxy avec **voms-proxy-init**). Les groupes dans les lignes marquées “attribute”. Une VO peut créer une hiérarchie dans les groupes.

```
|| voms-proxy-info -all
```

De quel(s) groupe(s) faites vous partie ?

2. Ajouter le rôle “Tutorial1” ou “Tutorial2” dans votre proxy. La moitié des participants à ce tutorial possède les droits pour utiliser le rôle “Tutorial1” et les autres pour utiliser “Tutorial2”. Si vous demandez un rôle pour lequel vous n’avez pas les droits, la commande **voms-proxy-init** échoue.

```
|| voms-proxy-init --voms vo.u-psud.fr:/Role=Tutorial1
```

Vérifiez le rôle dans votre proxy avec la commande **voms-proxy-info**. Vérifiez aussi que la commande génère une erreur quand vous demandez le mauvais rôle.

3 Système d’Information

Dans la grille EGEE/LCG, il existe deux systèmes pour trouver les états de services. Le plus ancien est basé sur LDAP et utilise une hiérarchie de serveurs. Vers le niveau “top” des serveurs, les utilisateurs peuvent trouver l’état de tous les services. Le plus récent est R-GMA (Relational Grid Monitoring Architecture). R-GMA utilise un modèle producteur-consommateur qui est plus distribué. Actuellement, R-GMA est principalement utilisé pour le monitoring des jobs et pour la comptabilité.

Le schéma des informations s'appelle "Glue". On peut trouver la norme⁶ sur le web.

3.1 Système basé sur LDAP

3.1.1 Usage Direct du LDAP

1. Allez sur le site web GStat⁷ et visualisez les différentes informations présentées pour chaque site. Cliquez sur un site puis regardez dans la section "Site General Information". Trouvez le site GRIF.
2. En utilisant la commande `ldapsearch`, interrogez le GRIS. Ce dernier tourne sur un "Computing Element" (CE). Utilisez le CE `grid10.lal.in2p3.fr`.

```
ldapsearch -x -h grid10.lal.in2p3.fr -p 2135 \  
-b 'mds-vo-name=local,o=grid'
```

Pour LDAP, chaque entrée possède un identifiant (dn) et une liste d'attributs avec leurs valeurs.

3. De la même façon, interrogez le GIIS tournant sur le site GRIF (utilisez la commande LDAP montrée sur la page web de la première étape). Combien de services différents pouvez-vous distinguer ?

```
ldapsearch -x -h bdii.grif.fr -p 2170 \  
-b 'mds-vo-name=GRIF,o=grid'
```

4. Interrogez le BDII qui tourne sur `grid01.lal.in2p3.fr` de manière à afficher la liste des CE répertoriés, ainsi que les "Storage Elements" (SEs) associés. Visualisez le résultat.

```
ldapsearch -x -h grid01.lal.in2p3.fr -p 2170 \  
-b 'mds-vo-name=local,o=grid' \  
'objectclass=GlueCESEBind' \  
GlueCESEBindCEUniqueID GlueCESEBindSEUniqueID
```

Les commandes pour la gestion de données utilisent ces informations lorsque vous ne spécifiez pas un SE spécifique.

5. Listez les CE répertoriés dans le BDII précédent à l'aide du tag `GLITE-3.0.9`.

```
ldapsearch -x -h grid01.lal.in2p3.fr -p 2170 \  
-b 'mds-vo-name=local,o=grid' \  
'objectclass=GlueSubCluster' GlueChunkKey \  
GlueHostApplicationsSoftwareRunTimeEnvironment: GLITE-3_0_9
```

Cet attribut est utilisé pour trouver les sites équipés avec les logiciels spécifiques. Dans les pages GStat, combien de "tags" différents peut-on trouver ?

⁶<http://glueschema.forge.cnaf.infn.it/Spec/V12>

⁷<http://goc.grid.sinica.edu.tw/gstat/>

- Listez ensuite les SE supportant la VO “dteam”.

```
ldapsearch -x -h grid01.lal.in2p3.fr -p 2170 \  
-b 'mds-vo-name=local,o=grid' \  
'objectclass=GlueSATop' GlueChunkKey \  
GlueSAAccessControlBaseRule | grep -B 4 dteam
```

Refaites la commande pour la VO vo.u-psud.fr.

3.1.2 La Commande lcg-infosites

La syntaxe de la commande **ldapsearch** est complexe. Il y a un outil **lcg-infosites** qui permet d’avoir les mêmes informations mais avec une syntaxe plus accessible.

- Utilisez cette commande pour trouver tous les “Computing Elements” accessibles à la VO vo.u-psud.fr.

```
lcg-infosites --vo vo.u-psud.fr ce
```

- On peut aussi trouver les “Storage Elements” (se), “LHC File Catalogs” (lfc), et “Resource Brokers” (rb). Visualisez toutes les options avec l’option **-help**.

```
lcg-infosites --vo vo.u-psud.fr se  
lcg-infosites --vo vo.u-psud.fr rb  
lcg-infosites --vo vo.u-psud.fr lfc
```

Pour les recherches simples, la commande **lcg-infosites** fonctionne très bien. La commande **ldapsearch** est plus adaptée aux recherches complexes ou pour cibler une ressource spécifique.

3.2 R-GMA (Relational Grid Monitoring Architecture)

Les informations disponibles dans R-GMA sont accessibles via un browser web et aussi via des lignes de commande. Les informations ne sont cependant pas complètes, mais l’outil est intéressant pour voir les types d’informations dans le schéma.

3.2.1 Browser R-GMA

Une page web permet de chercher les informations dans R-GMA. Utilisez un navigateur pour visualiser cette page⁸.

- La liste sur l’extrême gauche contient les requêtes prédéfinies. Pour voir les services publiés dans le système d’informations, cliquez sur le lien “GlueServices” (extrême gauche). Puis cliquez sur le bouton “Query”. La liste inclut les services R-GMA et quelques autres types de services. Si

⁸<https://grid08.lal.in2p3.fr:8443/R-GMA/>

rien ne s'affiche , ressayer avec un "timeout" plus grand, par exemple 15 secondes.

2. Pour voir les CEs dans le système d'information, cliquez sur "Table Sets" (extrême gauche) puis sur "GlueCE". Dans la liste "SELECT" on peut voir les champs de la table GlueCE. Par défaut, tous les champs sont sélectionnés. Cliquez sur le bouton "Query" pour visualiser les informations. Encore on peut-être doit utiliser un "timeout" plus grand.

3.2.2 R-GMA : la ligne des commandes

La ligne de commande offre la même fonctionnalité. Dans grid11.lal.in2p3.fr tapez la commande **rgma**. Cette commande démarre une console R-GMA qui permet d'interroger le système. La syntaxe pour les requêtes est très proche du SQL.

1. On peut visualiser l'aide avec la commande **help** et des exemples avec **help examples**.
2. Pour voir toutes les tables d'informations tapez :

```
|| show tables
```

Pour la description d'une table, utilisez la commande **describe** avec le nom de la table :

```
|| describe GlueService
```

Changez "GlueService" par le nom de table vous voulez voir.

3. Pour faire la liste des services utilisez la commande :

```
|| select * from GlueService
```

Pour voir les services dans un site (par exemple GRIF), utilisez la commande :

```
|| select * from GlueService where GlueSite_UniqueId="GRIF"
```

4. Pour voir les informations dans la table "GlueCE", utilisez une commande similaire :

```
|| select * from GlueCE
```

Comme précédemment, on peut limiter la réponse aux entrées intéressantes à l'aide du mot clé "where". On peut aussi lister les champs désirés.

```
|| select UniqueId,FreeCpus from GlueCE where FreeCpus > 10
```

Pour sortir, tapez la commande **quit**.

4 Soumission des Jobs

Tous les fichiers nécessaires se trouvent dans le répertoire *tutorial_material* que vous avez récupéré.

4.1 Suivi d'un job "Hello World"

1. Si ce n'est pas déjà fait, créez un proxy à l'aide de la fonction **voms-proxy-init**.
2. La description du job la plus simple possible se trouve dans le fichier *HelloWorld.jdl*. Le format de ce fichier est le couple clés/valeurs définis.
3. Soumettez le job *HelloWorld.jdl* en utilisant la commande **edg-job-submit**. Les commandes "workload management" comprennent les proxies VOMS et utilisent la VO indiquée dedans. On peut aussi spécifier la VO avec l'option "-vo".

```
|| edg-job-submit HelloWorld.jdl
```

La commande envoie une réponse contenant une URL qui débute par "https", quand elle se passe bien. Cet URL est l'identifiant du job ou "jobid" qui permet de manipuler le job.

4. Vérifiez le statut du job en utilisant la commande **edg-job-status**.⁹

```
|| edg-job-status <jobid>
```

Suivez les modifications de l'état du job jusqu'à l'état "Done (Success)". Combien d'états différents pouvez-vous distinguer ?

Si le job se termine dans l'état "Aborted", c'est qu'il ne passe pas. On peut trouver plus d'informations avec la commande **edg-job-get-logging-info**. Cette commande requiert le jobid.

5. Lorsque le job est terminé ("Done (Success)"), récupérez les données générées à l'aide de la commande **edg-job-get-output**. Le répertoire utilisé pour stocker les résultats dépend de la configuration du site. Au LAL, le répertoire est *~/JobOutput/*. Ce répertoire doit exister ; créez-le si nécessaire.

```
|| edg-job-get-output <jobid>
```

La réponse indique le répertoire qui contient les résultats du job.

6. Vérifiez que tout s'est déroulé correctement en consultant les fichiers *std.out* et *std.err*. Le fichier *std.err* doit être vide et *std.out* doit avoir "Hello World".

Cette procédure simple est utilisée pour le suivi de tous les jobs. Il est possible d'utiliser les options "-o", "-i", et "-noint" qui permettent de gérer un

⁹Sur linux on peut utiliser la commande **watch** pour exécuter une commande dans une boucle. Pour effectuer la commande **edg-job-submit** toutes les 5 secondes, utilisez **watch -interval 5 edg-job-status <jobid>**. Touchez ctrl-c pour sortir.

plus grand nombre de jobs avec les commandes **edg-job-***. Regardez les informations avec l'option **"-help"** ou la commande **man**.

Il existe également des APIs (java, c, c++) qui permettent la gestion des jobs vers un programme compilé. Ce tutorial ne couvre pas ces APIs.

4.2 Fichier JDL : modification et édition

1. Modifiez le fichier *HelloWorld.jdl* de manière à ce qu'il n'appelle plus **/bin/echo** mais le script *HelloWorldScript.sh*. Pour cela :
 - la ligne "Executable" doit être "HelloWorldScript.sh",
 - la ligne "Argument" peut rester avec "Hello World",
 - vous devez de plus définir le paramètre "InputSandbox".

Tous les fichiers listés dans "InputSandbox" sont transférés avec le job.

La syntaxe de cette ligne est :

```
InputSandbox = {"HelloWorldScript.sh"};
```

Si vous soumettez plusieurs jobs avec les mêmes fichiers d'entrée, une nouvelle copie de chaque fichier est créée.

Exécutez le job et vérifiez que tout fonctionne.

On peut utiliser n'importe quel script, cependant le shell utilisé par le script (indiqué dans la ligne "#!") doit exister dans le "Worker Node".

2. Modifier de nouveau *HelloWorld.jdl* de manière à ce qu'il appelle cette fois l'exécutable *myhostname*. Vous pouvez visualiser la source de cet exécutable, qui est un programme C : *myhostname.c*.

Vous n'avez cette fois pas besoin de définir d'argument. Il faut modifier la ligne "InputSandbox".

Exécutez le job et vérifiez que tout fonctionne. Sur quel ordinateur a tourné votre job ?

3. L'exécution d'un programme en C compilé n'est pas forcément pratique : l'exécutable peut être d'une grande taille, dépendre de plusieurs fichiers, ou dépendre d'un environnement d'exécution particulier.

Une solution consiste à compiler le programme directement sur le CE.

Modifier une nouvelle fois *HelloWorld.jdl* de manière à ce qu'il appelle le script *buildandrun.sh*, avec pour argument "myhostname". Testez ce script seul pour comprendre l'argument nécessaire.

Exécutez le job et vérifiez qu'il fonctionne toujours. Votre job a-t-il tourné sur le même ordinateur que précédemment ?

4.3 Pour comprendre les "Requirements" and "Rank"

Il y a deux clés très importantes dans les fichiers JDL : "Requirements" et "Rank". Les valeurs pour les clés Requirements et Rank sont des expressions. Votre job va tourner uniquement sur une ressource qui a une valeur "true" pour l'expression Requirements. S'il y a plusieurs ressources qui ont une valeur "true", le système utilise l'expression Rank pour choisir la meilleure ressource. La ressource qui a la plus grande valeur est choisie. S'il y a plusieurs ressources

avec la même valeur Rank, la ressource utilisée est choisie aléatoirement entre ces ressources de même valeur Rank.

On utilise la VO "dteam" ici pour voir plus de ressources. La VO vo.u-psud.fr est autorisée uniquement sur les deux sites de la GRIF. Réinitialisez votre proxy avec la commande **voms-proxy-init** mais *n'utilisez pas* l'option "-voms".

1. Plusieurs valeurs peuvent être utilisées pour définir les expressions Requirements et Rank. Par exemple, ajoutez l'expression ci-dessous dans le fichier *HelloWorld.jdl* pour choisir tous les sites qui permettent à un job d'utiliser plus d'une heure de temps CPU.

```
Requirements = (other.GlueCEPolicyMaxCPUTime > 60);
```

Pour voir la liste des ressources acceptables utilisez la commande **edg-job-list-match** et l'option "-vo dteam".

Combien de ressources autorisent les jobs qui utilisent plus d'une heure de temps CPU ? Plus de deux heures ? Plus de 10000 minutes ?

2. Pour choisir des sites spécifiques, on peut utiliser le nom de la ressource. Pour choisir tous les sites en France, changez la valeur de Requirements comme suit :

```
Requirements = RegExp(".*\.fr:.*", other.GlueCEUniqueID);
```

Le premier argument est une expression régulière. Combien y a-t-il de sites en France ? Combien de sites en France autorisent les jobs qui utilisent plus d'une heure de temps CPUs ? Modifiez la ligne Requirements pour choisir une ressource. Trouvez la syntaxe correcte pour exclure un site.

Une option "-r" existe pour la commande **edg-job-submit** qui permet de choisir une ressource spécifique. Cependant cette option évite tous le processus "MatchMaking" de Resource Broker et n'ajoute pas le fichier nécessaire (le fichier "BrokerInfo") pour la gestion de données. La technique avec other.GlueCEUniqueID est plus flexible et plus sûre.

3. Ajoutez les lignes suivantes pour utiliser la ressource avec le plus grand nombre des CPUs libres :

```
Rank = other.GlueCEStateFreeCPUs;
```

Utilisez la commande **edg-job-list-match** pour visualiser le résultat. (L'ordre indique le Rank. La ressource la plus intéressante est la première). Si on utilise le valeur "-other.GlueCEStateFreeCPUs", quelle ressource le Broker va t-il choisir ? Que fait le RB si on utilise le valeur "Rank = 1" ?

4.4 L'environnement d'exécution sur le Worker Node

Chaque utilisateur de la grille est mappé dans un compte local pour chaque site. Maintenant l'accès aux ressources locales est contrôlé par les droits de ce compte.

Réinitialisez votre proxy avec l'option "--voms vo.u-psud.fr".

1. Visualisez le contenu du fichier JDL *whoami.jdl*. Lancez le job et récupérez l'output. Visualisez le fichier *std.out*. Sur quel compte êtes-vous mappé ?
2. Visualisez le contenu du script *envvar.jdl*. Soumettez un job qui lance ce script dans la grille. Regardez la liste des variables. Combien de variables concernent la grille ?
3. Ecrivez un job qui liste les versions des logiciels disponibles dans le "Worker Node". On peut utiliser la commande **rpm** pour le faire.

4.5 Soumission d'une job MPI

Beaucoup de disciplines utilisent des jobs parallèles. MPI (Message Passing Interface) est un protocole qui permet la communication entre les tâches parallèles. Les jobs MPIs sont supportés dans la grille. La grille supporte deux installations différentes du MPI : LAM, et OpenMPI. Le support de MPICH et MPICH2 seront bientôt disponibles.

Le logiciel **mpi-start** développé par le projet européen int.eu.grid facilite l'utilisation du MPI dans la grille. L'utilisateur peut définir :

1. Un script à exécuter avant le binaire MPI (par exemple pour compiler le programme),
2. Le programme lui-même, et
3. Un script à exécuter après le binaire MPI (par exemple pour enregistrer les résultats dans un SE).

Le fichier *mpi-hooks.sh* définit les scripts à exécuter avant et après le binaire MPI. Le fichier JDL définit les paramètres comme des nombres de CPU pour le job.

1. Lancez un job avec *mpi-start-lam.jdl*. Regardez la sortie du job. S'il a fonctionné correctement, on pourra lire des lignes telles que :

```
||| Hello world! from processor 3 out of 8
```

2. Changez le nombre de CPUs utilisé par le job. Vérifiez que le job marche toujours. Fonctionne-t-il aussi si vous spécifiez 1000 CPU ?

La grille supporte les jobs MPI sur un seul site. MPI avec esclaves sur plusieurs sites n'est pas supporté. Le support du MPI évolue très rapidement dans la grille en ce moment. Ceci va se stabiliser dans quelques mois.

5 Gestion des Données

Les services qui permettent la gestion des fichiers sont le "replica catalog" (ou "file catalog") et le "storage element" (SE). Le "replica catalog" est une base de données qui contient l'association entre un identifiant et les locations des fichiers dans la grille. Le SE fournit l'espace pour stocker les fichiers.

5.1 Utilisation du Replica Manager

Il existe au moins trois versions différentes du “replica catalog” : le “EDG Replica Catalog”, le “FireMan Catalog”, et le “LHC File Catalog” (LFC). La version du “replica catalog” peut être différente pour chaque VO. Le “EDG Replication Catalog” et le “FireMan Catalog” sont périmés ; la plupart des VOs utilise le LFC.

Les commandes **lcg-*** marchent pour le EDG Replica Catalog et LFC. Le FireMan Catalog utilise d’autres commandes mais fonctionnent comme les commandes **lcg-***.

1. Si votre VO utilise le LFC (comme vo.u-psud.fr), il est nécessaire de définir la variable d’environnement LCG_CATALOG_TYPE. La VO vo.u-psud.fr utilise le LFC.

```
|| setenv LCG_CATALOG_TYPE lfc
```

pour le shell csh

```
|| export LCG_CATALOG_TYPE=lfc
```

pour le shell sh.

2. Visualisez les informations concernant les ressources disponibles pour la VO vo.u-psud.fr.

```
|| lcg-infosites --vo vo.u-psud.fr all
```

Combien de “storage elements” sont disponibles pour la VO vo.u-psud.fr ?
On peut voir plus de ressources pour la VO dteam. Combien de SEs sont disponibles pour la VO dteam ?

3. Créez un fichier texte, et copiez-le sur un SE à l’aide de la commande **lcg-cr**.

```
|| lcg-cr -d grid05.lal.in2p3.fr --vo \voname \  
|| file:'pwd'/toto.txt \  
|| -l lfn:/grid/\voname/myname-toto.txt
```

Remplacez “myname” par votre nom. L’option “-l” (la lettre) donne un nom logique à ce fichier. Pour la VO vo.u-psud.fr, ce nom doit être dans le format : “lfn :/grid/vo.u-psud.fr/...”.

De ce fait, les commandes **lcg-*** ne comprendront pas encore les proxies VOMS et la VO doit aussi être spécifiée.

Si l’option “-d” n’est pas spécifiée, la commande va utiliser le SE défini dans l’environnement. Utilisez **printenv** pour rechercher une variable contenant “PSUD”.

Lorsque la commande se passe bien, un GUID est retourné (Globally Unique Identifier) pour le fichier. Ce GUID peut être utilisé pour manipuler le fichier.

4. Vérifiez que le fichier copié est bien présent à l'aide de la commande **lcg-lr**.

```
lcg-lr --vo vo.u-psud.fr <GUID>
```

Si cela se passe bien, la commande affiche un URL avec un protocole "sfn".

On peut également remplacer le GUID par le nom logique (ou SURL).

```
lcg-lr --vo vo.u-psud.fr lfn:/grid/vo.u-psud.fr/myname-toto.txt
```

Le même URL qu'avant doit s'afficher.

5. On peut utiliser la commande **lcg-lg** pour trouver le GUID pour un nom logique ou SURL. Regardez l'usage de cette commande et vérifiez que le GUID est correct pour votre nom logique.
6. Choisissez un SE dans la liste des SE disponibles, et faites une copie de votre fichier sur ce SE à l'aide de la commande **lcg-rep**. Vérifiez que le fichier a bien été dupliqué avec la commande **lcg-lr**.

```
lcg-lr --vo vo.u-psud.fr <GUID>|<LFN>|<SURL>
```

7. Récupérez localement le fichier ou son replica à l'aide de la commande **lcg-cp**.

```
lcg-cp --vo vo.u-psud.fr <GUID>|<LFN>|<SURL> file:'pwd'/copie.txt
```

Vérifiez que la copie.txt et le fichier original sont identiques.

8. Supprimez ensuite le replica de ce fichier à l'aide de la commande **lcg-del**.

```
lcg-del --vo vo.u-psud.fr <SURL>
```

Vérifiez avec la commande **lcg-lr** que ce replica n'existe plus.

Ajoutez l'option "-a" pour supprimer le fichier de tous les SE. L'entrée correspondante au GUID alors dans le RC (Replica Catalog) est détruite. Cette fonctionnalité n'existera plus dans les sorties futures. *Ne le faites pas maintenant.*

5.2 Regarder les informations dans le LFC

Les commandes client ne trouvent pas automatiquement le serveur LFC pour votre VO. Pour la VO vo.u-psud.fr, il est nécessaire de définir la variable d'environnement LFC_HOST.

```
setenv LFC_HOST grid14.lal.in2p3.fr
```

pour le shell csh ou

```
export LFC_HOST=grid14.lal.in2p3.fr
```

pour le shell sh.

Normalement on ne change pas le contenu de ce catalogue manuellement. Si on supprime une entrée dans le catalogue, on peut laisser des fichiers qui seront introuvables dans la grille. Cependant, il peut être intéressant d’y regarder de temps en temps directement les informations et de créer des nouveaux répertoires.

1. Pour trouver les fichiers dans */grid/vo.u-psud.fr*, faites :

```
lfc-ls -l /grid/vo.u-psud.fr
```

On peut voir les permissions, la taille, la date de modification, et le nom avec l’option “-l”. Vous allez trouver votre fichier dans la liste.

2. Le LFC supporte les “Access Control Lists” (ACL). Pour regarder l’ACL pour un fichier ou un répertoire :

```
lfc-getacl /grid/vo.u-psud.fr
```

A chaque répertoire correspond deux ACLs : l’ACL pour le répertoire et une ACL par défaut. L’ACL par défaut est appliqué aux nouveaux fichiers du répertoire.

Normalement les commandes **lcg-*** mettent les bonnes permissions dans les fichiers et répertoires.

Le LFC utilise les ACLs mais malheureusement cette fonctionnalité n’est pas très utilisable car l’accès est toujours lié au compte unix. On ne peut donc pas encore utiliser les ACLs pour contrôler finement l’accès aux fichiers.

3. On peut créer un nouvelle arborescence avec **lfc-mkdir**. On doit alors vérifier que le nouveau répertoire possède les bons droits. Créez un nouveau répertoire et copiez-y un fichier. Ensuite utilisez **lcg-del** avec l’option “-a” et **lfc-rm** pour nettoyer le LFC.

5.3 Accéder depuis un job à des données stockées

1. Regardez les fichiers *InputData.jdl* et *InputData.sh*. Dans *InputData.jdl* il y a deux nouvelles lignes qui permettent d’ordonner les jobs près des fichiers de données :

```
DataAccessProtocol = {"file","gridftp","rfio"};  
InputData = "lfn:<LFN>";
```

Les protocoles sont les moyens par lesquels votre exécutable peut lire ce fichier. Tous les SEs supportent le protocole gridftp. Les autres ne sont pas obligatoires.

2. Modifiez le fichier *InputData.jdl*. Mettez le nom logique de votre fichier.

3. Affichez la liste des sites sur lesquels le job peut être soumis. Pourquoi est-elle si réduite ?
4. Lancez ce job et vérifiez que le fichier est bien lu.

5.4 Stockez des données depuis un job

1. Regardez le script *OutputData.sh* et le fichier *OutputData.jdl*. Ce job crée un fichier *toto.txt* et la grille transfère ce fichier dans un SE.
2. Changez le LFN dans le fichier JDL et lancez le job.
3. Regardez les fichiers de sortie. En particulier, le fichier *DSUpload*.out*. Pouvez-vous récupérer ce fichier dans grid11.lal.in2p3.fr ?