



La grille de calcul EGI

Jérôme PANSANEL

jerome.pansanel@iphc.cnrs.fr

Formation Administrateur Grille – Mai 2017 – LPNHE

Sommaire

- Introduction
- Fonctionnement des infrastructures
- Distribution du middleware

- Introduction
- Fonctionnement des infrastructures
- Distribution du middleware

Les grilles de calcul

Les grilles de calcul permettent de répartir les calculs et le stockage des données sur un ou plusieurs systèmes informatiques hétérogènes et distribués géographiquement.

→ *Station de travail* 

→ *Cluster de calcul* 

→ *Grille de calcul*



Quelques exemples de grille de calcul

- EGI - <http://www.egi.eu> (multidisciplinaire) ;
- Grid5000 - <http://www.grid5000.fr> (multidisciplinaire) ;
- Décrypton – <http://www.decrypthon.fr> ;
- World Community Grid – <http://www.worldcommunitygrid.org> ;
- GPUGRID – <http://www.gpugrid.net> ;
- SETI@home – <http://setiathome.berkeley.edu> ;
- ...

La grille de calcul EGI

Grille EGI

- Infrastructure pluridisciplinaire issue de WLCG en 2001
- Ressources distribuées dans 42 pays
- Une organisation officielle

Constitution

- Ressources (serveurs de calcul et de stockage) reliées par un réseau (Internet) ;
- Intergiciel (*middleware*) permettant l'interopérabilité des différents sites de grille
- Ensemble d'applications pour effectuer les calculs
- Services associés (surveillance, sécurité, *accounting*, ...)

EGI – <http://www.egi.eu>

Avantages

- Les données peuvent être copiées sur différents sites
- Les sites sont répartis sur plusieurs zone horaire permettant d'avoir toujours un expert en ligne
- Pas de « single points of failure »
- Les ressources gérées indépendamment ont encouragées l'émergence de nouvelles approches pour le calcul et le stockage
- Les chercheurs peuvent toujours accéder aux ressources, où qu'ils soient
- La grille est facilement extensible, permettant de répondre aux nouveaux besoins des utilisateurs
- Gestion des utilisateurs unifiés par certificat

Contraintes

- Qualité du réseau primordiale
- Abstraction et allocation des ressources grâce à un intergiciel (*middleware*)
- Sécurisation et unification des accès
- Gestion de l'utilisation des ressources / regroupement par VO
- Utilisation de standards ouverts imposés
- Structure coordonnant au niveau national les nœuds de grille en mode production (disponibilité > 90%)
- Gestion spécifique des licences logiciels

Organisation européenne : EGI

- Fondation créée en 2010 (à l'origine European Grid Infrastructure)
- Basée sur des projets européens :
 - 2001 → 2004 : Datagrid
 - 2004 → 2010 : EGEE
 - 2010 → 2014 : EGI-InSPIRE
 - 2015 → 2017 : EGI-Engage
- Coordonne les opérations au niveau européen (deux réunions mensuelles)
- Coordonne l'évolution du *middleware* (UMD, CMD)
- Organise le support aux utilisateurs
- Promotion du calcul et stockage distribués

Organisation nationale : France Grilles

- GIS créé en 2010 (NGI France)
- Établir et opérer une infrastructure nationale de grille de production, pour le traitement et le stockage de données scientifiques massives
- Animé par l'IdGC (jusqu'en 2017)
- Repris dans une future IR (avec FR-T2)
- Coordonne les opérations (réunion mensuelle)
- Support aux utilisateurs en France
- Lien entre les sites et EGI
- Liste de discussion :
<http://listserv.in2p3.fr/cgi-bin/wa?A0=operations-l>

France Grilles – <http://www.france-grilles.fr>

Organisation virtuelle (VO)

- Un ensemble d'individus et d'institutions organisé autour d'une thématique scientifique ou d'une zone géographique (enmr.eu, biomed, compchem, ...)
- Être membre d'une VO, c'est pouvoir accéder à un volume de ressources négociées sur la grille
- Le 10/05/2017, plus de 740.000 CPUs et 500 Po de stockage sont connectés à la grille EGI
- VO des opérations : ops
- VO dev : dteam
- VO nationale : `vo.france-grilles.fr`

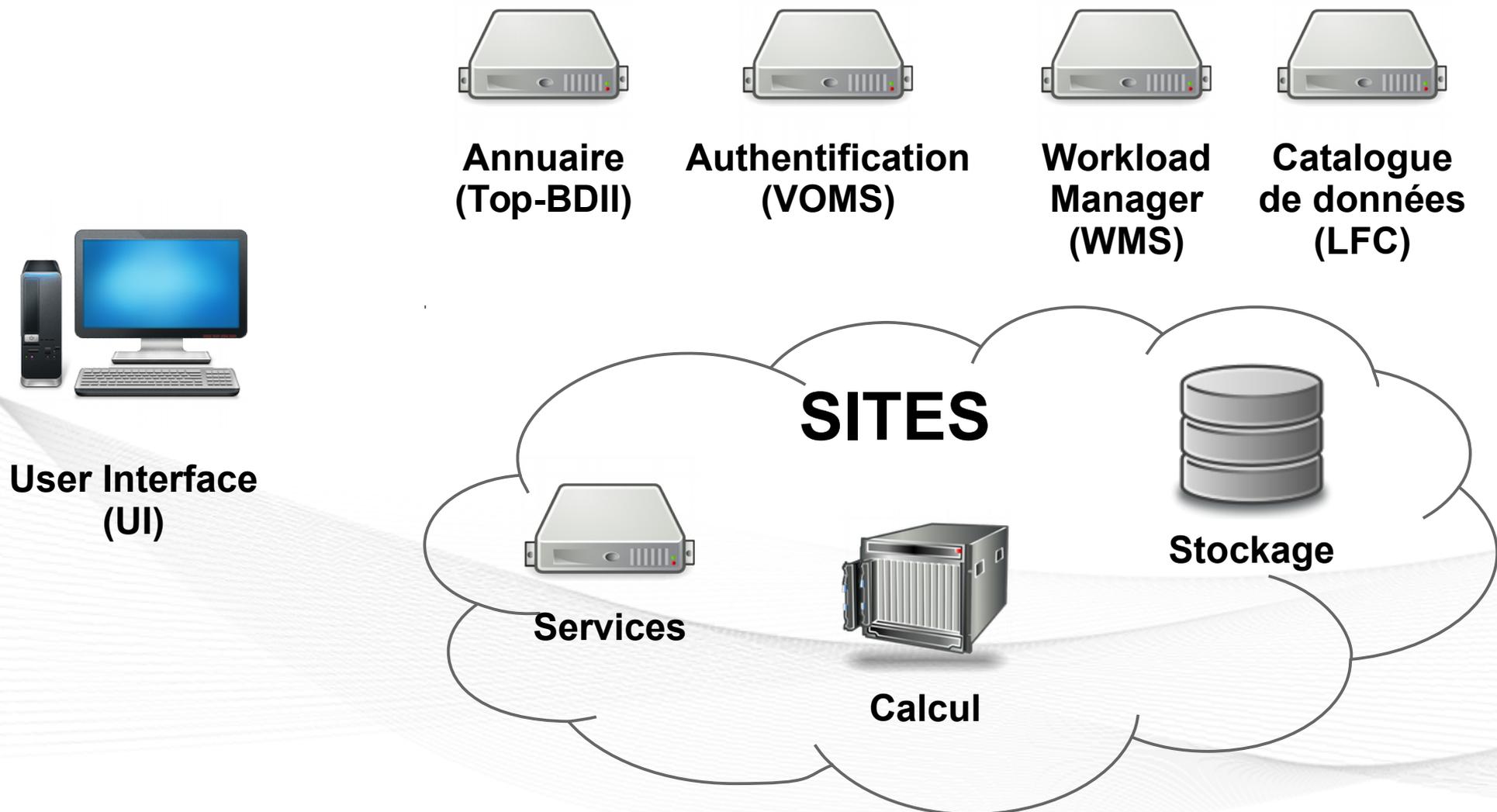


- Introduction
- **Fonctionnement des infrastructures**
- Distribution du middleware

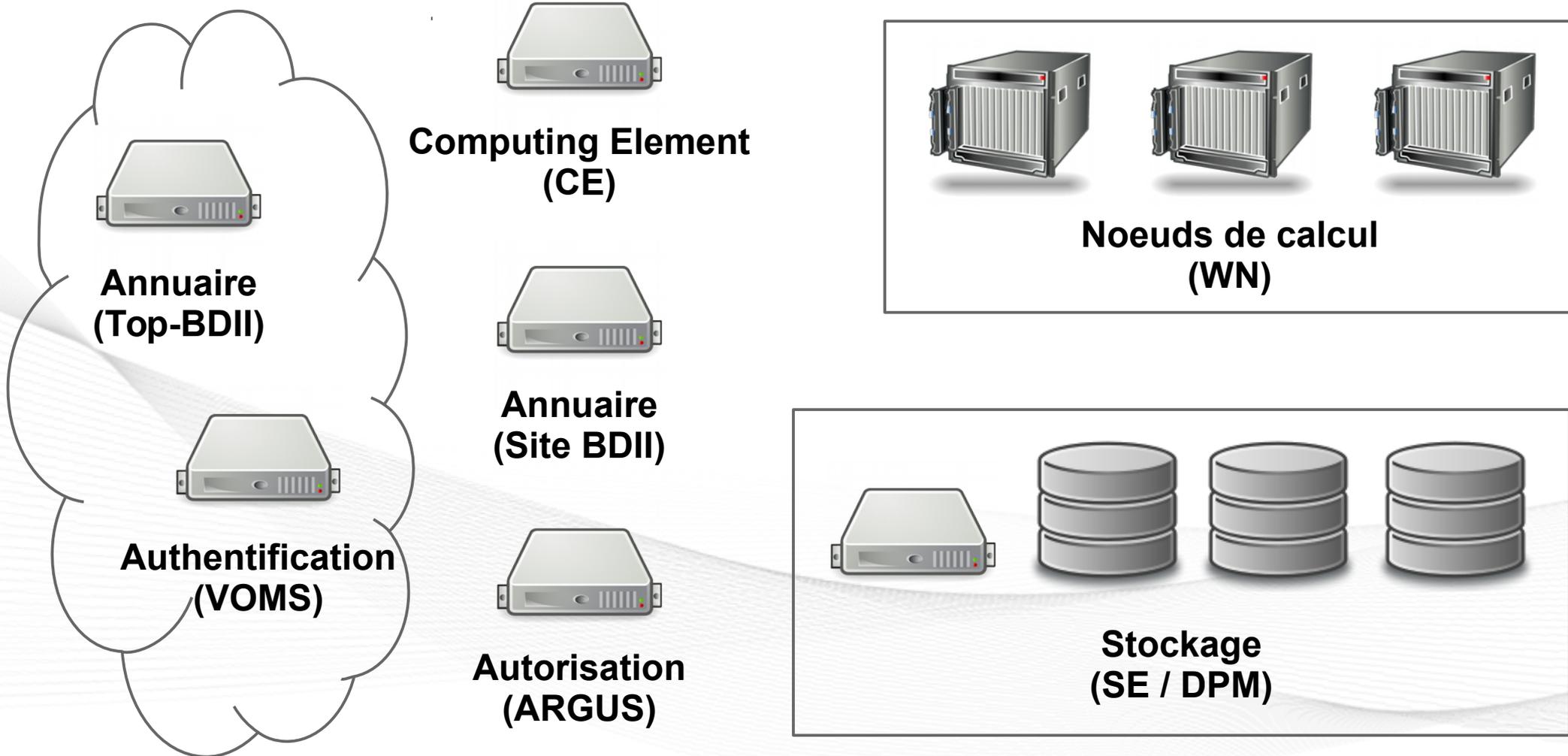
Le *middleware*

- Nécessaire pour que les différents sites puissent fonctionner ensemble
- Voir les sites comme une entité unique pour l'utilisateur
- Pas de middleware imposé
- Mais les services doivent être visibles (publication GLUE 1.3 / 2.0)
- gLite, ARC, UNICORE, QCG, Globus, ...
- → UMD

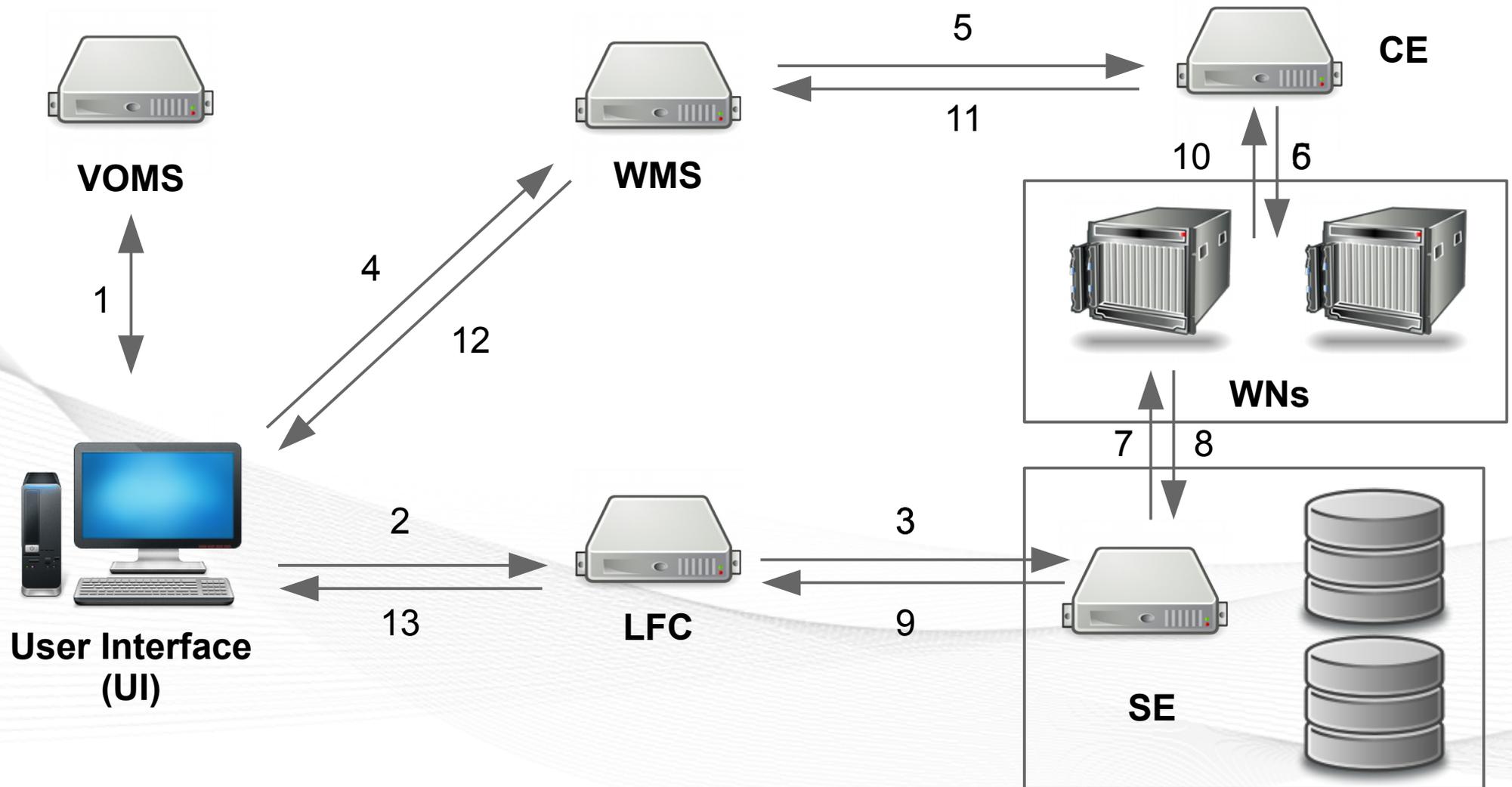
Middleware de grille (utilisateur)



Middleware de grille (site)



Cycle de vie d'un job



User Interface (UI)

- Ensemble d'outils pour la gestion des calculs et du stockage pour les utilisateurs
- Peut être configuré pour fonctionner avec plusieurs VOs
- Pré-configuré avec un WMS et un SE (*Close SE*).
- Méta-paquet : emi-ui
- Disponible pour SL6 (bientôt CentOS 7)
- Maintenance peu active
- Disponible en VM :
<http://grand-est.fr/resources/VM/FG-VirtUI-20161110.zip>
- Documentation :
<http://grand-est.fr/index.php/fr/support/documentation/6-installation-et-configuration-dune-interface-utilisateur-glite>
<http://grand-est.fr/index.php/fr/support/documentation/7-gestion-des-calculs-avec-les-outils-glite>

Workload Manager (WMS)

- Jusqu'à présent, assuré par le composant emi-wms
- Service proposé par les sites
- Plus de mises à jour depuis 2014
- En cours de remplacement vers DIRAC (<http://diracgrid.org/>)
- WMS développé à l'origine pour LHCb
- DIRAC est hébergé par les NGI ou par EGI
- Service France Grilles :
<http://www.france-grilles.fr/catalogue-de-services/fg-dirac/>



VOMS

- Authentification et autorisation de la grille EGI par certificat
- Le serveur VOMS permet d'ajouter une extension au proxy spécifique à la VO
- Permet d'attribuer un rôle et un groupe à un utilisateur
- Évolution vers CheckIn et OpenID au niveau d'EGI
- Service opéré pour les VOs par quelques sites seulement (GRIF, CC-IN2P3)
- Dernière version : v3.6.0 (mise à jour de sécurité critique)
- Documentation :
<http://italiangrid.github.io/voms/index.html>

LFC

- Méta-catalogue de fichiers
- Permet de faire le lien un nom de fichier logique et les réplicats disponibles
- Un serveur LFC par VO
- Une API et une base de données
- Maintenu dans le cadre du projet DPM (CERN) en mode *best-effort*
- Dernière version : v1.9.0
- Documentation :
<https://svnweb.cern.ch/trac/lcgdm/wiki/Lfc> (pas à jour)

Annuaire des services (BDII)

- Base de données LDAP peuplée par un outil externe (utilisation de fichiers LDIF) → dynamique
- Trois niveaux :
 - Ressource
 - Site
 - Top
- Élément essentiel de la grille de calcul
- Documentation :
<http://gridinfo.web.cern.ch/developers/resource-bdii>

Gestion de l'autorisation (ARGUS)

- Gestion de l'autorisation au niveau d'un site
- Système hiérarchique EGI → NGI → Site pour pouvoir bannir un utilisateur en central
- Fonctionne bien avec les Ces (CREAM, ARC et HTCONDOR)
- Pour DPM, il faut une colle :
https://www.gridpp.ac.uk/wiki/DPM_Argus_Integration
- Dernière version : 1.7.0 (la v1.7.1 ne va pas tarder)
- Maintenu par INFN, NIKHEF, CERN, IN2P3
- Documentation :
<http://argus-documentation.readthedocs.io/en/latest/>
<https://github.com/argus-authz>
https://www.gridpp.ac.uk/wiki/Argus_Server

CE

- Gestion de la soumission vers le système de batch
- Élément critique d'un site de grille
- Traduit le fichier JDL vers les instructions Torque / SLURM / HTCONDOR, ...
- Publie (ressource BDII) l'état actuel du système (queues, nombre de jobs / charge, ...)
- Trois types principaux :
 - ARC CE
 - CREAM CE
 - HT CONDOR CE
- Documentation

<http://www.nordugrid.org/arc/ce/>

<https://wiki.italiangrid.it/twiki/bin/view/CREAM/>

<https://wlcg-ops.web.cern.ch/htcondor-ce>

Systeme de batch

- Gestion des jobs sur les WNs (répartition, cycle de vie, ...)
- Plusieurs systèmes disponibles :
 - Torque / MAUI
 - SLURM
 - HTCONDOR
 - SGE
- Dépasse le scope de la formation

SE

- Gestion du stockage
- Principalement basé sur DPM et XrootD
- Élément critique d'un site de grille
- DPM est maintenu par le CERN
<https://svnweb.cern.ch/trac/lcgdm/wiki/Dpm>
- XRootD est maintenu par le SLAC et le CERN :
<https://github.com/xrootd/xrootd>

- Introduction
- Fonctionnement des infrastructures
- **Distribution du middleware**

Middleware



UMD (1 - 4)

Distribution du middleware

- La distribution du middleware de grille est coordonnée par EGI à travers les distributions UMD (Grille) et CMD (Cloud)
- https://wiki.egi.eu/wiki/UMD_products_ID_cards
- UMD-3 : SL5 (obsolète) et SL6 (uniquement maintenance)
- UMD-4 : SL6 et CentOS 7 (développement en cours)
- Conflit potentiel entre des paquets UMD (prioritaire) et EPEL
- Cycle des releases :
https://wiki.egi.eu/wiki/UMD_Release_Schedule
- Procédure générique d'installation :
http://repository.egi.eu/category/umd_releases/distribution/umd-4/
- Dépôts RPM :
<http://repository.egi.eu/sw/production/umd/4/>

Distribution du middleware

- Un dépôt *preview*
- Paquets *presque* stables
- Documentation :
https://wiki.egi.eu/wiki/Preview_Repository
- Liens :
http://repository.egi.eu/community/software/preview.repository/2.0/releases/centos/7/x86_64/RPMS/