



Introduction aux grilles de calcul et au projet EGI

David Bouvet, Catherine Biscarat

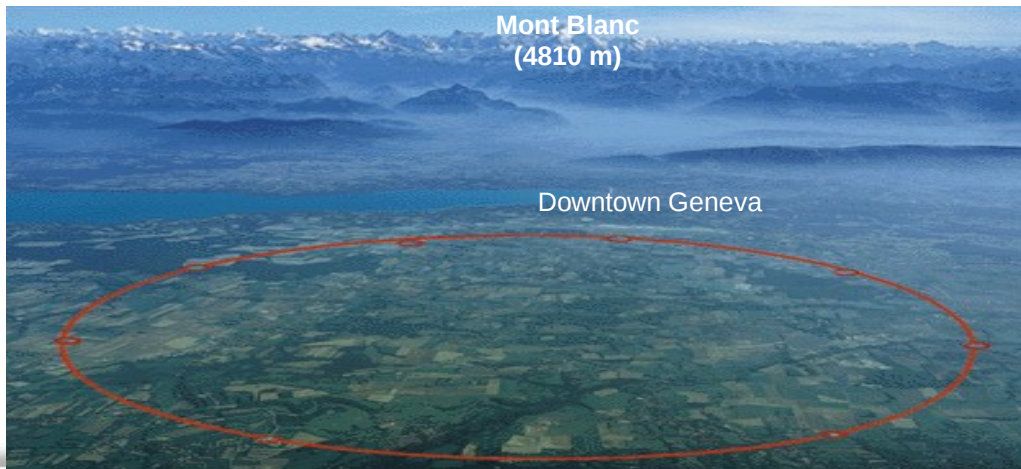
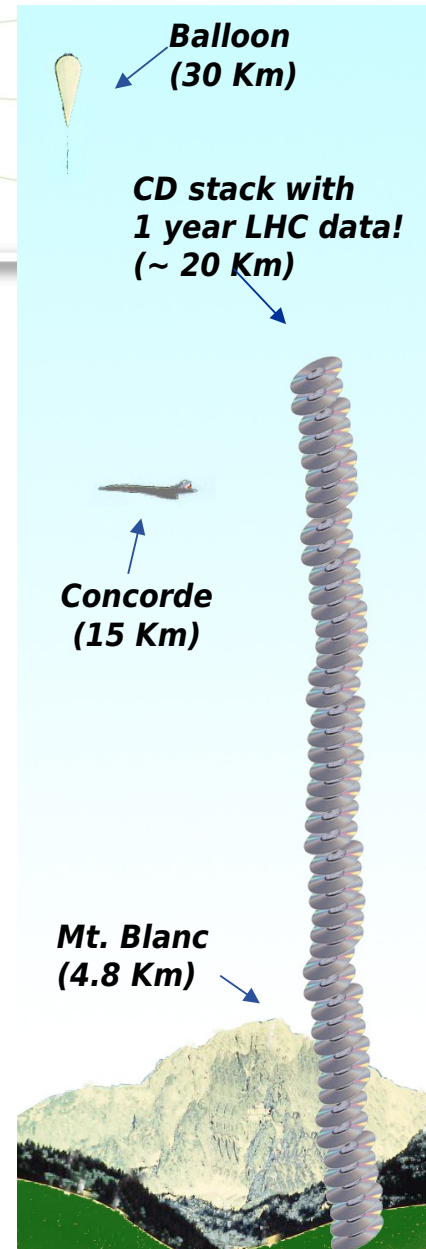
Centre de Calcul – Villeurbanne, 04-05/04/2012

(Basé sur une présentation de David Bouvet et David Weissenbach)

- **Introduction à la grille**
- **Le projet EGI**
- **Le middleware gLite**
 - Vue d'ensemble et architecture

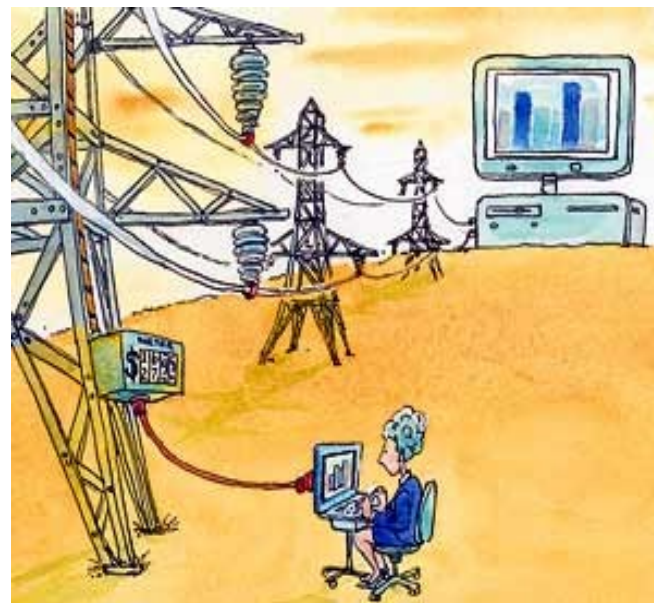
➤ Pourquoi la grille ?

- Les sciences sont devenues très gourmandes en données
 - beaucoup de données produites
 - grandes collaborations internationales
 - ex. Large Hadron Collider (LHC) au CERN :
~10 petabytes/an (~10 Million GBytes)



La solution : la grille ?

- ... ressources distribuées de manière sécurisée (calcul, stockage, etc) de façon que les utilisateurs peuvent collaborer de manière transparente au sein d'Organisations Virtuelles (VO)



... ou les grilleS ?

➤ Différents buts :

- De production (EGI, DEISA)
 - ressources coordonnées par un projet, utilisateurs connus
 - intergiciel commun, applications stables
- De recherche (ALADIN, Grid'5000)
 - ressources coordonnées par un projet, utilisateurs connus
 - tests d'intergiciel
- Bénévole (BOINC)
 - ressources mise à disposition d'un projet, utilisateurs inconnus

... ou les grillesS ?

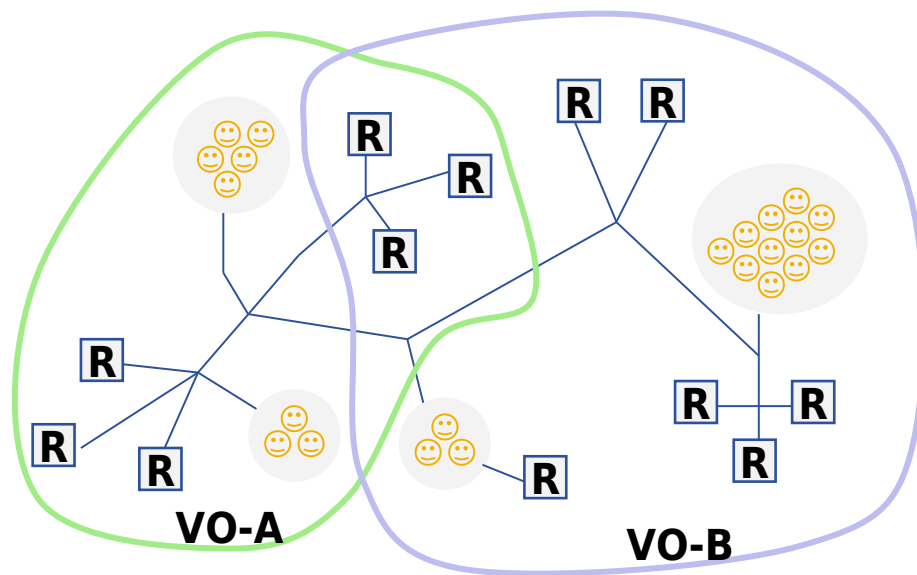
➤ Plus ou moins homogènes :

- desktop grids
 - BOINC / @home
 - XtremOS
- Grilles de service
 - EGI / OSG / NAREGI ...
 - Decryphon
- Grilles HPC
 - DEISA, TeraGrid, ...

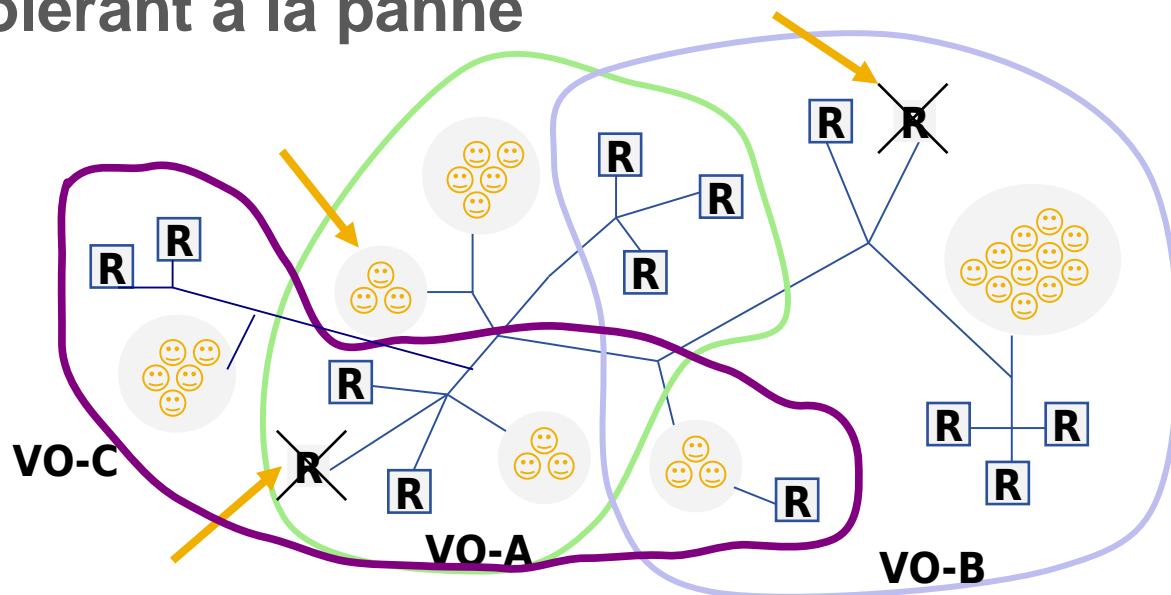
➤ Éventuellement interopérables.

- **Sur beaucoup de grilles, les utilisateurs sont regroupés en Organisations Virtuelles :**
 - Un groupe de personnes de différents instituts travaillant dans un but commun
 - Partageant des ressources distribuées de calcul et de stockage
 - ordinateurs
 - fichiers de données
 - instruments scientifiques
 - codes
 - ...

- Ressources et personnes distribuées
- Reliées par les réseaux, des structures administratives
- Partageant des ressources, buts communs



- Ressources et personnes distribuées
- Reliées par les réseaux, des structures administratives
- Partageant des ressources, buts communs
- Dynamique et tolérant à la panne



➤ Virtual Organization Membership Service (VOMS)

- Amélioration due à EGEE/gLite pour la gestion des VO
- Fournit les informations sur les liens entre les utilisateurs et l'Organisation Virtuelle(VO)
 - appartenance
 - membre d'un groupe
 - rôles de l'utilisateur
- Dispose d'une base de données des « comptes »
 - comparable à un serveur Kerberos
- Donne ces informations dans un format spécifique (VOMS credentials)
- Administration via ligne de commandes & interface web.

- **Connexion unique créant un *proxy* au début de la session**
 - Ajoute les attributs VOMS au *proxy* issu du certificat de l'utilisateur
- **Date d'expiration**
 - Autorisation à durée limitée (peut différer de la durée du *proxy*)
- **Multiples VO**
 - L'utilisateur peut être enregistré dans plusieurs VO
- **Compatibilité descendante**
 - Informations supplémentaires liées à la VO dans le *proxy*
 - Possibilité d'utiliser des services qui ne sont pas liés à VOMS
- **Sécurité**
 - Communication client/serveur sécurisée et authentifiée

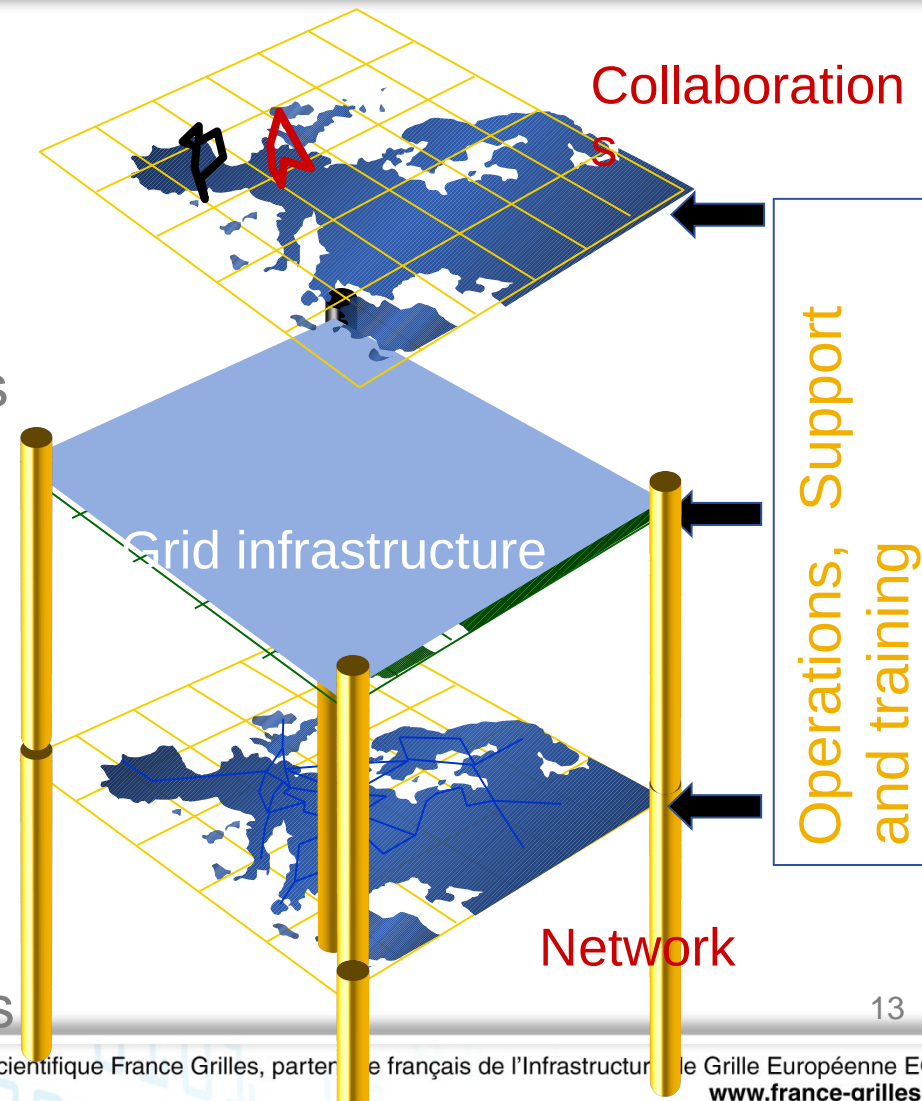
➤ EGEE : Enabling Grids for E-scienceE

- Projet phare d'infrastructure de grille de l'UE
- Financé par la Commission Européenne
- Objectifs principaux
 - infrastructure de grille de service cohérente, robuste et sécurisée
 - amélioration et continuation de l'intergiciel
 - attirer de nouveaux utilisateurs de l'industrie et des sciences
 - s'assurer qu'ils reçoivent une formation et un support de qualité
- Structure
 - plus de 250 centres de calcul dans plus de
 - 48 pays, fédérés en grilles régionales
 - 60.000 CPUs, > 5 Petabytes de stockage

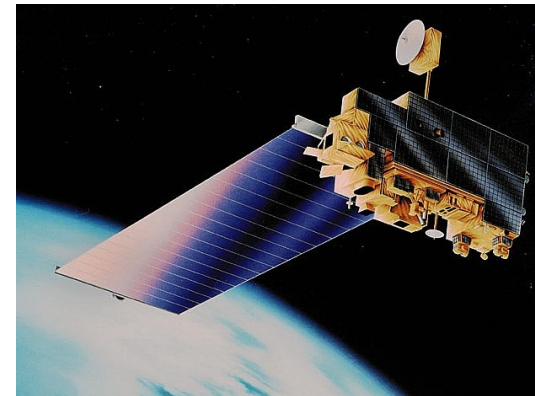
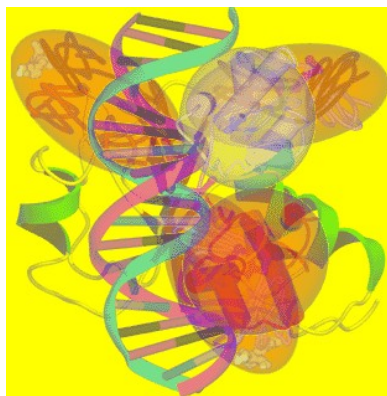


➤ **Construire une grille de production à grande échelle :**

- pour les sciences et technologies internationales
- pour promouvoir la coopération internationale dans la création et l'utilisation de l'e-infrastructure
- reliée à et basée sur les initiatives nationales, régionales et internationales



- Physique des particules
- Bioinformatique
- Industrie
- Astronomie
- Chimie
- Observation de la Terre
- Géophysique
- Biodiversité
- Nanotechnologie
- Modélisation climatique



- **multiparamétrique**
- **déterministe / probabiliste**
- **“Gridifier “ le code traditionnel / conçu pour les grilles**

Mais aussi :

- **standalone / parallèle (MPI)**
- **calculs courts / longs**

- ***data-flow* et *control-flow* souvent complexes**
- ***Grid workflow applications***

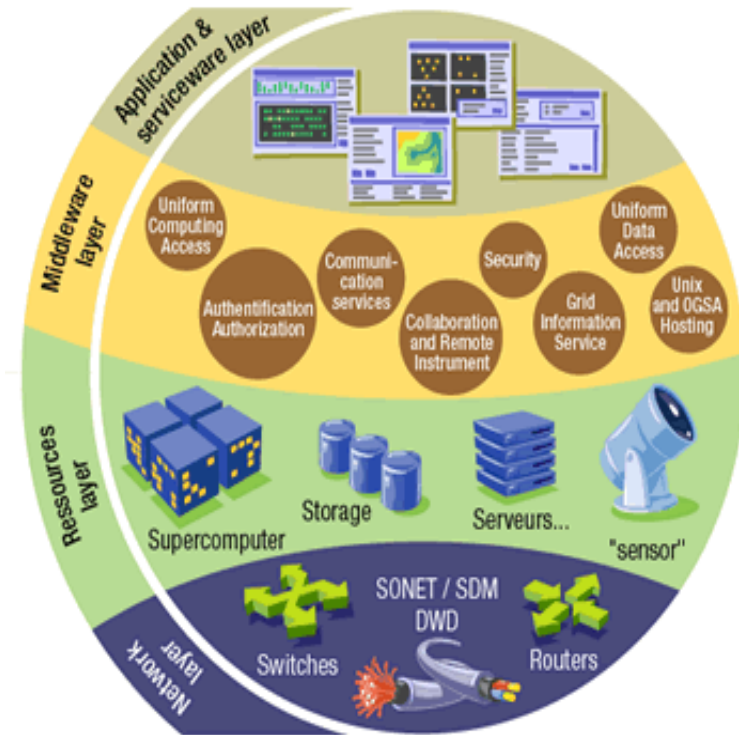
➤ EGI : European Grid Infrastructure

- Suite du projet EGEE, il vise à pérenniser la grille européenne.
 - a débuté en avril 2010
- S'appuie sur les NGI (National Grid Initiative)

➤ La grille repose sur une intergiciel qui fait office d'interface entre les ressources et les applications.

➤ L'intergiciel de la grille

- Trouve la meilleure place pour que l'application s'exécute
- Optimise l'utilisation des ressources
- Organise un accès efficace aux données
- Se charge de l'authentification avec les divers sites utilisés
- Lance les calculs et contrôle leur avancé
- Rapatrie les résultats des calculs



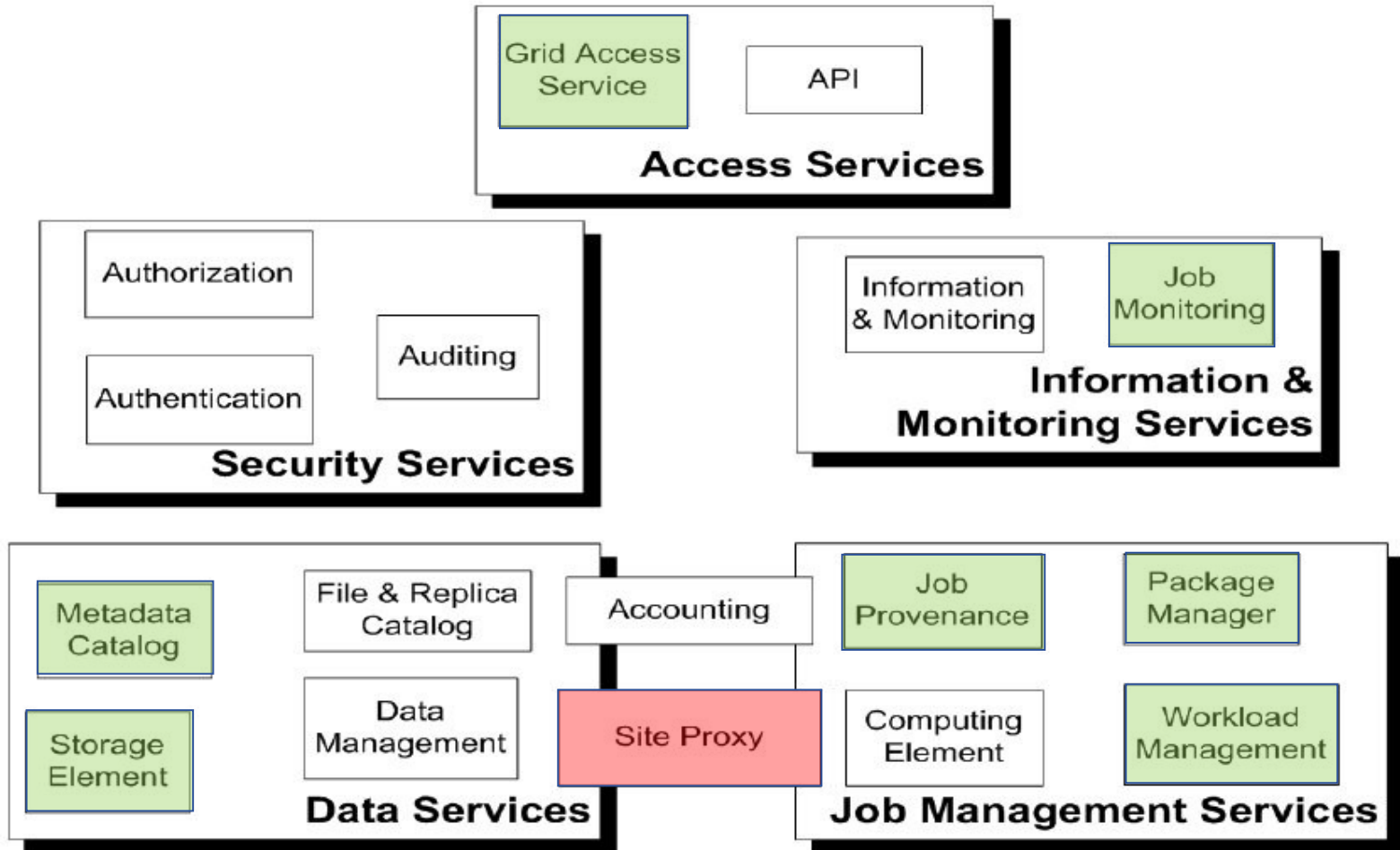
➤ **gLite**

- Première version 2005 (actuellement gLite 3.2) → fin de vie

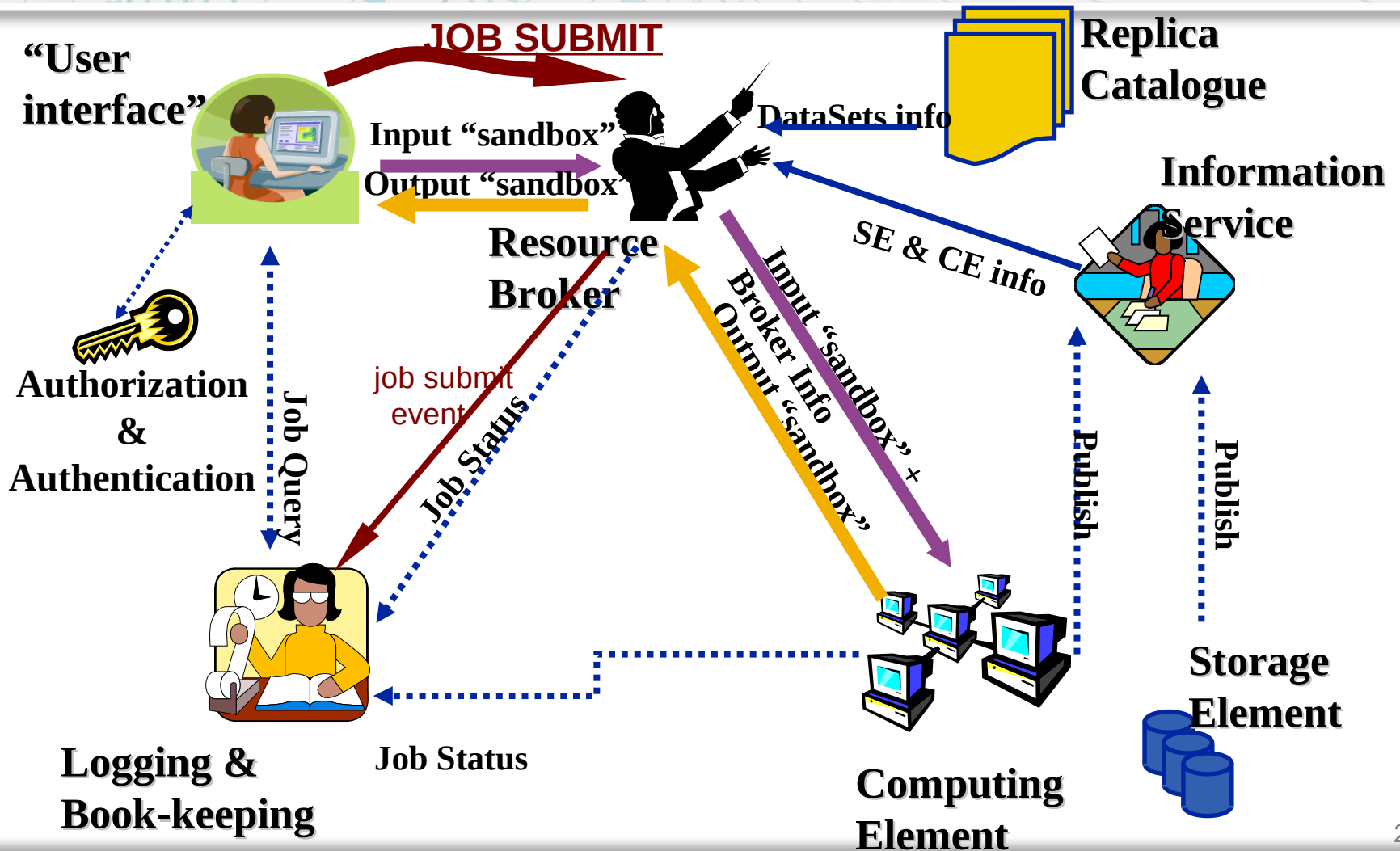
➤ **EMI**

- Première version 2010 : remplaçant de gLite
- Fournit des services ayant une qualité de production
- Développé à partir de composants existants (globus, condor...)
- Interopérable et co-existant avec l'infrastructure déployée
- Robuste : performant et tolérant aux pannes
- Sous licence open source
- Plateforme : actuellement seul Scientific Linux est supportée

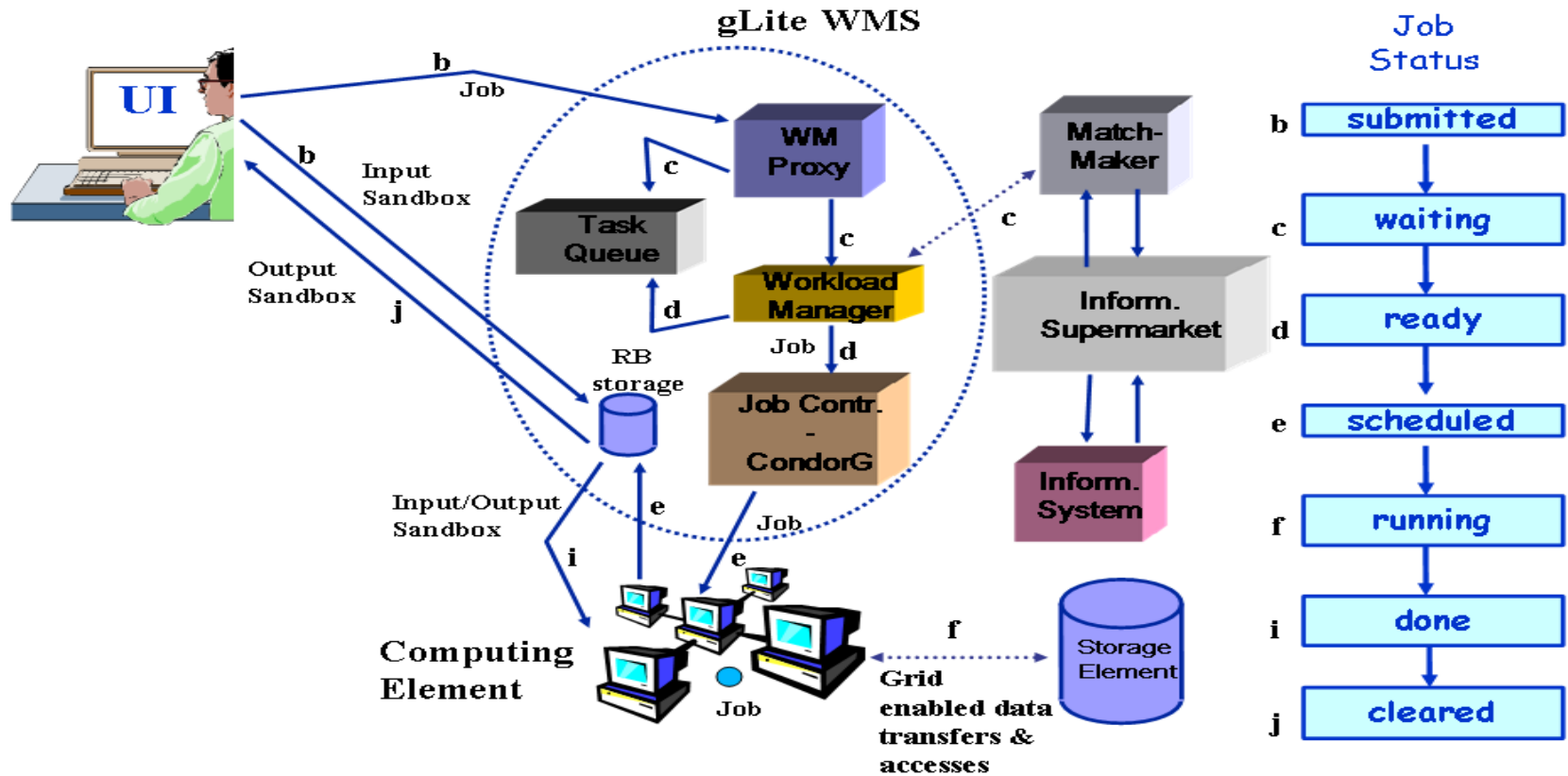
gLite/EMI : services



Glite/EMI : services



WMS : déroulement d'un calcul





Autre services d'intergiciel

- **FTS : File Transfer Service**
- **Gestion de métadonnées (interface aux systèmes de gestion de bases données)**
 - AMGA
 - GRELC / GDSE / OGSA-DAI
- **Alternative au WMS**
 - gridway

➤ Projets grille :

- EGI : <http://www.egi.eu/>
- France : <http://www.france-grilles.fr/> , <http://www.idgrilles.fr/>
- WLCG : <http://lcg.web.cern.ch/lcg/>
- EMI : <http://www.eu-emi.eu>
- BOINC : <http://boinc.berkeley.edu>
- Grid'5000 : <https://www.grid5000.fr>
- DEISA : <http://www.deisa.eu>

➤ Traitement d'incident

- GGUS : <https://ggus.eu>

➤ Activité sur la grille

- <http://gridview.cern.ch>
- <http://gridportal.hep.ph.ic.ac.uk/rtm/>
- <http://gstat-prod.cern.ch/gstat>