

Instituts
thématiques



Inserm

Institut national
de la santé et de la recherche médicale

1

Les grilles et le *Cloud Computing*

PARTIE 1 – La grille : Concepts, architectures et
fonctionnement général

Gilles Mathieu – gilles.mathieu@inserm.fr
Coordination de l'Informatique Scientifique de l'Inserm

Cette présentation...

... fait partie d'une série de 3 :

- **Partie 1 : La grille : Concepts, architectures et fonctionnement général**
- Partie 2 : La grille : Détails de fonctionnement, opérations et outils
- Partie 3 : Le *Cloud Computing* : principes et utilisation



La grille...

- **Généralités**
 - Concepts généraux et paradigmes
 - Historique
 - Panorama actuel
- **Fonctionnement**
 - Architecture globale
 - La vie d'un job : aperçu des services
- **La grille et le monde**
 - Bénéfices et limitations
 - Quelques chiffres et statistiques





Paradigmes et concepts

- **Une grille, c'est quoi ?**
 - Une infrastructure virtuelle constituée d'un ensemble de ressources, géographiquement éloignés mais fonctionnant en réseau pour fournir une puissance globale à des utilisateurs identifiés
- **What's in a name...**
 - Analogie avec le réseau électrique ("Power Grid")
 - *"Accessing computer power from a computer grid should be as simple as accessing electrical power from an electrical grid"*



Paradigmes et concepts

- **La boulette française**
 - Traduire "grid" en "grille"...



- *So much* pour l'image marketing et pour l'analogie avec le "power grid", qui perd son sens au passage



Paradigmes et concepts

- **Les 5 piliers de la grille**
 - Les ressources sont **partagées**
 - L'accès est **sécurisé**
 - L'usage des ressources est **optimisé**
 - La **distance** n'est pas importante
 - Les ressources sont **interopérables** au travers de **standards ouverts**

- *"The grid is about collaboration, about working together"*



Les différents types de grille

- **Grilles de données**
 - Distribution des données pour stockage à grande échelle (Grille de Data Centers)
- **Grilles de calcul**
 - Distribution de la puissance de calcul (Grille de Clusters)
- **Bien souvent, les deux sont couplés...**



Les différents types de grille

- **Grilles de production**
 - Objectif : servir à la production scientifique, ou industrielle
 - Ex: EGI, WLCG
- **Grilles de recherche**
 - Objectif : comprendre et expérimenter le fonctionnement de systèmes distribués
 - Ex: Grid5000



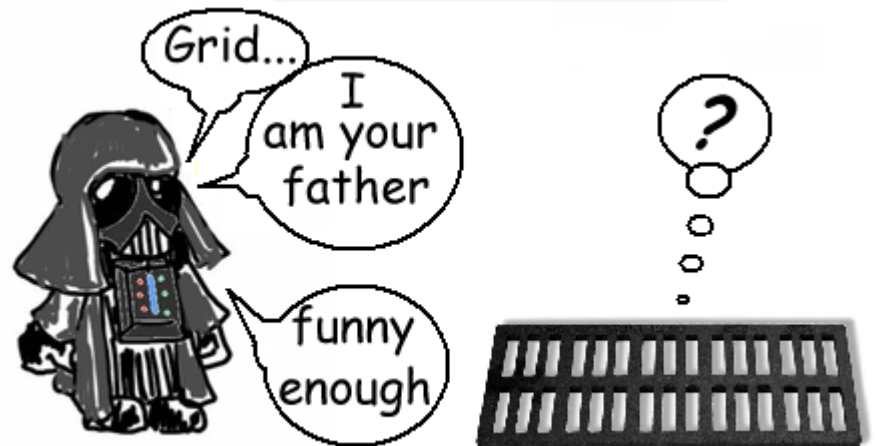
Un peu d'histoire...

- **Le principe n'est pas si nouveau**
 - SETI@home... L'une des premières grilles
 - Le web : une grille d'informations...
- **Les besoins croissants et l'amélioration des réseaux ont rendu ce principe réalisable**



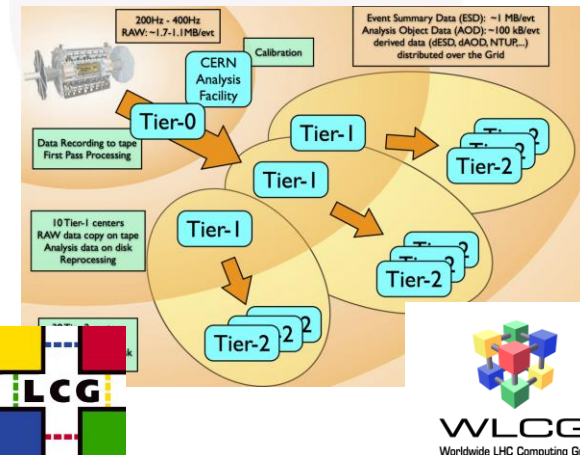
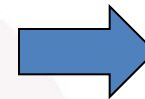
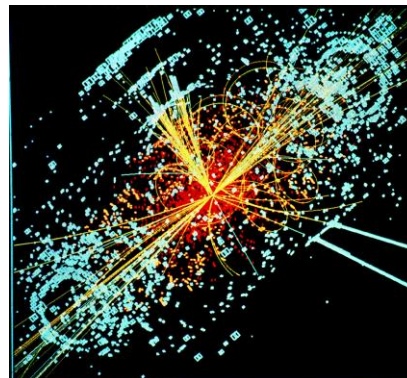
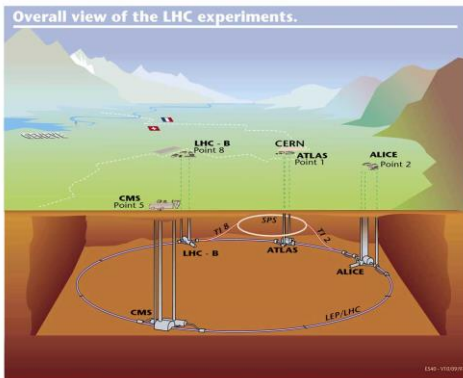
Un peu d'histoire...

- **1999 : Ian Foster et Carl Kesselman, les pères de la grille de calcul**
 - "The Grid: Blueprint for a new computing infrastructure"

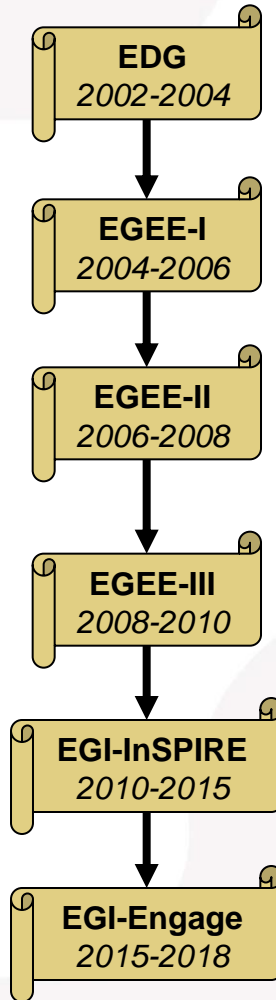


Les grands débuts

- **Fin des années 90: projet de créer le LHC au CERN**
 - Problématique: comment stocker et traiter les 15Po de données produites chaque année?
- **Un pari: créer une "grille" internationale**



Une lignée de projets européens



EDG, l'Ancêtre

EGEE 1^{er}, l'Explorateur

EGEE II, le Bâtitteur

EGEE III, le Fédéraliste

EGI-InSPIRE, l'Empereur

EGI-Engage, le Jeune



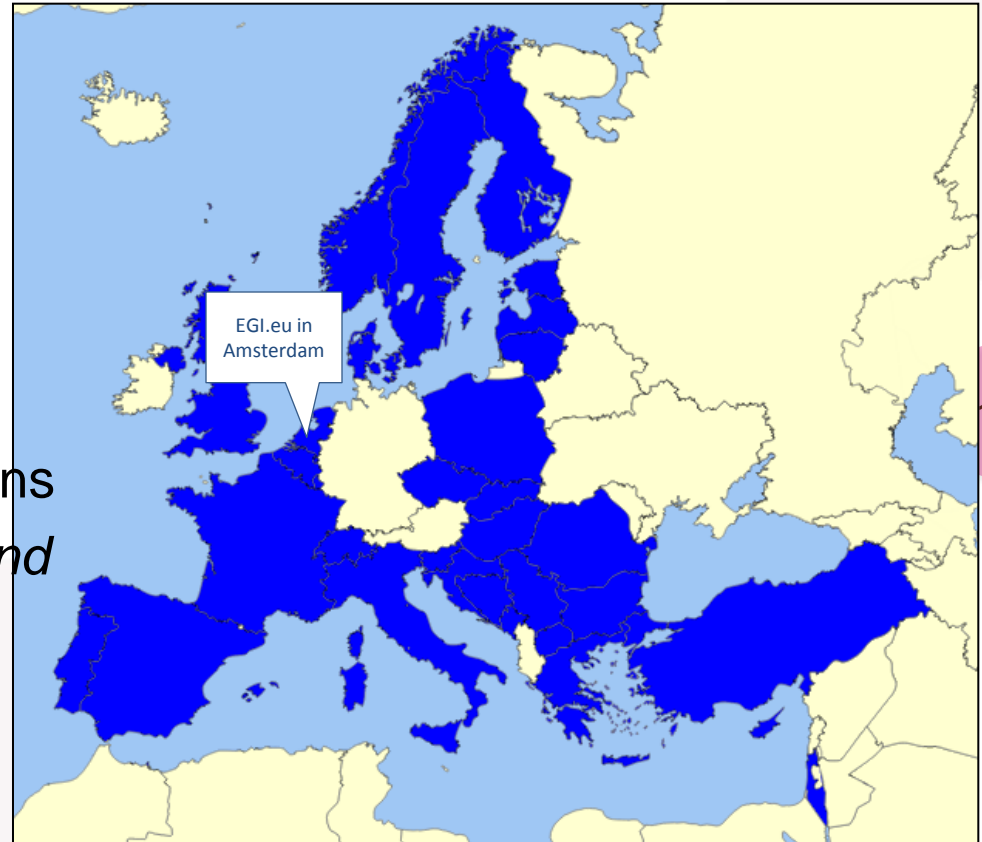
Panorama européen actuel

- **Une grille européenne de production : EGI**
- **Une fédération de grilles nationales : les *National Grid Initiatives* (NGI)**
 - NGI Française = France Grilles
- **Collaborations et interopérabilité bien au-delà de l'Europe**
 - Ressources "mondiales"
 - Accords entre projets et infrastructures



EGI : infrastructure internationale

- **European**
 - 32 countries
(National Grid Initiatives)
- **Grid**
 - **Federating** IT services
- **Infrastructure**
 - Computers, Data, Applications
 - Clouds, clusters, grids,..... *and beyond!!*
- **Sustainability**
 - Sustainable operation
 - Project-driven innovation
 - EGI-Engage, AARC, INDICO, etc.



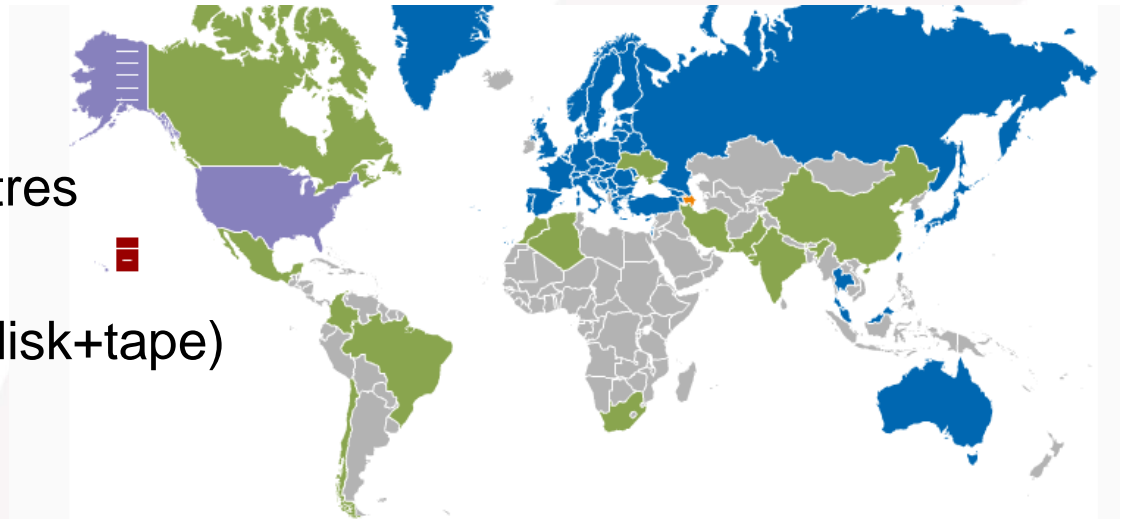
Crédits: G.Sipos (EGI.eu)

EGI

- Coordonnée par une organisation à but non lucratif (EGI.eu) basée à Amsterdam
- Partenariats et interopérations au-delà de l'Europe

- **EGI Resources**

- 364 resource centres
- 630 000 CPUs
- 375 PB storage (disk+tape)



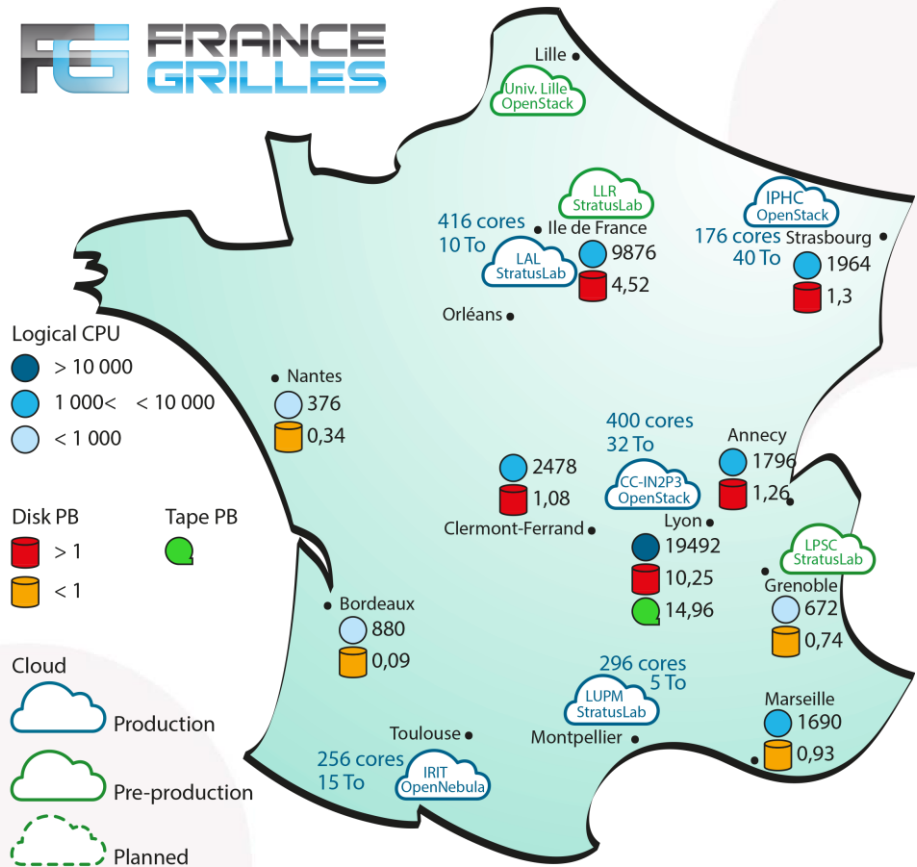
France Grilles (FG)



- **L'infrastructure de grille nationale (NGI)**
 - Grille de production intégrée à EGI
- **Groupement d'intérêt scientifique (GIS)**
 - Créé en juin 2010, renouvelé en 2014
 - 8 partenaires académiques

FG - Resources

- 37k logical CPUs
- 18.7 PB disk storage
- 15 PB tape storage
- 16 sites Grille
- 8 sites Cloud



- Les ressources sont fournies par les communautés et les partenaires



FONCTIONNEMENT GENERAL



Grid overview

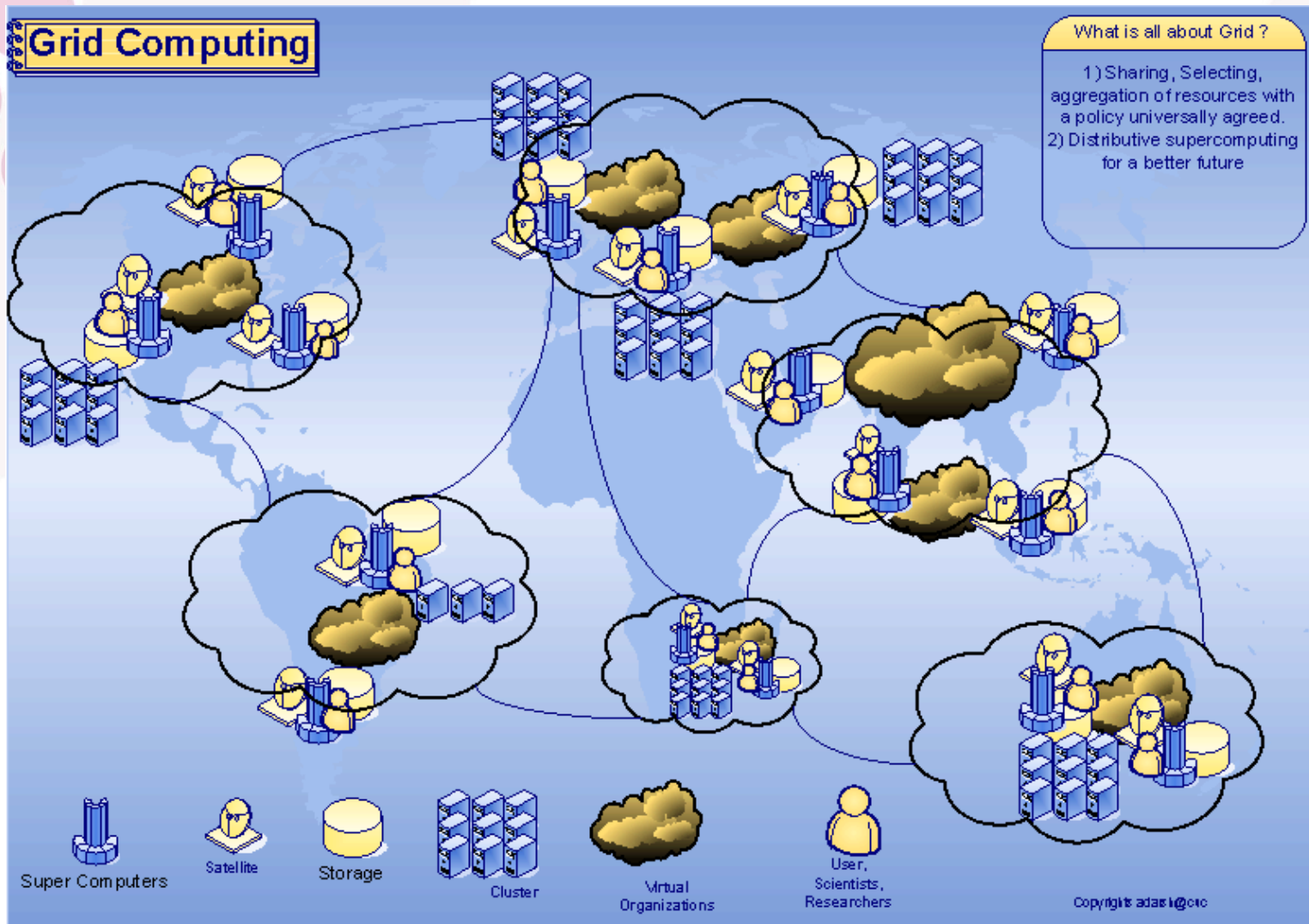
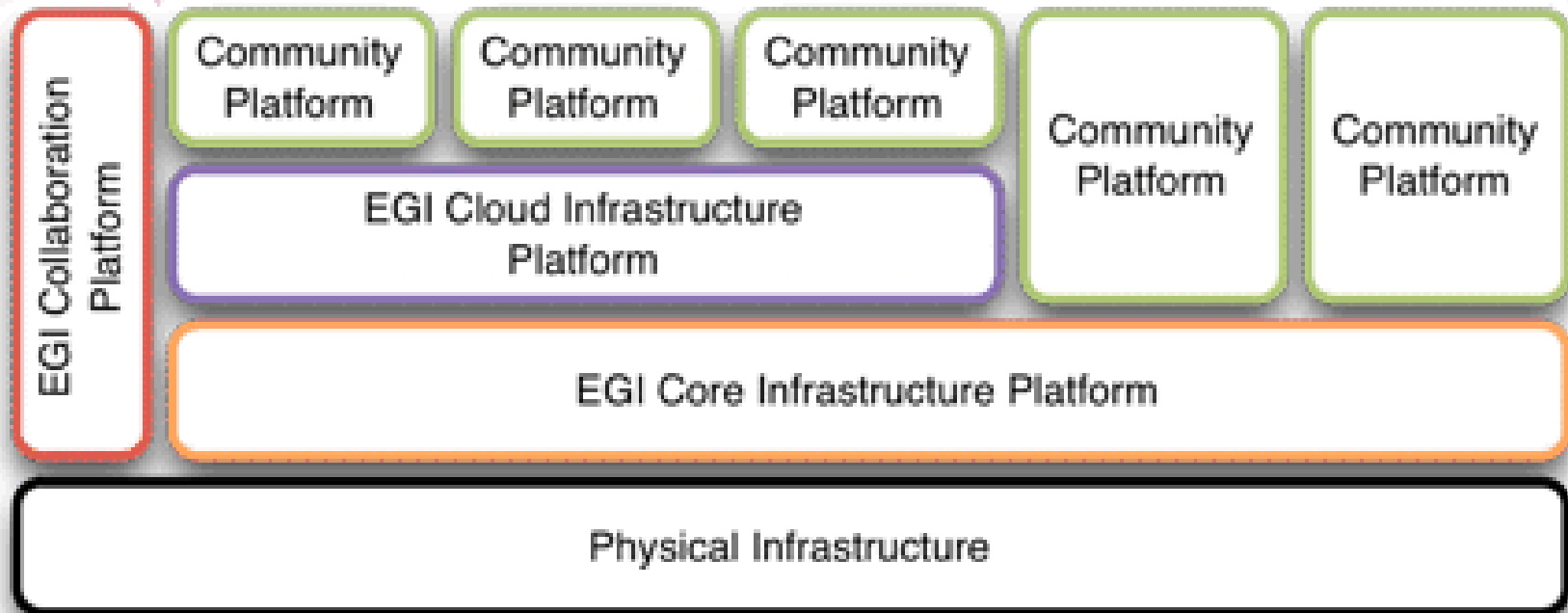


Image credits: <http://www.adarshpatil.com>

Une vue "plateforme"



La pile grille

Applications



Standalone apps



Portails & gateways

Middleware



File catalogs



Information System



Workload management



Storage Element



Computing Element



User Interface



Auth. service

Ressources



Disque/bande



Cluster



Serveur

Réseau



Internet public

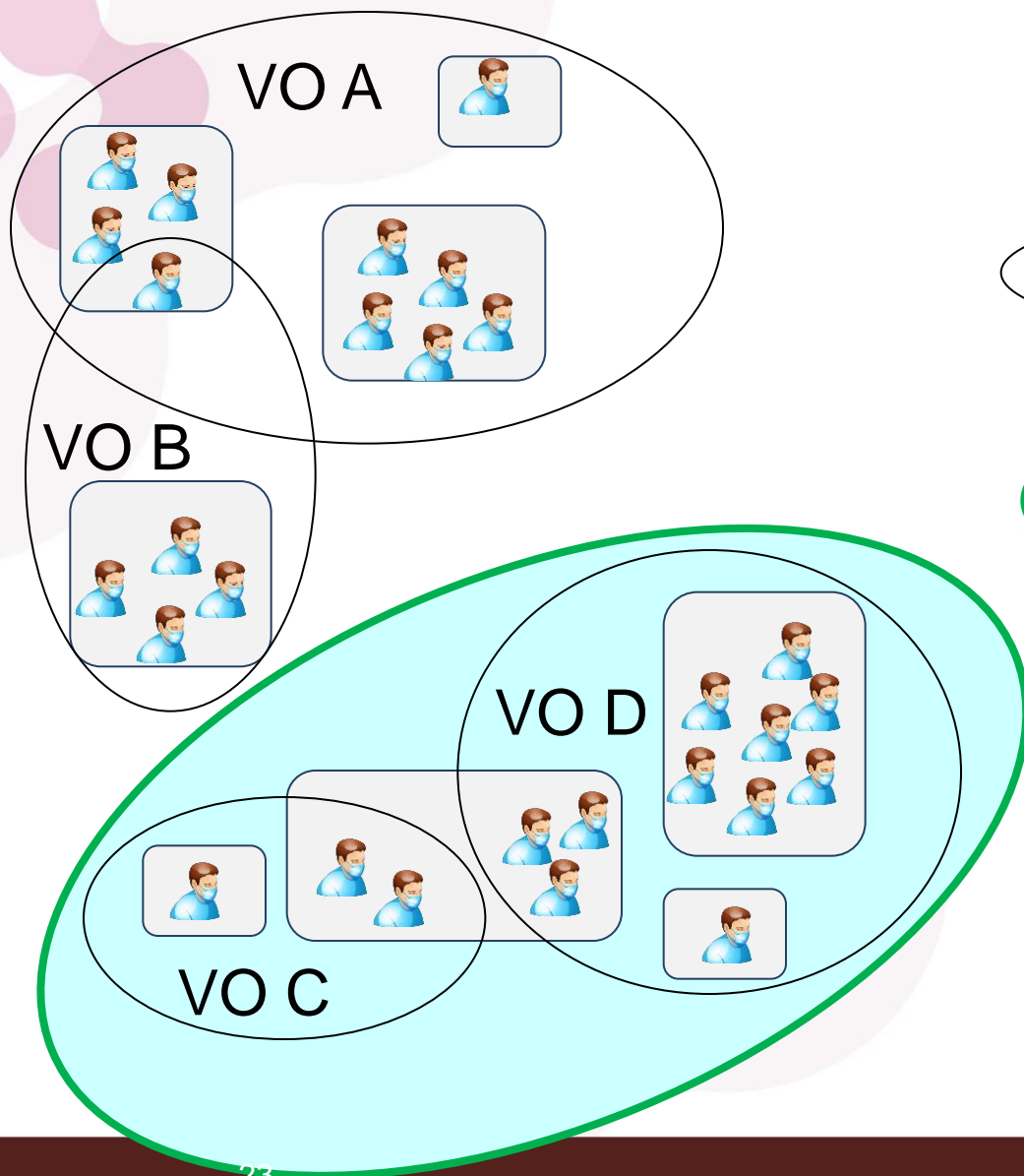



Réseaux académiques



Lignes dédiées

Organisation des utilisateurs



 Research teams

Virtual Organisations (VO)

Virtual Research Communities (VRC)

23

In EGI (Aug.2014):
- 213 VOs
- 5 VRCs

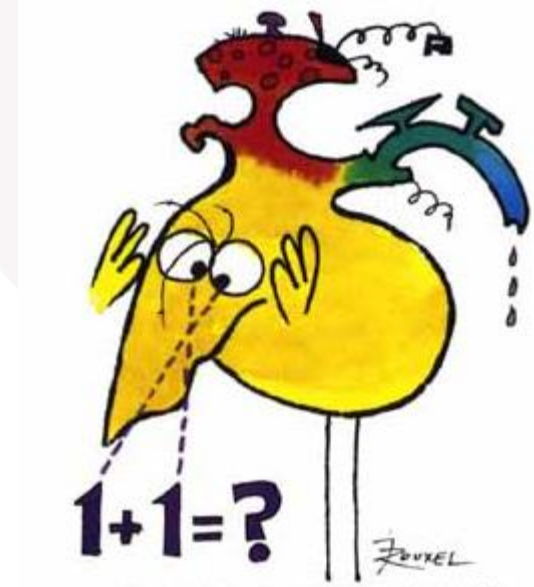


Pour commencer, une histoire simple

- Il était une fois, un chercheur qui voulait utiliser la grille pour lancer un programme compliqué (du genre ls -a)...

TO DO LIST

Ton calcul tu prépareras
 Patte blanche tu montreras
 Ton calcul tu soumettras
 Patiemment tu attendras
 Tes résultats tu récupèreras



Ton calcul tu prépareras

- **Il faut créer un "job"... c'est quoi ?**
 - un exécutable ou un script qui peut être lancé sur un (ou plusieurs) CPUs pour produire un *output*.
 - généralement, avec des paramètres et des données d'entrée
 - le tout encapsulé dans une couche de description écrite en JDL (*Job Description Language*)



Ton calcul tu prépareras

- **JDL = liste de type "key = value"**

```
[
Executable = "ls -al";
StdError = "stderr.log";
StdOutput = "stdout.log";
OutputSandbox = {"stderr.log", "stdout.log"};
]
```

- **C'est dans le JDL que sont spécifiés tous les *requirements* (librairies, OS, mémoire, données, etc.)**



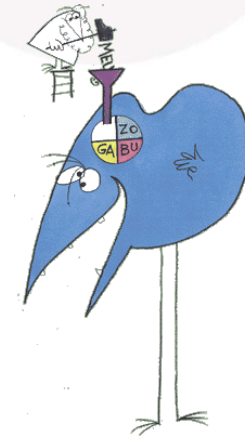
Patte blanche tu montreras

- **Passeport = certificat X509**
 - Délivré par une autorité reconnue (en France : Grid2FR, opérée par RENATER)
- **Visa = Proxy**
 - L'utilisateur doit demander un proxy (délégation) à durée de vie limitée (e.g. 24h)
 - Le service d'authentification ne délivre un proxy que si le certificat est valide et l'utilisateur est bien enregistré dans la VO correspondante
- **Authentification mutuelle**
 - Utilisateur et service doivent prouver leur identité



Ton calcul tu soumettras

- **Depuis où ?**
 - Une interface utilisateur en ligne de commande, un portail, un « job scheduler »...
- **Une commande, et hop**
 - Le job est soumis au système
 - Un mécanisme de *Workload Management* interroge le système d'information pour déterminer le meilleur élément de calcul (CE) disponible
 - Le job est mis en queue sur cet élément

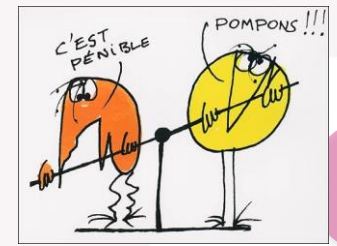
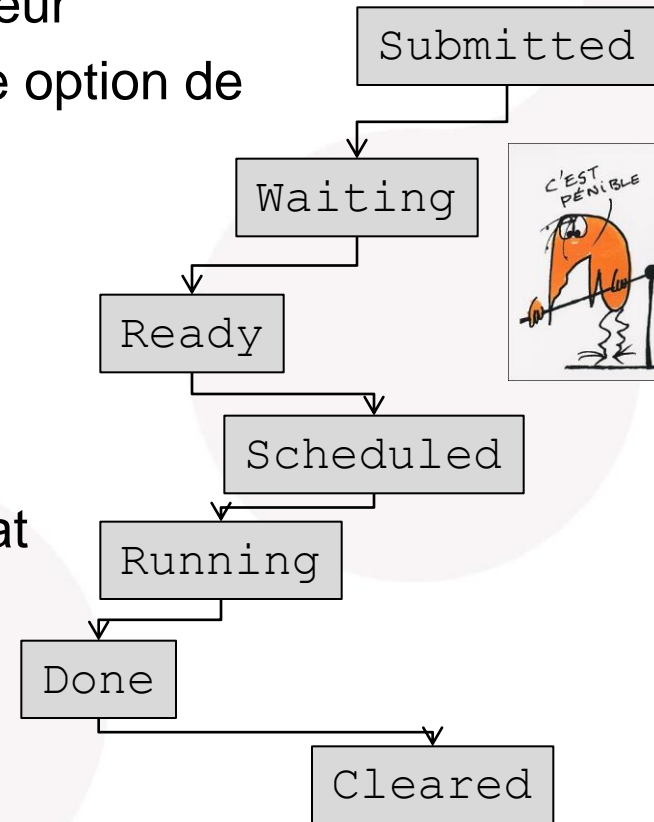


Patiemment tu attendras...

- **Pendant ce temps-là :**

- Le job a été envoyé par l'utilisateur
- Le système cherche la meilleure option de soumission
- Le job est transmis au CE
- Le job est en attente sur le CE
- le job tourne
- Il est terminé
- L'utilisateur a récupéré le résultat

Statut du job



Tes résultats tu récupèreras

- Dans notre exemple, le seul résultat est le stdout
- Dans la vraie vie, le "vrai" résultat est sans doute un set de données généré et déposé sur un élément de stockage
 - Données référencées dans un catalogue
 - La référence est passée dans le job Output



Et en détails ?

- Pour plus de détails sur le fonctionnement des différents services entrant en jeu dans le scénario précédent, voir la 2^{nde} partie de cette présentation...



LA GRILLE ET LE MONDE



Grille vs HPC

GRILLE

Architectures "classiques"
Répartition géographique
Modèle opportuniste
Unité: CPU.Hours
Parallélisme de données
Jobs indépendants

HPC

Supercalculateurs
Localisation géographique
Modèle d'allocation
Unité: Flop
Parallélisme de calcul
Jobs corrélés

Deux outils différents pour deux besoins différents!



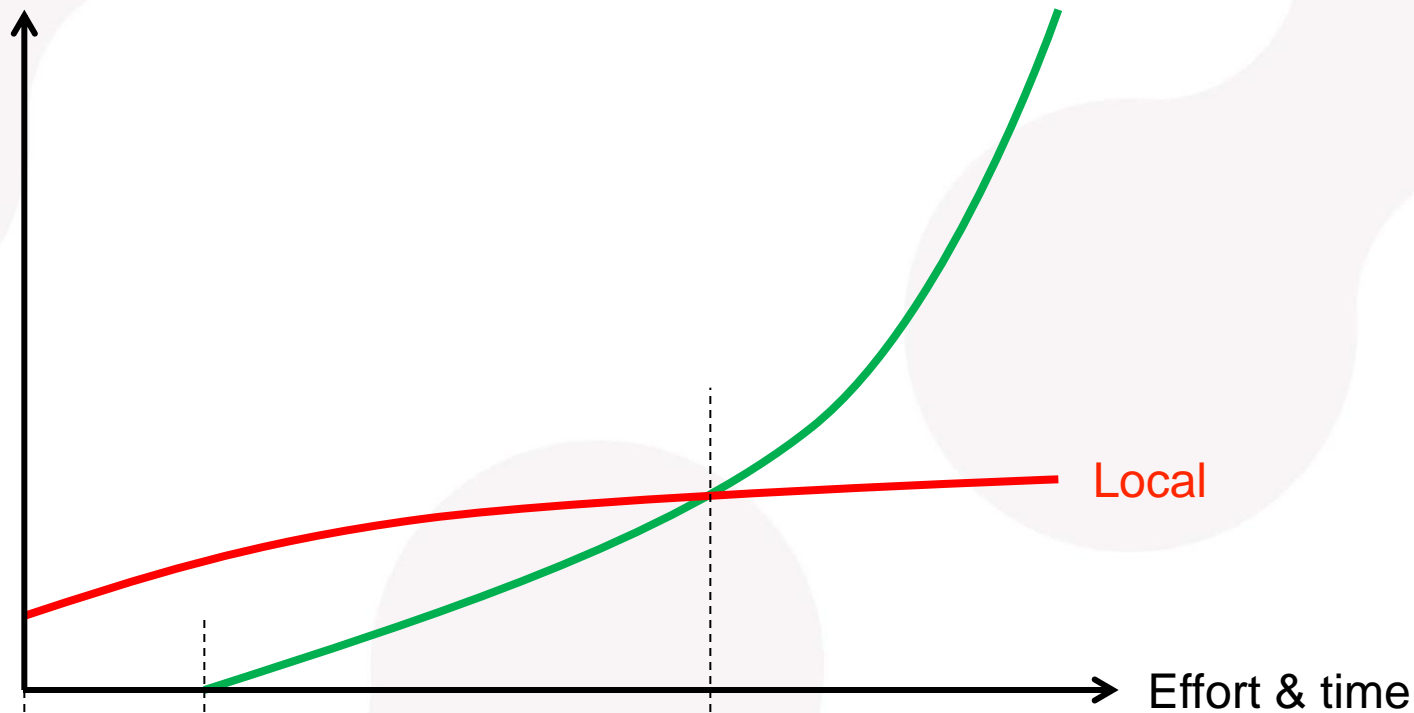
Bénéfices & limitations

- **La grille est faite pour l'analyse à haut débit (HTC). Elle fonctionne bien pour :**
 - Les tâches peu couplées, non communicantes
 - Les gros volumes de données distribuées
 - Les collaborations et utilisateurs dispersés
- **Les bénéfices dépendent aussi :**
 - Du niveau de support et des ressources disponibles pour la communauté considérée
 - Du ratio entre l'effort initial pour "gridifier" le code et l'usage attendu



Bénéfices et limitations

Benefits



"gridification"
overload



Quelques bénéfices

- **Optimisation de l'utilisation des ressources**
 - Les clusters de la grille tournent à bloc
- **Collaboration facilitée**
 - A travers le concept de VO
- **Partage d'expertise et support**
 - Création de nouvelles communautés
- **Faible coût économique**
 - Mutualisation et réutilisation



Quelques limites

- **Complexité et coût d'entrée**
 - Réécriture du code, intégration des concepts...
- **Taux d'erreur des jobs**
 - Le 100% de succès est difficile à atteindre
- **Files d'attente**
 - Surtout pour les petites VOs
- **Ne convient pas à tous les profils**
 - Ex: nombreux fichiers de grosse taille
 - Jobs très courts ou très longs



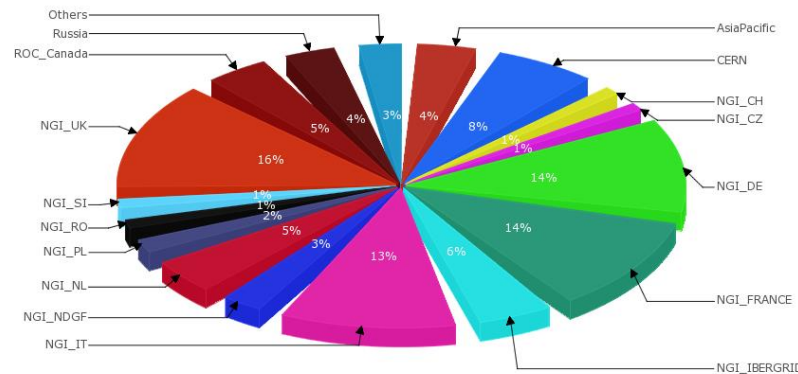
Utilisation des ressources dans EGI

- **2004-2014 : 57 milliards d'heures CPU**
 - Equivalent à 6,5 millions d'années CPU en 10 ans
 - 2^E10 heures (2.3 millions d'années) en 2014
 - 14% ont été utilisées sur France Grilles

Developed by CESGA 'EGI View': / normelap-HEPSPEC06 / 2004:1-2014:8 / REGION-VO / all (x) / GRBAR-LIN / 1

2014-08-12 07:05

Normalised Elapsed time (HEPSPEC06) per REGION



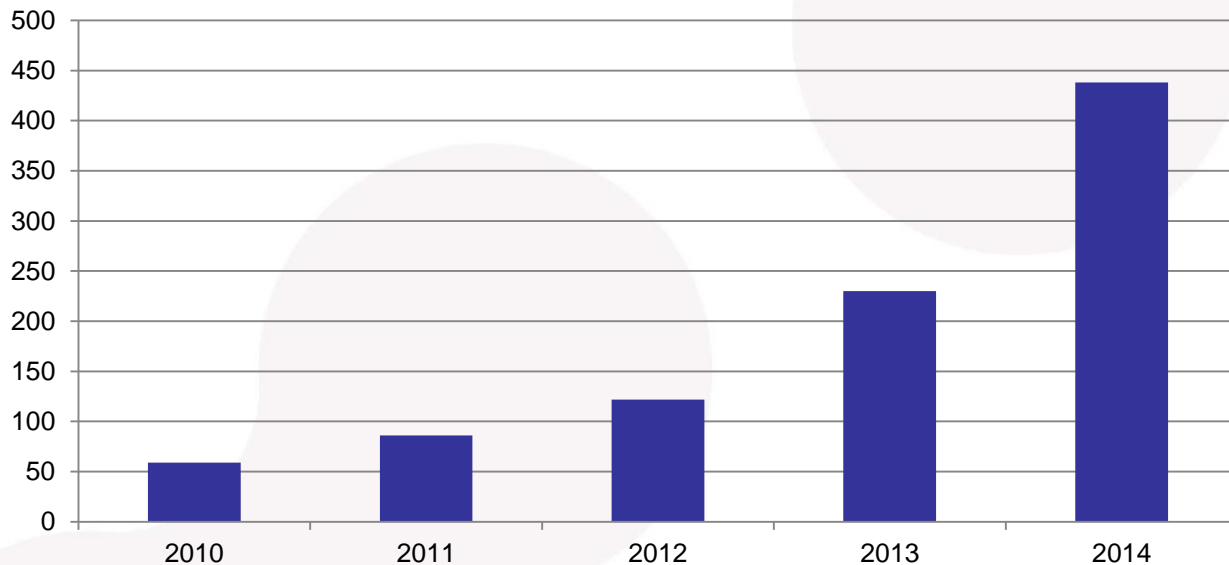
Source : EGI Accounting Portal



Sciences de la vie

- **438 millions d'heures CPU en 2014**
 - Seulement 2% de l'utilisation totale
 - Mais croissance exponentielle

Millions d'heures CPU normalisées

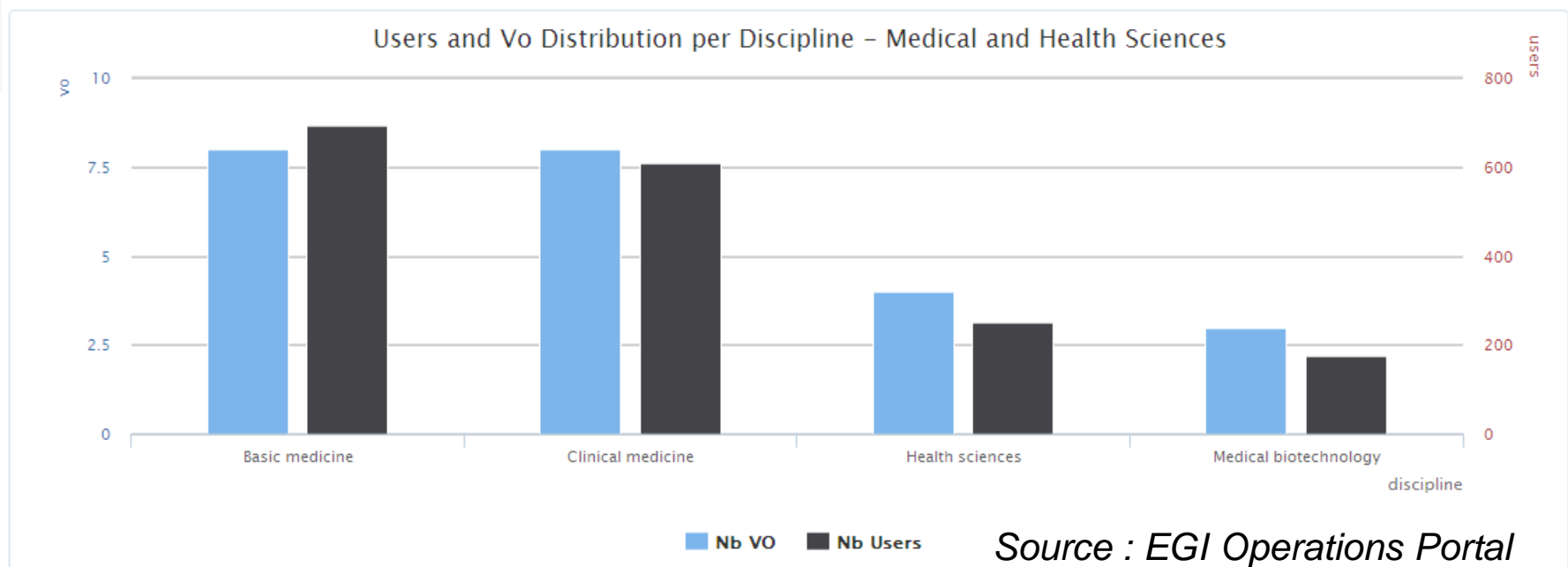


Source : EGI Accounting Portal



Sciences de la vie

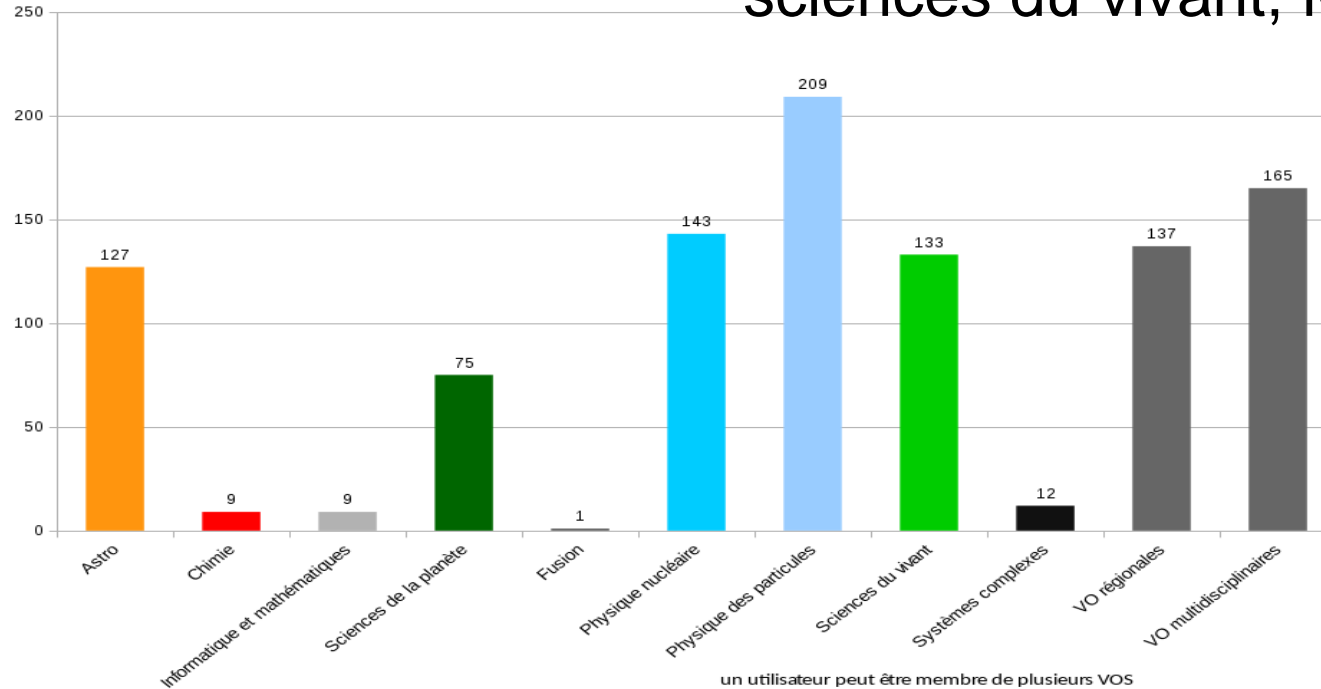
- **"medical and health sciences"**
 - 24 VOs (mai 2015)



Les utilisateurs en France

Nombre d'utilisateurs France Grilles par discipline selon leur inscription aux VOs décembre 2014

133 utilisateurs "directs" en sciences du vivant, MAIS:



- Certains utilisateurs font des sciences de la vie dans les VO pluri-disciplinaires
- D'autres sont "regroupés" derrière des portails





Questions ?

