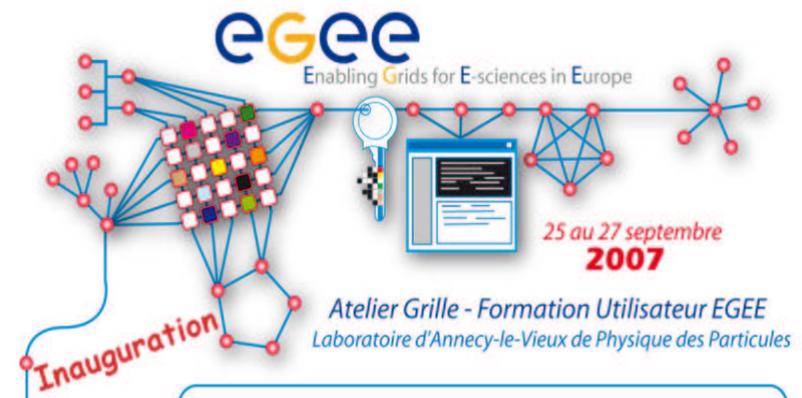




Enabling Grids for E-science

La grille EGEE : Nouvelle infrastructure pour les collaborations scientifiques et industrielles

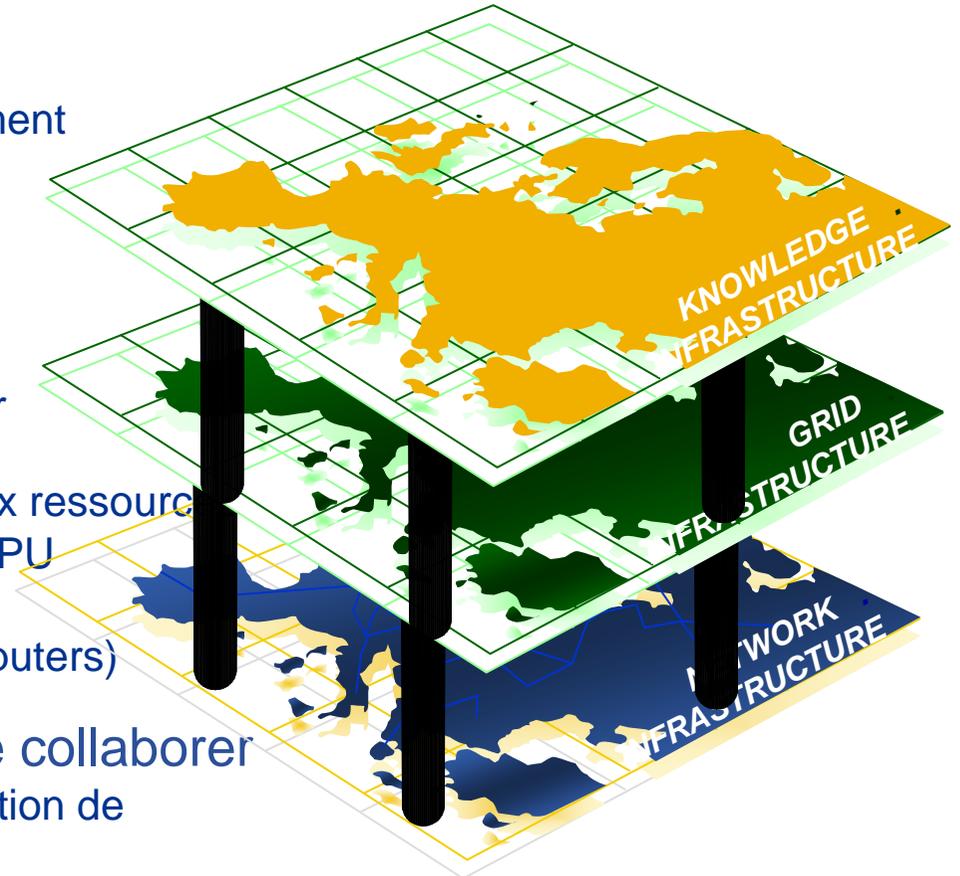
Bob Jones
Directeur du Projet EGEE-II
CERN



www.eu-egee.org



- Elles facilitent l'accès
 - aux petits groupes de recherche
 - aux scientifiques de différents domaines
 - aux pays isolés et en voie de développement
- Pour les nouvelles technologies
 - Permettent la production et le stockage de quantités massives de données
 - Permettent l'accès de façon transparente à plusieurs millions de fichiers répartis sur différents domaines administratifs
 - Permettent d'accéder dynamiquement aux ressources
 - Mobilisation d'un grand nombre de CPU & stockage à court terme (PC clusters)
 - Equipements haut de gamme (supercomputers)
- Elles créent de nouveaux moyens de collaborer
 - Développement d'applications par l'utilisation de workflows complexes
 - Facilité de Collaboration distribuée et de création de nouvelles Communautés
 - Facilitent l'accès à un enseignement de plus haut niveau



- **EGEE**

En mai 2007, la revue de la 1ere année de projet a été un succès
 « ...c'est une remarquable réussite d'avoir établi une infrastructure opérationnelle pouvant fournir des services de production pour une communauté d'utilisateurs d'aussi divers horizons. »

- **EGEE II**

- 1er Avril 2006 – 31 Mars 2008
- 91 partenaires dans 32 pays
- 13 fédérations

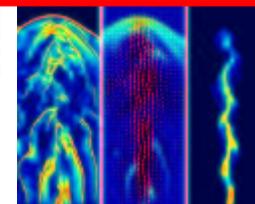


- **Objectifs**

- Offrir la qualité de service requise pour une infrastructure de production à grande échelle pour l'e-science
- Attirer de nouvelles ressources ainsi que les utilisateurs de l'industrie en plus des scientifiques
- Maintenir et faire évoluer le middleware "gLite"

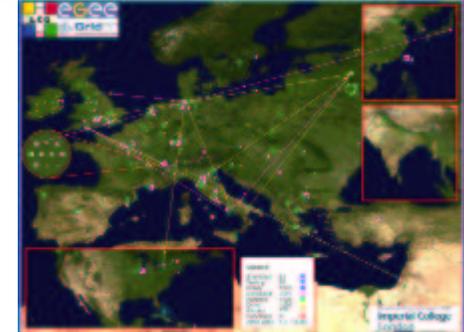
Partenaires en France:

CEA, CGG, CNES, CNRS, CRSA, CS SI, HEALTHGRID



- **Infrastructure de production**

- Comprend actuellement + de 220 sites dans + de 40 pays
- Monitoring continu des services de la grille
- Configuration/gestion automatique des sites



- **Middleware**

- Distribué sous une licence de logiciel «open source»



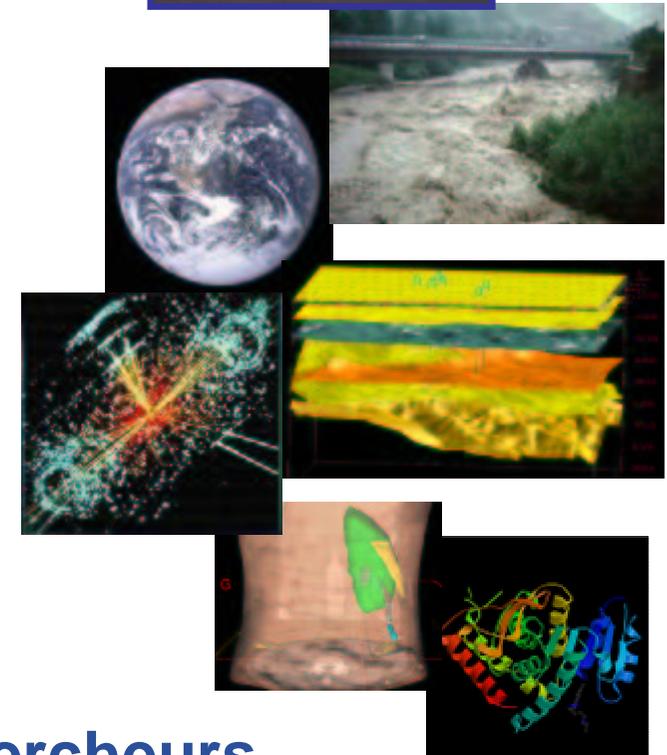
- **Support aux utilisateurs**

- Formation
- Documentation
- Support à la « gridification » des applications
- Helpdesk
- User Forum, Conférences, etc.



Applications dans un nombre de domaines croissants

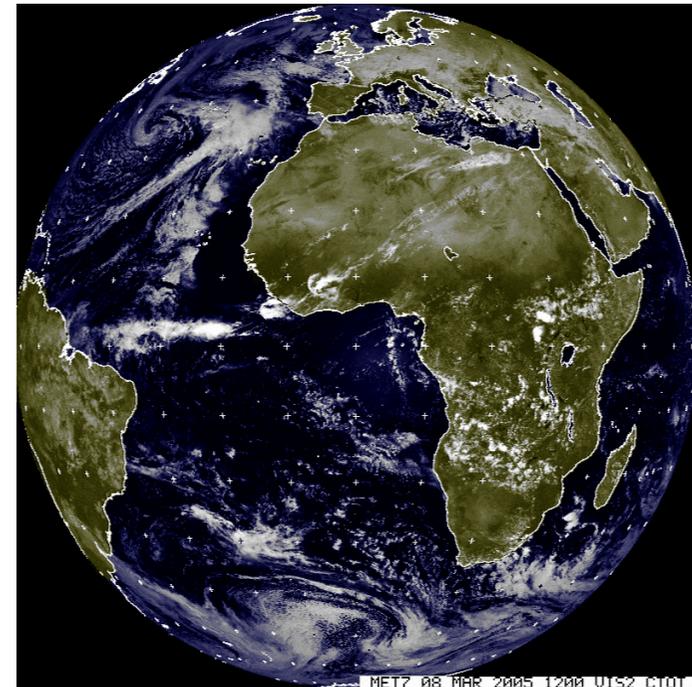
- Archéologie
- Astronomie & Astrophysique
- Protection civile
- Chimie
- Sciences de la Planète
- Simulation Financière
- Fusion
- Géophysique
- Physique des Hautes Energies
- Science de la vie
- MultiMedia
- Science des matériaux
- ...



**Infrastructure utilisée par >5000 chercheurs
~20 millions de jobs soumis en 2006**

- **Communauté**
 - Plusieurs petits groupes et projets qui collaborent

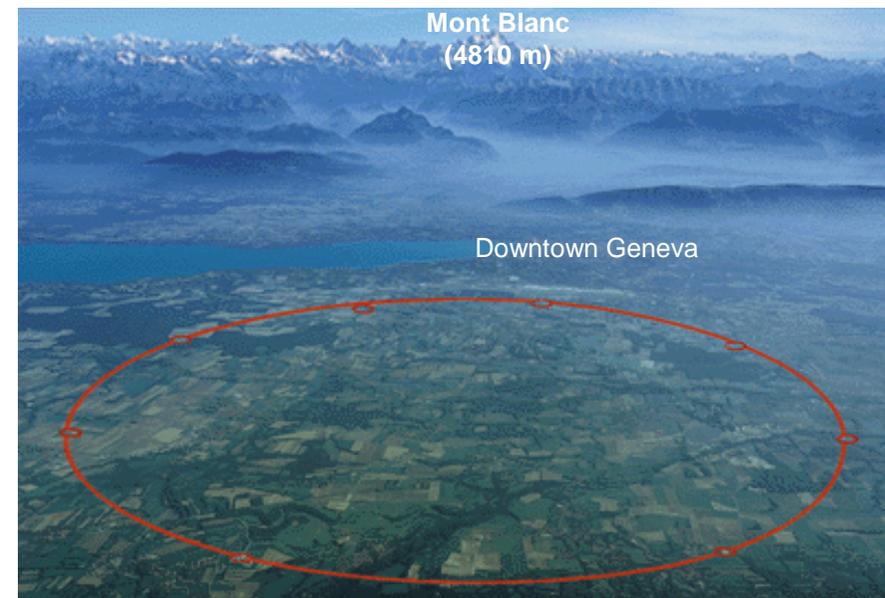
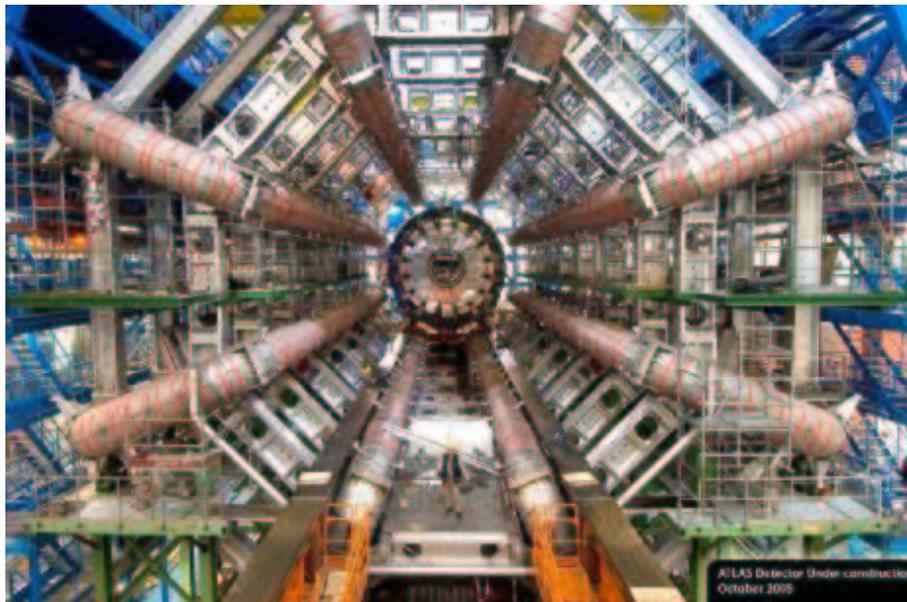
- **La terre**
 - Système complexe
 - Des domaines indépendants avec des interfaces
 - Croûte terrestre – Océan – Atmosphère
 - Physique, chimie et/ou biologie



- **Les Applications**
 - Observation de la terre par satellite
 - Physique du globe
 - Hydrologie
 - Climat
 - Géosciences
 - Pollution
- Météorologie, météorologie de l'espace
 - Atmosphère de Mars
 - Stockage de bases de données

Large Hadron Collider (LHC):

- L'un des instruments les plus puissants jamais construit pour la recherche sur la matière
- 40 Millions de collisions de particules par seconde
- 4 Expériences: ALICE, ATLAS, CMS, LHCb
- ~15 PetaBytes/an pour les 4 expériences
- Mise en route en 2008

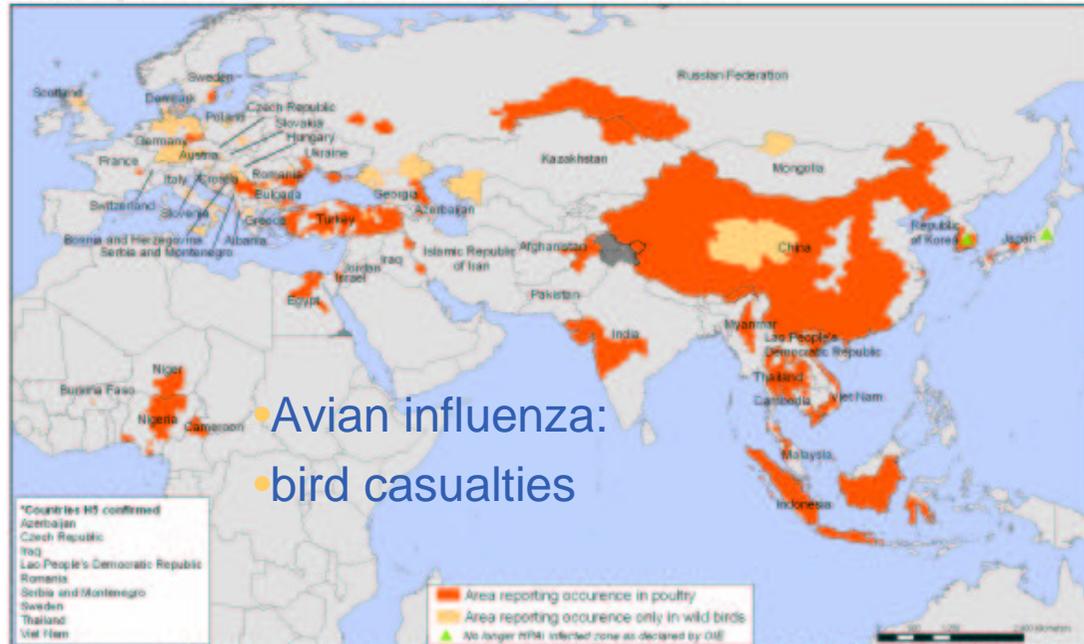


- Des pathologies comme le HIV/SIDA, la grippe aviaire, etc. sont une menace pour la santé publique en raison des échanges mondiaux et de la circulation des personnes
- La grille offre de nouvelles perspectives dans la recherche effectuée *in silico*
 - En réduisant les coûts et en améliorant la rapidité de recherche de nouveaux médicaments

Une collaboration Internationale est indispensable pour :

- La détection précoce
- La surveillance épidémiologique
- La prévention
- La recherche de nouveaux médicaments
- La recherche de vaccins

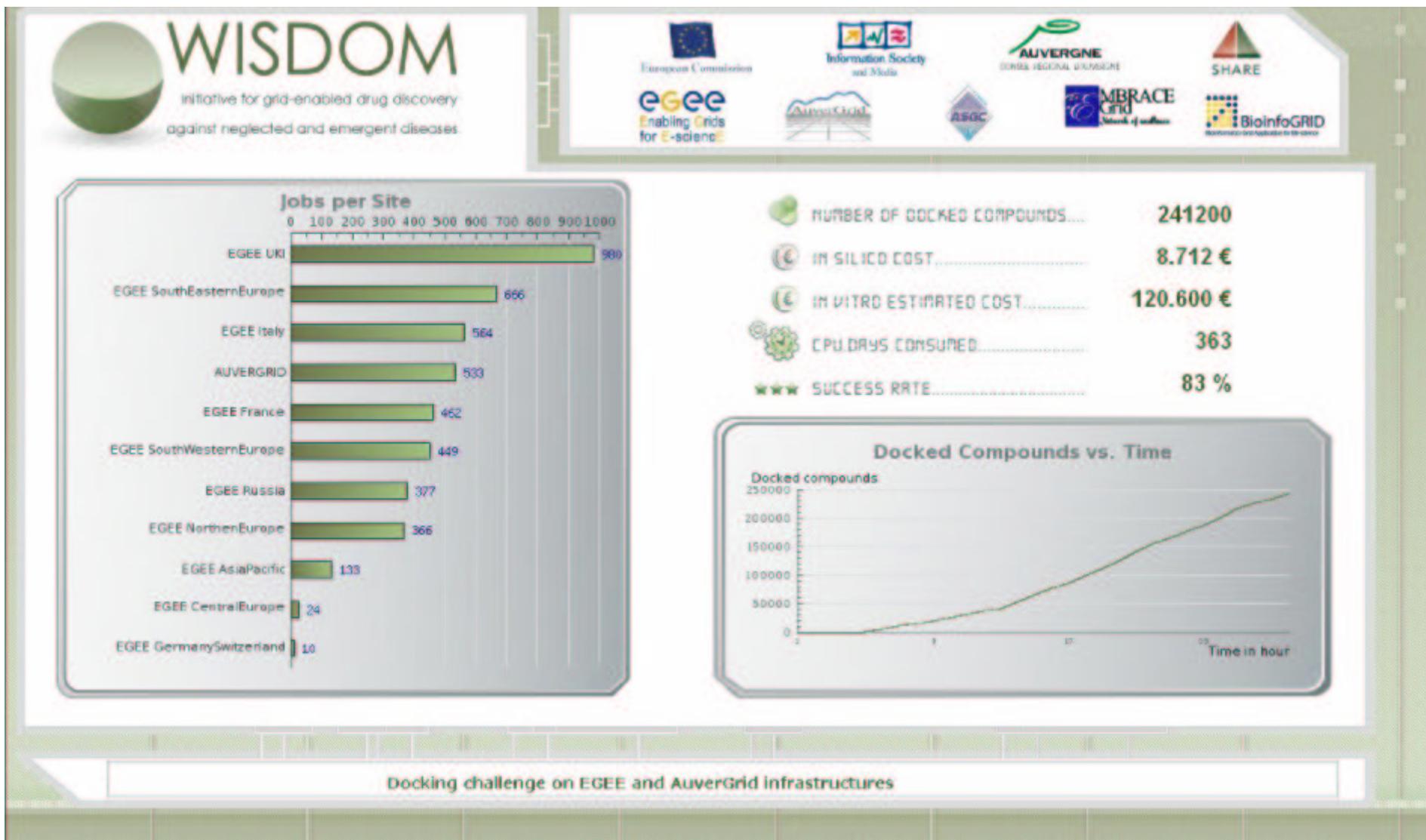
Areas reporting confirmed occurrence of H5N1* avian influenza in poultry and wild birds since 2003. Status as of 07 April 2005



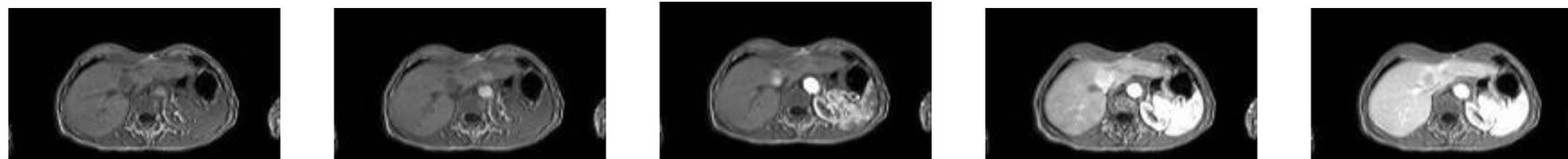
• Avian influenza:
• bird casualties

World Health Organization © WHO 2005. All rights reserved. The boundaries and names shown on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its borders or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement. Data Source: World Organisation for Animal Health (OIE) and national governments. Map Produced: Public Health Mapping and GIS Communicable Diseases (CD) World Health Organization

<http://wisdom.healthgrid.org/>

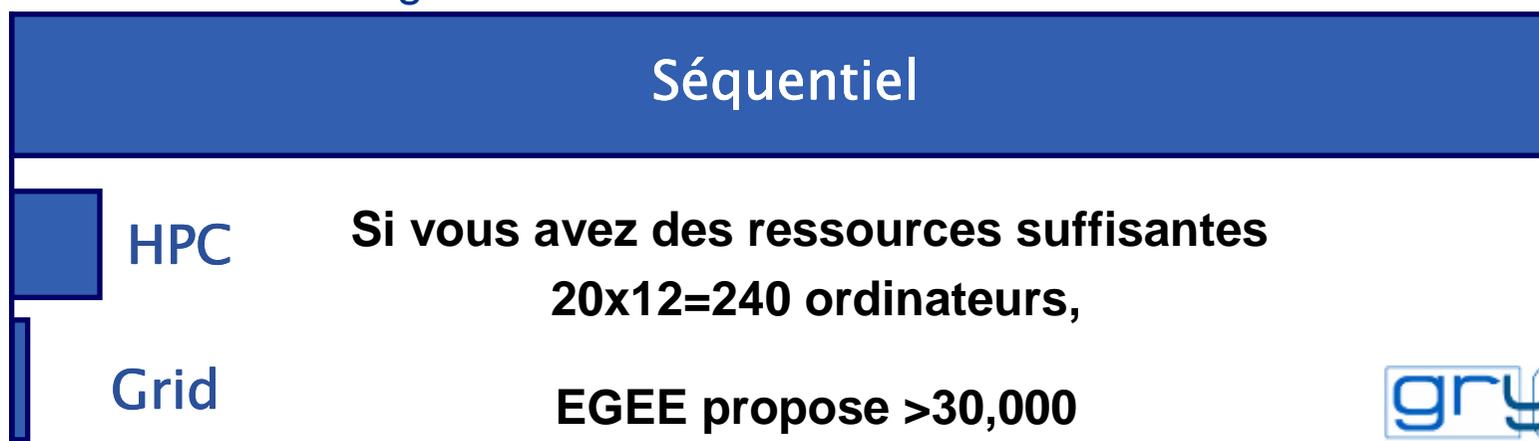


- **Pharmacocinétique : L'étude de la diffusion des agents de contraste**
 - Enregistrement simultané d'une série d'images médicales volumétriques dans l'analyse de l'évolution de la diffusion des agents de contraste.



- **Coûts informatiques**

- 20 Patients: 2623 heures (enregistrement + Image paramétrique)
- Utilisation d'une ferme informatique de 20-processeurs :146 heures
- Utilisation de la grille: <20 heures



GPS@: portail bioinformatique

- <http://gpsa.ibcp.fr/>
 - Donne accès à des séquences régulièrement mises à jour et à des banques de données en 3D (EMBL, GenBank, SWISS-PROT etc.)
 - Des dizaines de codes bioinformatiques
- Interface facile à utiliser avec accès à des banques de données réputées
 - Utilise les ressources de la grille pour analyser les séquences

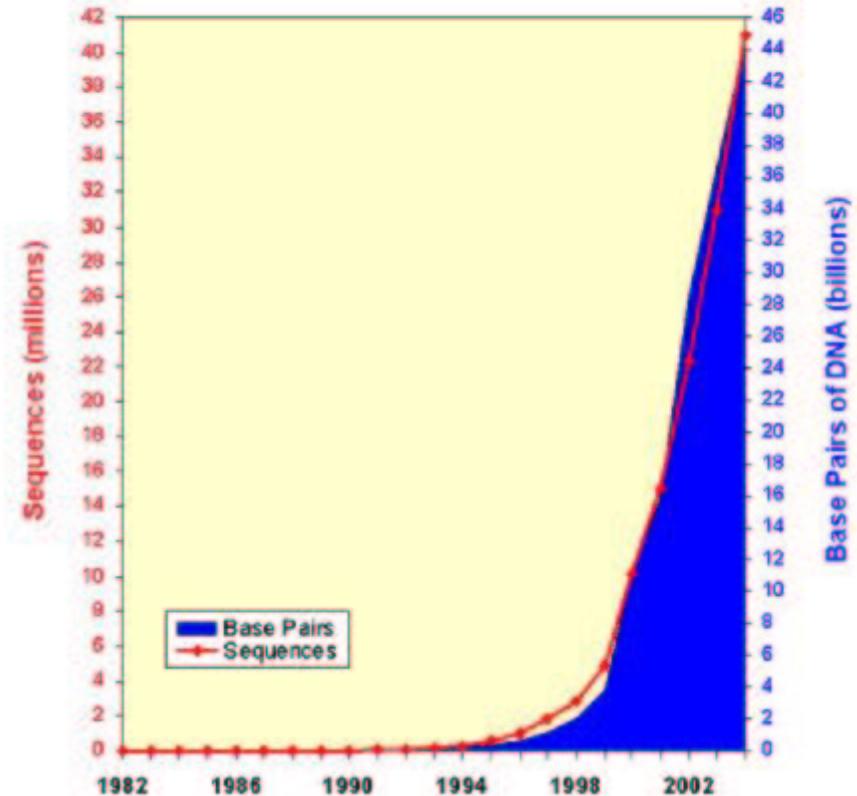


Data, Data, Data

Slide réalisée par Carole Goble



Growth of GenBank
(1982 - 2004)



www.eu-egee.org/uf2
Enabling Grids for E-science
EGEE USER FORUM
9 - 11 MAY 2007 MANCHESTER (UK)

Co-located with OGF20:

900+ attendees

~50 booths

user forum:

~30 sessions

100+ presentations

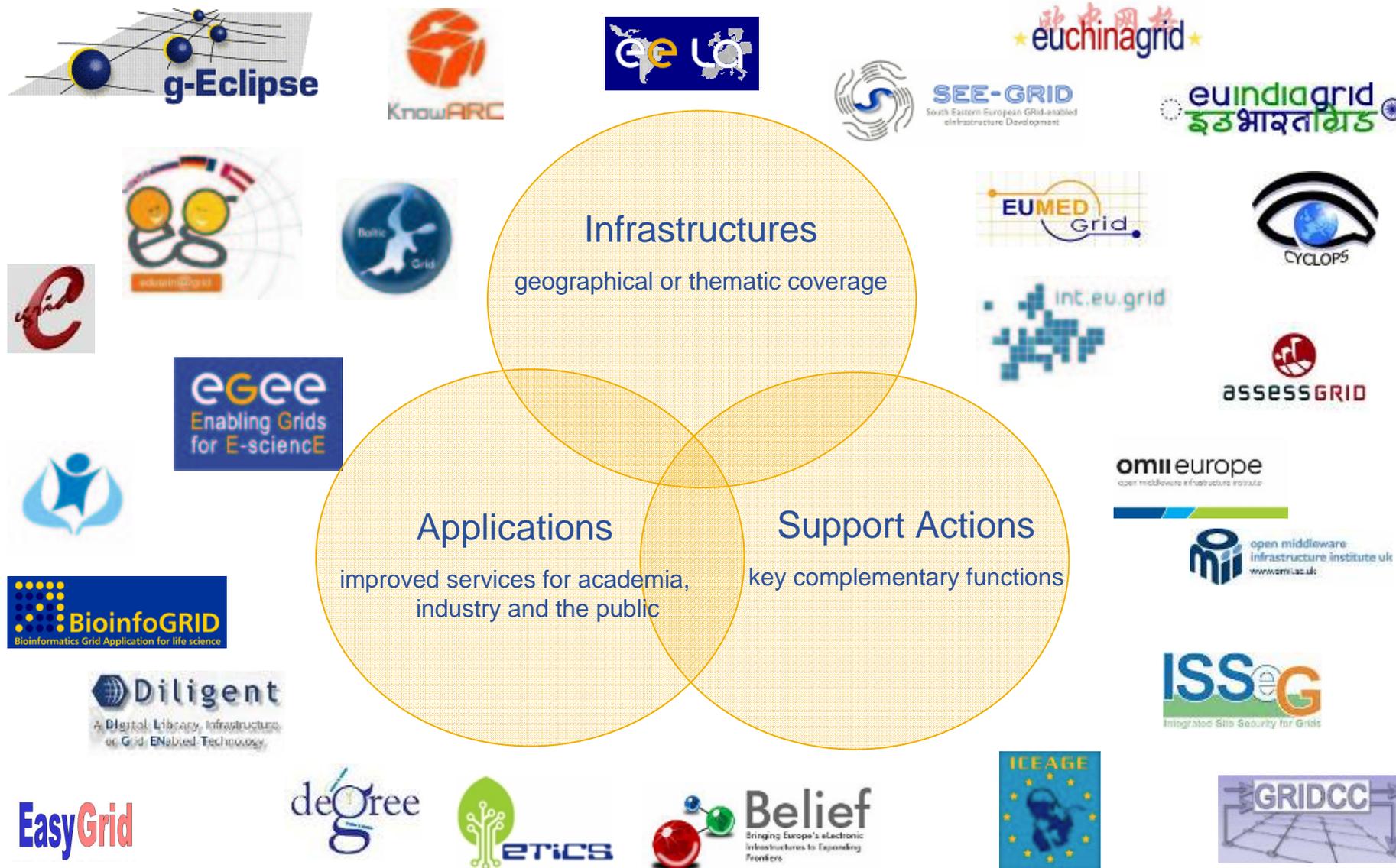
20 demos

~60 posters



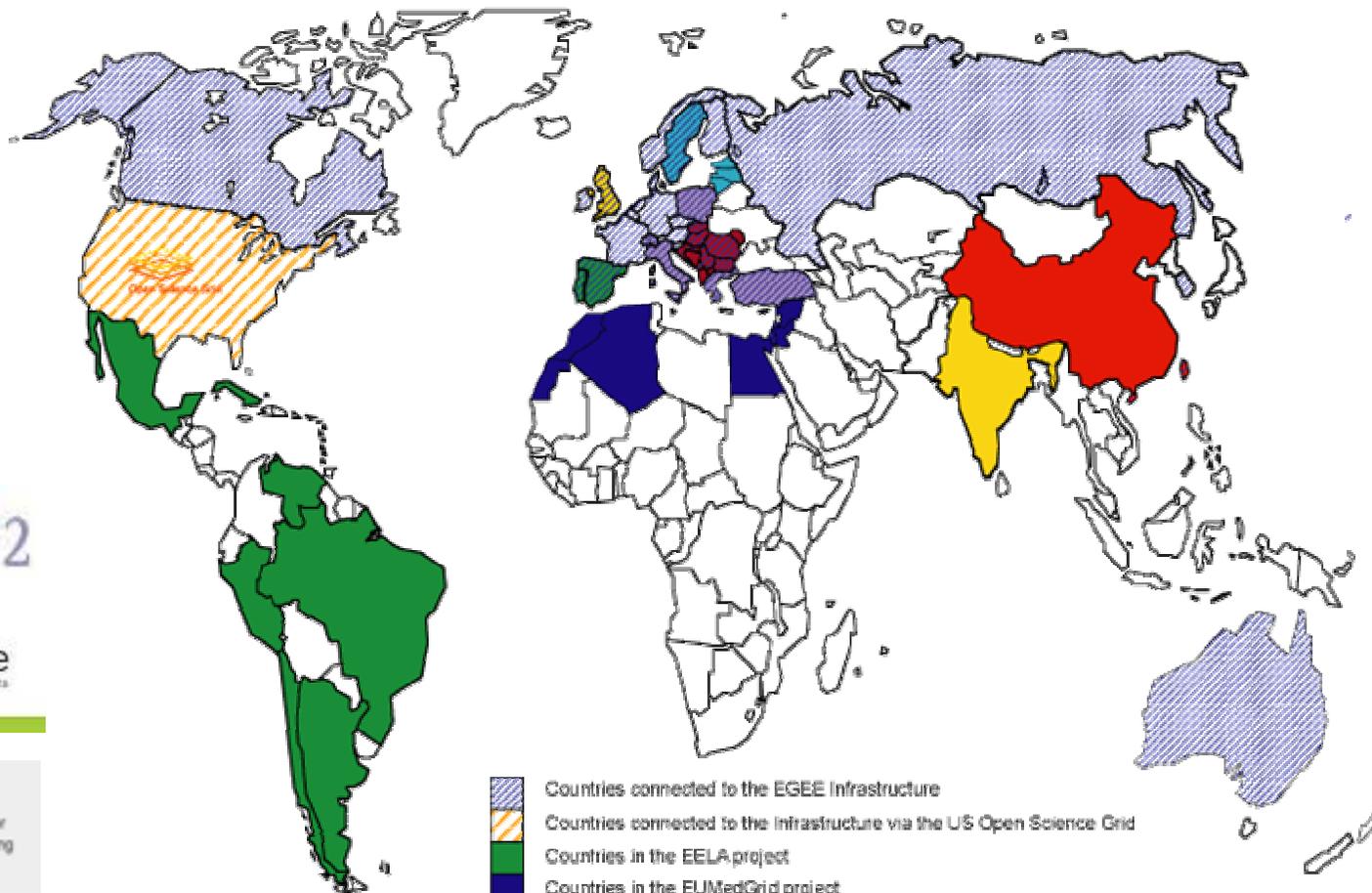
Le prochain User Forum aura
lieu à Clermont-Ferrand
en Février 2008

25 projects have registered as on June 2007: [web page](#)

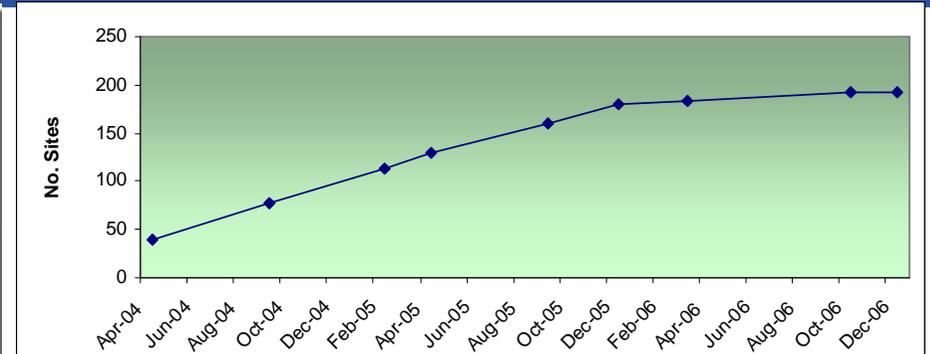
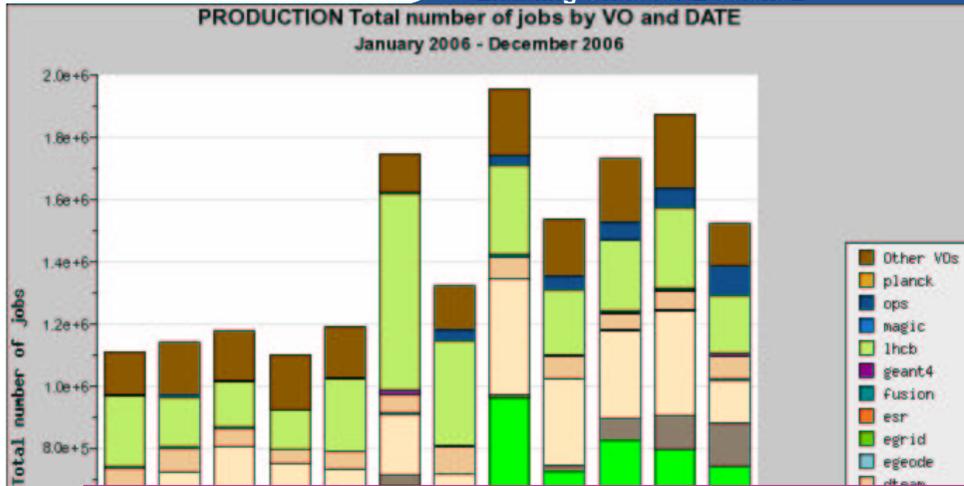




TeraGrid

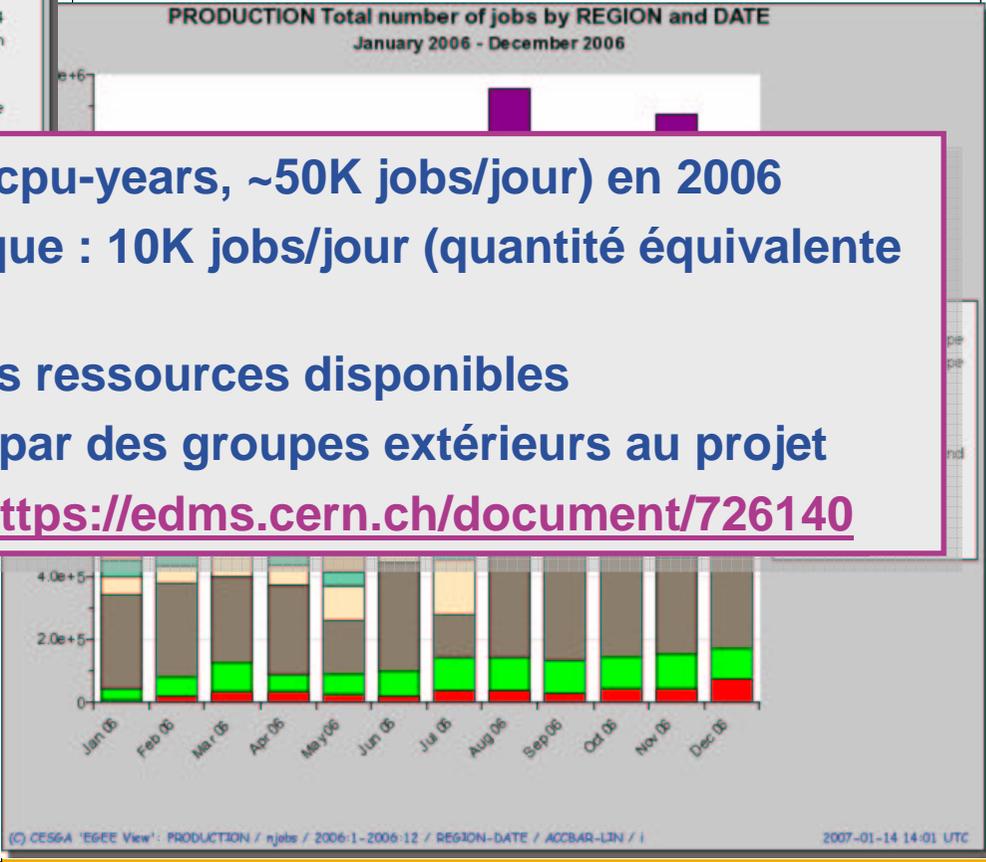


-  Countries connected to the EGEE Infrastructure
-  Countries connected to the infrastructure via the US Open Science Grid
-  Countries in the EELA project
-  Countries in the EUMedGrid project
-  Countries in the BalticGrid project
-  Countries in the SEE-GRID project
-  Countries in the EUIndiaGrid project
-  Countries in the EUChinaGrid project
-  Countries in several regional projects

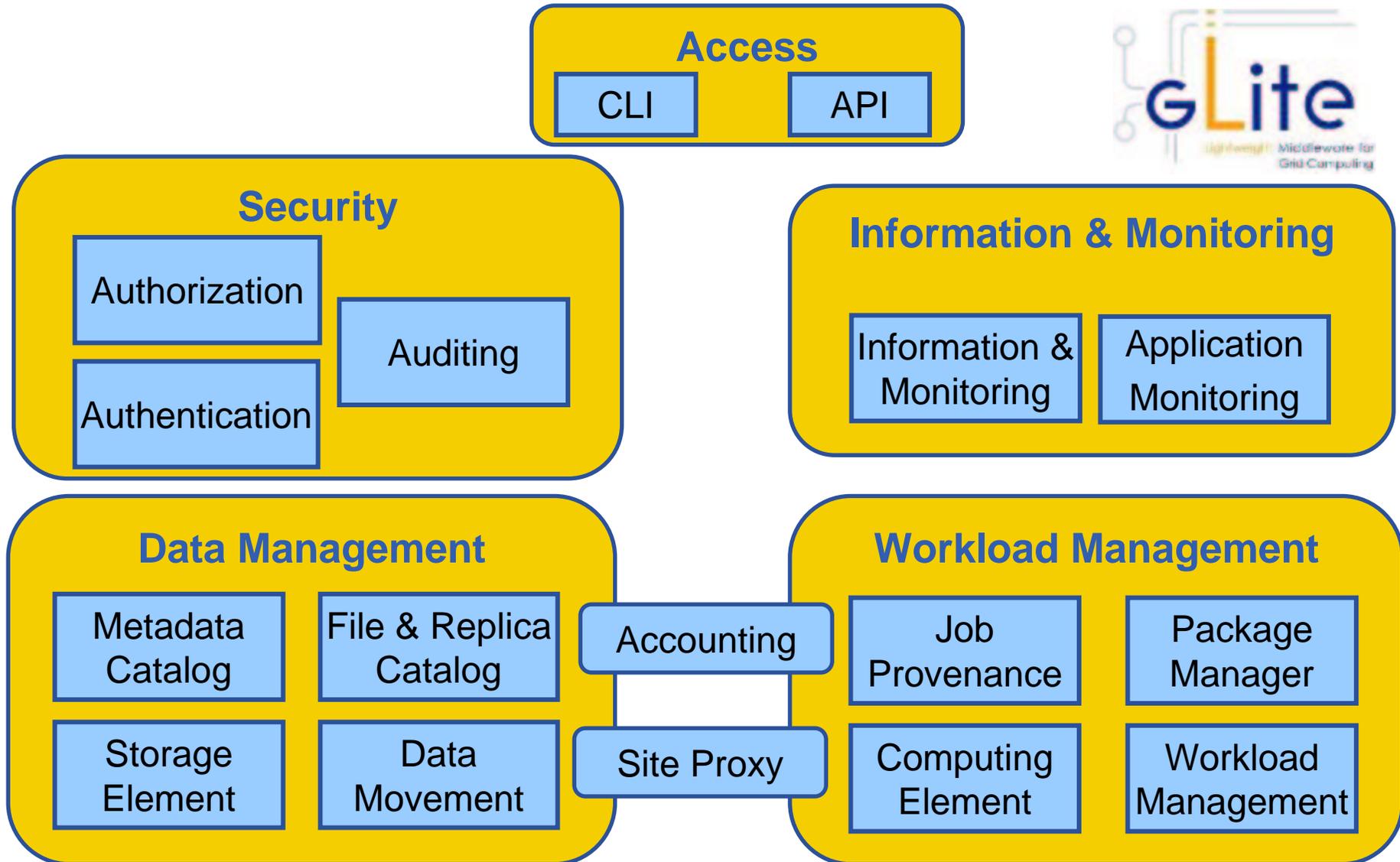


- ~20 millions de jobs soumis (8200 cpu-years, ~50K jobs/jour) en 2006
- Utilisation autre que pour la physique : 10K jobs/jour (quantité équivalente à celle de l'année 2005 pour EGEE)
- Utilisation permanente d' 1/4 à 1/3 des ressources disponibles
- 24% des ressources sont fournies par des groupes extérieurs au projet
- Rapport des activités de la grille: <https://edms.cern.ch/document/726140>

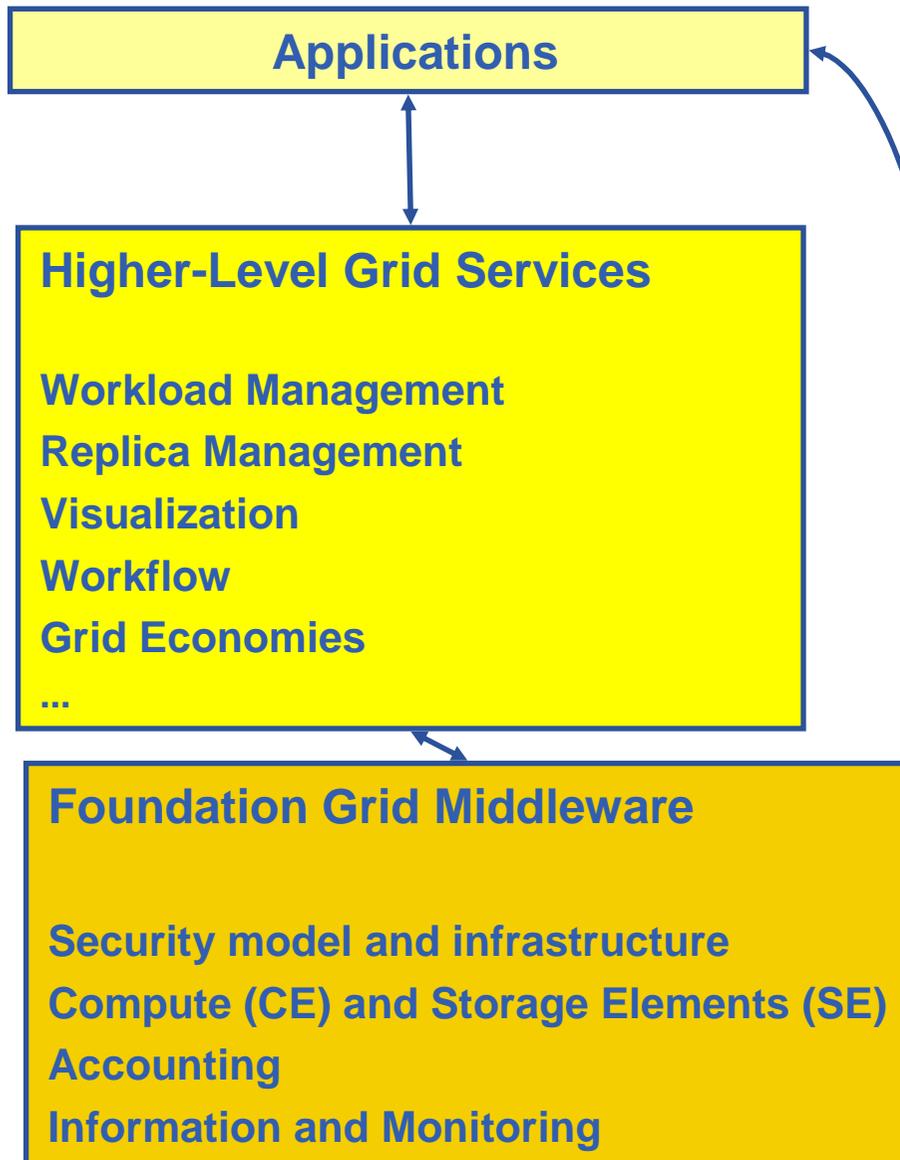
© CERN



© CERN 'EGEE View': PRODUCTION / njobs / 2006-1-2006-12 / REGION-DATE / ACCBAR-LIN / | 2007-01-14 14:01 UTC



<http://doc.cern.ch/archive/electronic/egee/tr/egee-tr-2006-001.pdf>

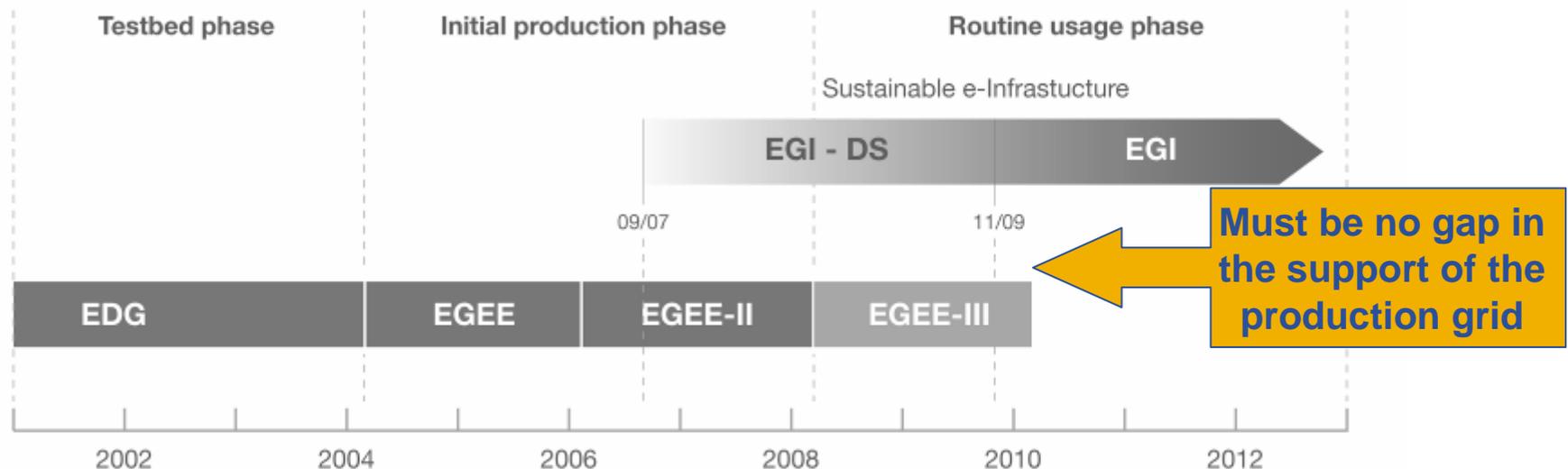


- Les applications accèdent à la fois aux services de haut niveau de la grille et aux Services de base du Middleware
- Les VOs complètent gLite avec d'autres services de haut niveau via le programme **RESPECT** (Recommended External Software for EGEE CommuniTies)
 - Identifient des logiciels externes fonctionnant avec gLite, apportant des services de haut niveaux aux utilisateurs
 - Informent les utilisateurs de l'existence de ces logiciels afin de leur éviter de les re-cr er

- **La proposition d'EGEE-III a été soumise le 20 septembre 07**
 - Commission Européenne INFRA-2007-1.2.3
- **Objectifs clés**
 - Etendre/optimiser l'existence de l' infrastructure d'EGEE, en incluant plus de ressources et de communautés d'utilisateurs
 - Préparer le passage d'un modèle en mode projet à une infrastructure durable et fédérative basée sur les Initiatives Nationales de la Grille
 - Période de 2 ans – printemps 2008 printemps 2010
 - Pas d'intervalle entre EGEE-II et EGEE-III
- **Consortium similaire**
 - Maintenant structuré sur une base nationale (National Grid Initiatives/Joint Research Units)

| <i>Networking activities</i> | <i>Specific Service Activities</i> |
|--|--|
| NA1: Management | SA1: Operations |
| NA2: Dissemination | SA2: Networking Support |
| NA3: Training | SA3: Integration, testing & Cert. |
| NA4: Applications | |
| NA5: International Coop. & Policy | <i>Joint Research Activities</i> |
| NA6: Business  | JRA1: Middleware engineering |

- **Besoin de créer une Infrastructure de Grille permanente et commune**
- **Assurer la pérennité de l'infrastructure européenne**
- **Indépendant des cycles de financement à court terme**
- **Coordonner l'intégration et l'interaction entre les Initiatives Nationales de Grille (NGIs)**
- **Au niveau européen: maintenir l'infrastructure en production pour un large éventail de domaines scientifiques**



Cyprus



Israel

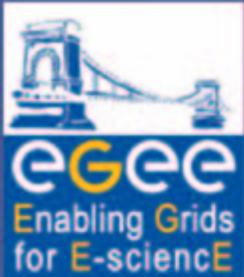


EGI - European Grid Initiative

Towards a sustainable production grid infrastructure

- **La proposition du projet EGI a été approuvée par la Commission Européenne (début 1^{er} Septembre 07)**
- **Le projet est soutenu par 31 initiatives nationales de Grille (NGIs)**
- **Projet de 2 ans pour préparer la mise en place d'un nouveau modèle organisationnel pour une infrastructure européenne durable de la grille.**
- **Modèle fédéré regroupant les NGIs dans la construction d'une organisation européenne.**
- **Responsabilités bien définies et complémentaires entre NGIs et EGI**

<http://www.eu-egi.org>



1-5 October, 2007
Europe Congress Center (ECC)
Budapest, Hungary

Building Bridges...

- Between Science and business
- Between users and infrastructures
- Between countries
- Between scientific disciplines
- Between projects
- Etc

<http://www.eu-egee.org/egee07>

- **Le Grid est synonyme de partage – c'est un moyen de travailler de façon coopérative à travers le monde**
- **L'infrastructure EGEE est « la plus grande grille mondiale multi-science en production »**
- **Le transfert de technologie vers les partenaires et vers l'industrie est une volonté d'EGEE et de la Communauté Européenne**
- **Le long terme en préparation**
 - EGEE, les projets associés, les projets d'initiatives nationales de grille, ainsi que les communautés d'utilisateurs travaillent ensemble afin de définir le modèle d'une infrastructure de Grille durable, indépendante des cycles de projet à court terme.