

# Εισαγωγή στο Grid, EGEE και το HellasGrid

## *Introduction to Grid, EGEE and HellasGrid*

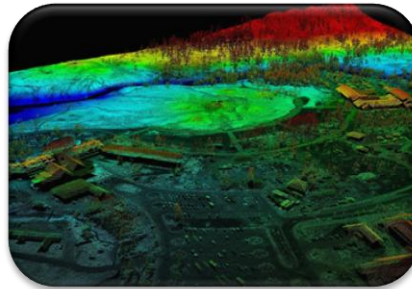
*Athanasia Asiki*

*[aassiki@cslab.ece.ntua.gr](mailto:aassiki@cslab.ece.ntua.gr)*

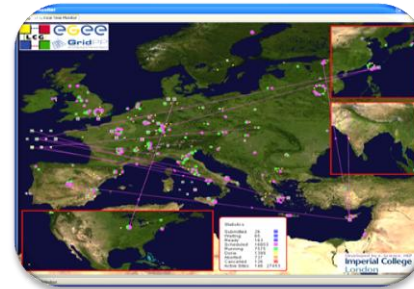
*Computing Systems Laboratory,  
National Technical University of Athens*



What is the Grid?



Grid paradigms



Enabling Grid for E-science (EGEE)



gLite middleware

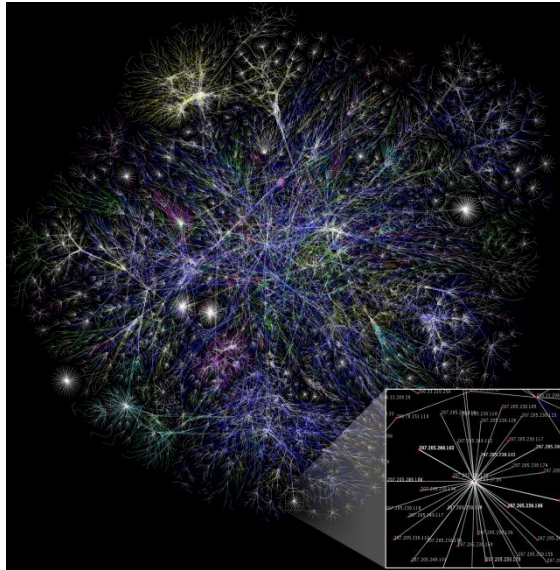


HellasGrid Taskforce



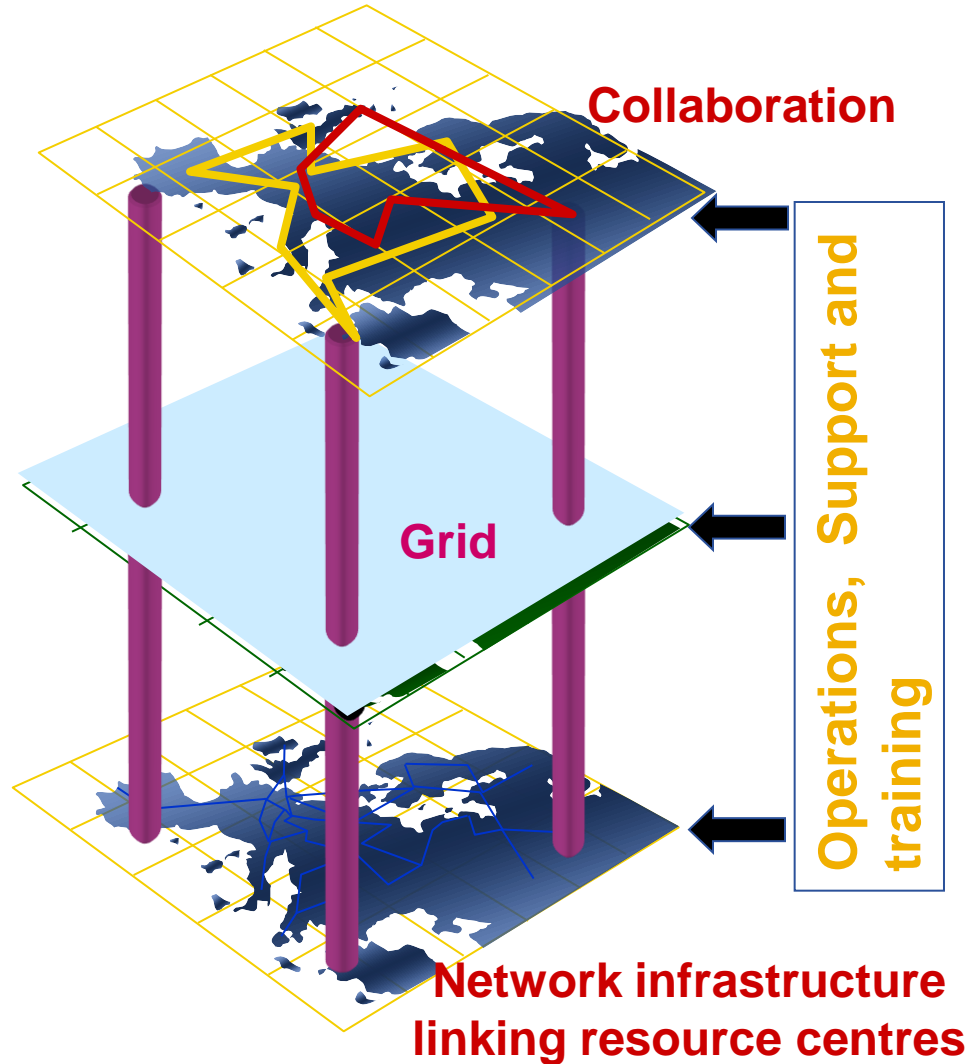
# What is the Grid?

- The *World Wide Web* provides seamless access to information that is stored in many millions of different geographical locations



- The *Grid* is an emerging infrastructure that provides seamless access to computing power and data storage capacity distributed over the globe





- **Collection of geographically distributed heterogeneous resources**

*“Most generalized, globalized form of distributed computing”*

- **“An infrastructure that enables flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources”**

***Ian Foster and Carl Kesselman***

- **Offers access to a virtual and very powerful computing system**
- **A user does not care, in which resource his / her job / jobs is going to be executed**

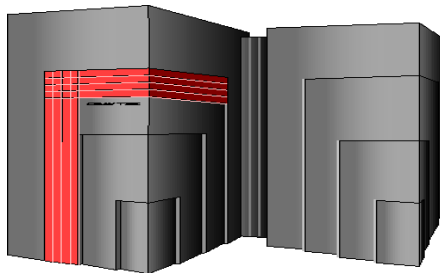
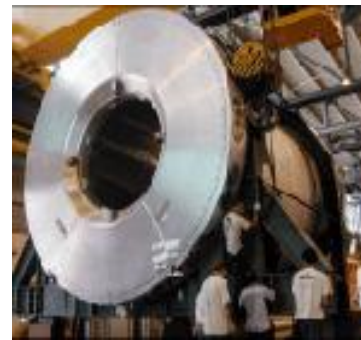
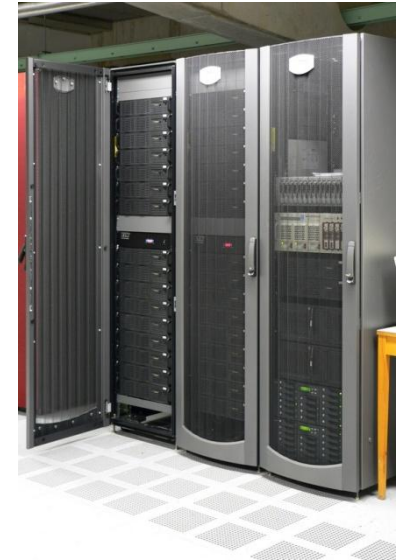
- In general terms:

A Grid is the combination of networked resources and the corresponding middleware, which provides services for the user.

- An entity that is going to be shared

such as:

- ✓ computational units
- ✓ storage units
- ✓ sensors
- ✓ visualization tools
- ✓ software
- ✓ licenses
- ✓ experience



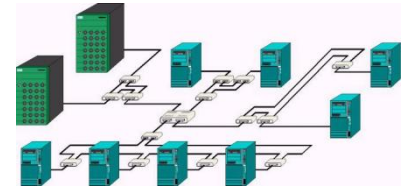
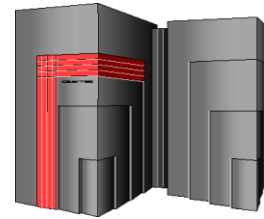
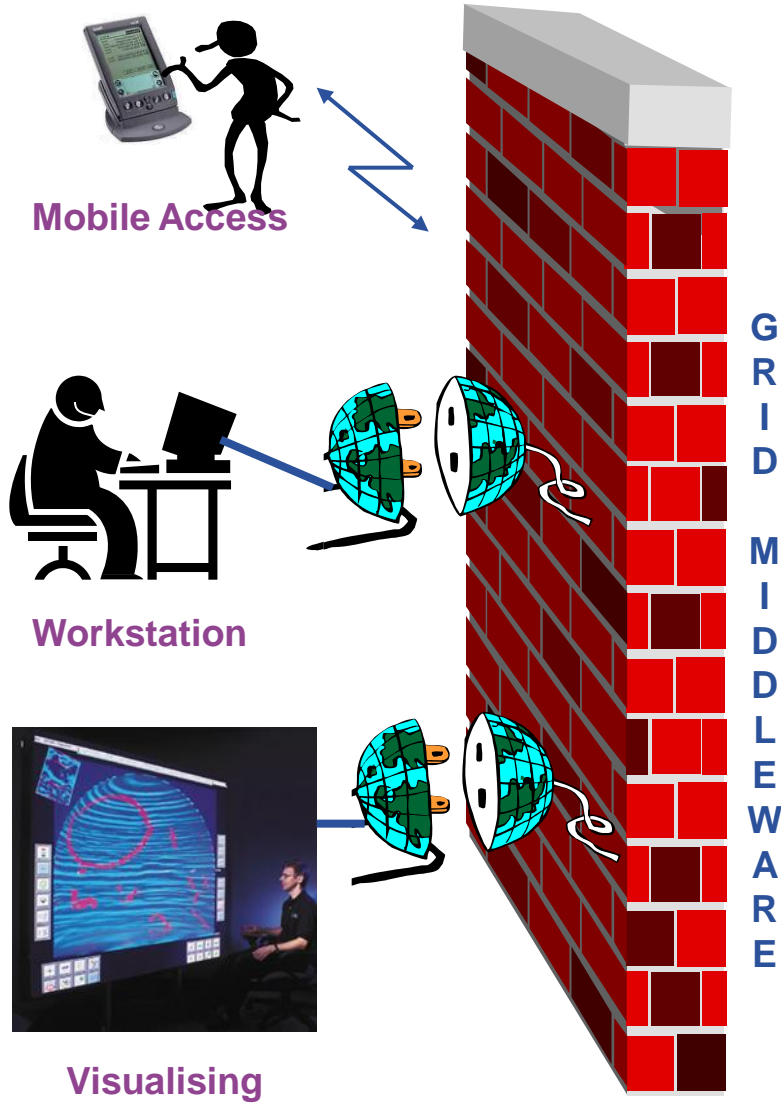


- **Resource sharing**
  - Geographically distributed resources offer computational power, storage capacity and bandwidth to the users
- **Secure and reliable access**
  - Authentication
  - Authorization
  - Access policy
- **Open standards**
- **Co-operation among people belonging to different organizations, institutes, groups**

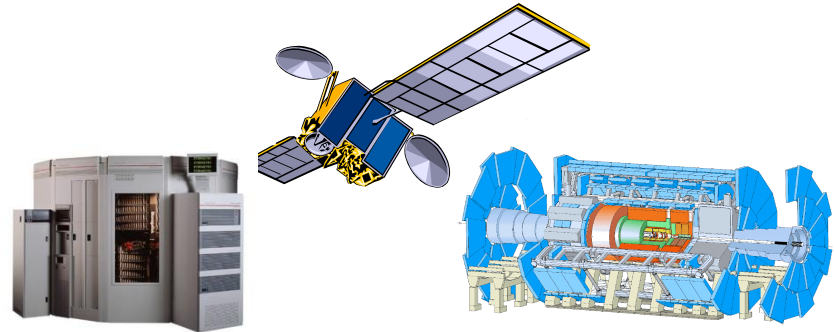
- **Lack of central control**
  - Where things run?
  - When jobs run?
  - What are the permissions of each user?
- **Shared resources**
  - Variability
- **Communication and coordination**
  - Different sites implies different administration rules, users, institutional goals and other constraints

- **Complex computation**
- **Data that can't fit on one site**
- **Data owned by multiple sites**
- **Applications that need to run faster and in a more complex manner**

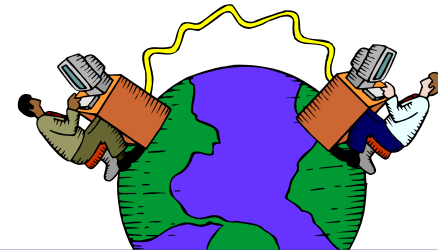
# Grid metaphorically ...



Supercomputer, PC-Cluster



Data-storage, Sensors, Experiments



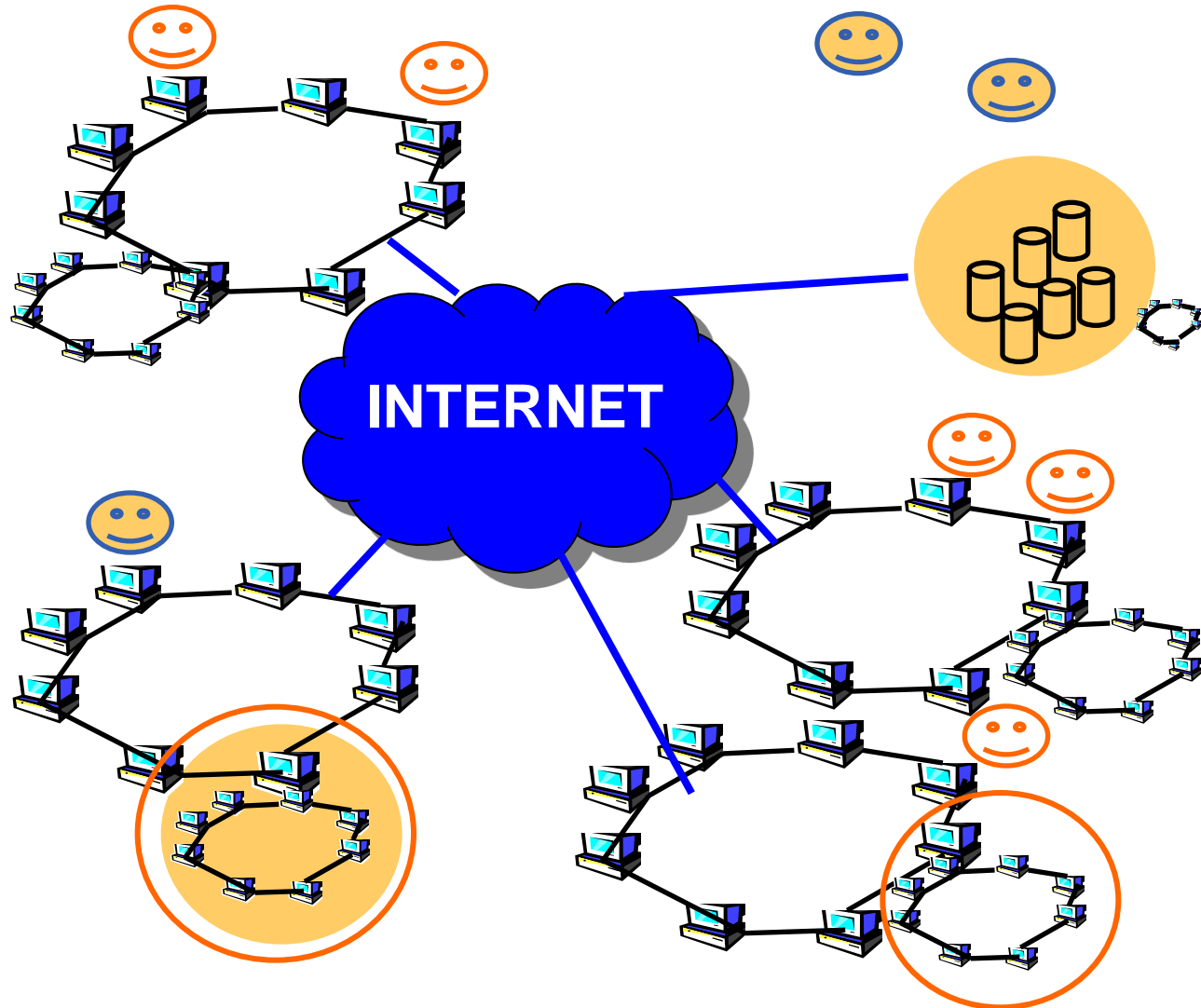
Internet, networks

GRID MIDDLEWARE

- **Development of networking technology (doubling every nine months or so over the last years) and high-speed networks**
  - ✓ **widespread penetration of optical fibers**
  - ✓ **wireless connections**
  - ✓ **new Internet technologies (ADSL, WiMax)**
  
- **Moore's law everywhere**
  - ✓ **Instruments, detectors, sensors, scanners, ...**
    - ⇒ **Organising their effective use is the challenge**
  
- **Applications require a huge amount of computations to be executed and the collaboration among scientists**

- **Science that became feasible and promiscuous by resource sharing (sharing of data, scientific instruments, computational resources, colleagues) across the Internet**
  - ✓ Often very compute intensive
  - ✓ Often very data intensive (both creating new data and accessing very large data collections) – data deluges from new technologies
  - ✓ **Crosses organisational and administrative boundaries**

- gLite middleware runs on each shared resource to provide
  - Data services
  - Computation services
  - Security service
- Resources and users form Virtual organisations: basis for collaboration
- Distributed services (both people and middleware) enable the grid



- **Virtual Organization**

“A set of individuals and / or institutions defined by highly controlled sharing rules, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share and the conditions under which sharing occurs”

*Ian Foster*

- **Abstract entities grouping users, institutions and resources in the same administrative domain**

↗ **What is going to be shared ?**

✓ resources

✓ software

✓ special equipment

✓ licenses

✓ services

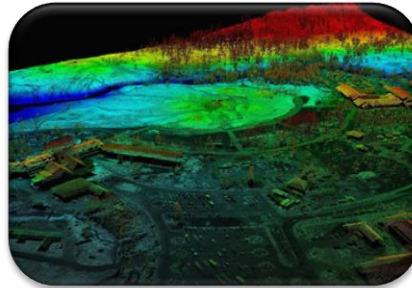
✓ Internet bandwidth



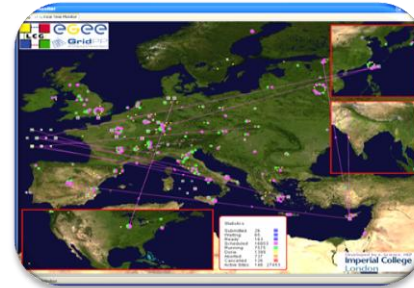
- Astrophysics, astro-particle physics
- Biomedical and Bioinformatic Applications
- Computational chemistry
- Earth sciences
- Finance
- Fusion
- Geophysics
- High-energy physics
- Infrastructure
- Other ...
  
- Our regional VO: SEE
- VO for trainings : hgdemo
  
- **List of existing VOs**
  - <http://cic.gridops.org/index.php?section=home&page=volist#1>



What is the Grid?



Grid paradigms



Enabling Grid for E-science (EGEE)

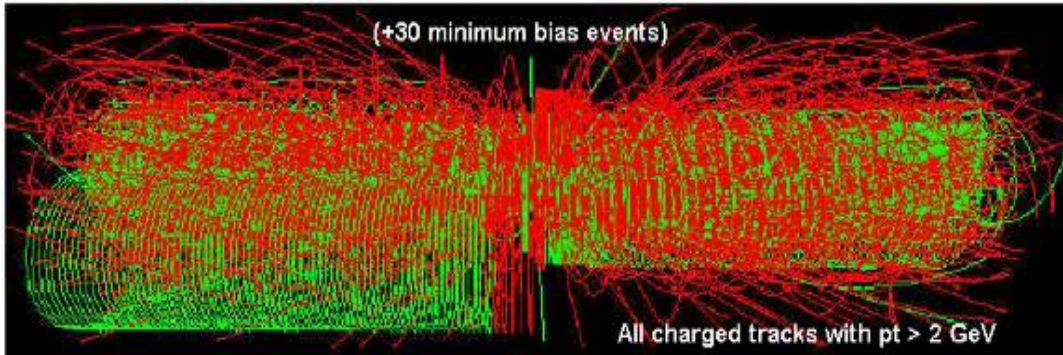


gLite middleware



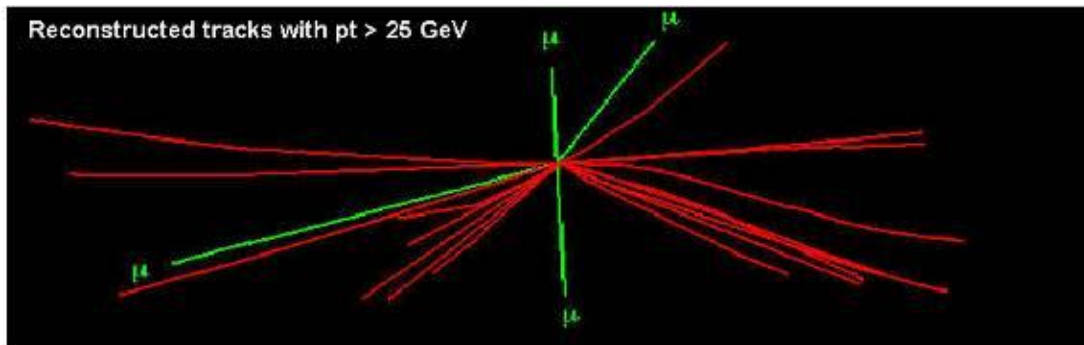
HellasGrid Taskforce

## Starting from this event (particle collision) ...

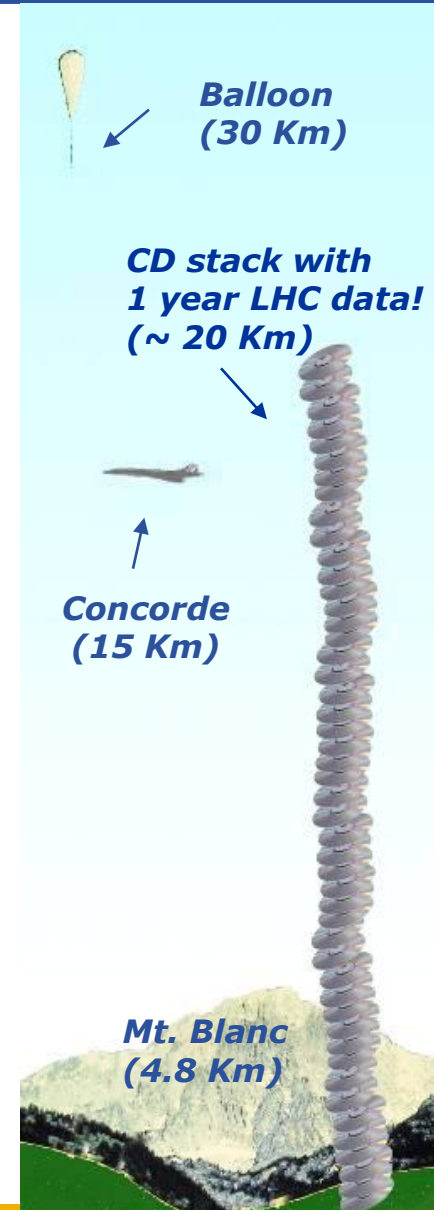
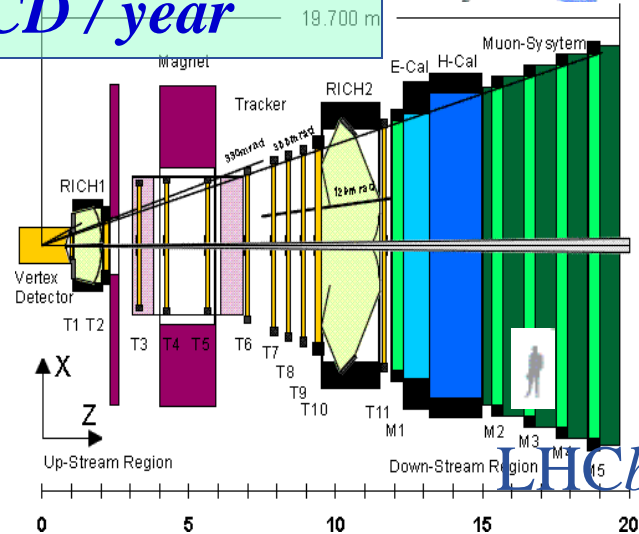
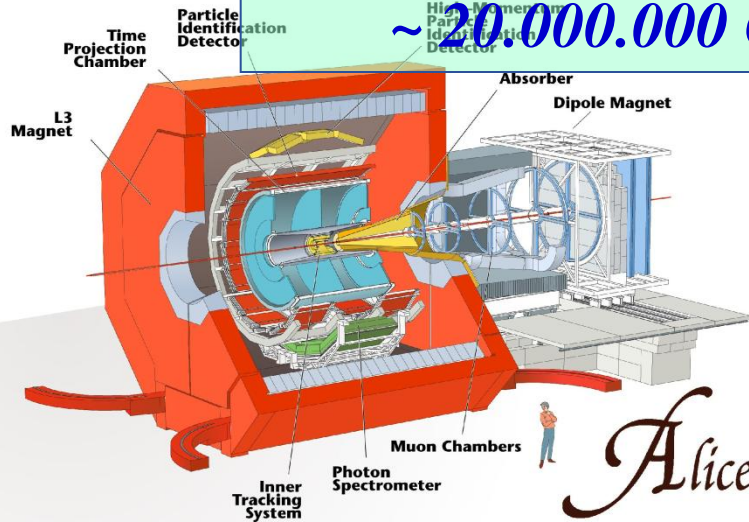
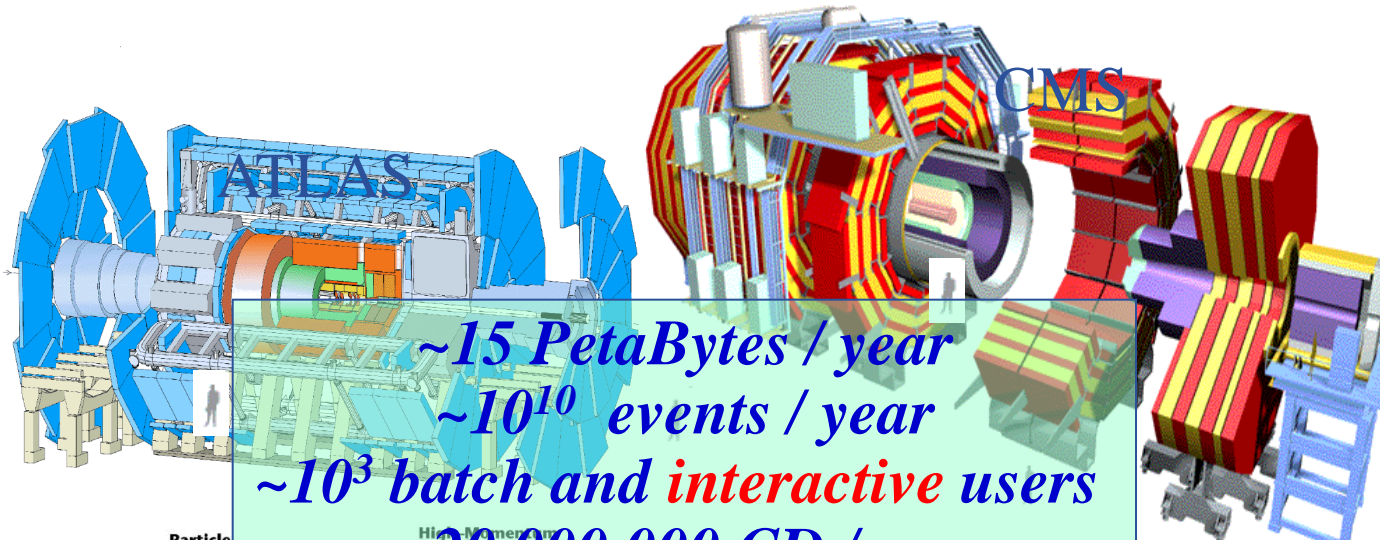


- ✓ Data Collection
- ✓ Data Storage
- ✓ Data Processing

## You are looking for this “signature”...

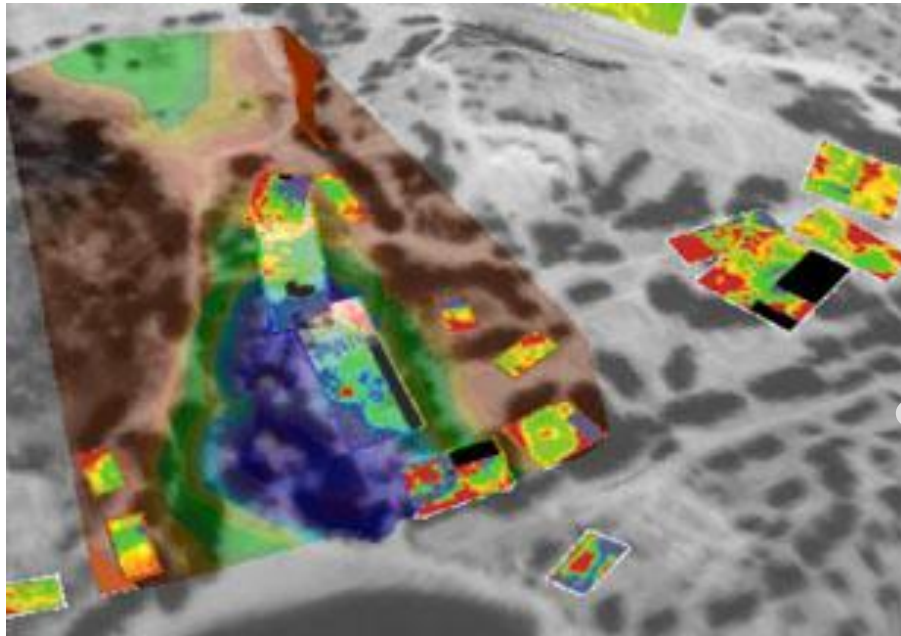


- *Selectivity: 1 in 10<sup>13</sup>*
- ✓ *Like looking for 1 person in a thousand world populations!*
- ✓ *Or for a needle in 20 million haystacks!*



- The LHC Computing Grid Project (LCG) was born to prepare the computing infrastructure for the simulation, processing and analysis of the data of the Large Hadron Collider (LHC) experiments.
- ⇒ The processing of the enormous amount of data, that will be generated, will require large computational and storage resources and the associated human resources for operation and support.
- ⇒ Preparation of a common infrastructure of
  - ✓ libraries
  - ✓ tools
  - ✓ frameworks
 required to support the physics application software



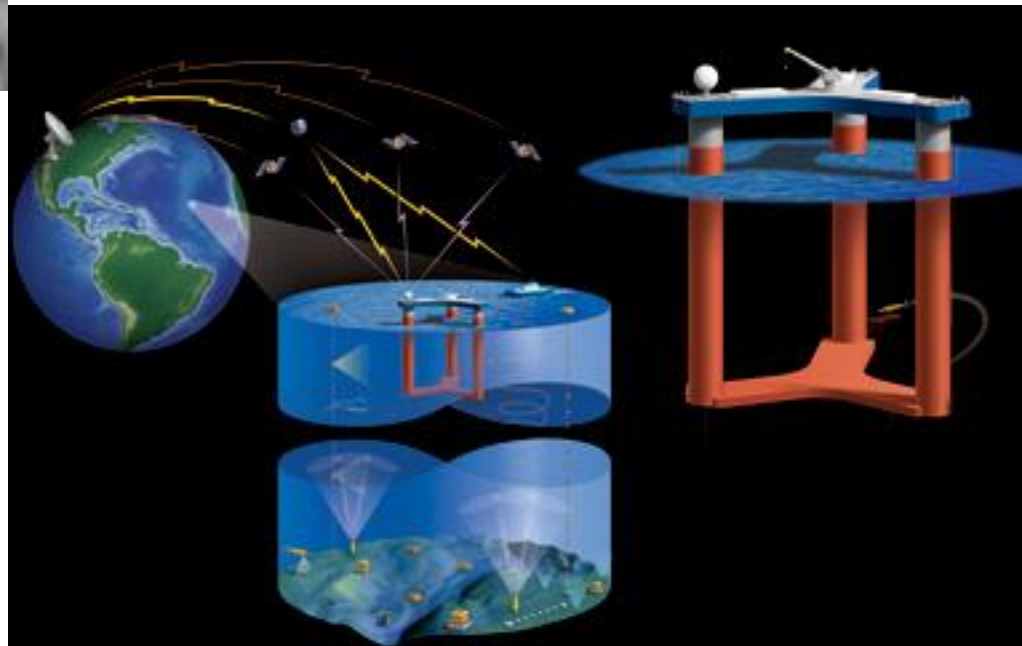


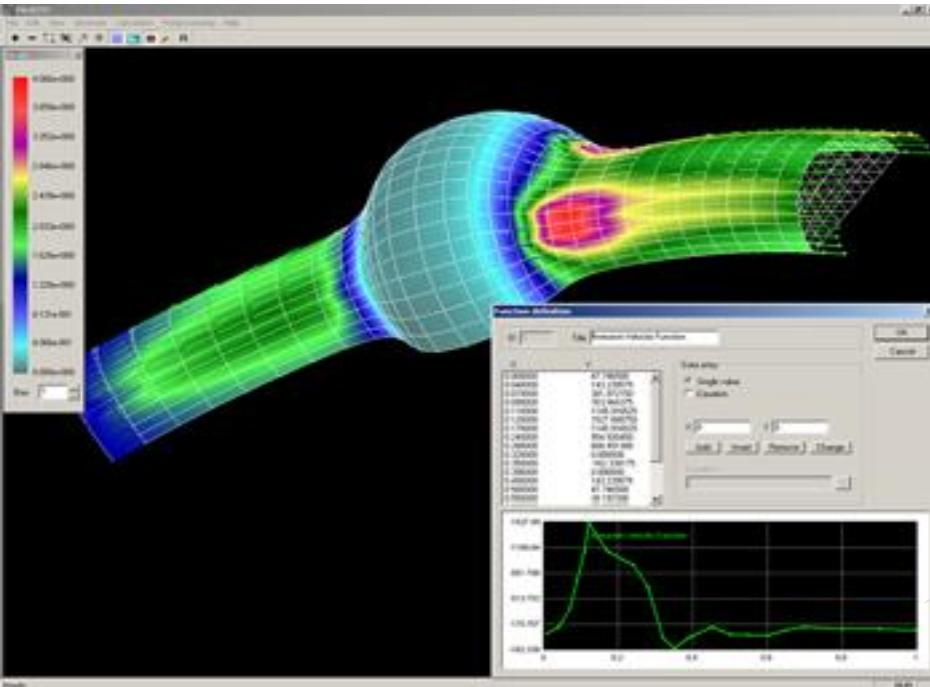
### ArchaeoGrid

Create a computer model that weaves together data from many sources and predicts feedback interaction

### LOOKING

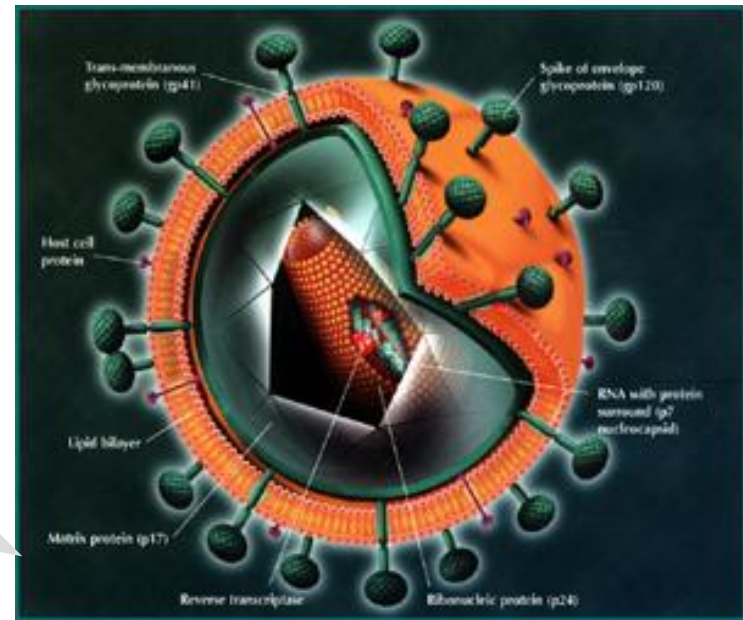
Observe and analyze data streams in real time. A sensor grid with thousand of different sensors providing real time data and measurements from ocean-going researchers enabling an enormous data grid infrastructure.





**Parallel Blood Flow Simulation**  
 Allows surgeons to perform virtual stent surgery until they get it just right. It combines parameters such as blood velocity and pressure with a series of medical images to automatically create a 3D computational model.

**ViroLab**  
 Aims to create a collaborative virtual laboratory for grid-based decision support for viral disease treatment. HIV treatment in the increasingly common case of HIV drug resistance is mainly studied. Virolab “vertically” integrates biomedical information relating to viruses, patients and literature resulting in a rule-based decision support system for drug ranking.

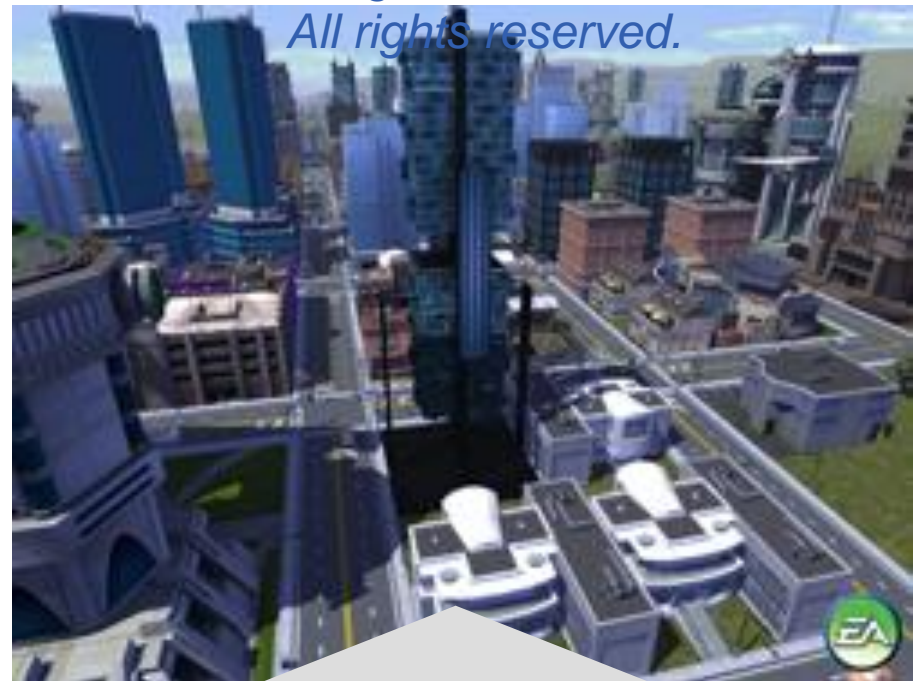




## Polar Grid

A planned project for an advance cyberinfrastructure, empowering smaller universities, and provide scientists with a gateway to teraflops of power: enough to drive new and improved high-performance simulations and enable measurement and prediction of ice sheet response to climate change and effect on ocean levels.

Image © *Electronic Arts Inc.*  
All rights reserved.



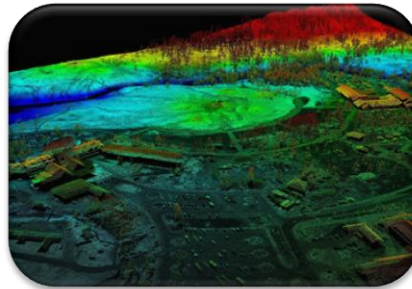
## MoSES (Modelling and Simulation for e-Social Science)

Runs predictive models integrating real Census data, survey data, healthcare data of UK population. Determine the impact of different policy decisions and various social aspects like increasing life expectancy, immigration, aging population.

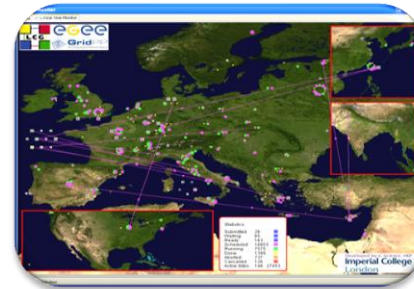




What is the Grid?



Grid paradigms



Enabling Grid for E-science (EGEE)



gLite middleware



HellasGrid Taskforce

- **EGEE objective:**

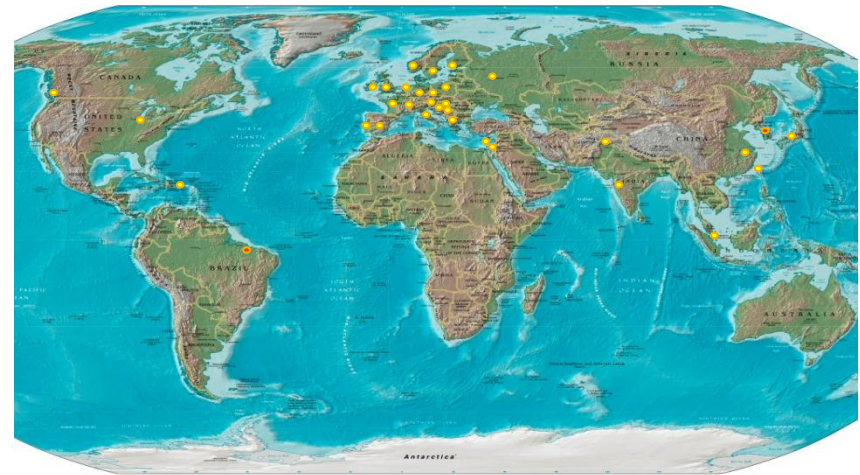
*“to establish a seamless European Grid infrastructure for the support of the European Research Area (ERA)”*

- **EGEE:**

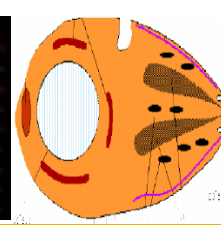
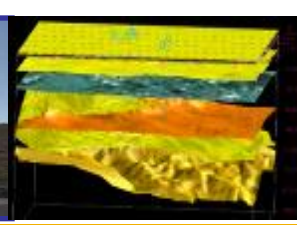
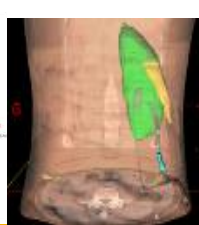
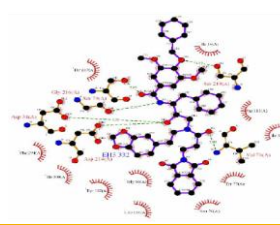
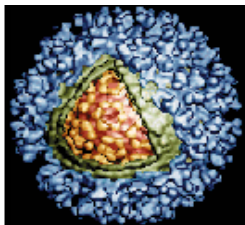
- Accomplished all of its objectives
- Scope expanded beyond Europe

- **EGEE-II :**

- Full capacity from day one
- Large-scale, production-quality infrastructure across the European Research Area and beyond
- Supporting a wide range of applications
- Staff with extensive knowledge of Grid technology



- **Mission:**
  - ✓ Manage and operate production Grid infrastructure for the European Research Area
  - ✓ Interoperate with e-Infrastructure projects around the globe (Open Standards-GGF) and Contribute to Grid standardisation efforts
  - ✓ Incorporate new users from the industry and from the research community as well assuring the best possible training and support
- **Support applications deployed from diverse scientific communities:**
  - ✓ High Energy Physics
  - ✓ Earth Sciences
  - ✓ Computational Chemistry
  - ✓ Fusion
  - ✓ Biomedicine
  - ✓ Astrophysics
  - ✓ Finance, Multimedia
  - ✓ Geophysics
  - ...
- Prepare for a permanent/sustainable European Grid Infrastructure (in a GÉANT2-like manner)



- EGEE III objectives:
  - expand and optimize EGEE, by continuous operation of the infrastructure
  - support for more user communities
  - add of further computational and data resources
  - prepare the migration of the existing production European Grid from a project-based model to a sustainable federated infrastructure based on National Grid Initiatives for multi-disciplinary use

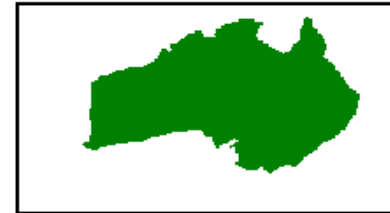
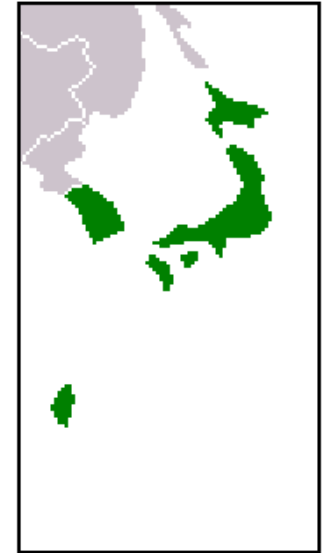
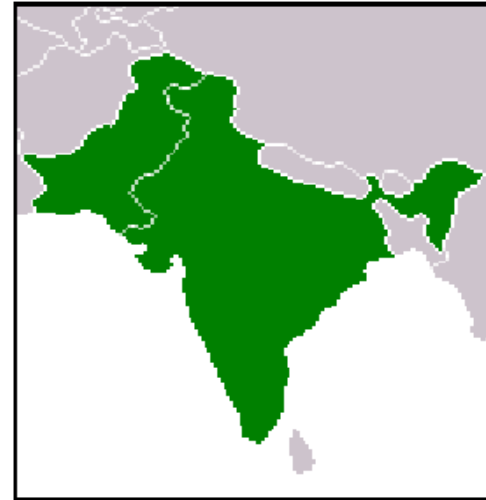
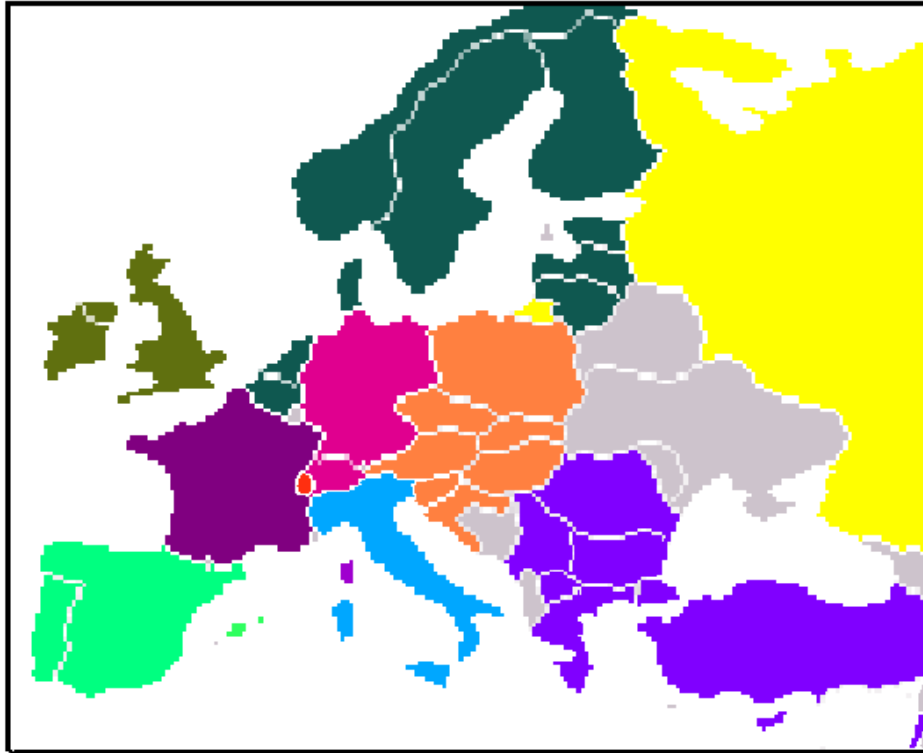
↘ Available infrastructure to the Research and Academic community 24 hours per day and 7 days per week

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

↘ Participants:  
✓ 50 countries

↘ Consists of:  
 ✓ 250 sites  
 ✓ ~ 68000 CPUs  
 ✓ ~ 20 PB  
 ✓ ~140 VOs  
 ✓ Massive data transfers  
 > 1.5 GB/s

(October 2009)



## Regional Operations Centres (ROC)

- Front-line support for user and operations issues
- Provide local knowledge and adaptations
- One in each region – many distributed

## User Support Centre (GGUS)

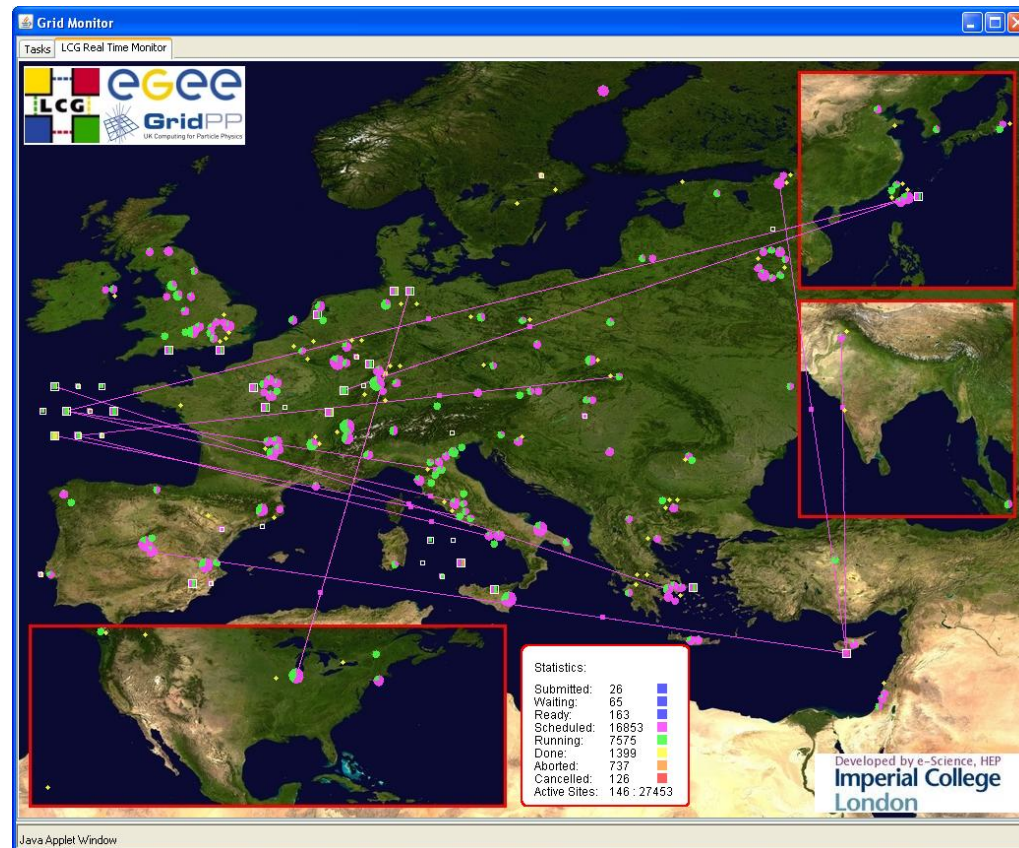
- In FZK: provide single point of contact (service desk), portal

• <https://gus.fzk.de/pages/home.php>

## Real Time Monitor

- Java tool
- Displays jobs running (submitted through RBs)
- Shows jobs moving around world map in real time, along with changes in status

<http://gridportal.hep.ph.ic.ac.uk/rtm/>



- **GÉANT2 is the seventh generation of pan-European research and education network, successor to the pan-European multi-gigabit research network GÉANT**
- 30 European National Research and Education Networks (NRENs) in 34 countries
- administrated by DANTE (*Delivery of Advanced Network Technology to Europe*)
- **It provides:**
  - Basic IPs services
  - Quality of service levels
- **Greece NREN**
  - GRNET

## GÉANT2

The world-leading research and education network for Europe.

★ Connect ★ Communicate ★ Collaborate

Grid paradigms

GEANT2 is operated by DANTE on behalf of Europe's NRENs.

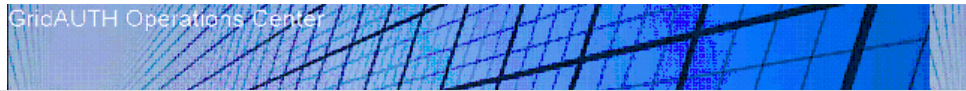
Austria	Belgium	Czech Republic	Denmark	France	Germany	Greece	Ireland	Italy	Netherlands	Poland	Portugal	Spain	Sweden	Switzerland	United Kingdom
---------	---------	----------------	---------	--------	---------	--------	---------	-------	-------------	--------	----------	-------	--------	-------------	----------------

\*Connections between these countries are part of GÉANT2 (the pan-European network)

GÉANT2 is co-funded by the European Commission  
its 6th R&D Framework Programme.

- **Operating system:**
  - Linux (+GNU utilities), usually a RHEL3-like, for example Scientific Linux
- **Middleware:**
  - gLite v3.x
- **Libraries and Applications**
  - Defined by the system and VOs administrators' foresight
  - The user can install and execute its own programmes





• <http://www.grid.auth.gr/pki/seegrid-ca/>

**GridAUTH Operations Center**

On July 2007, the GridAUTH Operations Center of the Aristotle University of Thessaloniki, implemented the SeeGrid Certification Authority, in order to facilitate the needs for Grid computing in the wider area of the Balkans.

The scope of the SeeGrid CA is to provide PKI services to the SEE countries - members of the SeeGrid project - that did not have the opportunity to establish their own national grid PKI infrastructure.

The SeeGrid CA is operated by the GridAUTH Operations Center at the Aristotle University of Thessaloniki under the supervision of GRNET as the SeeGrid Coordinator. GRNET is owned and supervised by the General Secretariat of Research and Technology, Greek Ministry of Development.

Find more info in the SeeGrid CA CP/CPS

EGEE Enabling Grids for E-science

SEE-GRID Wiki

Navigation: Main Page, Community portal, Current events, Recent changes, Random page, Help, Donations

Contents [hide]

- 1 News
- 2 SEE-GRID Infrastructure
  - 2.1 Monitoring and Operational tools
  - 2.2 Core Services
- 3 Site Admins
  - 3.1 For new sites
  - 3.2 Site Installation and Configuration
    - 3.2.1 Middleware guides
    - 3.2.2 Configuration guides
  - 3.3 Site Certification Procedure
  - 3.4 Installation of Specific Services and Tools
  - 3.5 Middleware Assessments
- 4 Users
  - 4.1 User Tools
- 5 Developers
- 6 SEE-GRID Operations Organization and Procedures
  - 6.1 SEE-GRID Operations
  - 6.2 CA, RA
  - 6.3 Support Organization
  - 6.4 Security Incidence Response
- 7 FAQs
  - 7.1 For Site Admins
  - 7.2 For Users
- 8 Contacts

• [http://wiki.egee-see.org/index.php/SEE-GRID\\_Wiki](http://wiki.egee-see.org/index.php/SEE-GRID_Wiki)



Path: Home page > Documentation > User documentation

**User Documentation**

**Overview**

Does the Grid sound complex and obscure? Want to know more? For complete beginners, please have a look at this introduction from CERN: <http://gridcafe.web.cern.ch/gridcafe/>

The following pages guide you through the process of using the Grid. The instructions cover three different kinds of perspective users:

1. New experimental users; and application.
2. New production users; those regional.
3. Experienced production users the regional Grid.

Also, this document gives the acceptable usage.

The general documentation for user access to [sa1.web.cern.ch/egee-sa1/using.htm](http://sa1.web.cern.ch/egee-sa1/using.htm). Re familiarised with how to access the grid. If you are a cluster administrator (and want to Europe cluster RC Administrator's documentation [here](#)

Search:  Google search

• **EGEE Helpdesk:**  
<https://helpdesk.egee-see.org/index2.php>

VOMS is the Virtual Organization Membership Service, a central database for VO membership information.

This is the web user interface of the VOMS Admin service for the seegrid VO. It provides services relating to VO membership for VO users and VO managers.

Please select an item from the services listed on the left side of this page.

- New user registration
- My requests
- FOR VO MANAGERS**
  - Administer the VO
  - Handle requests
  - Check audit data
- CONFIGURATION**
  - Configuration information
  - List all VOs on this server

VOMS Admin 1.2.14  
 Release 1  
 Copyright © 2005-2009 EGI, ETC  
 in support of the EGEE project.

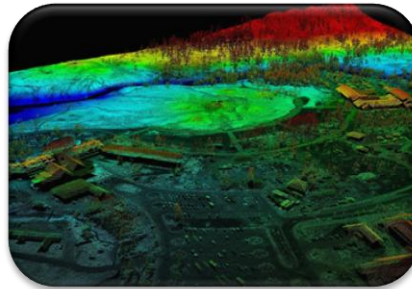
You are logged in as "IC=GR/O=HellasGrid/OU=rtus.gr/CN=Athanasia Assiki" certified by "IC=GR/O=HellasGrid/CN=HellasGrid CA".

• [http://www.egee-see.org/User\\_documentation.php](http://www.egee-see.org/User_documentation.php)

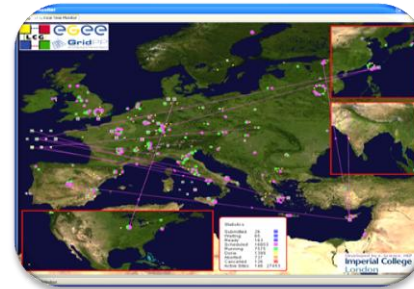
• <https://voms.irb.hr:8443/edg-voms-admin/seegrid/index.html>



What is the Grid?



Grid paradigms



Enabling Grid for E-science (EGEE)



gLite middleware

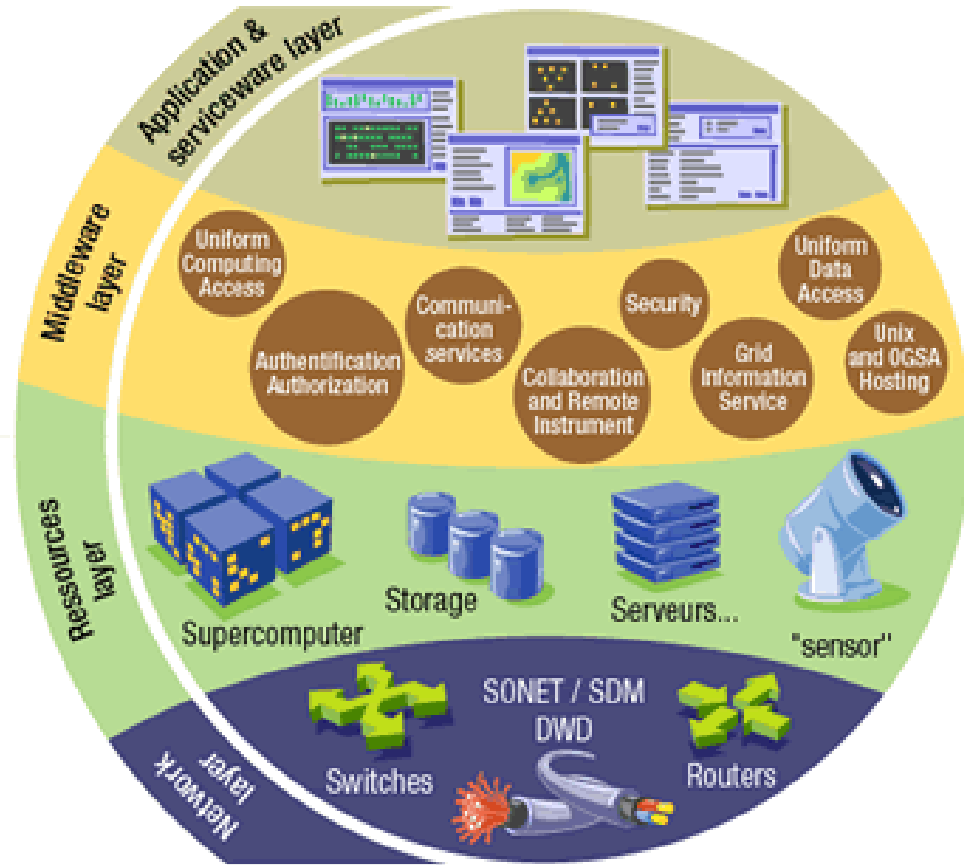


HellasGrid Taskforce

- The Grid relies on advanced software, called **middleware**, which interfaces between resources and the applications

- The Grid middleware:**

- Basic services
  - Secure and effective access to resources
- High level services
  - Optimal use of resources
  - Authentication to the different sites that are used
  - Job execution & monitoring of progress
  - Problem recovery
  - Transfer of results back to the user



- **Part of the EGEE project**
- **Next generation middleware for grid computing**
- **In its development participate from different academic and industrial European centers**
- *Provides services for computing element, data management, accounting, logging and bookeping, information and monitoring, service discovery, security, workload management*

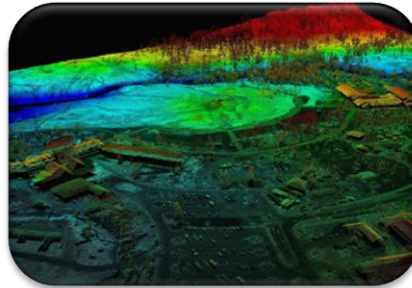


- **Applications running on a PC need to be adapted to run on a Grid**
  - Include new layers of grid-enabled software
- **Aspects to consider:**
  - Acquiring authentication credentials
  - Locating available data
  - Structure of jobs for computational tasks
  - Initiate computations
  - Monitor progress of computations and data transfers
  - Collection of output results

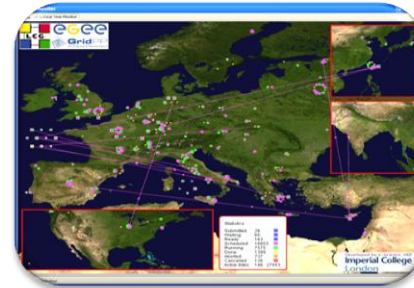




What is the Grid?



Grid paradigms



Enabling Grid for E-science (EGEE)



gLite middleware



HellasGrid Taskforce

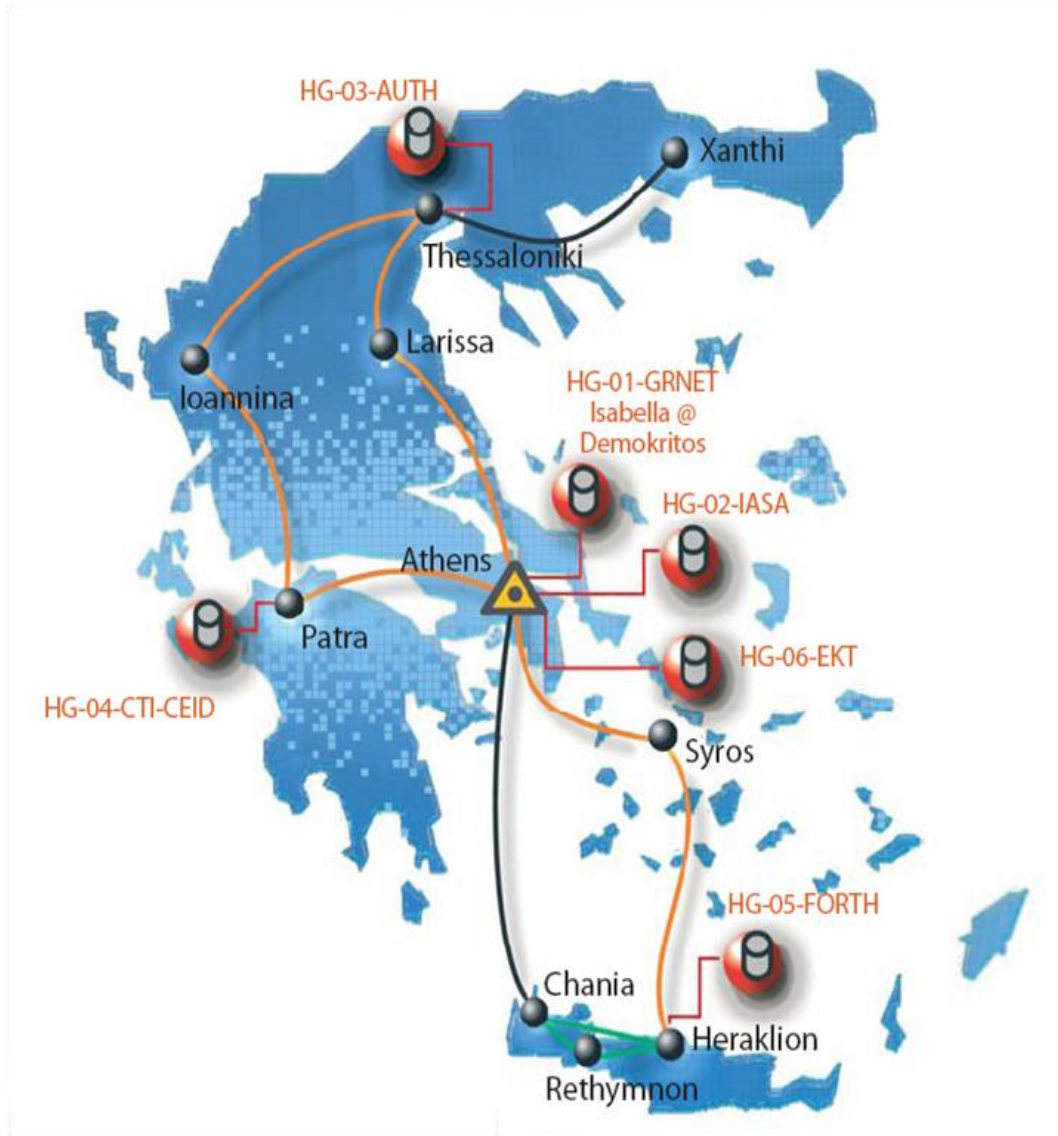
- **HellasGrid I**

- Located at N.C.S.R. Demokritos (a.k.a. Isabella)
- 34 dual Intel **P4 Xeon @ 2.8GHz, 1GB RAM, 2x 70GB SCSI HDD**, 2x Gbit
- IBM FAStT900 Storage Area Network
  - 2x Redundant Fiber Channel Controllers with 1Gbyte Cache each
  - 70x146.8GB= **10,276TB raw storage capability**, over 5 disk shelves
- Tape Library ~30 TBytes, integrated monitoring
- December 2004

- **HellasGrid II**

- 5 sites: EKT (>220), ΙΕΣΕ (48), ΑΠΘ (128), ΙΤΕ (128), ΙΤΥ (128)
- ~700 CPUs **x86\_64, 2 GB RAM, 1x 80GB SATA HDD**, 2x Gbit
- ~20 TBytes storage space in SAN (5x 4TBs)
- ~50 TBytes Tape Library in National Documentation Center

<http://www.hellasgrid.gr/>





- Main site: HG-01-GRNET (Isabella, [cslab@ICCS/NTUA](mailto:cslab@ICCS/NTUA))
- HG-02...HG-06 sites @ (NDC, IASA, AUTH, FORTH, CTI)

**CSLab**



- HG CA and VOMS (AUTH):

<http://www.grid.auth.gr/pki/seegrid-ca/>



- Helpdesk (CTI):

[user-support@hellasgrid.gr](mailto:user-support@hellasgrid.gr)



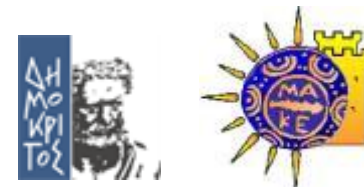
- Regional monitoring tools (FORTH):

<http://hellasgrid-ui.ics.forth.gr/acctROC/>



- Apps support (IASA):

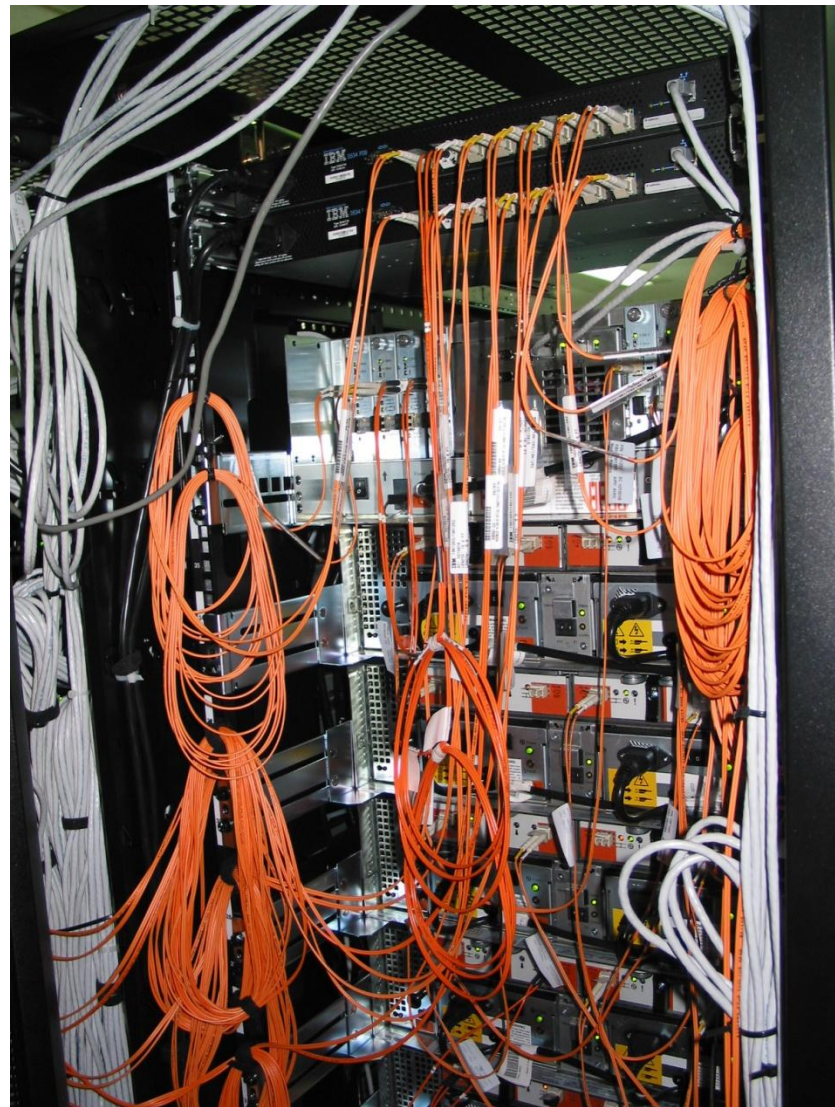
[application-support@hellasgrid.gr](mailto:application-support@hellasgrid.gr)





# HellasGrid I Infrastructure, Isabella

Enabling Grids for E-science





[https://access.hellasgrid.gr/register/registration\\_form](https://access.hellasgrid.gr/register/registration_form)

Not Logged In
Final Beta

## HellasGrid User Registration

HellasGrid  
National Grid Infrastructure

Εγγραφή νέων χρηστών > Φόρμα Εγγραφής Χρήστη

**Διαδικασία**

**Καταχώρηση προσωπικών στοιχείων**

Αίτηση ψηφιακού πιστοποιητικού

Αποστολή αιτήσεως

---

**Επικοινωνία**

GridAUTH Support

Εγγραφή νέου χρήστη

Όνομα	<input type="text" value="Ελληνικά"/>	<input type="text" value="Αγγλικά"/>
Επώνυμο	<input type="text" value="Ελληνικά"/>	<input type="text" value="Αγγλικά"/>
E-mail	<input type="text"/>	
Οργανισμός	<input type="text" value="Ανωτάτη Σχολή Καλών Τεχνών"/>	
Τηλέφωνο εργασίας	<input type="text"/>	
Επιστημονικός τομέας	<input type="text" value="Άλλο"/>	
Τμήμα	<input type="text"/>	
Ιδιότητα	<input type="text" value="Ερευνητής"/>	

Υπάρχοντες Χρήστες

Αν στο παρελθόν είχατε αποκτήσει ψηφιακό πιστοποιητικό από την Α.Π. HellasGrid CA το οποίο έχει πλέον λήξει, συμπληρώστε στο πεδίο που ακολουθεί το e-mail σας για να προχωρήσετε στη διαδικασία αίτησης καινούργιου ψηφιακού πιστοποιητικού.

**Αναζήτηση E-mail**

GridAUTH (HellasGrid User Registration)

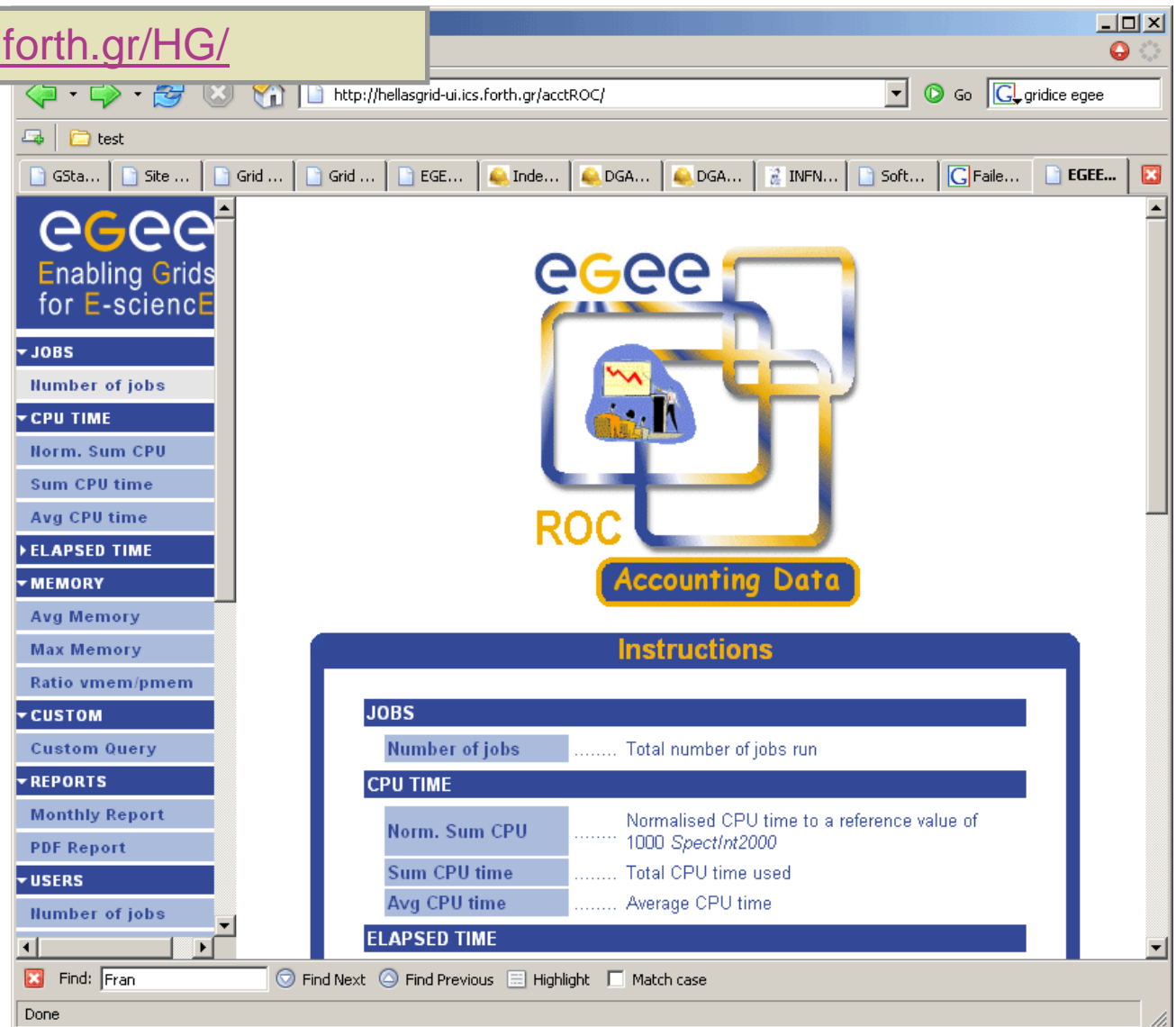








<http://hellasgrid-ui.ics.forth.gr/HG/>



The screenshot shows a web browser window displaying the HellasGrid Accounting Data interface. The browser address bar shows the URL <http://hellasgrid-ui.ics.forth.gr/acctROC/>. The page features a navigation menu on the left with categories like JOBS, CPU TIME, ELAPSED TIME, MEMORY, CUSTOM, REPORTS, and USERS. The main content area displays the 'ROC Accounting Data' logo and a table of statistics under the heading 'Instructions'.

Instructions	
<b>JOBS</b>	
Number of jobs	..... Total number of jobs run
<b>CPU TIME</b>	
Norm. Sum CPU	..... Normalised CPU time to a reference value of 1000 <i>SpectInt2000</i>
Sum CPU time	..... Total CPU time used
Avg CPU time	..... Average CPU time
<b>ELAPSED TIME</b>	

- **Core Services**

- Central LCG File Catalog (LFC) for the users of the VOs:
  - eumed, hgdemo, see
- Resource Broker and Information Index (BDII) which can be accessed by the users of the VOs:
  - atlas, alice, lhcb, cms, dteam, sixt, biomed, esr, magic, compchem, see, planck, hgdemo, eumed
- Catch-All User Interface for HellasGrid
  - Registration is handled through the Hellasgrid User-Support Team
  - UI services are offered by all HG sites

- **Service Availability Monitoring**

- <https://mon.isabella.grnet.gr/sft/lastreport.cgi> (Need a valid HellasGrid Certificate)



Thank you!

- **Grid café:**  
<http://www.gridcafe.org/>
- **Open Grid Forum:**  
<http://www.gridforum.org/>
- **HellasGrid Task Force**  
<http://www.hellasgrid.gr/>
- **EGEE (Enabling Grids for E-science)**  
<http://www.eu-egee.org/>
- **The Globus Alliance**  
<http://www.globus.org/>
- **Grid Operations Centre**  
<http://goc.grid-support.ac.uk/gridsite/gocmain/>
- **gLite UserGuide**  
<http://glite.web.cern.ch/glite/documentation/>



- **EGEE – South East Europe**  
<http://www.egee-see.org/>
- **SEE-GRID**  
<http://www.see-grid.org/>
- **GRNET**  
<http://www.grnet.gr/>
- **gLite**  
<http://glite.web.cern.ch/glite/>
- **SEE-GRID Wiki**  
[http://wiki.egee-see.org/index.php/SEE-GRID\\_Wiki](http://wiki.egee-see.org/index.php/SEE-GRID_Wiki)
- **GOC Wiki**  
<http://goc.grid.sinica.edu.tw/gocwiki/>
- **SEEREN2**  
<http://www.seeren.org/>