

Εισαγωγή στο Grid, EGEE και το HellasGrid

Introduction to Grid, EGEE and HellasGrid

Athanasia Asiki

aassiki@cslab.ece.ntua.gr

*Computing Systems Laboratory,
National Technical University of Athens*

What is the Grid?

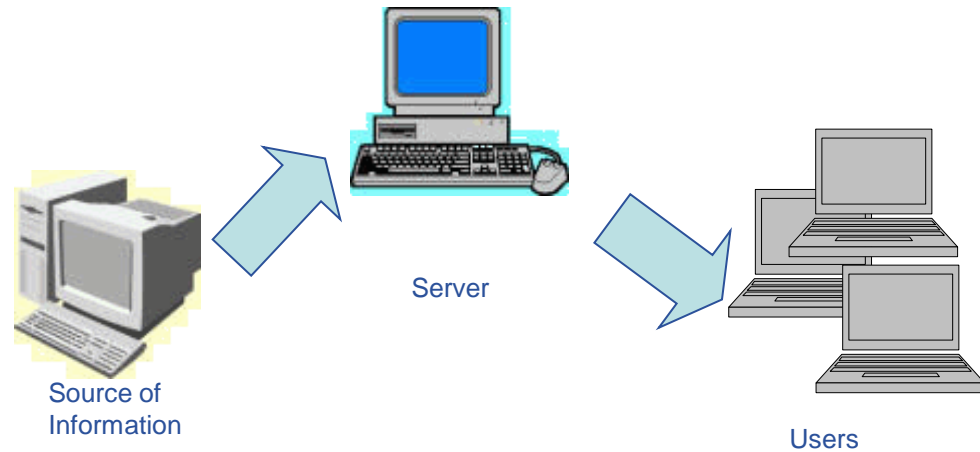
Grid paradigms

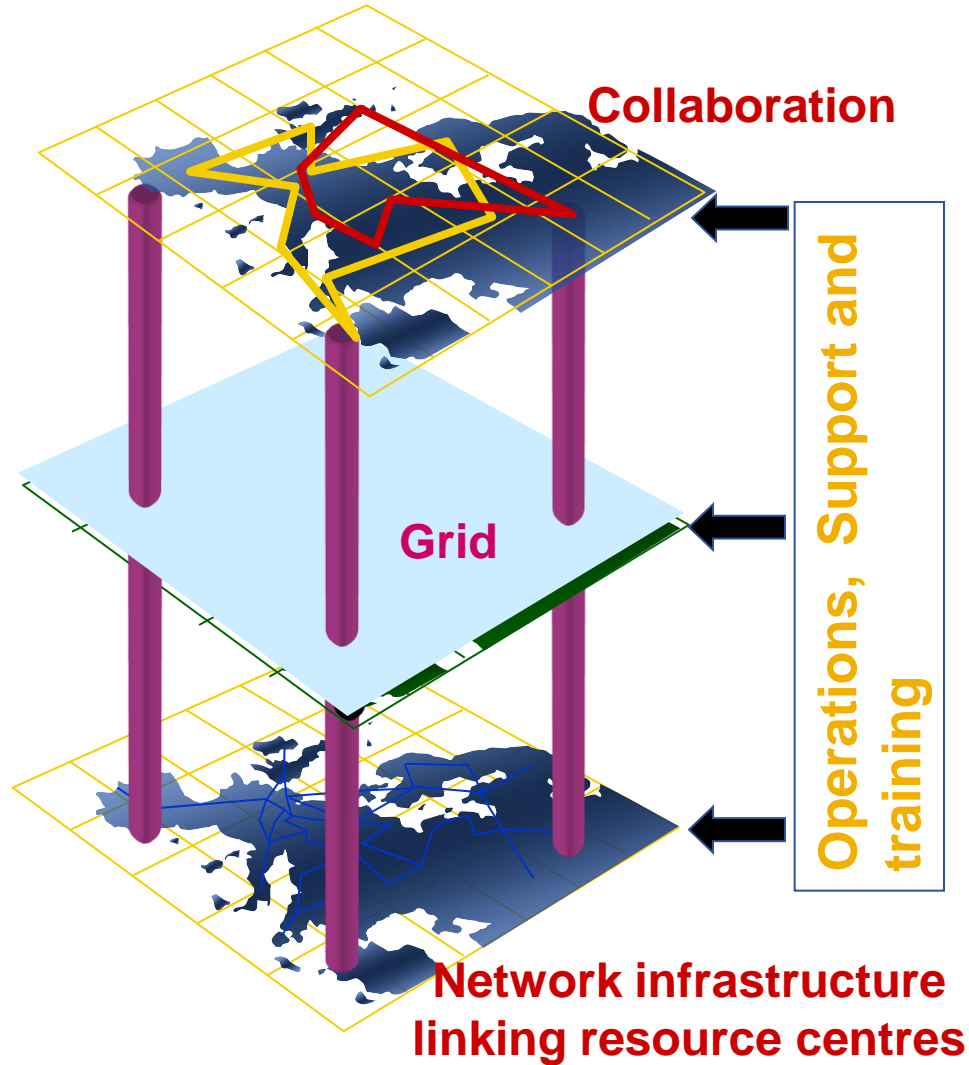
Enabling Grid for E-science (EGEE)

HellasGrid Taskforce



- 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

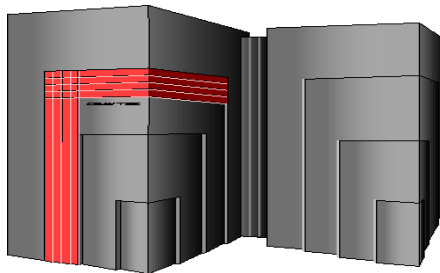
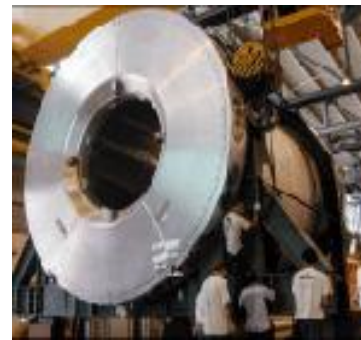
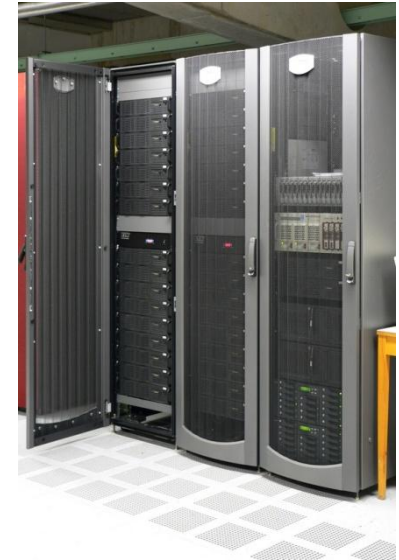
- **Offers access to a virtual and very powerful computing system**
- **A user does not care, in which resource his / her job / jobs is going to be executed**

- 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

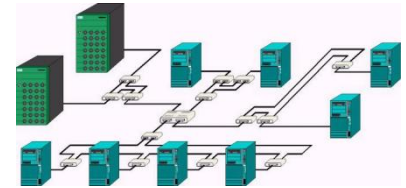
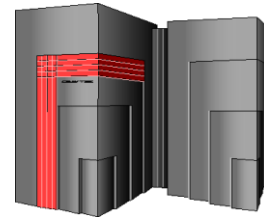
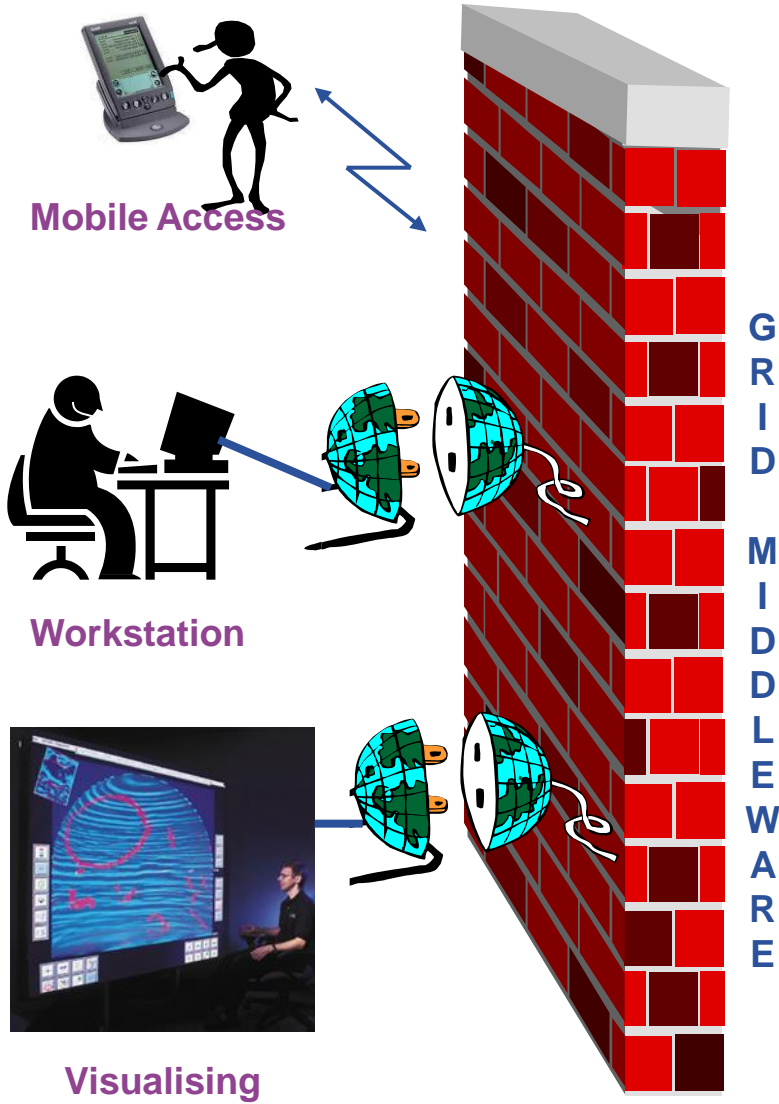
such as:

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

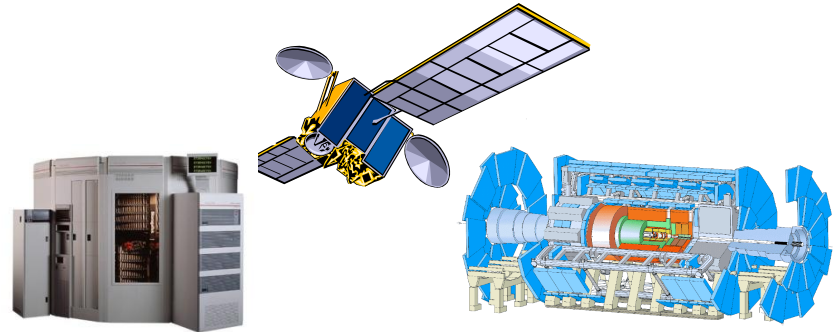


- **Resource sharing**
 - Geographically distributed resources offer computational power, storage capacity and bandwidth to the users
- **Secure and reliable access**
 - Authentication
 - Authorization
 - Access policy
- **Open standards**
- **Co-operation among people belonging to different organizations,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

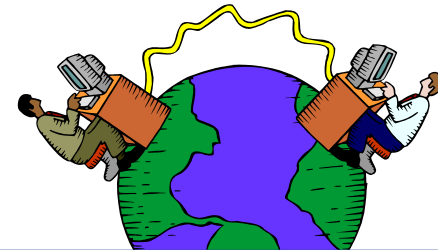
Grid metaphorically ...



Supercomputer, PC-Cluster



Data-storage, Sensors, Experiments

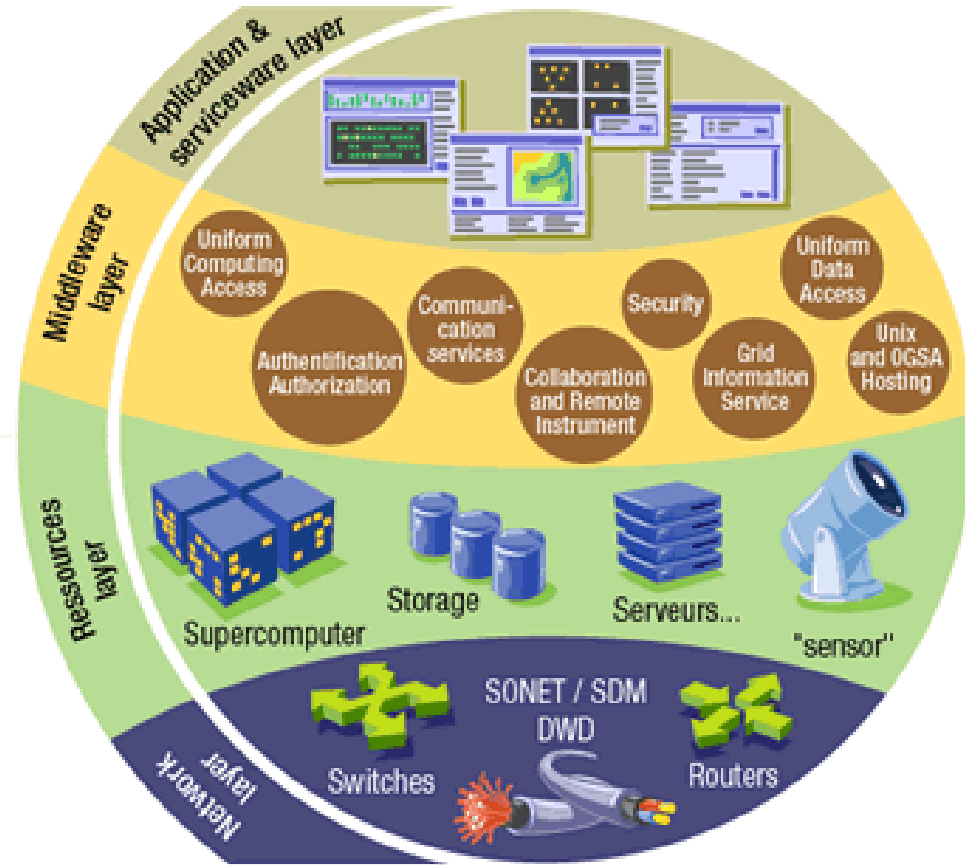


Internet, networks

- 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



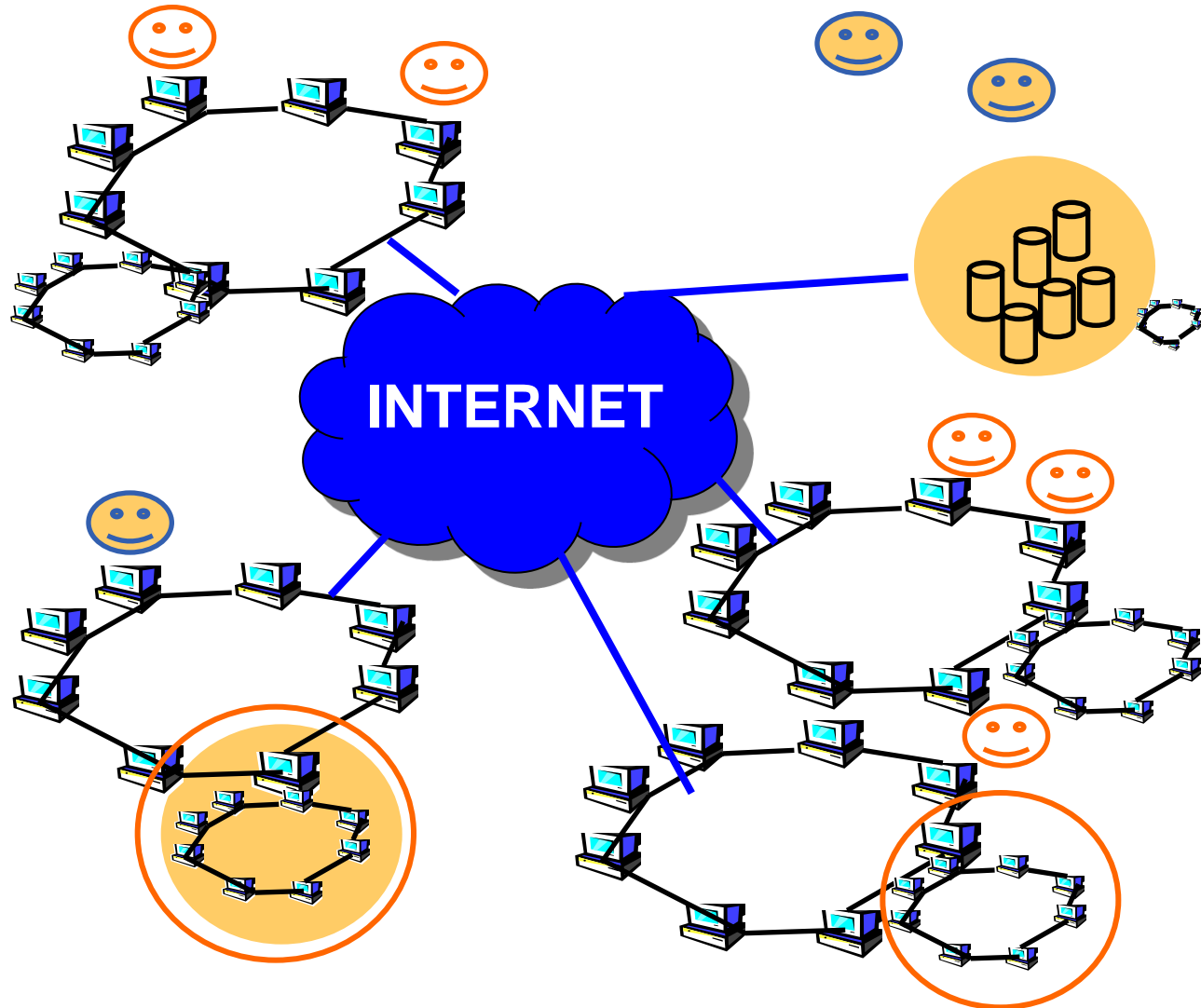
- **Development of networking technology (doubling every nine months or so over the last years) and high-speed networks**
 - ✓ **widespread penetration of optical fibers**
 - ✓ **wireless connections**
 - ✓ **new Internet technologies (ADSL, WiMax)**

- **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>

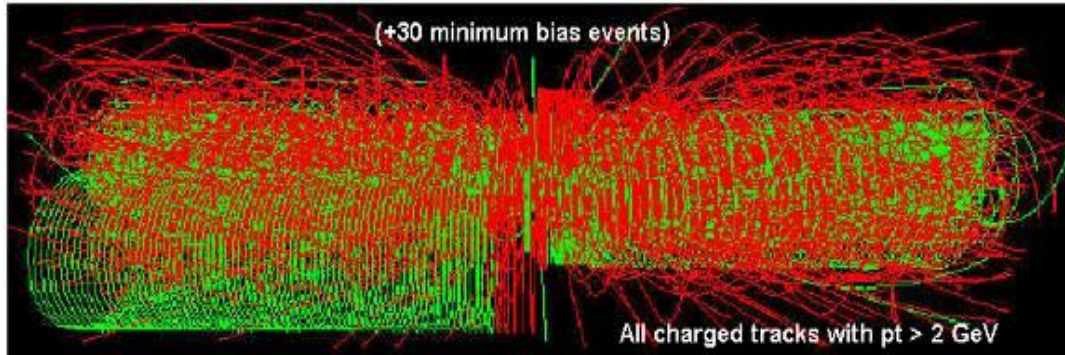
What is the Grid?

Grid paradigms

Enabling Grid for E-science (EGEE)

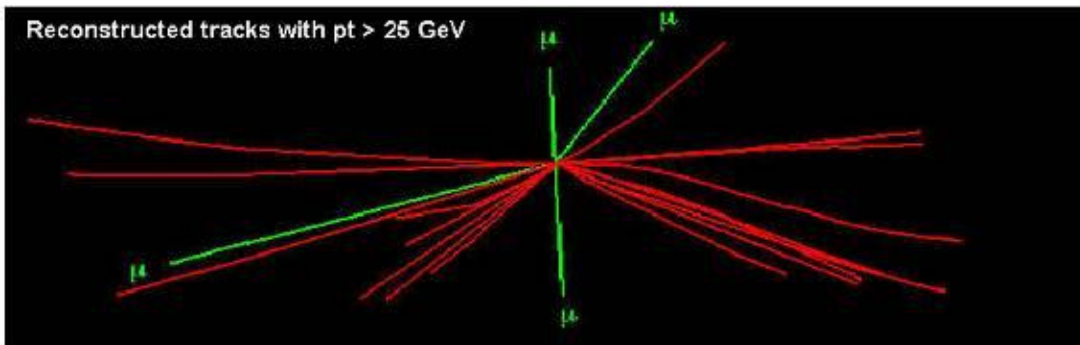
HellasGrid Taskforce

Starting from this event (particle collision) ...

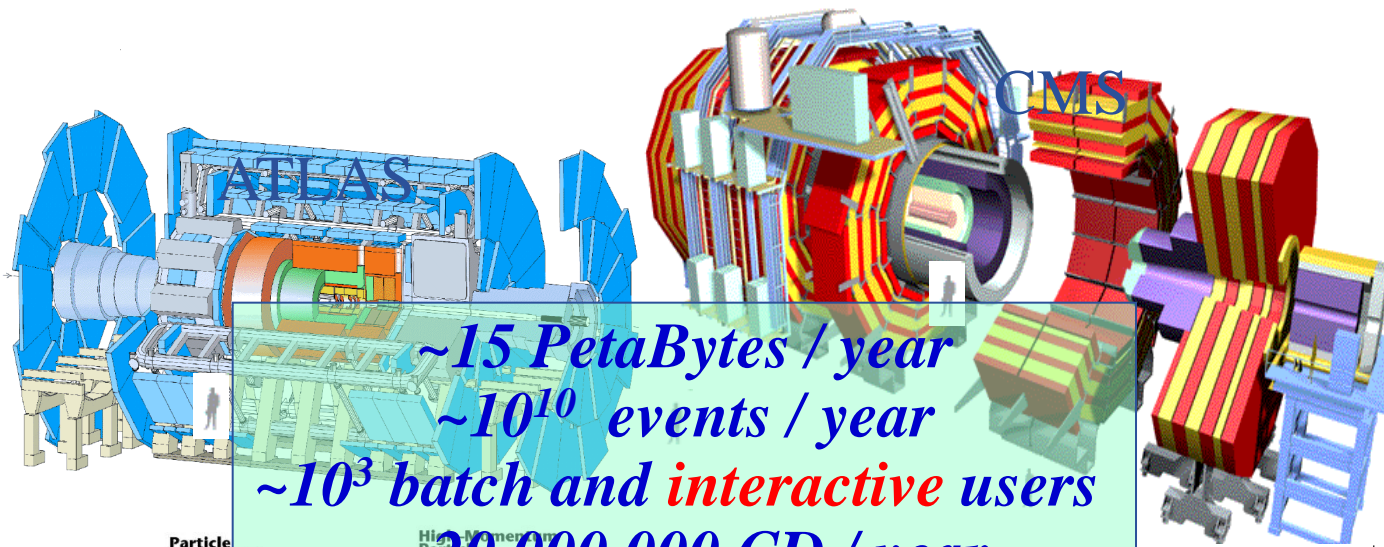


- ✓ Data Collection
- ✓ Data Storage
- ✓ Data Processing

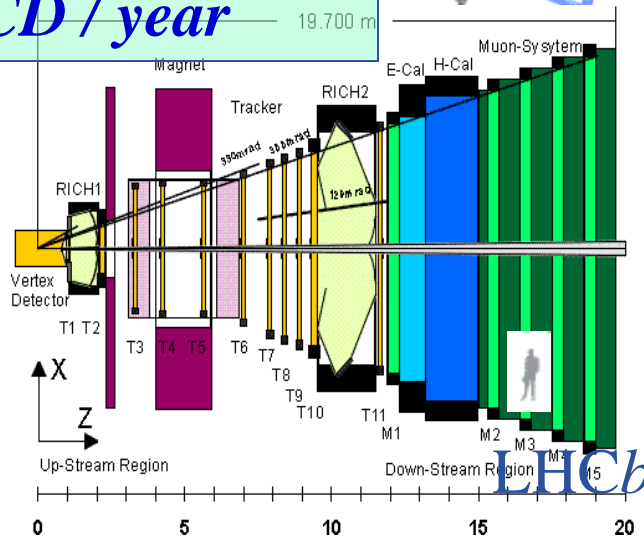
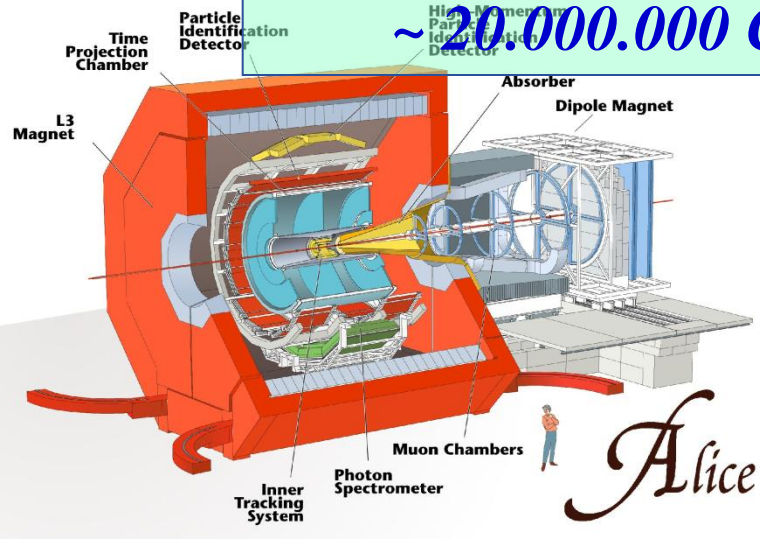
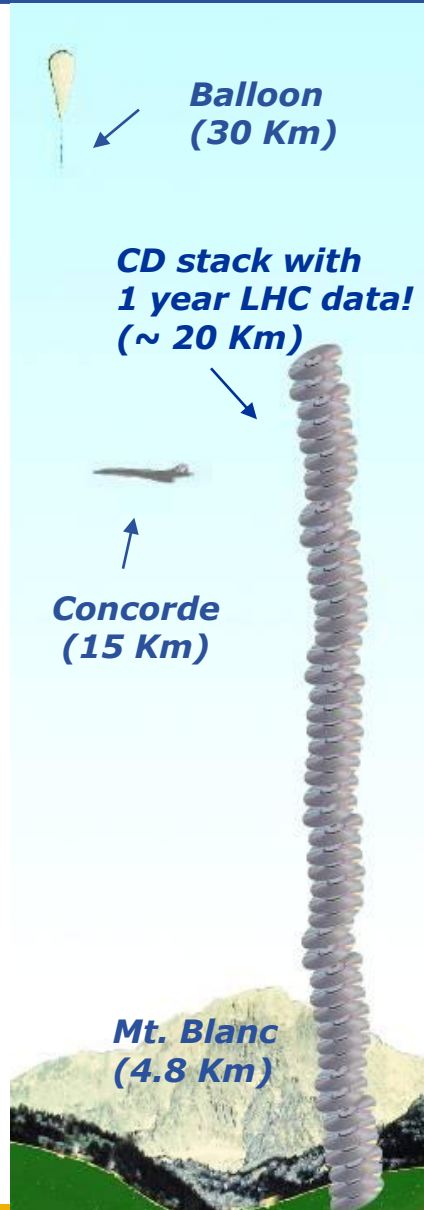
You are looking for this “signature”...



- *Selectivity: 1 in 10¹³*
- ✓ *Like looking for 1 person in a thousand world populations!*
- ✓ *Or for a needle in 20 million haystacks!*

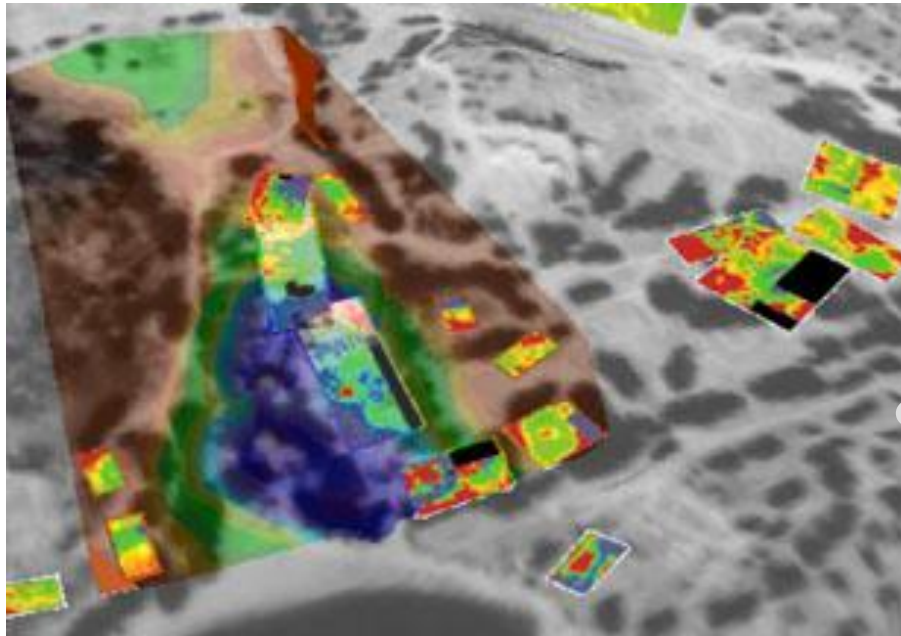


~15 PetaBytes / year
 ~ 10^{10} events / year
 ~ 10^3 batch and *interactive* users
 ~20.000.000 CD / year



- 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



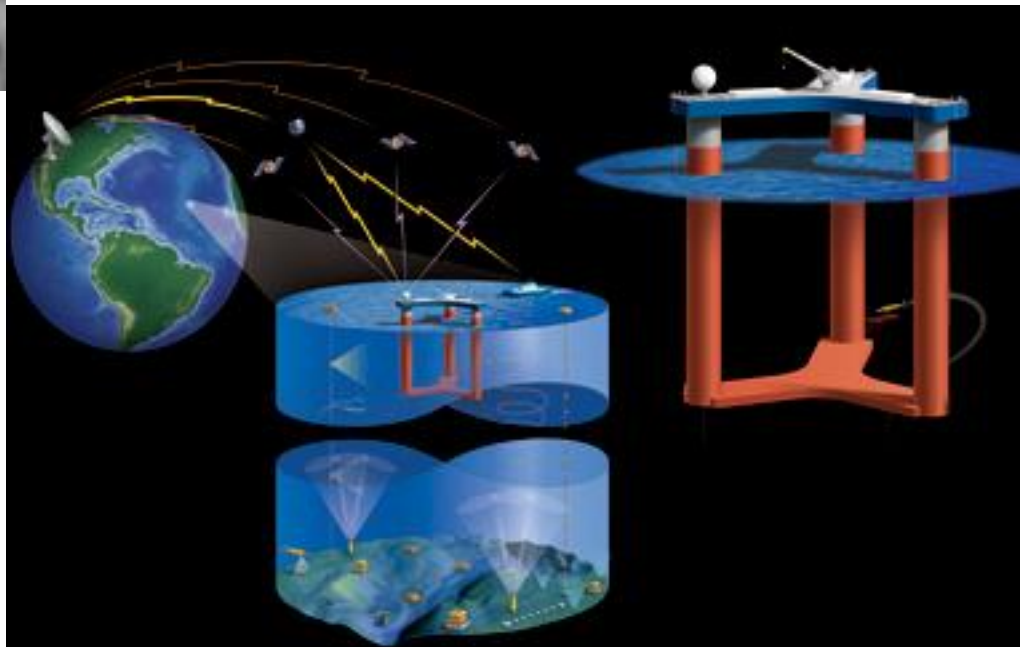


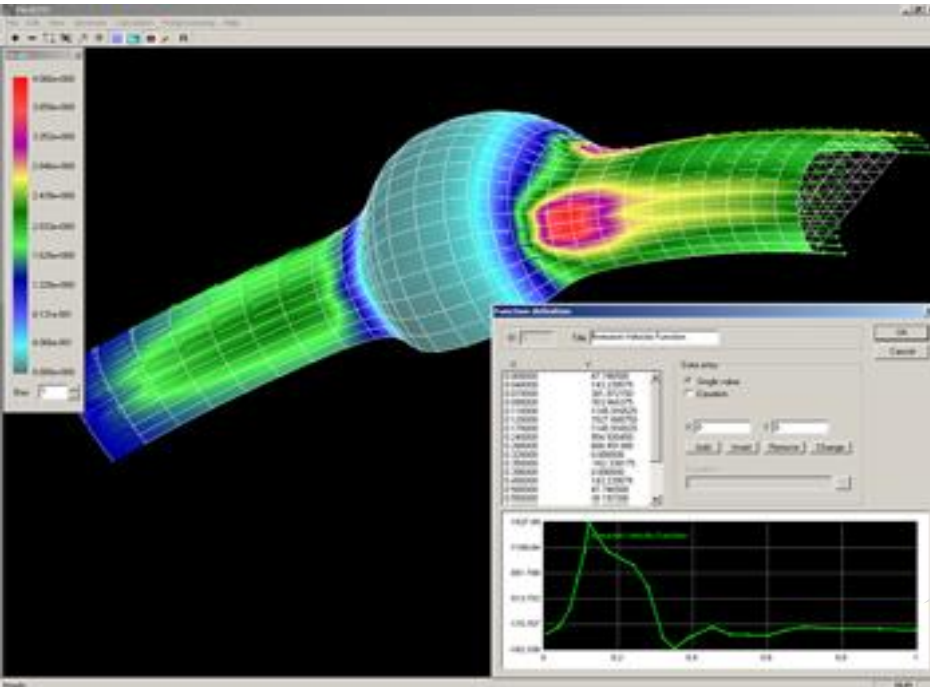
ArchaeoGrid

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

LOOKING

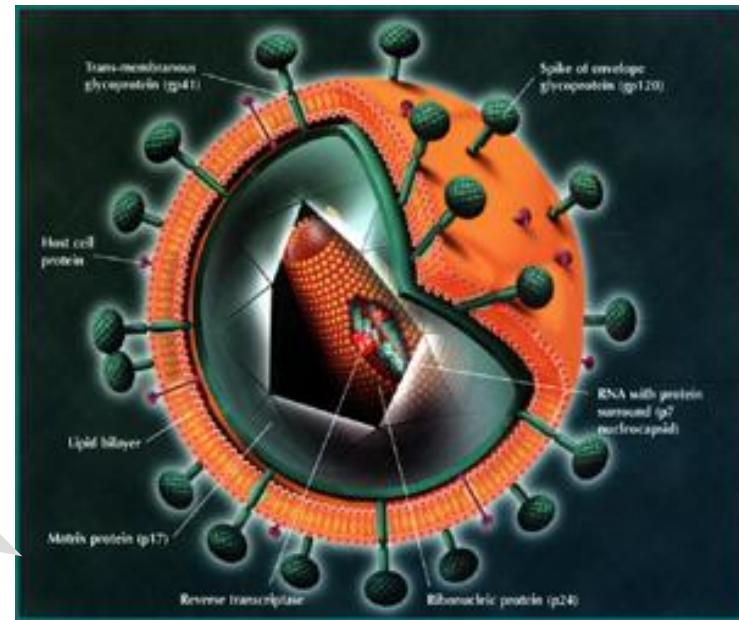
Observe and analyze data streams in real time. A sensor grid with thousand of different sensors providing real time data and measurements from ocean-going researchers enabling an enormous data grid infrastructure.





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)

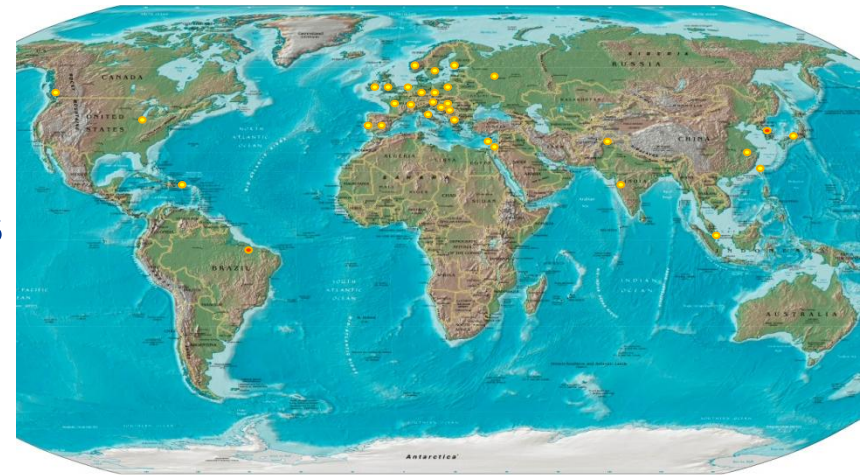
HellasGrid Taskforce

- **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
- Supporting a wide range of applications
- Staff with extensive knowledge of Grid technology

- **EGEE-II aims to provide a production quality Grid infrastructure across the European Research Area and beyond.**

↳ Started on 1 April 2006

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

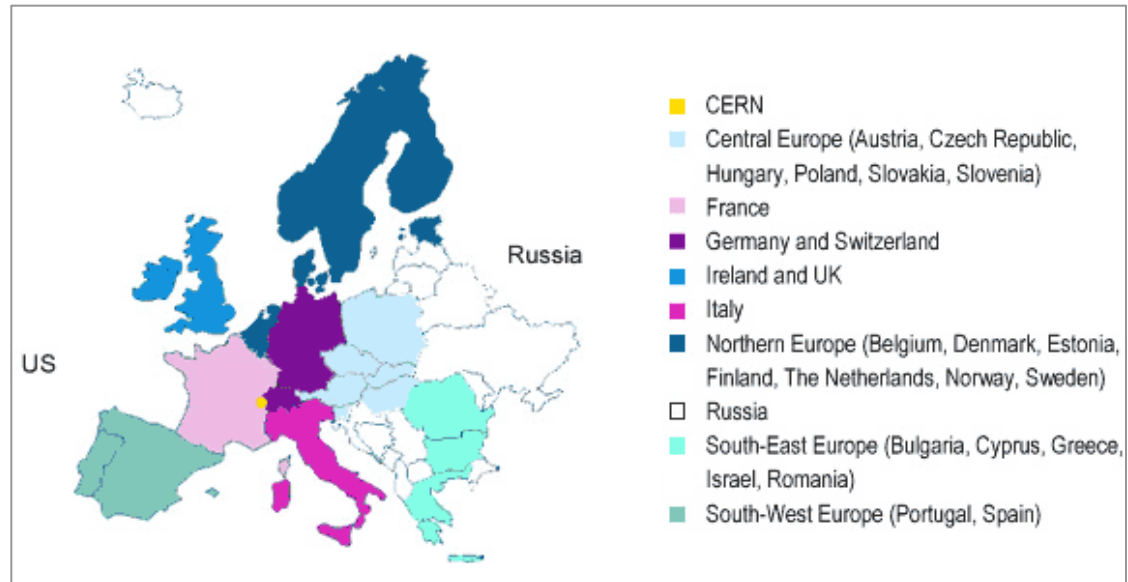
↳ Participants:

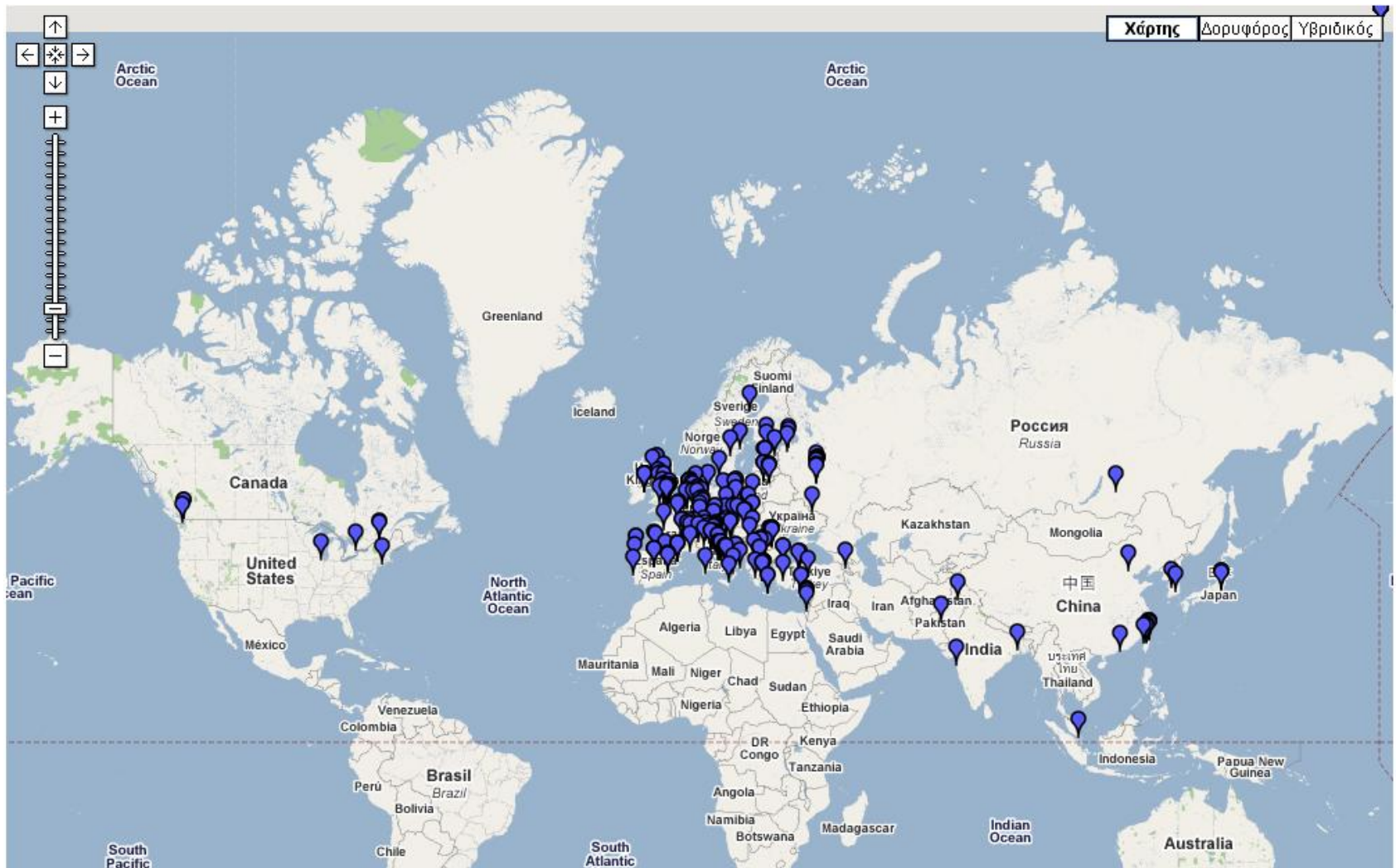
- ✓ 240+ institutions
- ✓ 45 countries

↳ Consists of:

- ✓ 202 sites
- ✓ ≈41.000 CPUs
- ✓ ≈5 PB
- ✓ 100,000 concurrent jobs

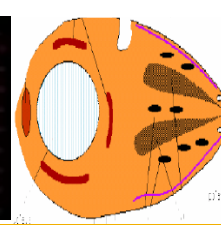
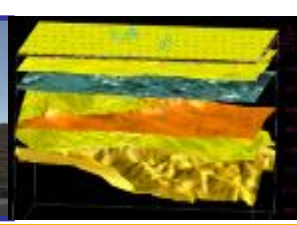
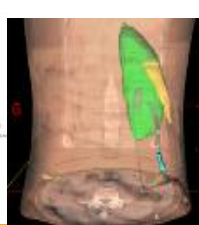
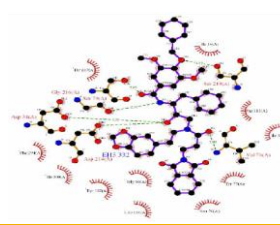
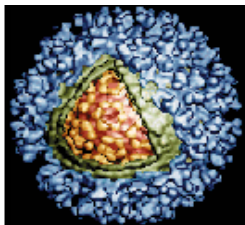
- <http://www.eu-egee.org/>

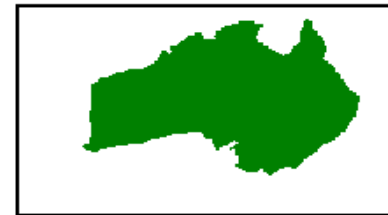
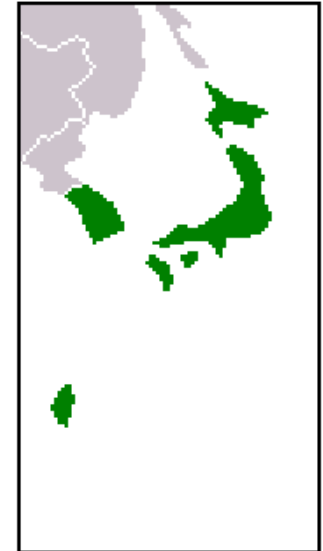
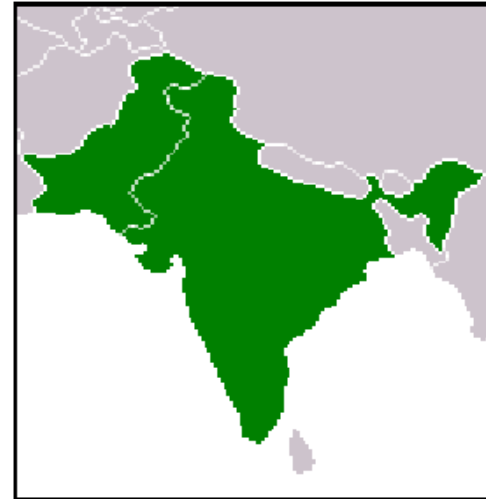
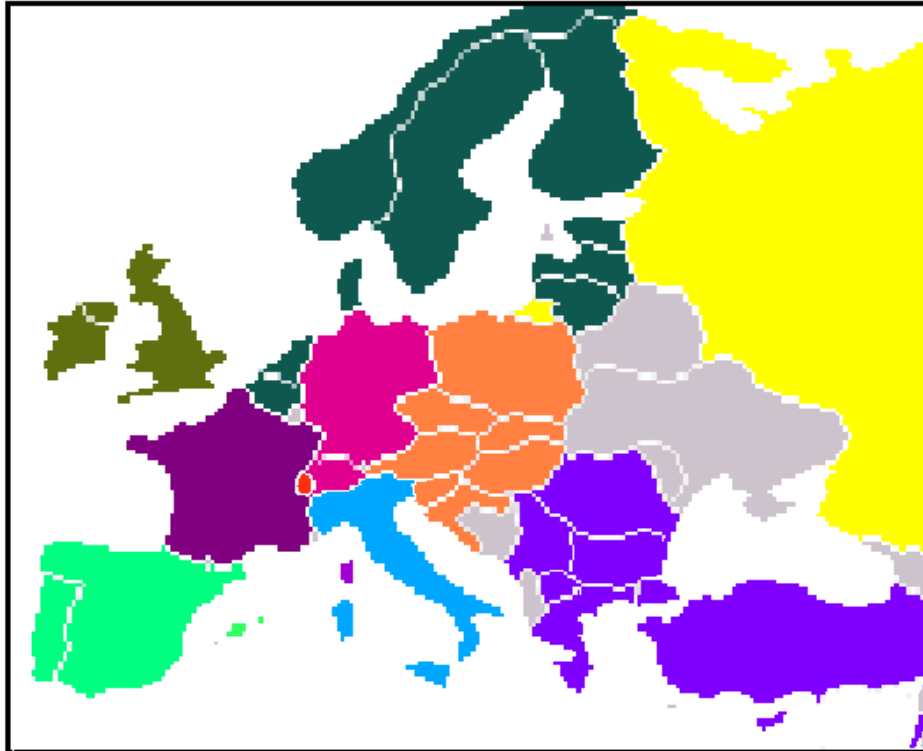




- Information about sites: <http://goc.grid.sinica.edu.tw/gstat/>

- **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)





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)

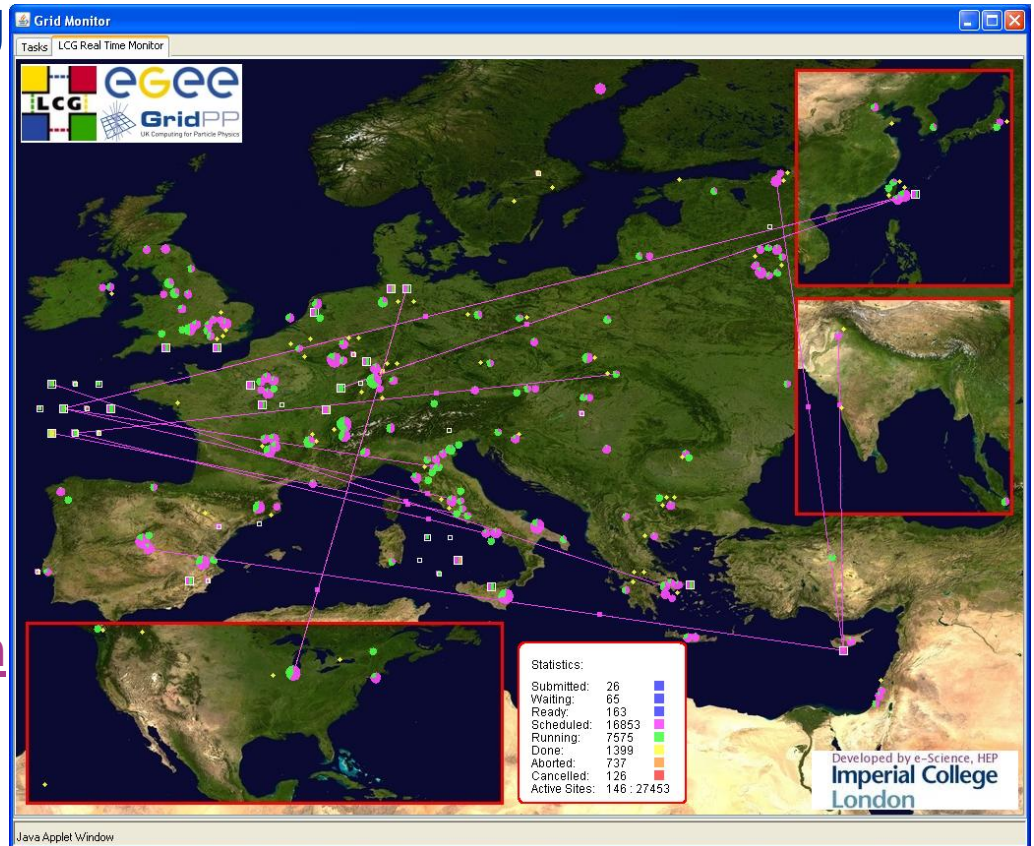
- In FZK: provide single point of contact (service desk), portal

• <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/>



- **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*





“GÉANT2 is the seventh generation of pan-European research and education network, successor to the pan-European multi-gigabit research network GÉANT”



European Commission
Information Society
and Media

<http://www.geant2.net/>

- The project officially started on 1 September 2004 and will continue to take place for the next 4 years
- The project is supported by the European Committee and by 30 European National Research and Education Networks (NRENs) in 34 countries and is administrated by DANTE (*Delivery of Advanced Network Technology to Europe*).
- It provides services of high quality and readability in the European Education and Research community and connects all the National Research Networks of European Union, Centre and East Europe, Israel and Cyprus



- It provides:
 - Basic IPs services
 - Quality of service levels
- Two main services:
 - Routed (Internet) and switched (L1-L2)
- Backbone mixed:
 - Part of will be based on dark fiber
 - Part of it on leased services
- Greece interconnection:
 - 2 * 10Gbps lambdas

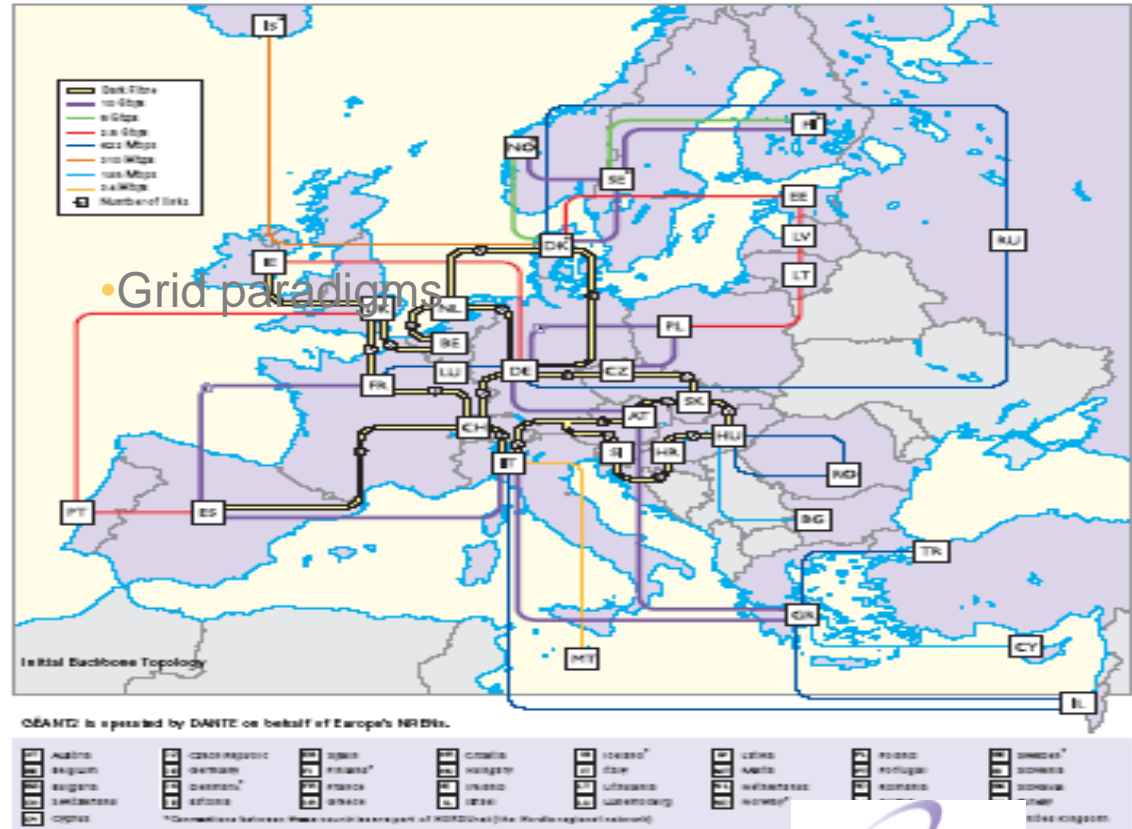


GÉANT2



The world-leading research and education network for Europe.

★ Connect ★ Communicate ★ Collaborate



• Grid paradigms



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

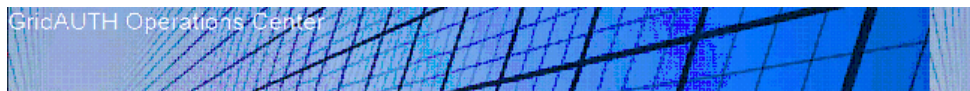


- **Operating system:**
 - Linux (+GNU utilities), usually a RHEL3-like, for example Scientific Linux, etc
- **Middleware:**
 - gLite v3.x
- **Libraries and Applications**
 - Defined by the system and VOs administrators' foresight
 - The user can install and execute its own programmes

Each VO according to its needs installs experimental software:

- **ATLAS:** atlas software (a big collection, v12.2.0 etc)
- **CMS:** cmkin, cobra, famos, geometry, ignominy, orca, oscar
- **ALICE:** alien, alice, root, proof
- **LHCb:** dirac, boole, DC, decfiles, gauss, paramfiles
- **BIOMED:** gate, cdss, gps@, gromacs, simri3d, gptm3d
- **ESR:** (earth science specific... eg, idl package)

- **The users can negotiate with their VOs for the installation of needed software**



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

- Seminars
- Certificate Authorities
 - SEE-Grid CA
 - Registration Authorities
 - Documents
 - Services
 - Downloads
 - HellasGrid Root CA
 - HellasGrid CA
 - Training CA
 - HellasGrid Demo CA
 - SEE-GRID Demo CA
 - HellasGrid CA 2003

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.

ROOT CA Certificate

Certificate Revocation Lists

Interesting Links

The SEE-GRID project

The EGEE project

EGEE-SEE

The European Grid Policy Management Authority (euGridPMA)

Find more info in the SeeGrid CA CP/CPS



Path: Home page > Documentation > User documentation

- Home page
- News
- Organization and Contacts
- Documentation
- Grid status
- Security
- Support
- Per-country Grid initiatives
- Events
- Training
- Newsletter
- Advanced Courses

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 u application.
2. New production users; those re
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/egge-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](#)

• **EGEE Helpdesk:**

<https://helpdesk.egge-see.org/index2.php>

• http://www.egge-see.org/User_documentation.php

The screenshot shows the SEE-GRID Wiki page. It includes a navigation menu with links to Main Page, Community portal, Current events, Recent changes, Random page, Help, and Donations. There is a search box with 'Go' and 'Search' buttons. A 'Contents' table of contents is visible, listing sections from 1 News to 8 Contacts. The page also features a 'toolbox' section with links for What links here, Related changes, Upload file, Special pages, Printable version, and Permanent link.

http://wiki.egge-see.org/index.php/SEE-GRID_Wiki

- 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 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.

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

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

What is the Grid?

Grid paradigms

Enabling Grid for E-science (EGEE)

HellasGrid Taskforce

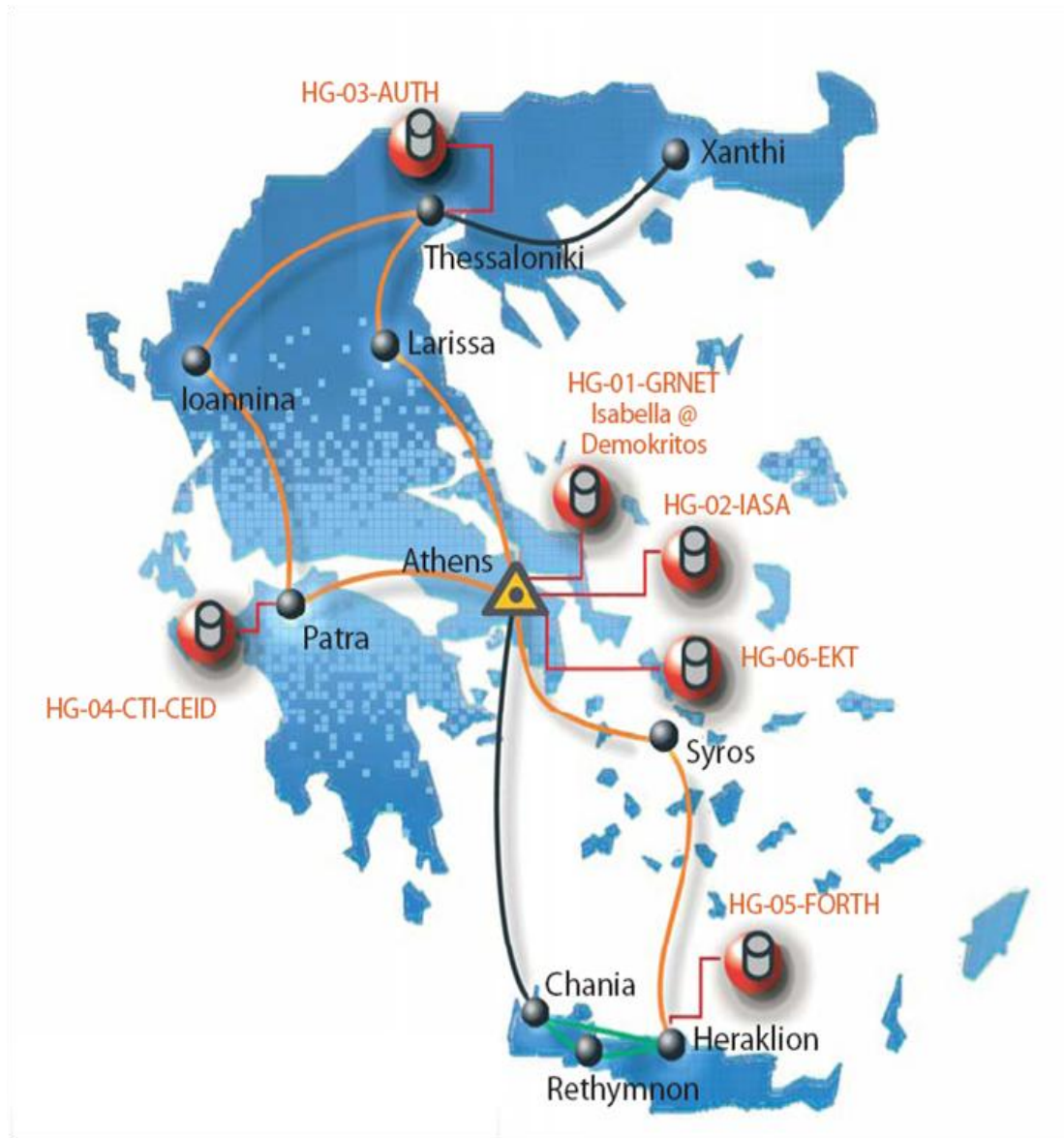
- **HellasGrid I**

- 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)
- 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



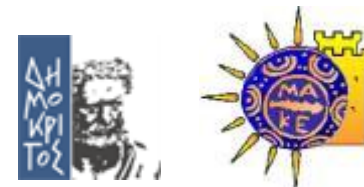
- Regional monitoring tools (FORTH):

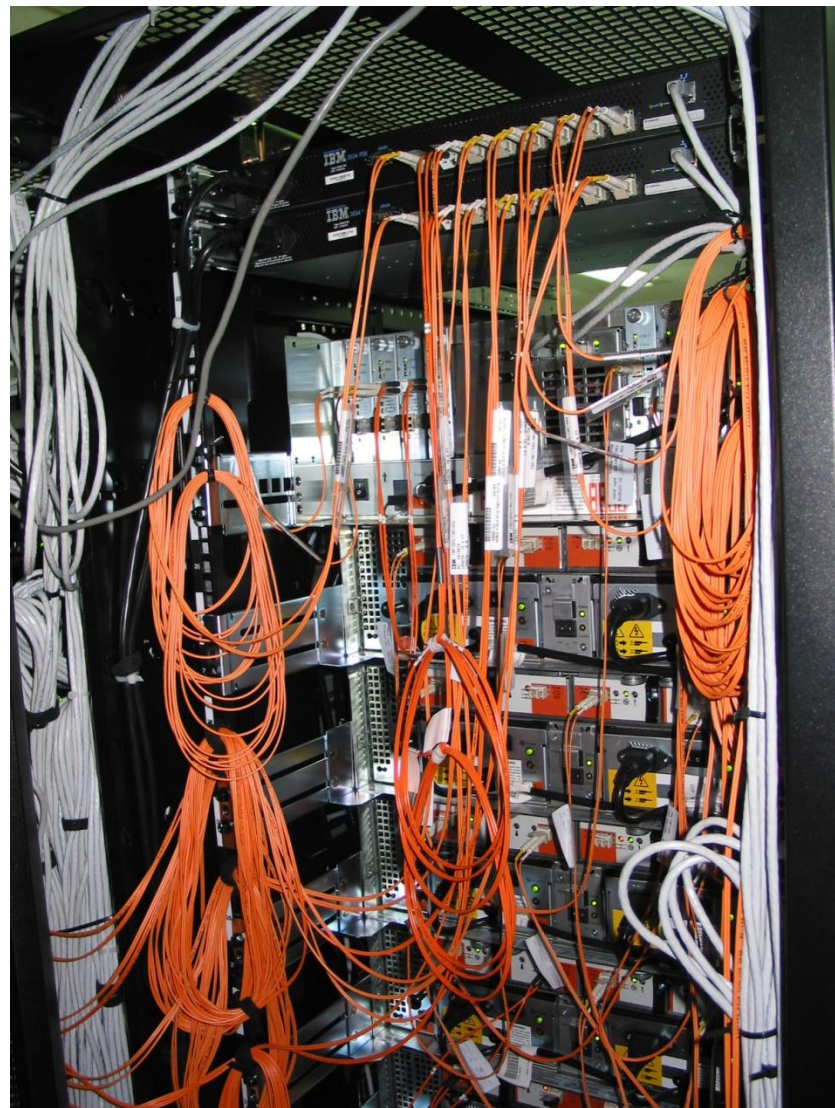
<http://hellasgrid-ui.ics.forth.gr/acctROC/>

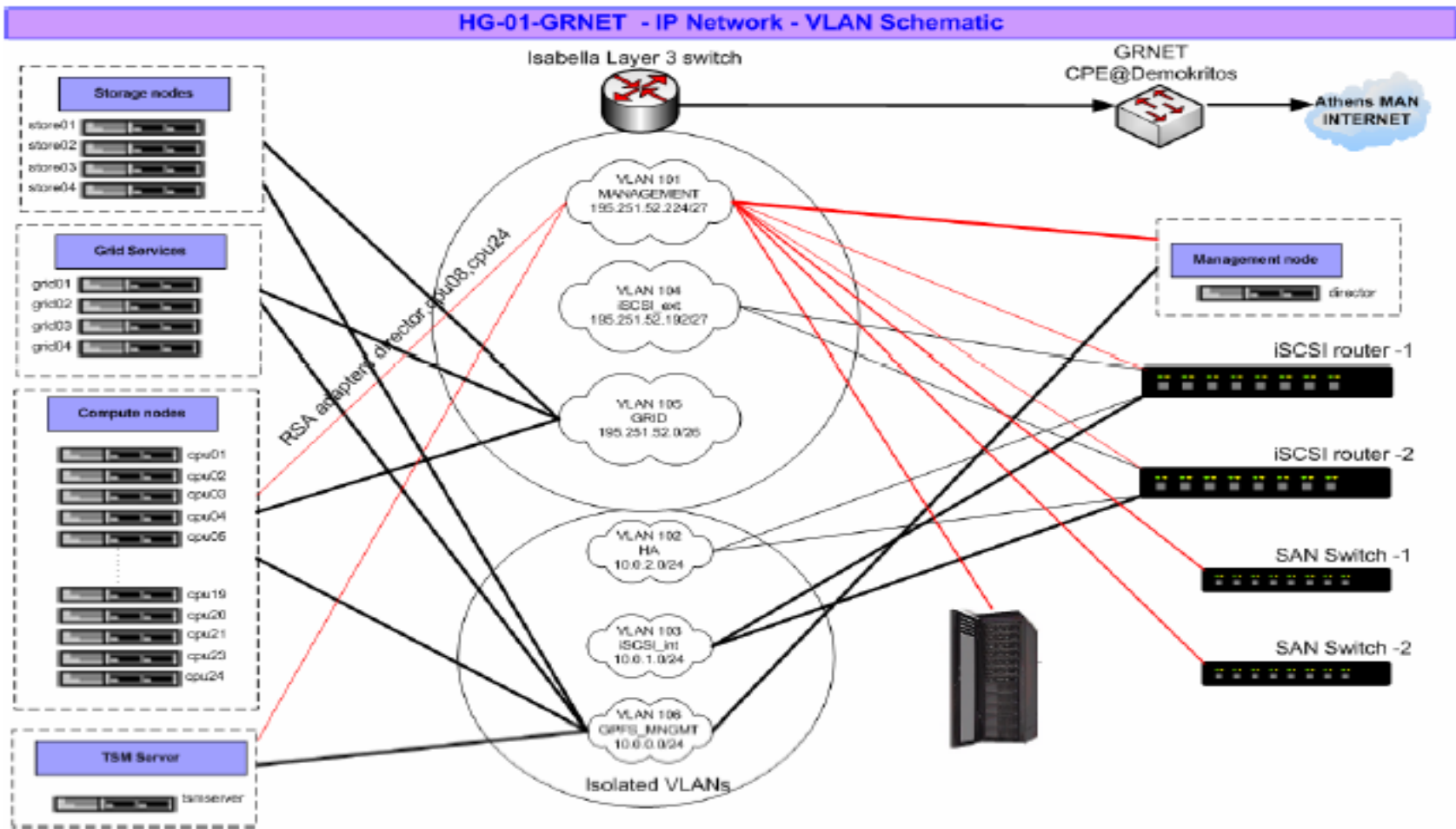


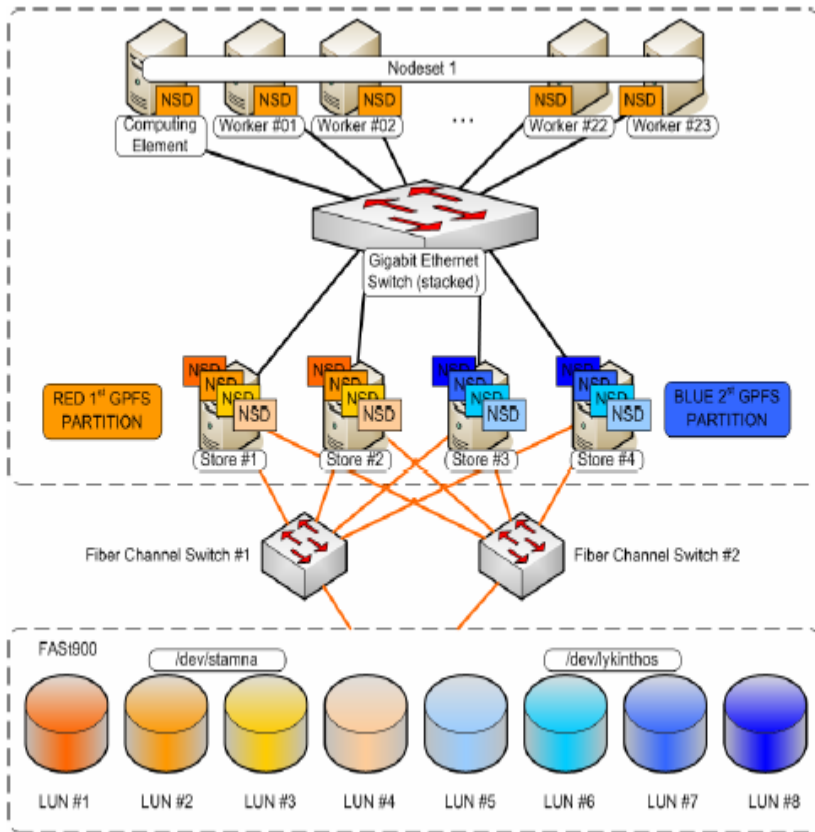
- Apps support (IASA):

application-support@hellasgrid.gr









- The first node of the Greek Grid Infrastructure, consisting a prototype for the next HellasGrid nodes
 - Remarkable and innovative organization of SAN and filesystems
- ⇒ Scientific Linux, gLite 3.0., LCG, MPICH, CODESA3D-1.0, VO-alice, VO-atlas, VO-biomed, VO-dteam, VO-cms, VO-esr, VO-lhcb, VO-see (octave), VO-seegrid ,...

- **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
- **Certification Services for new sites (SFTs)**
 - <https://mon.isabella.grnet.gr/sft/lastreport.cgi> (Need a valid HellasGrid Certificate)



https://access.hellasgrid.gr/register/registration_form

Not Logged In

HellasGrid User Registration





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)

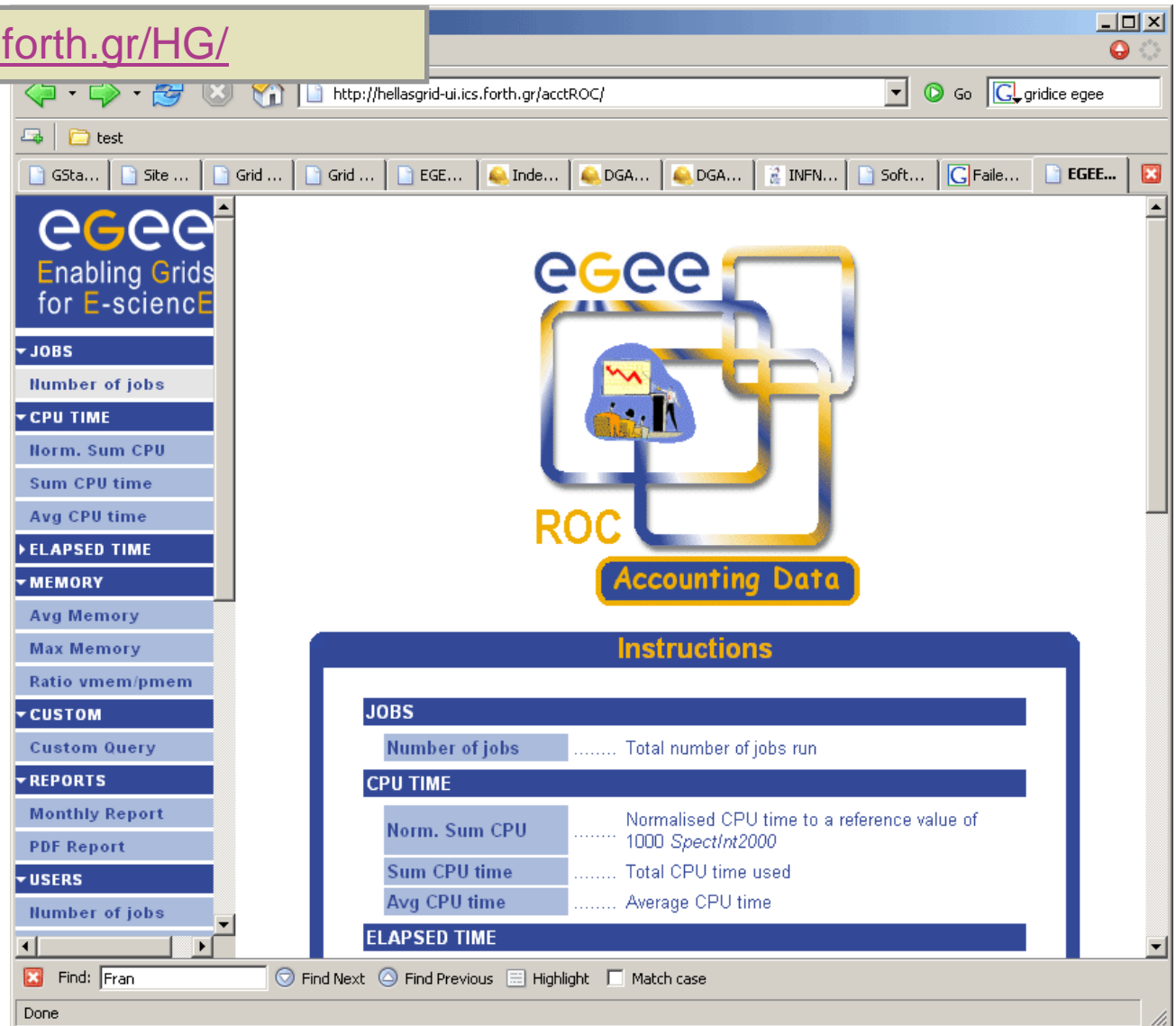








<http://hellasgrid-ui.ics.forth.gr/HG/>



The screenshot shows a web browser window displaying the HellasGrid Accounting Data interface. The browser's 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	
JOBS	
Number of jobs Total number of jobs run
CPU TIME	
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
ELAPSED TIME	



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
- **GOC Wiki**
<http://goc.grid.sinica.edu.tw/gocwiki/>
- **SEEREN2**
<http://www.seeren.org/>