

Εισαγωγή στο Grid,EGEE και το HellasGrid Introduction to Grid, EGEE and HellasGrid Athanasia Asiki

<u>aassiki@cslab.ece.ntua.gr</u>

Computing Systems Laboratory, National Technical University of Athens





www.eu-egee.org





What is the Grid?

Grid paradigms

Enabling Grid for E-sciencE (EGEE)

HellasGrid Taskforce



Grid Technologies (Grids)

Enabling Grids for E-sciencE

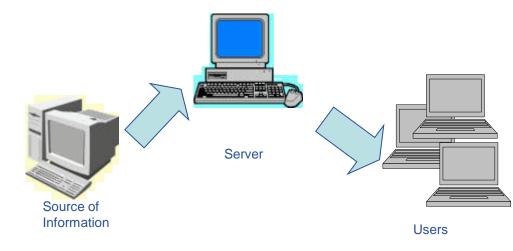




 The World Wide Web provides seamless access to information that is stored in many millions of different geographical locations

Enabling Grids for E-sciencE

eGee



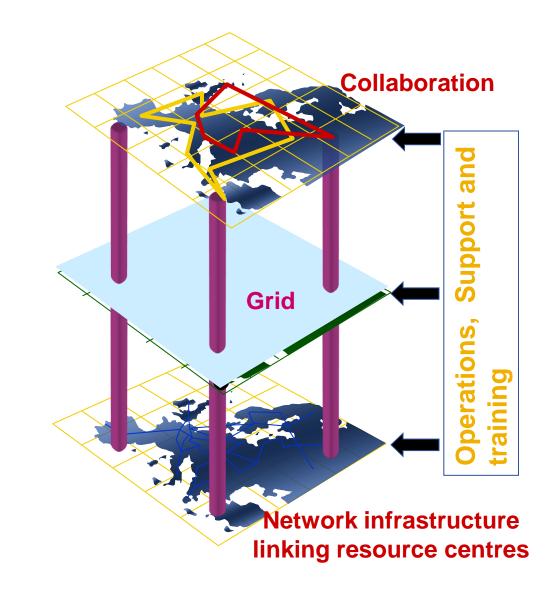
 The Grid is an emerging infrastructure that provides seamless access to computing power and data storage capacity distributed over the globe



eGee

The Grid

Enabling Grids for E-sciencE





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.

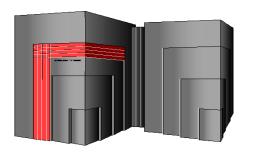


• An entity that is going to be shared

such as:

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











Resources



- Resource sharing
 - Geographically distributed resources offer computational power, storage capacity and bandwidth to the users

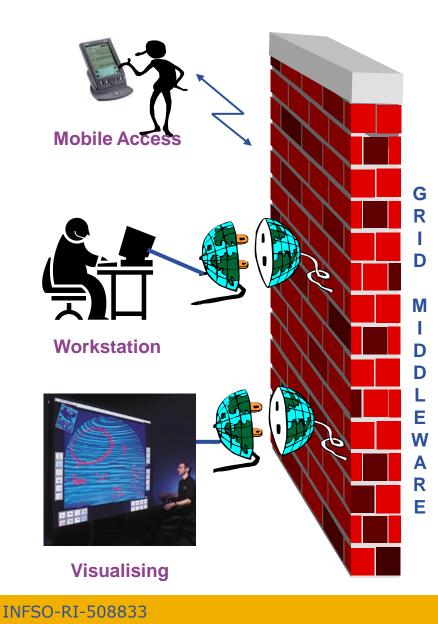
Secure and reliable access

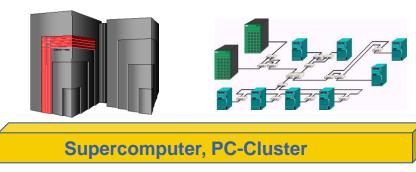
- Authentication
- Authorization
- Access policy
- Open standards
- Co-operation among people belonging to different organizations, institutes, groups

eGee

Grid metaphorically ...

Enabling Grids for E-sciencE







Data-storage, Sensors, Experiments

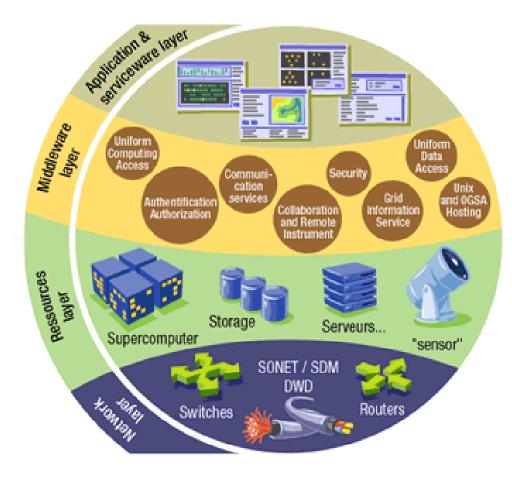


Internet, networks



Grid middleware

- 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)

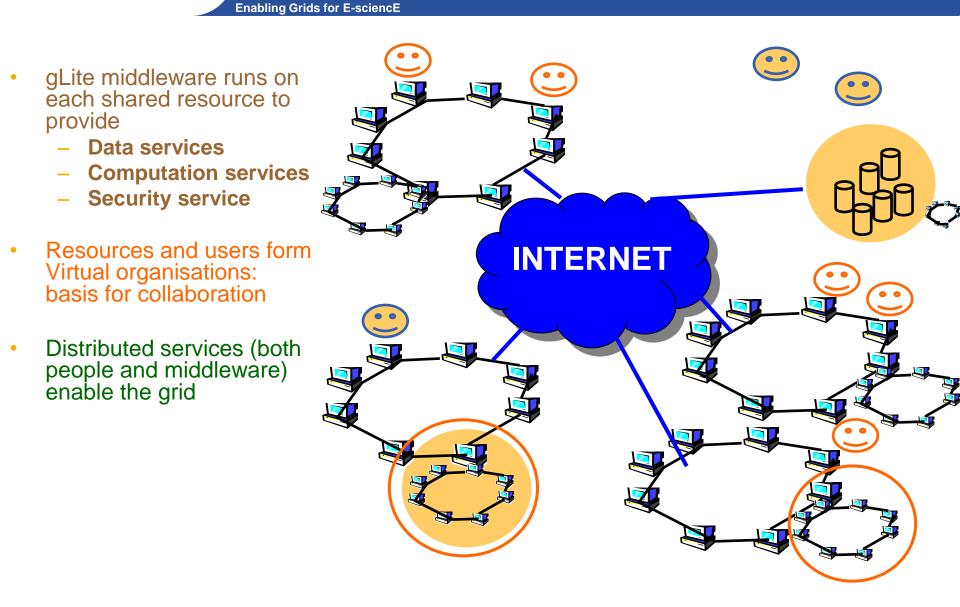
- Moore's law everywhere
 - ✓ Instruments, detectors, sensors, scanners, ...

 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





eGee



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" *lan 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



- Enabling Grids for E-sciencE
- 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
 - <u>http://cic.gridops.org/index.php?section=home&page=volist#1</u>





What is the Grid?

Grid paradigms

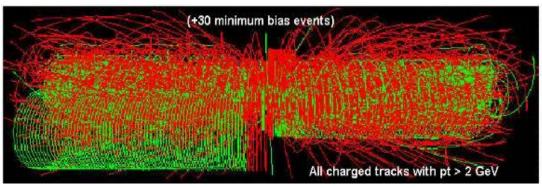
Enabling Grid for E-sciencE (EGEE)

HellasGrid Taskforce



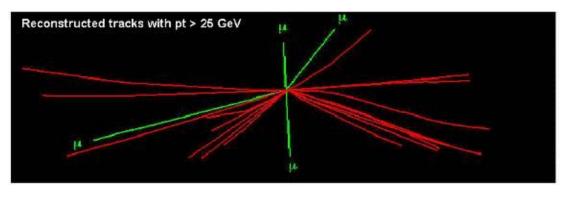
LHC Data Challenge

Starting from this event (particle collision) ...



 ✓ Data Collection
 ✓ Data Storage
 ✓ Data Processing

You are looking for this "signature"...



Selectivity: 1 in 1013

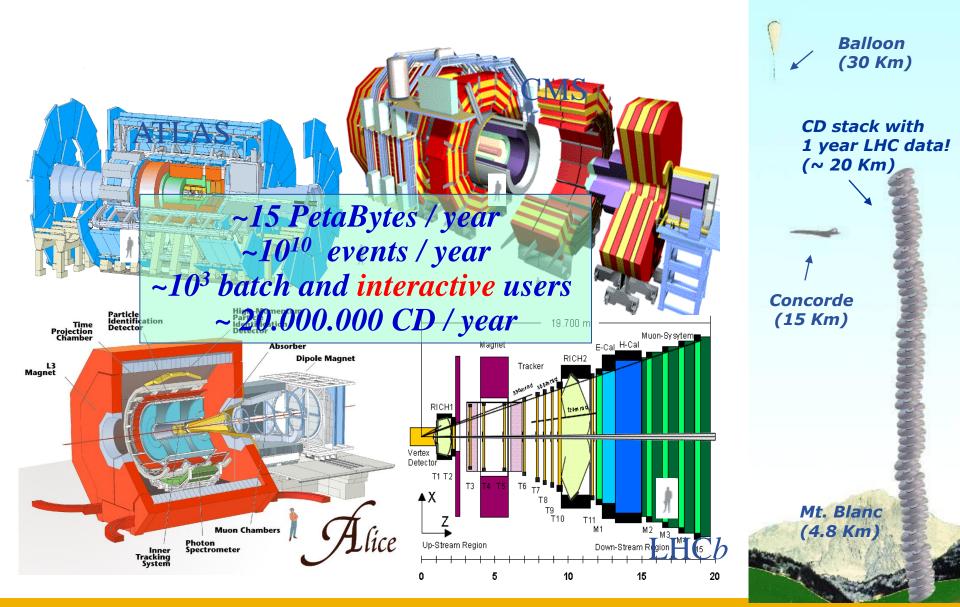
✓ Like looking for 1 person in a thousand world populations!

✓ Or for a needle in 20 million haystacks!



Amount of data from the LHC detectors

Enabling Grids for E-sciencE





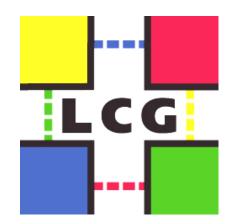
LHC Computing Grid

Enabling Grids for E-sciencE

- 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

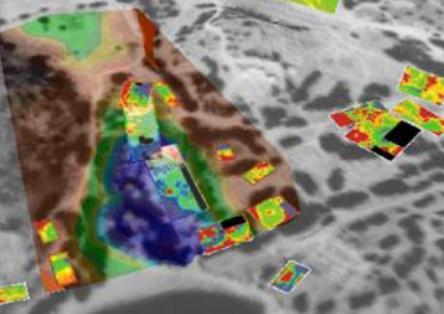
 Ibraries
 tools
 frameworks

required to support the physics application software



Some examples





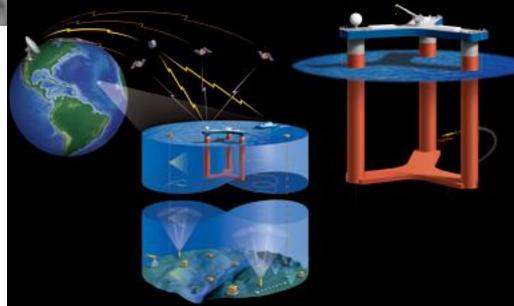
eGee

ArchaeoGrid

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

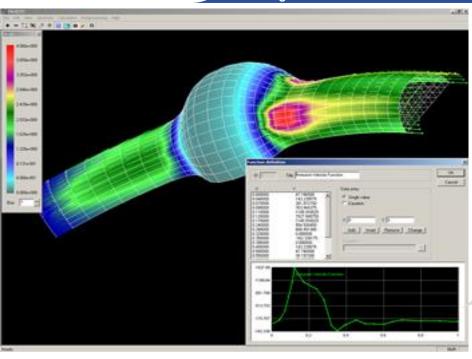
LOOKING

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.



Some examples

Enabling Grids for E-sciencE



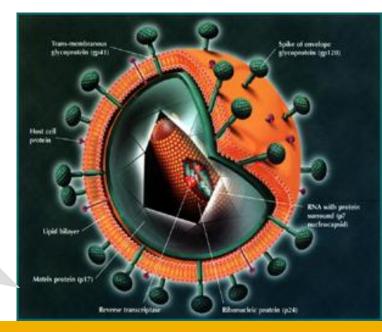
eGee

ViroLab

Aims to create a collaborative virtual labaratory 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 rulebased decision support system for drug ranking.

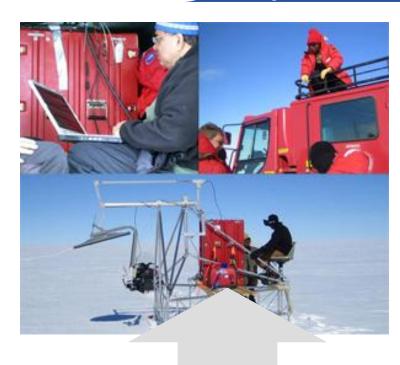
Parallel Blood Flow Simulation

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.



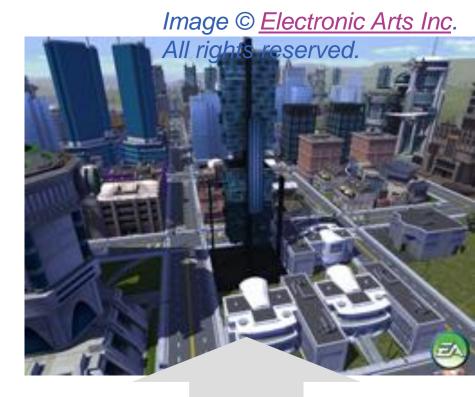


Enabling Grids for E-sciencE



eGee

INFSO-RI-508833



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.

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



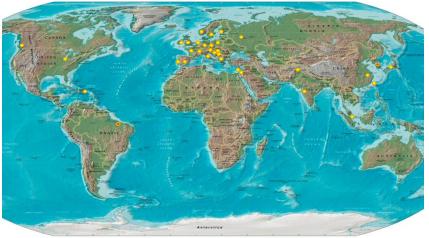
From EGEE to EGEE III

• 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



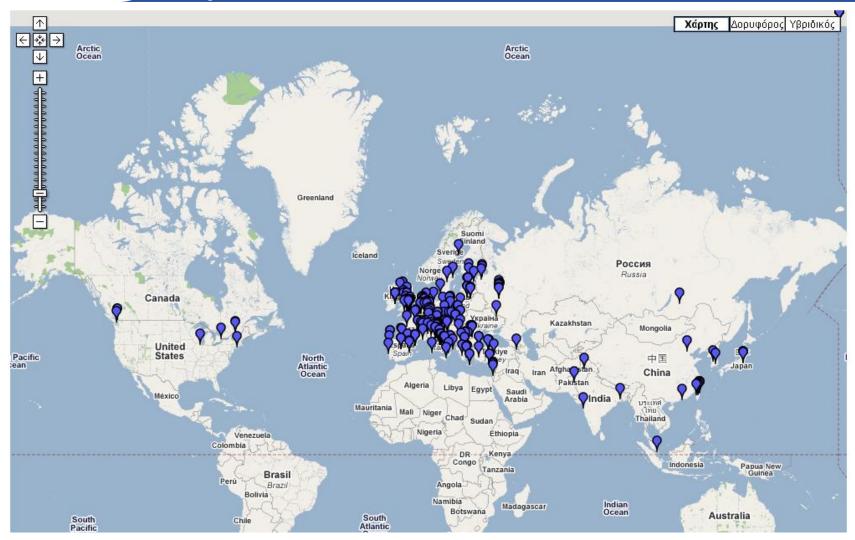
Enabling Grids for E-sciencE

- 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

- <u>http://www.eu-egee.org/</u>
- Participants:
 - ✓ 50 countries
- Consists of:
 - ✓ 250 sites
 - ✓ ~ 68000 CPUs
 - 🗸 ~ 20 PB
 - ✓ ~140 VOs
 - Massive data transfers> 1.5 GB/s

Infrastructure of EGEE

Enabling Grids for E-sciencE



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

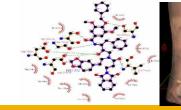
eeee



- 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)









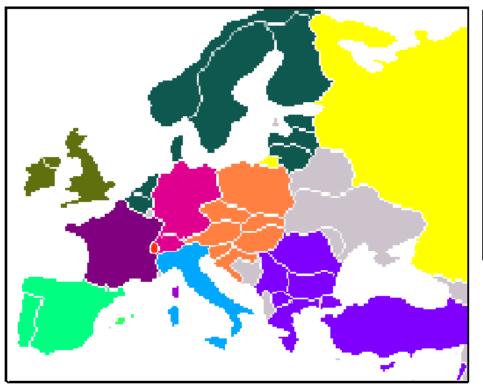






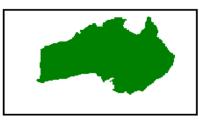
Operations centres in EGEE

Enabling Grids for E-sciencE









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

<u>https://gus.fzk.de/pages/home.php</u>



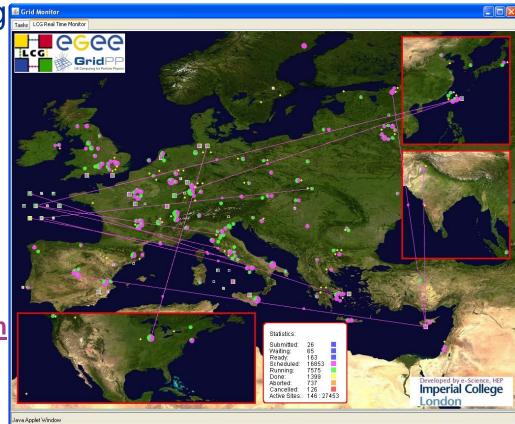
What is happening now?

Enabling Grids for E-sciencE

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



CGCC The European Network - GÉANT 2



"GÉANT2 is the seventh generation of pan-European research and education network, successor to the pan-European multigigabit research network GÉANT"

-800 - 1

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



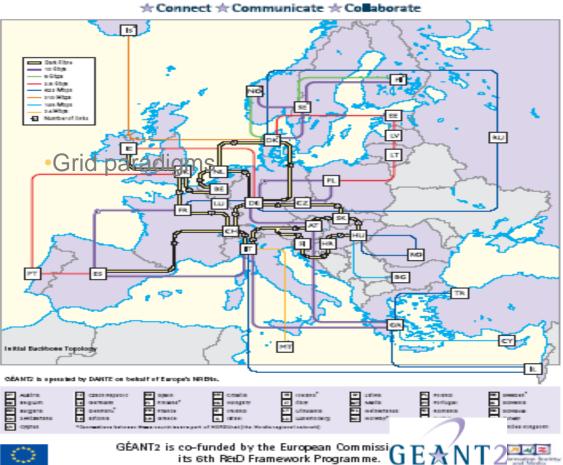
GÉANT 2

Enabling Grids for E-sciencE

- 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.





Grid systems' related software

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



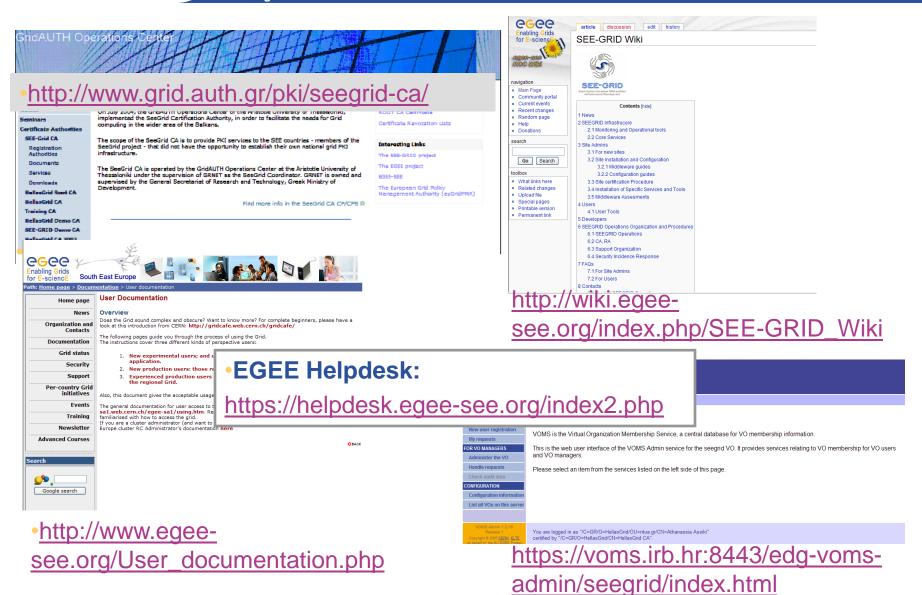
Each VO according to its needs installs experimental software:

- 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



Infrastructure Sites

Enabling Grids for E-sciencE





GridICE Monitoring for SEE

Enabling Grids for E-sciencE

http://mon.egee-see.org

the eyes of the Grid														🚡 for E
								Geo	view	Site	view	VO view	Help	
ridICE >> Site::ALL														
General Gris	Host		Job		Charts	Netw	ork							
					Over	view Co	mputing	Manageme	nt					
							Computing Reso					St	orage Resource	25
<u>ite</u> ▼		<u>Region</u>	<u>GK#</u>	<u>Q#</u>	<u>RunJob</u>	<u>WaitJob</u>	<u>JobLoad</u>	Power	<u>WN#</u>	<u>CPU#</u>	<u>CPULoad</u>	<u>Available</u>	<u>Total</u>	<u>%</u>
EGIS01-PHY-SCL		SEE	1	8	90	396	90%	95K	25	101	91%	Available	-	-
G-INRNE		SEE	-	-	- 50	-	16%	0	14	25	17%	_	-	-
G01-IPP		SEE	2	22	12	954	64%	0	10	11	61	916.2 GB	1 TB	11%
602-IM		SEE	1	6	2	22	100%	0	3	2	100%	15 GB	32.9 GB	54%
604-ACAD		SEE	1	12	54	241	98%	0	40	62	96%	24.8 GB	63.7 GB	615
605-SUGrid		SEE	1	8	10	208	69%	0	5	16	88%	35.2 GB	83.5 GB	58%
(-01-KIMON	S	SEE	1	10	37	11	5.x	0	37	74	<mark>4</mark> 4%	-	-	-
R-01-AUTH	重	SEE	1	12	10	341	92%	0	8	12	75%	165.2 GB	217.6 GB	24%
R-03-HEPNTUA		SEE	-	-	-	-	10%	0	15	30	7%	-	-	-
R-04-FORTH-ICS		SEE	1	10	2	429	100%	0	4	3	100%	-	-	-
R-05-DEMOKRITOS		<u>SEE</u>	-	-	-	-	-	-	-	-	-	50.1 GB	67.7 GB	26%
R-06-IASA		SEE	-	-	-	-	20%	0	10	20	0%	-	-	-
G-01-GRNET		<u>SEE</u>	-	-	-	-	100%	0	23	64	68%	-	-	-
G-02-IASA	1 () () () () () () () () () (<u>SEE</u>	-	-	-	-	7%	0	59	118	9%	-	-	-
G-03-AUTH		SEE	1	16	92	463	97%	0	58	118	90%	2.5 TB	2.7 TB	9%
G-04-CTI-CEID		SEE	1	15	113	71	97%	0	59	118	100%	2.3 TB	2.5 TB	6%
G-05-FORTH		SEE	1	15	92	78	92%	0	59	118	100%	-	-	-
G-06-EKT		SEE	1	17	225	280	100%	0	113	224	100%	-	-	-
-01-RBI		<u>SEE</u>	1	4	2	0	-	0	0	0	-	-	-	-
-BGU	0	<u>SEE</u>	1	2	0	8888	0%	0	5	5	5%	153.3 GB	154.9 GB	1%
CG-IL-OU	0	<u>SEE</u>	-	-	-	-	100%	0	5	8	100%	-	-	-
1K-01-UKIM_II		SEE	1	5	9	197	-	-	-	-	-	-	-	-





What is the Grid?

Grid paradigms

Enabling Grid for E-sciencE (EGEE)

HellasGrid Taskforce



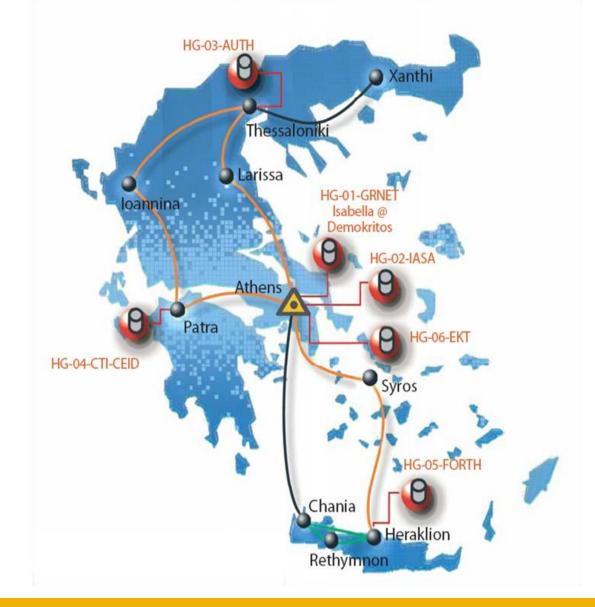
- HellasGrid I
 - Located at N.C.S.R. Demokritos (a.k.a. Isabella)
 - 34 dual Intel P4 Xeon @ 2.8GHz, 1GB RAM, 2x 70GB SCSI HDD, 2x Gbit
 - IBM FAStT900 Storage Area Network
 - 2x Redundant Fiber Channel Controllers with 1Gbyte Cache each
 - 70x146.8GB= 10,276TB raw storage capability, over 5 disk shelves
 - Tape Library ~30 TBytes, integrated monitoring
 - December 2004
- HellasGrid II
 - 5 sites: EKT (>220), IEΣE (48), AΠΘ (128), ITE (128), ITY (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/



HellasGrid e-Infrastructure

Enabling Grids for E-sciencE





HellasGrid structure

- Main site: HG-01-GRNET (Isabella, cslab@ICCS/NTUA)
- HG-02...HG-06 sites @ (NDC, IASA, AUTH, FORTH, CTI)
- Smaller sites (AUTH, UoM, FORTH, Demokritos, HEP-NTUA, IASA)
- HG CA and VOMS (AUTH): <u>http://www.grid.auth.gr/pki/seegrid-ca/</u>
- Helpdesk (CTI):

user-support@hellasgrid.gr

- Regional monitoring tools (FORTH): <u>http://hellasgrid-ui.ics.forth.gr/acctROC/</u>
- Apps support (IASA):
 <u>application-support@hellasgrid.gr</u>









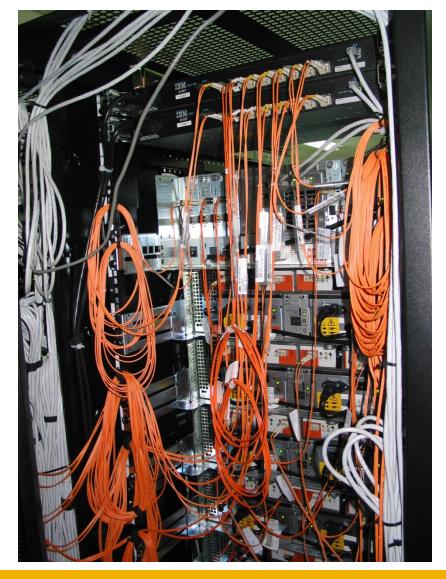




GGC HellasGrid I Infrastructure, Isabella

Enabling Grids for E-sciencE

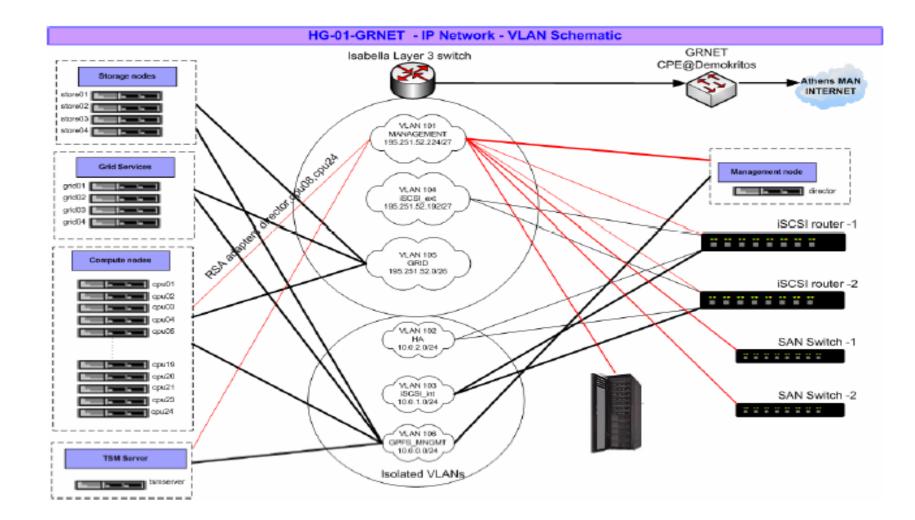




egee

HellasGrid I, Isabella: Network

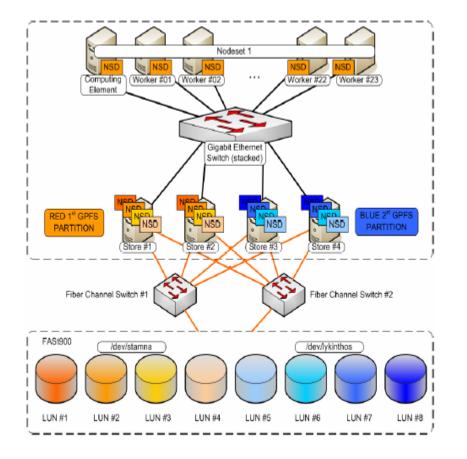
Enabling Grids for E-sciencE





HellasGrid I, Isabella: Storage

Enabling Grids for E-sciencE



- 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), VOseegrid



Core Services (HG-01-GRNET)

- Enabling Grids for E-sciencE
- 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)
 - <u>https://mon.isabella.grnet.gr/sft/lastreport.cgi</u> (Need a valid HellasGrid Certificate)



HellasGrid II Infrastructure

Enabling Grids for E-sciencE













Registration to HellasGrid

Enabling Grids for E-sciencE

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

				Not Logged In
HellasGrid U	ser Registi	ration		Final Bota
			the second	llasGrid id infrastructure
Εγγραφή νέων χρηστών > Φόρμα Εγ	γραφής Χρήστη			
<u>Διαδικασία</u> Καταχώρηση προσωπικών	Εγγραφή νέου Όνομα	χρήστη Ελληνικά Αγγλικά		Υ πάρχοντες Χρήστες Αν στο παρελθόν είχατε
στοιχείων Αίτηση ψηφιακού πιστοποιητικού Αποστολή αιτήσεως Επικοινωνία GridAUTH Support	Επώνυμο 	Ελληνικά Αγγλικά		αποκτήσει ψηφιακό πιστοποιητικό από την Α.Π. HellasGrid CA το οποίο έχει πλέον λήξει , συμπληρώστε στο πεδίο που ακολουθεί το e-mail σας για
	Οργάνισμός Τηλέφωνο	Ανωτάτη Σχολή Καλών Τεχνών	×	να προχωρήσετε στη διαδικασία αίτησης καινούργιου φηφιακού πιστοποιητικού.
	εργασίας Επιστημονικός τομέας Τμήμα	Άλλο		Αναζήτηση E-mail
	Ιδιότητα	Ερευνητής		
	KaraXaprior	GridAUTH (HellasGrid User Registration)		
	🚫 εγρωπαϊκή ε		HELLAS GR	ID
		EAET / GRNET		



Accounting statistics for the HellasGrid Infrastructure

Enabling Grids for E-sciencE







Enabling Grids for E-sciencE







- Grid café: <u>http://gridcafe.web.cern.ch/gridcafe</u>
- Open Grid Forum:
 http://www.gridforum.org/
- HellasGrid Task Force
 <u>http://www.hellasgrid.gr/</u>
- EGEE (Enabling Grids for E-science)
 <u>http://public.eu-egee.org/intro/</u>
- The Globus Alliance
 http://www.globus.org/
- Worldwide LHC Computing Grid
 - http://goc.grid.sinica.edu.tw/seegridwiki/
- Grid Operations Centre
 - http://goc.grid-support.ac.uk/gridsite/gocmain/
- gLite UserGuide

https://edms.cern.ch/file/722398//gLite-3-UserGuide.pdf



Χρήσιμα web links

• EGEE

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

- EGEE South East Europe http://www.egee-see.org/
- SEE-GRID
 http://www.see-grid.org/
- Hellas Grid Task Force
 http://www.hellasgrid.gr/
- GRNET

http://www.grnet.gr/

• gLite

http://glite.web.cern.ch/glite/

SEE-GRID Wiki

http://goc.grid.sinica.edu.tw/seegridwiki/

GOC Wiki

http://goc.grid.sinica.edu.tw/gocwiki/

• SEEREN2

http://www.seeren.org/