An overview of the EGEE project

MAGGI, Giorgio Pietro
INFN Bari
giorgio.maggi@ba.infn.it

BioinfoGRID
Bioinformatics Grid Application for life science

Outline

In this presentation an update on the status of the Grid in Europe (and in Italy), from the point of view of EGEE, will be given.

It will focus on the following subjects:

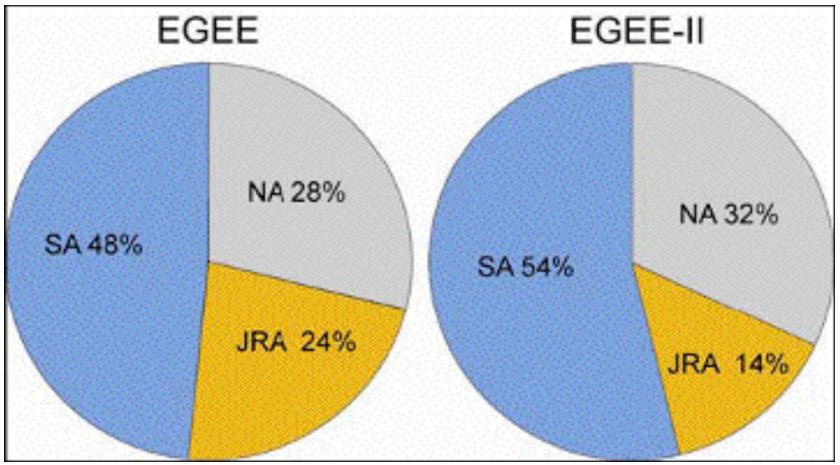
- The EGEE II project
- The related projects
- The EGEE infrastructure
  - The INFN production Grid
- The EGEE middleware (gLite 3.0)
- The GILDA testbed (the t-infrastructure)
- The applications
- The future
- Conclusions

The EGEE project

- EGEE
  - 1 April 2004 – 31 March 2006
  - 71 partners in 27 countries, federated in regional Grids

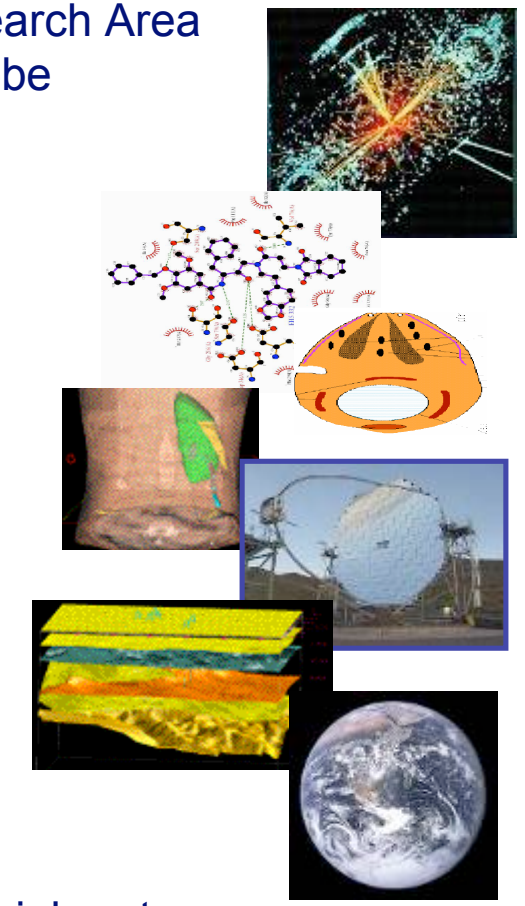
- EGEE-II
  - 1 April 2006 – 31 March 2008
  - Expanded consortium
    - 91 partners
    - 11 Joint Research Units (48 partners)
    - 13 Federation

  - Exploitation of EGEE results
  - Emphasis on providing production-level infrastructure
    - → increased support for applications
    - → interoperation with other Grid infrastructures
    - → more involvement from Industry



The EGEE II Mission

- Infrastructure
  - Manage and operate production Grid for European Research Area
  - Interoperate with e-Infrastructure projects around the globe
  - Contribute to Grid standardisation efforts
- Support applications from diverse communities
  - Astrophysics
  - Biomedicine
  - Computational Chemistry
  - Earth Sciences
  - Finance
  - Fusion
  - Geophysics (incl. industrial application EGEODE)
  - High Energy Physics
  - Multimedia
  - …
- Industry
  - Reinforce links with the full spectrum of interested industrial partners



+ Disseminate knowledge about the Grid through training
+ Prepare for sustainable European Grid Infrastructure

Outline

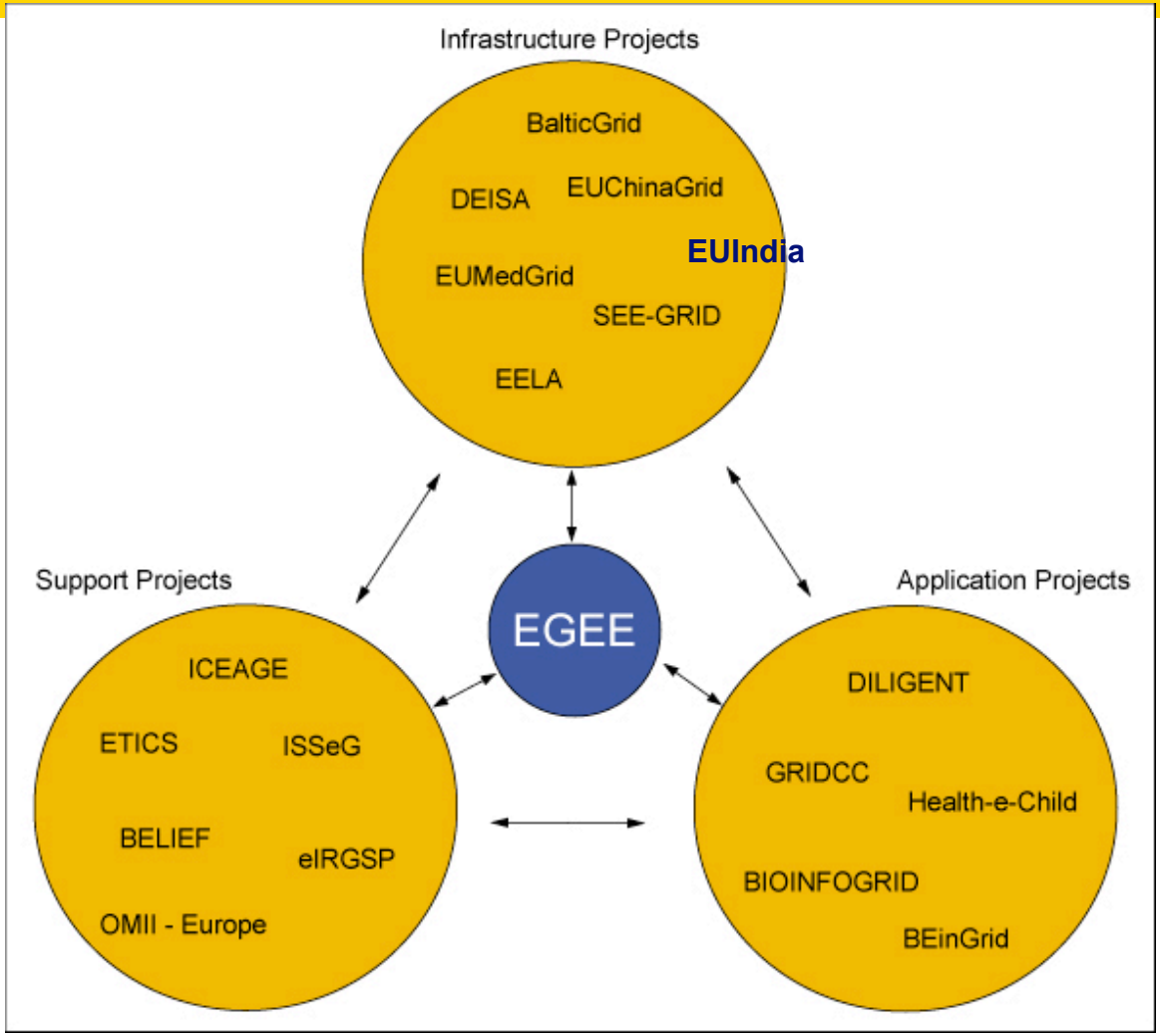
- The EGEE II project
- **The related projects**
- The EGEE infrastructure
  - The INFN production Grid
- The middleware (gLite 3.0)
- The GILDA testbed (the t-infrastructure)
- The applications
- The future
- Conclusions

Grids in Europe

- EGEE is not alone.
- Great investment all over Europe in developing Grid technology:
- A Sample of National Grid projects:
  - Austrian Grid Initiative
  - Belgium: BEGrid
  - DutchGrid
  - France: e-Toile; ACI Grid
  - Germany: D-Grid; Unicore
  - Greece: HellasGrid
  - Grid Ireland
  - Italy: INFNGrid; GRID.IT
  - NorduGrid
  - UK e-Science: National Grid Service; OMII; GridPP

- EGEE provides framework for national, regional and thematic Grids

Related EU projects



EGEE-II and Industry

- Industry Task Force
  - Group of industry partners in the project
  - Links related industry projects (NESSI, BEinGRID, …)
  - Works with EGEE's Technical Coordination Group

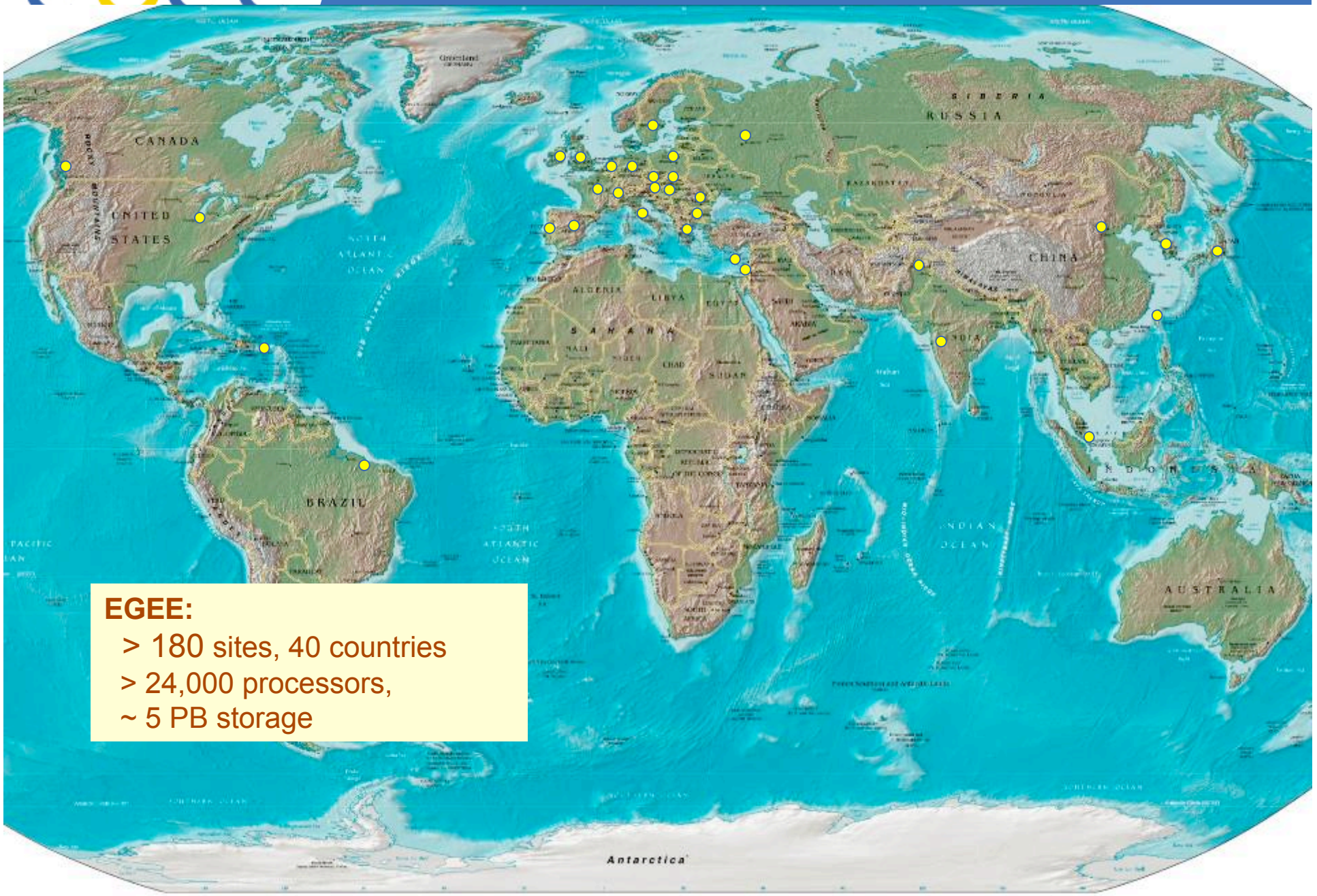
- Collaboration with CERN openlab project
  - IT industry partnerships for hardware and software development

- EGEE Industry Forum
  - Led by Industry to improve Grid take-up in Industry
  - Organises industry events and disseminates grid information
    - Last event: EGEE Industry Day 12 July 2006, Ischia ,Italy
      http://www.eu-egee.org/egee_events/industryday/IndustryDay_Ischia/
    - Next event: during fall in Catania
    - Further events across Europe are being planned

Outline

- The EGEE II project
- The related projects
- **The EGEE infrastructure**
  - The INFN production Grid
- The middleware (gLite 3.0)
- The GILDA testbed (the t-infrastructure)
- The applications
- The future
- Conclusions

EGEE Grid Sites : Q1 2006



EGEE:

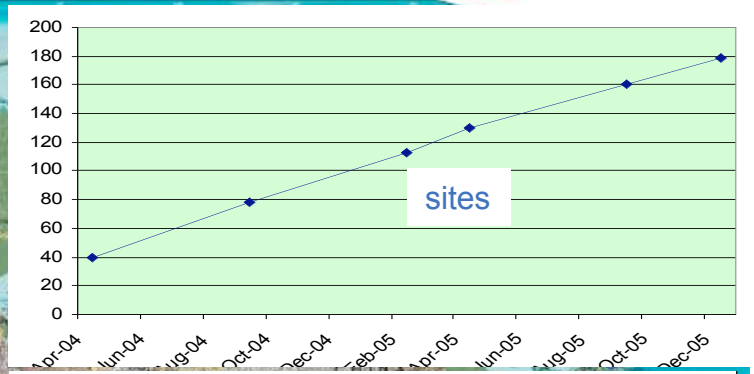
> 180 sites, 40 countries

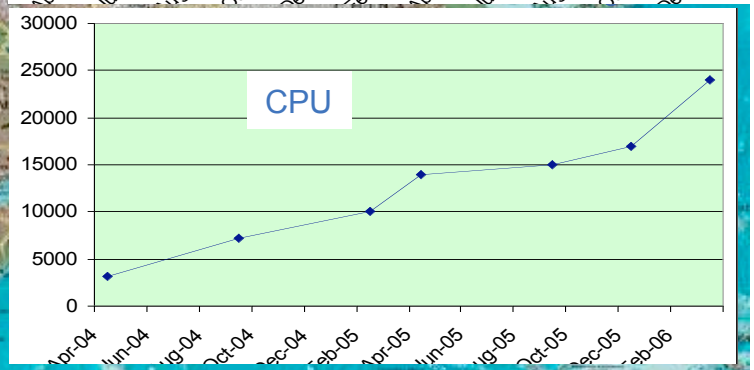
> 24,000 processors,

~ 5 PB storage

EGEE Grid Sites : Q1 2006

EGEE:
Steady growth over the lifetime of the project

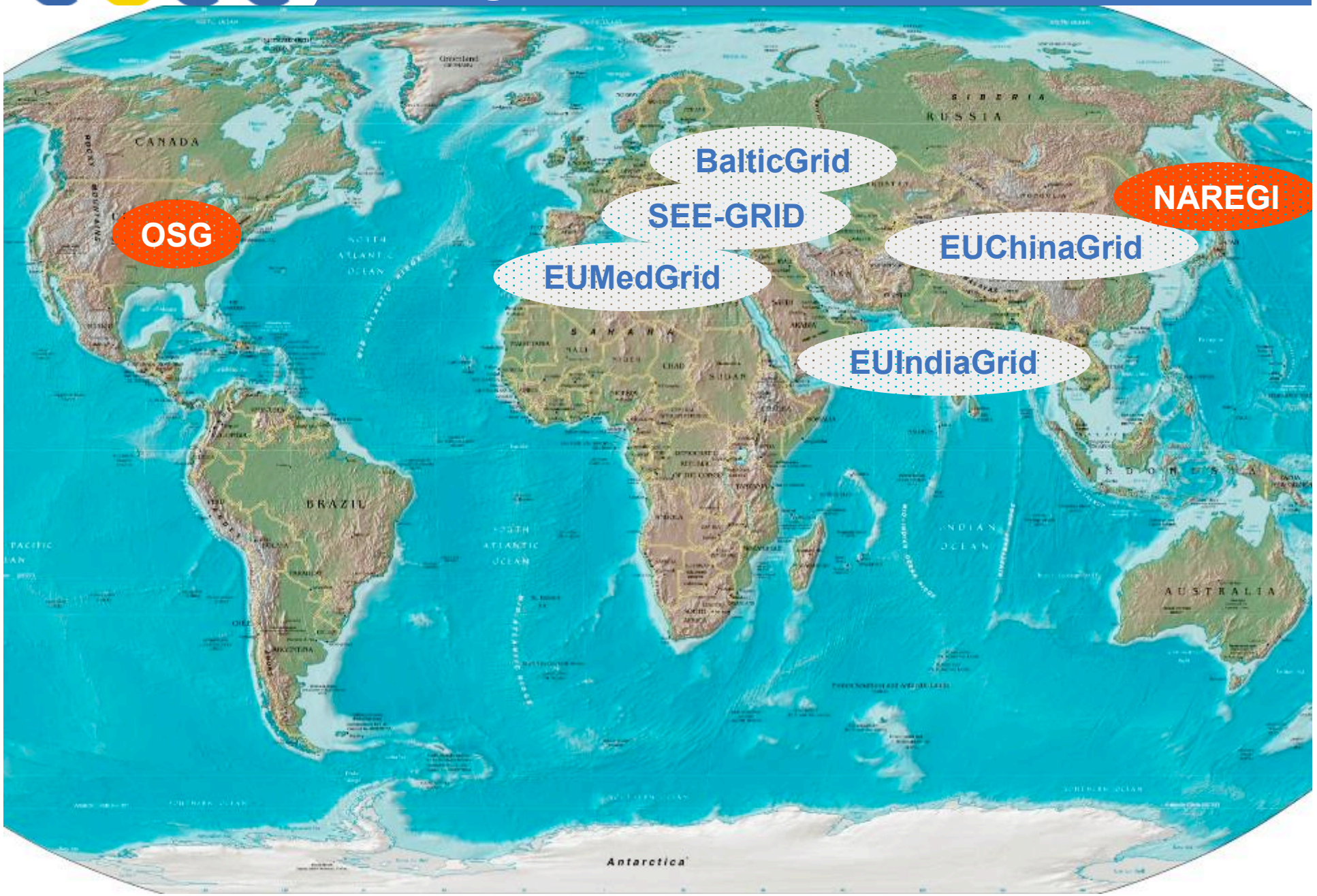




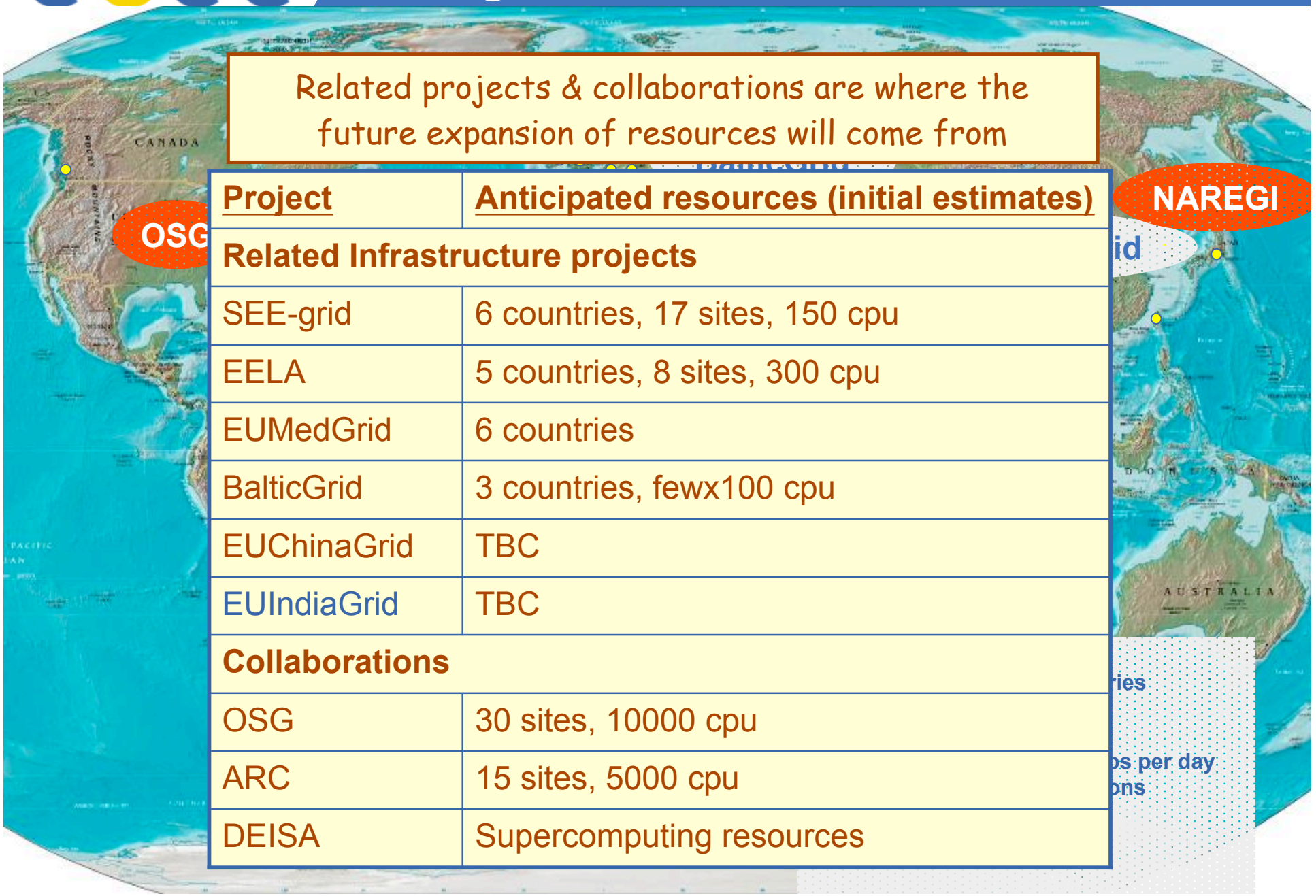
EGEE:
> 180 sites, 40 countries
> 24,000 processors,
~ 5 PB storage

country	sites	country	sites	country	sites
Austria	2	India	2	Russia	12
Belgium	3	Ireland	15	Serbia	1
Bulgaria	4	Israel	3	Singapore	1
Canada	7	Italy	25	Slovakia	4
China	3	Japan	1	Slovenia	1
Croatia	1	Korea	1	Spain	13
Cyprus	1	Netherlands	3	Sweden	4
Czech Republic	2	FYROM	1	Switzerland	1
Denmark	1	Pakistan	2	Taipei	4
France	8	Poland	5	Turkey	1
Germany	10	Portugal	1	UK	22
Greece	6	Puerto Rico	1	USA	4
Hungary	1	Romania	1	CERN	1

A global, federated e-Infrastructure



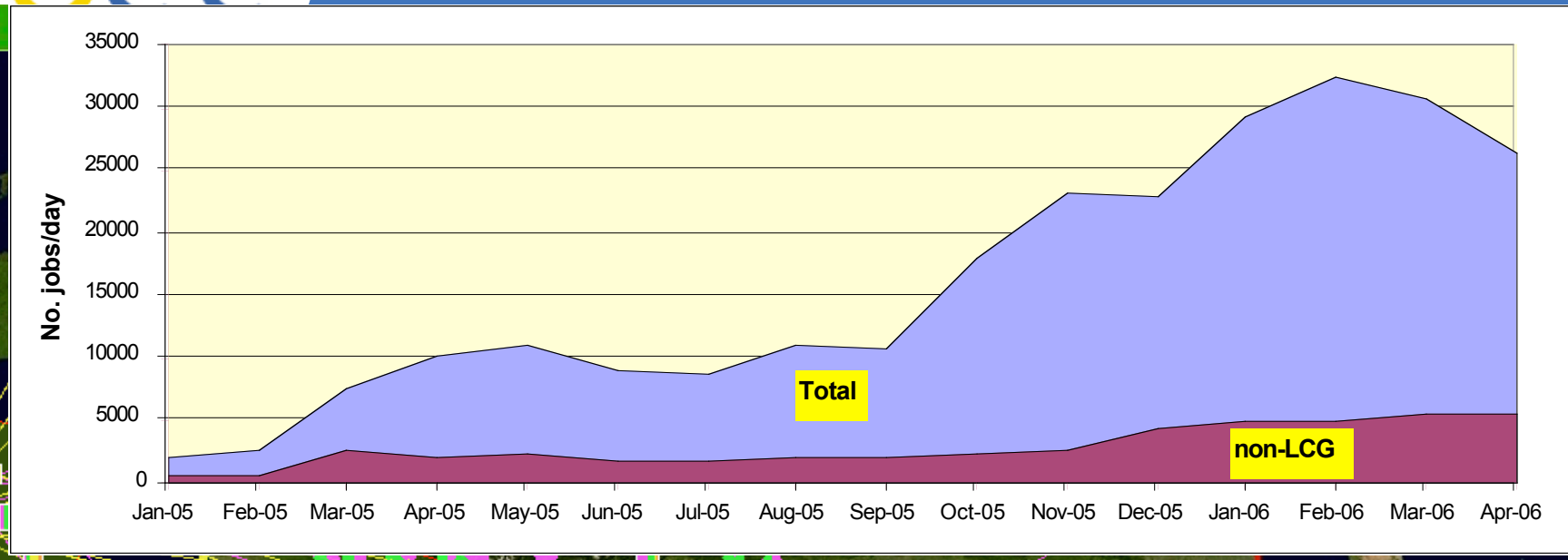
A global, federated e-Infrastructure



Related projects & collaborations are where the future expansion of resources will come from

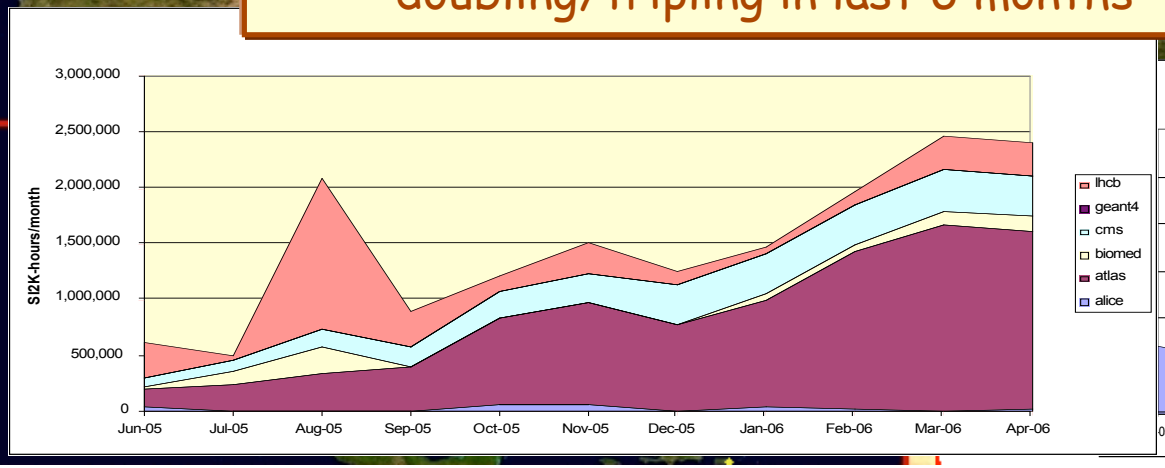
Project	Anticipated resources (initial estimates)
Related Infrastructure projects	
SEE-grid	6 countries, 17 sites, 150 cpu
EELA	5 countries, 8 sites, 300 cpu
EUMedGrid	6 countries
BalticGrid	3 countries, fewx100 cpu
EUChinaGrid	TBC
EUIndiaGrid	TBC
Collaborations	
OSG	30 sites, 10000 cpu
ARC	15 sites, 5000 cpu
DEISA	Supercomputing resources

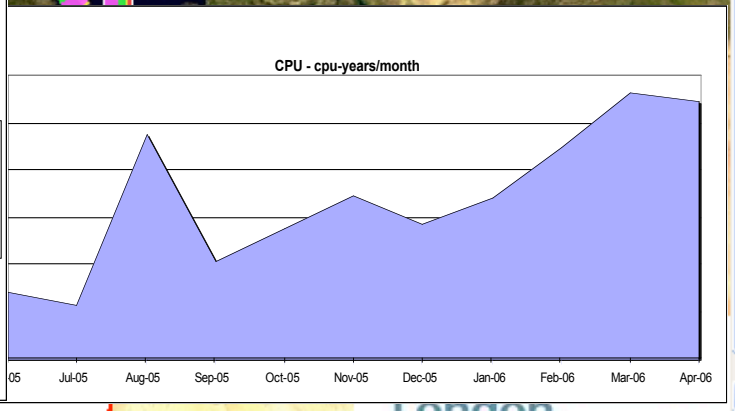
Use of the infrastructure



Sustained & regular workloads of >30K jobs/day

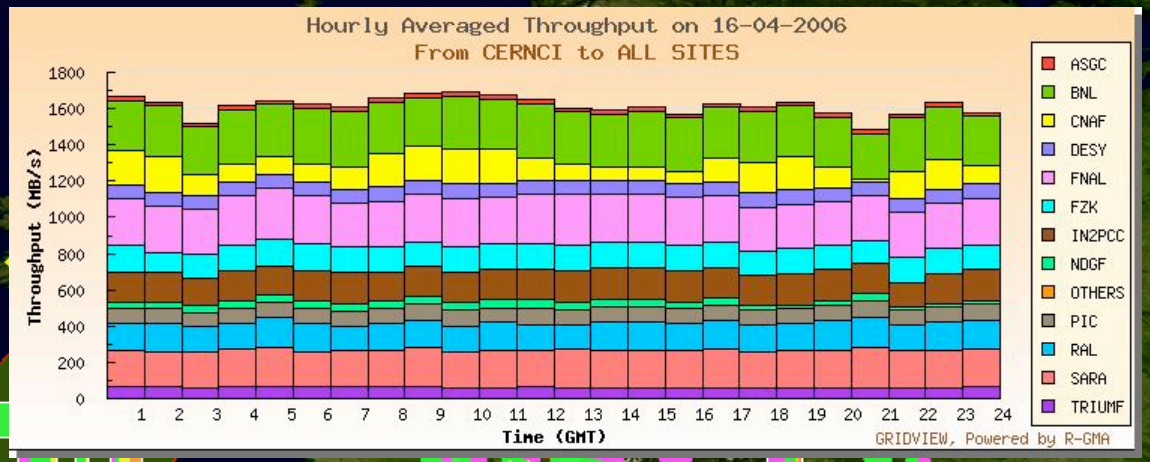
- spread across full infrastructure
- doubling/tripling in last 6 months – no effect on operations

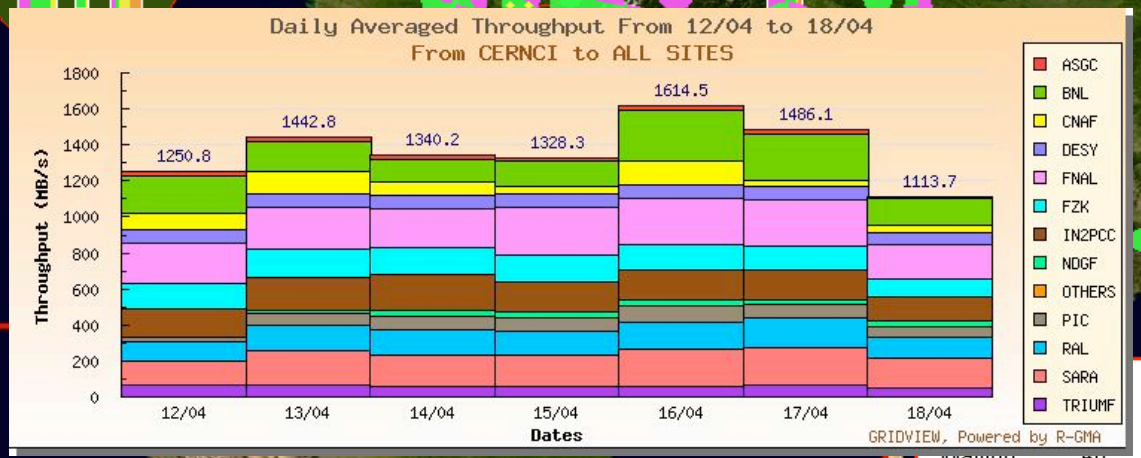


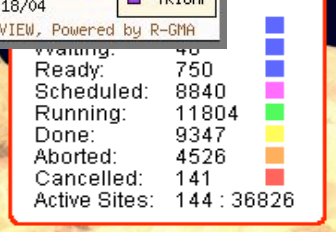


Use of the infrastructure

Massive data transfers > 1.5 GB/s



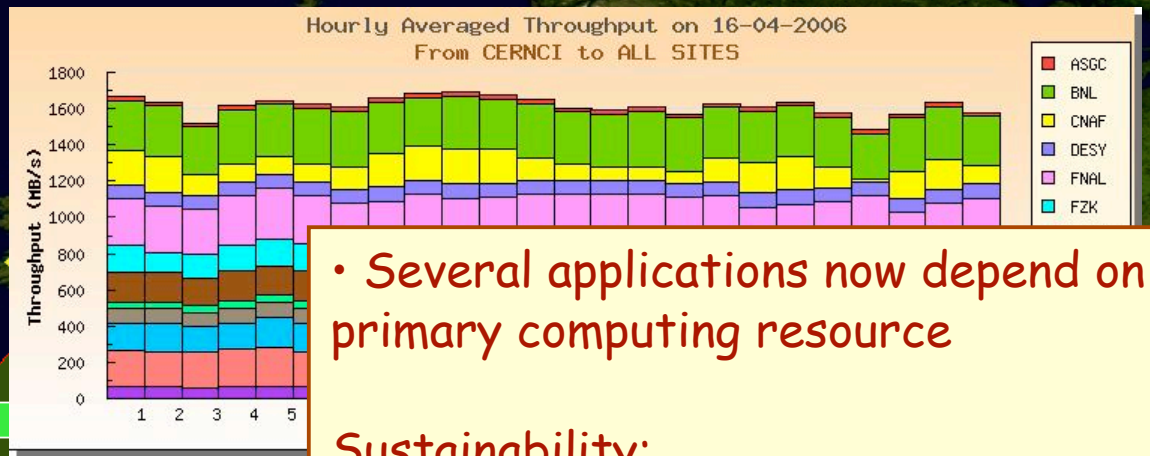


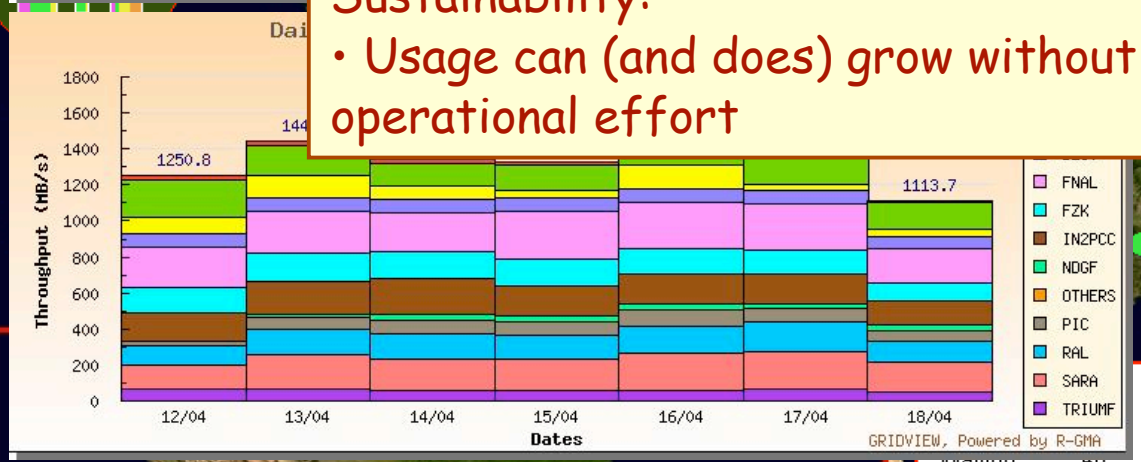


Developed by e-Science, HEP
Imperial College London

Use of the infrastructure

Massive data transfers > 1.5 GB/s

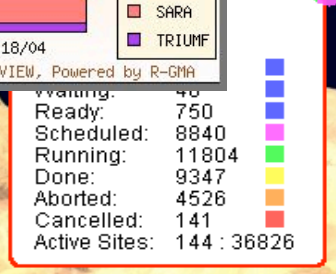




- Several applications now depend on EGEE as their primary computing resource

Sustainability:

- Usage can (and does) grow without need for additional operational effort

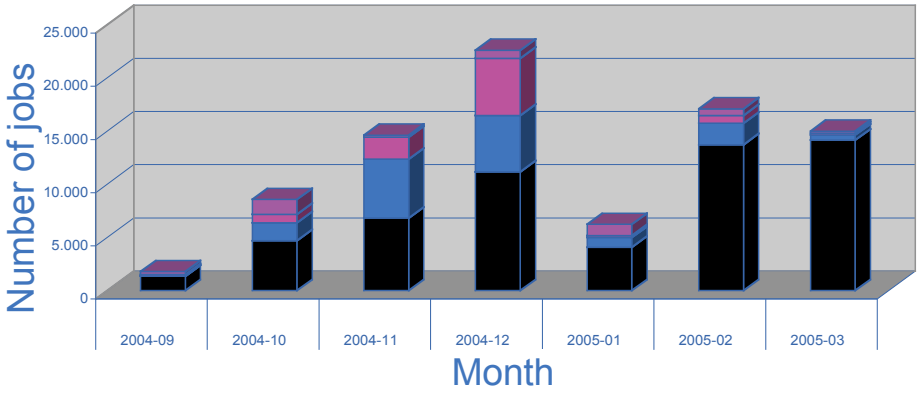


Developed by e-Science, HEP
Imperial College London

BioMed Overview

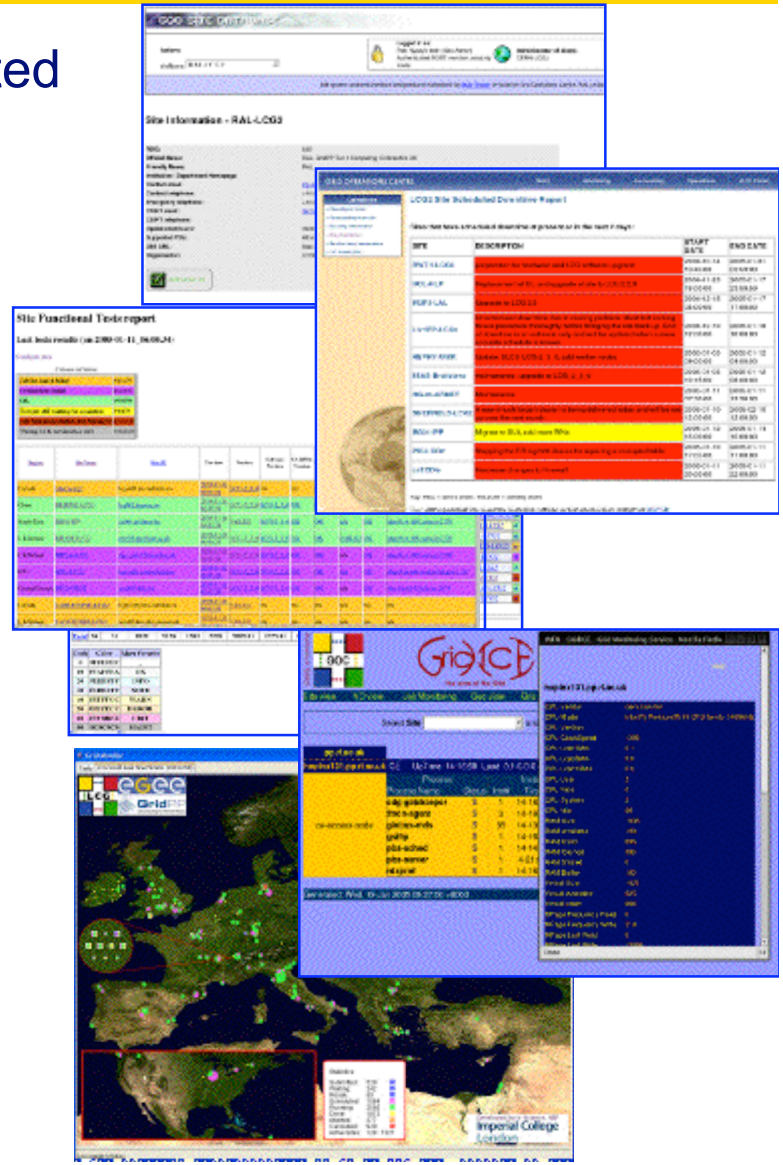
- Infrastructure
  - ~3.000 CPUs
  - ~12 TB of disk
  - in 9 countries

- >50 users in 7 countries working with 12 applications
- 18 research labs



EGEE Operations Process

- However the resources need to be operated
- Grid operator on duty
  - 6 teams working in weekly rotation
    - CERN, IN2P3, INFN, UK/I, Ru,Taipei
  - Crucial in improving site stability and management
  - Expanding to all ROCs in EGEE-II
- Responsibility for operations is geographically distributed:
  - There is no “central” operation
  - Operation Tools are developed/hosted at different sites:
    - GOC DB (RAL), SFT (CERN), GStat (Taipei), CIC Portal (Lyon)
- but strong Operations coordination
  - Weekly operations meetings
  - Regular ROC managers meetings
  - Series of EGEE Operations Workshops



Site Functional Tests

- **Site Functional Tests (SFT)**
  - Framework to test (sample) services at all sites
  - Shows results matrix
  - Detailed test log available for troubleshooting and debugging
  - History of individual tests is kept
  - Can include VO-specific tests (e.g. sw environment)
  - Normally >80% of sites pass SFTs
    - NB of 180 sites, some are not well managed

- Very important in stabilising sites:
  - Apps use only good sites
  - Bad sites are automatically excluded
  - Sites work hard to fix problems

Test summary

	SD	JL	JS	CT	OK	total
dteam	15	12	4	6	139	176
lhcb	15	81	5	35	39	175

Colours definition

SD	Scheduled downtime	#a3a3a3
JL	Job list match failed	#aab3ff
JS	Job submission failed	#f4876b
CT	Critical tests failed	#f9d48e
NT	Non-critical tests failed	#f2f98e
OK	OK	#b2f98e

Test abbreviations

csh	CSH test
swdir	VO software directory
rgma	R-GMA
dirac-test	Dirac full test
ver	Software Version (WN)
wn	WN host name
ca	CA certs version
crl	CRL timestamp test
rm	Replica Management
votag	VO Tag management
js	Job submission
bi	BrokerInfo

	St.	Site Name	Site CE	VO dteam												VO lhcb		
				St.	js	ver	wn	ca	rgma	bi	csh	rm	votag	swdir	crl	St.	js	dirac-test
	AsiaPacific																	
1.	CT	INDIACMS-TIFR	ce.indiacms.res.in	CT	O	2 6 0	I	O	O	O	O	X	O	O	!!!	JL	X	??
2.	OK	TW-NCUHEP	grid01.phy.ncu.edu.tw	OK	O	2 6 0	I	O	O	O	O	O	O	O	!!!	JL	X	??
3.	OK	TOKYO-LCG2	dgce0.icepp.jp	OK	O	2 4 0	I	O	O	O	O	O	O	O	!!!	JL	X	??
4.	OK	Taiwan-LCG2	lcg00125.grid.sinica.edu.tw	OK	O	2 6 0	I	O	O	O	O	O	O	O	!!!	JL	X	??
5.	OK	Taiwan-IPAS-LCG2	testbed001.phys.sinica.edu.tw	OK	O	2 6 0	I	O	O	O	O	O	O	O	!!!	JL	X	??
6.	OK	GOG-Singapore	melon.ngpp.ngp.org.sg	OK	O	2 6 0	I	O	O	O	O	O	O	O	!!!	JL	X	??
7.	OK	Taiwan-NCUCC-LCG2	ce.cc.ncu.edu.tw	OK	O	2 6 0	I	O	O	O	O	O	O	O	!!!	OK	O	O
8.	OK	LCG_KNU	cluster50.knu.ac.kr	OK	O	2 5 0	I	O	O	O	O	O	O	O	!!!	CT	O	!!!
	BNL																	
9.	SD	BNL-LCG2	lcg-ce01.usatlas.bnl.gov	SD	X	??	??	??	??	??	??	??	??	??	??	SD	X	??
	Canada																	
10.	JL	TORONTO-LCG2	bigmac-lcg-ce.physics.utoronto.ca	JL	X	2 6 0	I	O	O	O	O	O	W	O	!!!	OK	O	O
11.	SD	CARLETONU-LCG2	lcg02.physics.carleton.ca	SD	X	??	??	??	??	??	??	??	??	??	??	SD	X	??
12.	OK	TRIUMF-LCG2	lcgce01.triumf.ca	OK	O	2 6 0	I	O	O	O	O	O	O	O	O	OK	O	O
13.	OK	Umontreal-LCG2	lcg-ce.lps.umontreal.ca	OK	O	2 6 0	I	O	O	O	O	O	W	O	!!!	OK	O	O

Middleware Distributions and Stacks

- **Terminology:**
  - EGEE deploys a *middleware distribution*
    - Drawn from various middleware products, stacks, etc.
    - Do not confuse the *distribution* with development projects or with software packages
    - Count on 6 months from software developer "release" to production deployment
  - The EGEE distribution:
    - Moving from version labelled: LCG-2.7.0
    - to version labelled: gLite-3.0

- **EGEE distribution contents:**

❖ **LCG-2.7.0:**
- VDT – packaging Globus 2.4, Condor, MyProxy
- EDG workload management
- LCG components:
  - BDII (info sys),
  - catalogue (LFC),
  - DPM, data management libraries and CLI tools
  - monitoring tools
- gLite: R-GMA, VOMS, FTS

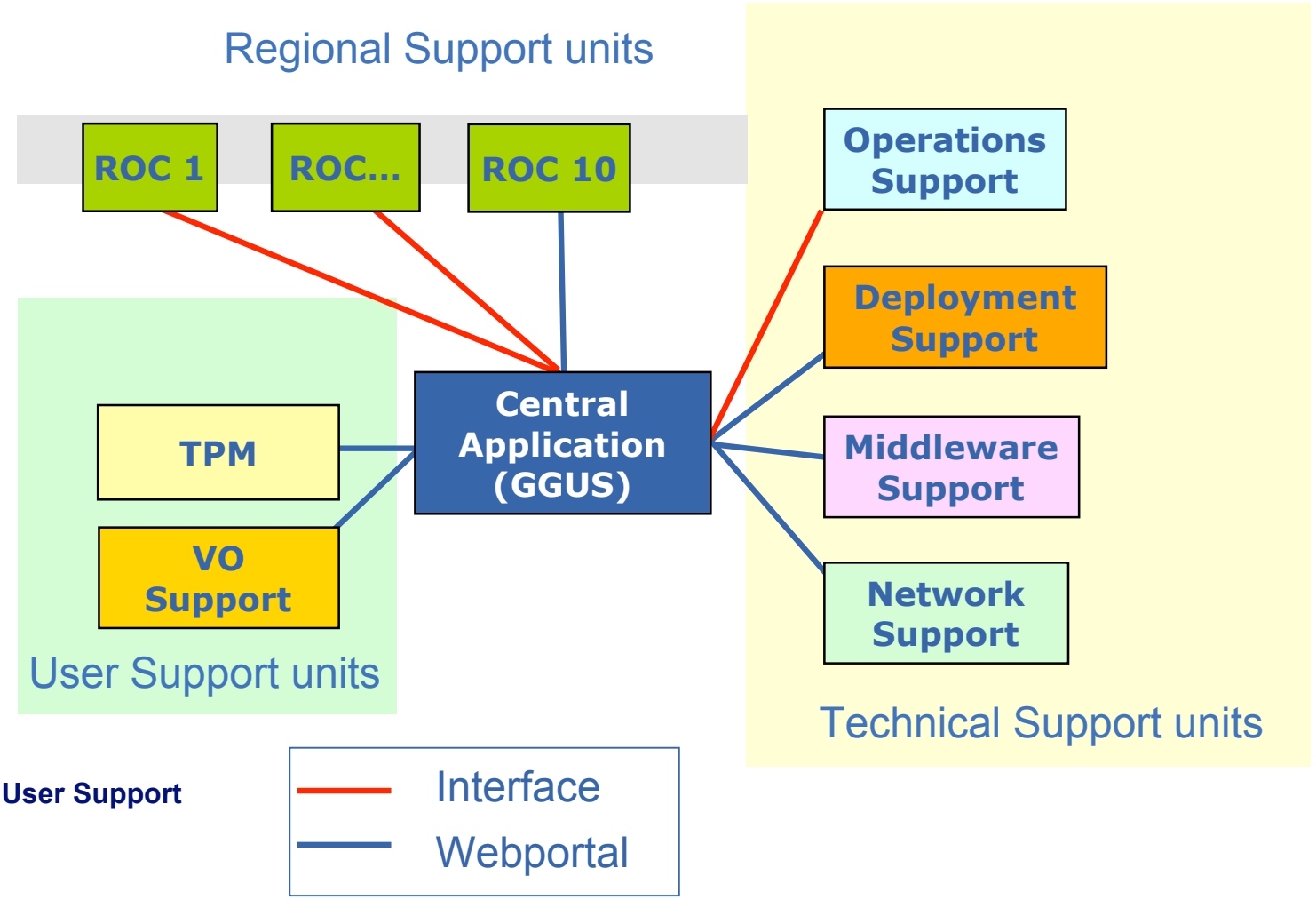
moving

❖ **gLite-3.0, released on May 4,2006:**
- Based on LCG-2.7.0, and
- gLite workload management
- Other gLite components (not in the distribution but provided as services):
  - AMGA, Hydra, Fireman
  - gLite-IO

The Support Model

"Regional Support with Central Coordination"

The ROCs, VOs and other project-wide groups such as the middleware groups (JRA), network groups (NA), service groups (SA) are connected via a central integration platform provided by GGUS.



GGUS = Global Grid User Support

The Support Model

"Regional Support with Central Coordination"

The ROCs, VOs and other wide groups as middle groups network (NA), service groups (SA) are connected via a central integration platform provided by GGUS.

Regional Support units

- GGUS is now being used for all problem reporting:
  - Operational, deployment and user support
  - VOs are using it for their support system
  - The use is growing steadily

perations Support

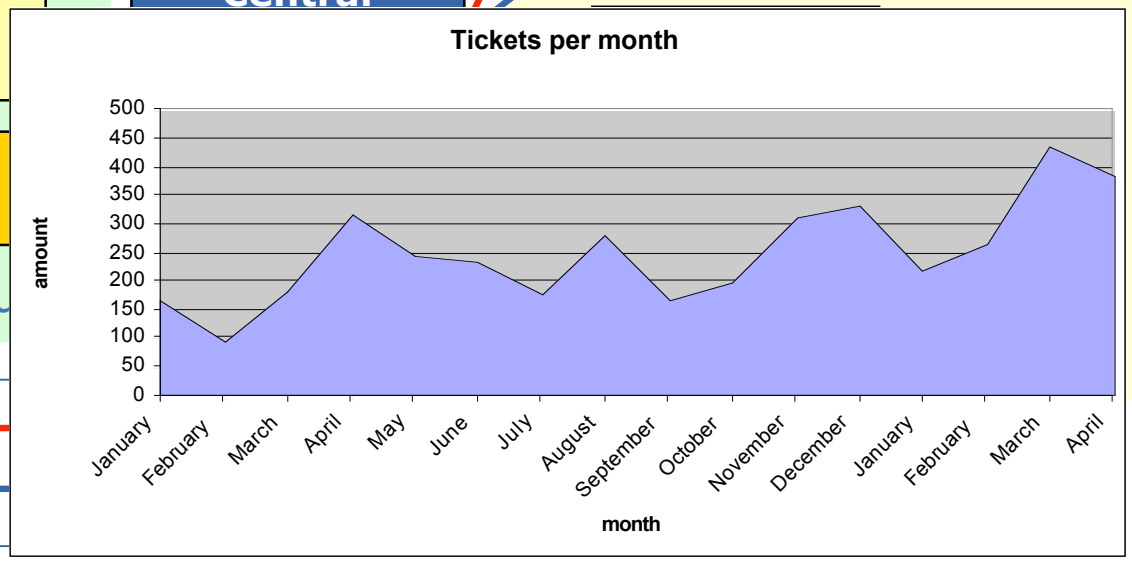
Deployment Support

Central

TPM

VO Support

User Support u



GGUS = Global Grid User Support

Outline

- The EGEE II project
- The related projects
- The EGEE infrastructure
  - **The INFN production Grid**
- The middleware (gLite 3.0)
- The GILDA testbed (the t-infrastructure)
- The applications
- The future
- Conclusions

The INFN production Grid

37 'resource centers':

25 site also part of the EGEE/LCG registered **GOCDB**

12 additional sites reachable only from inside Italy

The INFN production Grid : the resources

Site	Power	WN#	CPU#
CNR-ILC-PISA	18K	2	4
INAF-Trieste	35K	8	16
INFN-BARI	400K	38	76
INFN-BOLOGNA	98K	4	16
INFN-BOLOGNA-CMS	52K	7	14
INFN-CAGLIARI	92K	11	22
INFN-CATANIA	538K	43	104
INFN-CNAF	54K	6	11
INFN-FERRARA	91K	10	20
INFN-FIRENZE	379K	16	64
INFN-FRASCATI	34K	3	6
INFN-GENOVA	42K	4	8
INFN-LECCE	11K	2	4
INFN-LNL-2	776K	70	140
INFN-MILANO	264K	26	52
INFN-NAPOLI	229K	22	44
INFN-NAPOLI-ATLAS	-	-	-
INFN-NAPOLI-CMS	67K	7	14
INFN-NAPOLI-VIRGO	4K	1	2
INFN-PADOVA	531K	50	100
INFN-PERUGIA	200K	22	44
INFN-PISA	12K	2	4
INFN-PISA2	75K	13	23
INFN-ROMA1	253K	22	46
INFN-ROMA1-CMS	54K	5	12
INFN-ROMA1-VIRGO	31K	5	10
INFN-ROMA3	38K	4	8
INFN-T1	8269K	711	1422
INFN-TORINO	248K	24	48
INFN-TRIESTE	4K	1	2
ITB-BARI	294K	12	48
SNS-PISA	24K	3	6
SPACI-LECCE-IA64	15K	7	14
SPACI-NAPOLI	8K	3	3
SPACI-NAPOLI-IA64	29K	7	14
UNI-PERUGIA	-	-	-

2500 CPU’s in total

The INFN production Grid : the Supported VO's

http://grid-it.cnaf.infn.it/index.php?voregister&type=1

VO	Home page	LDAP/VOMS Server	Base DN/VOMS	Register (*)
alice	LHC Alice experiment	lcg-vo.cern.ch	ou=lcg1,o=alice,dc=eu-datagrid,dc=org	Click here!
atlas	LHC Atlas experiment	lcg-vo.cern.ch	ou=lcg1,o=atlas,dc=eu-datagrid, dc=org	Click here!
argo	INFN ARGO-YBJ experiment	voms.cnaf.infn.it	VOMS	Click here!
babar	Babar experiment	babar-vo.gridpp.ac.uk	ou=babar, dc=gridpp, dc=ac,dc=uk	Click here!
bio	Grid.it Biology group	voms.cnaf.infn.it	VOMS	Click here!
biomed	BIOMED	vo-biome.in2p3.fr	ou=lcg1,o=biomedical,dc=lcg,dc=org	Click here!
cms	LHC CMS experiment	lcg-vo.cern.ch	ou=lcg1,o=cms, dc=eu-datagrid, dc=org	Click here!
cdf	CDF experiment	voms.cnaf.infn.it	VOMS	Click here!
compchem	Dipartimento di Chimica-Universita' di Perugia	voms.cnaf.infn.it	VOMS	Click here!
dteam	LCG Deployment	lcg-vo.cern.ch	ou=lcg1, o=dteam, dc=lcg, dc=org	Click here!
egrid	EGRID experiment	voms.cnaf.infn.it	VOMS	Click here!
enea	ENEA	voms.cnaf.infn.it	VOMS	Click here!
esr	ESR Home	grid-vo.sara.nl	ou=lcgadmin,o=esr,dc=eu-egee,dc=org	Click here!
gridit	General Grid.it Project VO	voms.cnaf.infn.it	VOMS	Click here!
inaf	INAF	voms.cnaf.infn.it	VOMS	Click here!
infngrid	INFN-GRID project	voms.cnaf.infn.it	VOMS	Click here!
ingv	INGV Bologna	voms.cnaf.infn.it	VOMS	Click here!
lhcb	LHC LHCb experiment	lcg-vo.cern.ch	ou=lcg1,o=lhcb, dc=eu-datagrid, dc=org	Click here!
libi	**NEW VO**	voms.cnaf.infn.it	VOMS	Click here!
magic	MAGIC and grid	grid-vo.sara.nl	o=magic,dc=eu-egee,dc=org	
pamela	**NEW VO**	voms.cnaf.infn.it	VOMS	Click here!
planck	Planck Experiment	voms.cnaf.infn.it	VOMS	Click here!
theophys	INFN theorical physics group	voms.cnaf.infn.it	VOMS	Click here!
virgo	INFN Virgo experiment	voms.cnaf.infn.it	VOMS	Click here!
zeus	ZEUS experiment	grid-vo.desy.de	ou=zeus, ou=vo, o=desy, c=de	Click here!

(*) Your personal certificate must be already installed in your browser.

The INFN production Grid : the Central Management Team and the Regional Operation Center

- The CMT
  - The CMT is responsible of the certification: checking the functionalities of a site before to join the site to the production grid. In particular checks:
    - GIIS' information consistence
    - Local jobs submission (LRMS)
    - Grid submission with Globus (globus-job-run)
    - Grid submission with the ResorceBroker
    - ReplicaManager functionalities

- The ROC
  - First level support
    - Geographically based local front line support to Virtual Organization, Users and Resources Centres
    - Daily shift and Check list to be covered during the shift
    - Periodic (every 15 days) phone conference
      - ROC/CIC teams and site managers
    - Periodic reports to GDA
  - Second level support: CIC on Duty
    - Weekly shift
    - CIC tools use and development
  - Ticket system

Outline

This presentation will cover the following arguments:

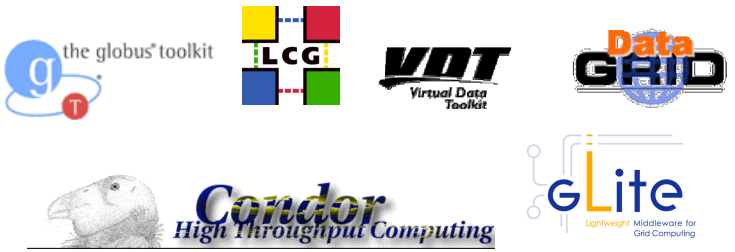
- The EGEE II project
- The Grid related projects
- The EGEE infrastructure
  - The INFN production Grid
- **The middleware (gLite 3.0)**
- The GILDA testbed (the t-infrastructure)
- The applications
- The evolution
- Conclusions

gLite 3.0 components

- Authentication
  - based on X.509 PKI infrastructure
  - uses (short lived) proxies
    - of the user certificates

- VO Membership Service: VOMS
  - Bare certificates are not enough for defining user capabilities on the Grid
  - Users belong to VO's, to groups inside a VO and may have special roles
  - VOMS provides a way to add attributes to a certificate proxy
  - The attributes are used to provide the user with additional capabilities according to the VO policies.

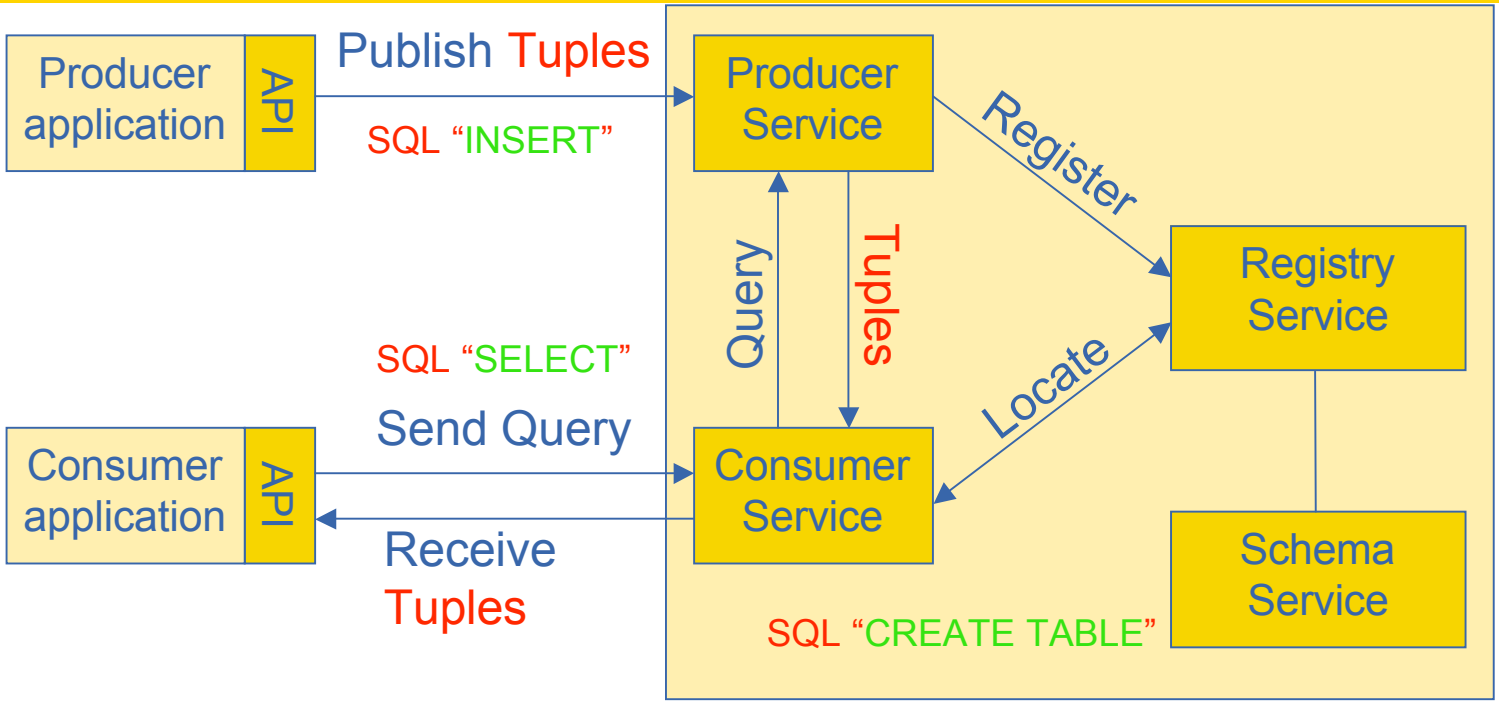
- gLite-3.0, released on May 4, 2006 :
  - Based on components from



gLite 3.0 components

- Accounting - DGAS
  - DGAS: accumulates Grid accounting information
    - *User, JobId, user VO, VOMS FQAN(role,capabilities), SI2K, SF2K, system usage (cpuTime, wallTime…),…*
    - allows billing and scheduling policies
    - levels of granularity: from single jobs to VO or grid aggregations

gLite 3.0 components: R-GMA



- The Relational Grid Monitoring Architecture (R-GMA) provides a uniform method to access and publish both information and monitoring data.
- From a user's perspective, an R-GMA installation currently appears similar to a single relational database.
- Relational implementation of the GGF's Grid Monitoring Architecture (GMA)

gLite 3.0 components: the Workload Management System

- Helps the user accessing computing resources
  - resource brokering
  - management of input and output
  - management of complex workflows
- Backward compatible with LCG-2
  - Capable of submitting to the LCG2.7 Computing Element

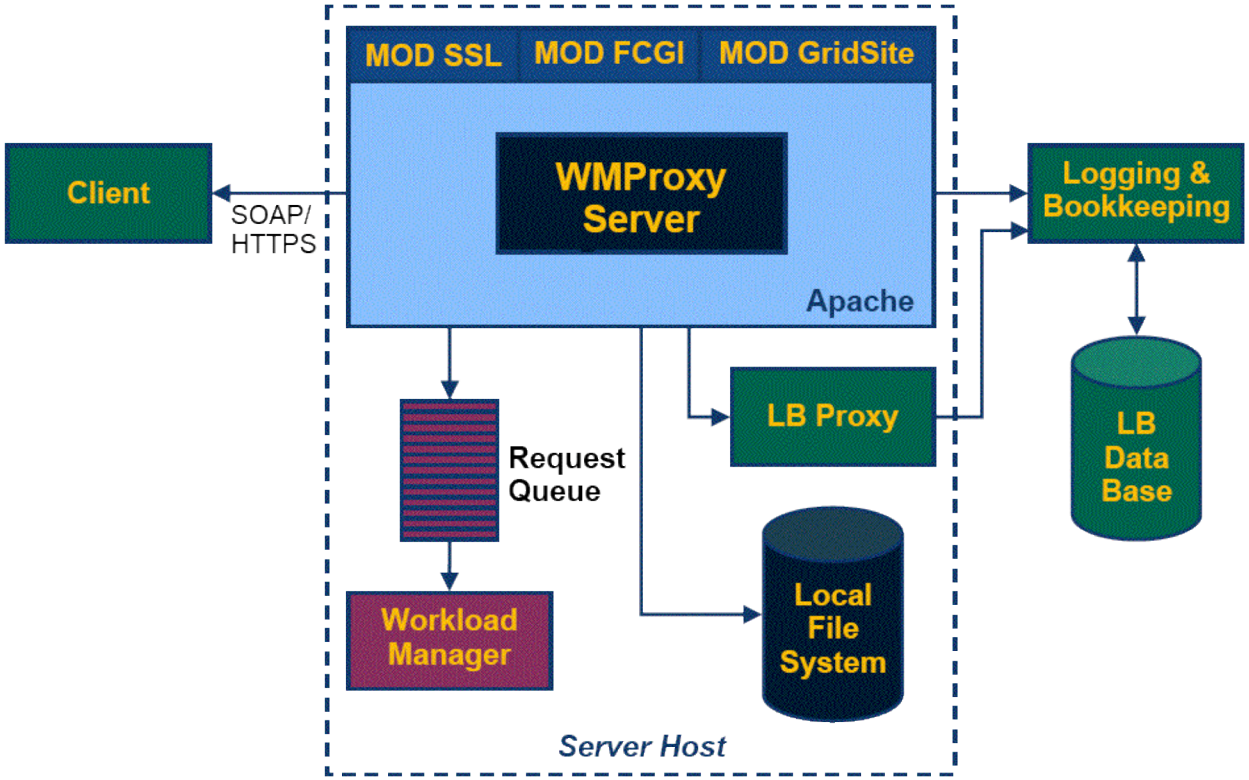
- Glite Computing Element
  - Condor-C GSI enabled
  - Uses CEMon to publish information (but also support provided for R-GMA and bdII)
  - and BLAH, Batch Local ASCII Helper, as Interface between the CE and the Local Resource Manager System

gLite 3.0 components: the Workload Management System

- Support provided for
  - *Bulk job submission*
  - *DAG job submission*
    - Support for execution of all DAG nodes within a single CE
      - chosen by user or by the WMS matchmaker
  - *deep or shallow resubmission*
    - Shallow Resubmission happens in case of failure only when the job didn't start
    - Deep Resubmission happens if the job didn't conclude all its operations
  - MPI jobs, even if the file system is not shared between CE and Worker Nodes (WN)
  - Push/pull mode job submission
  - collection of information from many sources
    - CEMon, bdII, R-GMA
  - Data management interfaces (DLI and StorageIndex)
  - file peeking during job execution (Job File Perusal)
  - for pilot job (initial support)
    - Pilot job which "prepare" the execution environment and then get and execute the actual user job

gLite 3.0 components: WMProxy

- WMProxy is a SOAP Web service providing access to the Workload Management System (WMS)
- Job characteristics specified via JDL



gLite 3.0 Data management components

Data Management

- File and replica catalogues (LFC)
  - Central or local (not distributed)
  - Replication via Oracle, or squid caches tested by LCG
  - Secure
- File Transfer Service (FTS)
  - Reliable data transfer
  - Uses gridftp or srmcopy as transport
- Storage Elements based on SRM interface
  - DPM: implements Posix ACLs, VOMS roles/groups
  - Other available SEs: dCache, Castor
  - Deprecated: “Classic SE” – basically just gridftp
- Metadata catalogue:
  - AMGA
- Secure Keystore (data encryption) :
  - Hydra
- Utilities and IO libraries:
  - Lcg-utils
  - GFAL – this is the SRM client library
  - gLiteIO – expect functionality to be replaced

gLite: components under development

- Job Provenance
  - Long term job information storage
  - Useful for debugging, post-mortem analysis, comparison of job executions in different environments
  - Useful for statistical analysis
- CREAM: web service Computing Element
- ICE: Interface to Cream Environment
- GPBOX: Distributed VO policy management tool

Outline

- The EGEE II project
- The Grid related projects
- The EGEE infrastructure
  - The INFN production Grid
- The middleware (gLite 3.0)
- **The GILDA testbed (the t-infrastructure)**
- The applications
- The future
- Conclusions

The t-Infrastructure

- Why a training infrastructure?
  - e-Infrastructure for production
  - t-Infrastructure for training

- Easy entry point for new communities
- Need safe resources for installation training
- Need guaranteed response for tutorials
- Need limit the vulnerability of production systems
  - use training grid
  - have training CA
  - able to change middleware to prepare participants for future releases on production system

The GILDA Project (https://gilda.ct.infn.it)

The GILDA Test-bed (https://gilda.ct.infn.it/testbed.html)

19 sites in 3 continents !

gLite and GILDA

- Grid tutorials
- GILDA Poster
- Video tutorials
- Live User Interface (NEW)
- User Interface PnP
- Instructions for users
- Instructions for sites (NEW)
- Useful links
- Sponsors
- Usage Statistics
- Old Usage Statistics

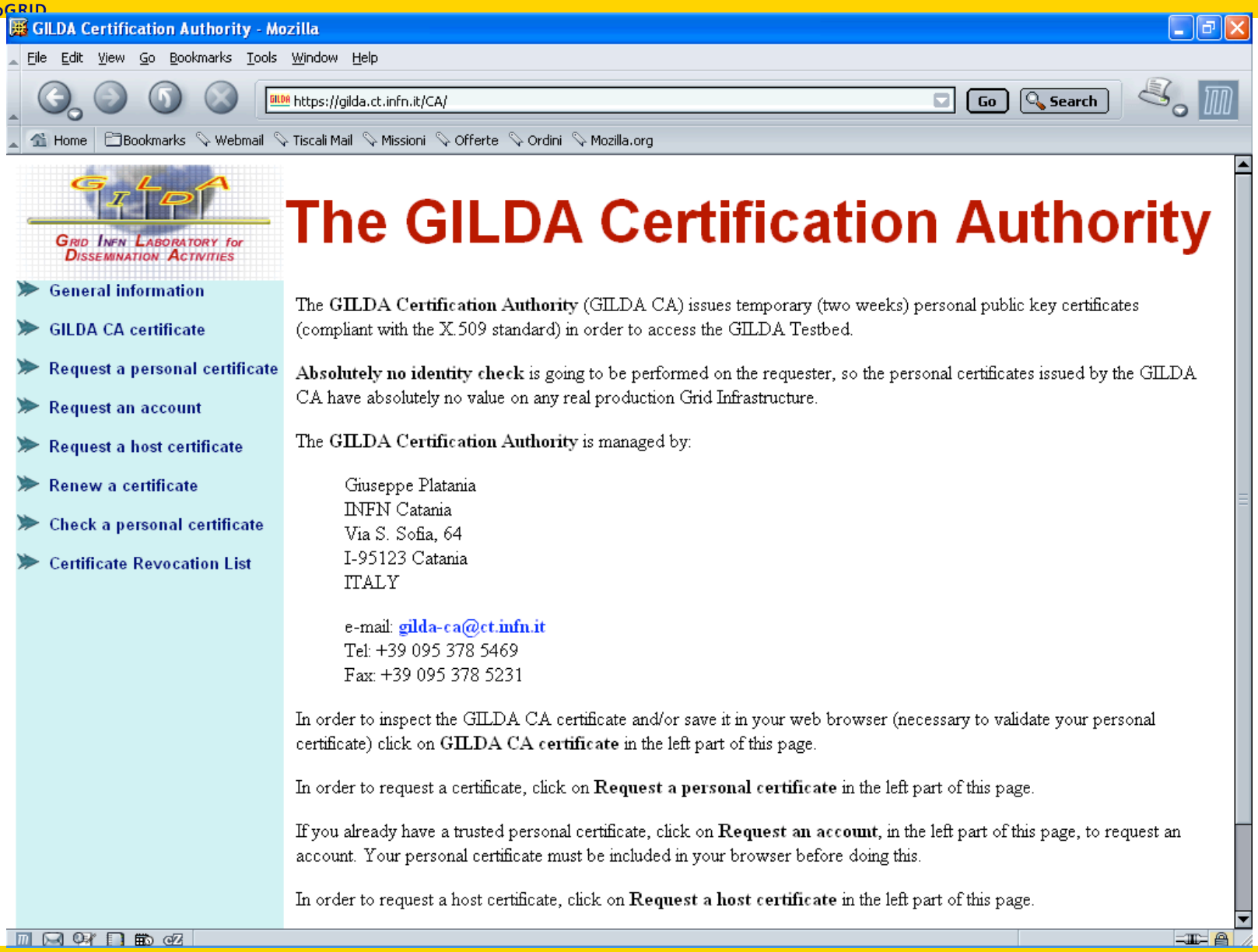
gLite Services

Grid services

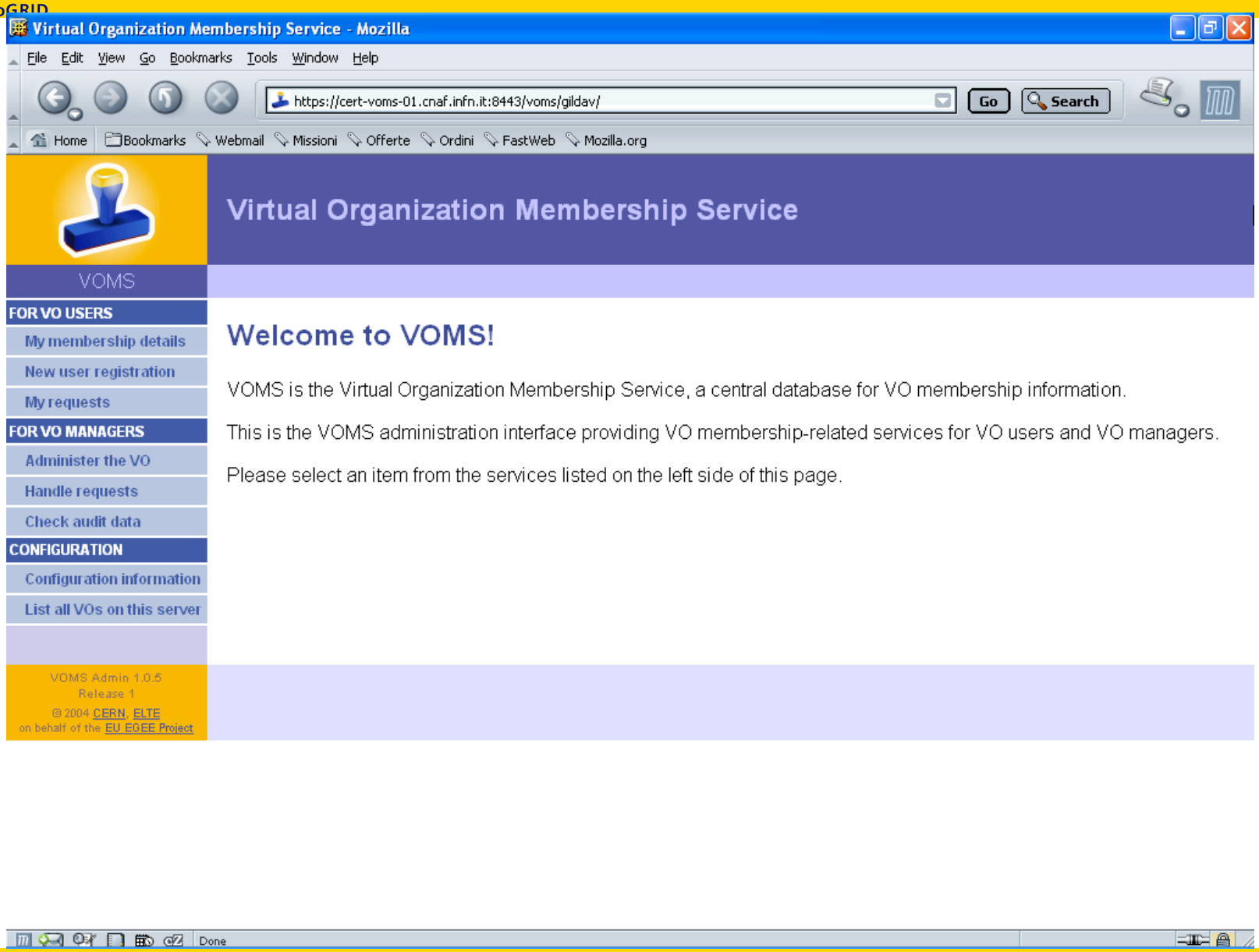
This is a table of the general Grid Services available on GILDA.

SERVICE	HOST
LCG-2 Resource Broker (RB)	grid004.ct.infn.it
LCG-2 Resource Broker (RB)	grid013.ct.infn.it
LCG-2 Resource Broker (RB)	skurut2.cesnet.cz
LCG-2 Resource Broker (RB)	rb.be.itu.edu.tr
gLite Resource Broker (RB)	glite-rb.ct.infn.it
gLite Resource Broker (RB)	glite-rb2.ct.infn.it
Information Index (BDII)	grid004.ct.infn.it
Backup Information Index (BDII)	grid013.ct.infn.it
GILDA VOMS server	voms.ct.infn.it:8443/voms/gilda
GILDA gLite DGAS Price Authority	grid-demo1.ct.infn.it:56568
GILDA gLite DGAS HLR	grid-demo1.ct.infn.it:56567
GridICE Monitoring System	alifarm7.ct.infn.it:50080
LCG-2 File Catalog (LFC)	lfc-gilda.ct.infn.it
Backup LCG-2 File Catalog (LFC)	lfc-gilda.cern.ch
gLite FiReMan Catalog	grid007.ct.infn.it
gLite File Transfer/Placement Service	fts.ct.infn.it:8443
AMGA Metadata Catalog	amga.ct.infn.it:8822
gLite R-GMA server	rgmasrv.ct.infn.it:8443
MyProxy Server	grid001.ct.infn.it
Backup MyProxy Server	grid014.ct.infn.it

The GILDA Certification Authority (https://gilda.ct.infn.it/CA)

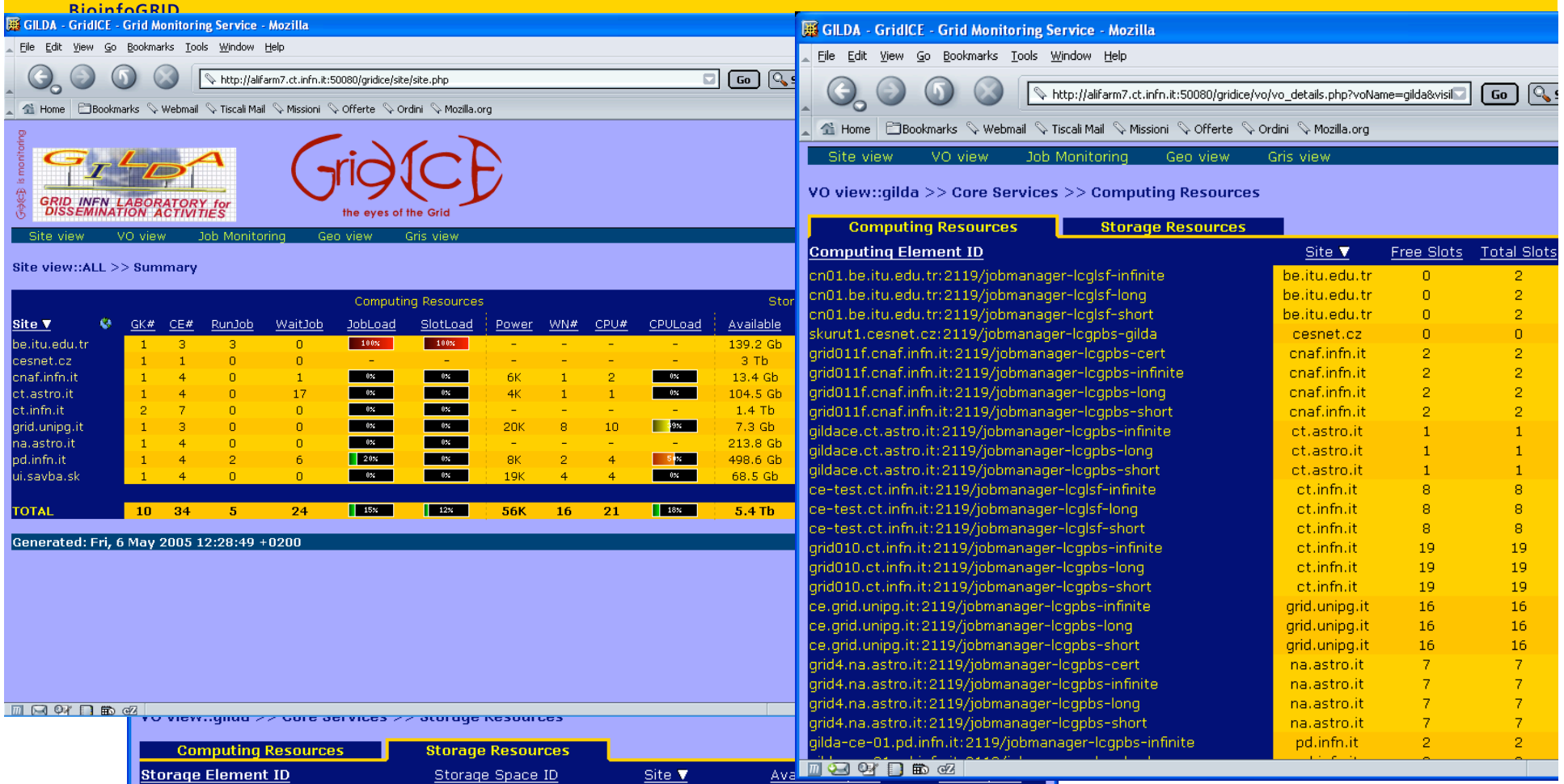


The GILDA VOMS

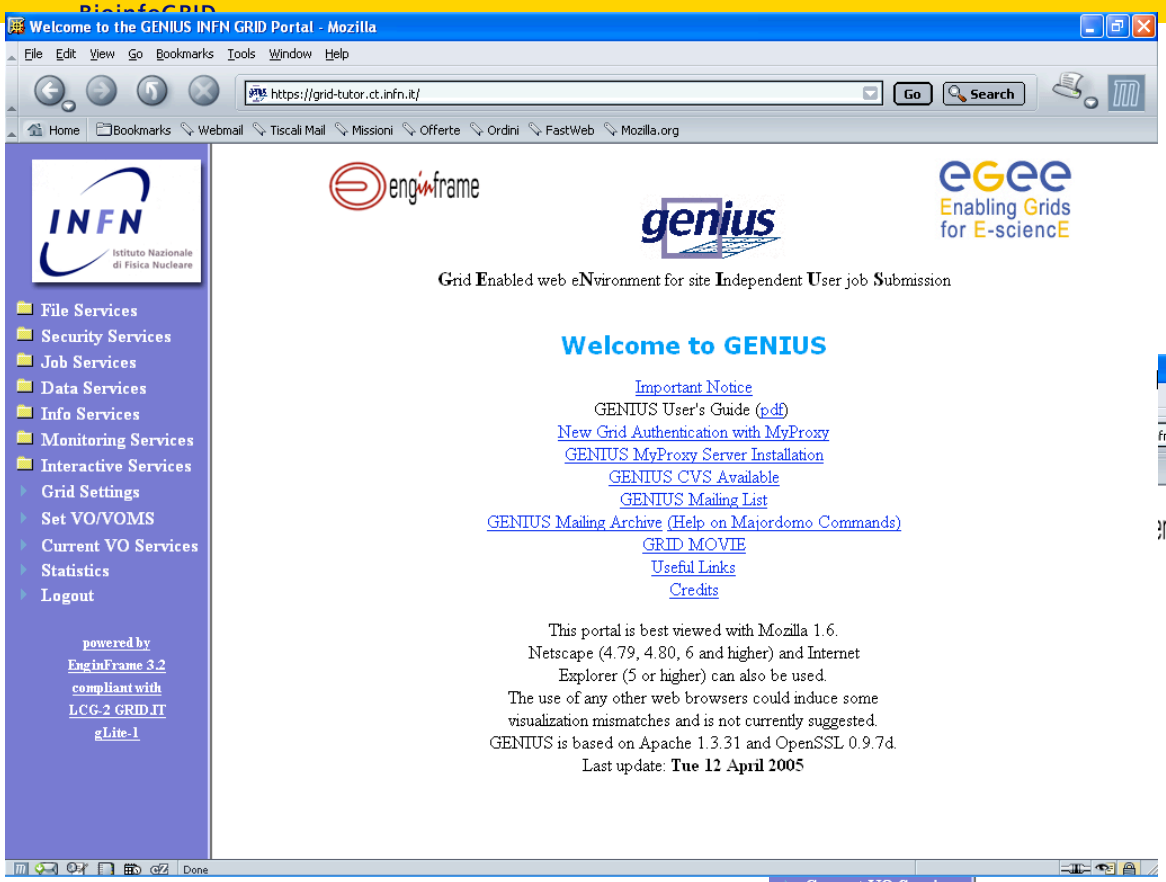


The GILDA Monitoring System (http://alifarm7.ct.infn.it:50080/gridice)

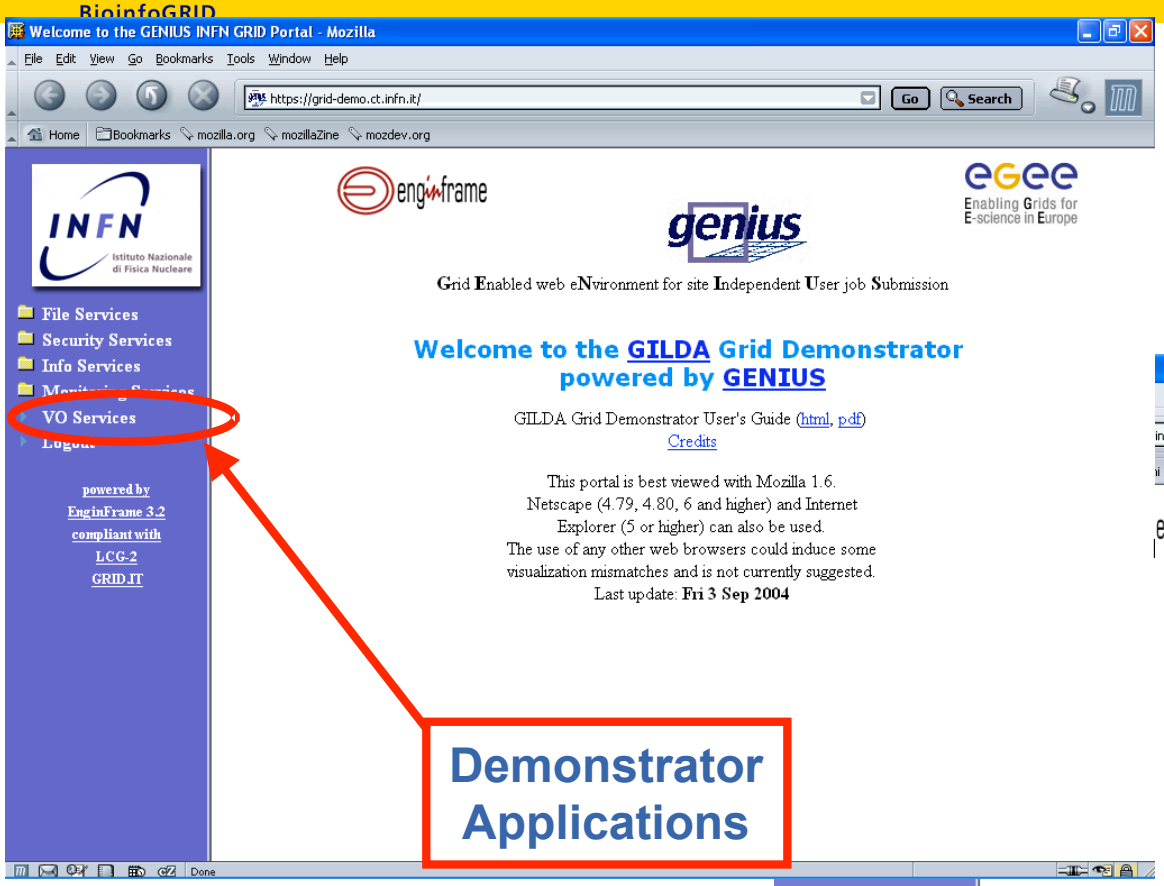
BioinfoGRID



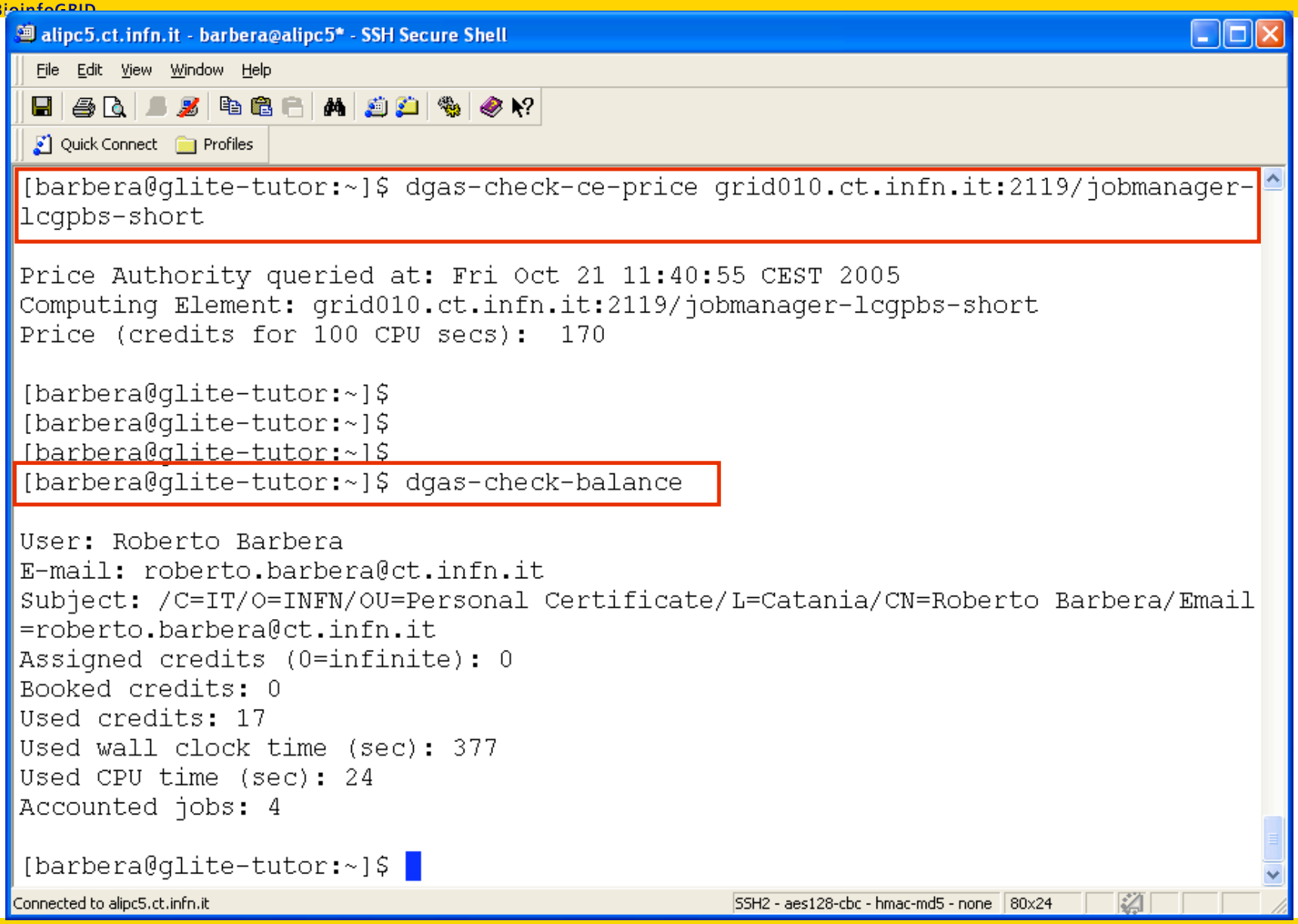
The Grid Tutor
(https://grid-tutor.ct.infn.it, https://glite-tutor.ct.infn.it)



The Grid Demonstrator (1/2)
(https://grid-demo.ct.infn.it, https://glite-demo.ct.infn.it)



The GILDA Accounting System (based on DGAS)



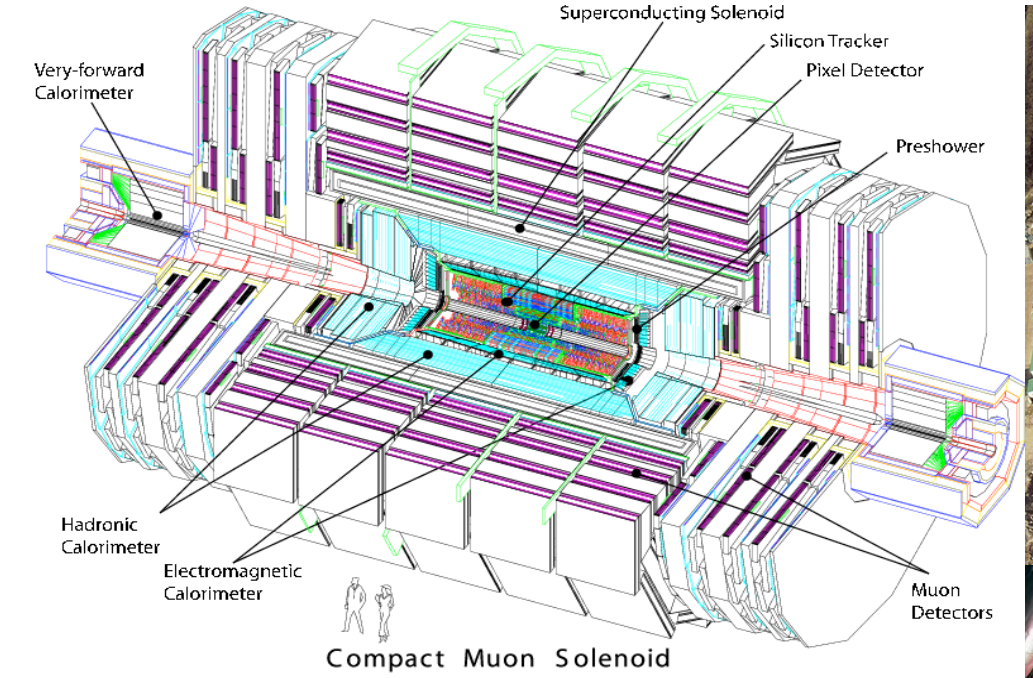
Outline

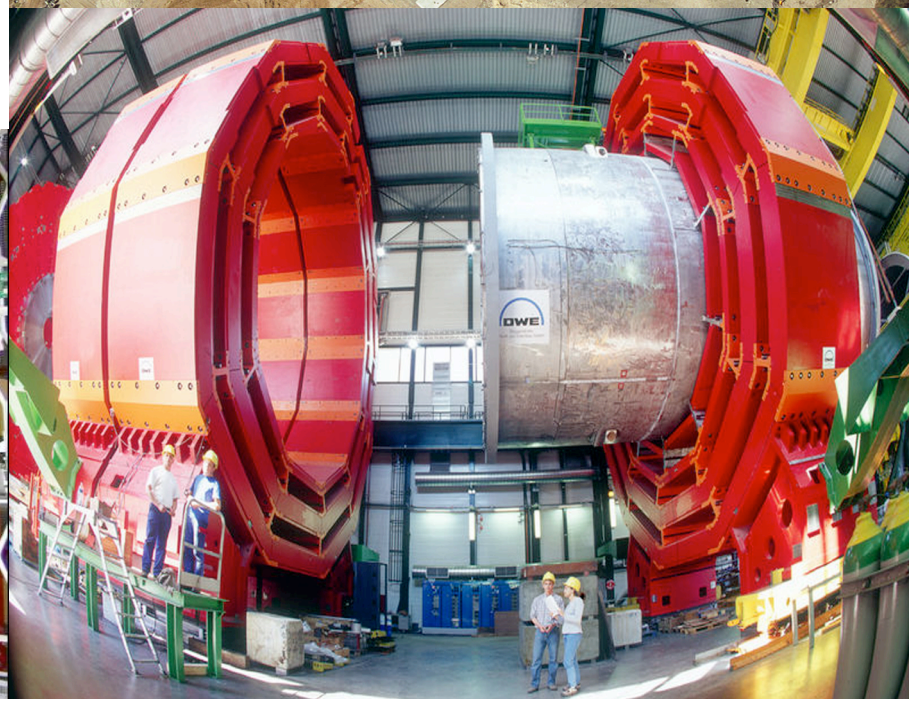
This presentation will cover the following arguments:

- The EGEE II project
- The Grid related projects
- The EGEE infrastructure
  - The INFN production Grid
- The middleware (gLite 3.0)
- The GILDA testbed (the t-infrastructure)
- **The applications**
- The future
- Conclusions

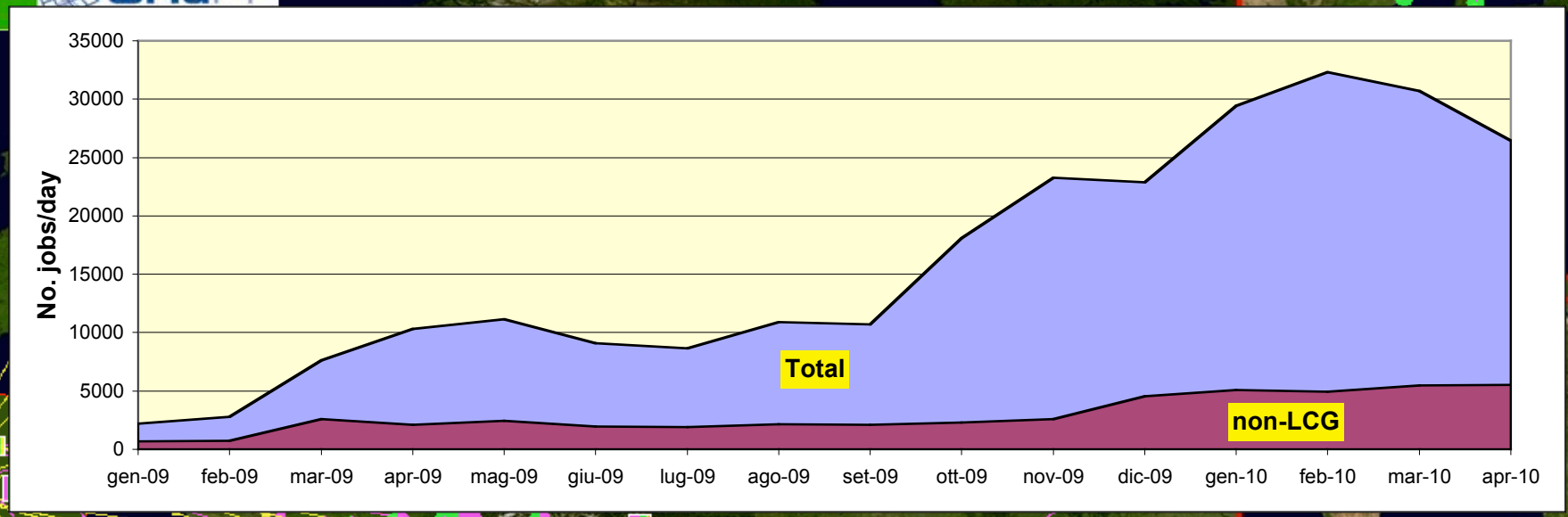
The LHC experiment:CMS

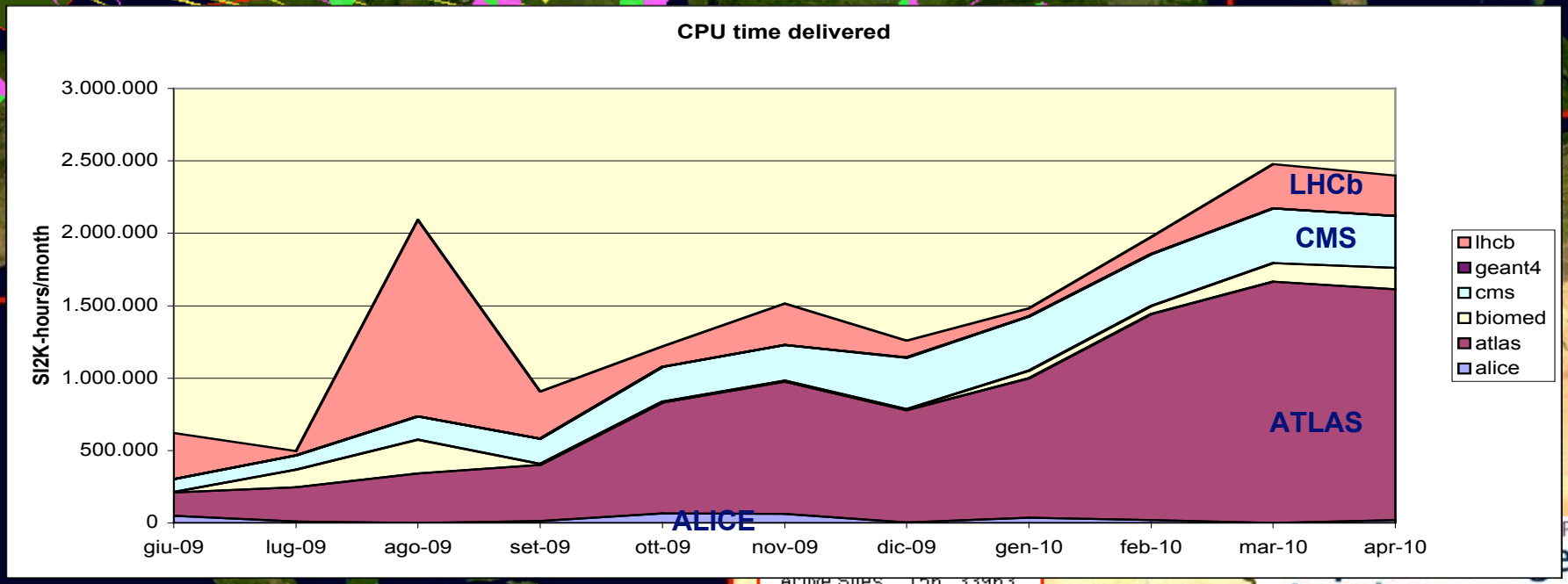
BioinfoGRID



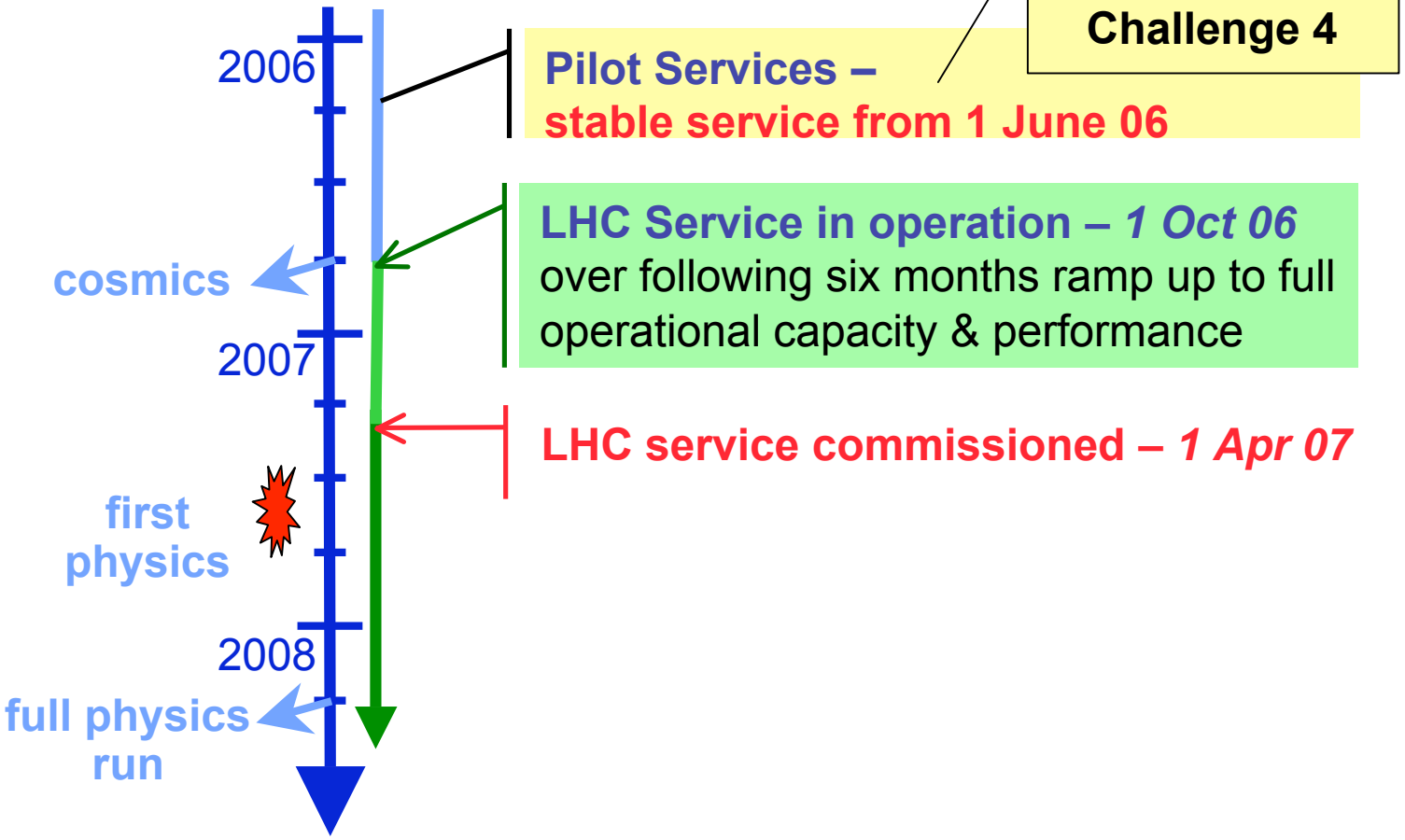


Use of the infrastructure





LCG Service Deadlines



les.robertson@cern.ch

NEWS RELEASE

Embargoed until 4 May 2006, 18:00 CCT (10:00 GMT, 12:00 MEST)

EGEE GRID ATTACKS AVIAN FLU

During April, a collaboration of Asian and European laboratories has analysed 300,000 possible drug components against the avian flu virus H5N1 using the EGEE Grid infrastructure. The goal was to find potential compounds that can inhibit the activities of an enzyme on the surface of the influenza virus, the so-called neuraminidase, subtype N1. Using the Grid to identify the most promising leads for biological tests could speed up the development process for drugs against the influenza virus.

One of the targets of existing drugs today on the market is viral neuraminidase, an enzyme that helps the virus to proliferate and infect more cells. As this protein is known to evolve into variants if it comes under drug stress, drug resistance becomes a potential concern in case of an influenza pandemic.

The challenge of the *in silico* drug discovery application is to identify those molecules which can dock on the active sites of the virus in order to inhibit its action. To study the impact of small scale mutations on drug resistance, a large set of compounds was screened against the same neuraminidase target but with various, slightly different structures. With the results from the *in silico* screening, researchers can predict which compounds and chemical fragments are most effective for blocking the active neuraminidases in case of mutations.

A GRID challenge to find functional analogous gene products

Donvito G.[2], Tulipano A.[1,2], Maggi G.[1,2], Gisel A.[3]

[1]Dipartimento Interateneo di Fisica, Università e Politecnico di Bari, via Amendola 173, 70126 Bari, Italy
[2]INFN Bari, Via Amendola 173, Bari, Italy
giorgio.maggi@ba.infn.it, giacinto.donvito@ba.infn.it, angelica.tulipano@ba.infn.it

[3]Istituto di Tecnologie Biomediche, CNR, Via Amendola 122/D, Bari, Italy
andreas.gisel@ba.itb.cnr.it

EGEE II new applications

- New application are moving to the GRID
  - Astrophysics
  - Computational Chemistry
  - Earth Sciences
  - Finance
  - Fusion
  - Geophysics (incl. industrial application EGEODE)
  - Multimedia
  - …

- In EGEE new application were selected (and followed) by an internal board (EGEE EGAAP)

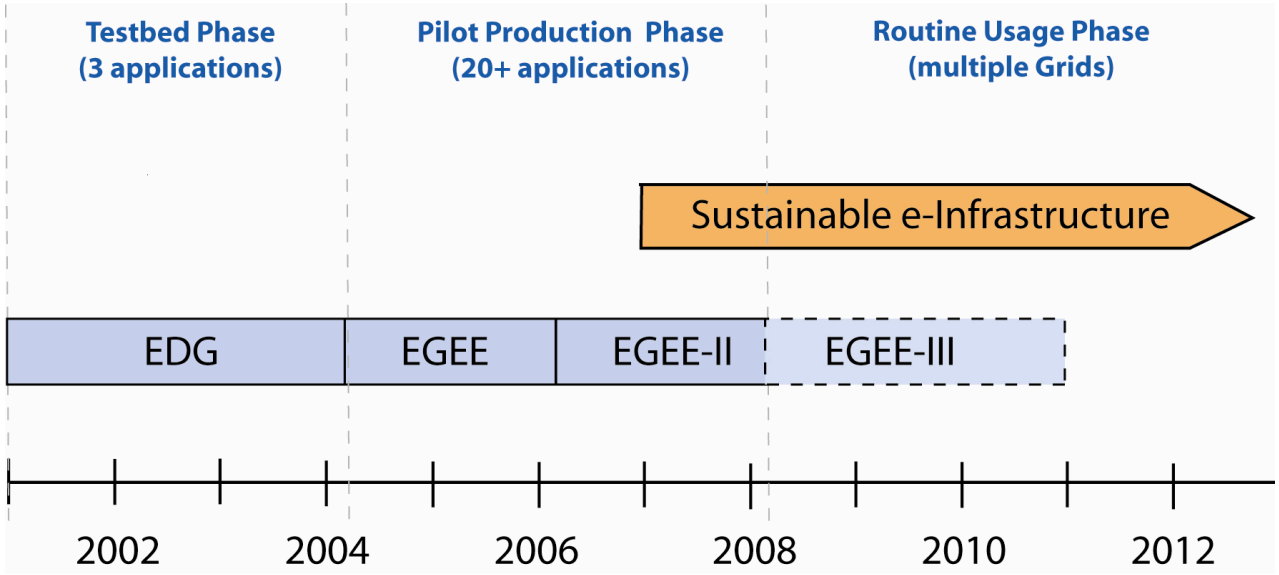
Outline

This presentation will cover the following arguments:

- The EGEE II project
- The Grid related projects
- The EGEE infrastructure
  - The INFN production Grid
- The middleware (gLite 3.0)
- The GILDA testbed (the t-infrastructure)
- The applications
- **The future**
- Conclusions

Sustainability: Beyond EGEE-II

- **Need to prepare for permanent Grid infrastructure**
  - Maintain Europe's leading position in global science Grids
  - Ensure a reliable and adaptive support for all sciences
  - Independent of project funding cycles
  - Modelled on success of GÉANT
    - Infrastructure managed centrally in collaboration with national bodies (in EGEE-II: **JRU**s)



A European R&E Networking Model

Connect. Communicate. Collaborate

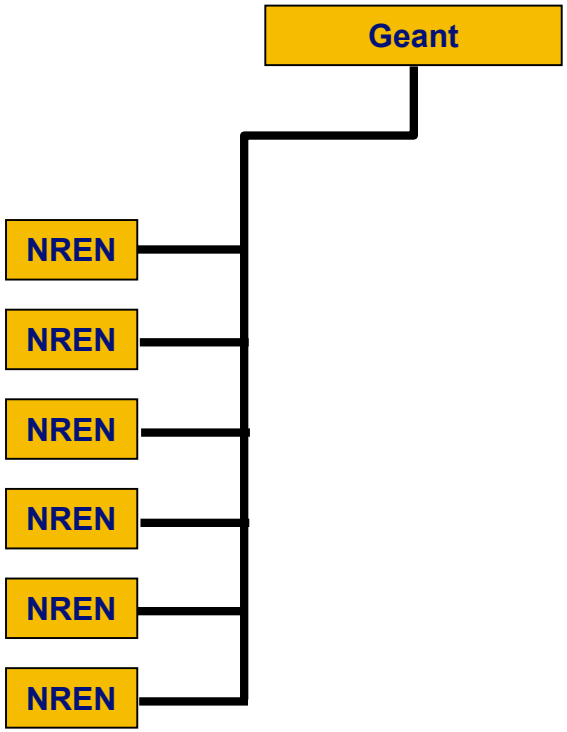
- Interconnects **34 National Research & Education Networks-NRENs** of the extended European Research Area (ERA)
- Connects more than **3500 Research & Education Institutions**
- Serves millions of end-users + ***eScience* Projects** (e.g. Grids) under *Accepted Usage Policy* (AUP) rules
- The model: **A 3-tier Federal Architecture**, partially subsidized by National and EU Research & Education funds:
  - The Campus Network (LAN/MAN)
  - The NREN (MAN/WAN)
  - The Pan-European Interconnection: **TEN34 → TEN155 → GÉANT** (GN1 in FP5) → **GÉANT2** (GN2 in FP6)**: Hybrid Optical Backbone (+ Cross Border Fibers)**

GN2 EC Subsidy < 10% of total European R&E Networking Cost

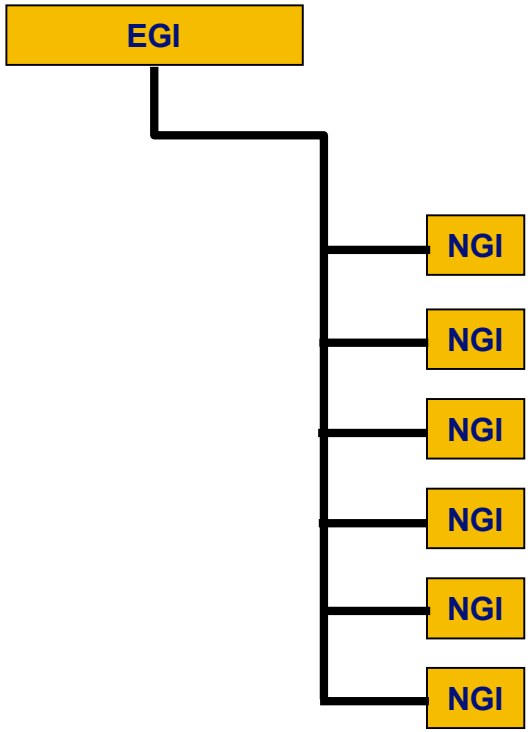
- **Governance:** NREN Policy Committee, GN2 Exec, DANTE, TERENA

A possible schema

- NETWORK



GRID



EGI = European Grid Infrastructure
NGI = National GRID Infrastructure

EGEE II as part of its program will prepare for sustainable European Grid Infrastructure
In each country a National Grid Infrastructure needs to be set up

IGI: the Italian Grid Infrastructure

- Discussion is going on between EPR, Consortia, GARR, etc,
  - to form a *temporary scope association*
  - as a first step toward the construction of an IGI (Italian Grid Infrastructure) for the public research.

- Possible IGI partners:
  - EPR's
    - CNR, ENEA, INAF, INFN, INGV, MIUR
  - Resource and network providers
    - UNINA, ICTP, ELETTRA, GARR, CASPUR, CINECA, CILEA, COMETA, COSMOLAB, SPACI, CRS4

The sustainability of the Middelware

- The OMII-EU project

- **The OMII-Europe vision is to harvest open-source, WebServices-based, grid software from across Europe and to supply these key grid services in a form that will allow them to interoperate across heterogeneous infrastructures, in particular EGEE, UNICORE and Globus.**

- The OMII-EU project is strategic to develop and maintain a interoperable and robust grid middelware.

Conclusions

- Unprecedented scale of a world-wide grid infrastructure, production quality, has been set up by the EGEE project and will be improved by EGEE 2
  - Proven working model for operations and deployment
- The gLite Middleware has been released, with new functionalities. It can be expected that soon will become as robust as the LCG middleware.
- There are already applications which depend on the grid as their primary computing resources.
  - As the time goes on more applications are moving to the grid.
- There is still hard work to do either at national and European level for:
  - Providing a permanent multidisciplinary production grid infrastructure
  - Protecting investments of application developers and users
  - Ensuring that the grid infrastructure used today will still be there tomorrow