



Enabling Grids for E-science

Overview of the EGEE project and the gLite middleware

Gergely Sipos
MTA SZTAKI
sipos@sztaki.hu

www.eu-egee.org



- **What is EGEE?**
 - The project
 - The infrastructure
- **gLite middleware**
- **EGEE applications**
- **Sources of further information**

- **Flagship European grid infrastructure project, now in 2nd phase with 91 partners in 32 countries**

- **Objectives**

- Large-scale, production-quality grid infrastructure for e-Science
- Attracting new resources and users from industry as well as science
- Maintain and further improve gLite Grid middleware

- **Structure**

EGEE: 1 April 2004 – 31 March 2006

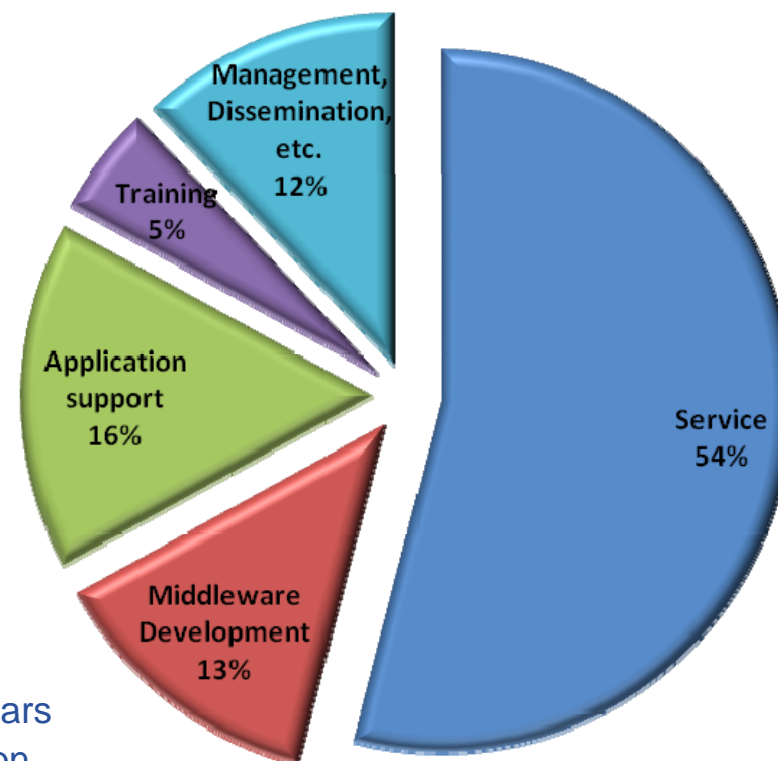
EGEE-II: 1 April 2006 – 31 March 2008

- Leveraging national and regional grid activities worldwide
- Funded by the EC at a level of ~37 M Euros for 2 years
- Support of related projects for infrastructure extension, application, specific services

- **EGEE-III**: 1 April 2008 – 31 March 2009

- Reaching self-sustainable state

EGEE Project Activities



- **From April 2006, natural continuation of EGEE**

- Expanded consortium
- Emphasis on providing an infrastructure
 - increased support for applications
 - interoperate with other infrastructures
 - more involvement from Industry

SA: service activities

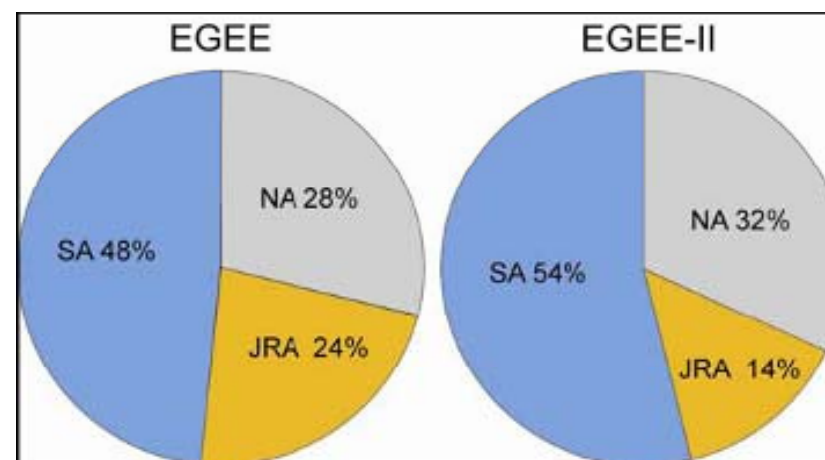
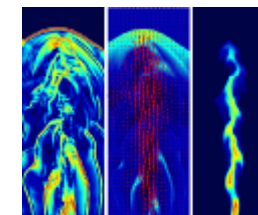
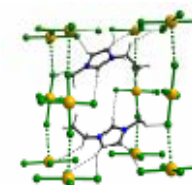
- establishing operations

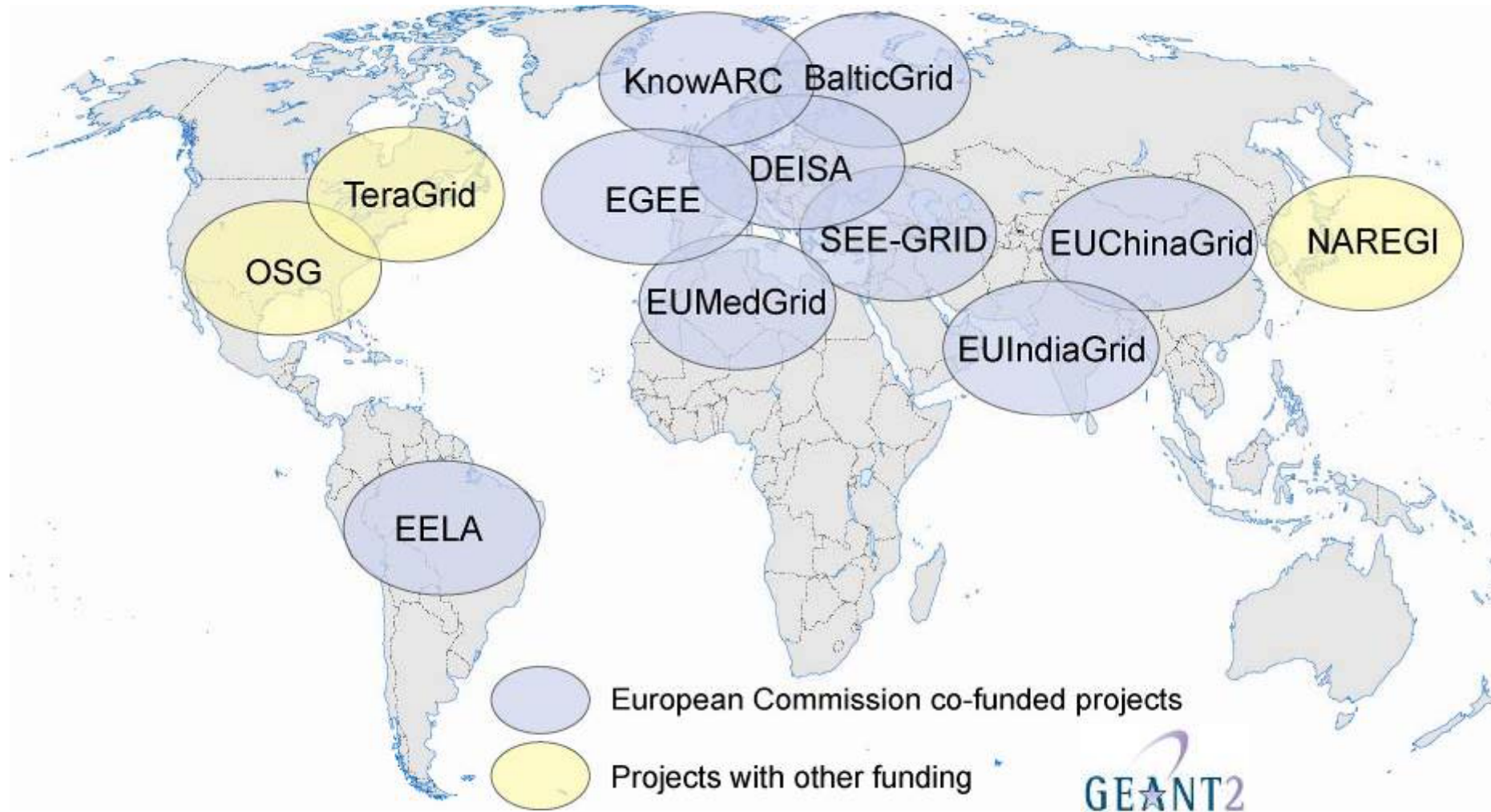
NA: network activities

- supporting VOs

JRA: “joint research activities”

- e.g. hardening middleware





Potential for linking ~80 countries by 2008

Test-beds & Services

Certification testbeds (SA3)

Pre-production service

Production service

Infrastructure:

- Physical test-beds & services
- Support organisations & procedures
- Policy groups

Support Structures

Operations Coordination Centre

Regional Operations Centres

Global Grid User Support

EGEE Network Operations Centre (SA2)

Operational Security Coordination Team

Security & Policy Groups

Joint Security Policy Group

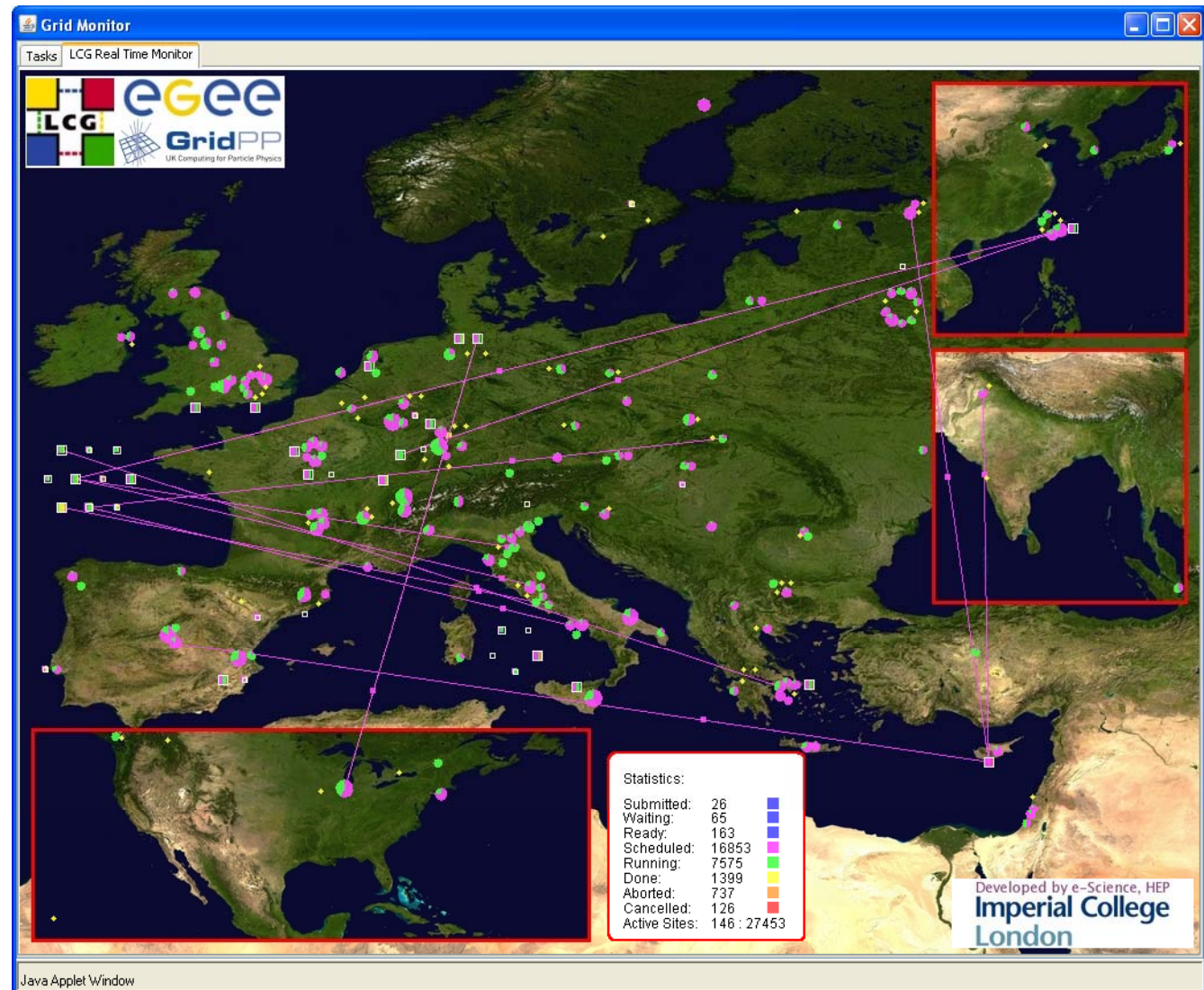
EuGridPMA (& IGTF)

Grid Security Vulnerability Group

Operations Advisory Group (+NA4)

Real Time Monitor

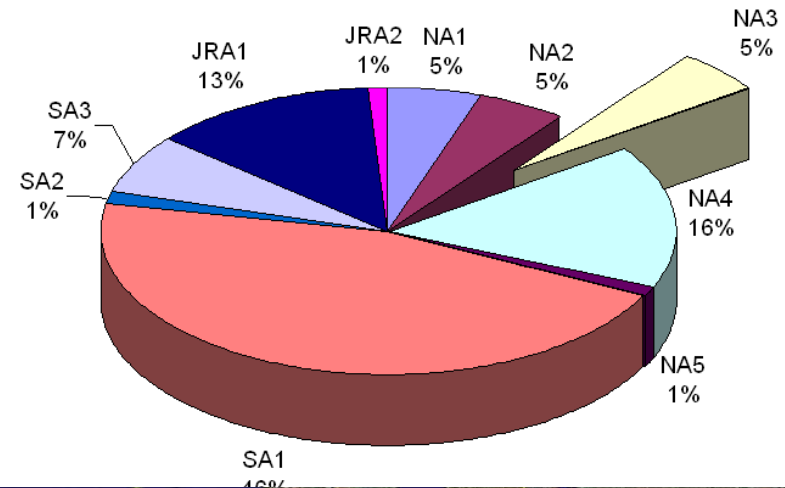
- Java tool
- Displays jobs running (submitted through RBs)
- Shows jobs moving around world map in real time, along with changes in status



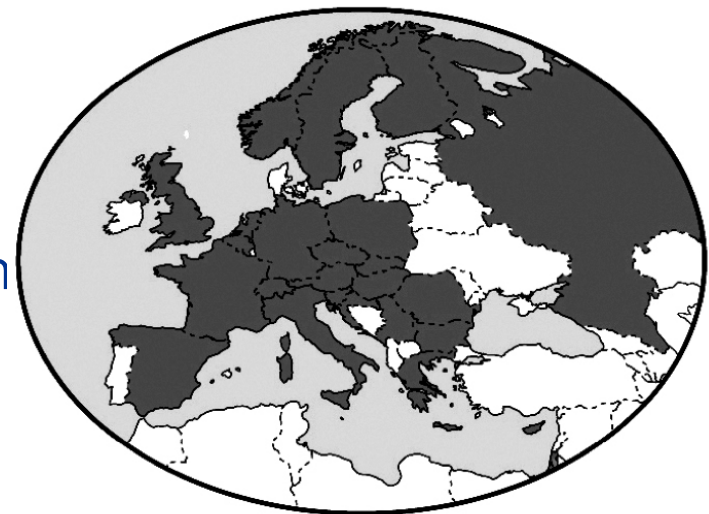
<http://gridportal.hep.ph.ic.ac.uk/rtm/>

(snapshot 16 January 2007)

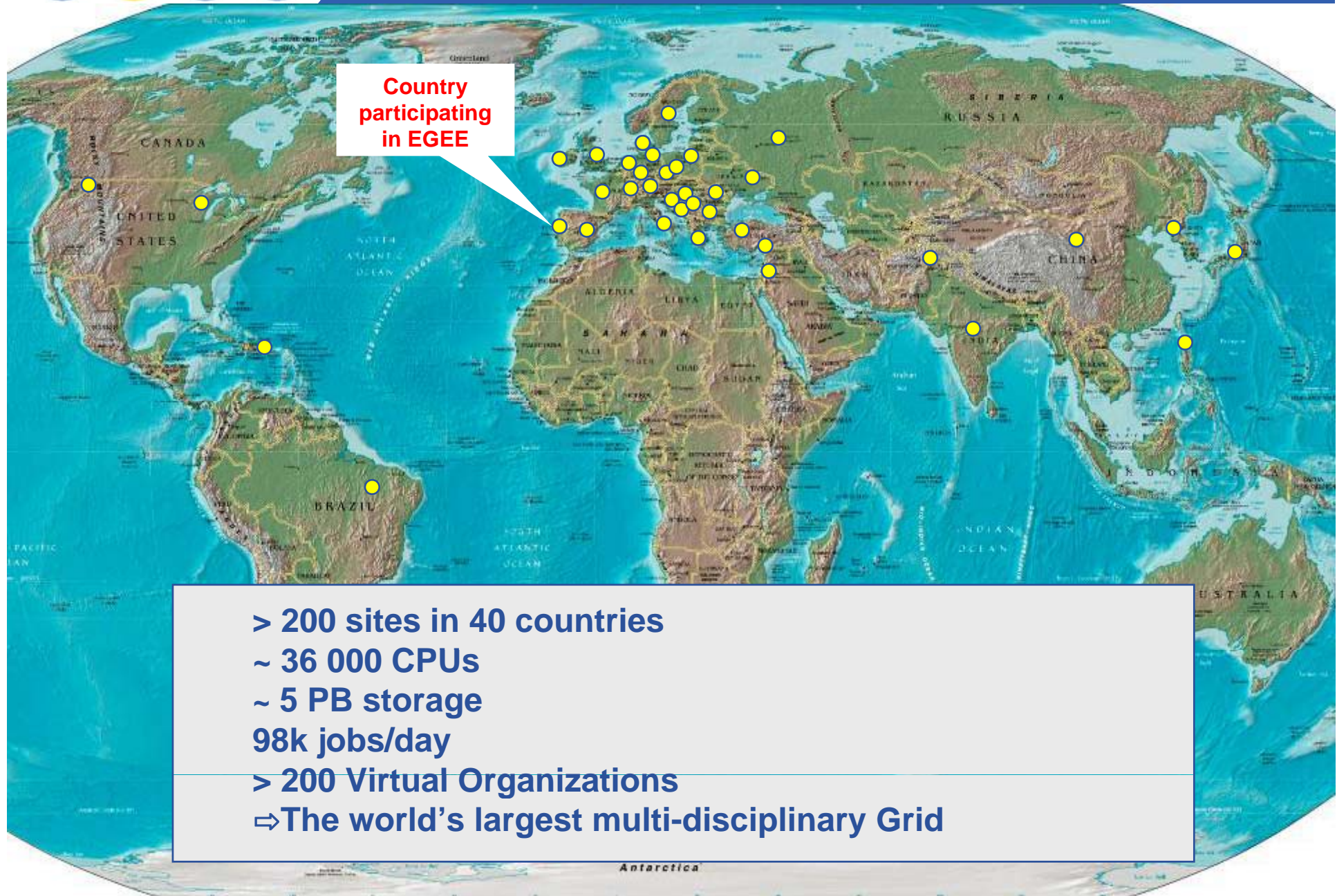
- Expand portfolio of training materials & courses
- Train a wide variety of EGEE users (internal/external)
- Develop effective mechanisms for training end-users of the EGEE infrastructure
- Enhance e-Learning structure and provision of t-Infrastructure
- Validate cohorts of trainers & compile directory
- Collaborate in cross-activity initiatives
- <http://www.egee.nesc.ac.uk/>



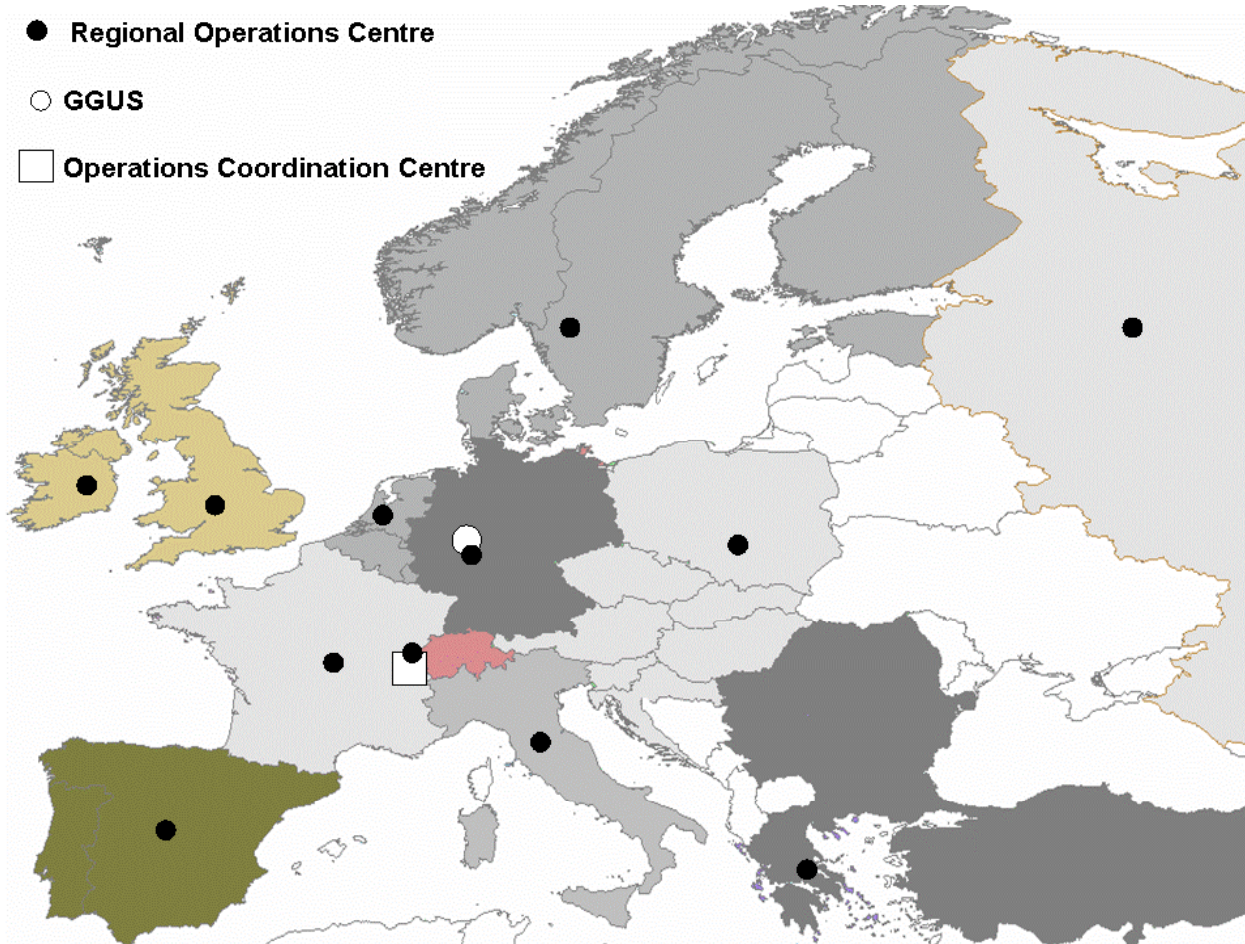
- **Application Identification and Support (NA4)**
 - 25 countries, 40 partners, 280+ participants, 1000s of users
- **Support the large and diverse EGEE user community:**
 - **Promote dialog:** Users' Forums & EGEE Conferences
 - **Technical Aid:** Porting code, procedural issues
 - **Liaison:** Software and operational requirements
- **Need active participation:**
 - **Feedback:** Infrastructure, configuration, and middleware
 - **Resources:** Hardware and human
- <http://egeena4.lal.in2p3.fr>



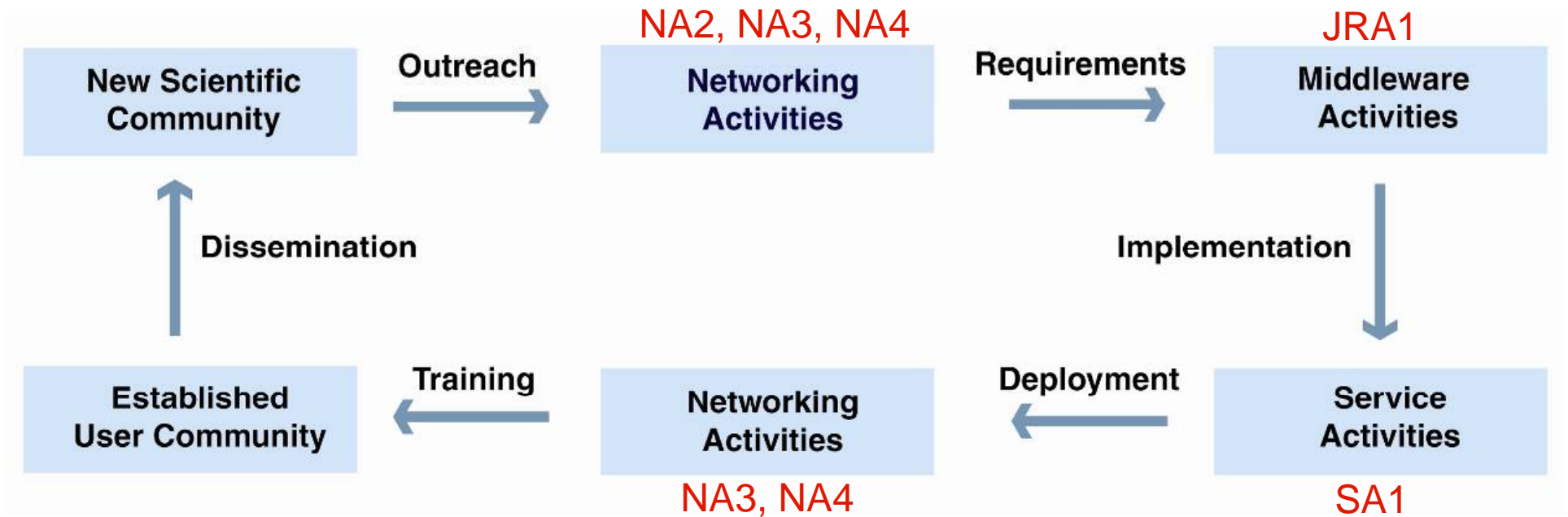
<i>Name</i>	<i>Description</i>
BalticGrid	EGEE extension to Estonia, Latvia, Lithuania
EELA	EGEE extension to Brazil, Chile, Cuba, Mexico, Argentina
EUChinaGRID	EGEE extension to China
EUMedGRID	EGEE extension to Malta, Algeria, Morocco, Egypt, Syria, Tunisia, Turkey
ISSeG	Site security
eIRGSP	Policies
ETICS	Repository, Testing
OMII-Europe	to provide key software components for building e-infrastructures;
BELIEF	Digital Library of Grid documentation, organisation of workshops, conferences
BIOINFOGRID	Biomedical
Health-e-Child	Biomedical – Integration of heterogeneous biomedical information for improved healthcare
ICEAGE	International Collaboration to Extend and Advance Grid Education



- Regional Operations Centre
- GGUS
- Operations Coordination Centre

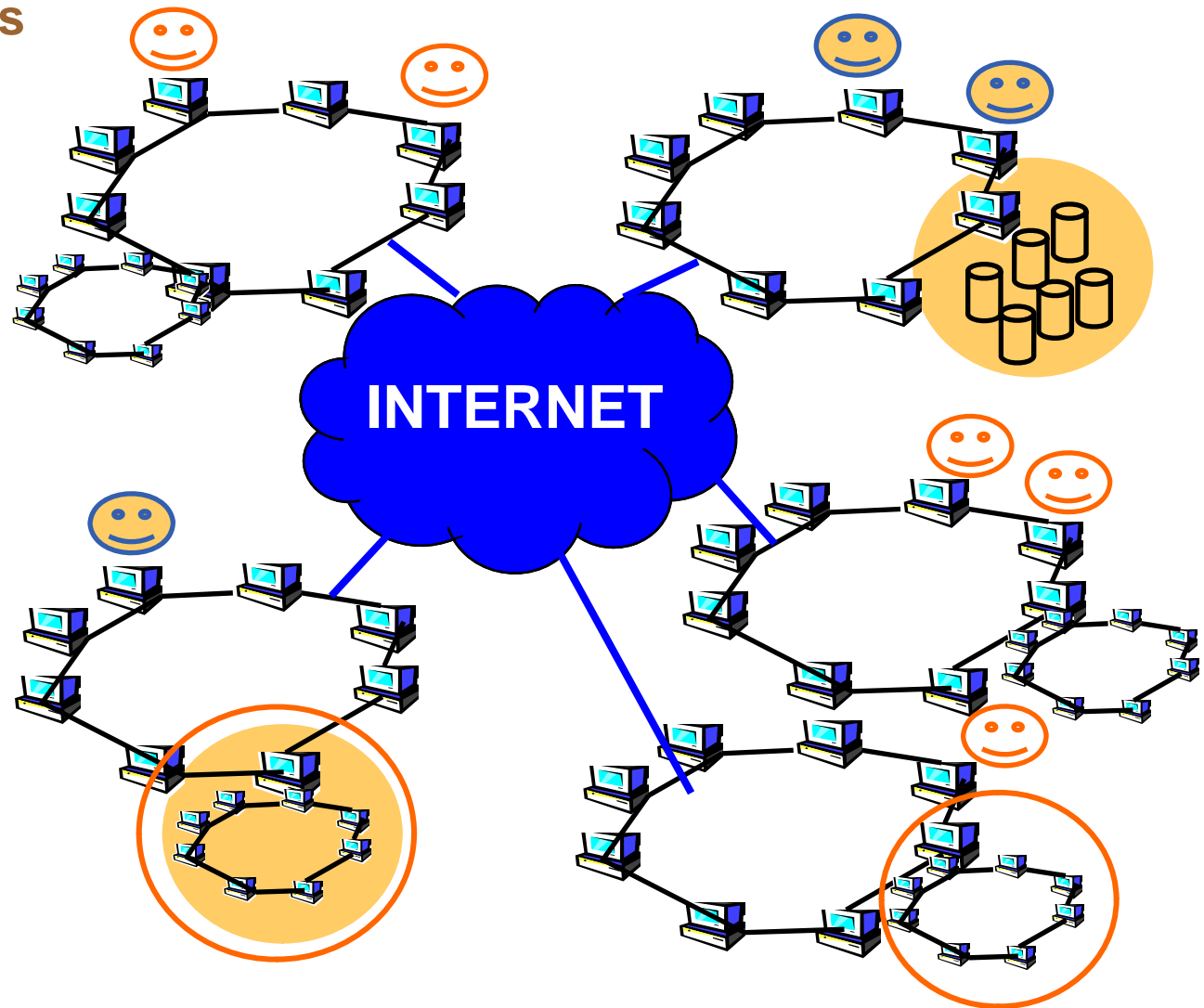


- **Operations Coordination Centre (OCC)**
 - management, oversight of all operational and support activities
- **Regional Operations Centres (ROC)**
 - providing the core of the support infrastructure, each supporting a number of resource centres within its region
 - **Grid Operator on Duty**
- **Resource centres**
 - providing resources (computing, storage, network, etc.);
- **Grid User Support (GGUS)**
 - At FZK, coordination and management of user support, single point of contact for users



Building effective user communities

- 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

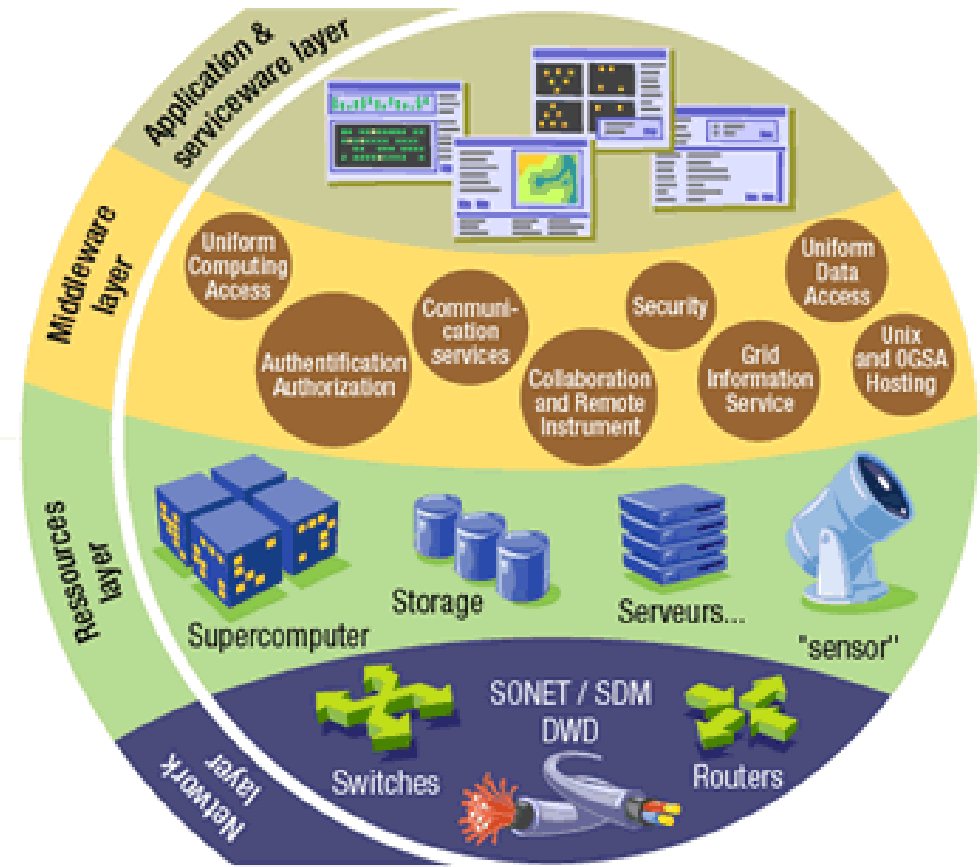


- **What is EGEE?**
 - The project
 - The infrastructure
- **gLite middleware**
- **EGEE applications**
- **Sources of further information**

- 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



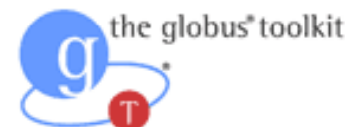
- **When using a PC or workstation you**
 - Login with a username and password (“Authentication”)
 - Use rights given to you (“Authorisation”)
 - Run jobs
 - Manage files: create them, read/write, list directories
- **Components are linked by a bus**
- **Operating system**
- **One admin. domain**
- **When using a Grid you**
 - Login with digital credentials – single sign-on (“Authentication”)
 - Use rights given you (“Authorisation”)
 - Run jobs
 - Manage files: create them, read/write, list directories
- **Services are linked by the Internet**
- **Middleware**
- **Many admin. domains**

- **gLite 3.0**

⇒ Merger of LCG 2.7 and GLite 1.5



- Exploit **experience and existing components** from VDT (Condor, Globus), EDG/LCG, and others
- Develop a **lightweight stack of generic middleware** useful to EGEE applications (HEP and Biomedics are pilot applications).
 - Should eventually deploy dynamically (e.g. as a globus job)
 - Pluggable components – cater for different implementations
- Focus is on providing a stable and usable infrastructure



User Interface



create
credential

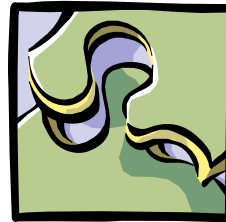


Authorization Service
(VO Management Service)

Retrieve
status & output

Submit job

Resource Broker

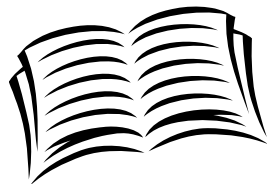


query

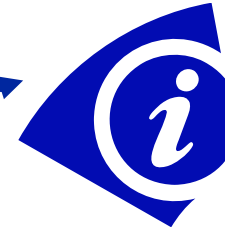
Retrieve
status & output

Submit job

File and Replica Catalog

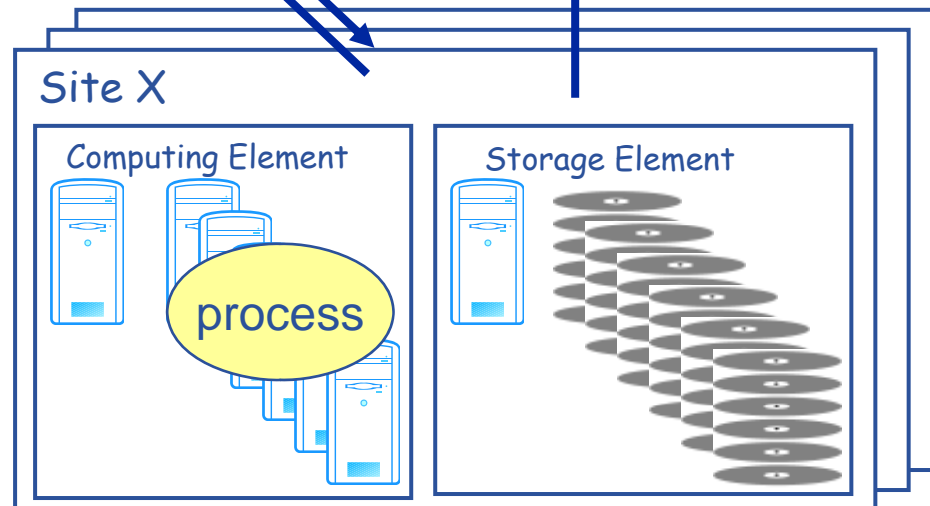


Information System



query

publish
state





User Interface (UI): The place where users logon to the Grid



Resource Broker (RB) (Workload Management System (WMS):
Matches the user requirements with the available resources on the Grid



Information System: Characteristics and status of CE and SE



File and replica catalog: Location of grid files and grid file replicas



Computing Element (CE): A batch queue on a site's computers where the user's job is executed



Storage Element (SE): provides (large-scale) storage for files



User Interface (UI): The place where users logon to the Grid



Resource Broker (RB) (Workload Management System (WMS):
Matches the user requirements with the available resources on the Grid



Int



File



Co



Storage Element (SE): provides (large-scale) storage for files

Grid Security Infrastructure: Single logon with security and trust

**All built upon
authorisation,
authentication,
security**



SE

eplicas

ere

Who provides the resources?!

<u>Service</u>	<u>Provider</u>	<u>Note</u>
<u>User interface</u>	User / institute / VO	Computer with client software
<u>Resource Broker (WMS)</u>	VOs - EGEE does not fund RBs	
<u>Information System</u>	Grid operations - EGEE funded effort	
<u>File and replica catalog</u>	VOs - EGEE does not fund catalogs	
<u>Computing Element (CE)</u>	VOs - EGEE does not fund CEs	Scalability requires that VOs provide resources to match average need
<u>Storage Element (SE)</u>	VOs - EGEE does not fund SEs	
<u>External services</u>	User / institute / VO	To extend the capabilities of the core infrastructure

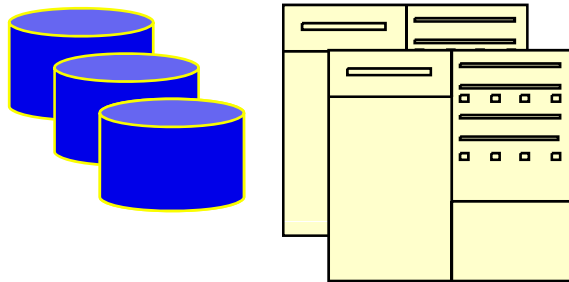
Application

Application toolkits

Command line & APIs

Higher-level gLite services
(WMS,...)

Basic gLite services:
CE, SE, info, security



Where computer science meets the application communities!

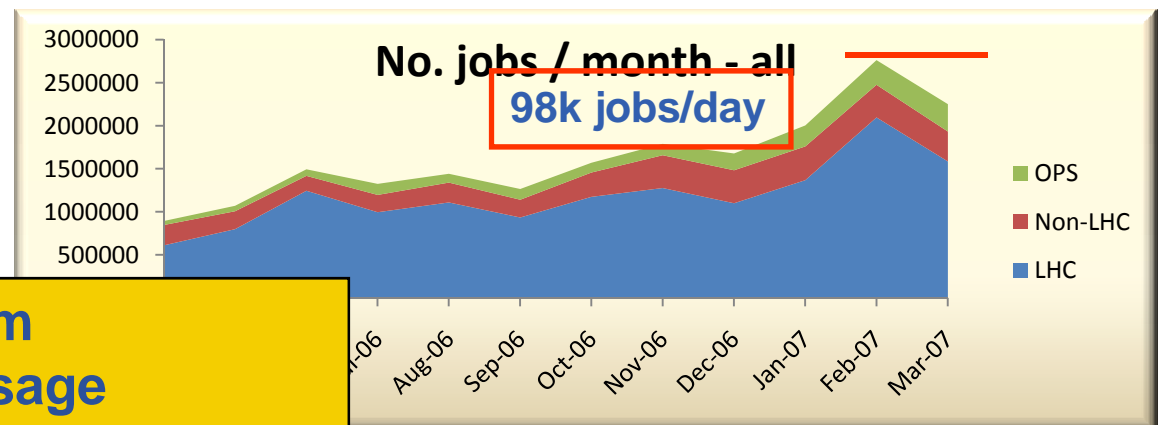
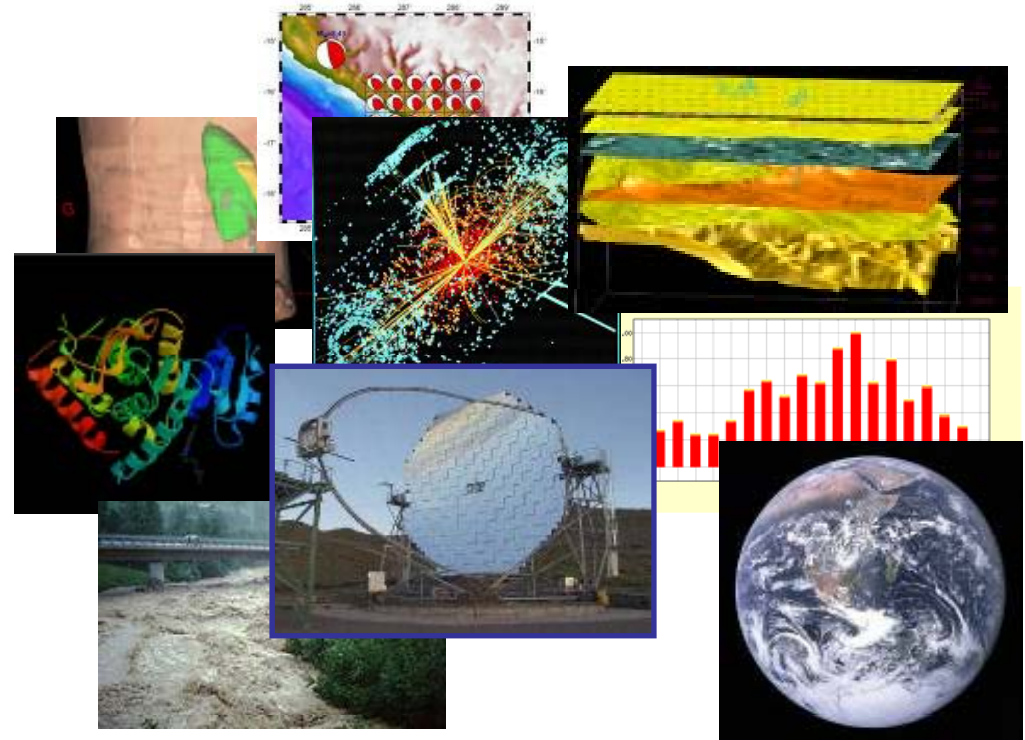
- Portals – P-GRADE, GENIUS
- Job management – GANGA
- Alternative WMS – GridWAY

Production infrastructure contains these services

- Basic services: Must be complete and robust; Should not assume the use of Higher-Level Grid Services
- High level services: help the users building their computing infrastructure but should not be mandatory

- **What is EGEE?**
 - The project
 - The infrastructure
- **gLite middleware**
- **EGEE applications**
- **Sources of further information**

- >200 VOs from several scientific domains
 - Astronomy & Astrophysics
 - Civil Protection
 - Computational Chemistry
 - Comp. Fluid Dynamics
 - Computer Science/Tools
 - Condensed Matter Physics
 - Earth Sciences
 - Fusion
 - High Energy Physics
 - Life Sciences
- Further applications under evaluation



Applications have moved from testing to routine and daily usage

~80-90% efficiency

- **Simulation**
 - Large number of similar, independent jobs – parameter study
- **Bulk Processing**
 - Widely-distributed input data, Sophisticated data management
- **Workflow**
 - Complex dependencies between individual tasks
- **Legacy Applications**
 - Licenses: control access to software on the grid
 - No recompilation \Rightarrow no direct use of grid APIs
- **Parallel Jobs**
 - Many CPUs needed simultaneously, Use of MPI libraries
 - *Limited support in gLite*: MPI configuration is not uniform
- **Responsive Apps.**
 - Short response time
 - *No real support in gLite* \rightarrow Interactive Grid FP6 project

- **EGEE**
 - <http://www.eu-egee.org/>
- **gLite middleware**
 - <http://www.glite.org>
- **gLite manuals, documentation**
 - <http://glite.web.cern.ch/glite/documentation/>
(gLite user guide)
- **Recommended External Software Packages for Egee CommuniTies (RESPECT)**
 - <http://egeena4.lal.in2p3.fr/>

- **EGEE is running the largest multi-VO grid in the world!**
 - Creating the “grid layer” in e-Infrastructure for research, public service and industry
- **Key concepts for EGEE**
 - Sustainability – planning for the long-term
 - Production quality
 - User support
- **EGEE’s middleware: gLite. Current version 3.0**
 - Basic middleware services
 - High level middleware services
- **External software to foster uptake of technology**