



European Grid Initiative



La Grid da EGEE a EGI e da infn-grid a IGI

A.Ghiselli
INFN-CNAF
Bologna

Calcolo Scientifico nella Fisica Italiana
Rimini 27-30 maggio 2008

What is a Grid

*In 1998, I. Foster and C. Kesselman wrote in **The Grid:***

Blueprint for a New Computing Infrastructure:

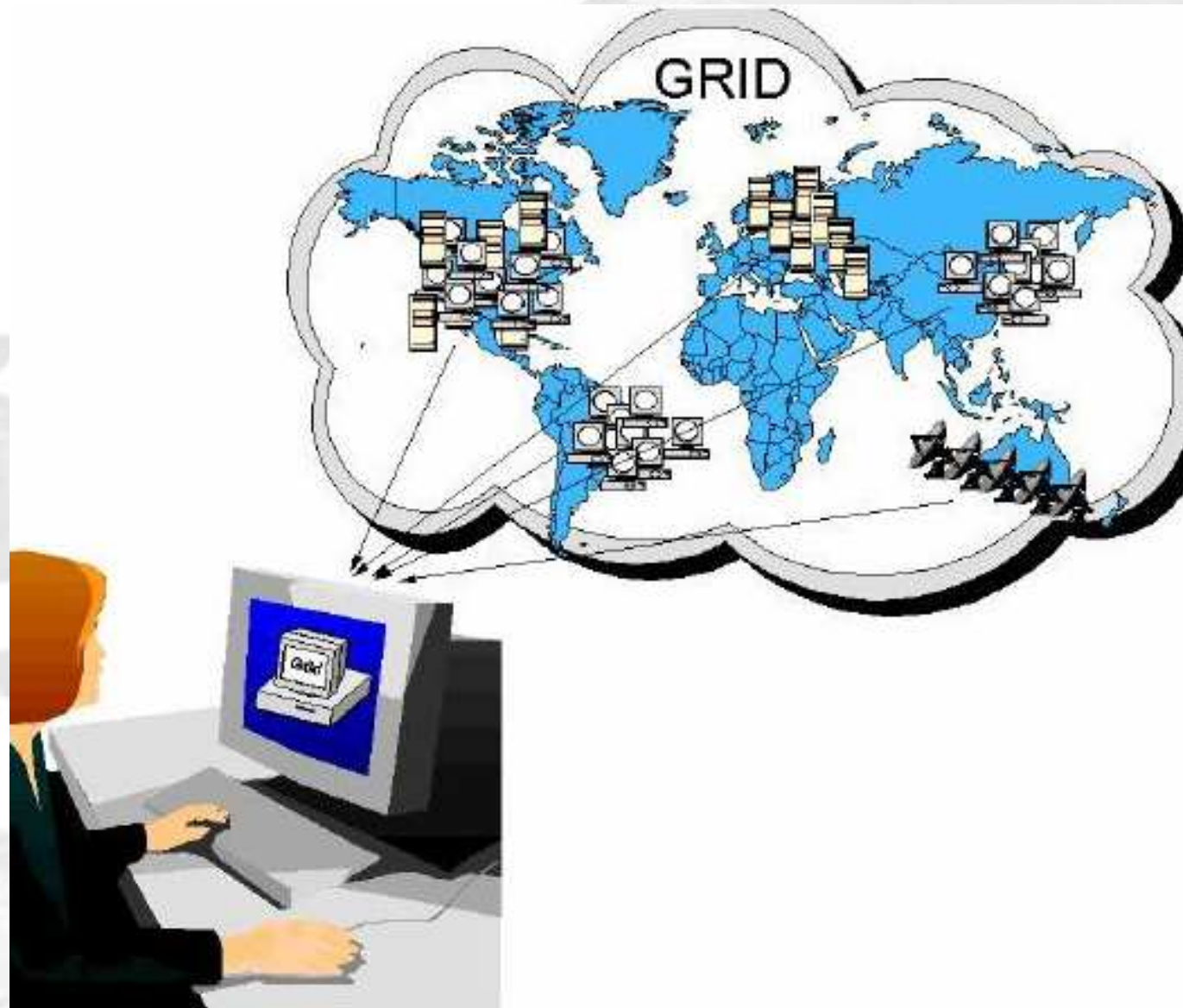
“A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive access to highend computational capabilities.”

*In 2002, I. Foster wrote in **What is the Grid? A Three***

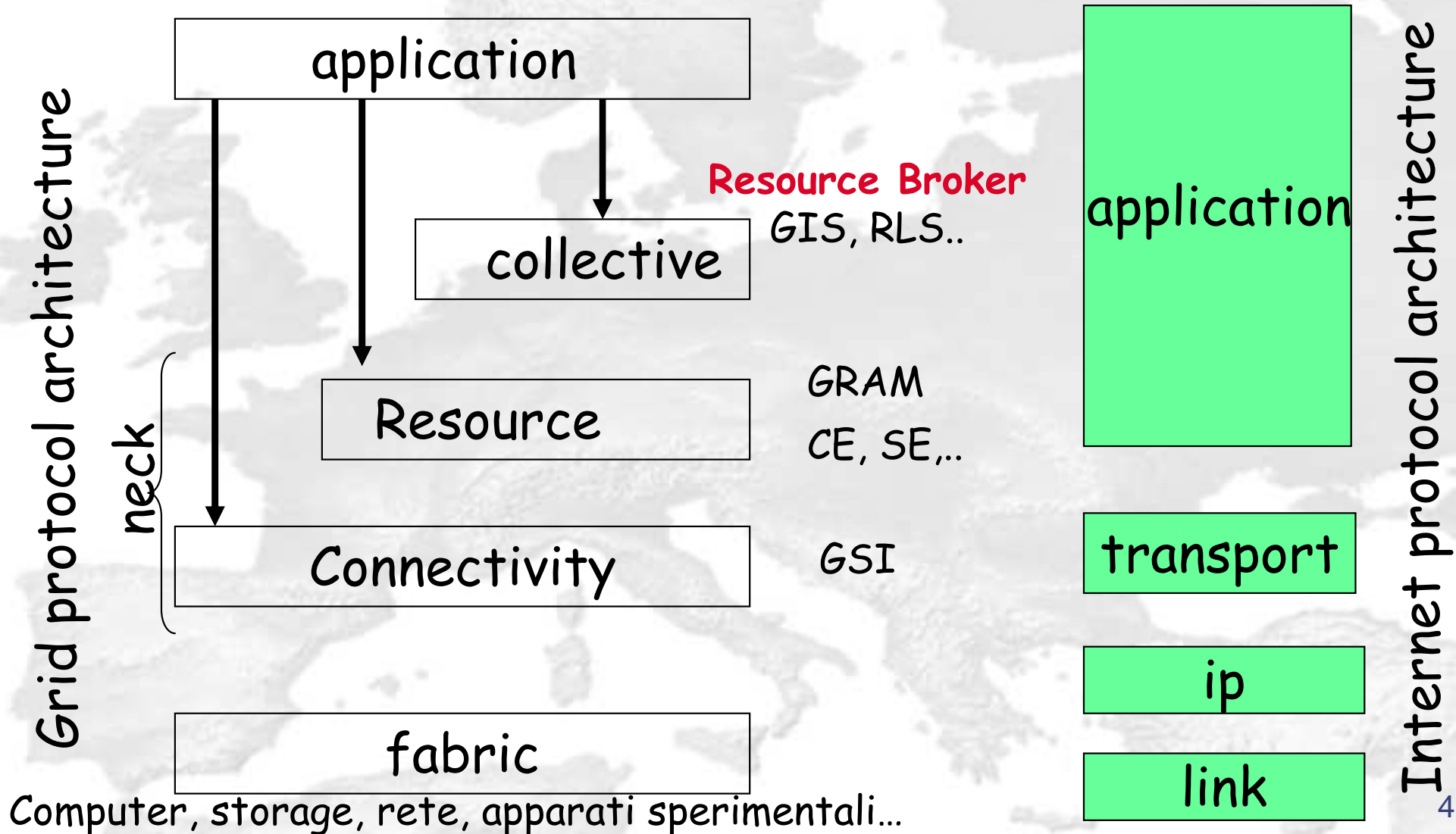
Point Checklist:

“... The essence of the [definition] can be captured in a simple checklist,... a Grid is a system that:

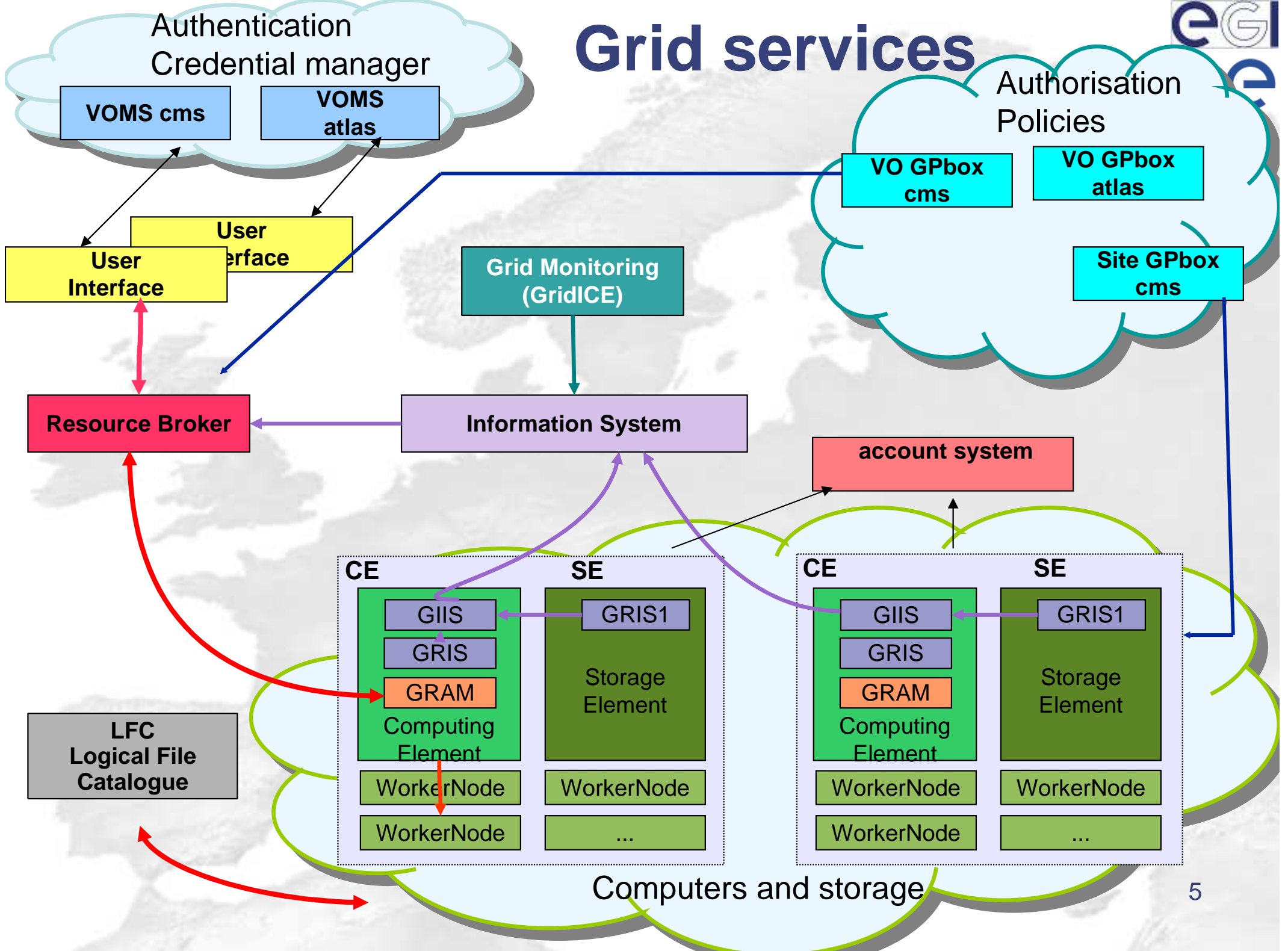
- 1. coordinates resources that are not subject to centralized control...*
- 2. ... using standard, open, general purpose protocols and interfaces...*
- 3. ... to deliver nontrivial qualities of service.*



Architettura e servizi di Grid e confronto con l'architettura di Internet

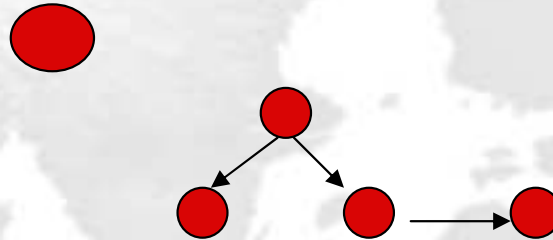


Grid services

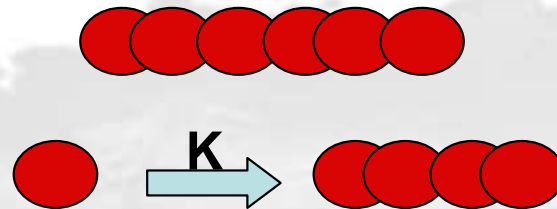


Supported Job Types

- Batch-like
- Dag workflow

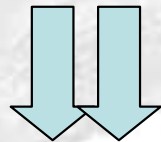


- Collection
- Parametric

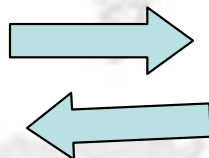


Compound

- MPI



- Interactive



EU e-Infrastructure: a success story



- Hundreds of M€ of EC and National funds (2001-08) have got a set of top level robust **Open Source Grid Middleware Services** (e.g.the gLite release) providing a solid foundation for the Researchers activities
- Deployed in the world largest **Grid e-Infrastructure** (= Internet + Grid): the one of EGEE (Enabling Grid and E-science ->EGEE II -> EGEE III),
- **Real Services, not prototypes**, daily offered by federations of well managed resource centers to thousands of users of several research communities
 - gLite today: 150K Jobs/day succesfull completed in 250 centers,
 - up to >10 GByte/sec of aggregated sustained access to scientific data
 - transparent access to ~ 20 PB of Storage
 - International Standards (Web Services, W3C, OGF, OASIS), high level of security (PKI, X509), interoperability of different implementations (ARC UNICORE)
- EU e-Infrastructure is ahead in this domain
- Members States and EC are now guaranteeing the **long term sustainability** of the baseline Grid MW and e-Infrastructures with the
 - **European Grid Initiative EGI** (FP7 Project: EGI-Design Study) and the **National Grid Initiatives (NGIs)**

Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
...

Scheduled = 21539
Running = 25374

>250 sites
48 countries
>50,000 CPUs
>20 PetaBytes
>10,000 users
>150 VOs
>150,000 jobs/day

21:13:50 UTC

Users and resources distribution



EGEE – What do we deliver?



- **Infrastructure operation**
 - Sites distributed across many countries
 - Large quantity of CPUs and storage
 - Continuous monitoring of grid services & automated site configuration/management
 - Support multiple Virtual Organisations from diverse research disciplines
- **Middleware**
 - Production quality middleware distributed under business friendly open source licence
 - Implements a service-oriented architecture that virtualises resources
 - Adheres to recommendations on web service interoperability and evolving towards emerging standards
- **User Support** - *Managed process from first contact through to production usage*
 - Training
 - Expertise in grid-enabling applications
 - Online helpdesk
 - Networking events (User Forum, Conferences etc.)



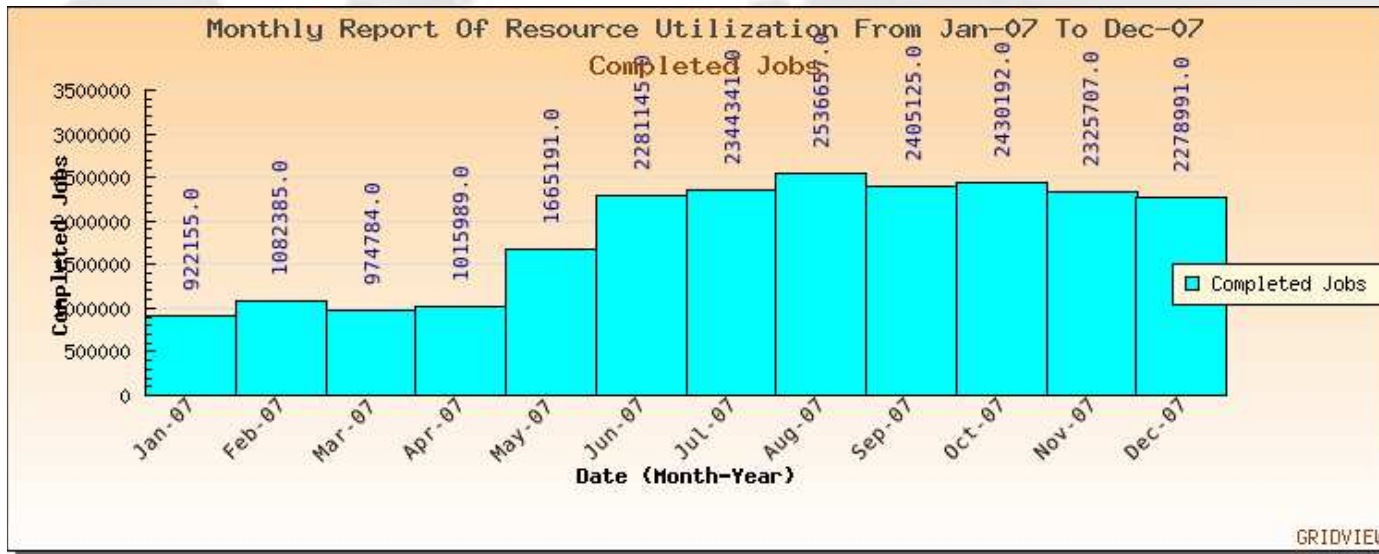
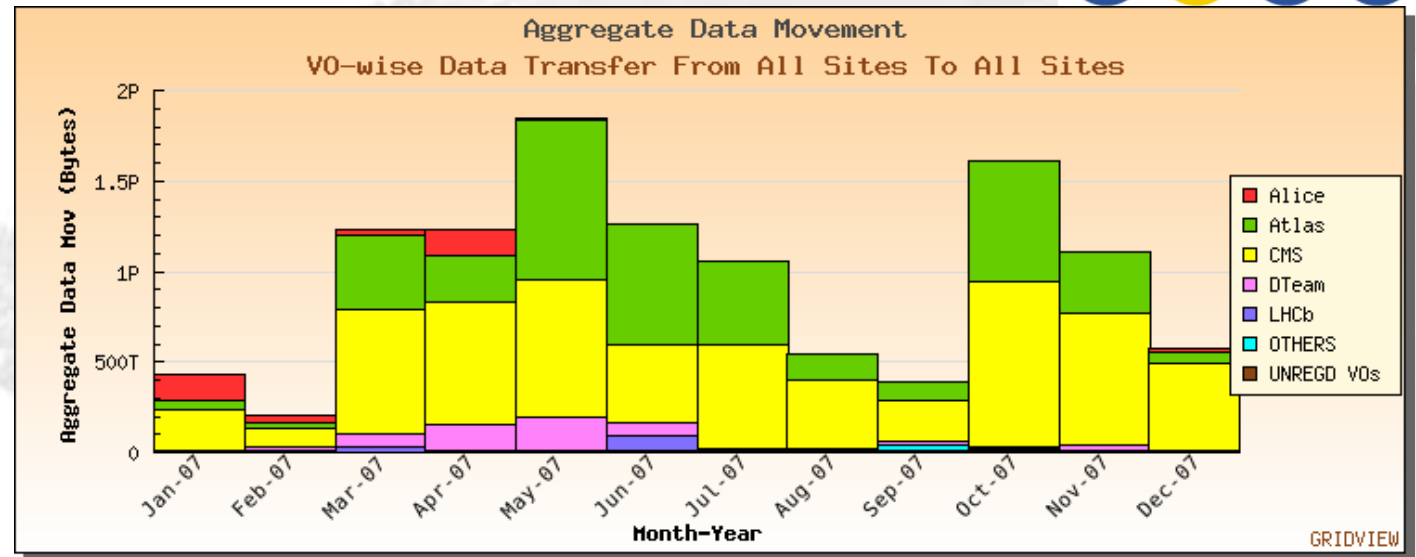
Bob Jones
EGEE project director

Types of applications

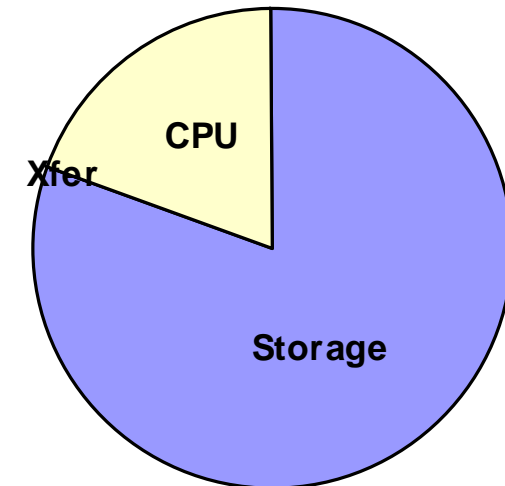
- **Simulation**
 - LHC Monte Carlo simulations; Fusion; WISDOM
 - Jobs needing significant processing power; Large number of independent jobs; limited input data; significant output data
- **Bulk Processing**
 - HEP ; Processing of satellite data
 - Distributed input data; Large amount of input and output data; Job management (WMS); Metadata services; complex data structures
- **Parallel Jobs**
 - Climate models, computational chemistry
 - Large number of independent but communicating jobs; Need for simultaneous access to large number of CPUs; MPI libraries
- **Short-response delays**
 - Prototyping new applications; grid Monitoring grid; Interactivity
 - Limited input & output data; processing needs but fast response and quality of service
- **Workflow**
 - Medical imaging; flood analysis
 - Complex analysis algorithms; complex dependencies between jobs
- **Commercial Applications**
 - Non-open source software; Geocluster (seismic platform); FlexX (molecular docking); Matlab, Mathematics; Idl, ...
 - License server associated to an application deployment model

EGEE workload in 2007

Data:
25Pb stored
11Pb transferred



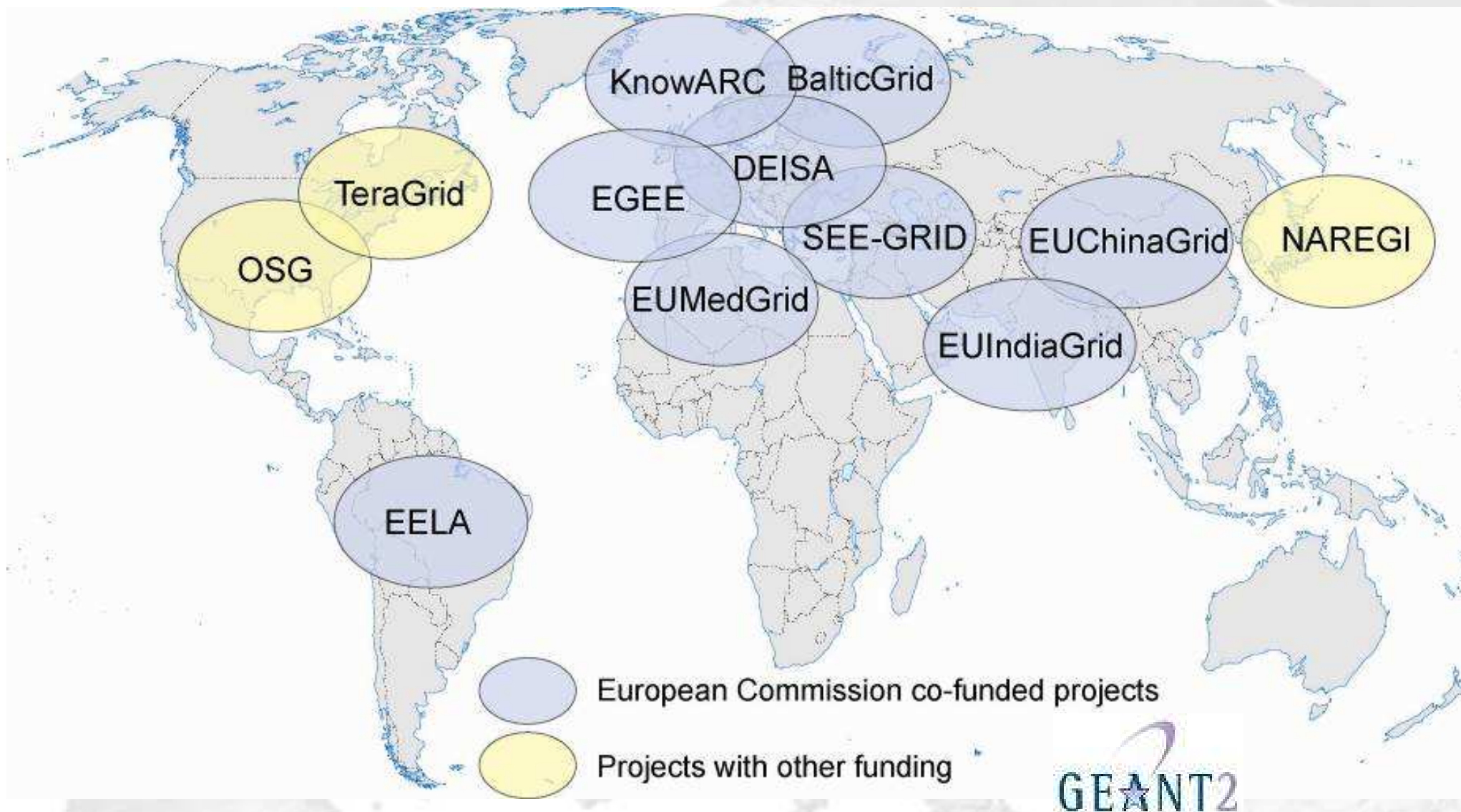
CPU: 114 Million hours



Estimated cost if performed with Amazon's EC2 and S3: £ 29,977,003

http://gridview.cern.ch/GRIDVIEW/same_index.php <http://calculator.s3.amazonaws.com/calc5.html?> 17/05/08 \$58688679.08

Collaborating e-Infrastructures

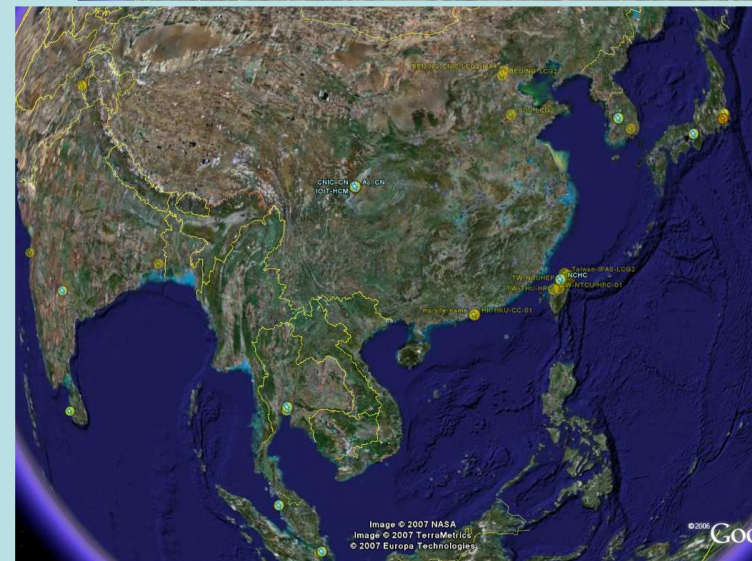


Potential for linking ~80 countries by 2008

Worldwide Grid Infrastructures



APAC
DEISA
EGEE
Naregi
NDGF
NGS
OSG
Pragma
Teragrid



EGEE-III Vision and mission

- The main vision of EGEE-III is to make a strong move towards a sustainable world-wide production quality Grid infrastructure by appropriate technical and organisational decisions.
- This must be capable of providing services to a rapidly increasing number of application areas, and make Grid technology easily accessible and usable for these communities.

INFN-Grid – goals (started at 2000)



1. To promote computational grid technologies research & development:
Middleware
 1. Through european and international projects
 1. **DataGrid (RB), DataTAG (VOMS, GLUESCHEMA, GridICE), EGEE I,II, OMII-Europe (interoperability and standardization), CoreGRID (R&D)**
 2. Internal R&D activities
2. To implement the INFN grid **infrastructure**
3. To set up the national Grid **Infrastructure** for the national research community
 1. **FIRB: Grid.it (Grid di produzione, G-DSE, StoRM, G-Pbox, GridICE II)**
4. To participate to the implementation of the global Grid infrastructure for the LHC community
 1. **LCG: Tier1 and n*Tier2**
5. To set up the **eInfrastructure** for the European Research Area
 1. **EU FP6-7: EGEE (European Production Infrastructure), EGI (Grid sustainability based on National Grids)**

Grid nazionale in EGEE oggi

IT federation

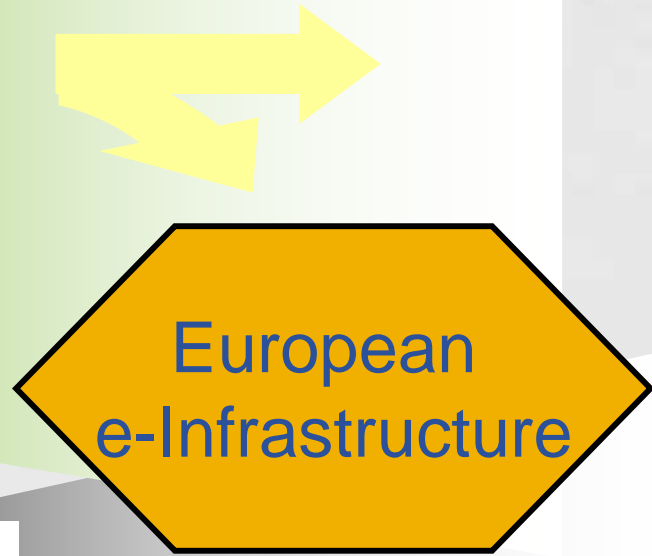
- 1 country
- 5 Partners:
 - INFN
 - ENEA
 - SPACI-UNILE
 - SPACI-UNINA
 - SPACI-UNICAL
- 5 more RCs:
 - Istituto Tecnologie Biomediche – CNR/BARI (LIBI Project)
 - PERUGIA University
 - Istituto Linguistica Computazionale CNR-PISA
 - Scuola Normale Superiore – PISA
 - ESA-ESRIN



EGI: The future evolution

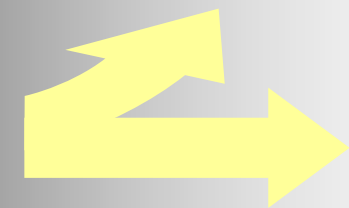
National

Logos of national grid services and infrastructure projects, including GridPP, UNICORE FORUM, Grid.it project, e-Toile, NGS National Grid Service, BEgrid, INFN GRID, NORDUGRID, ACI GRID, HELLAS GRID Task Force, and DutchGrid.



Global

Logos of global grid services and infrastructure projects, including DataTAG, Data GRID, SEE-GRID, eGee Enabling Grids for E-science, EUMED Grid, omii europe, Diligent, European Infrastructure for Supercomputing Applications, ISS@G, BEingRID, BioinfoGRID, and Health-e-Child.



Testbeds

Routine Usage

Utility Service



Why Sustainability TODAY?

- **Dependency:** some **application domains depend** on grids already today
- **Protection of Investment:** **Investment** in grids, both from funding organizations and from users, need to be **protected**
- **Perspective:** Today's grid users are **grid enthusiasts**, tomorrows grid users ask for a **longer term perspective**

EGL Design Study

Project proposal:

- submitted to FP7-INFRASTRUCTURES-2007-1,
1.2.1 Design Studies

Goal:

- Conceptual setup and operation of a new organizational model of a sustainable pan-European grid infrastructure

Consortium: 9 Partners → EGL Preparation Team

NGI Representatives (38) → EGL Advisory Board

Person months: ~300

Duration: 1 Sept 2007 – 30 Nov 2009 (27 Months)

EGI Objectives

- Ensure the **long-term sustainability** of the European e-infrastructure
- Coordinate the integration and interaction between **National Grid Infrastructures**
- Operate the European level of the **production Grid infrastructure** for a wide range of scientific disciplines to link National Grid Infrastructures

European Grid Initiative (EGI) = Sustainability of the e-Infrastructure = EGI.Org + National Grid Initiatives (NGI)



- Examples of **National Grid** Initiatives (NGIs) :

- Austrian Grid Initiative
- Belgium: BEgrid
- DutchGrid
- France: Institute des Grilles
- Germany: D-Grid; Unicore
- Greece: HellasGrid
- Grid Ireland
- Italy: IGI
- NDGF (Nord)
- Portuguese Grid
- Swiss Grid
- UK e-Science: National Grid Service; OMII; GridPP



- Within ~1 year EGI and NGI should become stable organizations as NRENS and DANTE

Characteristics of NGIs

Each NGI

- ... should be a recognized national body with a **single point-of-contact**
- ... should mobilise national funding and resources
- ... should operate the national e-Infrastructure
- ... should supports user communities (application independent, and open to new user communities and resource providers)
- ... should contribute and adhere to international standards and policies

Responsibilities between NGIs and EGI are split to be federated and complementary

La NGI Italiana :IGI

- Italian Grid Infrastructure (IGI) e'
 - una EU Joint Research Unit
 - Costituita con un MoU firmato da tutti i partner in dicembre 2007
 - Aperto al contributo di nuovi partner
 - Riconosciuta e sostenuta con lettera dal MUR
 - Riconosciuta dalla Commissione Europea
 - Interfaccia unica verso l'EU a livello Italiano
 - Embrione per una governance comune dell'e-Infrastruttura Italiana da parte di tutti gli Enti pubblici coinvolti
- Entro 3/2009 deve diventare un'entita' legale

I partners attuali di IGI

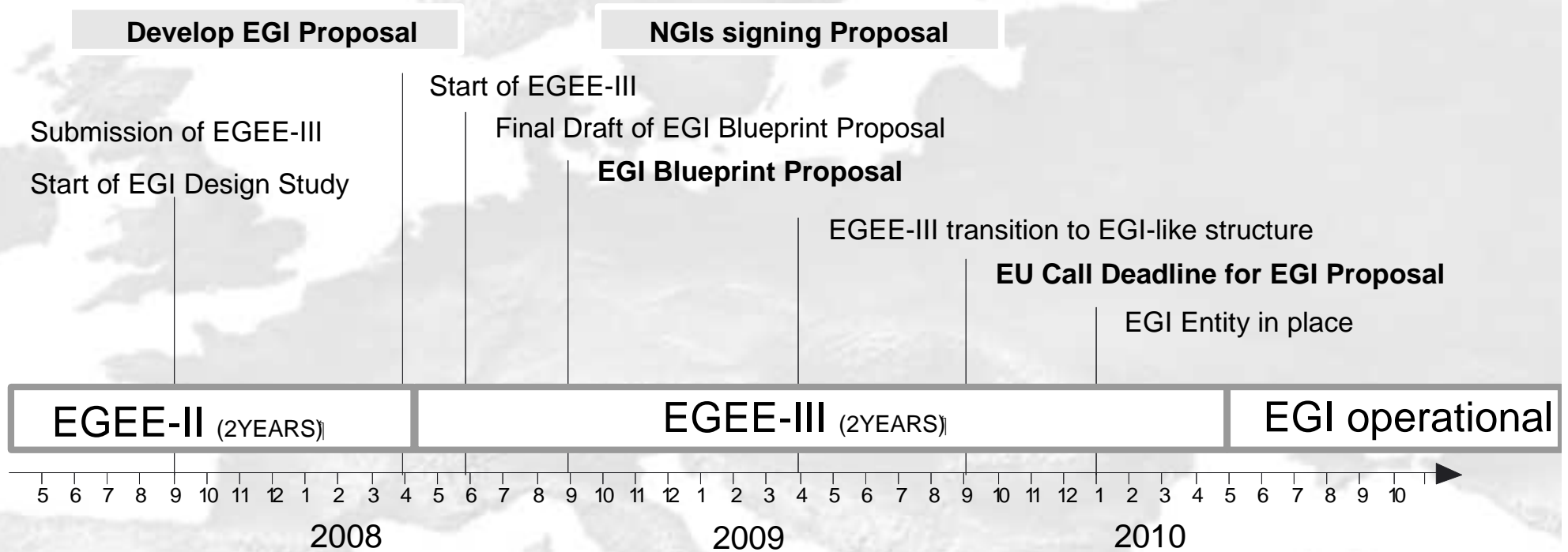
- Istituto Nazionale di Fisica Nucleare (INFN),
- Ente per le Nuove tecnologie, l'Energia e l'Ambiente (ENEA),
- Consiglio Nazionale delle Ricerche (CNR)
- Istituto Nazionale di Astrofisica (INAF),
- Istituto Nazionale di Geofisica e Vulcanologia (INGV),
- Università degli Studi di Napoli Federico II,
- Università degli Studi della Calabria,
- Sincrotrone Trieste S.C.p.A. (ELETTRA),
- Consorzio COMETA,
- Consorzio COSMOLAB,
- Consorzio SPACI,,
- Consortium GARR
- L'Università' di Perugia
- L'Università' del Piemonte Orientale

The supporting EU project : EGI Design Study

- Partners
- Johannes Kepler Universität Linz (GUP)
- Greek Research and Technology Network S.A. (GRNET)
- Istituto Nazionale di Fisica Nucleare (INFN)
- CSC - Scientific Computing Ltd. (CSC)
- CESNET, z.s.p.o. (CESNET)
- European Organization for Nuclear Research (CERN)
- Verein zur Förderung eines Deutschen Forschungsnetzes - DFN-Verein (DFN)
- Science & Technology Facilities Council (STFC)
- Centre National de la Recherche Scientifique (CNRS)
- -> In June-July the first proposal for EGI

EGI_DS Schedule

Duration 27 months:



EGI Blueprint

- EGI Geneva Workshop on Draft EGI Blueprint Presentation will be held on Monday, 30 June 2008 in the CERN



European Grid Initiative



II Middleware in EGI

Main Grid Middleware Stacks

Available Today

- From Europe
 - gLite universe
 - UNICORE
 - ARC
 - OMII-UK
- From US: Globus, Condor/VDT
- From Asia: NAREGI, CrownGrid

Europe plays a key role in middleware development

The Role of EGI for Grid Middleware

- Evolve current middleware stacks towards:
 - Achieving **standard-based interoperability**
 - systems/organizations able to provide/accept services from other systems/organizations and to use the services exchanged to enable them to operate effectively together
 - Adding **new functionalities** based on requirements from:
 - supported VOs
 - potentially new VOs
 - operations

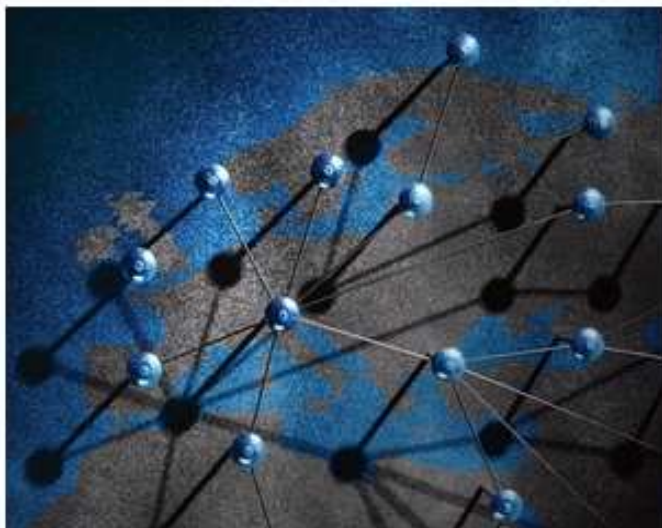
Middleware in EGI

(dal draft Blueprint)

- : **Universal Middleware Toolkit**
 - The main task of the Middleware function in EGI will be to offer as a general service to NGIs the release, maintenance and support of a new **Universal Middleware Toolkit**, similar to the Virtual Data Toolkit (VDT) in the US, which will contain the software components necessary to enable the NGIs to share their resources at the pan-European and International level.
- **Middleware stacks:**
 - The initial stacks will be gLite , ARC and UNICORE which provide quite complete solutions for an e-Infrastructure.
 - Other middleware projects (GridWay, pGrade, AssesGrid, GRIA...) in Europe, funded by the European Commission and by National funds, have provided high level services which complement the basic services provided by the 3 EU stacks.



WELCOME TO EGI



The European Grid Initiative (EGI) Design Study represents an effort to establish a sustainable grid infrastructure in Europe. Driven by the needs and requirements of the research community, it is expected to enable the next leap in research infrastructures, thereby supporting collaborative scientific discoveries in the European Research Area (ERA).

The main foundations of EGI are the National Grid Initiatives (NGIs), which operate the grid infrastructures in each country. EGI will link existing NGIs and will actively support the setup and initiation of new NGIs.

The goal of the EGI Design Study (EGI_DS) is to evaluate use cases for the applicability of a coordinated effort, to identify processes and mechanisms for establishing EGI, to define the structure of a corresponding body, and ultimately to initiate the construction of the EGI organization.

The EGI Design Study is a project funded by the European Commission's 7th Framework Program.

- ⇒ [Objectives of EGI Design Study](#)
- ⇒ [National Grid Initiatives](#)
- ⇒ [Press releases](#)
- ⇒ [Contact us](#)

EGI Webpage

www.eu-egi.org



The gLite Consortium

An opportunity for the EU
Companies and Research
Institutions

Why a gLite Consortium

- Need to preserve and evolve the EGEE middleware in the scenario of the European Grid Initiative (EGI) and National Grid Initiatives
- EGI will support more stacks and foster convergence: gLite, Unicore, ARC...
 - Globus, Unicore and ARC have already created their Consortia
- gLite need to speak with one voice and continue to evolve and support users

The initial membership

- The core of the gLite Open Consortium will be constituted by the same institutions that are involved with the gLite developments, i.e.: Elsig-Datamat (IT), CESNET (CZ), INFN (IT), SWITCH (CH), FOM (NL), UH.HIP, CERN, STFC, UNIMAN.
- The second level will be constituted by the Institution who operate or use the gLite services.
- Additional industrial or business partners should be included in order to create the conditions for an expansion of the Consortium's activities towards the commercial markets.

A new virtuous European Cycle

*EU MW Consortia:
General Services
for Research +
Industry, Commerce
Government*

*Research
e-Infrastructures:
Robustness
Standard adherence
scalability*

*Companies:
Specific
Internet Services
for Society*

Conclusions

- **Grids are all about sharing** – they are a means of working with groups around the world
- **EGEE** operates the world's largest multi-disciplinary grid infrastructure for scientific research
 - In constant and significant production use
- **EGI: need to prepare the long-term**
 - **Today we have a window of opportunity to move grids from research prototypes to permanent production systems (as networks did a few years ago)**
 - EGEE, collaborating projects, national grid initiatives and user communities are working to define a model for a sustainable grid infrastructure that is independent of short project cycles
- **IGI** deve diventare rapidamente un'Entita' legale, definire i servizi che puo' offrire, il modello di finanziamento e giocare un ruolo di primo piano nella costituzione di EGI
- Il nuovo **Consorzio gLite** offre un'opportunita' unica alle imprese per approfittare degli investimenti pubblici fatti fino ad oggi e sviluppare un modello di business basato sul suo supporto