eGee

International e-Infrastructure

Mike Mineter

mjm@nesc.ac.uk

www.eu-egee.org

Information Society

Enabling Grids for E-sciencE
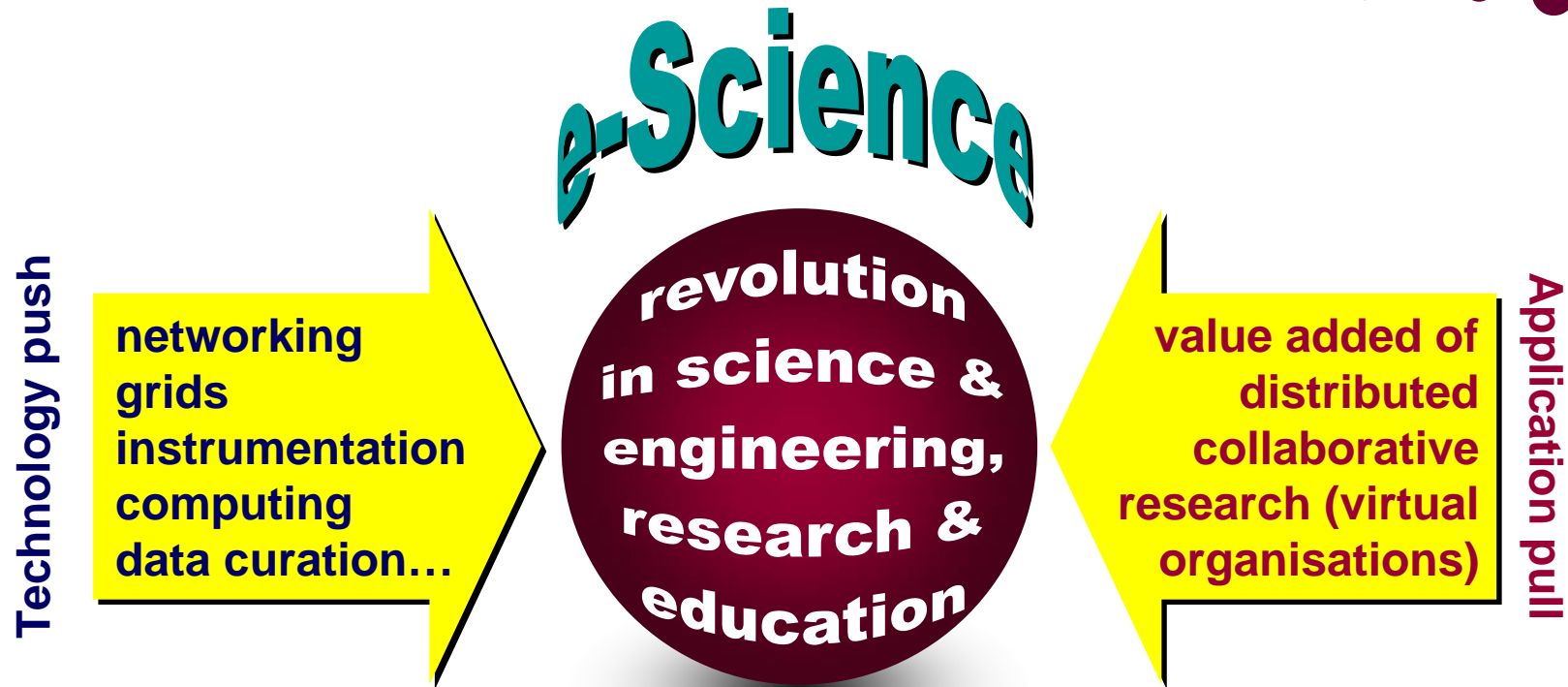
- **The view from the European Commission**

- **GEANT – European network**
- **DEISA – for when High Performance Computing is just not HP enough!**
- **EGEE – establishing grid e-Infrastructure**
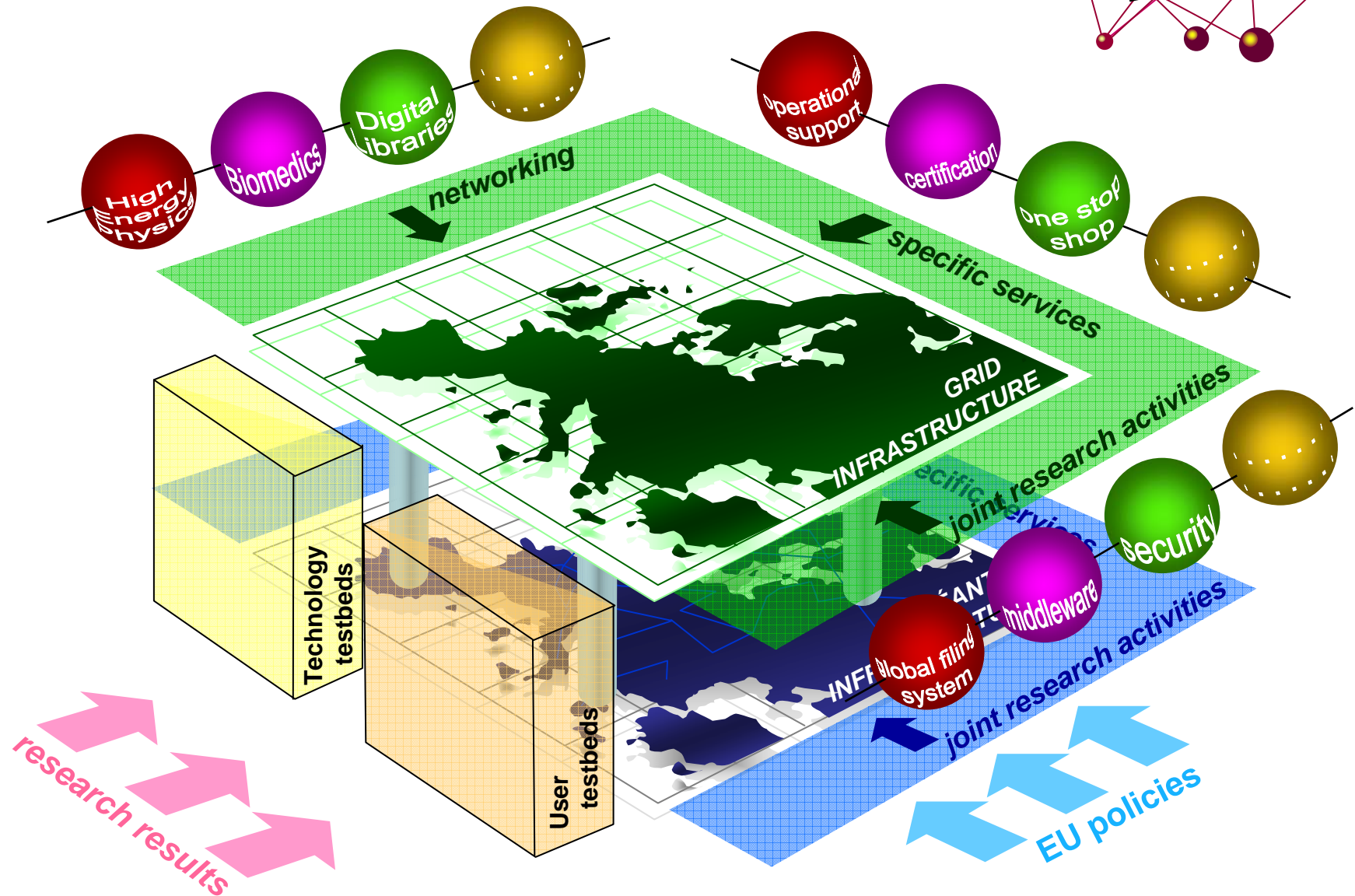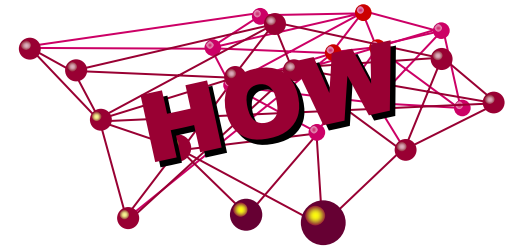 - WHY?!
  - HOW?!
  - With whom?

- **Entering the "knowledge society" from the "industrial society"**

  Industrial society = Transportation Infrastructure

  Knowledge society = Communications infrastructure

- **Lisbon strategy: Research and Innovation will be the most important factors in determining Europe's success through the next decades**

- **THE GOAL: "UNLEASH CREATIVITY"- by investment in**

  – Human skills

  – Infrastructures

- **Demands in growth of e-infrastructure**
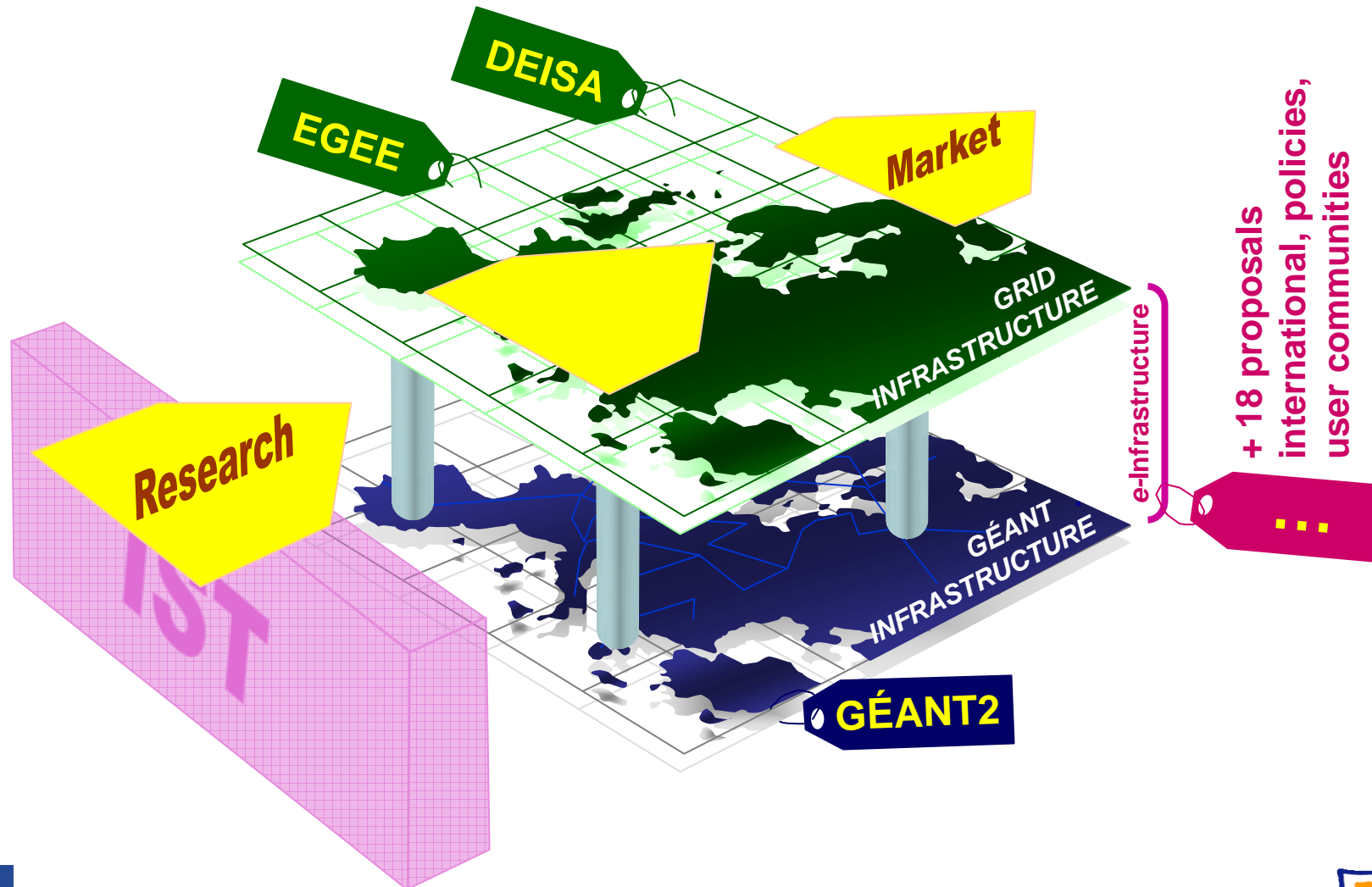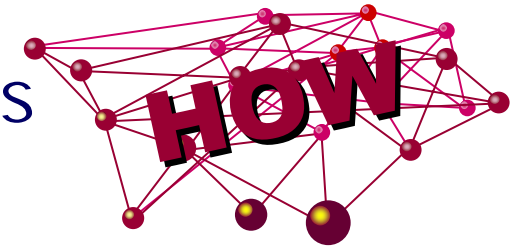
A new way of doing Science

WHY

e-Science

Technology push

networking
grids
instrumentation
computing
data curation…

revolution in science & engineering, research & education

value added of distributed collaborative research (virtual organisations)

Application pull

a new way for all scientists to work on research challenges
that would otherwise be difficult to address

European Commission

Information Society and Media

e-Infrastructure - Implementation blocks

HOW

High Energy Physics

Biomedics

Digital Libraries

networking

specific services

Operational support

certification

One stop shop

GRID INFRASTRUCTURE

joint research activities

Technology testbeds

User testbeds

specific services

joint research activities

Security

middleware

Global filing system

research results

EU policies

European Commission

Information Society and Media

e-Infrastructure - Strategic building blocks

HOW

EGEE

DEISA

Market

GRID INFRASTRUCTURE

Research

IST

GÉANT INFRASTRUCTURE

e-Infrastructure

+ 18 proposals international, policies, user communities

...

GÉANT2

Mário Campolargo DG INFSO F3, Pisa 24th October 2005

European Commission

Information Society and Media

- **"integration of existing national high-end platforms, tightly coupled by a dedicated network and supported by innovative system and grid software"**

- **Initial scientific applications include**
  - Material Sciences
  - Cosmology
  - Plasma Physics
  - Life Sciences

- **http://www.deisa.org**

Enabling Grids for E-sciencE

- **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**

- **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

- **GEANT2 en route**

- **http://www.geant.net/**
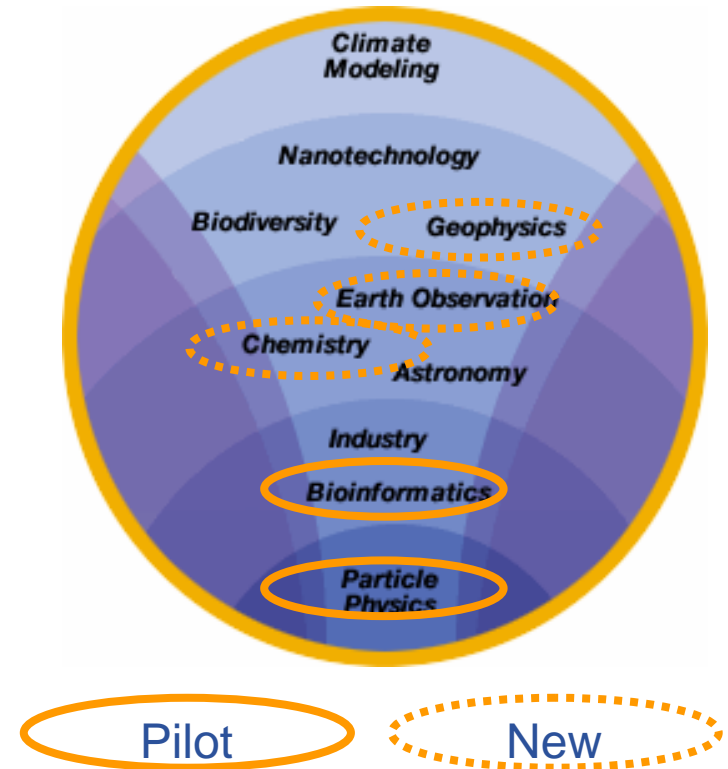
EGEE: Enabling Grids for E-sciencE

- **To underpin collaboration**

- **Link with and build on national, regional and international initiatives**

- **Foster world-wide international cooperation both in the creation and the use of the e-infrastructure**

Collaboration

Grid

Operations, Support and training

Network infrastructure (GÉANT)

EGEE Organisation
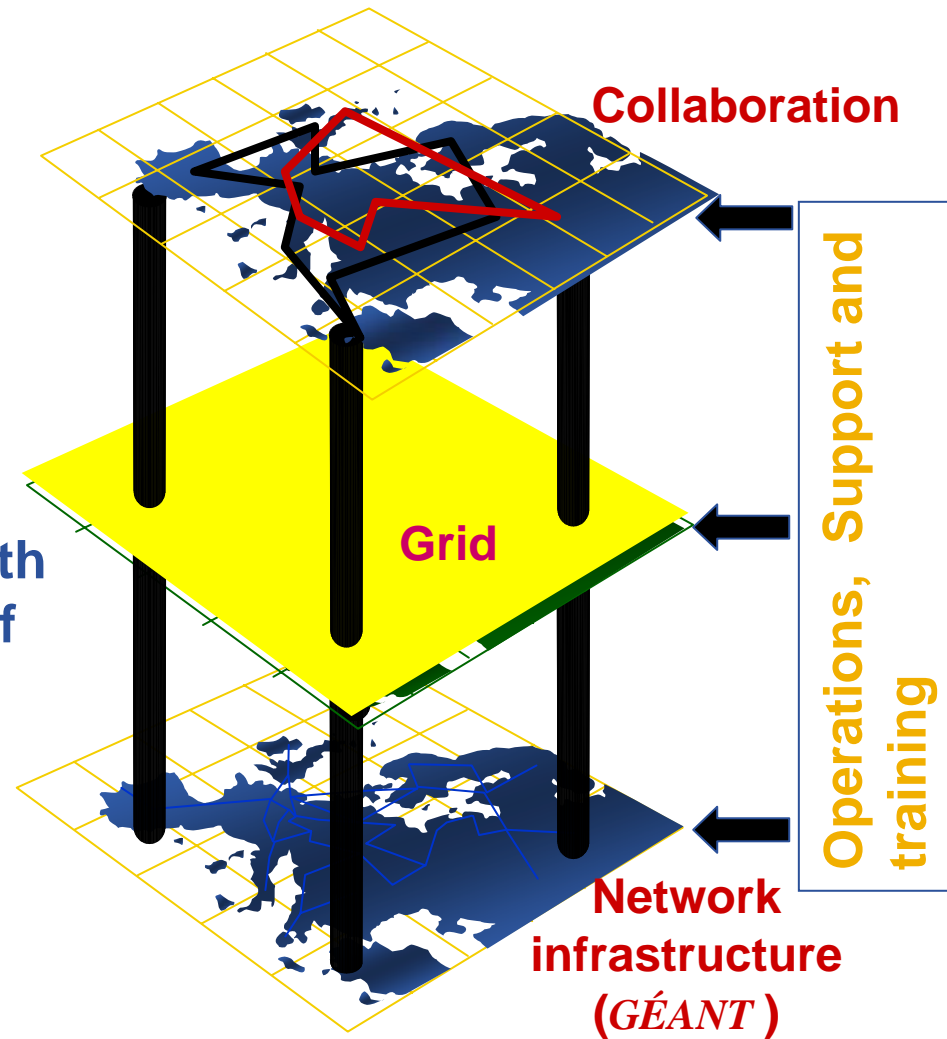
- **70 leading institutions in 27 countries, federated in regional Grids**

- **~32 M Euros EU funding for first 2 years starting April 2004 (matching funds from partners)**

- **Leveraging national and regional grid activities**

- **Promoting scientific partnership outside EU**

CERN
Central Europe
France
Germany and Switzerland
Ireland and UK
Italy
Networking
Northern Europe
Russia
South-East Europe
South-West Europe
USA

Enabling Grids for E-sciencE

- **Goal: establish production quality sustained Grid services with**
  - 3000 users from at least 5 disciplines
  - integrate 50 sites into a common infrastructure
  - offer 5 Petabytes (10^{15}) storage
- **Achieved**
  - **> 180 sites in 39 countries**
  - **~ 20 000 CPUs**
  - **> 5 PB storage**
  - **> 10 000 concurrent jobs per day**
  - **> 60 Virtual Organisations**

Climate Modeling
Nanotechnology
Biodiversity
Geophysics
Earth Observation
Chemistry
Astronomy
Industry
Bioinformatics
Particle Physics

Pilot   New

- **EGEE committed to "hit the ground running"**
- **EGEE profits from the resources - no funded computing/data resources in EGEE**
  - Provided by the VOs
- **LCG obtains additional production and operation efforts**
- **LCG experiments now comprise several of the many VOs in EGEE**
- **Current service ("LCG-2") based on work done in LCG**
  - Middleware components to be upgraded by "gLite" services as they are proven
  - "gLite 3" will be forged from LCG 2.7 + gLite services

LCG : Large Hadron Collider Compute Grid
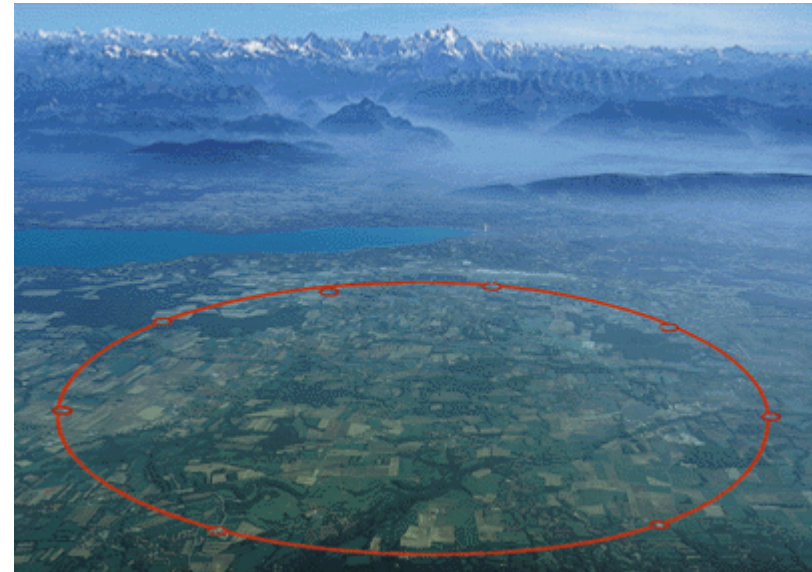
- **48 % service activities (Grid Operations, Support and Management, Network Resource Provision)**

- **24 % middleware re-engineering (Quality Assurance, Security, Network Services Development)**

- **28 % networking (Management, Dissemination and Outreach, User Training and Education, Application Identification and Support, Policy and International Cooperation)**

24% Joint Research

28% Networking

48% Services

Emphasis in EGEE is on operating a production grid and supporting the end-users

gLite: Guiding Principles

- **Service oriented approach**
  - Allow for multiple interoperable implementations
- **Lightweight (existing) services**
  - Easily and quickly deployable
  - Use existing services where possible
    - Condor, EDG, Globus, LCG, …
- **Portable**
  - Being built on Scientific Linux and Windows
- **Security**
  - Sites and Applications
- **Performance/Scalability & Resilience/Fault Tolerance**
  - Comparable to deployed infrastructure
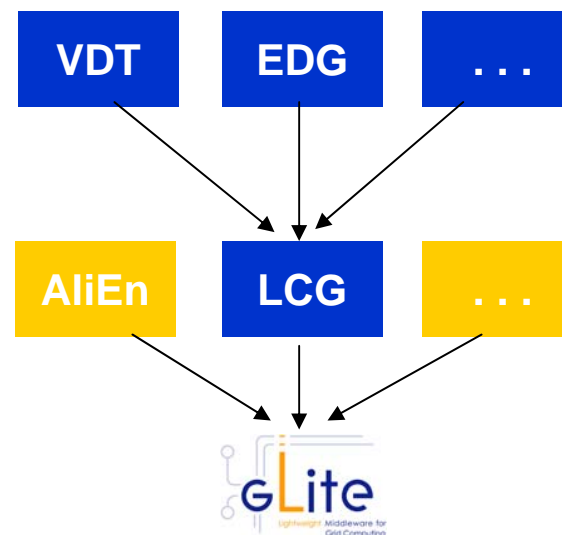
- **Co-existence with deployed infrastructure**
  - Co-existence with LCG-2 and OSG (US) are essential for the EGEE Grid services
- **Site autonomy**
  - Reduce dependence on 'global, central' services
- **Open source license**

Operations

- **Note the contrast between**
  - **"best-efforts" and production grids** for international collaborations… with hundreds of sites providing resources
    - Operational infrastructure (>40% of EGEE budget on operations)
    - Quality of service / policy issues
    - Focus on stability of sites
    - Support for VO's
  - **Research and production middleware**
    - procedures for upgrading middleware
      - *Pre-production grid – running many VO's applications*
  - **Project grids and international production grids**
    - Extent of international cooperation, policy agreement...
    - Multiple VO's

eGee

Enabling Grids for E-sciencE
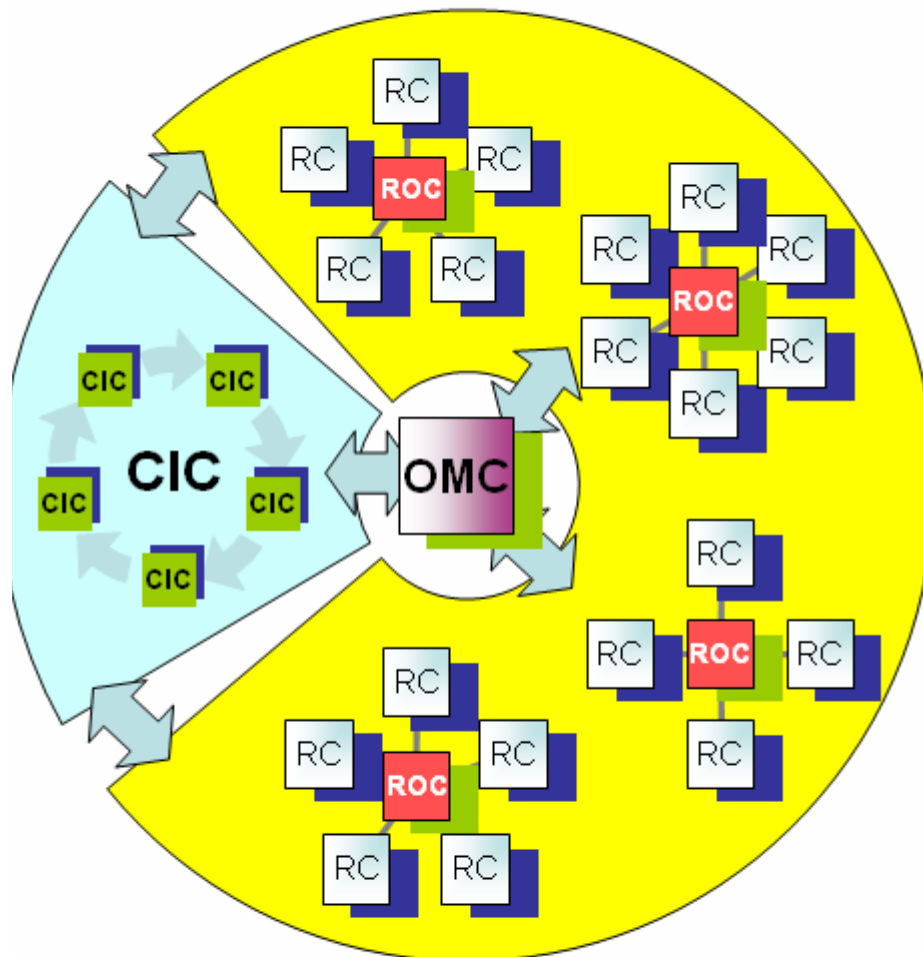
RC = Resource Centre
ROC = Regional Operations Centre
CIC = Core Infrastructure Centre
OMC = Operations Management Centre

- **CICs act as a single Operations Centre**
  - Operational oversight *(grid operator)* responsibility
  - rotates weekly between CICs
  - Report problems to ROC/RC
  - ROC is *responsible* for ensuring problem is resolved
  - ROC oversees regional RCs
- **ROCs responsible for organising the operations in a region**
  - Coordinate deployment of middleware, etc
- **CERN coordinates sites not associated with a ROC**
- **Global Grid User Support**

- **Grid operator on duty**
  - 6 teams working in weekly rotation
  - improving site stability
- **Operations coordination**
  - Weekly operations meetings
  - Regular ROC, CIC managers meetings
  - Series of Workshops
- **Geographically distributed responsibility for operations:**
  - There is no "central" operation
  - Tools are developed/hosted at different sites:
    - GOC DB (RAL), SFT (CERN), GStat (Taipei), CIC Portal (Lyon)
- **Procedures described in Operations Manual**

Global Grid User Support - first contact for users

http://www.ggus.org
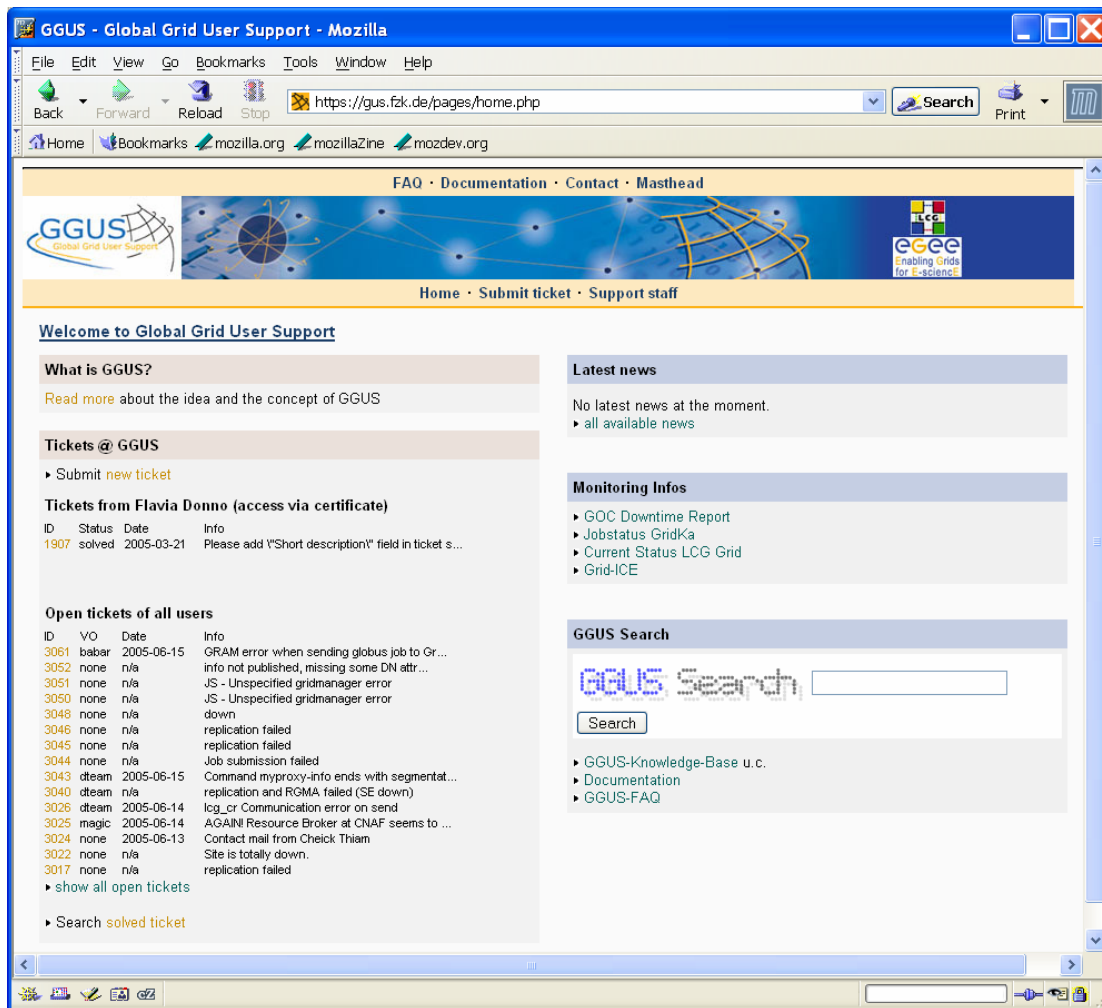
You need to **register**
in order to be able
To use this portal
(**GSI** or password based)

You can register
as **User** or as
Supporter.

Supporter ?
If you think you
have a good
knowledge in Grid
and have time
to provide support,
please contact
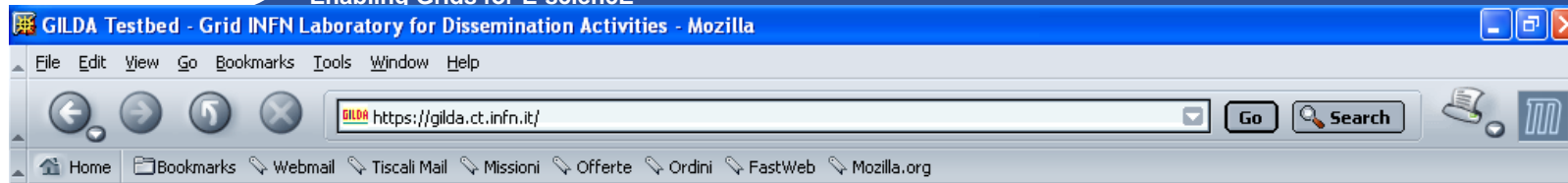your ROC or directly
ESC at:

Building user communities

- **High-Energy Physics (HEP)**
  - Provides computing infrastructure (LCG)
  - Challenging:
    - thousands of processors world-wide
    - generating petabytes of data
    - 'chaotic' use of grid with individual user analysis (thousands of users interactively operating within experiment VOs)

- **Biomedical Applications**
  - Similar computing and data storage requirements
  - Major additional challenge:
    security & privacy

- **Chemistry, Earth Observation, Astronomy, Geophysics, ...**

Enabling Grids for E-sciencE
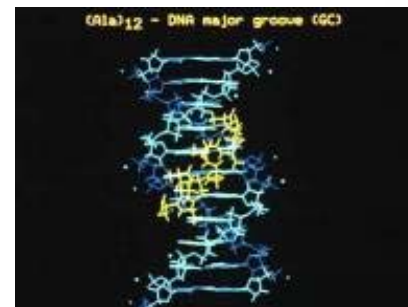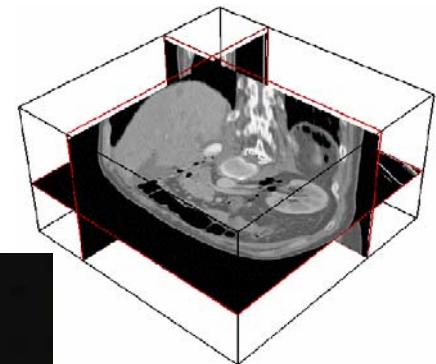
- **Why t-infrastructure?**
  - Training is necessary: personal + e-learning
  - e-Infrastructure for production
  - t-Infrastructure for training

- **Need guaranteed response for tutorials; 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
  - need safe resources for installation training
  - easy entry point for new communities

Enabling Grids for E-sciencE

International cooperation

2001

Condor   Globus   MyProxy   ...

EDG   ...

OSG, ...   VDT   DataTAG

CrossGrid   LCG   ...

SRM

2004

GridCC   NextGrid   EGEE   DEISA   ...

interactive

USA   EU

Used in

Future grids

- **Cooperation between EGEE and other Grid activities**
  - **Globus Alliance, Condor**
  - **Training/workshop events (International Summer School of Grid Computing, July… 2 intense weeks… registrations open! http://www.dma.unina.it/~murli/ISSGC06/)**
  - eInfrastructure reflection group in Europe http://www.e-irg.org/
- **Standard setting through attendance at global standard bodies such as the Global Grid Forum.**
  - Grid Storage Management GGF working group - http://sdm.lbl.gov/gsm/
  - Security, Authentication: US – EU cooperation
- **Mutual recognition of Certificate Authorities**
  - Requires collaboration to establish policy - and mutuality

- **Joint Security Policy Group**
 - EGEE with strong input from OSG
  - Policy Set:
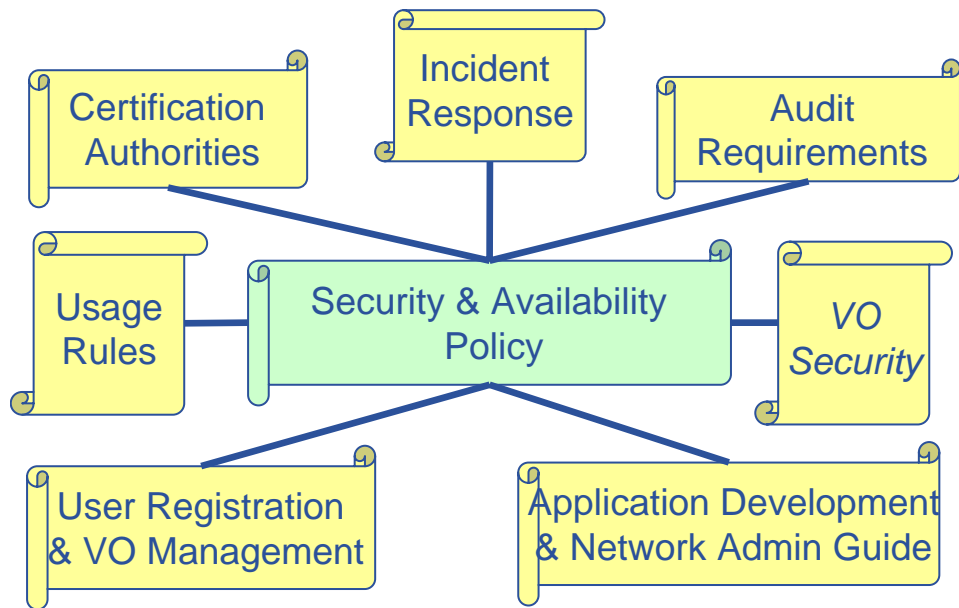
- **Policy Revisions**
  - Grid Acceptable Use Policy (AUP)
    - https://edms.cern.ch/document/428036/
    - common, general and simple AUP
    - for all VO members using many Grid infrastructures
      - *EGEE, OSG, SEE-GRID, DEISA, national Grids…*
  - VO Security
    - https://edms.cern.ch/document/573348/
    - responsibilities for VO managers and members
    - VO AUP to tie members to Grid AUP accepted at registration
  - Incident Handling and Response
    - https://edms.cern.ch/document/428035/
    - defines basic communications paths
    - defines requirements (MUSTs) for IR
      - *reporting*
      - *response*
      - *protection of data*
      - *analysis*
    - not to replace or interfere with local response plans

EGEE-II

Enabling Grids for E-sciencE

- **EGEE-II proposal submitted to the EU**
  - Proposed start 1 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

EGEE — SA 48%, NA 28%, JRA 24%

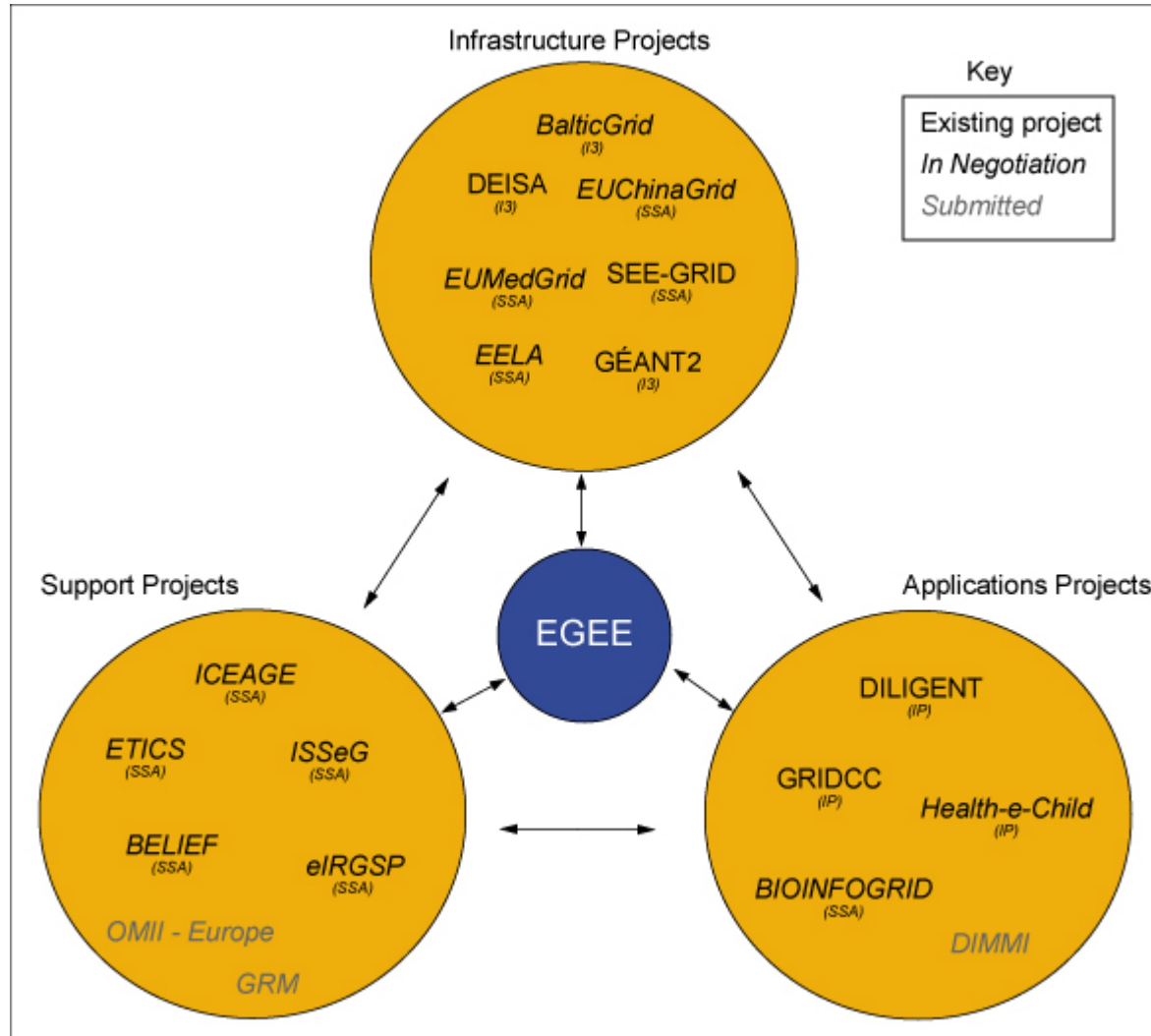EGEE-II — SA 54%, NA 32%, JRA 14%
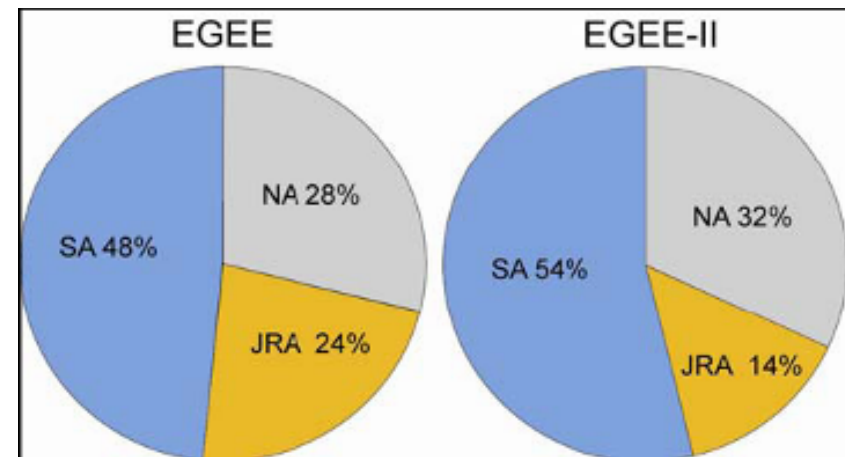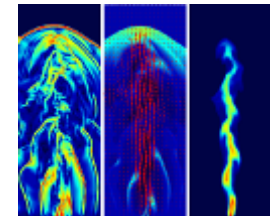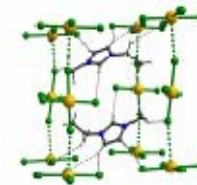
- **More than 90 partners**
- **32 countries**
- **12 federations**
→ **Major and national Grid projects in Europe, USA, Asia**

+ 27 countries through related projects:
  – BalticGrid
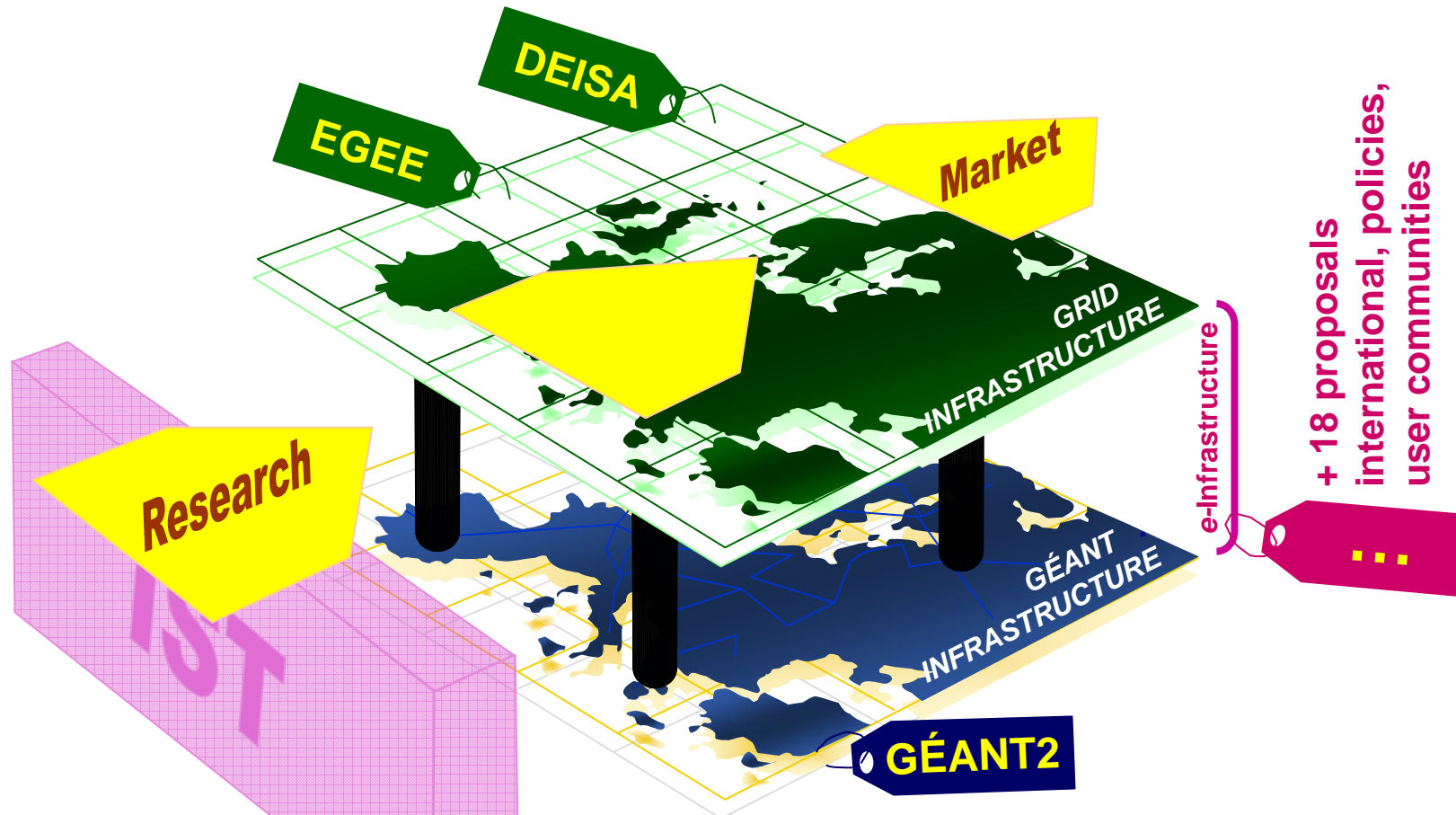 – SEE-GRID
  – EUMedGrid
  – EUChinaGrid
  – EELA

eGee

eGee

Enabling Grids for E-sciencE

Name	Description	Common partners with EGEE
BalticGrid	EGEE extension to Estonia, Latvia, Lithuania	KTH – PSNC – CERN
EELA	EGEE extension to Brazil, Chile, Cuba, Mexico, Argentina	CSIC – UPV – INFN – CERN – LIP – RED.ES
EUChinaGRID	EGEE extension to China	INFN – CERN – DANTE – GARR – GRNET
EUMedGRID	EGEE extension to Malta, Algeria, Morocco, Egypt, Syria, Tunisia, Turkey	INFN – CERN – DANTE – GARR – GRNET – RED.ES
ISSeG	Site security	CERN – CSSI – FZK – CCLRC
eIRGSP	Policies	CERN – GRNET
ETICS	Repository, Testing	CERN – INFN – UWM
ICEAGE	Repository for Training & Education, Schools on Grid Computing	UEDIN – CERN – KTH – SZTAKI
BELIEF	Digital Library of Grid documentation, organisation of workshops, conferences	UWM
BIOINFOGRID	Biomedical	INFN – CNRS
Health-e-Child	Biomedical – Integration of heterogeneous biomedical information for improved healthcare	CERN

- **… the largest multi-VO production grid in the world!**

- **What's happening now?**
  http://gridportal.hep.ph.ic.ac.uk/rtm/

- **What resources are connected?**
  http://goc.grid-support.ac.uk/gridsite/monitoring/

eGee

Enabling Grids for E-sciencE

- **See press release: „EGEE battles malaria with Grid wisdom" (over 46 million docked ligands)**

- **See press release: „EGEE makes rapid earth quake analysis possible" (analysis of large indonesian earth quake 28.03.05 within 30 hours, showed that it was not an aftershock of the tsunami)**

Enabling Grids for E-sciencE

- **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
  - And…

- **Grids are fundamentally about people**
- **… how people in different organisations commit to cooperate**
- **… and how that cooperation can be enabled by operations, training, support, and (most transient of all!) middleware**

egee

Enabling Grids for E-sciencE

Mário Campolargo DG INFSO F3, Pisa 24th October 2005

eGee
Enabling Grids for E-sciencE

- **EGEE Website**
  http://www.eu-egee.org

- **How to join**
  http://public.eu-egee.org/join/

- **EGEE Project Office**
  project-eu-egee-po@cern.ch

- **Global Grid Forum http://www.gridforum.org/**
- **Globus Alliance   http://www.globus.org/**

- **Condor http://www.cs.wisc.edu/condor/**
- **VDT http://www.cs.wisc.edu/vdt/**

- **Open Science Grid http://www.opensciencegrid.org/**
- **Grid Center http://www.grids-center.org/**

- **LCG http://lcg.web.cern.ch/LCG/**