Thanks to: V. Hernández, I. Campos, I. Martín Llorente, I.Blanquer, J.Gomes

Origin and Evolution of the

Spanish NGI

Cracow '08
Grid Workshop

Jesús Marco de Lucas *[marco (at) ifca.unican.es]*

CSIC Research Professor at Instituto de Física de Cantabria

Outline

- The e-Science Thematic Network in Spain
- Building the Spanish National Grid Infrastructure
- A relevant example: the initiative Grid-CSIC
- Towards an European Grid Infrastructure

Red española de e-Ciencia

Spanish e-Science Network

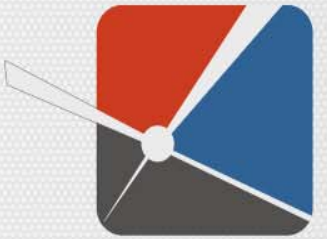
Acción financiada por:

GOBIERNO DE ESPAÑA

Entidad Coordinadora:

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Evolution of the Spanish e-Science Initiative

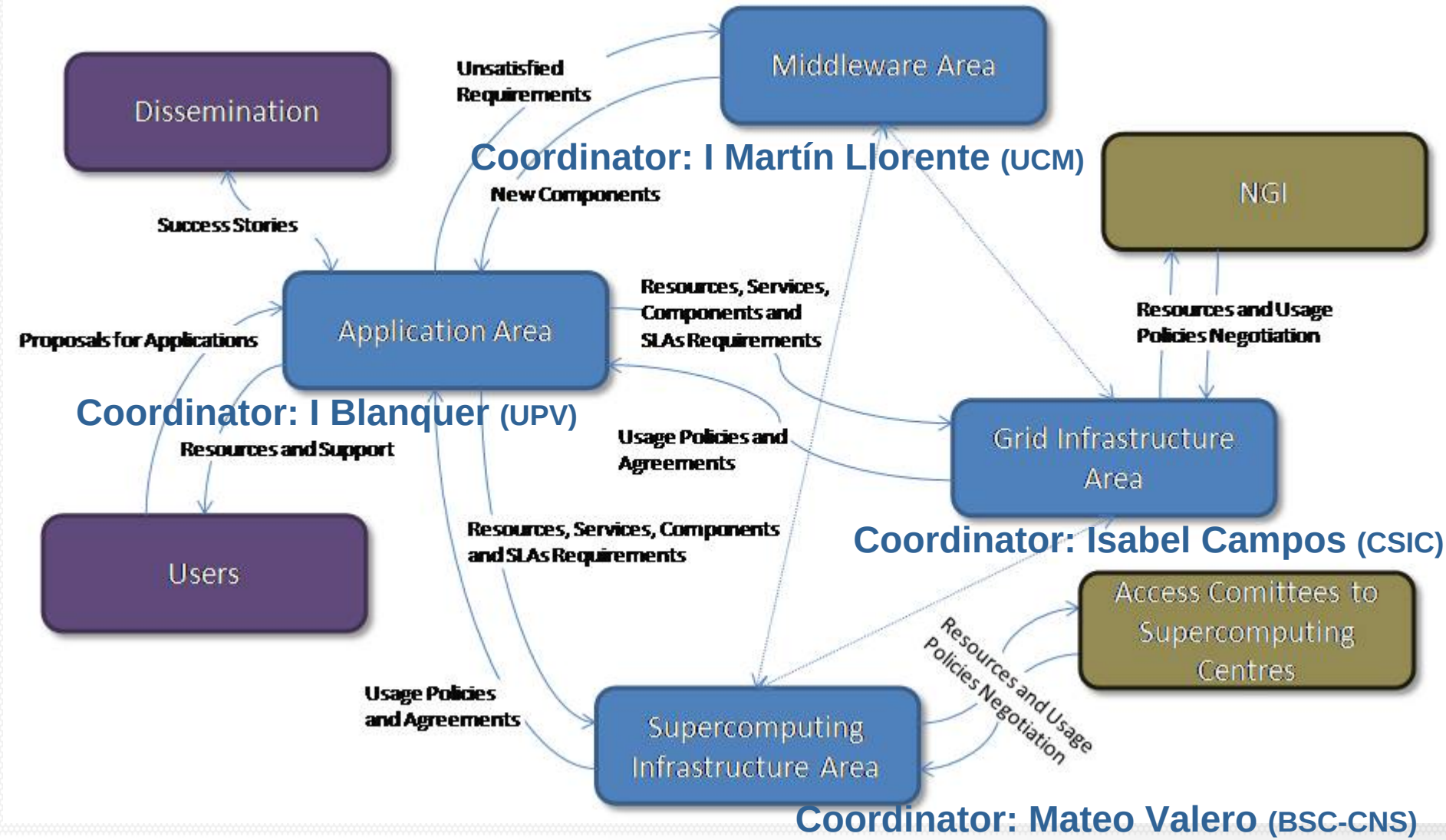


- In early 2003, around 20 Spanish research groups with interest in Grid Technologies, several of them participating in the **DATAGRID** and **CROSSGRID** projects, started a networking activity (IRISGRID)
  - objective: analyze, define and foster the consolidation of a scientific community of users, infrastructure providers, application and middleware developers with interest on Grids.
- This community started promoting the concept of a Spanish NGI at the Science Ministry
  - 2004: experts group prepared **White Paper on e-Science**
  - 2006: official responsible nominated by Ministry with consensus (Vicente Hernandez, UPV)
  - 2007: official support through "**e-Science Network**" and **Joint Research Unit for Grid** projects (namely ES-GRID)
  - 2008: first meeting in Madrid (February), start of funded activities
  - Next plenary meeting: Seville (23-24 October)
- Network has created a strong link with the Portuguese NGI (IBERGrid initiative)
- **It is open to, and welcomes, links with other initiatives**

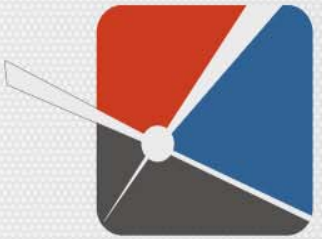
Spanish e-Science Network Organization

Red Española de e-Ciencia

Coordinator: Vicente Hernández (Universidad Politécnica Valencia)



Components in the Spanish NGI



- Core: **Spanish institutions participating in research and development projects about Grid research infrastructures** being the attractor for many other smaller-scale centers
  - **EGEE**, EUFORIA, DORII, EELA, i2g, … with common basic middleware (mainly gLite based)
  - RedIris (Spanish NReN) support EUGRIDPMA certificates
  - About 20 resource provider centers, with 2400 cores and 340 direct users
  - Communities of **High Energy Physics (LHC)**, Biomedicine, Fusion, Computational Chemistry, Astrophysics and Earth Sciences
- Relation with the **Spanish Supercomputing Network** (RES)
  - This network comprises several Spanish research centers that operate a common infrastructure of supercomputing.
  - The links between Supercomputing and Grid are one of the most important challenges of the Spanish Network for e-Science, and it is forecasted that many collaboration in terms of middleware, applications and even resources will be developed

Application Area

Red Española de e - Ciencia

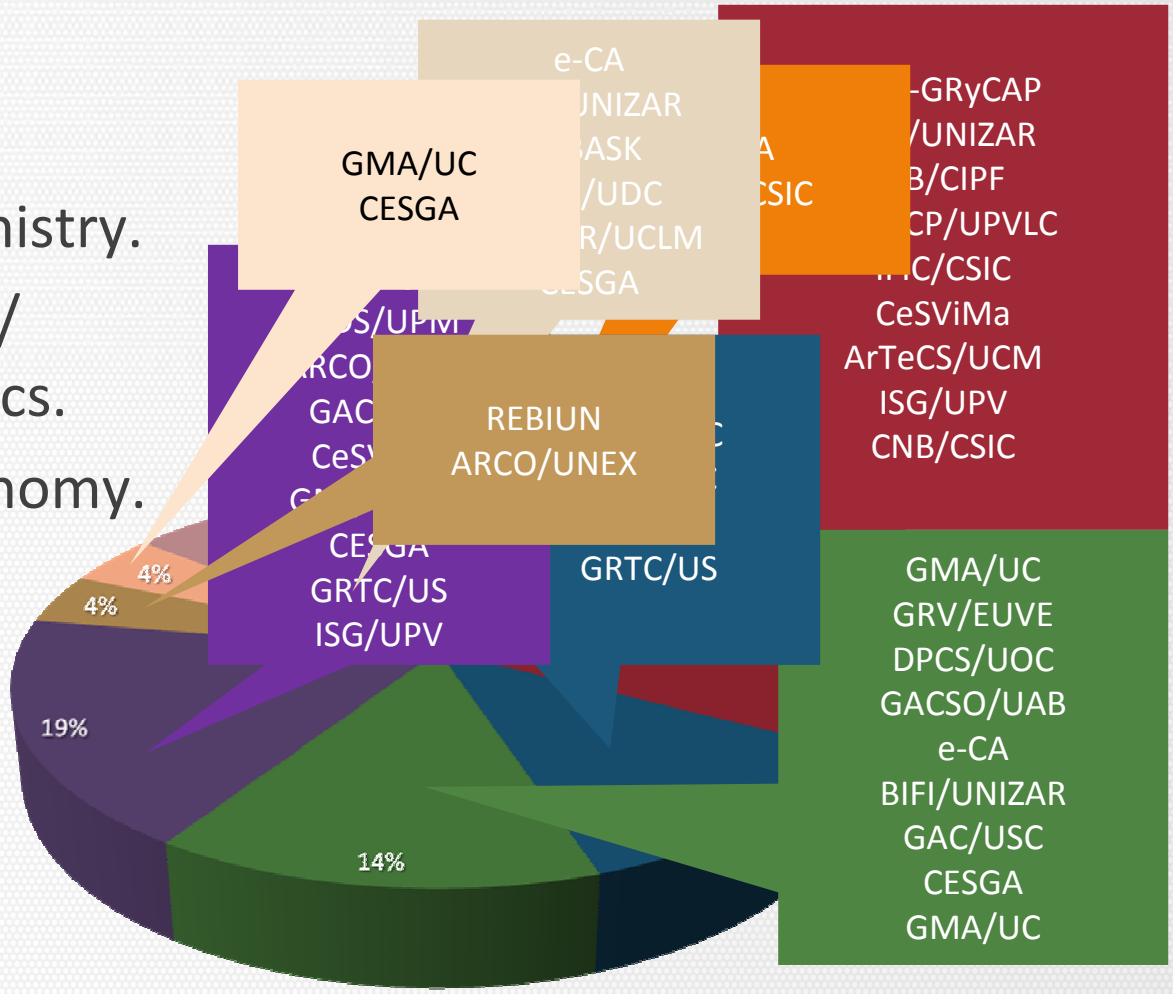
- **Consolidate existing Virtual Organizations, and integrate new research groups** interested in these areas.
- **Promote new areas:** create Virtual Organizations oriented to:
  - support applications of interest for Spanish research groups
  - address applications of social interest, with the help and implication of the administration.
- **Standardize the procedures** for creation, deployment and exploitation of applications in e-Science
- Promote the creation of **general interest software** useful in several applications, like data repositories, user interfaces…

Application Area

Research Groups

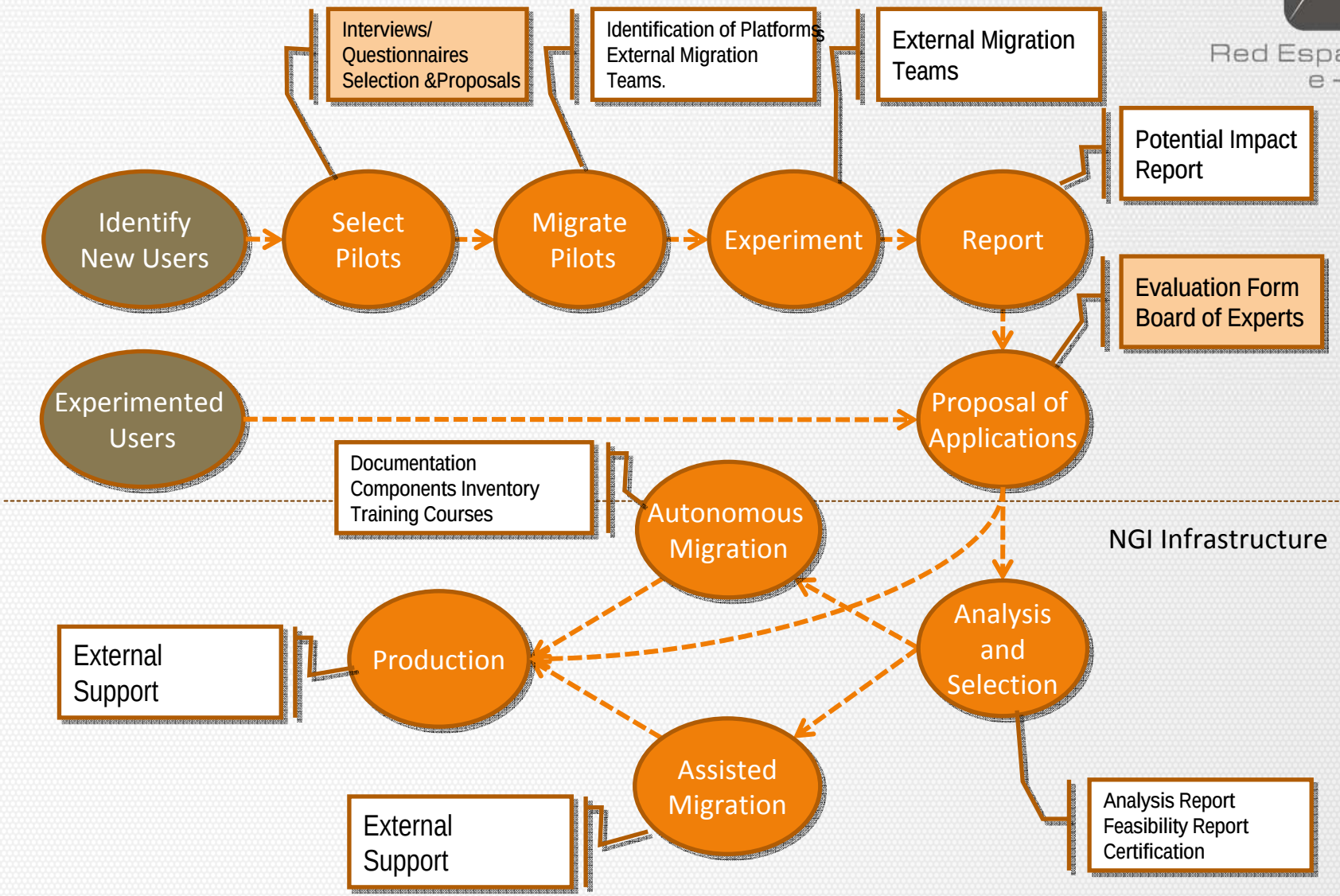
Red Española de e-Ciencia

- Biomedicine / Bioinformatics / Biotechnology.
- Engineering.
- Earth Science.
- Computational Chemistry.
- High-Energy Physics / Computational Physics.
- Astrophysics / Astronomy.
- Mathematics.
- Information and Communication Technologies.



Application Workflow

Long-term Workplan



Middleware Area

Red Española de e - Ciencia

Approach

- **Satisfy the requirements:** applications, infrastructure and interoperability

Constraints

- Existing infrastructures (DEISA, EGEE…) have their **own scheme** for test, validation, certification e integration
- Maturity of the **basic middleware** (GT, gLite & UNICORE)
- **Interests of the different groups and available funding**

Lines of action

- Focus on **high level components** close to the final user (applications) or managers (infrastructure)
- Impact of **infrastructure interoperability**
- Provide users with **models** (HTC, workflow…) **& local APIs** (DRMAA, MPI…)
- **Promote** developments and collect experience (web, wiki…) at Spanish level

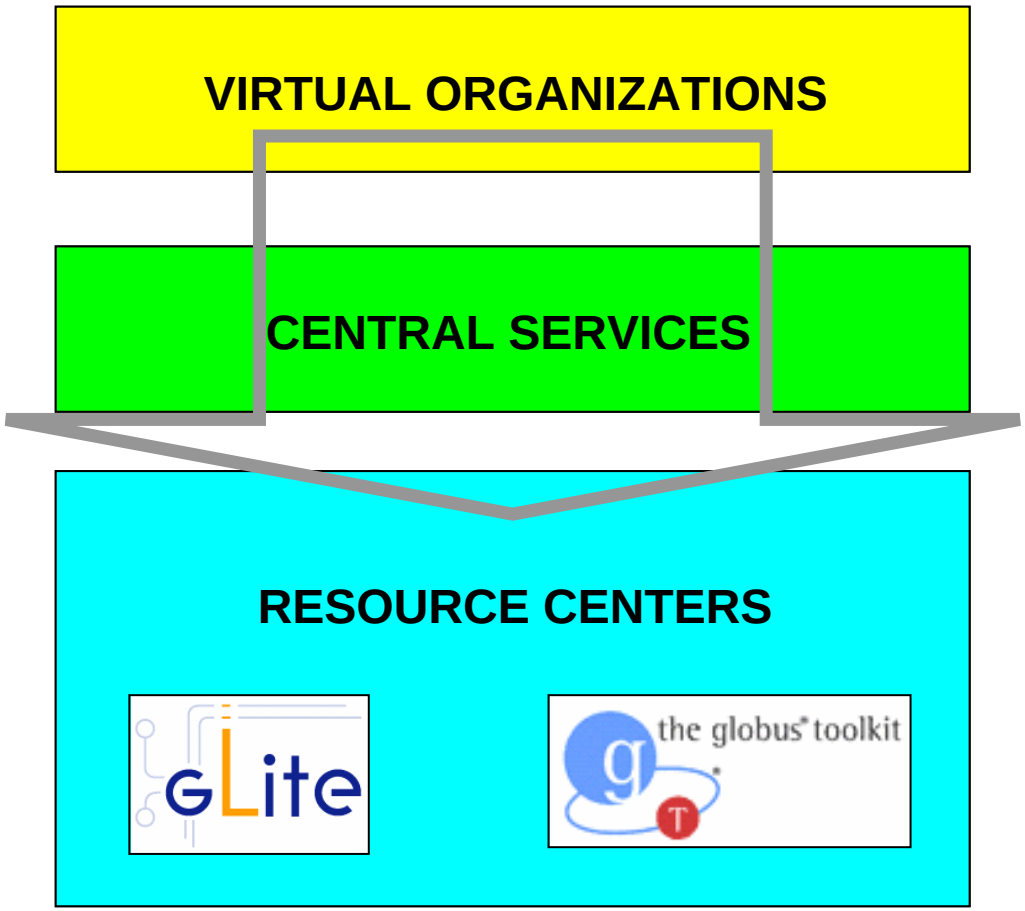
National Grid Infrastructure Architecture: oriented to support communities

VO Oriented

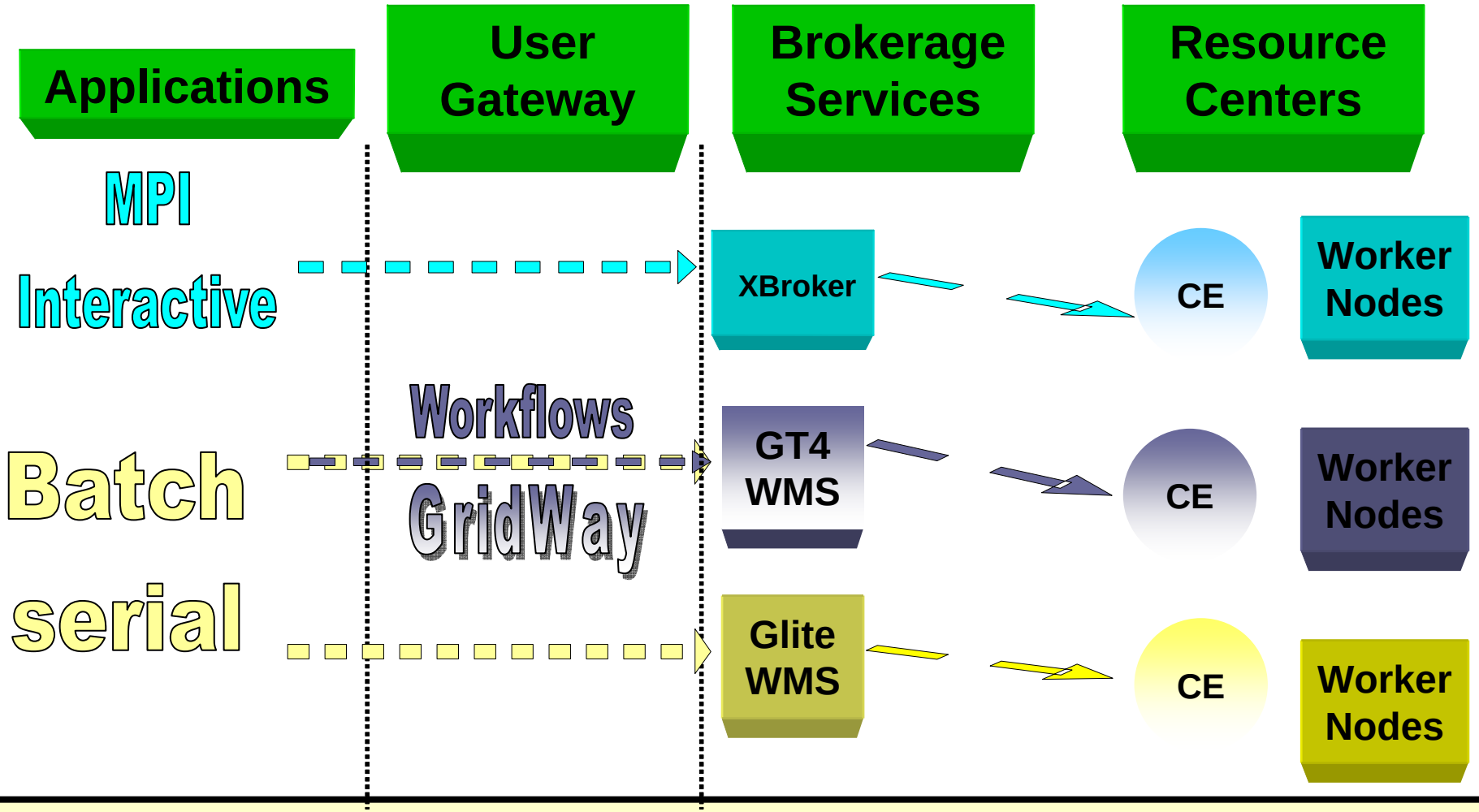
The Architecture of the NGI-ES is oriented to the support of Virtual Organizations

Key Issues

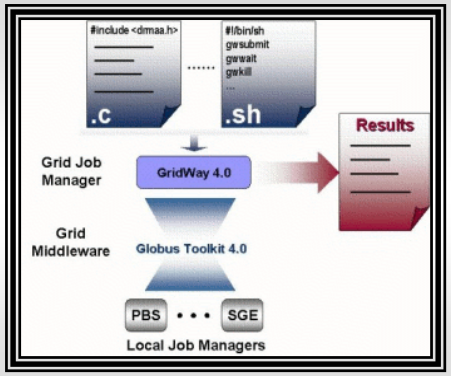
- ✓ Advanced VO services
  - ✓ User support
  - ✓ Monitoring & Accounting
- ✓ Application porting and support
- ✓ Middleware driven by application requirements



NGI Architecture:
making users life easier

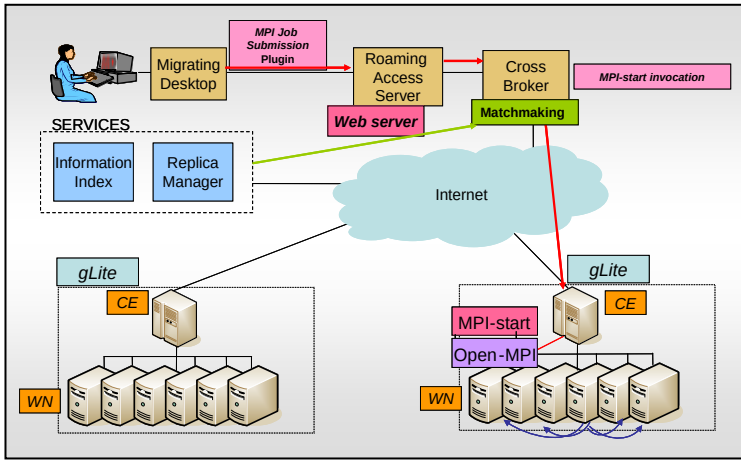


NGI-ES added value: Brokerage services



The Spanish NGI includes the experience in Scheduling and Job Management of the team from the University Complutense in Madrid which developed GRIDWAY

More information on http://www.gridway.org
Contact: Ignacio Martin Llorente (UCM)



The Spanish NGI supports MPI and interactive Jobs using the broker technology developed at the University of Autonoma of Barcelona, the CROSSBROKER

More information under http://www.i2g.eu
Contact: Enol Fernández (UAB)

The need for interoperation at the NGI level: important issues from different perspectives

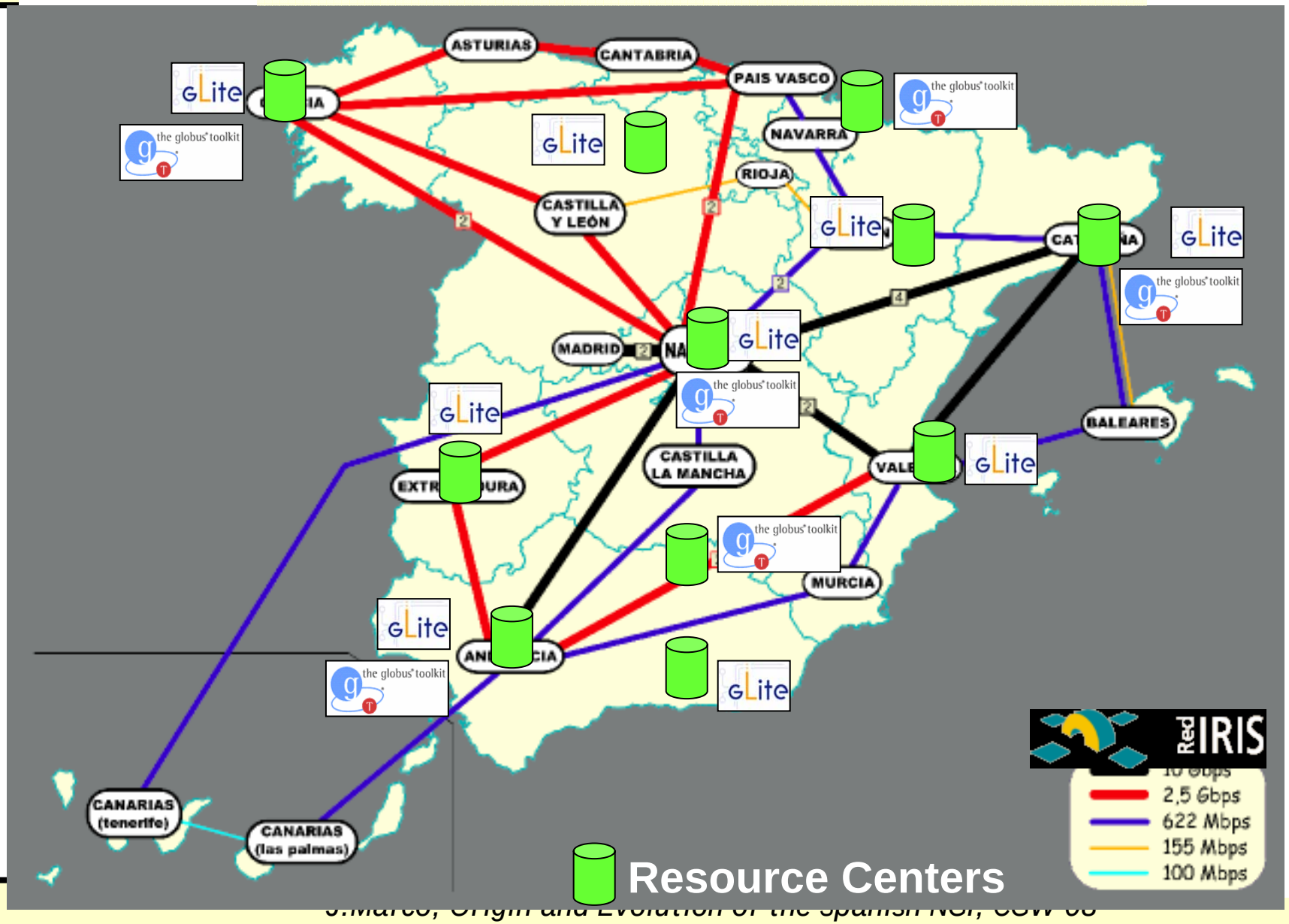
- **E-Science application users**
  - Common ways for accessing any e-infrastructure resources
  - Sometimes it is the same user that wants to access one of another type of resources depending on the application

  The access method should be transparent for the user
  - Potential access to a significantly larger set of resources
- **E-infrastructure owners (Resource Centers)**
  - Reduced management overheads if only a single Grid middleware system needs deployment
  - Potential for greater resource utilisation
- **Grid Middleware and Applications Developers**
  - An interoperable set of services
  - Applications portable across different Grid middleware systems

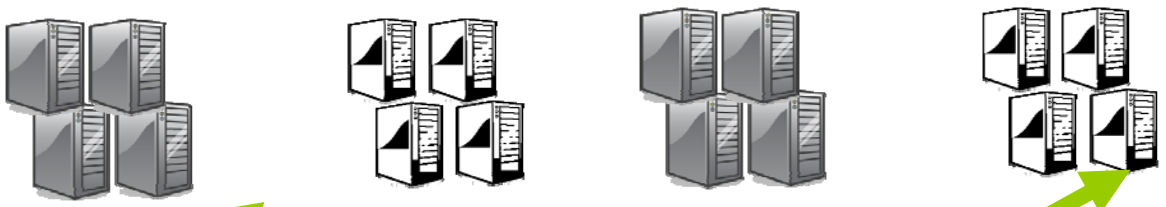
What do we need to interoperate ? NGI-ES Resource Centers

- So far 18 resource centers have answered to the call for infrastructures dedicated to the NGI
- Resources are a minimum of 1300 execution cores with a peak of 4300 at times when the local occupancy at the resource centers is lower than expected
- Resource Centers
  - **Grid Infrastructure of CSIC**
  - **Regional Computing Centers**
    - **CESGA (Galicia), CICA (Andalucia), CESCA (Cataluña)**
  - **Research Centers country wide involved in Grid projects**
    - **EGEE, i2G, EELA, Cytedgrid, WLCG,..**
  - **University departments with computational needs**
    - **Physics, Chemistry, Environment, Biology, Engineering,...**

Resource Centers Map

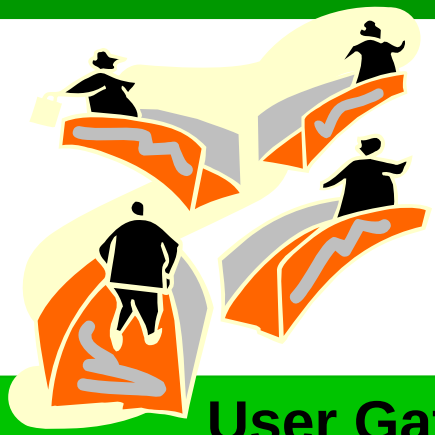


NGI Architecture
User oriented tools needed for interoperation



Resource Centers

Brokerage Services Interoperation Layer:
Global service of the NGI

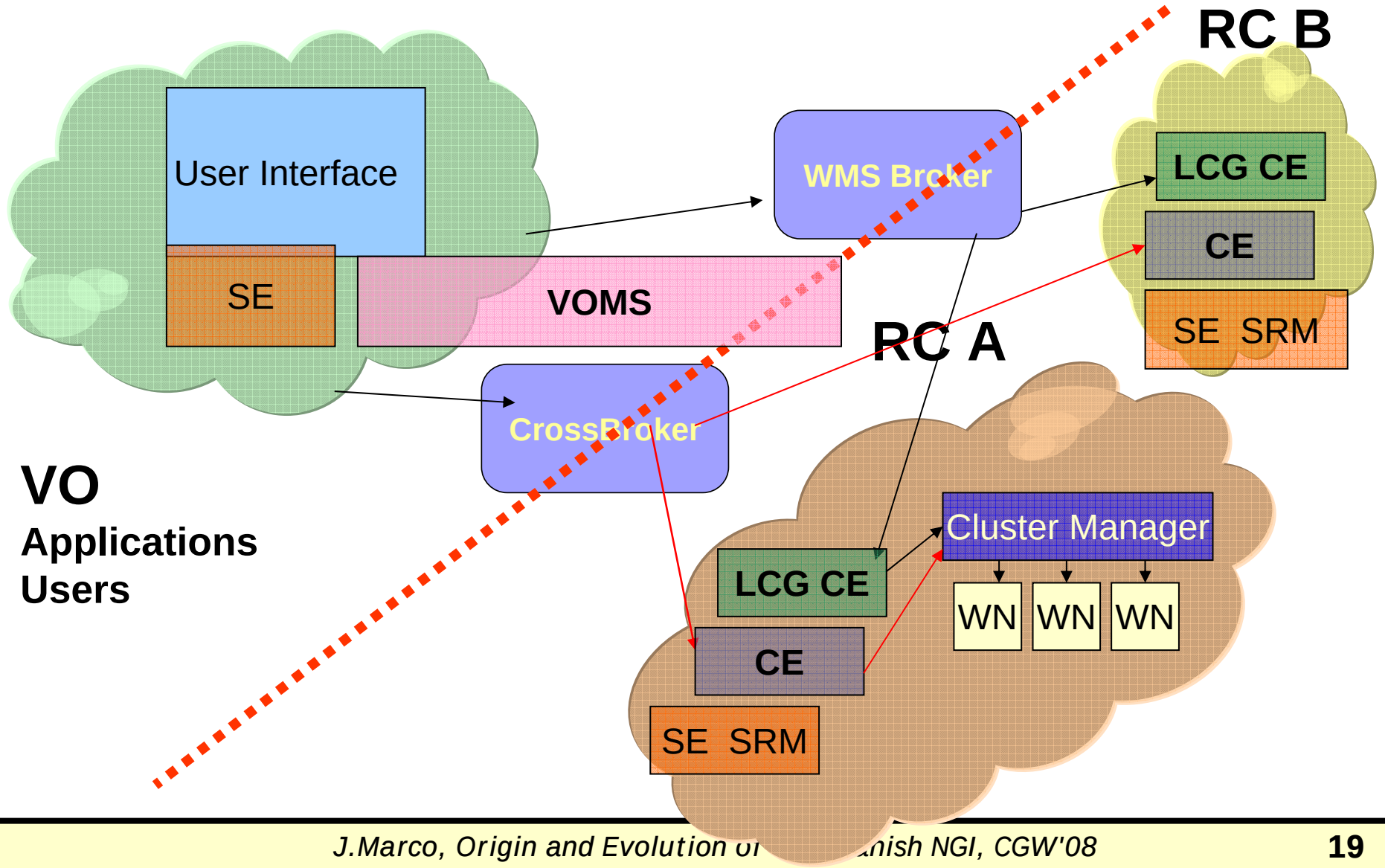


User Gateway:
Global Service of the NGI

Global Services needed for the interoperation layer

- **Authentication mechanism based on EUGRIDPMA**
  - **PKIRISGRID**
  - **Under consideration using ID Card based Certificates**
- **VOMS deployment**
  - VOMS are deployed by application area at the NGI support centers
  - Tied with the existence of a scientific community at the center
    - Proximity to users makes user support **easier**
- **Global Information System**
  - TOP-BDII based on standards:
    - OpenLDAP + GlueSchema implemented at CSIC
  - Integration work going on with the Portuguese NGI TOP-BDII
    - **IBERGRID initiative will model the union of the two NGI s to support common research projects between Spain and Portugal**
- **Portal for Monitoring and Accounting**
  - Accounting and monitoring portal based on the tools developed for EGEE by CESGA
- **Dedicated Brokerage Services**
  - *Glite WMS* services
  - *Gridway* to be used with glite and/or GT4
  - *CrossBroker* to be used with glite and/or GT4
- **Helpdesk for site administrators** based on RT
  - e-mail based ticketing system

VO ⇔ Services and Resource Centres relationship



Requirements and "agreements" RC ⇔ NGI

Requirements of the NGI to the Resource Centers

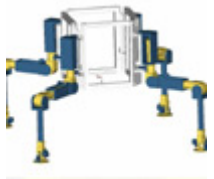
- **NGI-ES is built as a production infrastructure**
  - **In terms of QoS**
    - Follow the indications of site maintenance
    - Sites need to guarantee continuity at the level of human and material support
      - Minimum support hours: Monday – Friday from 9-17.
  - **Resource Accessibility**
    - Signature of a Resource Allocation Policy to guarantee users the access to the infrastructure
    - Sites will be included in the **monitoring system of the NGI..**
    - **...and will have to respect the directives**

Requirements of Resource Centers to the NGI

- **Resource Centers have already well defined access policies**
  - Mainly depending on the funding agency
  - How will those allocation policies will interact with NGI-ES is still to be defined
- Resource centers dedicated to general user support need a mechanism to **assign resources in a deterministic way**
  - Ex. The application of Bio Informatics from user A will use the resources of the Biomed VO for the next 6 months with a maximum of 300 cores and 2 TB / day
- **The NGI has to prove the advantages of accessing the resources in Grid mode**

GRID-CSIC Initiative

- **CSIC (Consejo Superior de Investigaciones Científicas) is the largest public research body in Spain.**
- **With >120 research centers throughout Spain, we play an active role in the scientific policy of all the country's autonomous regions.**
- **As a multidisciplinary body we cover all fields of knowledge, from basic research through to advanced technological development.**
- **CSIC is subdivided into eight science and technology areas:**
  1. Humanities and Social Sciences.
  2. Biology and Biomedicine.
  3. Natural Resources.
  4. Agricultural Sciences.
  5. Physical Sciences and Technologies.
  6. Chemical Sciences and Technologies.
  7. Materials Science and Technology.
  8. Food Science and Technology.

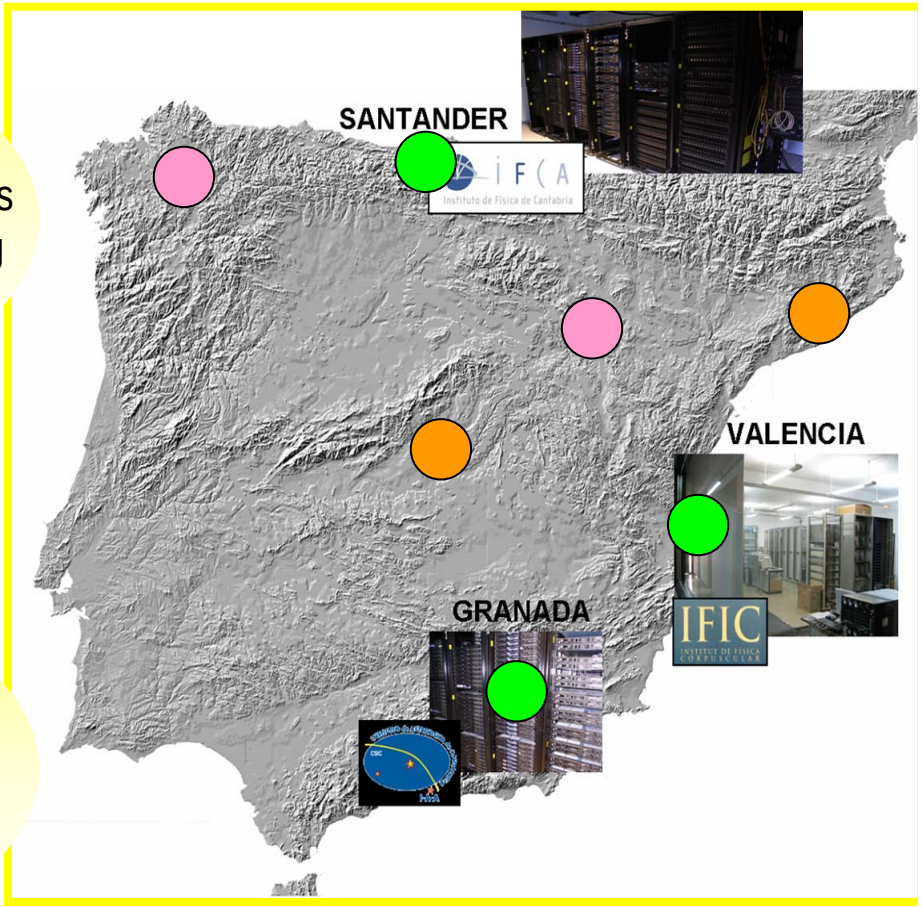



GRID-CSIC Initiative

- Origin:
  - experience at CSIC on GRID projects **(CROSSGRID, i2g, EGEE, EUFORIA, DORII)**
  - Opportunity: participation in e-Science network, NGI, EGI
  - Strategic objective
    - CSIC is also present in the Supercomputing Infrastructures in Spain: BSC & CESGA
  - Area of collaboration with other European institutions (like **CNRS**)
- Objective: deploy an advanced production Grid infrastructure to support research projects where the distributed computing needs exceed what available for a single user or group
- Oriented to promote **multidisciplinary** or **multi-center projects** where researchers need to simulate, analyze, process, distribute or access large data volumes.
- Examples (e-Science):
  - Experimental Particle Physics (CDF, CMS, ATLAS, ILC…)
  - Phenomenology (SUSY models) & Lattice
  - Space Missions (XMM, Planck…)
  - Astronomical Observations
  - Climate modeling
  - Computational Chemistry
  - Bio computing

GRID – CSIC Infrastructure

The Spanish National Research Council-CSIC is deploying the first stable distributed computing infrastructure in Spain to facilitate CSIC researchers the accomplishment of scientific projects requiring computing resources beyond the capabilities of a single user or research group



The GRID - CSIC infrastructure intends to foster multidisciplinary and joint projects between CSIC centers . Researchers from CSIC will have seamless access to a distributed infrastructure consisting in the first phase (2008 - 2010) of over **8000 cores** and a total storage of more than **1 Petabyte**

Why do we need/want/support EGI ?

- Similar reasons to NGI reasons…
- …at a different scale
- We are EU researchers!
  - Consolidate at EU scale
  - So we can address EU wide research efforts
  - Scale effect
    - Coordination (towards Worldwide)
    - Start/Support New Initiatives
    - Dissemination/Outreach/Training
  - Impact
    - EU can make it in 21st century ONLY if we are able to make EUROPEAN LEVEL RESEARCH

Summary

- **Spanish e-Science Network well active**!
  - Added value: broker middleware, support to MPI, etc.
  - Next meeting: 23-24 October in Sevilla
  - Deployment of NGI infrastructure going on
- Relevant Grid Initiatives in Spain
  - Existing RC in EGEE, i2g, EUFORIA, DORII, EELA…
  - New large initiatives: GRID-CSIC
- Interest in European Grid Infrastructure
- **GRID-CSIC project seeking collaboration at European level**
  - **Join us** at the **e-Science & Grids Workshop**, June 2009, Santander