

Status and Future of the EGEE project

Dr. Bob Jones
EGEE Technical Director
CERN
Geneva, Switzerland

OpenLab workshop, CERN

13th June 2005

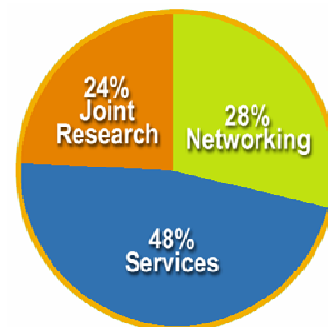
www.eu-egee.org



INFSO-RI-508833

The EGEE project

- **Objectives**
 - consistent, robust and secure service grid **infrastructure** for many applications
 - improving and maintaining the **middleware**
 - attracting **new resources and users** from industry as well as science
- **Structure**
 - 70 leading institutions in 27 countries, federated in regional Grids
 - leveraging national and regional grid activities worldwide
 - funded by the EU with ~32 M Euros for first 2 years starting 1st April 2004



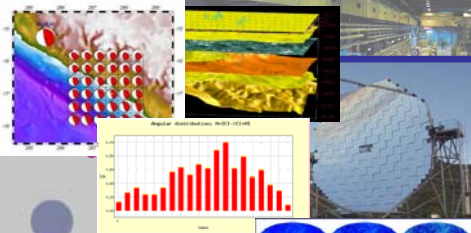
INFSO-RI-508833

OpenLab workshop, CERN, 13th June 2005

2

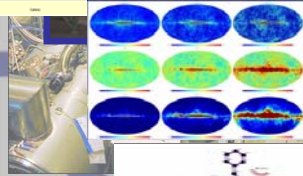
- **Infrastructure** (EGEE-0 / LCG-2):

- >130 sites
- >14 000 CPUs
- >5 PB storage
- >10 000 concurrent jobs
- >60 Virtual Organisations



- **Middleware**

- First release of gLite end of March 2005
 - Focus on basic services, easy installation and Management
 - Industry friendly open source license

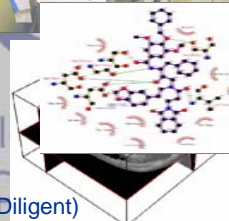


- **Applications**

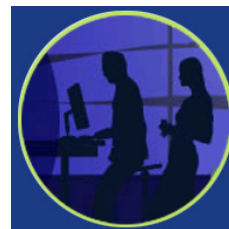
- Pilot Applications
 - HEP
 - Biomedical

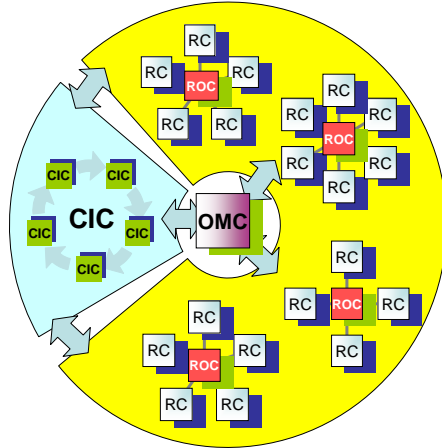
- Generic Applications

- Earth sciences
- MAGIC
- Computational Chemistry
- PLANCK
- Drug Discovery
- Digital Libraries (GRACE, Diligent)



- **Public and technical websites** constantly evolving to expand information available and keep it up to date
- **3 conferences** organised
 - ~ 300 @ Cork, ~ 400 @ Den Haag, ~450 @ Athens
- **Pisa 4th project conference** 24-28 October '05
- **More than 75 training events** (including the GGF grid school) across many countries
 - ~1000 people trained
 - induction; application developer; advanced; retreats
 - Material archive with more than 100 presentations
- **Strong links with GILDA testbed and GENIUS portal** developed in EU DataGrid

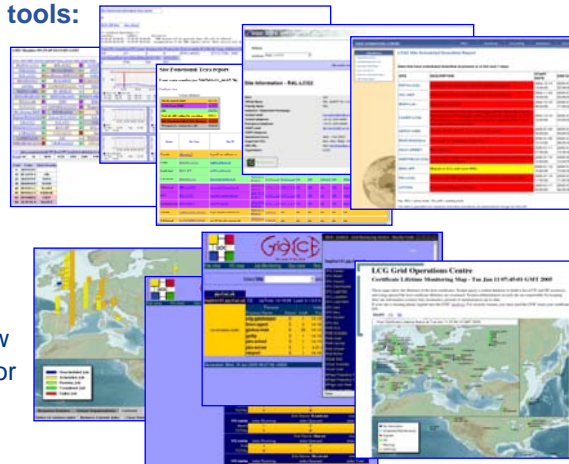




RC - Resource Centre
 ROC - Regional Operations Centre
 CIC - Core Infrastructure Centre

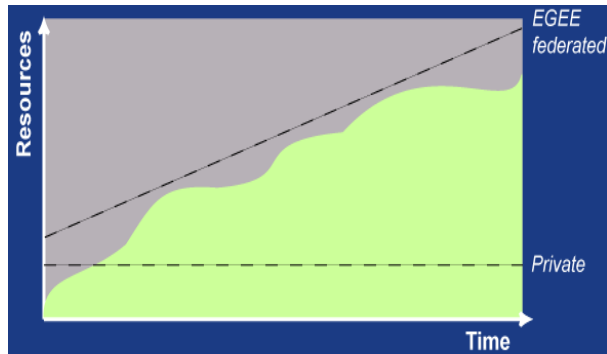
- The *grid* is flat, but
- **Hierarchy of responsibility**
 - Essential to **scale the operation**
- **CICs act as a single Operations Centre**
 - Operational oversight (**grid operator**) responsibility
 - Rotates weekly between CICs
 - Report problems to ROC/RC
 - ROC is *responsible* for ensuring 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**

- **Operation of Production Service: real-time display of grid operations**
- **Accounting information**
- **Selection of Monitoring tools:**
 - GIIS Monitor + Monitor Graphs
 - Sites Functional Tests
 - GOC Data Base
 - Scheduled Downtimes
 - Live Job Monitor
 - Gridlce – VO + fabric view
 - Certificate Lifetime Monitor



Private vs Federated Resources

For applications that must operate in a closed environment, EGEE middleware can be downloaded and installed on closed infrastructures

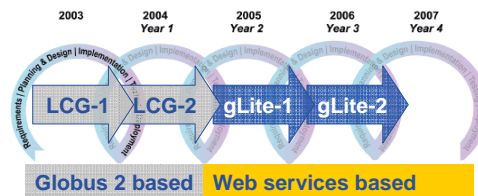


EGEE sites are administered/owned by different organisations

Sites have ultimate control over how their resources are used
Limiting the demands of your application will make it acceptable to more sites and hence make more resources available to you

Future EGEE Middleware - gLite

- Intended to replace present middleware (LCG-2)
- Developed mainly from **existing components**
- Aims to address present shortcomings and **advanced needs** from applications
- Regular, **iterative updates** for fast user feedback
- Makes use of **web-services** where currently feasible



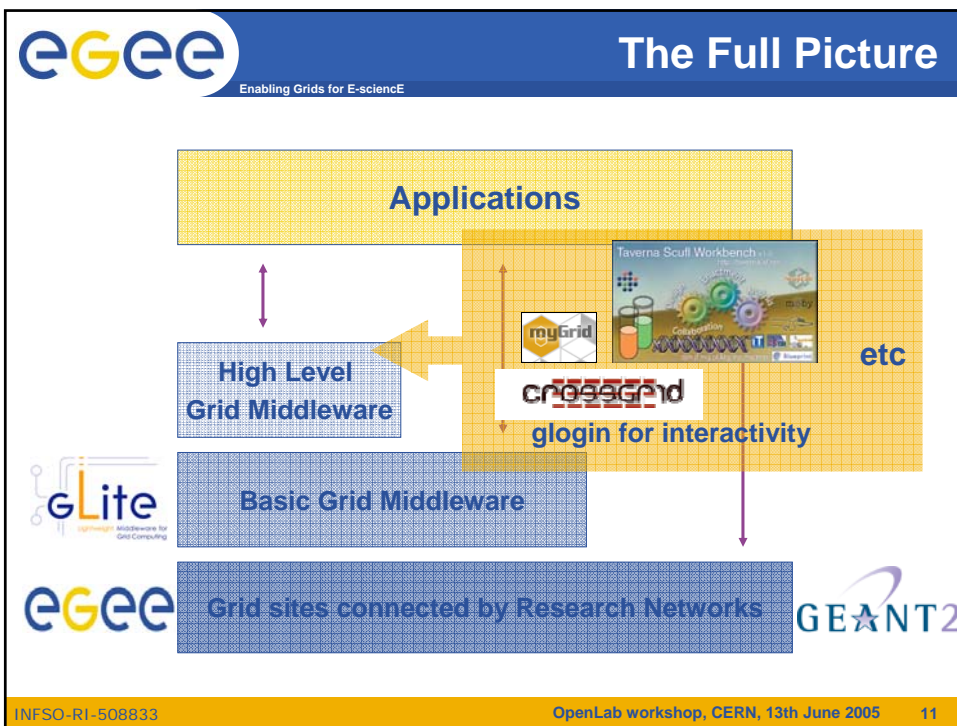
Application requirements <http://egee-na4.ct.infn.it/requirements/>

- The 1st release of gLite (v1.0) made end March'05
 - <http://glite.web.cern.ch/glite/packages/R1.0/R20050331>
 - <http://glite.web.cern.ch/glite/documentation>
- Lightweight services
- Interoperability & Co-existence with deployed infrastructure
- Performance & Fault Tolerance
- Portable
- Service oriented approach
- Site autonomy
- Open source license



- | | |
|---|--|
| <ul style="list-style-type: none"> • Computing Element <ul style="list-style-type: none"> - Gatekeeper (<i>Globus</i>) - Condor-C (<i>Condor</i>) - CE Monitor (<i>EGEE</i>) - Local batch system (<i>PBS, LSF, Condor</i>) • Workload Management <ul style="list-style-type: none"> - WMS (<i>EDG</i>) - Logging and bookkeeping (<i>EDG</i>) - Condor-C (<i>Condor</i>) • Information and Monitoring <ul style="list-style-type: none"> - R-GMA (<i>EDG</i>) | <ul style="list-style-type: none"> • Storage Element <ul style="list-style-type: none"> - glite-I/O (<i>AliEn</i>) - Reliable File Transfer (<i>EGEE</i>) - GridFTP (<i>Globus</i>) - SRM: Castor (<i>CERN</i>), dCache (<i>FNAL, DESY</i>), other SRMs • Catalog <ul style="list-style-type: none"> - File/Replica & Metadata Catalogs (<i>EGEE</i>) • Security <ul style="list-style-type: none"> - GSI (<i>Globus</i>) - VOMS (<i>DataTAG/EDG</i>) - Authentication for C and Java based (web) services (<i>EDG</i>) |
|---|--|

Now doing rigorous scalability and performance tests on pre-production service



EGEE Enabling Grids for E-scienceE

Open Source Software License

- **The existing EGEE grid middleware (LCG-2) is distributed under an Open Source License developed by EU DataGrid project**
 - Derived from modified BSD - no restriction on usage (academic or commercial) beyond acknowledgement
 - Approved by Open Source Initiative (OSI)
 - Same approach for new middleware (gLite)
 - New license agreed by partners is derived from the EDG license and takes into account feedback from the World Intellectual Property Office (WIPO)
- **Application software maintains its own licensing scheme**
 - Sites must obtain appropriate licenses before installation
 - EGEE will investigate policies for managing commercially licensed software

open source™

INFSO-RI-508833 OpenLab workshop, CERN, 13th June 2005 12



Why do people work with EGEE?

- **Transparent access to millions of files across different administration domains**
- **Low cost access to large computing resources**
 - Mobilise quickly large amounts of CPU on prompt basis
- **Produce and store massive amount of data**
- **Develop applications using distributed complex workflows**
- **Eases distributed collaborations**



How new communities join EGEE

- 1. New user community makes contact with EGEE application group**
 - For initial discussions <http://public.eu-egee.org/join/>
- 2. Clarifies needs and characteristics of application via a questionnaire**
- 3. Prepares submission to EGAAP (EGEE Generic Applications Advisory Panel) that makes recommendations taking into account**
 - Scientific interest of the proposed work and the grid added-value
 - Coordination and grid-awareness of the community
 - Agreement to the various EGEE policies and especially the security and resources allocation policies
- 4. Community and EGEE plan in greater detail the work to be performed**
 - Establishes a Memorandum of Understanding (MoU) signed by the community representatives and EGEE management formalising the engagements of each party
- 5. Progress of work is monitored regularly by the project**
 - Training
 - Porting of application (to GILDA, private infrastructure or production infrastructure)
 - Support for creation of virtual organisation and access to resources
 - Results achieved

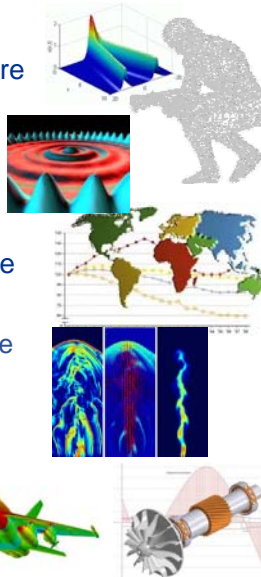
- **EGEE Industry Forum**
 - Raise awareness of the project in industry to encourage industrial participation in the project
 - Foster direct contact of the project partners with industry
 - Ensure that the project can benefit from practical experience of industrial applications



For more info:

<http://public.eu-egee.org/industry/>

- **EGEE I**
 - Large scale deployment of EGEE infrastructure to deliver production level Grid services with selected number of applications
- **EGEE II**
 - Natural continuation of the project's first phase
 - Opening up to a larger user community
 - increased multidisciplinary Grid infrastructure
 - More involvement from Industry
 - Extending the Grid infrastructure world-wide
 - increased international collaboration



- Ongoing **collaborations**

- with non EU partners in EGEE: US, Israel, Russia, Korea, Taiwan...
- with other European projects, in particular:
 - GÉANT
 - DEISA
 - SEE-GRID
- with non-European projects:
 - OSG: OpenScienceGrid (USA)
 - NAREGI



GEANT2



- EGEE as **incubator**

- 16 recently submitted EU proposals supported, among them:
 - Baltic states (Baltic Grid proposal to EU)
 - Latin America (EELA consortium on ALIS/CLARA networking)
 - Mediterranean Area (EUMedConnect)
 - China: EUGridChina

- From 1st EGEE EU Review in February 2005:
 - “The reviewers found the overall performance of the project very good.”
 - “... remarkable achievement to set up this consortium, to realize appropriate structures to provide the necessary leadership, and to cope with changing requirements.”
- Europe is strong in the development of e-Infrastructure also thanks to the initial success of EGEE
 - continue and reinforce this development in EGEE Phase II
- Applications are already benefiting from Grid technologies
 - more applications to come in Phase II
- Collaboration across national and international programmes is very important
 - will be extended in Phase II

- **EGEE Website**
<http://www.eu-egee.org>
- **EGEE Project Office**
project-eu-egee-po@cern.ch