



**ORGANISATION EUROPEENNE POUR LA RECHERCHE NUCLEAIRE  
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH**

Laboratoire Européen pour la Physique des Particules  
European Laboratory for Particle Physics

Position Paper

**Establishing an European Grid Organization (EGO)**

Geneva, June 2005

**Motivation**

Europe has invested heavily in developing Grid technology and testbeds during the past five years, with some impressive results. The European DataGrid (EDG) project, completed in early 2004, built a successful testbed to demonstrate the scientific potential of Grids. The European DataTAG project established new records in long-distance data transfers via international networks. The high-energy physics community was a major driving force behind these projects, motivated by the extreme computing, storage and network requirements of the Large Hadron Collider (LHC), due to start operation at CERN in 2007, for which Grid technologies have been selected.

The LHC Computing Grid (LCG) is now the world's largest scientific Grid, with over 130 participating sites in 31 countries. The EGEE project, which provides a coordinating framework for national Grids, has proved a vital catalyst for the success of the LCG, and is also helping other sciences, such as biomedicine and geophysics, to pilot Grid applications. EGEE is ensuring close coordination with developments in the US and other parts of the world, and as a result, Europe is now in a leading position in terms of deploying truly global science Grids. In parallel, the EU-funded DEISA project is developing Grid solutions for a supercomputer-based infrastructure in Europe, applicable to a complementary set of scientific problems.

As the Virtual Organizations (VOs) established by scientific communities move from a pilot phase of testing their applications on Grid infrastructures to a phase of more routine usage, it becomes increasingly urgent to ensure that reliable and adaptive Grid infrastructures are maintained, independent of project funding cycles. Without such continuity, scientific communities will hesitate to rely on this new computing paradigm, potentially jeopardising the large investments that have been made so far.

The current situation for Grid infrastructures is in some ways analogous to that of scientific networks fifteen years ago. Research networks started as incompatible testbeds, using different protocols and tools. Independent national initiatives led to the EU-EC promoting common standards, resulting ultimately in the GÉANT network infrastructure and the creation of the DANTE organization in 1993 to plan, build and operate networks. As a result, today a high-quality research network has become an essential part of the technological infrastructure for European and global scientific cooperation in many fields. A similar evolution needs to be planned now, to maintain Europe's pole position in global science Grids.

**Objectives of the European Grid Organization**

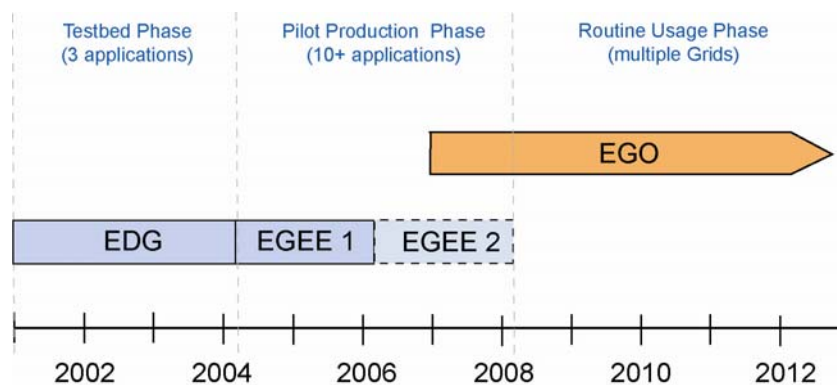
To ensure that Europe capitalises fully on its large investment in Grid infrastructures and applications, the objectives of the EGO should be to:

- Operate production Grid infrastructures for a wide range of scientific disciplines, in close collaboration with national Grid programmes and national and European research networks;
- Integrate, test, validate and package software from leading Grid middleware development projects, and distribute quality-assured middleware packages to sites supporting Grid infrastructures;
- Provide advice, training and support to communities wishing to migrate their applications to the general Grid infrastructure, and operate a test facility allowing such communities to evaluate Grid technologies;

- Promote Grid interface standards based on practical experience gained from Grid operations and middleware integration activities, in consultation with relevant standards organizations;
- Collaborate closely with industry as technology and service providers, as well as Grid users, to ensure a rapid and successful uptake of Grid technology by European industry.

## Launching the EGO

By 2007, when the EU 7<sup>th</sup> Framework programme will be launched, the LCG will be entering full production phase and begin to store and analyse data from the LHC for many years to come. The EGEE project will be in the middle of its second phase with a year to completion. Both the EGEE and DEISA projects will be supporting international Grid infrastructures, piloting applications for 10 or more VOs in various scientific disciplines. However, since these projects are necessarily time-limited, the VOs will need to have a clear indication of long-term commitment to support these Grid infrastructures. This is why, as indicated in the figure below, the EGO needs to be launched already in 2007, in time to provide the necessary continuity and stability of existing Grid infrastructures for their scientific users.



## Positioning the EGO

There are already a large number of organizations and projects active in developing Grid middleware, establishing standards for Grid technology and providing the underlying network connectivity for Grids. The figure overleaf places the EGO in the context of some of these existing activities.

The EGO's core mission corresponds to key roles played by the EGEE and LCG projects, namely providing a central operations support for a Grid infrastructure, and delivering middleware packages based on integration, validation and packaging of Grid middleware. Concerning operations, the EGO would run multiple Grid infrastructures for a wide variety of VOs, based on resources pooled by different scientific user communities. Branch offices of the EGO could be envisaged to support specialised Grid infrastructures where the relevant expertise is already at an established location.

Concerning the middleware, the EGO would not be an R&D facility or repository for Grid middleware, since there is already considerable and healthy competition in the academic sphere to ensure that a range of Grid middleware solutions can evolve. The UK e-science programme and Italy are probably the most predominant examples of large national investments in Grid developments. The EGO would instead evaluate best-of-breed solutions and re-engineer them if necessary to obtain production-quality software for use on the Grid infrastructures it operates. This is the role that EGEE's gLite middleware activity has successfully undertaken by putting emphasis on quality assurance rather than on new developments. The EGO would continue the active collaboration established by EGEE with projects such as VDT, Condor and Globus as well as national and EU-funded Grid R&D projects to foster the evolution of dependable grid components. The EGO should also collaborate closely with new initiatives such as C-OMEGA and OMII. A key element to achieve a more general Grid usage is better Grid security, which needs to be addressed by all Grid developers.

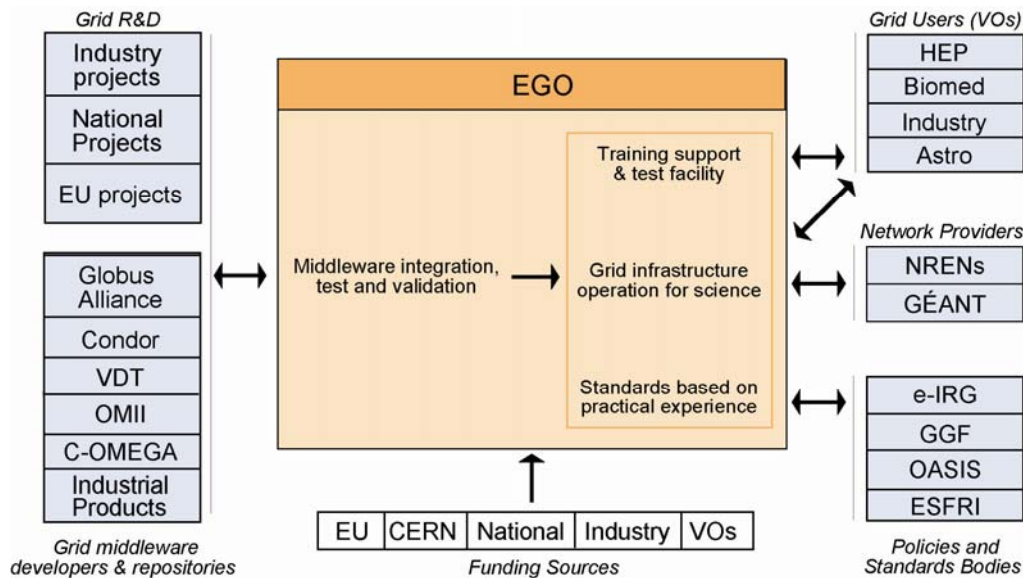
Providing training, support and a test facility for emerging communities is necessary to ensure the successful dissemination of the EGO's work. These activities should be carried out in collaboration with leading computing centres in Europe with established expertise to ensure adequate regional support.

The EGO would participate in known standards organizations such as GGF or OASIS, in order to influence Grid standards development based on practical experience from Grid operations and middleware testing and validation. The EGO could in addition provide a useful physical base for standards-related EU initiatives

such as the e-Infrastructures Reflection Group (e-IRG) or the European Strategy Forum on Research Infrastructures (ESFRI).

The EGO would work closely with industry as technology and service providers, as well as Grid users, to ensure a rapid and successful uptake of Grid technology by European industry. The EGEE industry forum, as well as the CERN openlab for DataGrid applications - a collaboration sponsored by leading IT companies to test advanced industrial solutions in a Grid environment - provide possible partnership models.

Finally, the EGO would coordinate with the network infrastructure providers (GÉANT/DANTE), to ensure optimal use of this infrastructure for Grid operations.



## Resources for the EGO

Based on estimates derived from current Grid activities, the EGO would require a staff of about 60 persons to manage its mission effectively: 20 persons for middleware re-engineering, testing and packaging, 20 persons for operations, 10 persons for training and support, and 10 persons for overall management, including managing relations with standards bodies and industry. These numbers are a first approximation and need to be revised in a detailed proposal.

To guarantee a smooth transition from earlier programmes the new organization should be hosted by an institution with the relevant experience and a proven track record in Grid operations, the real capital being top level experts trained during previous EU-funded projects and the ability to coordinate at an international level. The institution needs to provide the basic infrastructure as well as the necessary logistic, administrative and legal framework. The EGO should establish partnerships with leading national institutions to complement its work programme at the national level and with industry to transfer technologies in both directions.

In an initial five-year phase a large fraction of the EGO cost should be funded by the EU, complemented by nationally funded resource centres. Beyond the initial five-year time frame, a model where the end-user communities are charged for the services of the EGO on a non-profit basis should be developed, inspired by GÉANT/DANTE model.

## CERN as a host for the EGO

If this position paper proves of interest, CERN is ready to develop a detailed project proposal and implementation plan for the EGO and to host the new organization.

CERN is in an excellent position to implement the present proposal for a number of reasons:

- CERN is a recognised international meeting place for science and technology, with a track-record in IT innovation (“where the Web was born”), an extensive experience in coordinating large international projects involving hundreds or even thousands of scientists worldwide, and a proven ability to deliver advanced engineering solutions in collaboration with European industry.

- CERN is leading the major scientific Grid projects LCG and EGEE, with some 150 persons actively involved in Grid development and deployment on the CERN site, including some of the most talented software engineers and computing experts from Europe and the rest of the world.
- CERN's LHC experiments are an important driving force for Grid development and represent a highly organised worldwide community of very demanding users.
- CERN's status as an International Organization, combined with its attractive location in the heart of Europe, provides a globally competitive basis to attract and retain top international talent in Europe.
- CERN has a tradition of hosting new scientific and technical organizations in an early phase of their existence, such as the European Southern Observatory, and CERN's management is supportive of diversifying the portfolio of CERN's activities once the LHC becomes operational in 2007.
- CERN has a long tradition of close collaboration with leading firms in the IT industry, as CERN's openlab for DataGrid applications has successfully demonstrated for test and validation of cutting-edge commercial hardware and software solutions.
- CERN is currently pursuing plans to construct a new building with conference facilities which could host the EGO on the CERN site, and which would be integrated with the CERN Computer Centre, at the heart of the LHC Computing Grid. Funding for this is anticipated to come from private foundations.

## Conclusion

We propose to establish a European Organization as a service to the scientific community and to industry. The primary mission of the EGO will be to operate production quality Grid infrastructures, to re-engineer, integrate, test, validate and package middleware for different scientific disciplines, and to provide training and support to user communities from all sciences. The EGO will also work with standards bodies on establishing Grid standards based on practical experience from Grid operations, and will collaborate with leading IT companies to ensure an effective take-up of Grid technology by European industry, in particular by SMEs. The EGO will build on the success of current Grid projects such as EGEE and LCG, in order to maintain Europe's current leadership in global science Grids.

Considering that CERN

- has been a global player in science for fifty years and continues to be a world-leader,
- has a proven track record in multi-disciplinary projects,
- needs reliable Grids to produce scientific results from the LHC,
- leads today's largest and most successful Grid projects,
- has the required infrastructure,

CERN proposes to host the new organization at its premises in Geneva.

While the presents reflections of the e-IRG address the European Grid policies at a wider level and in much more detail, our proposal is perfectly in line with their ideas of creating an International Grid Organisation (IGO). It is not surprising that we are all coming to the same conclusion, given the experience gained over the past years in the Grid community. We prefer the name EGO to give a clear message of European leadership.

**Contact:** Wolfgang von Rueden  
 Head, IT Department  
 CERN  
 CH-1211 Geneva 23  
 email: wolfgang.von.rueden@cern.ch