

Table of contents

Editorial	p.1
EGEE technical overview	
✚ GENIUS, a grid portal dedicated both to users and science	p.2
✚ LCG: delivering the EGEE Production Service	p.3
✚ Try the Grid, using EGEE technologies	p.3
✚ Improving on the EGEE service: the daily operations portal	p.3
Applications	
✚ Important grid-added value for applications deployed on EGEE	p.4
News	
✚ News from the Industry Forum	p.6
✚ News from members	p.7
✚ Upcoming Grid events	p.7

Editorial

Now well into the 2nd year of EGEE and with a successful first review behind it, the project has demonstrated a clear direction and momentum. The Project Conference in Athens in April was a major success and, building on the review, has set course to meet the forthcoming challenges. Time is marching swiftly on and the end of the first phase of EGEE is already on the horizon. As I write, preparations are advancing at a pace in planning for the second phase of the project, anticipated to start in spring 2006.

A major driver of grid computing is to enable science that couldn't (easily) be done before and it is interesting to see, that with the expansion in the diversity of application communities in EGEE, tangible grid-added value is resulting in many areas. From handling peak demands in earthquake simulation, to the structuring effect for collaborations and on to providing access for important but neglected areas in drug discovery, the grid is being seen as having delivered real benefits. Access for such communities is greatly eased by tools like the Genius grid portal and the GILDA infrastructure.

The two major application communities in EGEE are stepping up a level in their respective service and data challenges. The LCG project is embarking on a service challenge as part of the build up to delivering grid services at scale and of production quality to the LHC experiments. In parallel the Bio-medical community are embarking on a major challenge in *in silico* drug discovery.

gLite, the first version of EGEE's next generation middleware, has been completed with substantial testing and validation on testbed infrastructures and the pre-production service. Rollout over the production service is about to commence.

The 4th EGEE Project Conference will be held in Pisa, Italy, from 24-28th October 2005. The theme of the conference will be creating and managing knowledge in the scientific community and building solutions which can be exported to industry and commerce. Specifically, the conference will ensure the project is ready for the focussed EU review scheduled for December. It promises to be an interesting conference. See you there!

Robin Middleton, CCLRC

GENIUS, a grid portal dedicated both to users and science

Genius Grid Portal: general presentation

The GENIUS Grid portal is a problem solving environment (<https://genius.ct.infn.it>) based on the European EGEE Project's middle-ware, which allows scientists to access, execute and monitor their own applications that exploit Grid resources only using a conventional web browser.

Most of the services provided by the EGEE middle-ware are currently exposed to end-users through rather complex Command Line Interfaces (CLI). There are tens of different commands with many options and rigid sequences and the description of jobs has even a dedicated language to be learned.

All this has the negative effect of discouraging many potential users from learning how to profit from the adoption of the Grid paradigm. A generic user would like to access the Grid services from all kind of appliances (desktop, laptop, PDA, cell phone, etc.) and from everywhere in the world, as he does with the World Wide Web. He would like to submit jobs and manage his data in a transparent way as if they all of them were local, i.e. independently of the physical computing and storage elements where jobs are executed and data are saved, respectively. It is just to realize these goals, exploiting the successful experiences made since the year 2000 within the ALICE Experiment at the CERN Large Hadron Collider, that, at the beginning of 2002, the Italian INFN Grid Project started the GENIUS Portal Project in conjunction with the Italian web technology company NICE srl.

Architecture and implementation of GENIUS

The present implementation of the GENIUS grid portal is carried out on top of the EGEE middle-ware services.

The layout of the portal can be described by a three-tier model (see Fig. 1):

- the client (top right in the figure): the user's workstation running a web browser;
- the server (left side of the figure): an EGEE User Interface (UI) machine (equipped with the EGEE middle-ware services able to

submit jobs and manage data on the Grid) which runs the Apache Web Server, the Java/XML portal framework EnginFrame (see next subsection) [10], developed by NICE srl, and GENIUS itself;

- the remote resources (bottom right in the figure): the Grid.

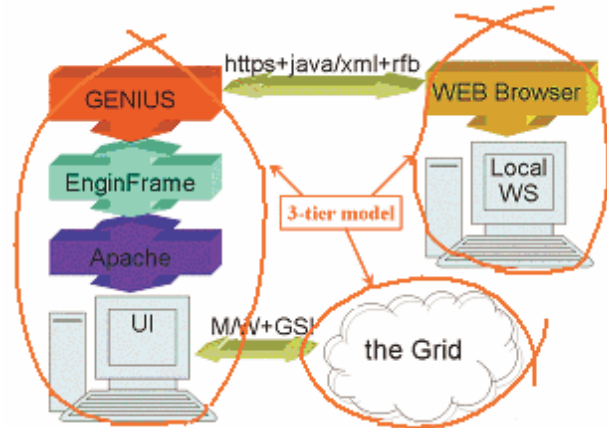


Fig. 1 - Three tier architecture of the GENIUS portal

Using the EnginFrame services the user can interact with files on the UI and, from there, the user can send jobs to the Grid and manage the data of the given Virtual Organization he belongs to. The use of the web interface eliminates any problem connected to the need of a particular Operating System and/or middle-ware running on the client, and to the locations themselves of the client and the server: the user can interact with the Grid from everywhere and with "everything".

Making use of the EnginFrame capability of services virtualization, GENIUS is transparently compliant with many versions of the EGEE middle-ware, including the latest version of brand new middle-ware gLite [11], and can be easily installed on a variety of Linux flavors, ranging from RedHat 7.3 to Scientific Linux .

GENIUS advantages

The current implementation of the GENIUS web portal is based on the middle-ware released within the context of the European EGEE Project. With GENIUS users can securely access the Grid, submit single and multi-jobs (including DAG's) prepared with graphic job creators to the distributed computing resources and take back their outputs. They can also interact, in a seamless way, both with the remote files located on the User Interface and with those of the Virtual

Organization they belong to as if they all were local ones.

Compared to other Grid portals available "on the market", GENIUS offers the following advantages:

- it is a complete production-ready environment which combines the concepts of "user portal" and "science portal";
- absolutely no client software needs to be installed on the user's workstation apart from the web browser with its usual plug-ins;
- it includes support for both single and composite jobs (including DAG's);
- interactive analysis and web access to personal spooling areas are possible.

Furthermore, due to the modularity and flexibility of EnginFrame, which acts as a general-purpose framework underneath GENIUS, the portal can be easily customized / adapted to interact with other Grids and/or new Virtual Organization using any other kind of middle-wares.

Thanks to the specific grant provided by the INFN Grid Project and to agreement between INFN and NICE srl, GENIUS code is open source and the EnginFrame license is free of charges for the academic and research world.

Acting as a simple and intuitive "gate" to access the Grid, the GENIUS portal brings with itself a tremendous dissemination power. In fact, GENIUS is the official portal of the GILDA Virtual Grid Laboratory, a complete t-Infrastructure for grid dissemination and training set up and managed by the INFN in the context of the EGEE Project.

EnginFrame has constituted a valid, secure, robust and flexible infrastructure on which we have built all of GENIUS Services used by EGEE community on various grid sites. In the meantime, the development of EnginFrame product, now at release 4.0, come with many new features that may help pushing further the vision of a usable and productive Grid interface. NICE and INFN are therefore planning the strategy for the next joint developments to turn GENIUS into a full-fledged "Science Gateway" for EGEE.

For further information, contact the EGEE Industry Forum or the authors.

R. Barbera (University of Catania and Istituto Nazionale di Fisica Nucleare, Italy), A. Falzone (NICE srl, Italy)

LCG: Delivering the EGEE Production Service



For EGEE, LCG provides more than 100 sites and is responsible for management of the overall production service (running LCG middleware at the beginning of the project) through a system of Core Infrastructure Centres and Regional Operation Centres. In return, EGEE provides LCG with gLite, its state of the art lightweight middleware solution, currently being deployed over the production service.

Extract of 'The EGEE Production Service – LHC Computing Grid'

This article about LCG and one about gLite will be soon available on the Industry Forum website.

Try the Grid, using EGEE technology

For a brief introductory walkthrough of the grid, try the GILDA Grid Demonstrator (<http://public.eu-egee.org/test/>).

For further information concerning GILDA, see <https://gilda.ct.infn.it/>

Improving on the EGEE service: the daily operations portal

CC-IN2P3

EGEE production service comprises 120 sites and 14,000 CPU, with a storage availability of roughly 5PB. Around 1100 users are registered through applications on the grid. The activity level to deal with is therefore quite high [1].

Monitoring of grid resources and core services: how to ensure an optimum availability of the grid

The current daily operations of EGEE are built mainly on the basis on weekly shifts insured by an international team from five of the federations within the project: CERN, France, Italy, UKI, and Russia. The main objective of this team, nicknamed CIC-on-duty (Core Infrastructure Centre on-duty), is to monitor grid

resources as well as core services to ensure an optimum availability of the grid. At the outcome of each of the team's shifts their feedback is analysed in weekly meetings to trigger some technical debriefings and proactive action from the EGEE operations management and from the federations' management.

Procedures and tests have been developed to monitor grid availability and functionality at first mainly by CERN. The monitoring tools check the status of the resources in production declared in a centralised database, developed and maintained at RAL, UK. Provided by FZK, Karlsruhe (Germany), since mid April 2005, GGUS tool – Global Grid User Support - ensures the tracking and the management of operational incidents. The collection of metrics is under way to estimate the level of reactivity of all the operational structures, in particular applications experts, site managers, grid experts and federation managers.

The CIC portal: a unique entry point for operational aspects

Regular EGEE operations result in the integration of several key components: centralised database containing site information, testing and monitoring services of the grid infrastructure, and end-user support. The Core Infrastructure Centre portal (<http://cic.in2p3.fr>) hosted and developed at CC-IN2P3 (<http://cc.in2p3.fr>), aggregates all these sources of operational information into a single, consistent and coherent tool.

This portal provides operational information and tools for several actors of EGEE grid, including grid-wide operations team, regional operations coordinators, site managers and end users.

For the day-to-day operations of the whole grid, the portal provides a single point for contact information of the operations people in every region, a detailed view of the operational status of core grid services (like workload resource management services, file catalogue services, virtual organisation management services, etc.), a dashboard of the operational status of the sites based on the results of the monitoring and testing service, an overview of the open incidents and their status, the collection of metrics regarding the operational performance of the infrastructure, and finally a tool for sending operational-related information to all the people involved in the EGEE grid.

Current end-users can find it useful for obtaining information about the set of sites providing services to his virtual organization, contact information on every site of the grid, and the operational status of each one of them. New virtual organizations are able to express their operational requirements and based on them, site managers can express interest in supporting the new grid users: the portal is then also useful as an operations planning tool.

Weekly operational reports are automatically generated and available through the portal for approval by each regional operations coordinator. Those reports are reviewed in the grid-wide weekly operations meeting and metrics are extracted from them regarding the availability of the sites for the grid.

Operating a service infrastructure the size and complexity of EGEE is an ongoing challenge. As little prior experience to base this work on exists, a dedicated team, coordinated by CC-IN2P3, was setup for continuously refining the processes and tools based on the feedback from all the actors of the EGEE operations area. All this valuable information is used to improve the service, which is the ultimate goal of the project.

EGEE is undergoing in 2006 an increase in the variety of the scientific disciplines supported by the infrastructure. With the variety of virtual organisations and their applications on one hand, and the dramatic increase of sites in other regions of the world currently not included in EGEE on the other hand, the coordination of the operations of the whole infrastructure and the associated tools will be of the utmost importance for reaching the quality of service the project aims at.

[1] Activity SA1, Quarterly report, SA1 - Quarter 5, Alistair Mills, 09/06/2005.

H. Cordier, F. Hernandez, CC-IN2P3

EGEE Applications

Important grid-added value for applications deployed on EGEE

Since the beginning of the EGEE project, several applications from various fields have been deployed in the Generic Application

branch of the EGEE Na4 activity (application identification and support). These applications are issued from the following communities: earth sciences, drug discovery, computational chemistry and cosmology.

During the third EGEE Conference, the EGEE Generic Applications Advisory Panel (EGAAP) held its third meeting, which was mainly devoted to the review of all the applications which have formally been approved to run on the EGEE infrastructure. The Panel paid special attention to identify the new potentialities that the grid paradigm enabled for these applications (so called grid-added value). Each application was able to convince EGAAP of very valuable grid-added value that, in most cases, goes way beyond the opportunistic usage of "free" CPUs.

In fact, many results were simply unattainable without the Grid:

- In Earth Science, a key point is the ability to securely and transparently access millions of files, regardless of their physical location, across different administrative domains. The data have been already processed and analyzed in the past. However in an easier and more extensive way EGEE permits to re-process the full ESA 10-year data archive using three different new algorithms, to compare and validate their results. This unique feature provided by the Data Grid infrastructure has brought about a first-time interactive collaboration among several Institutes and has led to the publication of new scientific results on global atmospheric Ozone.
- The possibility to mobilize quickly very large amounts of CPU on very *prompt peak use* basis was clearly illustrated by the computations performed by the Solid Earth group who computed in less than 24 hours (instead of a couple of weeks with local facilities), their original and ambitious milestone, all parameters relevant for major earthquakes. The final purpose is to explain precisely the source mechanisms and to better understand the geodynamical processes at work in some regions (Sumatra, West Indies,...). A better knowledge of all seismic sources is very valuable in the long-term prospect of earthquake prediction.

- *Distributed software* development of complex workflow is a difficult problem both for technical and human-related reasons. The Grid has been chosen by the Computational Chemistry community to solve it because they find it so hard to manage it otherwise. The best proof is given by their internal commitment: is community is investing ~20 FTE in grid-based software development and ask typically for the very modest 20 CPU in return. Allowing the various groups in charge of the various steps of the workflow to participate in the development and production on an equal footing thanks to the Grid paradigm has been found to be extremely efficient and motivating.
- ALL applications insisted very strongly on the benefits of *the collaboration-building tool* which the Grid represents. Thanks to the Grid, it is possible to **reflect the distributed nature of a scientific collaboration into its computing and storage model**. This proves to be highly beneficial on a human and financial point of view since it maximizes the spirit of collaboration and the flow of resources. It also allows, as pointed out by the Hydrology application, to include partners from less favoured countries (Tunisia in this case) in a way that was not possible before.
- Finally, one should not neglect the fact that minimizing the cost of access to large CPU resources is also key for some applications where *by definition*, the corresponding appropriate funding will *never* materialize otherwise. This is for instance the case for Drug discovery for neglected diseases, where the word "neglected" in the title tells immediately why it is the case. In such applications, many thousands of human lives can be hopefully be saved by offering low cost access to large CPU quantities than will never be accessible otherwise.

All communities expressed their happiness to run their applications on EGEE. This shows that the hard work they have to invest to become familiar with this new techniques, which in the present state are far to be as user friendly as one could wish, is more than compensated by the benefit they are driving out of it for their scientific work

Encouraged by the positive developments mentioned above and by the very significant improvements made since their first proposal, EGAAP recommends to approve EGRID as an official EGEE generic application. The main role of EGAAP is indeed the evaluation of all proposals from external groups, to make prioritized recommendations to the Na4 and to the EGEE management. EGRID, financial application, now includes several important collaborators coming from the academic and industrial financial sector, included in the large international collaboration. Many grid-added values features mentioned above directly apply to this project: vast volume of distributed data (coming from the various stock exchanges) across different administrative domains, distributed nature of the collaboration, important of peak capacity due to the cyclic nature of the financial world, collaboration building between many partners that have never worked together. This application can also be used as a test case for the new gLite middleware, since some of its new functionalities are needed for this application.

Finally EGAAP expresses its interest to the potential new application in the nuclear fusion domain and encourages the group to respond to the call for proposal EGAAP will issue in the summer.

For further information concerning the applications deployed on EGEE, see also <http://public.eu-egee.org/sheets/> and <http://public.eu-egee.org/news/>.

Guy Wormser, In2p3 (wormser@lal.in2p3.fr)

News

News from the Industry Forum

Industry Forum meeting in Athens

The third EGEE Industry Forum meeting took place on Thursday 21st April 2005 in Athens, Greece, during the third EGEE conference. The presentations of the participants are accessible on:

<http://indico.cern.ch/conferenceDisplay.py?confId=0513>

Industry Forum, fr-grid and Gus'G meeting in Lyon

The Grid user's Group of the Aristote association [1], the fr-grid consortium [2] and the EGEE industry forum organized on June 17th, in Lyon France, a working day on "Operational and Industrial uses of Grids".

The day started with a tutorial on EGEE "from research to deployment" by Ruediger Berlich from Forschungszentrum Karlsruhe, Germany. After a summary of the requirements for the large applications targeted in EGEE, he exposed the technical aspects of the project not forgetting issues to be raised to run an operational service.

The other sessions alternatively targeted the "users" or the "providers" point of view.

Large companies (IBM and Platform) presented their view of the Grid based on a high level industrial expertise, whereas software companies (GridXpert and C-S) tackled the problem of industrialization of research tools (e.g. middlewares).

On the other hands, users' presentations - Finance applications (BNP, ECP), and PEP which is a consortium of SME's- showed that there are still many technological barriers to break. Most important is security, but interoperability and standards should not be forgotten.

One major difference between EGEE and operational or industrial uses is that while EGEE aims at deploying a platform as a support for applications, operational grids grow dynamically on a project, collaboration or commercial basis. It is the application -through a Virtual Organization of partners- which "makes" the underlying Grid. This makes more challenging the emergence of generic solutions in a much more heterogeneous context. More information is available on the fr-grid web site [3].

[1] <http://www.aristote.asso.fr/GUSG>

[2] <http://www.fr-grid.org> "french grid enabling initiative"

[3] <http://www.fr-grid.org/17-Juin/17-juin.html>

Philippe d'Anfray, CEA (philippe.anfray@cea.fr)

Industry Forum website is to be revamped

The EGEE Industry Forum will take advantage of the summer to update its website <http://public.eu-egee.org/industry/>

Next Industry Forum Meeting

The next Industry Forum meeting will take place during the EGEE Conference in Pisa, October 26th.

A round table is planned, during which some industrials will share with us their own experience and their expectations about Grid.

For more information, contact brun@mas.ecp.fr

News from members

The Fraunhofer Institute SCAI

In several European and German projects, like Unicore, UnicorePlus, Eurogrid and Grip, the grid middleware UNICORE (www.unicore.org, unicore.sourceforge.net) was established and is now available as open source. It has proven to be a robust and secure grid system, so it is ready to use in production. Therefore the German weather service (DWD) decided to build mostly all of their grid initiatives upon UNICORE. Partner institutes willing to use the compute resources of the DWD have now the opportunity to access them via UNICORE.

To guarantee a smoothly flow of UNICORE use, the Fraunhofer institute SCAI (partner in the EGEE project) together with T-System SfR provide user and administration support for the German weather service and user from partner institutes as the "Bundesanstalt für Wasserbau" (BAW). A support chain utilizing also the user help desk of the DWD was established to provide all kinds of support up to collaboration with the Unicore Technical Board.

SCAI also offers implementation of all kind of tools and plug-ins within the UNICORE concept. In addition to chemistry application plug-ins for the graphical user interface SCAI prepared an ssh-like access within the UNICORE environment. It was build to allow BAW user to connect to computing resources of the DWD in an interactive way utilizing the single sign on mechanisms and services of UNICORE.

In the future UNICORE-GS, a WSRF compliant version of UNICORE, will give the possibility to join UNICORE with other WSRF grid middleware's like gLite, to create a grid, which is suitable for a wider range of use. In this way limitations of grid systems may be overcome and they benefit of each other. For example,

having a UNICORE like fully integrated graphical grid client, being able to create, submit and control jobs and handle data within a gLite and a UNICORE environment, will lead to better acceptance of grid computing in new areas. On the other side sophisticated data handling and virtual organization management like in the EGEE environment are developments that could be additional to UNICORE.

*H. Schwichtenberg and Marc Lob, SCAI
(schwichtenberg@scai.fraunhofer.de or
marc.lob@scai.fraunhofer.de)*

Upcoming Grid events

3rd International Summer School on Grid Computing 2005

Location: Vico Equense, Naples, Italy

Date: 10-22 Jul 2005

The curriculum of the school will include lectures on the main topics of Grid development and technology, as well as emerging key grid applications. Lectures will be given in the mornings with practical exercises taking place in the afternoons.

For more information please visit:

<http://www.dma.unina.it/~murli/GridSummerSchool2005/>

5th WSEAS International Conference on Simulation, Modeling and Optimization

Location: Corfu Island, Greece

Date: 17-19 Aug 2005

This event will cover a range of Grid topics including the Mathematical Foundation of Grid Computing; Management and Security Aspects on Grid Computing; Human Resources and Grid Computing; and Architectures for Grid Computing.

For more information please visit:

<http://www.worldses.org/conferences/2005/corfu/smo/grid/index.html>

CERN School of Computing 2005

Location: Saint Malo, France

Date: 4-17 Sep 2005

Since the early eighties CERN has organized every year the CERN School of Computing (CSC), in one of the CERN Member States. CERN Schools of Computing are open to postgraduate students and research workers with a few years of experience in elementary

particle physics, in computing or in related fields. The participants come mainly from the CERN Member States or from laboratories in countries associated with CERN. However, a fraction of the students come from outside the particle physics community, generally attracted by the advanced topics that are taught. Attendance ranges from 70 to 80 students, typically of 15 to 25 different nationalities. The schools last two weeks, and are generally organized on the campus of a university or in a hotel with close networking connections to a university or laboratory.

The CERN School of Computing 2005 is organised around 3 thematic tracks (Grid Technologies, Software Technologies, Physics Computing, each track comprising lectures and exercises.

For more information please visit:

<http://www.cern.ch/CSC>

EuroPVMMPI 2005

Location: Capri, Italy

Date: 18-21 Sep 2005

The 12th European PVM/MPI Users' Group conference will be a forum for the users and developers of PVM, MPI, and other message-passing programming environments. They will have the opportunity to meet each other, share ideas and experiences, and meet members of the PVM and MPI teams.

For more information please visit:

<http://www.pvmmpi05.jeanmonnet.unina2.it/>

GGF 15

Location: Boston, United States

Date: 9-12 Oct 2005

GGF 15 will likely be the Global Grid Forum's global plenary event. Working groups and research groups will also have the opportunity to meet.

eChallenges e-2005 Conference and Exhibition

Location: Ljubljana, Slovenia

Date: 19-21 Oct 2005

This is the fifteenth in a series of Annual Conferences supported by the European Commission, which regularly attracts over 500 delegates from commercial, government & research organisations around the world to share knowledge, experience, lessons learnt and good practice.

Focusing on eBusiness, eGovernment, eWork, eEurope beyond 2005 and ICT take-up by

SMEs and International Co-operation on IST, the goal of e-2005 is to stimulate take-up of Research & Technology Development (RTD) results by industry, in particular SMEs, and the European public sector.

For more information please visit:

<http://www.echallenges.org/2005/default.asp?page=home&#>

Anyone interested in joining the EGEE Industry Forum should contact either Christian Saguez (christian.saguez@ecp.fr), or Guy Wormser (wormser@lal.in2p3.fr) or Myriam Brun (brun@mas.ecp.fr)



EGEE
Enabling Grids
for E-science