

The EPIKH Project

Roberto BARBERA¹, on behalf of the EPIKH consortium
¹University of Catania, INFN, and Consorzio COMETA,
 Via S. Sofia, 64, Catania, 95123, Italy

Tel: +39 095 3785504, Fax: +39 095 3785231, Email: roberto.barbera@ct.infn.it

Abstract: Nowadays, e-Science and e-Infrastructures are considered key enablers of the progress and sustainable development of a country and are concrete means to address the problems of the “digital divide” and the “brain drain” which are endemic in large parts of the world. In the context of its 6th and 7th Framework Programmes, the European Commission has already co-funded several projects to stimulate and foster e-Science and Grids well outside its borders and in several parts of the world such as Asia, Latin America and the Mediterranean. However, the adoption of the “Grid paradigm” and the effective usage of e-Infrastructures require a capillary activity of knowledge dissemination and training to help scientists to make use of distributed computing capabilities for/in their scientific applications. This paper presents the EPIKH project that will start on the 1st of January of 2009 and discusses how EPIKH addresses the issues outlined above.

Keywords: e-Infrastructure, Training, Education, Grid Computing.

1. Introduction

1.1 – (Some of) The “divides” of the world

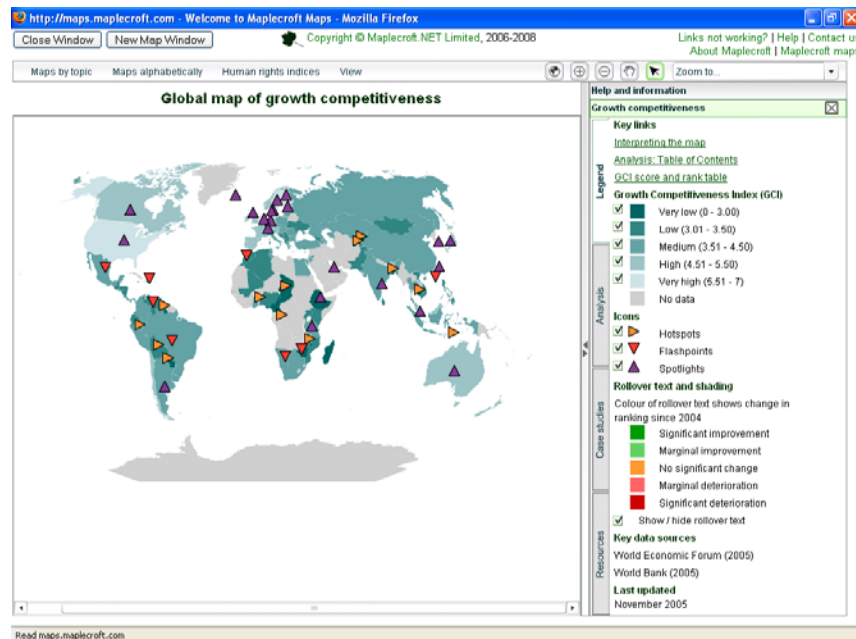


Figure 1: Geographical distribution of growth competitiveness in the world.

Almost 250 years after the publication of the illuministic and equalitarian theories of J. Rousseau [1], today’s world still suffers from a very uneven distribution of “opportunities”. Figures 1, 2 and 3 show, respectively, the geographic maps of growth competitiveness, education attainment, and digital inclusion [2].

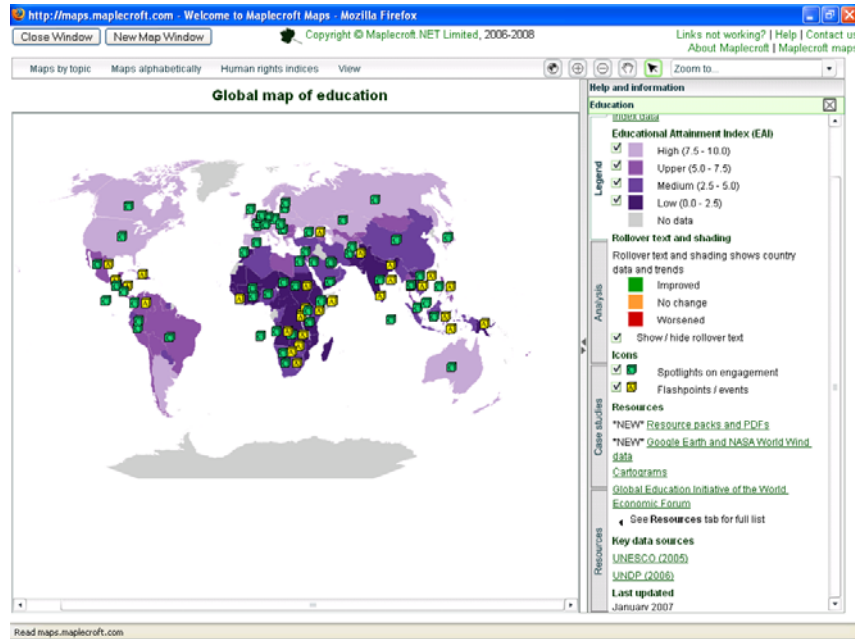


Figure 2: Geographical distribution of education attainment in the world.

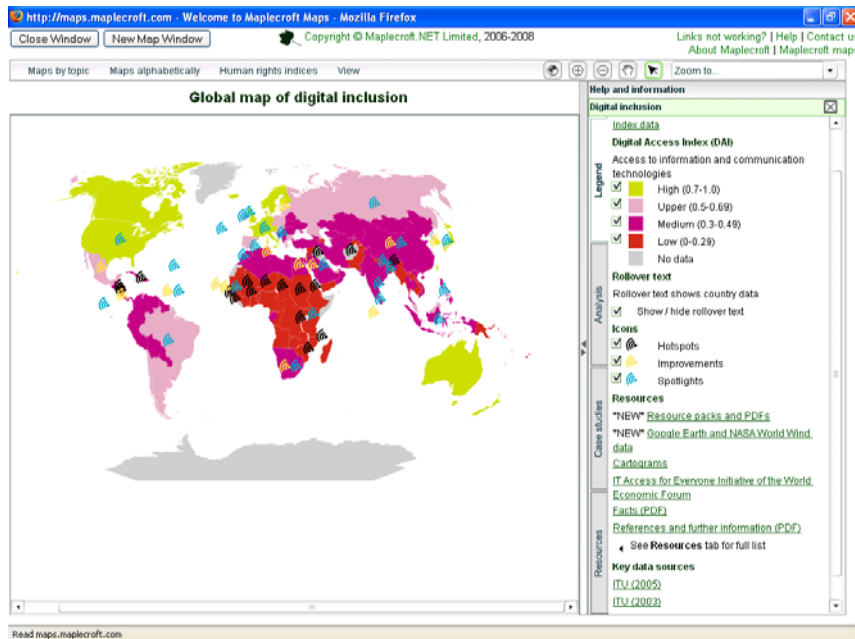


Figure 3: Geographical distribution of digital inclusion in the world.

Looking at the maps above, two considerations are in order:

1. First, there is a quite strong correlation among the three quantities reported in the figures: this means that several factors contribute “in parallel” to make the gap between more advanced and less advanced countries deeper and deeper every day inducing endemic problems like massive migrations, under-development, alienation, and poverty. Along the same reasoning, fighting against more than one problem at the same time could then help to alleviate the others. As reported by the Education and Training Task Force (ETTF) [3] of the e-Infrastructure Reflection Group (e-IRG) [4], country studies carried out both by the Organisation for Economic Co-operation and Development (OECD) [5] and the World Bank [6] have confirmed an obvious correlation between investment in education and quality of life and GDP.

2. Second, there are several centres of excellence and “hot-spots” in many of the countries suffering from the above mentioned “divides”.

It is just in these respects that the European Commission funded “Exchange Programme to advance e-Infrastructure Know-How” (EPIKH) project, as it will “unfolded” and expanded throughout this paper, aims at addressing the issues of digital divide and education on new technologies to foster scientific competitiveness/collaboration between several regions of the world so reducing both the “digital divide” and the “brain drain”. This will be done identifying and gathering, since the creation of the consortium meant to manage the project and the preparation of the proposal, a number of trusted and renowned partners operating some of the centres of excellence mentioned above.

1.2 – The European and “global” Research Areas

At the onset of the 21st Century, the way scientific research is carried out in many parts of the world is rapidly evolving to what is nowadays called e-Science, i.e. a “scientific method” which foresees the adoption of bleeding-edge digital platforms known as e-Infrastructures throughout the process from the idea to the production of the scientific result. The e-Science vision is depicted in figure 4.

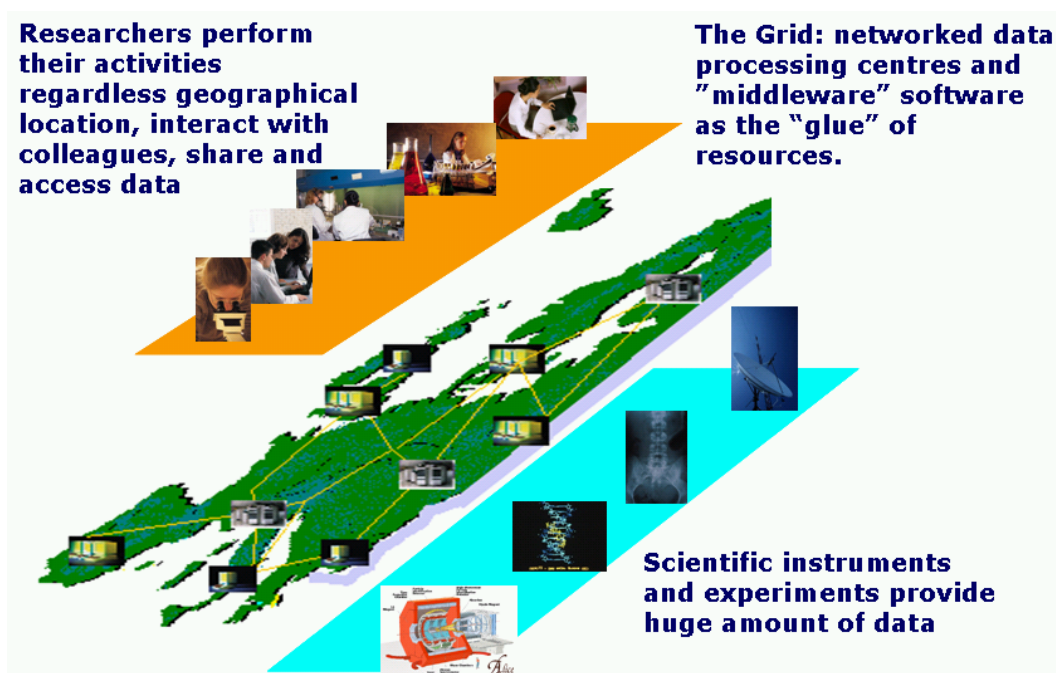


Figure 4: The “vision” of e-Science.

Scientific instruments are becoming increasingly complex and produce huge amounts of data which are in the order of a large fraction of the whole quantity of “information” produced by all human beings by all means. These data are often relative to inter/multi-disciplinary analyses and have to be analyzed by ever-increasing communities of scientists and researchers, called Virtual Organisations (VOs), whose members are distributed all over the world and belong to different geographical, administrative, scientific, and cultural domains. The emerging computing model which is being developed since a decade or so is what is called the “Grid”, i.e. a large number of computing and storage devices, linked among them by huge-bandwidth networks, on which a special software called middleware (intermediate between the hardware and the operating system and the codes of the applications) is installed and make them behave as a single huge “distributed” computer

which “dissolves” in the fabric of the Internet and can be accessed ubiquitously through virtual services and high-level user interfaces.
 The Grid and the underlying network constitute the e-Infrastructure (see figure 5).

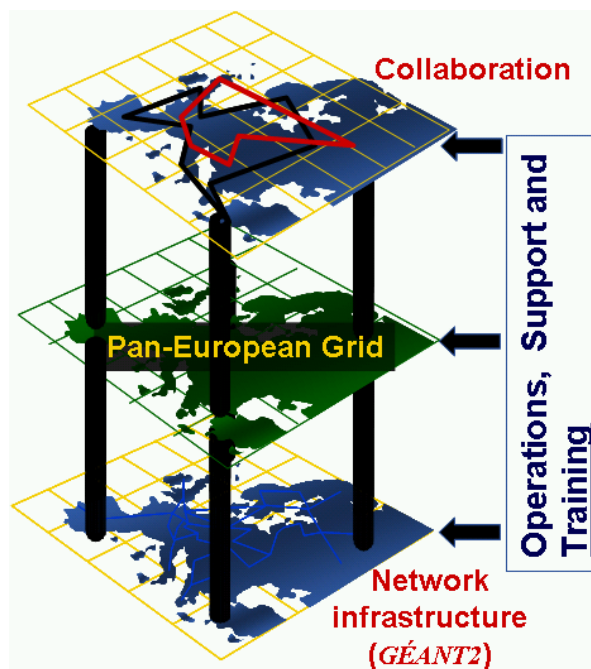


Figure 5: e-Infrastructure model of the European Research Area.

The European Commission (EC) is heavily investing through its Framework Programmes in e-Infrastructures and this platform is by now considered as one of the key enablers of the European Research Area (ERA). In fact, at the top of the three-layers model of an e-Infrastructure there is the most important “network”: the human collaboration among scientific communities of researchers that work together on unprecedented complex multi-disciplinary problems whose solutions have very positive effects for the society and the progress at large.

The European Research Education Network, which connects about 3900 Institutions in more than 40 countries in the continent, and support the work of more than 30 millions of students, teachers, and researchers, is realized in the context of the GÉANT2 project [7] coordinated by DANTE [8].

The pan-European Grid is realized by flagship projects like EGEE [9] and DEISA [10].

In order to bridge the digital divide between Europe and other less developed regions of the world, in the last 3-4 years the European network and the European Grid have been expanded well outside the borders of the “old continent” in the context of several successful EC co-funded projects such as ALICE (network project for Latin America) [11], EUMEDCONNECT (network project for the Mediterranean region) [12], GÉANT2-ERNET (network collaboration for India,) [13], ORIENT (network project for China) [14], SEE-REN (network project for the South-Eastern European region) [15], TEIN2 (network project for the Asia-Pacific region) [16], BalticGrid (Grid project for the Baltic region) [17], EELA (Grid project for Latin America) [18], EUAsiaGrid (Grid project for South East Asia and Australia) [19], EUChinaGRID (Grid project for China) [20], EU-IndiaGrid (Grid project for India) [21], EUMEDGRID (Grid project for the Mediterranean region) [22], and SEE-GRID-SCI (Grid project for the South-Eastern European region) [23].

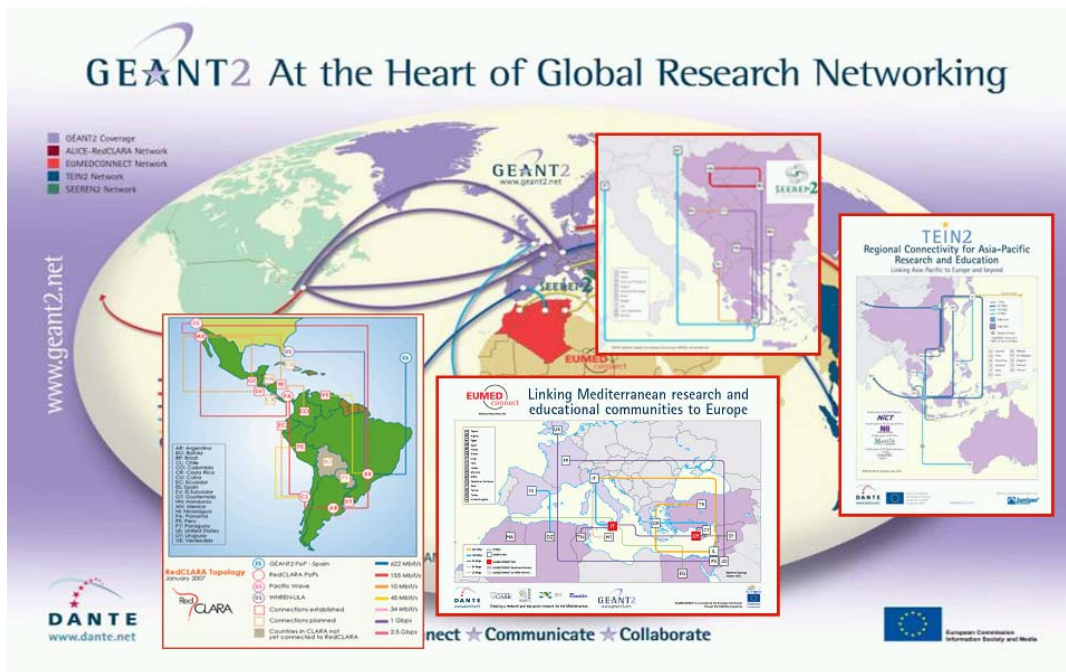


Figure 6: The “global” network.

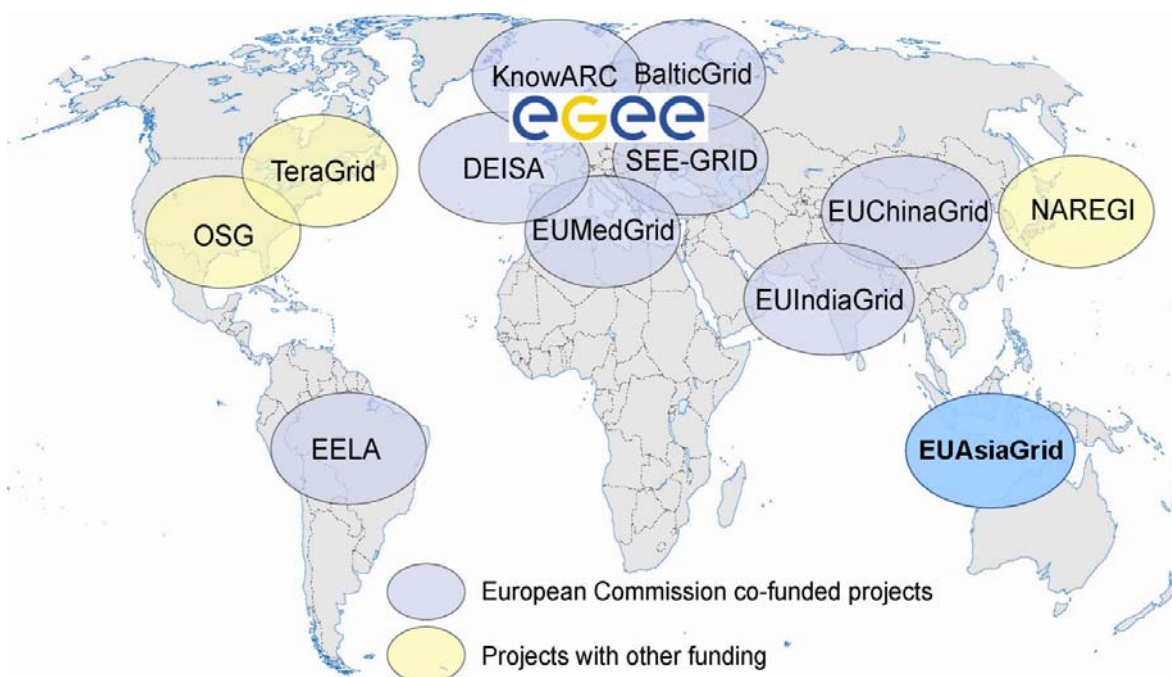


Figure 7: The “global” Grid.

All the so-called “regional” Grid projects mentioned above share the same kind of work plan whose “virtuous cycle” is depicted, in a graphical way, in figure 8.

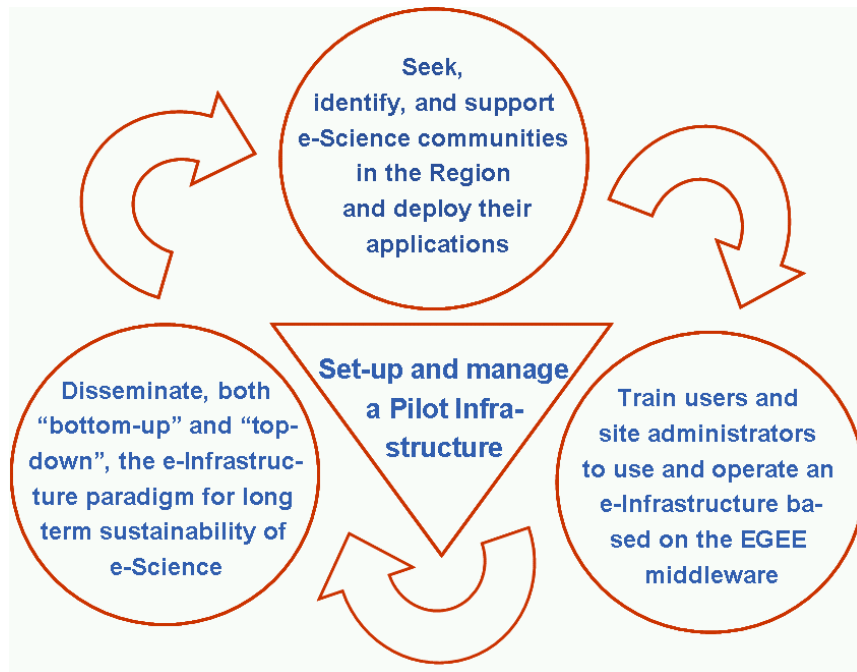


Figure 8: The “virtuous cycle” of regional Grid projects.

Dissemination activities, which all these Support Actions focus on, help to identify new communities of scientists and researchers and those applications. Users are then trained to access and use a pilot e-Infrastructure set-up to let applications to be deployed and developed and then be used as exemplar use cases and success stories in further dissemination events.

So far, however, due to the fact that in the context of many of the regional projects Grid infrastructures have been deployed from scratch, training activities have been focused on induction courses where Grid middleware services and their interplay have been the main part of the curriculum taught. With EPIKH we plan to shift the educative mission and develop an intensive and diversified training programme where Grids are not the “goal” but just the “mean” to develop “e-scientific” applications and gather together scientific communities from four continents letting them access production quality e-Infrastructures distributed world-wide. The EPIKH joint research programme and its relevance for the progress of science in those areas will be developed in the following sections.

2. Objectives

The EPIKH project [24] will be led by the Consorzio COMETA [25] which will act as coordinator of 22 more Organisations belonging to 4 continents. All the partners of EPIKH are listed in [26].

The strategic objectives of EPIKH are to:

- O1) Reinforce the impact of e-Infrastructures in scientific research defining and delivering stimulating programme of thematic educational events, including Grid Schools and High Performance Computing courses;
- O2) Broaden the engagement in e-Science activities and collaborations both geographically and across disciplines.

These ambitious goals translate into the following specific actions:

- A1) Spreading the knowledge about the “*Grid Paradigm*” to all potential users: both system administrators and application developers through an extensive training programme;

- A2) Easing the access of the trained people to the e-Infrastructures existing in the areas of action of the project;
- A3) Fostering the establishment of scientific collaborations among the countries/continents involved in the project.

3. Methodology

The exchange programme will be implemented in phases that will alternate between each other along the time span of the project. First, a selected team of brilliant young researchers will be invited to visit the EU partners of EPIKH for periods in the order of 1 month to get trained as trainers of Grid technology for what concerns both administration of Grid sites and application support (“gridification”).

Then, at least 2 Grid educational events per year and per continent (Africa, Asia, and Latin America) will be organized and run. Each of them will consist of the following phases:

1. Organization phase (2-3 months):
 - based on available *ex-ante* analysis, 1 or 2 complementary scientific disciplines, strategic for the region of the world where the EPIKH educational event are expected to take place, will be selected;
 - a “call for applications”, belonging to the scientific domains chosen, will be open and largely advertised; interested users and communities will be requested to fill detailed questionnaires about the scientific and technical details of the applications they want to propose;
 - received applications will be selected against public and well described criteria by a Selection Committee that will be appointed by the management of EPIKH;
 - authors of selected applications will then be invited to come to the EPIKH Grid school;
2. Execution phase (1 month):
 - the school to port the selected applications on a Grid environment will be preceded by a school for Grid system administrators; the school, that will have a duration of 7-10 days will have the objective to form skilled people and set-up the Grid training infrastructure (t-Infrastructure) to be used for the subsequent school for application porting;
 - in the next 14-16 days, authors of selected applications and tutors (both from EU and non-EU countries participating to EPIKH) will get together and “gridify” the applications; the school will have “theoretical”, where the Grid principles and applications’ scientific details will be explained, followed by “hands-on” sessions where tutors will work together with applications’ experts to adapt their codes to run on the t-Infrastructure; trainees will also be taught how to access the e-Infrastructure present in their continent after the end of the school;
 - at the end of the schools, a scientific workshop of 2-3 days will be held where the problems of the scientific disciplines chosen as topics of the school will be debated and how the adoption of the Grid “paradigm” can represent a big advantage. Decision makers and stakeholder from politics and industry will be invited whenever and wherever possible and applications “gridified” during the school will be shown as use cases and success stories; this will allow to pursue a combined “top-down/”bottom-up” approach for the long-term sustainability of e-Infrastructures in the region.

Local network conditions (bandwidth, latency, etc.) permitting, all the lectures of the schools and the presentations of the workshop will be broadcasted live and saved in e-

Learning Management Systems (LMS) to allow more users to benefit, with their own pace, from the EPIKH educational event, even when it will be over.

The relevance of the proposed joint research programme, “framed” in what has been said in Section 1, is clear:

- Groups active in strategic scientific domains will be indentified early on and put in contact with colleagues in Europe and other parts of the world widening, at a global scale, the diffusion of scientific (in)formation and best practices;
- Grid technology will be used as a powerful “tool” to impart/improve education on e-Science;
- The t-Infrastructure built during the first school can be left as a seed of a Grid infrastructure in regions where e-Infrastructure sites are not yet present;
- At the end of the second school, more applications will be ready to run on large e-Infrastructures and more users will get aware of the benefits of this technology for the progress of science and society.

In order to implement its work plan and reach its objectives, EPIKH will mobilize in more than 500 secondments about 115 people for a total of over 650 researchers-months, not counting of course the people that will be outreached by the project. These are huge figures, even larger than those of many much richer projects, confirming the strong interest of the 4 continents of the world involved in the project in setting up an exchange programme to improve the dissemination of the know-how about Grid and e-Infrastructures.

4. Technology Description

4.1 – The gLite middleware

During the training events of the EPIKH, the gLite middleware [27] developed in the context of the EGEE project will be used. gLite is the middleware deployed on world largest e-Infrastructure and it is becoming a “de facto” standard in the Grid Computing. Default gLite services have been extended by the Consorzio COMETA to conjugate Grid Computing with High Performance Computing and this will increase the number of scientific communities EPIKH can reach during its time span.

4.2 – The GILDA t-Infrastructure

The term t-Infrastructure indicates an e-Infrastructure adapted to the needs of education, trainers and students. Shared t-Infrastructure would be usable by students and teachers internationally, providing easy access to educational exercises running on e-Infrastructure.

GILDA [28] officially started at the beginning of 2004 as an initiative of INFN [29] in the context of the INFN Grid Project [30] and the European EGEE project. The purpose of GILDA was to create a test-bed entirely dedicated to training and dissemination, based on the gLite middleware services, comprising of the most useful facilities such as a dedicated Certification Authority and a Virtual Organization and monitoring and support systems for users. Besides the support for training events, GILDA has also been used around the clock in the years by beginners, users or sites, wishing to start their Grid experience.

During these four years, more than 250 training and dissemination events have been supported by GILDA. The level of support went from the simple release of certificates (more than 13000 certificates have been issued since the beginning) to the creation of accounts and demonstrative applications on the resources for the full use of the training infrastructure during the events.

While supporting this large number of events, lots of problems have been faced and their solutions have become a series of best-practices in Grid education now usually adopted worldwide.

- Loose identification procedures for hosts and personal user certificates: the strict identification required from a “real” Certification Authority can be discouraging for new users and it can be considered not really necessary when approaching the Grid for the first time. Users have just to fill in a web form, and the certificate will be signed and sent to the email address of the requestor. This practice has clearly increased the number of certificates released, reducing the errors to which a non-experienced user is exposed and making this first, and usually error prone, step a lot easier. The risks that could derive from this non-strict identification, such as a misuse of the certificate, are mitigated by the small scope and time duration of these certificates: typically, just the GILDA test-bed itself, and two weeks, by default.
- Use of generic certificates and accounts for tutorials: in the first supported tutorials, users were contacted one or two days before the tutorial, and they were requested to ask for a personal GILDA certificate. As a matter of fact, most of them ignored these email. Many who did complete the certificate request correctly subsequently forgot to bring their certificates with them as advised, leaving them on their own machines. Requesting the certificate during the tutorial was not an effective solution, because the requests, and the various problems which may arise, made it impossible to support all participants quickly. This problem has been solved by forcing the tutorial organizer to specify the expected number of participants. Then, the GILDA CA manager issues the number of generic certificates requested and these are also exported in the requested format. System accounts are created on the official GILDA User Interface machines, and the certificates are copied in there. If the tutorial organizers plan to use different UI's, they can even request for the certificates to be sent separately. This practice also has the risk that certificates can be misused but this is mitigated by the fact that certificates are valid just for the duration of the tutorial, and they also have a limited scope.
- Use of wiki pages for on-off/line training: the most used training instruments in the beginning of GILDA have been transparencies. This practice proved to be not very effective, especially for exercises, for at least a couple of reasons. First, they offer a limited space for editing text, which is really a problem when reporting long option commands or command outputs. Second, it is hard to continuously maintain and update them in case of errors or changes due to new middleware releases. To face this issue, a wiki site has been setup in GILDA. The choice of the wiki site was also motivated by the fact that it enhances the collaboration among trainers without requiring physical access to the web server.
- Use of virtual machines for training: virtual machines have been proven to be a very effective instrument for Grid site administration tutorials, i.e. those training events where attendees learn how to install and operate Grid services. Since this exercise needs always a machine installed from scratch, the use of real machines is a clear limitation, because requires both a large number of available boxes and a huge amount of time to reinstall the machine from scratch for each Grid element. Virtual machines allow learners to start always from a preloaded image containing just the flat operating system. Students install the Grid service and then, once the exercise is finished, they have just to shutdown the virtual machine, reload the flat image, and they can start the next exercise in just a few minutes. Also, virtual machines are effectively used for dissemination purposes since they have been made available from the GILDA web site with several preinstalled Grid elements ready to be downloaded by users who wish to play with Grid elements even if they can't install a real machine or they don't have the opportunity to set up a full featured Grid.

4.3 – The GENIUS portal

GENIUS [31] is a web portal jointly developed since 2002 by INFN and NICE [32] with the goal to create a simple, though powerful and customizable, instrument for teaching Grid computing. Many Grid beginners are in fact discouraged by the complexity of the standard command line based interface (CLI) offered by the Grid middleware on a UNIX-like environment, that is hostile to a large part of potential users, that are not skilled computer scientists. GENIUS, which is usually installed on top of a gLite User Interface, offers a graphical, simple, and intuitive interface to the Grid services accessible from a common web browser without any additional requirements. When used during training events, GENIUS proves to be very effective in introducing Grid concepts, because Command Line Interface snags are hidden to users who do not have to check command syntax and can easily abstract their meaning.

GENIUS is available in GILDA in two flavours: the first is a full-featured installation, which requires a personal account, and has the same potentialities as a standard user interface. The other installation has been set up for demonstrative purposes. Although this has restrictions, like reduced capabilities for job submissions, it's available to everyone, including those having not a personal certificate or account, and so it allows broad dissemination of the strong capabilities of Grid computing.

4.4 – The P-GRADE portal

The P-GRADE Portal [33] is a web portal including workflow management capabilities developed by MTA-SZTAKI [34]. A P-GRADE installation has been set up for the international GILDA training infrastructure in December 2006. The environment serves as a demonstration, dissemination and learning environment for everybody who is interested in the usage and capabilities of GILDA, the gLite middleware and the P-GRADE Portal itself. During the roughly one year the GILDA P-GRADE Portal has been used during every mayor Grid training events worldwide. The GILDA P-GRADE Portal is P-GRADE Portal 2.5 installation being connected to the GILDA training infrastructure. It provides graphical environment to perform certificate management, job submission, file transfer, information system browsing, application monitoring on GILDA, eliminating the sometimes cumbersome and hard to memorize commands from the learning curve. As a result the learning time required for Grids can be significantly shortened by the tool. Besides providing graphical interfaces for GILDA middleware services, the GILDA P-GRADE Portal also contains high level tools that extend the capabilities of gLite. Workflows and parameter studies can be defined and managed by the graphical editor the integrated workflow manager components.

Attendees of P-GRADE courses are provided with pre-defined exercises that introduce the general concept of parallel Grid application development and use the gLite middleware to demonstrate them in practice. These exercises stress gLite middleware services and they are formulated as data parallel parametric studies, functional parallel workflows or some combination of these two. Based on the examples students can understand and distinguish the generic concepts of Grid applications from implementation details specific to a gLite VO or on a Globus VO. (P-GRADE Portal is also compatible with Globus middleware [35] based Grids, however this is not used in the GILDA P-GRADE Portal installation.) We can declare that P-GRADE courses are very beneficial for attendees.

On the other hand, such events are prosperous for the developers of P-GRADE and gLite too. As tutors meet existing and potential new Grid users who are committed to understanding the capabilities and limitations of Grid systems they can easily collect feedback. Feedbacks collected during training are very valuable in order to set priorities and lay down roadmaps for future developments. Consequently, tutors of P-GRADE Portal

sessions take extra care to take note of the experience of every course, to categorize and then to forward these notes to appropriate Grid developer and operator groups.

In order to demonstrate the maturity of Grid and the benefits of the technology we use real life applications during P-GRADE Portal tutorials wherever possible. However, it is necessary to decrease the size of those applications in order to fit them into the time and resource constraints of the events. This is typically achieved by reducing the size of the input data sets – resulting shorter execution times and smaller resource demands – or dropping some application components. Since 2006 it is used during P-GRADE Portal events to demonstrate the concept of data driven workflows and basic steps of workflow management. As the original workflow was running for about an hour the input data set had to be reduced and now it finishes in about 10 minutes on GILDA thus suitable for training purposes.

Although students work more eagerly on exercises that are based on real life case studies, sometimes setting up exercises from scratch solely for the sake of training is inevitable. This is the situation when such capabilities of P-GRADE must be demonstrated that are new, or simply not used by any production application yet. As P-GRADE aims to serve as a general Grid application developer and executor environment its new features target potential application classes. Meanwhile other Grid tools are typically designed for one particular experiment, most of the P-GRADE features are not used right after they become part of a release, their take up by Grid communities take some time until new features become known and understood. Some artificial exercises shorten this take up period focusing onto such new features and putting them into theoretical but possible use case scenarios.

The GILDA P-GRADE Portal provides permanent service that anybody can access any time. Having the training environment and infrastructure publicly available all the time is very beneficial, as in this way participants of organized events can continue their P-GRADE studies after the course. These people can stay in the same environment and can reuse applications that they designed during the tutorial. However, as the certificates they use during tutorials expire on the same day (or within a few days) these people must obtain new GILDA certificates and must register to the GILDA VO again. Even if GILDA provides lightweight mechanisms through its web to accomplish these steps, users loose a few days and must invest a few hours extra work just to get GILDA access again. Tutorial certificates with extended lifetime or some simplified GILDA registration procedure could address this issue. (For example in a simplified GILDA registration procedure the user could obtain a GILDA certificate, GILDA VO membership and user accounts on GILDA UI, GENIUS and P-GRADE Portal in a single step, filling out only one web form).

4.5 – The Grid2Win User Interface

Today's e-Infrastructures are mostly made of machines running Linux or other Unix-like operating systems. This holds also for those machines that act as User Interface. This constitutes in many cases a "learning barrier" and prevents many people used to Microsoft Windows platforms to access the Grid. In order to address and overcome this problem and to complement the formative offer of the Grid portals mentioned above, the Consorzio COMETA and the INFN are developing since more than a year Grid2Win [36] which is a graphic Windows application implementing a fully-fledged gLite User Interface. The use of Grid2Win will be shown and suggested in all educational events that will be done by EPIKH and the tool will be made available to everybody outreached by the project.

5. Developments

The work plan of EPIKH is articulated into the three Work Packages (WPs) listed in table 2.

Work Package no.	Work package title
WP1	Management of the project
WP2	Dissemination and outreach
WP3	Organisation and running of events

Table 2: Work packages of the EPIKH project.

The objectives of the three WPs are the following:

- WP1:
 - Administrative Coordination:
 - Internal management dealing with baseline administrative matters;
 - Project reporting (deliverables, reports, etc.);
 - Finances: cost statement collection and analysis, interim payment coordination and transfer, budget follow-up;
 - Liaison with members, related projects and other Grid initiatives to organise common events;
 - Relations with the EC;
 - Technical Coordination:
 - Day-by-day technical coordination;
 - Regular assessment of progress achieved in preparing project deliverables, to ensure that commitments are met on time;
 - Interface with other projects on technical matters;
 - Documentation and distribution of all relevant technical discussions, decisions.
- WP2:
 - Promote and spread the project to the scientific, academic and industrial communities as well as among politicians and decision makers. The Work Package will also be responsible for the creation and maintenance of the project website and all the web tools (agenda, document server, video-conferencing system, e-learning system, etc.) necessary for remote collaboration;
- WP3:
 - Organisation and running of the events planned by the project.

6. Results

EPIKH will start on the 1st of March 2009 and will last for 4 years. At the time of writing, then, we can not talk about results but rather about the impact of the expected results and how they match the expectations and the recommendations of the European Commission. The EPIKH project addresses many of the issues and implements many of the recommendations that appear in the EC document “Towards a European Research Area” [37].

In the rest of this section, some parts of that document are reported verbatim (in *italic*) together with short paragraphs explaining how EPIKH is relevant for them.

- *“Even more so than the century that has just finished the XXIst century we are now entering will be the century of science and technology. More than ever, investing in research and technological development offers the most promise for the future.”*
 - EPIKH does invest in R&D development investing in a bleeding-edge technology such as Grid computing through the most important and long-lasting “tool”: education.
- *“In the final years of the XXth century we entered a knowledge-based society. Economic and social development will depend essentially on knowledge in its different forms, on the production, acquisition and use of knowledge. Scientific research and technological development more particularly are at the heart of what makes society tick. [...] By creating new products, processes and markets research and technology provide one of the principal driving forces of economic growth, competitiveness and employment.”*
 - Scientific research is one of the most important “providers” of knowledge and e-Infrastructures are one of the most important “enablers” of e-Science, a new way of “doing science” in a global and more collaborative framework. EPIKH implements well definite training processes and policies and addresses education to spread the know-how on how to use and profit from e-Infrastructures.
- *“[...] Networking of existing centres of excellence in Europe and the creation of virtual centres through the use of new interactive communication tools.”*
 - In the context of EPIKH, several EU and non-EU stakeholders of e-Infrastructure management and operation are sharing and “merging” their experiences to create a virtual laboratory, based on Grids, which will act as a new, strong tool to “communicate” science, technology and development in general.
- *“More use should be made in future, at national and European level, of the possibility of using mobility as an instrument of information and technology transfer.”*
 - This is indeed at the very base of the EPIKH idea. EPIKH will use the mobility funds made available for information and technology transfer.
- *“Information, training and familiarisation projects for researchers [...] should, moreover, be undertaken in tandem by the Member States and the Commission [...]”*
 - This is exactly what EPIKH plans to implement.
- *“In the case of developing countries, to guarantee the development of local research potential, [...] system should be such as to encourage the beneficiaries to return to their countries in order to take advantage of their experience and to spread the knowledge they have acquired.”*
 - Indeed, EPIKH plans to train in Europe future experts of e-Infrastructures from beneficiary countries and then perform all the scheduled education events in the continents interested by the project (Africa, Asia and Latin America) using both EU and non-EU trainers. This will be essential to ensure long-term sustainability to the action even beyond the end of the project.
- *“The full panoply of instruments available to the Union should be brought into play: [...] Structures and mechanisms of exchange of information and experience: working groups, networks of experts and operators; [...]”*

- The creation of education working groups and the “exchange” of experts and operators between Europe and the non-EU countries involved in EPIKH is one of the main goals of the project.

The EPIKH project has also a big potential to create and/or develop lasting collaborations with the eligible third country partners.

This for the following reasons:

- several scientific collaborations are already ongoing among many of the partners of EPIKH and the project, lasting for 4 years, will reinforce and extend them;
- the programs of the Grid schools will be very much driven by the scientific field and by the applications selected; this will foster synergies among groups working on the same subject and ease the establishment of lasting collaborations;
- almost all the partners of the EPIKH consortium are also members of large e-Infrastructures existing in their continents; they will then act as “gateways” or “catalysts” of further collaborations between Institutions and/or Scientific Communities involved in EPIKH and Institutions and/or Scientific Communities not involved in EPIKH but already using those e-Infrastructures; figure 9 shows in detail the work flow that will be put in place to induce this “catalysis”.

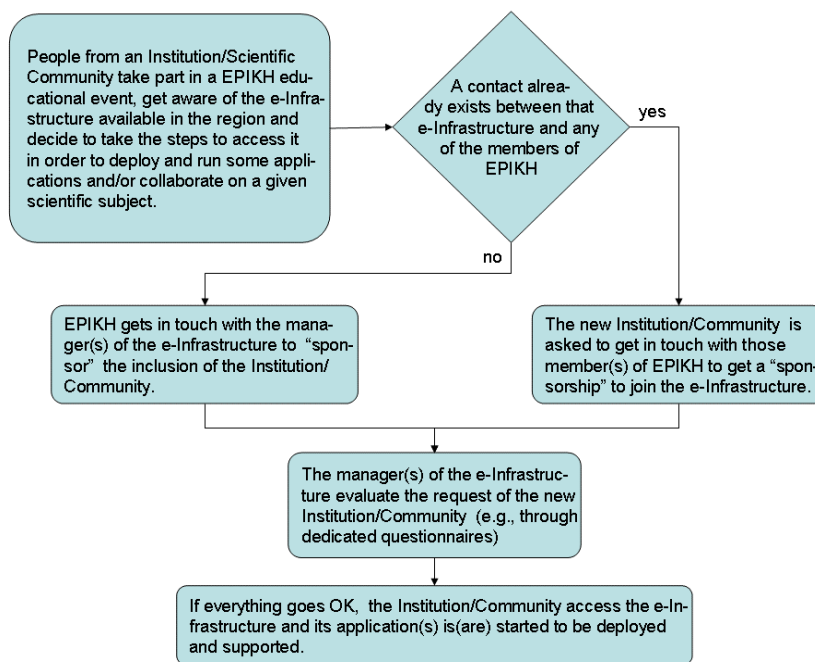


Figure 9: “Workflow” to attract and support new communities through EPIKH.

As mentioned above, the partners of EPIKH are also involved in projects/consortia like EELA, EGEE, EUAsiaGrid, EUChinaGRID, EU-IndiaGrid, and EUMEDGRID which last from at least two years and will continue for at least two more years so the synergies that will be put in place between EPIKH and those projects will ensure lasting collaborations and will strengthen the scientific relationships among the four continents involved in EPIKH.

7. Business Benefits

Although the primary beneficiaries of the EPIKH project will be researchers and technologists belonging to the academic world, the EPIKH training events will also be open to the participation of people coming both from the industry and business world.

The adoption of grid technology by industry is slower than it could be (and than what happened with the web 15 years ago). Recent surveys have confirmed that the most important “brakes” are:

- Security:
 - Strict Access Control Lists and strong “privacy” rules for sensible data;
- Standards:
 - Their adoption would encourage long term investments;
- Training:
 - Need of a formative offer “tailored” for the business world;
- Management of commercial software licences:
 - Usage control and protection of revenues of commercial software providers;
- The full support for applications needing high performance computing:
 - Availability of MPI-1 and MPI-2 protocols on low-latency networks;
- The level of quality of service:
 - Need of dynamic systems so define and enforce Quality of Service (QoS) and Service Level Agreements (SLAs) which foresee the possibility to change contracts as a function and time and external conditions;
- Availability and reliability of accounting and billing systems:
 - Creation of computing and storage accounting systems able to work in a distributed grid environment and to implement different economic models in a dynamic way.

Whenever and wherever possible and applicable to the audience foresee, the above concepts will be included as integral part of the programmes of the EPIKH schools and thematic workshops.

8. Conclusions

The “Exchange Programme to advance e-Infrastructure Know-How” (EPIKH) project will start on the 1st of January 2009 and will last for 4 years involving, in more than 500 secondments, about 115 people for a total of over 650 researchers-months (not counting the people that will be outreached by the project’s dissemination activities).

EPIKH will be carried out by a consortium of 23 Organizations belonging to 4 continents and can be considered one of the largest and widest exchange programmes ever in the field of training and education on research e-Infrastructures.

Eight Grid schools and as many thematic workshops will be organized and run in Africa during the life time of the project. These can then represent a big opportunity for many African scientists and technologists to learn a “bleeding edge” technology and access and use large, geographically distributed, computing and storage infrastructures. This will bridge the “digital divide” between Africa and the rest of the world in the field of scientific research and can turn out into a key enabler of the so called “brain gain”.

References

- [1] http://en.wikipedia.org/wiki/Jean-Jacques_Rousseau.
- [2] Maps presented in section 1.1 are taken from <http://maps.maplecroft.com>. Data sources are reported in the figures.
- [3] http://www.e-irg.eu/images/stories/publ/task_force_reports/ettf_long_report_final_july08.pdf.

- [4] <http://www.e-irg.org>.
- [5] <http://www.oecd.org>.
- [6] <http://www.worldbank.org>.
- [7] <http://www.geant2.net>.
- [8] <http://www.dante.net>.
- [9] <http://www.eu-egee.org>.
- [10] <http://www.deisa.org>.
- [11] <http://alice.dante.net>.
- [12] <http://www.eumedconnect.net> and <http://www.eumedconnect2.net>.
- [13] <http://global.dante.net/server/show/nav.1416>.
- [14] <http://global.dante.net/server/show/nav.1418>.
- [15] <http://www.seeren.org>.
- [16] <http://www.tein2.net>.
- [17] <http://www.balticgrid.eu>.
- [18] <http://www.eu-eela.eu>.
- [19] <http://www.euasiagrid.eu>.
- [20] <http://www.euchinagrid.eu>.
- [21] <http://www.euindiagrid.eu>.
- [22] <http://www.eumedgrid.eu>.
- [23] <http://www.see-grid-sci.eu>.
- [24] <http://www.epikh.eu>.
- [25] <http://www.consortio-cometa.it>.
- [26] http://wiki.epikh.eu/index.php?option=com_content&view=article&id=47&Itemid=55.
- [27] <http://www.glite.org>.
- [28] <https://gilda.ct.infn.it>.
- [29] <http://www.infn.it>.
- [30] <http://grid.infn.it>.
- [31] <https://genius.ct.infn.it>.
- [32] <http://www.nice-italy.com>.
- [33] <http://portal.p-grade.hu>.
- [34] <http://www.sztaki.hu/?en>.
- [35] <http://www.globus.org>.
- [36] <http://grid2win.gilda-forge.ct.infn.it>.
- [37] <http://ec.europa.eu/research/era/pdf/com2000-6-en.pdf>.