e-Infrastructures for e-Science: A Global View

**Giuseppe Andronico · Valeria Ardizzone · Roberto Barbera · Bruce Becker ·
Riccardo Bruno · Antonio Calanducci · Diego Carvalho · Leandro Ciuffo ·
Marco Fargetta · Emidio Giorgio · Giuseppe La Rocca · Alberto Masoni ·
Marco Paganoni · Federico Ruggieri · Diego Scardaci**

Abstract In the last 10 years, a new way of doing
science is spreading in the world thank to the de-
velopment of virtual research communities across
many geographic and administrative boundaries.
A virtual research community is a widely dispersed
group of researchers and associated scientific
instruments working together in a common virtual
environment. This new kind of scientific environ-
ment, usually addressed as a "collaboratory", is
based on the availability of high-speed networks
and broadband access, advanced virtual tools and
Grid-middleware technologies which, altogether,

G. Andronico · V. Ardizzone · R. Bruno ·
A. Calanducci · M. Fargetta · E. Giorgio ·
G. La Rocca · D. Scardaci (✉)
Division of Catania, Italian National Institute
of Nuclear Physics, Via S. Sofia, 64,
95123 Catania, Italy
e-mail: diego.scardaci@ct.infn.it

G. Andronico
e-mail: giuseppe.andronico@ct.infn.it

V. Ardizzone
e-mail: valeria.ardizzone@ct.infn.it

R. Bruno
e-mail: riccardo.bruno@ct.infn.it

A. Calanducci
e-mail: antonio.calanducci@ct.infn.it

M. Fargetta
e-mail: marco.fargetta@ct.infn.it

E. Giorgio
e-mail: emidio.giorgio@ct.infn.it

G. La Rocca
e-mail: giuseppe.larocca@ct.infn.it

R. Barbera
Department of Physics and Astronomy
of the University of Catania, Viale A. Doria, 6,
95125 Catania, Italy
e-mail: roberto.barbera@ct.infn.it

R. Barbera
Italian National Institute of Nuclear Physics - Division
of Catania, Viale A. Doria, 6,
95125 Catania, Italy

B. Becker
Meraka Institute, Building 43, CISR,
Meiring Naude Road, PO Box 395 Pretoria 0001,
Pretoria, South Africa
e-mail: bbecker@csir.co.za

D. Carvalho
Centro Federal de Educaçãao Tecnológica Celso
Suckow da Fonseca, Av. Maracanã, 229, 20271-110
Rio de Janeiro, Brazil
e-mail: d.carvalho@ieee.org

L. Ciuffo
Rede Nacional de Ensino e Pesquisa, R. Lauro Müller
116/1103, 22290-906 Rio de Janeiro, Brazil
e-mail: leandro.ciuffo@rnp.br

are the elements of the e-Infrastructures. The European Commission has heavily invested in promoting this new way of collaboration among scientists funding several international projects with the aim of creating e-Infrastructures to enable the European Research Area and connect the European researchers with their colleagues based in Africa, Asia and Latin America. In this paper we describe the actual status of these e-Infrastructures and present a complete picture of the virtual research communities currently using them. Information on the scientific domains and on the applications supported are provided together with their geographic distribution.

Keywords e-Infrastructure · e-Science · Virtual research communities · Virtual organization · Gridification

1 Introduction on the Concept of e-Infrastructure

Across the end of the Second Millennium and the beginning of the Third, the way scientific research is carried out in many parts of the world is rapidly evolving to what is nowadays referred to as e-Science, i.e. a "scientific method" which foresees the adoption of cutting-edge digital platforms known as e-Infrastructures [1] throughout the process from the idea to the production of the scientific result. Conceptually, e-Infrastructures can be represented by three layers: (1) the bottom layer is made up by the scientific instruments and

A. Masoni
Italian National Institute of Nuclear Physics, Division of Cagliari, Complesso Universitario di Monserrato, S.P. per Sestu km. 0.700, 09042 Monserrato (CA), Italy
e-mail: alberto.masoni@ca.infn.it

M. Paganoni
Department of Physics of the University of Milano Bicocca, Piazza della Scienza, 3, 20126 Milan, Italy
e-mail: marco.paganoni@cern.ch

F. Ruggieri
Division of Roma Tre, Italian National Institute of Nuclear Physics, Via della Vasca Navale, 84, 00146 Rome, Italy
e-mail: federico.ruggieri@roma3.infn.it

experiments providing huge amount of data; (2) then, on the top of the previous one, we have the Grid layer, networked data processing centres and middleware software as the "glue" of resources; (3) the third and highest level includes researchers that perform their activities regardless geographical location, interact with colleagues, share and access data.

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) [2, 3], 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" [4], i.e. a large number of computing and storage devices, linked among them by high-bandwidth networks, on which a special software called middleware (intermediate between the hardware and the operating system and the codes of the applications) [5] is installed, allowing the resources to 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 what is commonly called an e-Infrastructure. On the top of the e-Infrastructure we have the e-Science Collaborations, virtual research communities composed by widely dispersed group of researchers working together in the common virtual environment.

The European Commission (EC) is heavily investing through its Framework Programmes in e-Infrastructures and these platforms are by now considered as 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

are highly beneficial for the society and the progress at large.

EC investments in e-Infrastructures have gone well beyond its own borders and many project have been funded to create/extend e-Infrastructures in/to other regions of the world and make them connected and interoperable [6] with the European one. All together, these projects/initiatives have contributed to create a "global" network and a "global" Grid as shown in Fig. 1.

All these projects share the same work plan whose "virtuous cycle" can be summarized as follow:

- Set-up and manage a pilot infrastructure;
- Seek, identify and support e-Science communities in the region and deploy their applications;
- Train users and site administrators to use and operate an e-Infrastructure based on the European middleware (e.g. gLite [7, 8]);
- Disseminate, both bottom-up and top-down, the e-Infrastructure paradigm for long term sustainability of e-Science.

In this paper we not only provide a global view of the e-Infrastructures, based on the European middleware gLite, currently operational in the regions of the world but we also show, for the first time,

a global picture of how the research communities take advantages of this new scientific environment. After a description of the e-Infrastructures deployed in the last years in the world, we present a picture of the virtual research communities that exploit them worldwide with their scientific applications classified by scientific domain and geographic location. In this way, the impact of the e-Infrastructure model to the scientific research can be evaluated as a function of the scientific domain and the geographical distribution depicting how this new way of doing science spreads over the world.

The paper is organized as follow. Section 1 introduces the concept of e-Infrastructure. Section 2 describes the e-Infrastructure status in all the regions involved. Section 3 explains the Gridification Process, addressing the common problems encountered in deploying e-Science applications on e-Infrastructures and describing the procedure adopted. Section 4 shows a worldwide overview of the virtual research communities exploiting the existing e-Infrastructures with their scientific applications. Applications have been classified according to their scientific domains and their geographical origin and a short description of some relevant pilot applications for each of the scientific domains considered is provided. Finally, in Section 5, conclusions are drawn.

Fig. 1 The "global" Grid
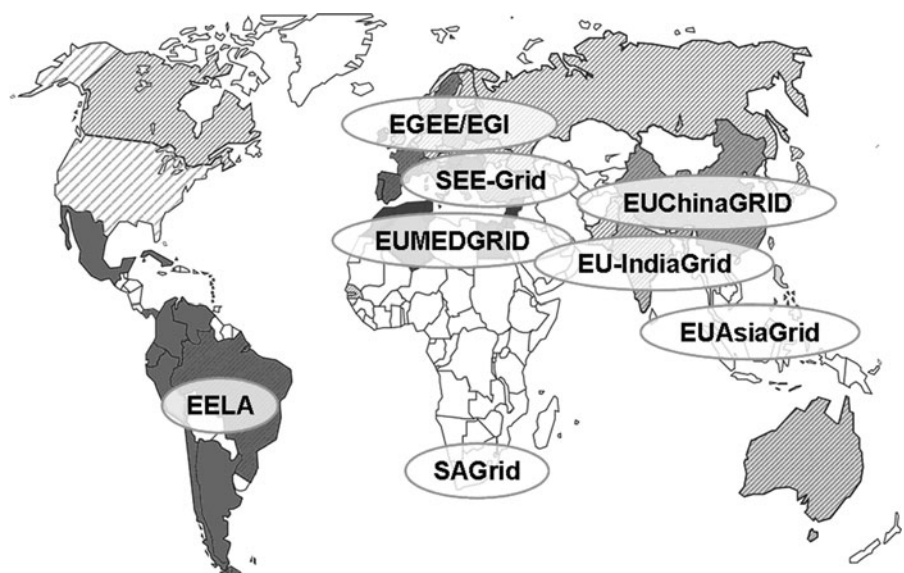
2 Regional e-Infrastructures

This section describes the status of the e-Infrastructures currently operational in the regions of the world. The picture depicted in this section take in account only the e-Infrastructures based on the European Middleware, gLite (with the only exception of the Indian area). Several other important e-Infrastructures, based on different middleware, have been deployed and are currently running in the world like the Open Science Grid [9], DEISA [10], TeraGrid [11], NAREGI [12], the China National Grid [13], etc. They are not in the scope of this paper.

2.1 Europe

The Grid infrastructure for e-Science in Europe is based on the GÉANT network [14] and has been established through a series of projects started in early 2001 with the European DataGrid (EDG) project [15] followed by the three phases of the Enabling Grids for E-SciencE (EGEE) project [16]. The EGEE project, which was based on national and thematic Grid efforts, as well as the pan-European network provided by GÉANT and the National Research and Education Networks (NRENs), brought together experts from all over Europe and the rest of the world (more than 50 countries) with the common aim of building on recent advances in Grid technology and developing a service Grid infrastructure which is available to scientists around the clock. Since the beginning of the second phase of the project, several scientists, grouped into Virtual Organisations (VOs), started to access and use the virtual services exposed by the Grid framework to carry on collaborative research and gain access to shared computing and data resources. The up to date list of all registered VOs is available from the CIC portal (http://cic.gridops.org). EGEE provided a variety of services to scientists, ranging from training and user support, through to the software infrastructure necessary to access the resources. Thanks to the reliable services developed by the project, several scientific communities such as: Life Sciences and High Energy Physics (HEP) now depend on the EGEE Grid infrastructure, which is an essential and crucial part

of their large-scale data processing. Other disciplines, notably astrophysics, Earth sciences, computational chemistry, and fusion are increasingly using it for production processing. In 6 years, the EGEE project has also established several relationships with other Grid-related projects such as: EUMEDGRID [17], EUChinaGRID [18], and EUAsiaGrid [19] projects, to name just a few, in order to extend the infrastructure, and its paradigm, to foster the induction of new scientific communities, to establish a seamless computing infrastructure around the globe and reduce the digital divide. Currently, there are 10 disciplines and 36 scientific areas that are being supported. Among them:

- Astronomy and Astrophysics;
- Computational Chemistry;
- Earth Sciences;
- Fusion;
- High-Energy Physics;
- Life Sciences.

2.1.1 The European Grid Infrastructure

The European Grid infrastructure (EGI) is a federation of Resource Infrastructures (typically National Grid Initiatives but not only) with a central coordinating body (EGI.eu, incorporated in Amsterdam). EGI is based on the gLite middleware services deployed on a worldwide collection of computational and storage resources, plus the services and support structures put in place to operate them. As of today, the European Grid Infrastructure, shown in Fig. 2, consists of more than 330 Resource Centres for a total of about 330,000 CPU cores and 80,000 TB of storage.

2.2 Latin America

The Latin American Grid Infrastructure (LGI), based on the RedCLARA backbone [20], has been developed thanks to several projects co-funded by the European Commission since 2006 (EELA, EELA-2 and GISELA).

The RedCLARA backbone, implemented and managed by CLARA (Cooperación Latino Americana de Redes Avanzadas), is composed of 9 nodes, interconnects 12 Latin American

Fig. 2 EGI infrastructure. GStat Geo View (Gstat is a web application which aim is to display information about Grid services, the Grid information system itself and related metrics (http://gstat.egi.eu/gstat/geo/openlayers). Sites under maintenance could not appear on the figure.)

NRENs: RNP (Brazil), InnovaRed (Argentina), REUNA (Chile), RENATA (Colombia), CEDIA (Ecuador), RAICES (El Salvador), RAGIE (Guatemala), CUDI (Mexico), RedCyT (Panama), RAAP (Perú), RAU2 (Uruguay) and REACCIUN (Venezuela). It is also connected to the pan-European network GÉANT, as well as to other international research networks.

LGI deployment started in 2006 thanks to the EELA first-phase project [21] that provided its users with a stable and well supported Grid which proved, in the two year lifetime of the project, that the deployment of an European–Latin American e-Infrastructure was not only viable but also responding to the real needs of a significant part of the Scientific Community. After then, the EELA-2 Project [22] (E-science Grid facility for Europe and Latin America) started and carried on the work performed during EELA with the aim to build a high capacity, production-quality, scalable Grid to answer the needs of a wide spectrum of applications from European–Latin American scientific collaborations.

The focus of EELA-2 was on:

- Offering a complete set of versatile services fulfilling Applications requirements;
- Ensuring the long-term sustainability of the e-Infrastructure beyond the term of the project

EELA-2 can be considered as the first and most inclusive initiative that has been carried out in Latin America in the area of distributed computing infrastructures.

One of the main outcomes of the EELA-2 project was the definition of the LGI sustainability model, rooted on National Grid Initiatives (NGI) or Equivalent Domestic Grid Structures (EDGS), that will be implemented, in collaboration with CLARA, during the GISELA project, the third phase of the LGI deployment process. The

ambitious target of GISELA is to make LGI self-sustainable by the end of the project. In the same time GISELA works to provide Virtual Research Communities (VRCs) with the e-Infrastructure and Application-related Services required to improve the effectiveness of their research and to create a collaboration channel between LGI and the European Grid Initiative (EGI).

2.2.1 The Latin Grid Infrastructure (LGI)

The Latin American Grid Infrastructure, shown in Fig. 3, consists, as of today, of 28 Resource Centres for a total of more than 6000 CPU cores and about 500 TB of disk storage.

2.3 Mediterranean and Middle-East

Co-funded by the European Commission within the Sixth and Seventh Framework Programmes, the EUMEDGRID projects (EUMEDEGRID and EUMEDGRID-Support [23]) have run in parallel but in conjunction with EUMEDCONNECT projects [24, 25] (the EUMEDCONNECT projects have played a pioneeristic role in the promotion of Communication Networks, as fundamental components of e-Infrastructures in the Mediterranean), and have supported the development of a Grid infrastructure in the Mediterranean area. EUMEDGRID promoted the porting of new applications on the Grid platform, thus allowing Mediterranean scientists to collaborate more closely with their European colleagues. EUMEDGRID has disseminated Grid awareness and competences across the Mediterranean and, in parallel, identified new research groups to be involved in the project, helping them to exploit Grids' enormous potential to improve their own research activities.

The implementation and coordination of a Grid infrastructure at a national (or wider) level can be regarded as an opportunity to optimize the usage of existing, limited storage and computing resources and to enhance their accessibility by all research groups. This is particularly relevant for the non-EU countries involved in the project.

Many research fields have indeed very demanding needs in terms of computing power and storage capacity, which are normally provided by large computing systems or supercomputing

Fig. 3 LGI infrastructure. GStat Geo View
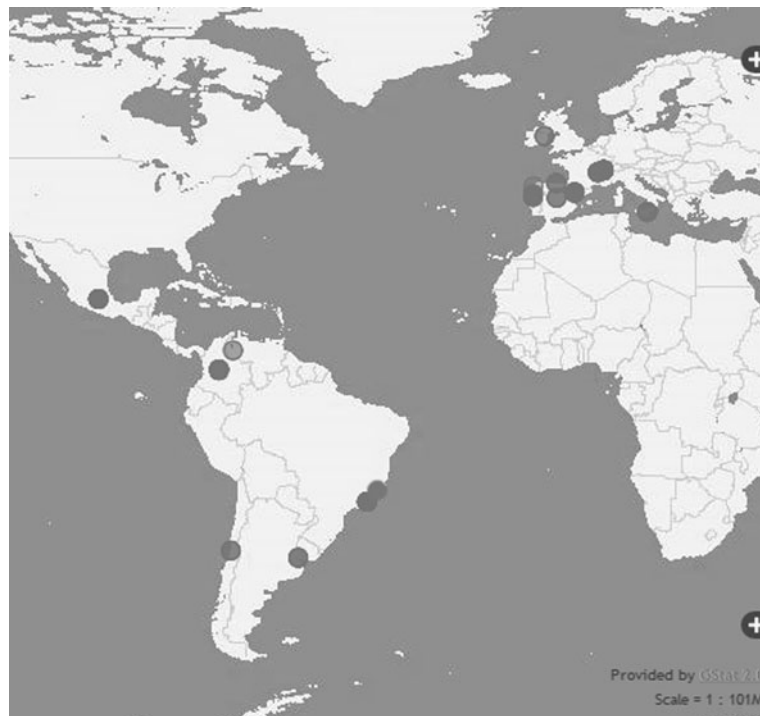
centres. Furthermore, sophisticated instruments may be needed to perform specific studies. Such resources pose different challenges to developing economies: they are expensive, they need to be geographically located in a specific place and they cannot attract a critical mass of users because they are usually very specific and are relevant only for small communities of researchers scattered across the country/region. This is the case even in some strategic domains such as water management, climate change, biodiversity and biomedical activities on neglected or emerging diseases. Thus, a significant part of researchers is forced to emigrate to more developed countries to be able to continue their scientific careers. However, thanks to the creation of global virtual research communities and distributed e-Infrastructure environments, all these drawbacks can be overcome: through an appropriate access policy, different user groups can use resources wherever dispersed, according to their availability. Furthermore, geographically distributed communities working on the same problem can collaborate in real time thus optimizing not only hardware and software resources but also human effort and "brainware".

The EUMEDGRID projects have been conceived in this perspective and have set up a pilot Grid infrastructure for research in the Mediterranean region which is interoperable and compatible with that of the EGI initiative. The EUMEDGRID's vision focused on improving both the technological level and the know-how of networking and computing professionals across the Mediterranean thus fostering the introduction of an effective Mediterranean Grid infrastructure for the benefits of e-Science.

The EUMEDGRID achievements can be categorised into two main areas:

- The creation of a human network in e-Science across the Mediterranean;
- The implementation of a pilot Grid infrastructure, with gridified applications, in the area.

Cooperation among all the participants has been demonstrated by the enthusiastic participation to common workshops and meetings organized during the lifetime of EUMEDGRID and the success obtained fostering the creation of National Grid Initiatives and national Certification Authorities (CAs) officially recognized by IGTF. Impressive results were also obtained in the events of knowledge dissemination on Grid technology and services.

The promotion of National Grid Initiatives carried out in all non-EGI partner countries registered a good level of success with programmes already operational in Algeria, Egypt, Morocco and Tunisia and well advanced plans in Jordan and Syria. The national impact and policy level awareness in some of these countries has led to an initial financial support of the initiatives.

Besides its scientific mission, EUMEDGRID had also a significant socio-economic impact in the beneficiary countries. Fostering Grid awareness and the growth of new competences in EU Neighbours' scientific communities is a concrete initiative towards bridging the digital divide and the development of a peaceful and effective collaboration among all partners.

e-Infrastructures also contribute to mitigate the so-called "brain-drain" allowing brilliant minds in the area to stay in their regions and contribute significantly to cutting edge scientific activities, concretely enlarging the European Research Area (ERA). Research and Education Networks and Grids are fundamental infrastructures that will allow non-EU researchers to carry out high quality work in their home laboratories without the need to migrate in more advanced countries.

An extended Mediterranean Research Area could thus be seen as a first step towards the realisation of more politically ambitious plans of open market, open transportation infrastructures, free circulation of citizens, etc.

At the present time a big effort is devoted to push for a consolidation of the existing EUMEDGRID infrastructure and for the development of sustainable e-Infrastructures in the Mediterranean region in a broad, general, meaning.

2.3.1 The Mediterranean Grid Infrastructure

The Mediterranean Grid Infrastructure, shown in Fig. 4, consists of 30 Resource Centres for a total of about 1,600 CPU cores and about 600 TB of storage.

Fig. 4 EUMEDGRID infrastructure. GStat Geo View

2.4 Far East Asia

The development of Grid infrastructures in the Asia-Pacific region has been driven by the participation of local scientists to the CERN Large Hadron Collider experiments as well as a number of applications of specific interest such as biomedical research, engineering applications and disaster mitigation. After then, the EUAsiaGrid project [19], co-funded by the European Commission in the context of its Seventh Framework Program, starting from the Grid sites already available in the region, managed to build a production Grid infrastructure increasing the capacity currently available in a number of countries such as Malaysia, Indonesia, the Philippines, Thailand and Vietnam. Nowadays, this infrastructure ensures that local support and key services such as user interface nodes, compute and storage elements can be taken for granted by researchers and that the core of a sustainable e-Infrastructure in the region is put in place. The Grid infrastructure is built on top of the third generation of the Trans-Eurasia Information Network (TEIN3), maaged by the Asia-Pacific Ad-

vanced Network [26] organisation and the TEIN3 project [27], providing today a dedicated high-capacity Internet network for research and education communities across Asia-Pacific. It currently connects 11 countries in the region (Australia, China, Indonesia, Japan, Korea, Laos, Malaysia, Philippines, Singapore, Thailand and Vietnam) and a direct connectivity to the GÉANT network in Europe.

While a number of countries in the Asia-Pacific region have done significant investments in e-Infrastructures for research, the level of funding is still very heterogeneous and only a few countries have National Grid Initiatives (NGIs) that provide the necessary coordination at the national level to leverage the capability of Grids to provide persistent and sustainable e-Infrastructures that can be taken for granted by researchers and that enable them to focus on their substantive research. Until recently, most Grid-related initiatives were based at individual institutions that sought to build up capacity to support specific research projects and application areas. As a consequence, many resource providers ended up trying to support installations with different

middleware stacks, stretching their resources. Clearly, a coordinated approach to the development of a persistent and sustainable e-Infrastructure could not only maximise the return on investment by enabling a wider range of researchers to benefit from the resources but also help resource providers cope with the heterogeneity and continuous evolution of Grid technologies.

Through the coordination and support provided by EUAsiaGrid, much needed local capacity has been developed, both in terms of resources available as part of the world-wide EGI infrastructure and in terms of the supporting human infrastructure that is needed to carry on their ongoing operation and effective exploitation by researchers. To minimize barriers to access the Grid infrastructure, the EUAsiaGrid project also created and maintained a catch-all, application neutral, Virtual Organisation called EUAsia. Furthermore, several Certification Authorities (CAs) approved by the International Grid Trust Federation already operate in the region, with the Academia Sinica Grid Computing one serving as a catch-all CA and taking care of users of those countries that do not yet have their national CA.

Any researcher from the region interested in trying the Grid for his/her research can get a certificate through a nearby Registration Authority and immediately subscribe to the EUAsia VO. Although application neutral, nodes serving this VO have installed many application packages to be easily available. Also, each partner has set up a user interface to provide local access to the Grid.

2.4.1 The EUAsia-Grid Infrastructure

The EUAsia-Grid Infrastructure, shown in Fig. 5, consists, as of today, of 16 Resource Centres for a total of 880 CPU cores and more than 60 TB of storage.

The EUAsiaGrid project came to an end in June 2010 but the EUAsiaGrid Consortium has agreed to keep open, on a best effort basis, the existing infrastructure.

2.5 India

In India, two main Grid Initiatives have been taken at governmental level: Regional WLCG [28] set up by the Department of Atomic Energy (DAE),

Fig. 5 EUAsiaGrid Grid infrastructure. Gstat Geo View

in coordination with the Department of Science & Technology (DST), and the GARUDA National Grid Initiative [29]. GARUDA benefits from the creation of the National Knowledge Network (NKN), a new backbone created by the National Knowledge Commission (NKC) with the aim to bring together all the stakeholders in Science, Technology, Higher Education, Research & Development, and Governance with speeds in the order of tens of Gb/s coupled with extremely low latencies. The regional WLCG in India is going to be migrated to NKN that will provide transport to ERNET (the Indian National Research and Education Network) replacing its existing backbone.

The Indian Department of Atomic Energy (DAE) is actively participating to the scientific program of LHC Computing Grid project taking active part in CMS and ALICE experiments, devoted to find answers to the most fundamental questions at the foundations of matter constituents. The data from the LHC experiments will be distributed around the globe, according to a four-tiered model. Within this model, to support researchers with required infrastructure, India has also setup regional Tier-2 centres connected to CERN. In India there are two Tier-2 centres: one for CMS at TIFR in Mumbai and one for ALICE at Saha-VECC in Kolkata. These centres provide access to CMS and ALICE users working from Tier-3 centres at Universities and national labs and LCG Data Grid services for analysis.

Specific activities are also ongoing in the area of Grid Middleware Software development, devoted to ensuring Grid enabling of IT systems. These activities cover the area of Grid Fabric management, Grid Data management, Data Security, Grid workload scheduling and monitoring services, fault tolerant systems, etc. DAE developed number of Grid based Tools in the area of Fabric management, AFS file system, Grid View and Data Management, which are being deployed by CERN in their LHC Grid operations since September 2002. So far, the number of software tools and packages such as a correlation engine, Grid operations monitoring, problem-tracking system, Pool Database Backend Prototype, Scientific Library Evaluation and Development of Routines, AliEn Storage System and Andrews File System, were

developed by DAE team members under a Computing Software agreement. Currently, BARC engineers are working on the enhancement of the Grid View software tool.

GARUDA is a collaboration of scientific and technological researchers for a nation wide Grid comprising computational nodes, mass storage systems and scientific instruments. It aims to provide the technological advances required to enable data and compute intensive science for the 21st century.

GARUDA has transitioned form the Proof of Concept phase to the Foundation Phase in April 2008 and it is currently its third phase: Grid Technology Services for Operational GARUDA. This phase has been approved and funded for 3 years until July 2012.

The GARUDA project coordinator, CDAC, established in November 2008 an IGTF recognized Certification Authority which allows access to worldwide Grids for Indian Researchers.

GARUDA aims at strengthening and advancing scientific and technological excellence in the area of Grid and Peer-to-Peer technologies. It will also create the foundation for the next generation Grids by addressing long term research issues in the strategic areas of: knowledge and data management, programming models, architectures, Grid management and monitoring, problem solving environments, tools and Grid services.

The EU-IndiaGrid project [30], operating within the Sixth Framework Program of the European Commission, has played a bridging role between European and Indian Grid infrastructures and its successor, EU-IndiaGrid2 [30], aims at increasing the cooperation between European and Indian e-Infrastructures capitalizing on the EU-IndiaGrid achievements.

EU-IndiaGrid and EU-IndiaGrid2 are part of a group of projects, funded within the Sixth and Seventh Framework Programs for Research and Scientific Development of the European Commission, which aim at integrating the European Grid infrastructure with other regions in order to create one broad resource for scientists working on existing or future collaboration.

The EU-IndiaGrid project ran from 2006 to 2009. The leading responsibilities of the EU-IndiaGrid

Indian partners and the project bridging role between European and Indian e-Infrastructures gave to EU-IndiaGrid project the opportunity to be at the core of the impressive developments in India in the e-Infrastructures domain and to effectively contribute at improving cooperation between Europe and India in this area. In all these activities, the role and the contribution of EU-IndiaGrid partners, as well as the bridging role of the EU-IndiaGrid project, was particularly relevant and obtained full recognition at the highest level by representatives of the Indian Government and of the European Commission so contributing at supporting the improvement of the e-Infrastructures capabilities. Its successor EU-IndiaGrid2 will run from January 2010 to December 2011 and capitalizes on the EU-IndiaGrid achievements by acting as a bridge across European and Indian e-infrastructure to ensure sustainable scientific, educational and technological collaboration. The EU-IndiaGrid contribution in the Euro-India e-Infrastructures cooperation's scenario and the perspectives for EU-IndiaGrid2 can be well resumed in the words of Dr. Chidambaram, Principal Scientific Advisor to the Govt. of India, who gave the opening speech at the EU-IndiaGrid2 project launch: "I am happy to learn about the second phase EU-IndiaGrid2 project–Sustainable e-Infrastructures across Europe and India. The first phase has benefited immensely a variety of scientific disciplines including biology, earth science and the Indian collaboration for the Large Hadron Collider (LHC). The successful working of the initial phase of multi-gigabit National Knowledge Network, Indian Certification Authority, and participation in Trans-Eurasia Information Network (TEIN3) phase 3 are some of the important building blocks for supporting virtual research communities in India and their collaboration work with other countries."

Fig. 6 Map of the GARUDA infrastructure
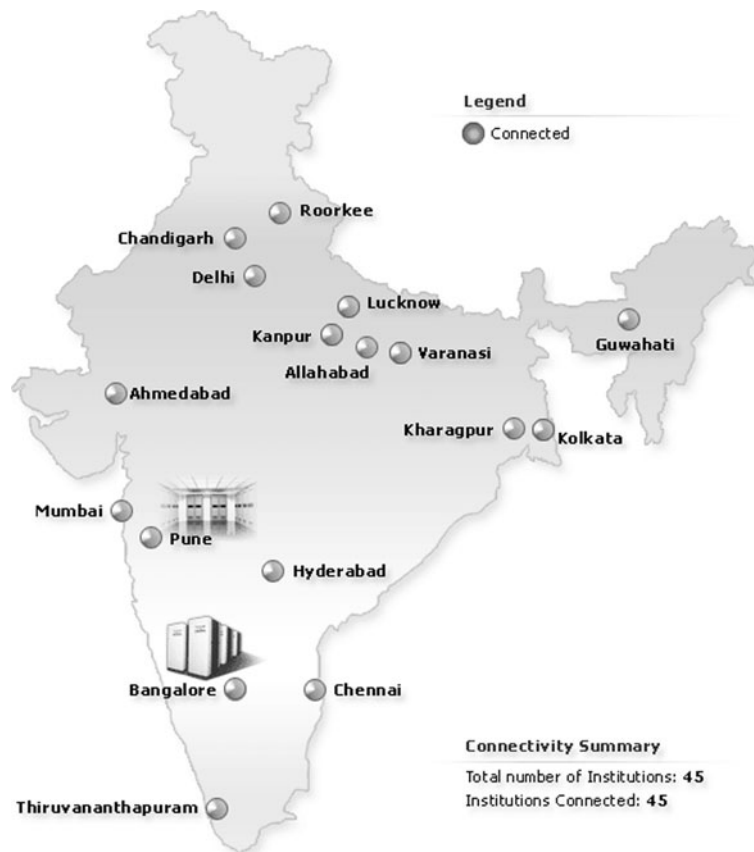
2.5.1 The GARUDA Grid Infrastructure

GARUDA, the National Grid Initiative of India is a collaboration of scientific and technological researchers on a nationwide Grid comprising of computational nodes, storage devices and scientific instruments. GARUDA today connects 45 institutions cross 17 cities in its Proof of Concept (PoC) phase with an aim to bring "Grid" networked computing to research labs and industry. Figure 6 shows the actual status of the e-Infrastructure.

2.6 Sub-Saharan Africa

With few exceptions, African universities and research centres lack access to dedicated global research and education resources because they are not connected to the global infrastructure consisting of dedicated high capacity regional networks. The consequence is that research and higher education requiring such access can currently not be conducted in Africa and the continent is not well represented in the global research community. This is witnessed by the world map of scientific divide (see Fig. 7) where territory size shows the proportion of all scientific papers (published in 2001) written by authors living there.

An important bottleneck is the lack of direct peering with other research and higher education networks. This bottleneck can be removed only by creating dedicated National Research and Education Networks (NRENs) connecting research and tertiary education institutions in each African country to a Regional Research and Education Network (RREN) interconnected to the peer infrastructures on other continents. In this context, a pioneering and very important role has been played by the Ubuntunet Alliance [31]. Incorporated in 2006, Ubuntunet gathers the following 12 NRENs in Eastern and Southern Africa: Eb@le (Democratic Republic of Congo), EthERNet (Ethiopia), KENET (Kenya), MAREN (Malawi), MoRENet (Mozambique), RwEdNet (Rwanda), SomaliREN (Somalia), SUIN (Sudan), TENET (South Africa), TERNET (Tanzania), RENU (Uganda), and ZAMREN (Zambia) and it is fostering the creation of new ones in Botswana, Burundi, Lesotho, Namibia, Mauritius, Swaziland, and Zimbabwe.

The mission of the Alliance is to secure affordable high speed international connectivity and efficient ICT access and usage for African
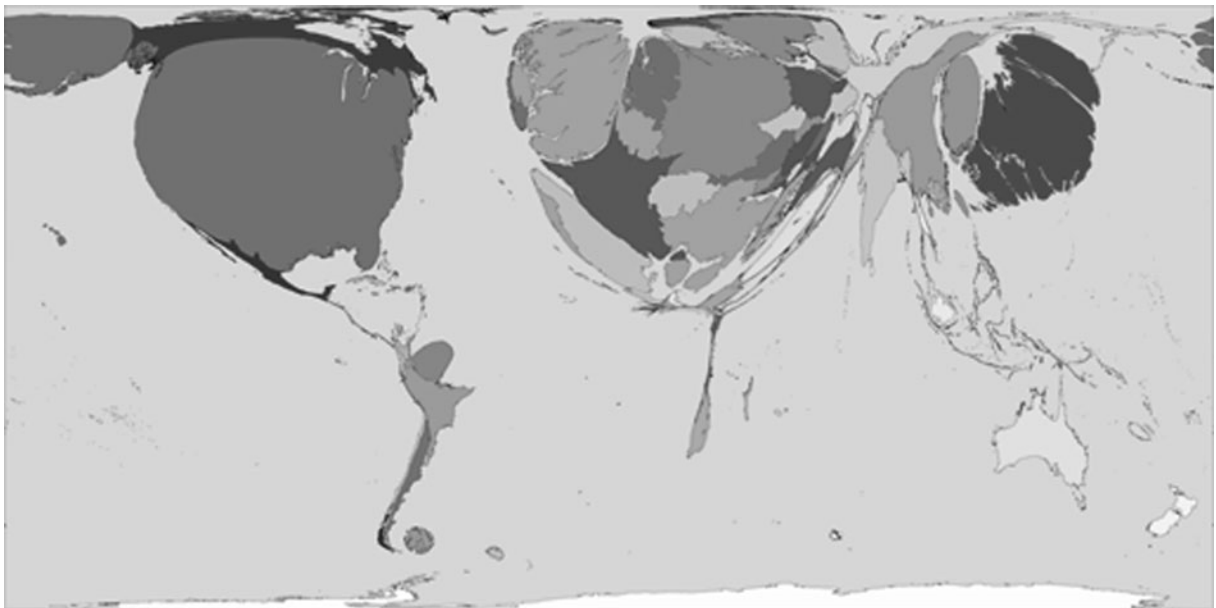

Fig. 7 World map of the scientific divide

NRENs. In this respect, Ubuntunet has been one of the stakeholders of the FEAST project [32] (Feasibility Study for African–European Research and Education Network Interconnection) that, between December 2008 and December 2009, has studied the feasibility of connecting African NRENs to the GÉANT network and has documented the relevant issues in the region inhibiting these enabling technologies.

FEAST has also paved the way for the creation of the AfricaConnect consortium that should take care, under the coordination of DANTE, of the creation, in the next 3–4 years, of a RREN in Sub-Saharan Africa at a total cost of 15 M€, 80% funded by the European Commission and the rest co-funded by the beneficiary countries.

Notwithstanding the large dissemination activities of strategic projects, such as IST-Africa [33], EuroAfrica-ICT [34], and eI-Africa [35], co-funded by the European Commission in the context of its Sixth and Seventh Framework Programs, the Sub-Saharan region of Africa has seen the least amount of activity in distributed computing initiatives. However, the recent advent of affordable international bandwidth, the reform of national telcoms policies and the subsequent construction of high-bandwidth national research networks in the early part of the first decade of the century has had a catalytic effect on interest in deploying e-Infrastructures in the region. These naturally have a scope well-beyond that of Grid computing projects for scientific research, but have been identified by researchers, higher-learning institutions, and governments in the region as enablers of collaboration and tools to reduce the effect of the digital divide discussed above.

As in other cases discussed in this chapter, scientific projects requiring significant infrastructure—in particular the Southern African Large Telescope (SALT, 2010) and the Karoo Array Telescope—were great stimuli of the interest in deploying networks and Grids in the region. The remote location of the scientific equipment and the wide geographic separation of the members of the collaborations using it were prime motivators, for example, for the development of the South African NREN. Data sharing considerations were long a concern, too, for the South

African participation to two experiments of the Large Hadron Collider. Two groups of research centres participate to the ALICE and ATLAS experiment, respectively, and the hub of medical and fundamental nuclear physics research undertaken at the iThemba Laboratories was one of the original drivers for experimenting with a national data and compute Grid.

South Africa is the only country in the Sub-Saharan region with a dedicated activity to coordinate distributed computing, which started with two projects centrally funded by Department of Science and Technology. These were the SANREN national research and education network and the Centre for High-Performance Computing, which was inaugurated in 2006. The plan for a high-speed network connecting the country's universities and national laboratories generated interest in the creation of a federated distributed computing infrastructure based on the Grid paradigm. The creation of a Joint Research Unit in mid-2008 was the start of this project, which aimed to integrate existing computing clusters and storage distributed in the institutes into a national Grid computing platform.

The South African National Grid (SAGrid) [36] by the start of 2010 consisted of a federation of seven institutes taking part in Grid operations and belonging to the SAGrid Join Research Unit with open activities under way for futher inclusion of other universities in the country. SA-Grid leans upon the high-bandwidth SANReN and its development was based in many ways on the experience acquired in Europe, starting with the EGI model and regional activities. The EGI middleware (gLite) and the EGI operational tools were adopted as standards at all sites, ensuring that the infrastructure would be easily used by Virtual Organisations operating on the EGI resources.

A major obstacle in the Sub-Saharan region was the lack of a CA accredited by IGTF. Since there is indeed no region of the IGTF responsible for Sub-Saharan Africa, the nearest Policy Management Authority (PMA) is that responsible for Europe and the Near East : EUGridPMA. A proposal to accredit a new CA for South Africa, the SAGrid CA, was accepted by EUGridPMA in 2009 and full accreditation is expected by mid 2011. To avoid delays, the INFN CA has assigned
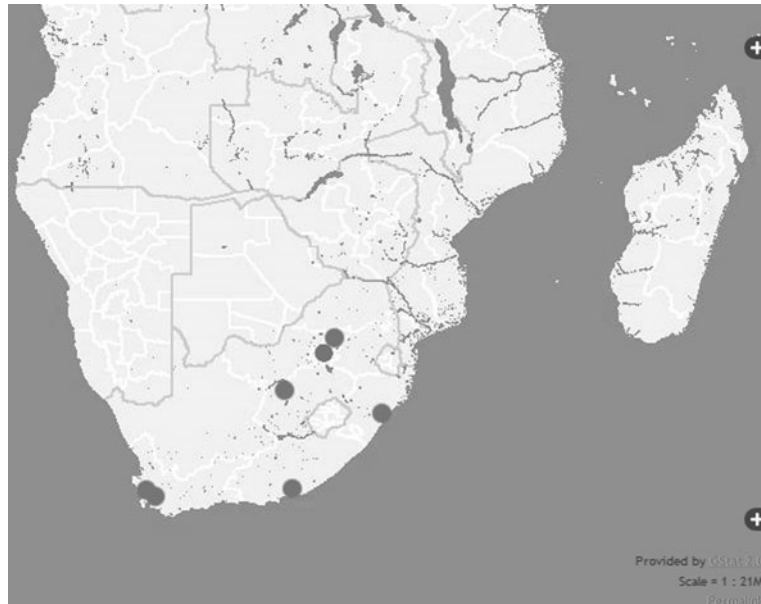
Registration Authorities in several South African institutes which are able to issue digital certificates for individuals and services locally.

Coordinated training and development events both in South Africa and the broader region, undertaken in collaboration with the GILDA t-Infrastructure [37], have expanded the base of competent site administrators and users, in concert with similar activities undertaken by the EUMEDGRID projects (see above). This foundation work is essential in developing the base of applications, technical experts and eventually (and most importantly) users in the region.

2.6.1 The SA-Grid Infrastructure

The SA-Grid Infrastructure, shown in Fig. 8, consists of 7 Resource Centres for a total of about 1,300 CPU cores and about 12 TB of storage.

3 Deploy e-Science Applications on e-Infrastructures: The Gridification Process

Application porting on e-Infrastructures is a highly important support service for the successful and efficient penetration of Grid technologies into the various research disciplines. Most of the major Grid projects recognized this and setup some form

of porting support for their user communities. Porting an application to an e-Infrastructure is a rather technology-oriented step that can be carried out only after the beneficiaries of an application have recognized that Grids could improve the application in some way. Furthermore, the beneficiaries must also identify the Grid type to which their application is suited. Reaching this level of understanding is difficult and one needs to analyze the problem from several perspectives. The porting process itself then can be carried out in a rather mechanical way which can be described with the following workflow:

- Carrying out a technical analysis of the current form of the application: how it is used on a single computer/supercomputer/cluster/Grid, what is its structure, the interfaces between components, between the application and its environment, between the application and its users. How the application now is embedded into the daily routine work of users.
- Understanding the "Grid vision" of the community: what do the beneficiaries expect from the Grid-enabled version of the application, what are their motivations and priorities for porting this application to the Grid, e.g. better performance, scalability, accessibility, fault tolerance, quality of service, etc. What changes

they can tolerate in their routine work to get a Grid-enabled version of the application.

- Analysing the application structure from the Grid point of view: identifying the compute/data intensive parts, analysing the dataflow, collecting or measuring required hardware capacities, understanding licensing schemes, understanding compilation methods, programming interfaces, etc.
- Defining one (or more) porting scenarios by which the application could be ported to the Grid in order to meet the users' demands. Discussing the scenarios with the application porting team and with the end users. Usually several iterations are required to select the final scenarios that represent the most optimal solution for every involved party. The scenario makes recommendations on Grid services that are relevant to the application, identifies infrastructural and software components that current Grids don't provide but are required for the successful adoption and so need to be developed. The foreseen steps of porting and development are associated with estimated human resource requirements. The advantages and disadvantages of the scenario over other solutions should be provided too, and this can later serve as a reference for the future extension of already Grid-enabled (gridified) application.
- Completing the practical development steps of the porting. This may include the extension of the original code with new APIs, wrapping the code with scripts, adapting the code to the Grid with porting tools, development of new infrastructure services, new interfaces, integrating various components. Component and application level testing and debugging, performance analysis, documentation closes the core porting cycle.

This application porting methodology has been applied in the e-Infrastructures covered in this paper with some slight difference according to their different contexts. In the next sub-sections we describe how the "gridification" process has been implemented in Europe, where the Grid knowledge could be just considered high, and in the other regions of the world. It's important to notice that the best way to promote the Grid model outside Europe is the identification and gridification of scientific applications that could have a great benefit from the e-Infrastructure. For this reason, a big effort has been devoted to look for scientific applications able to satisfy this requirement selecting, in a proper way, which applications to support.

3.1 The Gridification Process in Europe

At the beginning of the EGEE project, two pilot applications were identified to test the infrastructure and its services. These applications came from the High Energy Physics (HEP) [38] and the Life Sciences domains. Now, many scientific applications exploit the distributed computing and storage resources of the EGEE infrastructure to effectively run data challenges and share knowledge with other partners regardless their geographical location.

In its third phase, the EGEE project flagged application porting as a critical service for the successful adoption of Grid solutions by new users and new communities. One of the activities of EGEE-III project (User Community Support and Expansion) includes "Application Porting Support" among its generic user support activities (besides Virtual Organization Support and Direct User Support). The goal of the EGEE Application Porting Support group was to aid developers in effectively porting Virtual Organizations' (VO) applications to the Grid.

The EGEE Grid Application Support Centre (GASuC) [39] was implemented as a distributed network of centres coordinated by the Hungarian partner MTA SZTAKI and involved researchers with long-term experience in applications porting from the following institutes:

- Universidad Complutense de Madrid (UCM), Madrid, Spain;
- Instituto Nazionale di Fisica Nucleare (INFN), Catania, Italy;
- Commissariat a l'Energie Atomique (CEA), Paris, France ;
- Consejo Superior de Investigaciones Cientificas (CSIC), Madrid, Spain;

- Academia Sinica Grid Computing Centre (ASGC), Taipei, Taiwan;
- University of Melbourne, Melbourne, Australia.

The Grid Application Support Centre (GASuC) provided assistance to Grid users and application developers during the application gridification process. GASuC helped to identify suitable approaches, tools for the porting process and apply best patterns, and practices in order to get codes running on production Grids infrastructures. In order to speed up the porting of scientific applications in the production quality infrastructures, intensive workshops and personalized training events were organized by the team to ensure that application owners became expert on porting tools and could perform the porting process with the help of the porting team quickly and efficiently. After the complexity of the porting problem was understood, a porting expert was assigned to the client. The expert's responsibility was to define the roadmap for the porting, to allocate resources, to coordinate the porting process and to ensure that the final application will meet the targeted requirement. For further information about the support cycle implemented by GASuC, "seen" by the owner of an application, please refer to the "Gridification process" section (http://www.lpds.sztaki.hu/gasuc/?m=1).

3.2 The Gridification Process in Other Regions of the World

A slightly different procedure has been adopted by other projects, such as EELA-2, EUAsiaGrid and EUMEDGRID, to identify and support new applications on their Grid infrastructures. An application questionnaire has been created to gather information from the various potential partners about their applications. Applications are then selected taking in consideration several criteria as number of involved institutions from Europe and region of interest for the project (Latin America, Mediterranean, etc.), suitability for Grid deployment, easiness of gridification, Grid added value, resources (CPU, storage) commitments of the Institutions involved, usage of the infrastructure (number of jobs and frequency of runs) and potential outreach / impact (in the scien-

tific community, industry, socially in the country, towards policy / decision makers). These criteria, mapped on quantitative ranges and properly weighted, allowed evaluating and ranking each application received through the questionnaire. Then, selected applications are supported to be deployed in the Grid infrastructure and, hence, it is assumed that a selected application might go through the following stages:

1. *Ungridified / ready for standalone use*: The application has been developed and runs on a local cluster. The application cannot run unmodified or without adaptations in a GRID for multiple reasons: needs recompilation against specific OS or libraries available in GRID environments, its data needs to be prepared for parallel processing, the algorithms need to be adapted to combine results obtained in a parallel manner, etc.;
2. *Gridified / ready for gLite*: The application has been adapted so that it can be deployed in a gLite-based Grid environment. This usually consists on (a) defining how data is cut up to be processed in a parallel manner and (b) creating scripts and adapter programs to coordinate launching jobs and assembling resulting data. In more complicated cases, this might require actual modification of the applications, using Grid APIs, etc.;
3. *Deployed / ready for validation*: The application has been deployed in the Development Grid environment (GILDA), is ready to accept job submissions and start validation. At this stage applications must conform with specific GRID deployment procedures and standards. For instance, sites usually have deployment procedures for applications to be run in test mode different than for applications to run in production mode. Executable and data files must be placed in locations accessible to jobs whenever they are submitted to the queuing systems, etc.;
4. *Ready for deployment:* the application has been deployed and validated on the development infrastructure (GILDA) and is ready to be installed into the production infrastructure;
5. *Production*: The application is deployed and validated and is ready to start scientific

Fig. 9 Application life-cycle towards the production stage
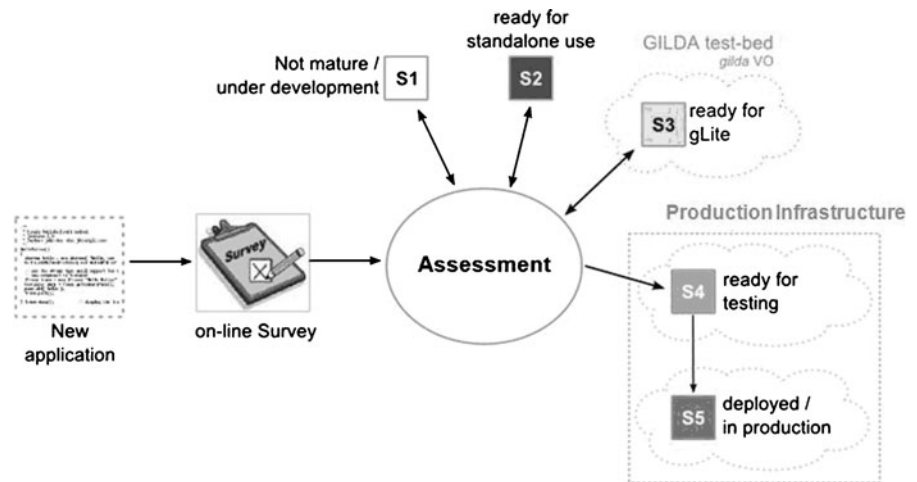
production. Jobs are being lunched in a regular manner, data is being produced and systemic non-functional qualities must be monitored (performance, accessibility, security, etc.).

Different applications may enter into this life cycle at different stages, depending on their degree of maturity (see Fig. 9).

The applications deployment was supported by a group of very skilled people, the "gridification team", in charge of helping the application developers to interface their codes with the Grid middleware in the e-Infrastructure. Such "helpdesk" approach was/is carried out through both virtual meetings and specific on-demand retreats of one or two weeks. During such retreats, called Grid Schools, application experts work in close collaboration with tutors to interface their applications with the Grid middleware. These events proved to be very effective and decisively helped in speeding-up the porting of several scientific applications on the projects' e-Infrastructure.

4 Scientific Domains and Applications Supported: A Worldwide Overview

The evolution of the e-Infrastructure model around the world has reached a high maturity level and, nowadays, there are several virtual research communities, spread in all the continents, profiting of this new kind of scientific environment. This section shows, per scientific domain

and per geographic area, the distribution of the applications deployed on the e-Infrastructures described in the Section 1 providing, for the first time, a worldwide overview of the virtual research communities exploiting the existing e-Infrastructures with their scientific applications. Applications have been classified according to their scientific domains and their geographical origin (continent and country of the scientists developing/using them).

In the following sub-sections, a short description of some relevant pilot applications for each of the scientific domains considered is provided. These applications benefits of the e-Infrastructure in several ways, from the large amounts of computation time available to the large amounts of storage available, from the creation of human collaboration among scientific communities of researchers that work together to allow brilliant scientists to stay in their regions and contribute significantly to cutting edge scientific activities and reducing the digital divide. The number of the applications profiting of the e-Infrastructures and their spread in several scientific domains could be considered two important clues suggesting that scientific research will be mainly based on e-Infrastructures in the near future.

4.1 Virtual Research Communities and Scientific Domains Involved

The applications using the e-Infrastructures described in Section 1 have been classified

Fig. 10 Applications distributed per Scientific Domain
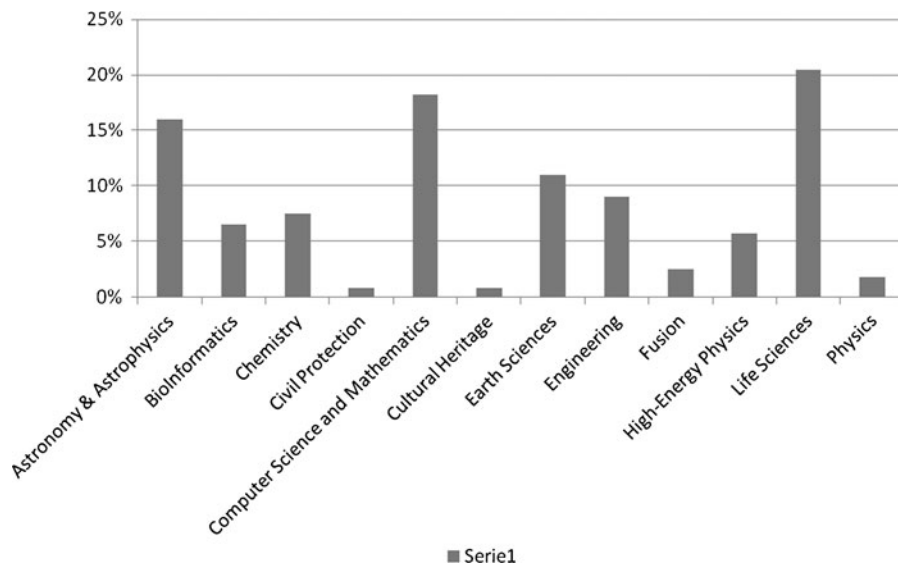
according to the following taxonomy: Astronomy & Astrophysics, Bioinformatics, Chemistry, Civil Protection, Computer Science and Mathematics, Cultural Heritage, Earth Sciences, Engineering, Fusion, High-Energy Physics, Life Sciences and Physics.

Figure 10 shows the applications distribution per scientific domain. The distribution per geographic origin is shown in Fig. 11 (a world-wide vision) and in Figs. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 and 23 (per geographic region).

4.2 Astronomy and Astrophysics

4.2.1 CORSIKA (Europe, Germany–Switzerland–Italy)

CORSIKA (COsmic Ray SImulations for KAscade) [40] is a program for detailed simulation of extensive air showers initiated by high energy cosmic ray particles. Protons, light nuclei up to iron, photons, and many other particles may be treated as primaries. The particles are tracked through the atmosphere until they undergo reactions with the air nuclei or—in the case of instable secondaries—decay. The hadronic interactions at high energies may be described by several reaction models alternatively: The VENUS, QGSJET, and DPMJET models are based on the Gribov–

Regge theory, while SIBYLL is a minijet model. The neXus model extends far above a simple combination of QGSJET and VENUS routines. The most recent EPOS model is based on the neXus framework but with important improvements concerning hard interactions and nuclear and high-density effect. HDPM is inspired by findings of the Dual Parton Model and tries to reproduce relevant kinematical distributions being measured at colliders. Hadronic interactions at lower energies are described either by the GHEISHA interaction routines, by a link to FLUKA, or by the microscopic UrQMD model. In particle decays all decay branches down to the 1% level are taken
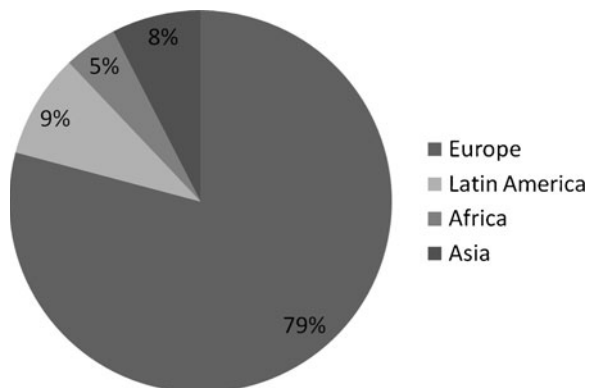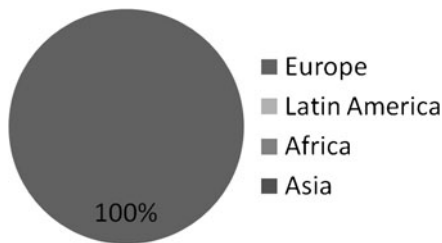

Fig. 11 Applications distributed per continents

Fig. 12 Astronomy and astrophysics

Fig. 14 Chemistry

into account. For electromagnetic interactions a tailor-made version of the shower program EGS4 or the analytical NKG formulas may be used. Options for the generation of Cherenkov radiation and neutrinos exist. CORSIKA may be used up to and beyond the highest energies of 100 EeV.

4.2.2 LOFAR (Europe, Netherlands)

LOFAR (Low Frequency Array) [41] is a unique instrument developed from the ambition of Dutch astronomers to observe the earliest stages of our universe. For this purpose, a radio telescope is needed that is a hundred times more sensitive than existing telescopes. The LOFAR radio telescope consists of thousands of sensors in the form of simple antennas. These antennas are distributed over Europe and connected to a supercomputer via an extensive glass fiber network. It was soon realised that LOFAR could be turned into a more generic Wide Area Sensor Network and now sensors for geophysical research and studies in precision agriculture have been incorporated in LOFAR as well.
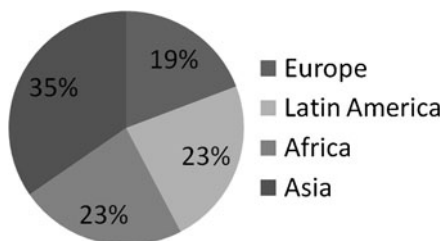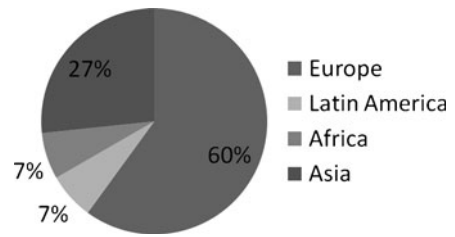
4.3 Bioinformatics

4.3.1 GAP Virtual Screening Service (Asia, Taiwan)

Dengue fever is an epidemic disease in more than 100 countries. The World Health Organization confirmed that around 2.5 billion people, about 40% of the world's population, are now at risk. Fifty million cases of Dengue infection are estimated to occur annually world-widely and around 95% of those are related to children less than 15 years old in South-Eastern Asian countries. EUAsiaGrid has successfully deployed the Virtual Screening Service (VSS) based the Grid Application Platform (GAP) [42] for the Avian Flu DC2 Refine disease. In March 2009 only, a total of 1,111 CPU-days were consumed on the EUAsiaGrid infrastructure and more than 160,000 output files with a data volume of 12.8 GB were created and stored in a relational database. The scientific collaboration composed by EUAsiaGrid partners now seek for potential drugs to cure the Dengue Fever using the GVSS (GAP Virtual Screening Service; Fig. 24).

The target for Dengue Fever Activity in EUAsiaGrid that has been focused on is the NS3
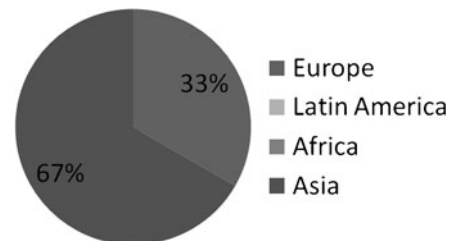

Fig. 13 Bioinformatics
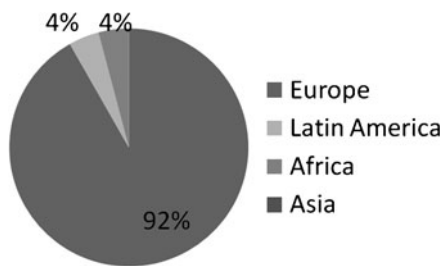

Fig. 15 Civil protection

Fig. 16 Computer science and mathematics

Fig. 18 Earth sciences

protease (Nonstructural protein 3), also known as an enzyme critical for virus replication, since its aminoacid sequence and atomic structure are very similar among the different diseases caused by flaviviruses. Grid-enabled high-throughput screening for structure-based computational methods has been used to identify small molecule protease inhibitors. ASGC has developed the GVSS application package integrated with the gLite-based services DIANE2 and AMGA, and used the Autodock as the simulation docking engine. User can submit Grid jobs as well as select compounds and targets in the application. GVSS also enables the user to upload their compounds and targets to do the same docking process, compare the significant components, and verify them trough in-vitro experiments.

4.3.2 The Rosetta "ab initio" Software to Predict the 3D-structure of Never Born Proteins (NBP; Europe, Italy)

Rosetta is an "ab-initio" protein structure prediction software [43] which is based on the assumption that in a polypeptide chain local interactions
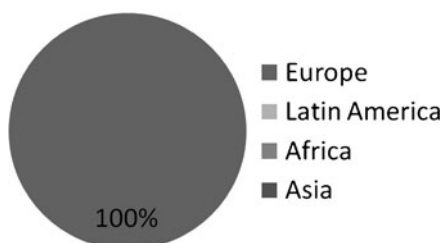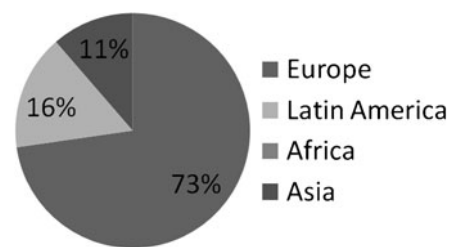
bias the conformation of sequence fragments, while global interactions determine the three-dimensional structure with the minimal energy compatible with the local biases. To derive the local sequence–structure relationships for a given aminoacid sequence (the query sequence) Rosetta ab-initio uses the Protein Data Bank to extract the distribution of configurations adopted by short segments in known structures. The latter is taken as an approximation of the distribution adopted by the query sequence segments during the folding process.

The protein structures predicted using Rosetta ab-initio during some data challenges have been compared with the structures predicted for the same proteins using the Early/Late stage software developed by Roterman et al. at the Medical College of the Jagiellonian University (Krakow, Poland). One of these proteins, protein number 4091was selected to be produced and characterized experimentally. The protein was expressed in the bacterium Escherichia coli in a laboratory of the Peking University in China and its structure is in the process of being determined experimentally by nuclear magnetic resonance methods.
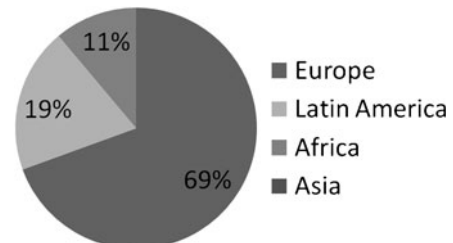

Fig. 17 Cultural heritage
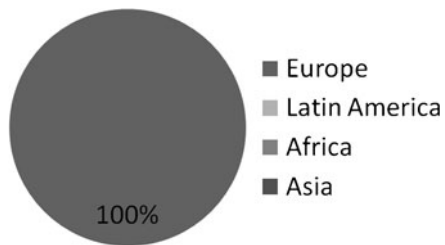

Fig. 19 Engineering

Fig. 20 Fusion

Fig. 22 Life sciences

4.4 Chemistry

4.4.1 Gromacs (Latin America, Brazil)

GROMACS [44] is a versatile package to perform molecular dynamics, i.e. simulate the Newtonian equations of motion for systems with hundreds to millions of particles. It is primarily designed for biochemical molecules like proteins and lipids that have a lot of complicated bonded interactions, but since GROMACS is extremely fast at calculating the non-bonded interactions (that usually dominate simulations) many groups are also using it for research on non-biological systems, e.g. polymers. GROMACS supports all the usual algorithms that researchers would expect from a modern molecular dynamics implementation. The input/output data depends on the complexity of the molecule being studied and depends on the number of steps configured for the experiment. For an average molecule, the input size is around 100 MB and the output data is around 10 GB.

4.4.2 Haddock (Africa, South Africa)

HADDOCK (High Ambiguity Driven protein–protein DOCKing) [45] is an information-driven flexible docking approach for the modeling of bio-
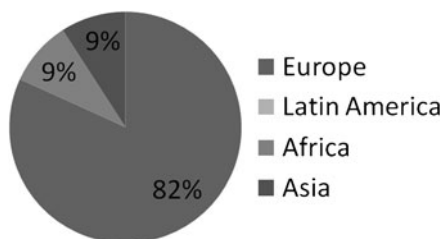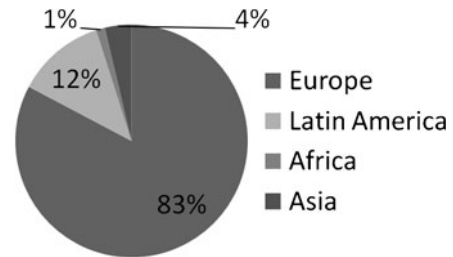
molecular complexes. HADDOCK differs from ab-initio docking methods because it encodes information from identified or predicted protein interfaces in ambiguous interaction restraints (AIRs) to drive the docking process. HADDOCK can deal with a large class of modeling problems. It has first been developed for protein–protein docking but has since been applied to a variety of protein–ligands, protein–peptides, and protein–DNA/RNA complexes.

4.5 Computer Science and Mathematics

4.5.1 MeshStudio (Europe, Italy)

3D object analysis is of particular interest in many research fields and the popularity of 3D computer models has grown during the last decades due to the advance of 3D scanning technology. Mesh-Studio [46] allows the interaction with 3D objects, represented by triangular meshes, by means of a GUI (Graphical User Interface). The main features of the interface are: (1) control of the orientation of the triangle facets of the mesh; (2) control of the polygons rendering; (3) control of
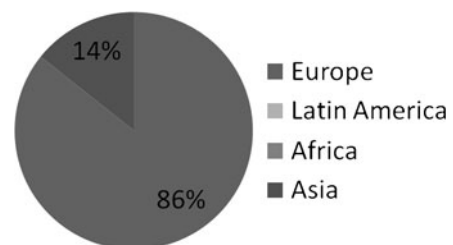

Fig. 21 High-energy physics
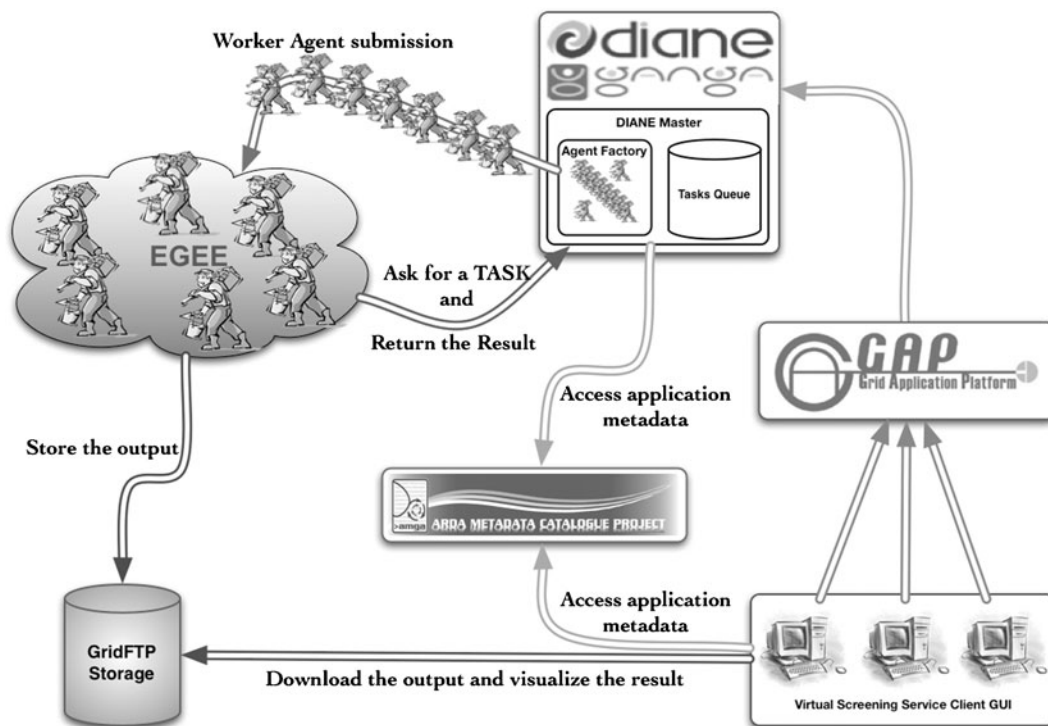

Fig. 23 Physics

Fig. 24 GAP virtual screening service deployed in the Grid infrastructure

the lightning; (4) control of rotation, translation and scaling.

MeshStudio implements a new automatic algorithm for 3D edge detection on large objects. The detection step is performed by evaluating the curvature of the surface considering those elements falling within a neighborhood, defined by a set of radius values. The algorithm does not require any input parameter and returns a set of elements patches that represent the salient features on the model surface regarding the best radius value. The computational complexity of the algorithm required a parallel implementation in order to process high resolution meshes in an acceptable time and, for this reason, the application has been deployed in a Grid infrastructure.

4.6 Cultural Heritage

4.6.1 Archaeogrid (Europe, Italy)

The approach of modern archaeology to the study of the evolution of ancient human societies is

based on the acquisition and analysis of many types of data. It is well known that in archaeology large use is made of digital technologies and computer applications for data acquisition, storage, analysis and visualization. The amount of information coming from remote sensing, from acquisition of 3D artifacts' images by scanners laser, from GPS precise reference of geographical points, and from other sciences are increasing at a large extent the amount of data that are needed to be stored and made available for analysis. Such data must, however, be analyzed if they have to become valuable information and knowledge. The data analysis use advanced methods developed in mathematics, informatics, and physics and in other natural and human sciences. Moreover, the use of Virtual Archaeology as a new approach to the narration and visualization in Archaeology, is expanding rapidly, not only in the museum and archaeology professions, but also in the broadcast media, tourism and heritage industries. The inevitable result of this is an exponential increase of the amount and complexity of information that

must be acquired, transferred, stored, processed and analyzed.

ArchaeoGRID [47] is a very powerful application for the archaeological research helping to manage this information using the power of Grid infrastructures and implementing the most common data analysis methods.

4.6.2 Astra (Europe, Italy–Spain)

The ASTRA (Ancient instruments Sound/Timbre Reconstruction Application) project [48] runs since 2006 as a collaboration among the Conservatory of Music of Parma the Conservatory of Music of Salerno, the Division of Catania of the Italian National Institute of Nuclear Physics and other distinguished international partners. Since 2010 it is hosted at the University of Malaga. The project's aim is to bring back to life the sound/timbre of ancient musical instruments. By applying the physical modelling synthesis, a complex digital audio rendering technique which allows modelling the time-domain physics of an instrument, the experts who are carrying out the project can recreate models of some musical instruments that were lost for ages (since hundreds and hundreds of years) and reproduce their sounds by simulating their behaviours as mechanical systems. So far, ASTRA effort was focused in the reconstruction of the sound of plucked ancient string instruments. The history of the plucked string instruments dates back since 6 millennia ago in Mesopotamia, where the instruments were built by limited technical skills, craftsmanship, and material knowledge of the ancient times. In the course of history, plucked string instruments were first constructed by joining natural environmental objects such as wooden sticks, gourds, and turtle shells, and later, gradually, by manufacturing selected processed and specifically designed parts. The idea and the mathematical concepts behind remodelling early instruments have been around since the 1970s but the amount of information that needed to be processed to put these theories into practice was too great for earlier computers. Nowadays, researchers can access the distributed computing power of the Grid, based over research and academic networks, and speed up the sound reconstructions. The starting point is represented by all archaeological findings about the instrument such as fragments from excavations, written descriptions, pictures on ancient urns, etc. The PhM Synthesis uses a set of mathematical equations and algorithms that describe the physical materials used in the ancient instruments to generate the physical sources of sound. For instance, to model the sound of a drum, a formula for how striking the drumhead injects energy into a two dimensional membrane would be created. Then, the properties of the membrane (mass density, stiffness, etc.), its coupling with the resonance of the cylindrical body of the drum, and the conditions at its boundaries (e.g., a rigid termination to the drum's body) would describe its movement over time and how it would sound. Similar modelling can be carried out with instruments such as violins, though replicating the slip–stick behaviour of the bow against the string, the width of the bow, the resonance and damping behaviour of the strings, the transfer of string vibrations through the bridge, and, finally, the resonance of the soundboard in response to those vibrations.

4.7 Earth Science

4.7.1 WRF4G (Asia/Europe, Thailand–Spain)

WRF (Weather Research and Forecasting) for GRID (WRF4G) [49] is a GRID-enabled framework for the WRF Modeling System. The Weather Research and Forecasting Model [50] is a next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. It features multiple dynamical cores, a 3-dimensional variational (3DVAR) data assimilation system, and a software architecture allowing for computational parallelism and system extensibility. WRF is suitable for a broad spectrum of applications across scales ranging from meters to thousands of kilometers.

WRF allows researchers the ability to conduct simulations reflecting either real data or idealized configurations. WRF provides operational forecasting a model that is flexible and efficient computationally, while offering the advances in physics, numeric, and data assimilation

contributed by the research community. WRF has a rapidly growing community of users.

The deployment of WRF to the GRID has been an achievement allowing to the WRF's user community exploiting the computational and storage resources provided by the GRID to carry on more ambitious experiments. The typical experiments are, multi-physics, parametric sweep, initial condition perturbations that suit well for GRID computing.

4.7.2 Seismic Travel Time Tomography (Asia, India)

The application "Seismic Travel Time Tomography" known as SeisTom of Earth Sciences [51], has been developed by the Seismic Data Processing Team, Scientific Engineering & Computing Group at the Centre for Development of Advanced Computing of Pune (India). It reconstructs Earth velocity model from travel time. The reconstructed model heavily depends on the correctness of the travel time picked for the purpose of reconstruction.

Seismic imaging, popularly known as tomography, borrowing the term from medical sciences, produces the raster image of the internal structure by combining information from a set of projections obtained at different viewing angles. To engineer high resolution and accurate solutions to mining problems, seismic travel time tomography can yield a high resolution image of the subsurface to provide information about lithologic characterization, fracture void or gallery detection, fluid monitoring, etc. The measured characteristics of received wave energy include amplitude and travel time.

A number of seismic tomographic studies have explained how seismic velocities can be estimated along raypaths between boreholes and demonstrated their results with real field examples. Imaging techniques used in these applications include back projection, back propagation and ray tracing.

The resolution of the estimated images in seismic tomography depends on the inversion schemes used, completeness of the data sets and the initially assigned values of the medium parameters in the iterative reconstruction techniques. Since optimal solutions are always aimed at, even from poor initial models, global optimization and search technique such as real-coded genetic algorithm is applied for solving the present imaging problem.

Using seismic travel time tomography by parallel real-coded genetic algorithm, the subsurface can be imaged successfully. The complex synthetic examples demonstrated the authenticity of this transmission seismic imaging technique. This tomography algorithm using RCGA is highly compute intensive, because travel time computation must be done for thousands of models. The advantage with the GA based tomography algorithm is that it can be easily distributed on HPC and Grid platforms.

While most of the existing tomographic imaging techniques are quite versatile, the RCGA algorithm is found to work over a large domain of problems even with very poor initial guess.

The application was tested on Param Padma (1 TFlop machine). It gave satisfactory result but high resolution of reconstructed model was required for greater details and better visualization. But the estimated time on Param Padma for such detailed model output was quiet large. So it is has been ported on the Grid exploiting the powerful of the EU-IndiaGrid infrastructure.

4.8 Engineering

4.8.1 Industry@Grid (Latin America, Brazil)

Developed by the Production Engineering Department of the Federal Center for Technological Education Celso Suckow da Fonseca of Rio de Janeriro (Brazil), Industry@Grid [52] is a Grid application composed by a set of Grid modules aiming at solving industrial operations problems. Gridified during the first year of the EELA-2 Project, the Product Mix Module of Industry@Grid is currently used by the plastic film industry in order to achieve a more profitable product mix distribution. The Product Mix Problem is to find out which products to include in the production plan and in what quantities these should be produced in order to maximise profit, market share or some other goal.

In a real case, a plastic film production line is able to produce a few hundreds of distinct

products, varying weight, painting and material cutting. In the real life, the optimisation of the product mix yields to a problem that has thousand variables and more than tens of hundreds constraints. This kind of problem can be solved by an ordinary personal computer in a matter of seconds. However, besides the large number of variables and constraints, each production line generates scrap during its production process through extrusion and cutting, which increases the complexity, since variability is introduced. This new problem needs now a reasonable amount of time and computer power to be tackled.

One alternative to solve this new problem is to use Monte Carlo simulation methods, where we use the computer to create several instances or configurations of the scrap generated during the production process. So, the Industry@Grid module generates several instances of the new problem in order to maximise contribution margin of each item to the overall profit, simulating the scrap distribution for every instance and solving them using the Grid. In the context of the company environment, every product mix sample needs an execution of a 50,000 instances, amounting to over 16 CPU days for a single study cycle. Thanks to the Latin American Grid infrastructure, this special case of the product mix problem can be solved in an affordable time, allowing an operational margin increase.

4.8.2 MAVs-Study (Latin America, Argentina)

MAVs-Study (Biologically Inspired, Super Maneuverable, Flapping Wing Micro-Air-Vehicles) [53] is an application jointly developed by the Argentinian Universities of Río Cuarto and Córdoba to study the flight of insects and small birds in order to design, in the future, micro air-vehicles able to be used in different contexts such as atmospheric studies, fire detection, inspections of collapsed buildings, hazardous spill monitoring, inspections in places too dangerous for humans, and planetary exploration, among others. Flying robots, that have maneuvering capabilities similar to insects, the more agile flying creatures on Earth, have ceased to be a subject of science fiction and are becoming real. However, there are still major engineering problems to be solved such

as, for instance, miniature energy sources, ultra-light materials, and the complete understanding of unsteady aerodynamics at low Reynolds numbers.

Unlike conventional vehicles such as airplanes, dirigible airships, and helicopters, which are studied by steady and linear aerodynamics theories, insects use unsteady and highly non-linear unconventional aerodynamic mechanisms, which are effective at low Reynolds numbers. They are, among others, the delayed stall, the rotational lift generated by the rotation of the wings, and the wake capture.

Due to the unsteady, non-linear, and three-dimensional nature of the flow generated by the flapping of the wings, the aerodynamic model used in MAVs-Study is a modified version of the method known as "unsteady vortex lattice method" (UVLM), a generalization of the "vortex lattice method". MAVS-Study allows to study the non-steady aerodynamics associated with the flow generated by a fly (*Drosophila melanogaster*) in "hover". The combination of the kinematics model with the aerodynamics model, together with a pre-processor able to generate the geometry of the insect wings, allow to predict the flow field around the flapping wings, estimate the space-time distribution of vorticity of the bounded sheet of the wings, estimate the distribution of vorticity and shape of the wake emitted from the sharp edges of the wings, predict the aerodynamic loads on the wings, and take into account all possible aerodynamic interferences.

MAVS-Study has been developed in a Windows Operating System and it could be deployed in a Grid infrastructure thanks to the successful porting of the gLite middleware to the Microsoft Windows platform, done by the Italian National Institute of Nuclear Physics (INFN) in the context of EELA-2 project. Thanks to this work, a completely new set of e-Science applications, until now excluded by the Linux-based Grid world, can today benefit of the powerful Grid infrastructures.

4.8.3 SACATRIGA (Africa, Morocco)

The Sub-channel Analysis Code for Application to Triga research reactors [54] is a preliminary code devoted to study the thermal hydraulic behavior of Moroccan TRIGA MARK II research

reactor. Sub-channel approach is used to model the geometry of the reactor core. Over each sub-channel temperature, pressure and flow rate are predicted by solving Navier–Stokes equations coupled to the energy conservation equation. The application requires MPI compiler for Fortran90 and it has been deployed on the EUMEDGRID infrastructure. A graphic interface is also available on the infrastructure which makes use of the GENIUS portal.

4.9 High-Energy Physics

4.9.1 LHC Experiments (International)

All the experiments at the CERN Large Hadron Collider are all run by international collaborations, bringing together scientists from institutes all over the world. Each experiment is distinct, characterized by its unique particle detector.

The two large experiments, ATLAS [55] and CMS, are based on general-purpose detectors to analyze the myriad of particles produced by the collisions in the accelerator. They are designed to investigate the largest range of physics possible. Having two independently designed detectors is vital for cross-confirmation of any new discoveries made.

Two medium-size experiments, ALICE and LHCb [56], have specialized detectors for analyzing the LHC collisions in relation to specific phenomena.

The last two experiments, TOTEM [57] and LHCf, are much smaller in size. They are designed to focus on 'forward particles' (protons or heavy ions). These are particles that just brush past each other as the beams collide, rather than meeting head-on

The ATLAS, CMS, ALICE and LHCb detectors are installed in four huge underground caverns located around the ring of the LHC. The detectors used by the TOTEM experiment are positioned near the CMS detector, whereas those used by LHCf are near the ATLAS detector. All the LHC experiments make heavy use of the Grid infrastructures available in the various regions of the world and represent their most important customers.

4.9.2 The Pierre Auger Observatory (Europe/Latin America, Argentina–Italy–Mexico)

The Pierre Auger Experiment [58] is designed to study the cosmic rays radiation in the ultra high-energy regime (10^{18}–10^{20} eV) and is currently taking data as the deployment goes on. An intense simulation activity is planned in order to analyze the recorded data, which show a large variability in the primary energy (over 2 orders of magnitude), in zenith angles, presumably also in the primary mass. In addition, showers with the same primary energy–zenith–mass have different longitudinal and lateral profiles, because of the intrinsic stochastic development of the cascade in the atmosphere.

The southern Auger Observatory, located near the small town of Malargüe (in the Argentinian province of Mendoza) is nearly completed and is taking data regularly.

A large number of simulated and reconstructed showers is in this case required by the need of covering all the available phase space (energy–zenith–mass) with simulated data in order to understand the real data behavior.

The number of required simulated data increases further on, if one wants to use different hadronic models to describe the data, or if one wants to face particular tasks (such as trigger or acceptance studies, in which a time-dependent detector simulation is desired).

At the present status, two simulation codes, CORSIKA (COsmic Ray SImulations for KAscade) and Aires (AIRshower Extended Simulations), allow a complete tri-dimensional treatment of the development of the cascade, but need long CPU times to generate a shower, especially at the highest energies. Other codes, such as Conex, need shorter CPU times, but allow a mono-dimensional simulation of the cascade, which is suitable only for fluorescence detector analysis or for fast hybrid simulations. The cited codes are all used inside the Pierre Auger collaboration together with the official simulation and reconstruction tool (Auger Offline).

The possibility of using this specific software in a Grid environment allowed to speed up every step of the simulation–reconstruction

chain including the so-called hybrid analysis, never performed before. In a final perspective, e-Infrastructures will speed up the analysis of the real data recorded at the Pierre Auger Observatory.

4.10 Life Sciences

4.10.1 Wisdom (Europe, France–UK–Ireland)

The WISDOM initiative [59] started in 2004 as a collaboration between Fraunhofer Institute SCAI and CNRS: the initial idea was to use the computing resources offered by the EGEE project to select drug-like molecules active on a biological target to fight malaria. The choice of malaria is due to the fact that tropical and neglected diseases suffer from a lack of Research and Development because of the cost to bring to market new drugs. A way to reduce the cost is to develop in silico approaches, which are significantly cheaper than in vitro approaches. An important step in the drug discovery process is virtual screening, which is about selecting drug-like molecules active on a specific biological target by computing the binding energy of the molecule to the target active site. The prerequisite for the use of virtual screening is to know the three-dimensional structure of both the drug-like molecules and the target active site. The 3D structures of more than 3 million chemical compounds are now available in public databases like ChemBridge and ZINC while the Protein Data Base provides the structure of more than 50,000 proteins of biological interest. Since 2004, the WISDOM initiative has successfully deployed large scale virtual screening computations on Grid infrastructures in order to find new drugs against malaria, avian flu and diabetes. Meanwhile, it has also grown into a multidisciplinary collaboration of biologists, biochemists, bioinformaticians and e-scientists from Africa, Asia and Europe. Several thousands of CPUs have been used on e-infrastructures like AuverGrid in France, EUChinaGrid and TWGrid in Asia, EUMedGrid in Africa, OSG and Digital Ribbon in the United States and of course EGEE which has provided the majority of the resources. About 20% of the drug-like molecules selected in silico have been confirmed by in vitro tests to be active inhibitors and most promising molecules have been patented.

4.10.2 GRIP—Grid Image Processing for Biomedical Diagnosis (Latin America, Chile)

Patterns and structure identification in images for diagnostic purposes in physiology and pathology represent a key step in the development of strategies of early detection and timely treatment of human diseases.

Images come from a variety of devices (X-Ray, NMR, MRI, FMR, Confocal microscopy, Immunohistochemistry, Immunofluorescence, among other) and their handling require sophisticated modeling and numerical processing in order to enhance the features of significant content for diagnosis.

Research in Biomedical diagnosis encompass functional analysis of activity/structure of isolated tissues (cell cultures) as in Immunohistochemistry of breast cancer, fluorescence and isotope luminescence, etc. It allows identification of highly specific functional mechanisms of cell membranes, receptors and ion channel identification, intracellular waves of messengers (Ca2+, AMPc, IP3, RyR) and nuclear CREBS at DNA level. Research in these fields require processing of a significant amount of fluorescence images with different contrast levels, intensities, diffusion, etc., together with the development of advanced image processing algorithms and tools adapted to the needs of the morphologist.

The GRIP project [60] is an initiative for the design and implementation of Grid-based image processing tools for the advancement of research in pathology and physiology. Its applications have been deployed and executed on the EELA and EELA-2 infrastructure.

5 Conclusion

E-Infrastructures have demonstrated to be instrumental platforms for the development of e-Science and enablers of progress in many areas of the world affected by the "digital divide".

In this paper we have described the actual status of these e-Infrastructures, based on the gLite middleware, and presented a complete picture of the virtual research communities currently using them. Information on the scientific domains and on the applications supported were provided, for the first time, together with their geographic distribution.

The big number of virtual research communities, coming from several scientific domains, profiting of the e-Infrastructures can be considered as a proof of how this new scientific model has changed the way to do science in the world. The use of e-Infrastructures allows to remove many hindrances that delayed the scientific evolution until now. Indeed, the huge quantity of processors and storage space available on an e-Infrastructure make possible to solve very complex problems,

Moreover, the e-Infrastructure model had a strong social impact on the less developed countries. Providing e-infrastructure access to these countries mitigates the so-called "brain-drain" allowing brilliant minds in the area to stay in their regions and contribute significantly to cutting edge scientific activities reducing the "digital divide".

Presently, the size of the e-Infrastructures and the geographical distribution of the applications running on them show a predominance of Europe but, day after day, the relative weight of the other continents is steadily increasing.

All these observations can be considered as an important clues about the success of the e-Infrastructures for e-Science and we can easily imagine a future where this model will spread over all countries and scientific domains, definitely changing the way to do science.

References

1. Foster, I.: The Grid: a new infrastructure for 21st century science. Phys. Today **55**, 27–42 (2002)
2. Foster, I., Kesselman, C., Tueckle, S.: The anatomy of the Grid: enabling scalable virtual organizations. Lect. Notes Comput. Sci., Springer, **2150** (2001)
3. Cornwall, L.A., Jensen, J., Kelsey, D.P., Frohner, Á., Kouřil, D., Bonnassieux, F., Nicoud, S., Lőrentey, K., Hahkala, J., Silander, M., et al.: Authentication and authorization mechanisms for multi-domain Grid environments. J. Grid Computing **2**(4), 301–311 (2004)
4. Foster, I., Kesselman, C.: The Grid: blueprint for a new computing infrastructure. In: Kauffman, M. (ed.) Book (1997)
5. Baker, M., Buyya, R., Laforenza, D.: Grids and Grid technologies for wide-area distributed computing. Softw: Pract Exp **32**(15), 1437–1466 (2002)
6. Catlett, C.: Standards for Grid computing: global Grid forum. J. Grid Computing **1**(1), 3–7 (2003)
7. Laure, E., Hemmer, F., Aimar, A., Barroso, M., Buncic, P., Di Meglio, A., Guy, L., Kunszt, P., Beco, S., Pacini, F., Prelz, F., Sgaravatto, M., Edlund, A., Mulmo, O., Groep, D., Fisher, S.M., Livny, M.: Middleware for the next generation Grid infrastructure. In: Proceedings of Computing in High Energy Physics and Nuclear Physics 2004, p. 826. Interlaken, Switzerland, 27 Sep–1 Oct (2004)
8. Laure, E., Gr, C., Fisher, S., Frohner, A., Kunszt, P., Krenek, A., Mulmo, O., Pacini, F., Prelz, F., White, J., Barroso, M., Buncic, P., Byrom, R., Cornwall, L., Craig, M., Di Meglio, A., Djaoui, A., Giacomini, F., Hahkala, J., Hemmer, F., Hicks, S., Edlund, A., Maraschini, A., Middleton, R., Sgaravatto, M., Steenbakkers, M., Walk, J., Wilson, A.: Programming the Grid with gLite. In: Computational Methods in Science and Technology, vol. 12 (2006)
9. Pordes, R., Petravick, D., Kramer, B., Olson, D., Livny, M., Roy, A., Avery, P., Blackburn, K., Wenaus, T., Würthwein, F., Foster, I., Gardner, R., Wilde, M., Blatecky, A., McGee, J., Quick, R.: The open science Grid. Journal of Physics: Conference Series, vol. 78 (2007)
10. Lederer, H., Pringle, G.J., Girou, D., Hermanns, M.-A., Erbacci, G.: DEISA: extreme omputing in an advanced supercomputing environment. In: Bischof, C., B̈ucker, M., Gibbon, P., Joubert, G.R., Lippert, T., Mohr, B., Peters, F. (eds.) Parallel Computing: Architectures, Algorithms and Applications. John von Neumann Institute for Computing, Jülich, NIC Series **38**, 687–688 (2007)
11. Reed, D.A.: Grids, the TeraGrid and beyond. In Comput., IEEE Comput. Soc. **36**(1) (2003)
12. Matsuoka, S., Shinjo, S., Aoyagi, M., Sekiguchi, S., Usami, H., Miura, K.: Japanese computational Grid research project: NAREGI. Proc. IEEE **93**(3), 522–533 (2005)
13. Yang, G., Jin, H., Li, M., Xiao, N., Li, W., Wu, Z., Wu, Y., Tang, F.: Grid computing in China. J. Grid Computing **2**(2), 193–206 (2004)
14. GÉANT Network Web Site. Available at http://www.geant.net. Accessed Feb 2011
15. Gagliardi, F., Jones, B., Reale, M., Burke, S.: European datagrid project: experiences of deploying a large scale testbed for E-science applications. In: Performance Evaluation of Complex Systems: Techniques and Tools. Lect. Notes Comput. Sci. **2459**, 255–264 (2002)
16. Gagliardi, F., Jones, B., Grey, F., Bégin, M.-E., Heikkurinen, M.: Building an infrastructure for scientific Grid computing: status and goals of the EGEE

project. Phil. Trans. R. Soc. A **363**(1833), 1729–1742 (2005)

17. Andronico, G., Barbera, R., Koumantaros, K., Ruggieri, F., Tanlongo, F., Vella, K.: Grid infrastructures as catalysts for development on e-science: experiences in the Mediterranean. In: Bio-Algorithms & Med-Systems, 5 Grid in Science (2007)

18. Wang, Y., Scardaci, D., Yan, B., Huang, Y.: Interconnect EGEE and CNGRID e-Infrastructures through Interoperability between gLite and GOS Middlewares. IGIIW 2007—International Grid Interoperability and Interoperation Workshop 2007 in conjunction with e-Science 2007—3nd IEEE International Conference on e-Science and Grid Computing, Bangalore, India, 10–13 December (2007)

19. Paganoni, M.: The EUAsiaGrid project. In: Production Grids in Asia, vol. 1, pp. 41–45. Springer US (2010)

20. CLARA Web Site. Available at http://www.redclara.net/index.php?lang=en. Accessed Feb 2011

21. Marechal, B., Bello, P.H.R., Carvalho, D.: Building a Grid in Latin America: the EELA project e-infrastructure. In: Seventh IEEE International Symposium on Cluster Computing and the Grid, 2007. CCGRID 2007, pp. 835–839 (2007)

22. Barbera, R., Brasileiro, F., Bruno, R., Ciuffo, L., Scardaci, D.: Increasing e-infrastructure usability: the EELA-2 experience. 3pgcic 2010 International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, pp. 332–337 (2010)

23. EUMEDGRID-Support Web Site. Available at http://www.eumedgrid.eu. Accessed Feb 2011

24. EUMEDCONNECT Web Site. Available at http://www.eumedconnect.net. Accessed Feb 2011

25. EUMEDCONNECT2 Web Site. Available at http://www.eumedconnect2.net. Accessed Feb 2011

26. APAN Web Site. Available at http://www.apan.net. Accessed Feb 2011

27. TEIN3 Web Site. Available at http://www.tein3.net. Accessed Feb 2011

28. Venton, D.: WLCG improves India's access to LHC data. In: CERN Computer Newsletter No.2007-002 in Section 'Grid news' (2007)

29. Prahlada Rao, B.B., Ramakrishnan, S., Raja Gopalan, M.R., Subrata, C., Mangala, N., Sridharan, R.: e-Infrastructures in IT: a case study on Indian national Grid computing initiative—GARUDA. In: Computer Science—Research and Development, vol. 23, pp. 283–290. Springer Berlin (2009)

30. Masoni, A.: EU-IndiaGrid2 sustainable e-infrastructures across Europe and India. In: Lin, S.C. (ed.) Data Driven e-Science: Use Cases and Successful Applications of Distributed Computing Infrastructures (ISGC 2010) (2010)

31. Ubuntunet Alliance Web Site. Available at http://www.ubuntunet.net. Accessed Feb 2011

32. FEAST Web Site. Available at http://www.feast-project.eu. Accessed Feb 2011

33. IST-Africa Web Site. Available at http://www.ist-africa.org. Accessed Feb 2011

34. EuroAfrica-ICT Web Site. Available at http://www.euroafrica-ict.org. Accessed Feb 2011

35. eI-Africa Web Site. Available at http://www.ei-africa.eu. Accessed Feb 2011

36. SAGRID Web Site. Available at http://www.sagrid.ac.za. Accessed Feb 2011

37. Fargetta, M., Scardaci, D., Barbera, R.: GILDA status and recent activities in Grid training. Book chapter: Managed Grids and Cloud Systems in the Asia-Pacific Research Community, pp. 311–322. Springer US (2010)

38. Andreeva, J., Campana, S., Fanzago, F., Herrala, J.: High-energy physics on the Grid: the ATLAS and CMS experience. J. Grid Computing **6**(1), 3–13 (2008)

39. GASuC Web Site. Available at http://www.lpds.sztaki.hu/gasuc. Accessed Feb 2011

40. Heck, D.: The air shower simulation program CORSIKA. Contribution to Int. Workshop on Simulation and Analysis Methods for Large Neutrino Detectors, DESY Zeuthen (Germany), Report DESY-PROC-1999-01, 228 (1998). More information at https://na4rs.marie.hellasgrid.gr/na4-EGEE/index.php

41. De Vos, M., Gunst, A.W., Nijboer, R.: The LOFAR telescope: system architecture and signal processing. In: Proceedings of the IEEE, Special Issue Advances in Radio Telescopes, vol. 97, no. 8, pp. 1431–1437 (2009). More information at https://na4rs.marie.hellasgrid.gr/na4-EGEE/index.php

42. GAP Virtual Screening Service Web Site. Available at http://gap.Grid.sinica.edu.tw/index.php?option=com_frontpage&Itemid=1. More information at http://www.euasiagrid.org/index.php?option=com_content&task=view&id=116&Itemid=146. Accessed Feb 2011

43. Rohl, C.A., Strauss, C.E., Misura, K.M., Baker, D.: Protein structure prediction using Rosetta. Methods Enzymol. **383**, 66–93 (2004)

44. Ribeiro, A.S.T.: The GROMACS package. In: Proceedings of the First EELA-2 Conference Bogota (Colombia), CIEMAT (2009). More information at http://applications.eu-eela.eu/application_details.php?l=20&ID=54

45. Dominguez, C., Boelens, R., Bonvin, A.M.J.J.: HADDOCK: a protein–protein docking approach based on biochemical and/or biophysical information. J. Am. Chem. Soc. **125**, 1731–1737 (2003)

46. Cipolla, M., Valenti, C.: A new automatic algorithm for three-dimensional edge detection. In: Proceedings of the Symposium GRID Open Days at the University of Palermo (Italy), pp. 221–225 (2007). More information at https://na4rs.marie.hellasgrid.gr/na4-EGEE/index.php

47. Pelfer, G., Pelfer, P.G.: From WEB to GRID, a new perspective for archaeology. Proc. Nucl. Sci. Symp. Conf. **2**, 834–834 (2004)

48. Astra Web Site. Available at http://www.astraproject.org/. More information at http://applications.eumedgrid.eu/application_details.php?l=20&ID=2. Accessed Feb 2011

49. WRF4G Web Site. Available at http://www.meteo.unican.es/software/wrf4g/. More information at http://www.euasiagrid.org/index.php?option=com_content&task=view&id=116&Itemid=146. Accessed Feb 2011

50. Michalakes, J., Dudhia, J., Gill, D., Henderson, T., Klemp, J., Skamarock, W., Wang, W.: The weather

research and forecast model: software architecture and performance. In: Zwieflhofer, W., Mozdzynski, G. (eds.) Proceedings of the Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology, pp. 156–168. World Scientific (2005)

51. Rastogi, R., Srivastava, A., Majumder, S., Gholap, S.: Multiple realisation of real-coded genetic algorithm: a tool for 2D traveltime tomographic inversion. SEG Exp. Abst. **28**, 2491 (2009). doi:10.1190/1.3255362

52. Barbastefano, R., Carvalho, D., Rossi De Souza, L.: Industry@Grid Demo. Presented at First EELA-2 Review, Brussels (Belgium) (2009). More information at http://applications.eu-eela.eu/application_details.php?l=20&ID=14

53. Roccia, B., Preidikman, S., Massa, J.: De la Biología a los Insectos Robots: Desarrollo de un Código Computacional Interactivo para Estudiar la Cinemática de Alas Batientes. Mecánica Computacional **27**, 3041–3058 (2009). More information at http://applications.eu-eela.eu/application_details.php?l=20&ID=72

54. Meroun, O., Al Mers, A., El Bardouni, T.: Preliminary subchannel analysis code for application to TRIGA. Presented at EGSAP-1, the First EUMedGrid School for Application Porting (2007)

55. ATLAS Experiment Web Site. Available at http://www.cern.ch/atlas. Accessed Feb 2011

56. LHCb Experiment Web Site. Available at http://www.cern.ch/lhcb. Accessed Feb 2011

57. TOTEM Experiment Web Site. Available at http://totem.web.cern.ch/Totem. Accessed Feb 2011

58. Pierre Auger Web Site. Available at http://www.auger.org. Accessed Feb 2011

59. WISDOM Web Site. Available at http://wiki.healthgrid.org/Wisdom:web_site. Accessed Feb 2011

60. Lopez Fenner, J., Abarzua, J.P., Morales, D.V., Fuentes, A.L., Romero, F., Roa, J.C., Rivera, R.P., Salinas, L.: Development of a parallel algorithm for detection of cancerous cells in inmunohistochemical samples. Presented at 2nd EELA-2 Grid School, Queretaro (Mexico) (2009). More information at http://applications.eu-eela.eu/application_details.php?l=20&ID=60