e-Infrastructures Reflection Group White Paper

Version 1.6
22 August 2005

Work in Progress

Subject to Further Discussion and Improvement

0. Preface

1. Executive summary

1.1. Background

The main objective of the EU e-Infrastructure initiative is to support the creation of a policy framework to enable the shared use of distributed electronic resources across Europe and beyond – particularly for grid computing, storage and networking. The e-Infrastructures Reflection Group (e-IRG – www.e-irg.org) was founded to define and recommend best practices and policies for e-Infrastructures. It consists of official government delegates from all EU countries.

The e-IRG White Paper is a live document that captures best practices and experiences in the sharing of electronic resources in Europe. These form the basis for the policy recommendations that are endorsed by the e-IRG delegates. The White papers are produced every six months following the corresponding EU presidencies. Similarly, the e-IRG meetings are also usually hosted by the EU presidencies.

1.2. About the structure and the procedure of the White Paper

The structure of the document has been modified slightly compared to the previous versions of the White Paper. As mentioned several times in the e-IRG meetings and workshops, it is important to evaluate the policies and the necessary changes to them in the light of the current state-of-the-art technical implementations. This combination of technical state of the art and policy references facilitates the analysis of the necessary technical developments in light of the current and planned policy initiatives.

The new structure of the White Paper aims to emphasize this duality by explicitly separating the rapidly – and sometimes in revolutionary fashion – developing technical basis of the e-Infrastructure on the one hand and the more incrementally evolving policy-framework on the other.. Thus the main part of the paper will be focusing on policy issues, while technical background has been moved to the end labeled as technical appendices. This way the document will be easier to maintain and the future version will be able to refer to the policy part of the previous White Papers without having to deal with partially out-of-date technical content. Finally, this approach makes it possible to augment the technical background without changing the compact policy description in the beginning of the document.

As far as the procedure to produce the "Luxembourg" White Paper is concerned, the main innovation is that significant input has been received during the e-Infrastructure workshop[1] that was held on 13 May in Amsterdam, one month before the e-IRG meeting.

[1] http://www.e-irg.org/meetings/2005-NL

This one month timeframe, although short, has enabled the ripening of different opinions and lead to the current White Paper. Note that previous e-Infrastructures workshops were held jointly with the e-IRG meetings, a fact that did not contribute much to the enrichment of the White Paper. It is recommended, however, that in future a timeframe of at least two months should separate the two events, so as to leave enough time for adequate processing and expansion of ideas.

1.3. Summary

This document draft acts – like its predecessors – as a tool for the dialogue between technical experts and policy makers, aiming to aid in the development of the more inclusive and efficient e-Infrastructures. At least as important, however, is the role of this document as a catalyst so that the communities participating in the work of the e-IRG can engage a much wider group of audiences, which is needed to develop a common vision for the next 10 to 20 years.

The current White Paper studies a series of areas where policies need to be developed, including AAA (Authentication, Authorization and Accounting), Acceptable Usage Policies (AUPs), User Support, Optical Networking and related Grid Requirements, Middleware, Legal Issues, Advanced Computing Facilities, Storage and Data Services, as well as the role of Industry. In parallel, attempts to tackle issues like Generic vs. Disciplinary Grids, the Governance of the Grid Infrastructure and of the corresponding Middleware are being made. Note that the latter issues need further brainstorming and discussions among all relevant stakeholders.

Section 2 demonstrates the importance of the **e-IRG** during the past couple of years in order to shape the European Research Area, highlights some of the group's achievements and reveals the role that it has played in the development of the 7th European Research Framework Programme (FP7). This is reflected in particular in the creation of a roadmap for e-Infrastructures in Europe for the next 10-20 years, which has been drafted as a separate document and was presented in the Luxembourg e-IRG meeting. The roadmap is based on a short document called the "opportunities list", which was compiled from the replies to a structured questionnaire filled in by the e-IRG national delegates.

Section 3 looks at the issue of "**General purpose vs. Disciplinary Grids**", which was initially brought up in Den Haag. In this section, an effort is made to clarify the terminology of the Grid and its components. This terminology allows a "factorization" of the problem of building Grids for various communities and varying disciplines. "Grid paradigms", a rather confusing concept, are replaced by three concepts which can be analyzed from technical and economic points of views: "desired organizational schemes and requirements", "identification and availability of digital resources" and "capability, efficiency and cost of a particular Grid or Network Infrastructure given an organizational scheme". The proposed analysis needs to be elaborated further in order to provide more tangible results. However, it seems that the Den Haag recommendations remain valid, i.e.

that a “general purpose” Grid is technically and policy-wise difficult to establish; multi-disciplinary grids are a good step forward, since they contribute to the elimination of duplication of efforts.

Section 4 raises the issue of whether the Grid Infrastructure should follow the GEANT-NREN model in the future, i.e. a central co-ordinating body and one national body per country. The **European Grid Organization – National Grid Initiative scheme** are defined and briefly analysed, identifying some basic issues. It is felt though that further brainstorming and discussions among the different stakeholders is needed.

Section 5 analyses both the **Authentication & Authorization Infrastructure** (AAI) updates and the Accounting novelties since the Den Haag White Paper. As far AAI is concerned, it was felt that a superstructure integrating all the nation- or community-based AA federations is essential and therefore it needs to be explored how to make current practices and systems compatible. For **Accounting,** the situation is still technically immature; furthermore accounting requires some new real life experience to be gained to establish the feasibility of potential policies.

Section 6, for the first time in the White Paper series, deals with the **Legal Issues** in electronic infrastructures as summarized in the Amsterdam workshop. It is recommended that a dialogue between legal and technical experts should be kicked-off to prepare an inventory of legal issues that are currently being encountered or that are foreseen to become relevant. A dialogue could be initiated by means of a workshop, taking as a starting point the Luxembourg White Paper. If deemed appropriate, a more permanent group might be established.

Section 7 exemplifies how **Network Research** is imperative to meet next generation Grid requirements. Integrating evolving optical technologies with e-Science applications and Grid middleware is critical. Thus e-Science networking requirements should be studied carefully, new technologies and control-plane[2] solutions should be investigated and possibly integrated in the applications/middleware in order to automate the networking requests. As far as **International Global Connectivity** is concerned the major conclusion is that this is very much a multi-domain activity. Global connectivity will never mean a global network, but rather inter-domain cooperation and interoperability. This must be taken into account for all planning that strives towards global end-to-end connectivity for researchers.

Section 8 continues the Den Haag survey on **User Support Policies** in a more generic way identifying four main areas of support: education, simple access to a broad range of

---

[2] Defined as the "Infrastructure and distributed intelligence that controls the establishment and maintenance of connections in the network, including protocols and mechanisms to disseminate this information; and algorithms for engineering an optimal path between end points". The recent developments in this domain will allow the netwerk logic to interact with the Grid middlewarr, e.g. for optimization proces. (see: http://www.glif.is/working_groups/controlplane/controlplane.html)

information, application integration & support, and round the clock support for the users of grid data, computing and networking services. The main message in all four areas is that plans for support of the grid infrastructure should allow for the continuity of support in national structures beyond the end of major grid projects.

Section 9 considers Grid middleware as a fundamental component of e-Infrastructures across the European Research Area. The possibility of establishing a **Federated Middleware Institute** to ensure the development of a production quality Grid middleware leveraging European as well as national efforts should be investigated.

Section 10 reports on updates in the **Usage Policies** area. The new trend is to have a single common Acceptable Usage Policy (AUP), agreed to by all Grid users. This fairly abstract and liberal definition of acceptable use is further qualified in a VO-specific way. Thus, each VO prepares also its own acceptable use policy, which clearly expresses the aims of the VO, together with any necessary details of the applications and/or data involved. Grid infrastructure management and site/resource managers are then able to evaluate the merits of each VO before deciding whether or not to allow the VO to operate on the Grid, safe in the knowledge that all members of the VO will have agreed to the common AUP at the time they registered with the VO.

Section 11 summarizes national large scale **supercomputing** background and facilities identifying two different visions: infrastructure based on disciplinary needs (e.g. a system for climatology) or infrastructure based on computer architectures (64-bit vs. 32-bit addressing and arithmetic, low latency vs. high bandwidth, capability vs. capacity computing, etc.) The outcome of the Amsterdam e-IRG workshop in May 2005 clearly was in favour of a model based on computer architectures, rather than one based on disciplines, provided that the issue of user support was well addressed in a European science Grid environment.

Section 12 covers **storage and data services** and recommends the establishment of a distributed shared network of European data centres, maintaining digital research data and other digital materials over their entire life-cycle and making them available to current and future generations of users. It is also recommended to co-ordinate thedevelopment efforts of data management software across Europe and to stimulate P2P and Google-like technologies to be applied in research for data management purposes.

Section 13 pinpoints the role of the EC in supporting **Industry** in joining the Grid arena. European enterprises including SMEs and other commercial stakeholders are invited to come forward and identify both the expectations from the business community and the contributions they expect to be able to make in return, considering the long term goals of the European e-Infrastructures.

Section 14 provides a list of **policy roadmap** statements per section, gathering the most important future steps, action points and recommendations.

Sections 15-20 provide some background information on the technologies in use for some of the above sections in the form of **technical appendices**. This is part of the effort of separating policy information from technical background.

Finally, *sections 21-24* are the other **appendices** as they have been introduced in the Den Haag version, with the history of the white papers, a list of previous endorsement statements by the eIRG, the eIRG members list and a list of acronyms.

2. The role of the e-IRG in shaping the ICT European Research Area

2.1. Background

Two major initiatives have had a key impact on the European research and education policies: The special meeting of the European Council in Lisbon in March 2000 focused on employment, economic reform and social cohesion for a Europe of innovation and knowledge; the Bologna process promotes the creation of the European Higher Education Area. Europe is gearing up its efforts to become the world leader in terms of the quality of its education and training systems. Thus it is of utmost importance to understand the central roles of research, innovation and education as the vehicles to reach the vision of a flourishing European society with economic growth and welfare. The realisation of a modern information society in Europe should continuously be based on more tightly networked research and innovation, within the framework of a European research and innovation area.

The European Research Area (ERA) – complemented with education and other policies – aims at co-ordinating national research policies to share objectives, expertise and resources. With the ERA, Europe provides itself with the resources to exploit its potential to become the world's most competitive and dynamic economy.

In this context, the *raison d'être* of the current European Research Framework Programme (FP6) is to help to make the European Research Area a reality, already with a view to step up innovation in Europe – taking into account all the efforts made to this end at national, regional and European level.

2.2. The European e-Infrastructure and the role of the e-IRG

The Europeans have recognized that the building of an innovation infrastructure based on a new generation of ICT Research Infrastructures (e-Infrastructure) is a key factor that can determine Europe's path to innovation in the 21st century. This vision puts an emphasis on high-speed and high-QoS networks, to enable the "unlimited" transmission of information, and on Grid-based tools and services, to enable "unlimited" world-scale collaborations, sharing of knowledge, and effective use of resources. Today, research and scientific explorations are increasingly based on wide-spread scientific collaborations – following a rather revolutionary trend in creating, disseminating, and preserving scientific and engineering knowledge. Large-scale collaborations create the need to share and to process (often in parallel and in real time) data and other electronic resources (both hardware and software) that are distributed in hundreds or even thousands of centres and institutes in Europe and around the world.

In the beginning, technology seemed to play the pivotal role in the process of users adopting new ICT-infrastructures, in particular those that implement a model of shared use of distributed electronic resources and knowledge – across organizations, scientific disciplines, countries and other administrative and technological domains. However, quite early in the process it became clear that different administrative policies and social factors across the above domains pose significant barriers in the adoption of the new infrastructures. Therefore, policies and social aspects are of equal importance to technology.

The establishment in 2003 of the e-Infrastructures Reflection Group (e-IRG) by the EU Member States – in collaboration with the Commission – reflects the strong commitment of the European research authorities to address the above policy level challenge on a European scale and to move fast towards the adoption of the new generation of Grid-based ICT-infrastructures in Europe.

Since its establishment, the e-IRG found a continuous support for the development of its processes and content from EU (but also non-EU) states and, in particular, from the research authorities of the EU countries that held the EU-presidency since the beginning of 2003 (namely Greece, Italy, Ireland, the Netherlands and Luxembourg).

In its 2.5 years of its existence, the e-IRG has contributed decisively to the creation of subsequent awareness across research communities in Europe (and to an extent also around the world), has created five versions of the so-called White Paper[3] and has established itself, in short, as a reference group on policy creation in the field of ICT research infrastructures. Perhaps most of all, the e-IRG has attracted – through its work – the attention of the public research funding authorities in Europe and is stimulating today their increased support and emphasis in the deployment of new ICT-infrastructures and the alignment of their actions on the level of policy.

In addition, through the e-IRG activity the need to closely co-ordinate and align policies, operation, and administrative actions between all layers and components of the e-Infrastructure (networking, Grid-middleware, security etc.) became more apparent and all stakeholders got involved, with a view to ensure the effective creation, economic operation, and best service provision to the users of such an e-Infrastructure (one-stop-shop e-Infrastructure service).

Therefore, the e-IRG initiative today is playing a significant role in strengthening the foundations of the ERA and in assisting the EU to reach the eEurope[4] and eEurope+ objectives.

---

[3] The e-IRG White Paper is a live document that captures best practices and experiences in Europe on the above discussed policy-level aspects including relevant recommendations. The latest version of the White Paper can be found on the e-IRG website at http://www.e-irg.org/publ/2004-Den-Haag-eIRG-whitepaper.pdf .

[4] The eEurope 2005 Action Plan was launched at the Seville European Council in June 2002 and endorsed by the Council of Ministers in the eEurope Resolution of January 2003. It aims to develop modern public services and a dynamic environment for e-Business through widespread availability of broadband access at competitive prices and a secure information infrastructure.

The successful presence of the e-IRG in the policy field called for an extensive role of the group in the context of the development of the 7th European Research Framework Programme (FP7), in particular to create a roadmap for e-Infrastructure in Europe for the next 10-20 years. Following an invitation by the Commission (in the group's November 2004 meeting in Den Haag) to develop this roadmap, the e-IRG produced in March 2005 a brief document called "Opportunities List", i.e. a list of ICT research infrastructures, for which the funding in FP7 was seen as important and potentially highly beneficial for Europe. This "opportunities list" can be considered as a precursor to the e-Infrastructure roadmap, the first version of which is being worked in parallel and has been published as a separate document. (Draft version at http://www.e-irg.org/roadmap/roadmap.0.7.pdf)

2.3. e-IRG related initiatives

The preparation work for the 7th Framework Programme for 2007-2013 is rapidly shaping up to concrete actions. The goal is to look for a constructive approach where the joint wish and interest of a variety of groups are included into a common will. Besides the e-IRG, other groups are also contributing to shaping the science infrastructure in Europe, notably:

- ESFRI – European Strategy Forum on Research Infrastructures (the forum for Research Infrastructures in general).
- ENPG – the European Networking Policy Group. The ENPG was founded in 1995 as a forum for the national policy and funding authorities for research networking in Europe. In the ENPG, the ministries funding the National Research and Education Networks (NRENs) throughout Europe co-ordinate their efforts.
- TERENA – the Trans-European Research and Education Networking Association and its related policy bodies, such as the General Assembly. TERENA was formed in October 1994 to promote and participate in the development of a high-quality, international information and telecommunications infrastructure for the benefit of research and education. TERENA carries out technical path finding activities and provides a platform for discussion to encourage the further development of a high-quality networking infrastructure for the European research community. Worthwhile mentioning, in the context of the e-IRG, is the TERENA TACAR[5], a common repository for storing and validating the Certificate Authority root certificates and policies established by TERENA, which was formally recommended by the e-IRG in its Dublin meeting as a best practice in the field that could be deployed and followed by all relevant initiatives.
- NRENs – National Research and Education Networks (the research communities operating the research networks) and related policy bodies, namely Policy Committee (NREN PC) and Executive Committee, steering bodies of the GN2 project and GÉANT2 network. There is one NREN in each country that is tasked

[5] www.terena.nl/task-forces/tf-aace

in some case by the ministry for science and technology in other cases on behalf of the research community (Universities, research institutes and so on), which usually is/are paying for the network services to the NREN to provide the communications infrastructure for the totality of the research community in the country. Increasingly the NRENs are also taking over the responsibility for providing connectivity to schools (typically via the Ministries of Education). In the process, the NRENs are more and more regarded as part of the national infrastructure (like roads and electricity).

ESFRI has produced the document "Towards New Research Infrastructures for Europe: the ESFRI – List of Opportunities"[6], which includes concrete examples of new, large-scale Research Infrastructures in the areas of physical sciences and engineering, biological and medical sciences, and social sciences and humanities. These infrastructures will be needed by the scientific community in Europe in the coming decade.

It is obvious that the work of the above groups should be geared towards providing the e-Infrastructure solutions, which will enable Europe to exploit the opportunities that ESFRI and the e-IRG outlined in the context of their list of opportunities.

2.4. Future roles of the e-IRG

It is important to note that the e-IRG is expected to continue to make significant contributions in the context of both the policy-framework (for the shared use of electronic resources) and the roadmap. Regarding the roadmap activity, closer links are sought between the e-IRG and ESFRI.

An aspect related to the future work of the e-IRG is whether the focus of its policy-work will continue to be on public research infrastructures (as it has been so far) or if industrial and commercial concerns (in the same technical field) should also be considered. It is worthwhile mentioning here that more and more people suggest lately that the e-IRG should play a more (pro-)active role in helping industry to deploy similar policies in the field of sharing of electronic resources and data and that relevant actions towards industry should be taken. This last issue is expected to rank high on the agenda of discussions in the coming months about the future roles of the e-IRG.

[6] Available from ESFRI reports page: http://www.cordis.lu/era/esfri_working_groups.htm; direct link to the document: ftp://ftp.cordis.lu/pub/era/docs/esfri_list_opportunities_290305.pdf

3. Generic vs. disciplinary grids

3.1. Introduction

The previous e-IRG white paper (the "Den Haag paper") points out that Grid technology is still in its infancy. This is also reflected in the fact that the terminology used is often rather imprecise or that the same term is used in different contexts to mean different things. It is important to start refining the terminology and to keep an agreed vocabulary, whilst allowing for the evolution that will undoubtedly take place. The Den Haag paper presented also a taxonomy of Grid organizational schemes and called for them to be further analysed in order to identify re-useable pieces and to avoid duplication of effort. In this paper, we try to take a step forward from this initial analysis.

3.2. Terminology

A first immediate problem is that the word "infrastructure" is used to mean many different things. We start from the definition in the online edition of the Oxford Dictionary of the English language:

> ***Infrastructure***: (noun); the basic physical and organizational structures (e.g. buildings, roads, power supplies) needed for the operation of a society or enterprise.

Note that in this definition, the people occupying the buildings, the cars using the roads and the appliances using the power from the supplies are not part of the infrastructure. We propose to adapt this definition to the needs of the Information Age as follows:

> ***Grid Infrastructure***: The basic (persistent) digital and organizational structures (e.g. Certification Authorities, trust agreements, Virtual Organization membership, Catalog and Information services) needed to support operation of distributed digital resources for use by non-geographically defined teams.

> ***Network Infrastructure***: The basic (persistent) digital and organizational structures (e.g. Domain Name Services, routing agreements, Internet name and numbering schemes) needed to support operation of multiple networks to provide seamless connectivity to end users.

We propose to use the term "digital resources" to describe the computers, storage systems, etc. that are attached to the infrastructures. Note that the resources and the infrastructures may be owned by different entities, which themselves may be different from the end users. This approach now allows us to define:

> ***Grids***: The set of all instances of "a Grid"

> ***A Grid***: The set of (dynamically reconfigurable) digital resources that a team is using at any given time to support their e-Activity across administrative domains.

The following are corollaries of these definitions:

> Grids are "disciplinary", or rather a Grid is always devoted to a particular purpose.
>
> Grid and Network Infrastructures may be totally general or may be designed with particular purposes in mind.

3.3. Discussion

It immediately follows from the refined terminology that in an ideal world it should be possible for a given team to organize their Grid in whatever way is most suitable to its needs. It also follows that in an ideal world a single general-purpose Grid Infrastructure and a single general-purpose Network Infrastructure should be able to support all imaginable Grids.

Actual practical implementations will, of course, place certain limitations. Nevertheless, it seems reasonable to expect many Grids corresponding to many teams, yet only a few Grid and Network Infrastructures with different degrees of generality or specialization.

The Den Haag paper identifies a number of "Grid paradigms" (flat, collaborative, power plant, hierarchical, etc.). Using the refined terminology, these paradigms correspond first and foremost to instances of organizational schemes which are best suited for particular teams. For any particular paradigm and its corresponding detailed requirements (levels of security and reliability, data rates, response times, etc.) it is now possible to ask concrete key questions, which - if they are answered positively – will enable the team to successfully construct and use "their Grid" under this organizational scheme:

- Can the correct *digital resources* be identified and made available with the correct capacity vs. time profiles in order to fulfil the requirements of this particular team?
- Are the *Grid and Network Infrastructures*, by which the above mentioned digital resources are made to behave as a Grid, capable of supporting the desired organizational scheme and with what efficiency and cost?

In this way, a "factorization" of the problem is achieved while retaining the very desirable user-centric approach for the services to be delivered. In a situation where Grids and Grid Infrastructures are expected to evolve, this factorization allows for migration of functionality between the components. For example, a general-purpose Grid infrastructure may today require a very sophisticated Grid application in order to support a desired organizational scheme. But in some future version, functionality may be added to the infrastructure, which may allow simplification of a given class of Grid applications corresponding to similar organizational schemes or requirements.

3.4. Summary

A refined terminology is proposed to define Grids, Grid and Network Infrastructures, and digital resources.

This terminology allows a factorization of the problem of building Grids for the various communities with varying disciplines.

Clear ownership of the components can be defined, without forcing artificial separations.

“Grid paradigms”, a rather confusing concept, is replaced by three concepts which can be analysed from technical and economic points of views: “desired organizational schemes and requirements”, “identification and availability of digital resources” and “capability, efficiency and cost of a particular Grid or Network Infrastructure given an organizational scheme”.

3.5. Conclusions

The refined terminology and decomposition proposed above could lead the way towards clearer answers towards the relative effort that should be devoted in future to general purpose or disciplinary Grids. Unfortunately this is not a trivial task. In the meantime and based on current practices and achievements, we believe that the Den Haag approach is still valid. The main outcome of the Den Haag and Amsterdam discussions was that the available technology and especially the policy is not yet ready for realization of a single general-purpose-Grid vision and there is currently an emphasis on disciplinary and in some cases multi-disciplinary Grids with user communities of manageable size and disciplinary width. Many aspects of the underlying technology (middleware and some other elements of the policy architecture) are common between disciplinary and multi-disciplinary Grids, and it is important not to duplicate efforts. At the same time it is important to encourage bottom-up development that has proven to be very successful in several cases. Standardization efforts are thus needed but should not be started prematurely.

4. Towards National Grid Initiatives

4.1. Introduction

This section examines whether the ultimate vision of organizational structure supporting the Grid-enabled e-Infrastructures should follow the Research Networking paradigm. The Research Networking model is based on central co-ordinating body, being DANTE (Delivery of Advanced Network Technology to Europe - www.dante.net), a non-for-profit limited liability company co-ordinating the pan-European Research Networking infrastructure GEANT- www.geant2.net acting as a "hub" integrating the National Research and Education Networks (NREN) of all countries. An NREN is a "public body" responsible for developing the national research networking backbone in its country, and interconnecting it to the pan-European network GEANT. The NREN is in most cases established by a related ministry (Research, Education, Development, etc.) and offers its services to all Research and Academic institutes in the country, and in some cases even to the schools. The service being offered is infrastructure-oriented and application-neutral. Applied to the Grid domain, there would be an European Grid Organization (called hereafter EGO), co-ordinating the pan-European Grid infrastructure and one official National Grid Initiative (NGI) per country providing the national Grid infrastructure and co-ordinating all the national Research Grid efforts. Following the eIRG discussions in Mondorf, Luxembourg, it seems that NGI is a more mature idea that should be taken forward, while the EGO idea is still not mature enough. This section examines some of the pros of cons for this vision.

4.2. Background

The current landscape in the Grid Research Infrastructure projects area is quite clear. Following the successful paradigms of the 5th Framework Programme projects such as Datagrid www.eu-datagrid.org and Crossgrid www.crossgrid.org, the EGEE project (Enabling Grids for EsciencE www.eu-egee.org) emerged as the first project running a large-scale production-quality Grid infrastructure offering mainly computing and data resources all over Europe and beyond, enabling scientists to cooperate around the clock. Note that EGEE, through the LCG2 and soon the Glite middleware, integrates mainly PC-based clusters running at its vast majority on Linux-based operating systems. DEISA is a complementary project currently deploying and operating a persistent, production quality, distributed supercomputing environment with continental scope integrating the biggest supercomputers in Europe. SEEGRID (South Eastern European GRid-enabled eInfrastructure Development) through training, dissemination and other support actions, aims at extending EGEE to South East Europe, easing the digital divide in Europe. The successful SEEGRID model has been adopted by other regional areas such as Baltics, Mediterranean, Latin America and China, submitting proposals to cover these areas. EGEE plans to co-ordinate the above extensions aiming at gradually integrating them into its infrastructure.

Concerning the internal organization of EGEE, there is no strict rule on who participates from each country. In some cases there is only one leading contractor from each country gathering behind all other institutes (with new FP6 mechanisms such as the so-called *"Third parties"* and *"Joint Research Unit"* structures), while there are cases with multiple institutions per country. This has resulted to a consortium of 70 partners, which could have been even bigger without the current Third parties. The dense and thorny consortium has led to the organization of geographical and other-type-of *"federations"*, grouping partners together in order to form domains with common objectives and workplans. An example of other-type of federation is the networking one, gathering the DANTE and a group of NRENs.

Note that the EGO-NGI model is not meant to reduce the number of partners of the leading Grid infrastructure project, rather than to provide a sustainable (if not permanent) scheme to offer Grid services to all application communities in a neutral and professional way similar to the GEANT-NREN model has been doing for many years.

4.3. The European Grid Organization concept

An **European Grid Organization** *(EGO) aims at providing a long-term framework for maintaining and evolving a Grid infrastructure for science.* Europe is currently in a lead position in piloting scientific Grid infrastructures. The objective of the EGO is to ensure that Europe capitalizes fully on these efforts by mainly operating production Grid infrastructures for a wide range of scientific disciplines, in close collaboration with national Grid programmes and national and European research networks;

The EGO should be an independent entity working in the interest of all sciences. The EGO should have the appropriate support in order to provide the stability and continuity of the Grid infrastructure necessary from the user perspective. The EGO could potentially be partly funded by the communities for the services it provides or by the National Programmes.

Note that the EGO is more oriented towards the running of the infrastructure and its day-to-day operations including user and applications support, rather than to develop and maintain a reliable and robust middleware as this will be expressed in section 9, describing the idea of a *Federated Middleware Institute*. It seems that these ideas seem with a first look complementary rather than competing, but obviously a thorough analysis is recommended.

Finally the EGO concept is also relevant to the previous dilemma on whether we should envisage a general-purpose versus disciplinary Grids. Obviously the EGO is more compatible with the general-purpose Grid infrastructure, but ways of achieving disciplinary grids over a common infrastructure is also viable. This conclusion is also drawn from the recent development in the research networking arena, where parallel to the general-purpose GEANT network end-to-end optical connections are being planned, in order to offer specific disciplinary requirements.

4.4. National Grid Initiatives

A **National Grid Initiative** (NGI) is effectively a governance model to guide Grid deployment and operation at country-level and can be defined as "*concertated efforts taken at National level in order to deploy, operate, and expand grid infrastructures in a coherent and co-ordinated way. This usually involves the physical (network) and policy-based (e.g. AAA) inter-connection and inter-operation of multiple research and academic organizations (resource centers) under an umbrella of a national programme aiming to integrate the available resources in order to establish an e-Infrastructure for the benefit of the society at large.*[7] "

The acceptance and use of such a governance model was discussed in the SEE-GRID Policy Workshop, where it was deemed that the NGI model, if implemented appropriately per country, can help achieve the five key recommendations[8] that are important for deploying Grids in "greenfield" (developing) regions as well as aid the governance of grids expansion and operations also in established regions.

The main advantage of the NGI is that it assigns a formal (i.e. "state-sealed") stature to the group of the various research and academic institutes and organizations that collectively represent the web of grid resources at national level. Furthermore, it identifies a representative organization that co-ordinates the NGI activities and acts on behalf of and for the benefit of all NGI members, expressing in a single voice and in communication at all levels the interests and concerns of the NGI group. By doing so, there is increased effectiveness in actions involving political, technical, financial, or other issues of concern to the NGI, such as the examples mentioned below:

---

[7] "SEE-GRID Policy Workshop: A roadmap for establishing National Grid Initiatives", N. Vogiatzis et al., Den Haag, Netherlands, 24 Nov 2004, http://www.see-grid.org/Policy.pdf

[8] The five key recommendations of the SEE-GRID Policy Workshop were:

1. There can be no (product-quality, high-performing) grid without a (product-quality, high-capacity, reliable) network; GEANT extensions are critical in order for EU to support expansion of eInfrastructures beyond its borders.
2. Grid-expansion efforts in "greenfield" regions can benefit equally by studying "best-practices" as well as "bad-practices" followed in more advanced, eInfrastructure-established regions and countries.
3. There is no unique "success technical reference model": alternative technical roadmaps must be made available to fit to the needs and available infrastructures of each region.
4. There is also no "silver bullet" to guarantee long-term growth of the deployed eInfrastructures; sustainability can only be achieved by persistently undertaking several complementary and well-co-ordinated actions.
5. Establishing a "web of trust" that brings the right mixture of human capital together must be a key priority; the sociological element in the deployment of eInfrastructures can be as challenging as the technological one.

- NGIs can provide in a formal way the necessary fostering Political environment and support, and help establish the country grid governance model before reaching out at regional and inter-national level participation in grid efforts.
- NGIs can increase the impact of grid dissemination activities (especially training) by co-ordinating them and addressing the needs of the scientific communities and users at National level.
- Having an "early champion" to co-ordinate the whole set of activities, such as an NGI, can help integrate and maximize the usage of available resources at country level, while keeping in mind that things have to be organized in a way that is optimal for the specific area and not blindly copy-paste success stories from other areas. Especially for regions were existing user groups are relatively small, an NGI can help pull them together and work on the long-term outlook of integration.
- NGIs provide a formal governance model where all Grid user communities can be in the steering committee of the deployment effort, thus increasing the probability that the end-result would indeed serve the needs of the communities in question, while aiming set-up national multi-disciplinary projects involving multi-domain scientists.
- NGIs can help address in a unified and coherent way delicate policy issues in the operation of the infrastructure, most notably the ratio between those resource groups/centers who provide substantial funding and those who get access to the resources

Essentially, an NGI helps co-ordinate the country-wide collaboration of involved members by preventing having everyone trying to put their "stamp" on things, which means that in order to be effective, an NGI should:

- guarantee "apolitical" nature of the core infrastructure needed in order to operate and provide grid services at a national level
- facilitate collaborations and not impede other actions, yet monitor that other efforts are not doing everything from scratch
- ensure that "buy-in" in the grid deployment is achieved from all involved parties
- include key personnel, e.g.:
    - people who are dedicated to make the expansion and deployment of the infrastructure a national success story
    - stakeholders that share an inclusive vision that does not prevents any relying/interested party from being involved
    - people that can help the society-at-large clearly understand the benefits of "trail blazers" / pioneers, thus investing in long-term results rather than short-term gains

On the other side, even though with an NGI it is possible to have a very structured organized model at national level, this may prevent the materialization of a grid-based economy where the market is open and everyone is allowed to contribute resources to the grid. The idea of grid economy is deemed quite promising by its advocates Nonetheless, especially for “greenfield” regions, it is important first to achieve a critical mass -from both quantitative and qualitative point of view- in order to be able to switch to economical models at a later stage.

4.5. Discussion and recommendations

There is a delicate balance between initial fast progress with strong, centrally driven uniform approach and long-term sustainability that requires a buy-in from as large community of users as possible. Due to fragmented funding situation even in national scale, it is likely that also in the so called “Greenfield” regions there are ongoing Grid initiatives that represent a sizeable investment. Even conservative estimates see the aggregate national investment in the EU member states exceeding the EC funding for the Grid research. Thus there is a challenge for gaining the acceptance and legitimacy of each individual NGI which influences also how EGO is seen by the stakeholders that see themselves being bypassed from the initial decision making. For optimal balance between quick start-up and long-term efficiency, it is thus recommended that:

- Each e-IRG member state would evaluate the feasibility of establishing a NGI and planning for an open invitation for all the national projects to join in.
- A European policy for determining the common approach to the co-ordination functions of the EGO and NGIs should be established. The emphasis should be in maintaining local control of the resources, protecting the investments of the existing and potentially unknown projects (both in terms of technology and the social capital that is surrounding a specific Grid implementation). The emphasis should be in standardizing interfaces and protocols, not implementations.

In any case the issue needs to further be analysed in the next White papers and discussed in future e-IRG meetings, while future Research Infrastructure projects like the second phase of EGEE and DEISA could provide support in the documentation and analysis of the feasibility and effectiveness of the EGO-NGI model. In this effort, cooperation with the NREN community should be sought, as the experience of the NREN community could be exploited analysing the commonalities and differences with the GEANT-NREN model.

4.5.1. Recommendation

The eIRG would like to promote the idea of the National Grid Initiatives i.e. a governance model to guide Grid infrastructure deployment and operation at country-level in an application-neutral way. The eIRG believes that the adoption of the NGI idea across Europe would be a requirement for the evolution to the next phase of the eInfrastructures that is expected to be implemented in FP7, and encourages all countries to work on this.

5. Authentication, Authorization, Accounting policies

5.1. Authentication and Authorization Infrastructure

The situation with respect to Authorization was thoroughly analyzed in the Den Haag version of the White Paper and the following recommendation was made:

"The eIRG encourages work towards a common federation for academia and research institutes that ensures mutual recognition of the strength and validity of their authorization assertions."

The following paragraphs give an update on the steps taken in this direction.

TF-EMC2 (the TERENA Task Force on middleware[9]) has recently started an effort on harmonising schemas in order to enable inter-institutional data exchange. The aim of the group is to produce a set of schemas that make identity and attribute assertions, which are used to authorize users, employ well-understood and agreed syntax and semantics. A common understanding of attributes will definitely help in filling the gaps among different federations and systems. This is especially true, if we take into account that most are based on essentially compatible technical foundations. This effort is called SCHAC[10] and has initially been concentrated on individual data but the group is starting to consider the inclusion of access entitlements and attributes supporting VO management.

The TF-EMC2 is also working on AA-RR[11], a tool oriented to describe and exercise AAI components. AA-RR is intended primarily as an interoperability assessment tool, although some other applications are being explored, such as diagnostic facilities in a multi-system AA environment or inter-federation connection,. A modified version of AA-RR is planned to be used as a basic building block of certain GEANT2 AAI components.

The GEANT2 AAI group is finishing its architecture specification. The architecture proposal is aimed to cover use cases like network resource allocation, web application access, web service use and Grid AA. GEANT2 intends to offer full AA services in the cases of either:

---

[9] http://www.terena.nl/tech/task-forces/tf-emc2/

[10] http://www.terena.nl/tech/task-forces/tf-emc2/schac.html

[11] http://www.rediris.es/app/aarr/

- The GEANT2 AAI is acting as a connector of pre-existing infrastructures either at a national or group-wide (national or international) level ("established infrastructures"). In this case, the GEANT2 AAI will provide the superstructure for federating the established infrastructures, and the GEANT2 AAI interfaces will be used by their (infrastructures') components through appropriate adaptors. These adaptors may be located at any point in the infrastructure, so the GEANT2 AAI can be used either to route and translate all interactions or to establish the trust links that enable the infrastructure components to directly interact.

- The GEANT2 AAI is acting as a direct provider of authentication and authorisation information and decisions to the services requesting them, according to the general service framework described in the Den Haag version of this paper. In this case, the GEANT2 AAI will expose some of its interfaces to the user services.

Since AA interactions always imply at least two actors, namely the e-Infrastructure resource being accessed and the acknowledged source of identity being used, the GEANT2 AAI architecture will support any combination of the above cases. It is worth noting that the GEANT2 AAI team has committed to building the interfaces fully *Shibboleth*-compatible with respect to attribute queries and responses.

The *GridShib* activity has gained important momentum, and there are several integration-experiments reported so far. The GridShib model coincides with the one sought by the GEANT2 AAI for Grid integration, and can be simply described by the following steps:

1. The proxy certificate that the user employs in accessing a certain grid resource (generated by *VOMS*, for example) contains a reference to an Attribute Authority (a link to what Shibboleth calls an Identity Provider and what GEANT2 AAI calls a Connector).

2. The grid resource queries the Attribute Authority using this reference and the trust fabric and operations offered by the AAI. The Attribute Authority can be built on top of campus directories, VO repositories or whatever other data source the VO is willing to be based on, as long as it provides an interface to the attribute exchange operations of GEANT2 AAI/Shibboleth.

3. The grid resource may even use the GEANT2 AAI Authorisation services (which could use, of course, specific Grid Authorisation engines) to help it to make the final decision.

The EUGridPMA has had a noticeable impact in forming of the International Grid Trust Federation (IGTF), consisting of European Grid PMA, the Asia Pacific Grid PMA and the Americas Grid PMA. The structure and governance of the IGTF is at the moment being in late stages of negotiations, based on the draft charter document[12]

5.1.1. Recommendation

Support the establishment of frameworks able to integrate all the (nation- or community-based) AA federations, in the spirit of the achievements of the EUGridPMA with respect to PKI policies and practices, and promote the necessary steps to compatibilize current authorisation practices and systems with these frameworks. The group acknowledges the steps in this direction taken by the GEANT2 AAI and the Cotswolds Group initiative.

[12] Latest version of the draft charter has been produced on the 14th of July2005, available from http://www.eugridpma.org/igtf/IGTF-Federation-20050714-0-6.pdf.

5.2. Accounting

5.2.1. Introduction

The e-IRG meeting in Den Haag made the first concrete proposals for accounting policy development. Nevertheless the statement made at the Rome meeting that accounting for the grid is technically immature is still mainly true.

The Den Haag White Paper gave a good background definition of frameworks, economic models, and the state of the art. Since then a few new projects have started (or come to light) but only a fraction of the possible phase space has been investigated. Unlike some other areas covered in this White Paper, accounting requires some new real life experience to be gained to establish the feasibility of potential policies. It is possible some of this can be gained from other fields but this has not yet been done.

5.2.2. Current status and standardisation

Based on the analysis that has been done in the previous White papers it is obvious that accounting is still immature. A small number of projects have made good progress in implementing usage recording systems (metering). This work has fed back in the GGF standards procedure and standards are appearing. There has been little work however on economic models, banking, charging, pricing, therefore little progress on future polices.

Most **standardisation**, work is carried out in the Global Grid Forum. A series of the relevant GGF working groups is listed below:

Usage Record-WG

The GGF UR-WG usage record has now been adopted by a number of grid projects (EGEE, TeraGrid, NGS, OSG, SweGrid, NextGrid, LCG, NorduGrid, GridIT) and appears to be approaching a de facto standard. It will be interesting to compare their implementations for interoperability, comparison of extensions, subsets etc.

GESA-WG

The GGF Grid Economic Services Architecture GESA-WG aims to define the protocols and service interfaces needed to extensibility support a variety of economic models for the charging of Grid Services in the OGSA. It has been in hibernation for some time and no progress has been made since November.

RUS-WG

The revised spec for the Resource Usage Service is on track for submission to the GGF editor in the near future.

The MCS project implemented a web-services RUS and have been testing it on the UK National Grid Service. This experience was fed back into the specification and has updated it from OGSI to WS-I standards.

GRAAP-WG

The relevant GGF work on negotiation is taking place in the Grid Resource Allocation Agreement Policy Working Group.

5.2.3. Challenges

It is clear that work needs to continue to develop mature accounting schemes and gather experience on the grid market as a whole. Further work and studies need to be carried out like:

- gathering best practices in accounting from existing consortia and computing centres and transferring them to solutions based on open standards;
- detailed exploration of accounting use cases;
- Grid market pilot studies to identify bottlenecks and practical challenges in the economic models;
- pilot studies for the utilisation of banking systems with Grid resource billing;
- studies on legislation issues concerning resource trading;
- development of common vocabulary for accounting terminology;
- intensifying international collaboration in defining common accounting standards and interoperability between existing and future solutions.

In addition, it is a challenge to build an accounting system that can respect most - or possibly all - of the legal issues regarding privacy. We can learn from mobile phone and banking sector about legal issues but also about technical issues.

5.2.4. Recommendations

During the UK chairmanship the following points should be discussed:-

1. The Legal aspects of accounting such as:
    - Access to accounting data allowed by data privacy laws
    - Arbitration in the event of disagreements on charges between user and resource providers
2. The interaction of accounting with banking systems

How to manage the trust relationships between user, resource provider, metering systems, and financial brokers.

6. Legal issues in e-Infrastructures

6.1. Introduction

Ever since the 1980's, legal scholars and practitioners have been paying significant attention to the (potential) legal aspects of information technology. Starting with issues such as privacy and the legal protection of computer software, the debate and the statutory state of the art has grown to include as well topics such as electronic commerce, domain names, telecommunications, identity, security, liability, computer crime and legal protection of databases. So far, e-Infrastructures have not been touched upon as a separate subject in the legal debate. This will have to change, given the growing size and societal importance of e-Infrastructures.

This contribution reflects the discussion that took place in a legal session during the e-Infrastructures workshop of the e-Infrastructure Reflection Group as held in Amsterdam on May 13, 2005. This text reflects the outcome of the discussion between mainly technicians and policy makers. In describing the results of the discussion, next to an overview of the legal issues that are currently being perceived, also attention will be given to the way the legal discussion relating to e-Infrastructure could be brought forward.

6.2. Legal Issues currently perceived

e-Infrastructures as currently known offer the possibility to exchange and process electronic data with unprecedented speed and capacity.

Basically this seems to constitute a mixture of on one hand well known functionalities (electronic data processing, electronic communication), structures (electronic network) and parties involved (users, suppliers, intermediaries, governing bodies), and on the other hand new aspects such as sharing resources (scale, consequences, preconditions) and the level of communication and processing capabilities.

If we can take this as a starting point, an initial legal analysis of today's e-Infrastructures will lead to:

> (a) The identification of legal issues that have been identified previously in relation to the Internet, mobile communications etc;
>
> (b) The identification of legal issues that, although not entirely new, have a greater significance in an e-Infrastructure environment than in a more 'traditional' computer network environment;
>
> (c) New issues not encountered before in studying data processing and communication facilities from a legal perspective.

During the Amsterdam workshop session the following legal issues were perceived to be particularly relevant in relation to current e-Infrastructures:

- *Intellectual Property Rights*: Copyright and patent protection of specifically software might impede the development and adjustment of tools for operating and governing e-Infrastructures;
- *Relations*: e-Infrastructures are enabled by multi party cooperation. This requires adequate attention for the structure, legal form and the contents (roles and responsibilities) of the various relationships that can be distinguished; *Governance*: This includes issues such as responsibility, control, reporting and transparency;
  Rules for sharing resources, including Acceptable Use Policies; *Enforcement*: Rules have to be implemented as well enforced in order to be effective between the parties involved;
- *Security*: Which measures should and can be taken to ensure the integrity of the infrastructure and the data processing; who is responsible?
- *Authentication*: the authenticity of instructions, statements, reports etc. is essential for safeguarding amongst other thing the integrity of e-Infrastructures; *Accounting*: facilitating the economic aspects of the use of e-Infrastructures is key for its further development;
- *Competition law issues*: the (potential) economic significance of e-Infrastructure requires careful consideration of competition law issues that might be arise in the course of further development of e-Infrastructures. The capacity as offered by means of e-Infrastructures could lead to a distortion of the market of communication and processing capacity. This will especially be relevant when e-Infrastructure facilities are made accessible for organizations with (as well) commercial interests. State aid could be a relevant issue here. Also in building e-Infrastructures, the rules for public procurement can be relevant;
- *Responsibility and liability*: virtual organizations play an important role in creating, operating and controlling e-Infrastructures. This results in interesting questions into the extent to which they could be held responsible or liable for the way the e-Infrastructures is being (mis)used;
- *Criminal law issues*: processing of data in geographically distributed networks spread across various jurisdictions automatically leads to questions regarding criminal law responsibilities for the processing of data (e.g. applicable standards, which person or entity is responsible, jurisdiction); *Data protection*: e-Infrastructures enable both the export of data, as well as extremely powerful data analysis; both might endanger the privacy of individuals; *Applicable law and jurisdiction*: e-Infrastructures are by nature cross border and thereby multi jurisdiction.

The participants in the legal session identified the following issues as being currently of specific importance: *intellectual property*, *data protection*, *governance (including use policy's and the role of governments)* and, *the validation of data*.

We should of course keep in mind that the concept of e-Infrastructures is still developing which may lead to (significant) new insights.

6.3. Conclusions and the way forward

In which way should the legal discussion with regard to e-Infrastructures be taken forward? Discussion of the legal issues relating to e-Infrastructures could be problem driven on the basis of practical issues currently envisaged by the scientific community, which can impede the use of today's e-Infrastructures. However, next to this there should be a more fundamental approach aimed at analyzing the more elementary issues involved in sharing computing resources and expansion of the scope of application of this phenomenon beyond the scientific community, both on the demand side (users) and supply side (those making resources available).

During the Amsterdam workshop it was concluded that a mixture of both approaches seems most adequate – for now. A first step would be to make a description of the baseline, i.e. an inventory of those issues that are currently being encountered in practice or could be perceived to be relevant.

A next step would be to evaluate these issues on the basis of the current knowledge of the legal aspects of the on-line world. Such analysis would help in identifying the issues that could be held to be specific for e-Infrastructures. On the basis of such an overview it could be determined which approach should be followed to design solutions. These could range from (European) regulatory initiatives to standard setting, developing certification schemes, codes of conduct, model contracts and other forms of self regulation. In all cases, the issues should be addressed on the European level, and in some cases may be even on a global level (treaties, standardization initiatives, declarations etc.)

In further developing e-Infrastructures, legal issues should be taken into account in a structural way and as a part of a multidisciplinary approach. Incorporation of legal aspects in the debate will be beneficiary to the further development of e-Infrastructures. In that process, the first step will be to bridge the (knowledge) gap between those involved in the development and current application of e-Infrastructures and the community of legal scholars in the field of information technology and law.

6.4. Recommendations

Legal issues in electronic infrastructures are of vital importance and specific steps should be taken in order to make progress. It is advised that a dialogue between legal and technical experts should be kicked-off preparing an inventory of legal issues that are currently being encountered or foreseen to be relevant taking as a starting point the Luxembourg White Paper. A dialogue could be initiated by means of a workshop, and on the basis thereof when deemed appropriate a more permanent group might be established. A proper analysis and evaluation of the above inventory list would have to follow, on one hand providing priorities and classifying the issues and on the other hand paving the way for the integration of legal issues in the next generation of e-Infrastructures.

7. Network developments & grid requirements

7.1. Introduction

This section demonstrates the potential key role of advanced optical networking for future scientific discovery and collaboration. Dedicated optical connections that run in parallel with the traditional Research Internet would greatly facilitate scientific collaboration and will bring researchers much closer than before. However there are still many technical and policy issues to be resolved in order to enable this end-to-end connectivity. The use of dark fibre, which in some cases spans multiple nations, poses some resource sharing and access issues, as well as security concerns. In addition, there are several non-compliant or always interoperable technical ways of achieving the end-to-end mission, and policy decisions have to be taken. Scheduling of the connections is another non-trivial issue. Optical control plane advances could greatly influence towards one or another direction; while integration of a user-driven connectivity request into Grid applications and corresponding middleware will simplify the whole process. All the above issues sketch out the current research networking landscape.

In the following paragraphs we will provide the corresponding GÉANT/NREN-PC views as submitted to the eIRG in the framework of the eIRG European eInfrastructure Roadmap, and then briefly analyse the above-mentioned factors. Emphasis will be given to the International nature of connectivity and the issues that have to be addressed.

7.2. The GÉANT *roadmap*[13]

A major policy statement of the eIRG meeting in The Hague, regarding networking was the need to deploy a next generation optical pan-European network platform to support the needs of the Research & Academic Community. Specifically, this e-Infrastructure should integrate advanced IP based services with lower layer manageable "lambda" and/or Ethernet switched Gigabit provisioning for the support of eScience initiatives (e.g. Grids, collaborative research etc.)

This vision reinforces and coincides with NREN and GÉANT2/NREN-PC policy decisions, leading to the pan-European optical infrastructure currently under procurement. GÉANT2will be based on a wide Dark Fibre (DF) footprint able to provide switched "lambda" services, complemented with leased wavelength- and SDH-based provisions for regions where DF solutions are currently not available, or are prohibited from a techno-economic view point. As a result, with the roll-out of GEANT2 (3Q05) the extended European Research Area (ERA) will be enabled with an abundance of configurable 10 Gigabit/sec connectivity as a combination of national, cross-border and inter-national fibre paths.

[13] Based on the GÉANT2 Input to eIRG 2015 European eInfrastructure Roadmap, 22 February 2005

It is further expected that DF provision will be available in all European regions within the next five years, thus the near-term vision of the NREN community is the deployment of ubiquitous e2e services based on the lowest possible protocol layers. This vision includes most underdeveloped areas in an effort to ease the “digital divide” in Europe.

The trans-European network GÉANT/GÉANT2/GÉANT3 will take advantage of NREN optical facilities and will proceed in dynamic provisioning of production quality seamless connectivity, based on carefully drafted and continuously enforced SLAs among NRENs, optical fibre providers and DANTE (the co-ordinator of GÉANT).

In fact, all players involved have only to gain from this initiative, thus creating a true “win-win” situation. Complemented by distributed computing platforms (e.g. LHC computing & storage, FP6 grid initiatives such as EGEE, DEISA and SEE-GRID, and projects such as LOFAR and ITER), GEANT2 and its successor networks will greatly enhance the human re-search & development potential of the extended ERA, towards the fulfilment of the Lisbon objectives for a competitive knowledge society.

The networking facilities being developed within the GÉANT2 project and by Europe’s NRENs will equip Europe with the networking infrastructure and services to support grids until 2008.

Regarding services, GÉANT2 will develop, test and deploy the technologies to provide bandwidth allocation and reservation services, network performance diagnostic and enhancement services (including network monitoring facilities), and authentication and authorisation services for network access. These services will be adopted by the GÉÁNT2 network operation function and rolled out to users, providing an advanced platform on which to operate grid technologies.

In developing a European Science Grid, the ability to provide and control advanced high-performance networking facilities will be crucial to the stability and performance of the grid, as well as the definition of interfaces for the inter-operability of grid middleware with the network control and management planes. The use of dark fibre acquired from the “new market” implements a new model of “ownership” of the networking re-source, as it decouples the provision of the network from bandwidth provision – and the related pricing – by traditional carriers. This opens a completely new and innovative perspective for applications (like grids), as the cost of bandwidth is no longer a serious bottleneck for network provision.

Thus, longer term strategic issues not directly dependent on current practices and cutting edge technologies must drive e-Infrastructure planning, including research & education networking. The emerging business model should resolve fundamental questions like ownership of infrastructures, sharing policies, foresight of capital investment, consequences of technology driven choices etc. In this version of the White paper some policy issues related to the multi-domain aspect of the international connectivity requiring certain steps to accomplish end-to-end dedicated connections, as well as other related issues are being tackled.

7.3. International Connectivity

7.3.1. Current Status

The pan-European research networking arena has dramatically evolved the last ten years. From TEN-34 and TEN-155 European funded projects based on ATM technology and low speeds (34-155 Mbps) to the GÉANT10 Gigabit network was a huge move forward. The currently running project GN2 is the ambitious next step towards the next generation Pan-European optical network GÉANT2 that will fulfill advanced networking needs of the research and academic community, including seamless support for eScience initiatives. The GÉANT2 project supports the European Union's vision of a European e-Infrastructure. This e-Infrastructure will be a major contribution towards the realisation of the extended European Research Area. It further enhances European competitiveness and serves as a leading paradigm towards the realisation of the Global Terabit Research & Education Network (GTREN) with actual and planned high-speed connections within our Continent but also to the Americas, Japan, China, Korea, India, the North African Mediterranean countries etc.

Network research and test-bed networks are also essential. GN2 is working towards a separate test-bed network in parallel with the production one. The same is valid for the US, where there is separate network for research, HOPI, for testing disruptive technologies).

In short we see that shared and agreed organizations have greatly improved the worldwide landscape. Telecom market deregulations did help a lot, but still there are discrepancies among world regions concerning economics, policy, regulatory, and the like, resulting in still hard-to-reach places.

7.3.2. The future

We see as building blocks for global connectivity Research and Education (R&E) Networks (as alternative to commercial Internet only) and concertation between user representatives. These are important issues for sharing resources, tools etc. Equally important point is joint financing of links between countries and continents, especially in relation to developing countries. Discussions about what is fair share are on-going, however.

It is no longer sufficient to connect researchers to the Internet; they have to be connected to each other with a high-quality link with a QoS that is not disrupted by the normal Internet activities of some third parties. So the challenge is to provide seamless end-to-end connectivity over a multi-domain infrastructure and to make sure that there is a smooth migration from pioneering services in the R&E community into the commercial arena.

GLIF community shares a common vision of building a new grid-computing paradigm, in which the central architectural element is optical networks, not computers, to support this decade’s most demanding e-science applications. The challenge is to bring light paths to the desktop of the researchers and to their scientific instruments.

Today we see already hybrid networks: packet switched Internet for regular many-to-many usage and light paths for new high-speed few-to-few usage. This is now becoming mainstream in R&E networking, e.g. GEANT2 will be a hybrid network.

7.3.3. Recommendations

The major conclusion at which both workshops independently arrived was that it is important to realise that R&E networking is very much a multi-domain activity.

The eIRG would like to stress that global connectivity will never mean global network, but rather inter-domain cooperation and attention for interoperability. This must be taken into account for all planning that strives towards global end-to end connectivity for researchers.

Of course end-to-end connectivity as such and the smooth migration of users to it are still serious challenges that we have to overcome. The solution here will be to generate and agree on the necessary interconnection and middleware/AAA issues. Any notion that a single network will be able to solve this should be avoided, since it may slow down working on the real issues and diverges energy in the wrong direction.

7.4. The effect of network research

Network research is vital to meeting future generation Grid computing - with a strong focus on “vertical integration”. Integrating advanced optical technologies while advancing e-science applications and Grid middleware is critical. One shall not assume network technology remains static while research is conducted for advancing grid middleware and applications. Research projects will benefit the e-science community if the effort included advancing the network infrastructures/protocols to meet the new demands of these evolving E-science apps integrated with middleware.

Currently, we have a view of the behaviour of potential future enterprise applications by focusing on the needs of Big E-science applications. It could be argued that the enterprise applications will follow today’s E-science application’s use of the advanced network infrastructure and Grid integration. Such markets as the banking industry or medical are already taking steps towards this direction.

Collection of requirements for Enterprise is also very necessary for the development of the networks.

Interdisciplinary research is vital from application to the network - vertical integration.

Next generation networks could be vastly different than today’s mode of operation - should not constrain research to today’s model

7.4.1. Global e-Science network requirements

The basic e-Science requirements in terms of networking characteristics can be summarised in the following list:

- High bandwidth pipes along very long distances – terabyte transfers, petabyte, etc
- Network resources co-ordinated with other vital Grid resources – CPU, and Storage
- Advanced reservation of networking resources
- Deterministic end-to-end connections – low jitter, low latency
- Applications requesting optical networking resources host-to-host connections - on demand
- Near-real-time feedback of network performance measurements to the applications and middleware
- Exchange data with sensors via potentially other physical resources

7.4.2. The role of the Control plane

Automated and highly dynamic reconfiguration of optical end-to-end connection may play an important role in NRENs. It should be verified if dynamic reconfigurability can bring down cost. Thus the Research Networking community should follow closely and possibly contribute to the development, in-field testing, standardisation and adoption of control plane solutions, in close cooperation with industry. In parallel, intermediate short or mid term solutions should be adopted (e.g. based on the management plane) until the point in time when a unified, seamless and interoperable control plane is in place.

7.4.3. Recommendations

The eIRG would like to emphasise that network research is imperative to meeting next generation Grid requirements. Integrating advanced optical technologies while advancing e-science applications and Grid middleware is critical. One shall not assume network technology remains static while research is conducted for advancing grid middleware and applications. Thus e-Science networking requirements should be studied carefully, new technologies and control plane solutions should be investigated and possible integrated in the applications/middleware in order to automate the networking requests.

8. User Support policies

8.1. Introduction of the overall model for grid user support

8.1.1. Users and user support services

Users of grid architectures come from a wide area of disciplines and have a range of roles within these disciplines. Thus, for example, a user may be a doctor using grid technology to access confidential patient information, or an engineering application developer with highly developed computer skills trying to interface his application to grid services. In both cases they require basic support services appropriate to their requirements. These may be broken down into the following categories:

- –**Education.** Each user will require some basic education in getting started, and in developing skills for use of the grid. Thus the doctor would need instruction for the use of an appropriate interface for his application, with appropriate updates as it develops. The compute specialist would need instruction in interfacing his application to the grid, backed up by appropriate workshops and discussion groups. In both cases the educational process is essential and ongoing.

- –**Simple access to a broad range of information.** All users face the problem of navigating through the large volume of information available on all aspects of using the grid – getting started in a VO, simple user manuals, detailed technical manuals, self instruction courses, documentation on running applications etc. All users demand that this information can be accessed in a clear manner through appropriate portals and search engines. This is a key area in current large grid projects typically demanding cooperation between users, application developers, educators, deployment teams and the dissemination teams, in order to develop a simple and extensible system.

- –**Application integration and support.** All application areas attempting to use grid technology need skilled support in order to interface the application to grid services. Following the basic interfacing work there is then an ongoing period of supporting the application running in a grid environment that is evolving rapidly. The availability of such expertise is a key factor in the rapid development of the use of grids by applications.

- –**Round the clock support for the users of grid data, compute and networking services.** This area was discussed in some detail in the Den Haag paper, and is updated in this document. The recommendations from Den Haag for this area are given below.

8.1.2. Reminder of the recommendations from Den Haag

The chapter on User Support Policies in the Den Haag paper covered only the area of "round the clock support for grid users". It said

"The analysis here suggests that the federated model is better suited to the problem domain than the hierarchical model."

A few technical recommendations were made which are discussed in the section on grid user support that presents some experiences and lessons learned in the development of a federated support model. We now present more detail on the needs for user services in the key areas, together with technical and policy recommendations as appropriate.

8.2. The current experience and recommendations in the key areas

Below we draw largely on the experience in the first year of running of the EGEE project, and attempt to draw lessons for the support of large scale grid infrastructures.

8.2.1. User education

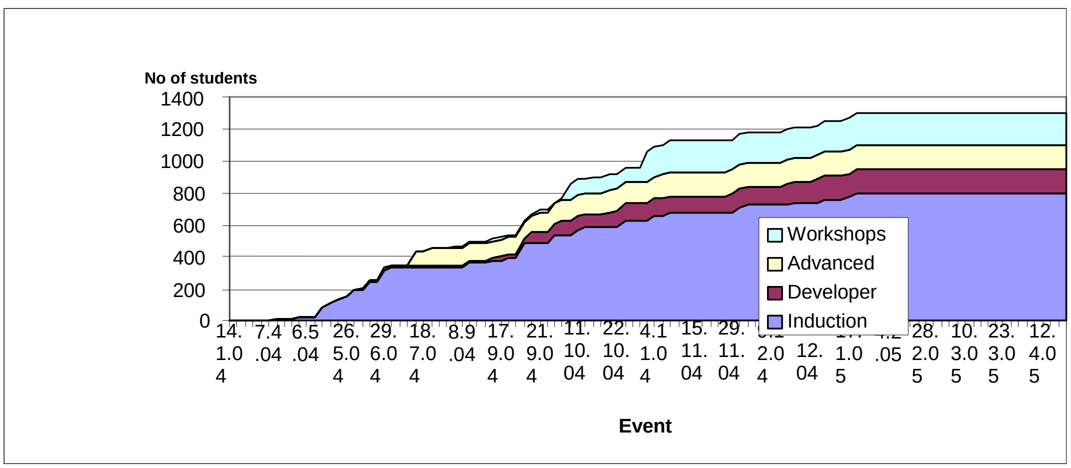


Figure 8-1: Total attendance at courses in the first year of EGEE

Figure 8-1 shows the development of the educational process within the EGEE project from April 2004 to April 2005. There has been heavy demand for courses distributed throughout the federations. This training activity has covered all levels from the basic grid starter induction course to a range of developer courses and intensive retreats. This is very much evolving with more applications using the infrastructure, and with the evolution of the grid services.

The demand for training is growing rapidly, and it is proving difficult to sustain high-quality training throughout all regions. This requires continuous collaboration and communication, and the provision of the appropriate level of human and technical resources. It is clear that training for any user is an ongoing service developing with the user knowledge and skills, and with the grid technology itself. It is clear that this is an essential area for the efficient use of grid technology in any major grid infrastructure.

Recommendation

Infrastructure planning should include the provision of resources for the co-ordination and delivery of an educational programme covering all aspects of grid usage.

8.2.2. Provision of information to users and information maintenance

The provision of all the information relevant to people working on different aspects of a Grid is a major undertaking. The vast number of people involved in such a project has a wide range in familiarity and knowledge about Grids. A huge number of documents aiming at different audiences has to be produced, checked and distributed.

We identify the following major user communities that have to be serviced with documentation:

- users:
    - Grid/distributed computing naive users;
    - Grid aware users;
- developers:
    - Grid-naive developers;
    - Grid-aware developers;
- advanced users/operations staff:
    - central project operations staff;
    - external operations staff.

Finding a solution to the problem of having to produce such a large amount of targeted documentation cannot be given to a specialized group of people. It is impossible for them to gain enough knowledge on all the topics that need to be covered. The problem can only be solved by a joint effort of all the stakeholders (users, applications support, education, system deployment, operations, dissemination).

A small team consisting of key people from the groups mainly involved in this effort should be steering the effort. They must make sure that all topics are covered and that the material is kept up-to-date.

Another big issue is the clear presentation of the material and easy user access. We opt for building a single point of entry for documentation, with subsections for the different communities addressed.

Documentation should be presented in a user-focused way. It should include the following material:

- for users:
    - quick start guides;
    - manuals;
    - security information;
- for developers:
    - quick start guides;
    - API reference documentation;
- for advanced users/operations staff:
    - organizational reference;
    - technical reference;
    - security information.

There is work in progress along these lines within the EGEE project, which needs to take into account similar work going on in other major projects, such as OSG[14], and the GGF working groups.

Summary comments on future developments

- The provision of appropriate information material to the whole range of users can only succeed in the long-term if there is a co-ordinated effort from all stakeholders with the appropriate knowledge and motivation, namely applications support, education, system deployment, dissemination and the users themselves. This work should be presented in the GGF environment to enable co-operation and cross-fertilization between major international grid projects.
- The information material should be presented via a single point of entry for all users. The emphasis must be on simple, users focused interfaces.

8.2.3. Provision of support for migration of applications to grid environments and ongoing support

Presently, there is a rapid growth in the number of applications using grid infrastructures. Interfacing applications to grid services requires skilled support, and appropriate, easily accessible resources. New application communities can be intimidated by the hurdles they must face in order to get their applications running on a grid infrastructure. In addition to learning about how to interface to grid services users must also find sites where they have the appropriate running rights.

Technical Appendix F summarizes the successful experience of the EGEE project with the provision of the GILDA test-bed for getting new applications started in grid infrastructures. This includes hardware and software resources supported by skilled personnel.

---

[14] Open Science Grid, http://www.opensciencegrid.org/

In addition to consultancy services, an application area needs expertise within its own community to provide ongoing support in helping to solve running problems, and in developing the applications themselves. The level of support required here is largely dependent on the scope and complexity of the computing models within the discipline.

Policy Recommendation

We recommend that the infrastructure planning includes the provision of resources for the support of application areas, both in the form of support for the first application migration, and with dedicated ongoing support for application areas according to their scope and complexity.

8.2.4. The day-to-day support for the users of grids using a federated model

The Den Haag Version of the White Paper was mainly focused on day-to-day grid user support. In this area, it favoured a federated approach to user support over a hierarchic model, which could be the right choice for smaller sized projects, or projects with a limited user community.

We will summarise our view of a federated support model and its advantages and problems, including some updates from the Den Haag version of the White Paper. We will comment on the recommendations from the Den Haag White Paper and present recommendations based on current experiences.

The description of a federated user support model is given using the EGEE project as an example. The EGEE project aims at a wide range of users from different science communities, including the LHC experiments, biomedical, astro-particle physics, computational chemistry and other applications. This diversity of usage led to a different approach in building a user support structure. To meet the needs of a diverse user community consisting of people from all kinds of sciences it was decided to go for a federated approach in user support, following the federated operational structure of the project.

Each EGEE federation has a Regional Operations Centre (ROC) which has the mandate to provide support to people within its assigned geographical region. To capitalize on this structure, user support was implemented by connecting the ROCs through a central helpdesk application. The Central Infrastructure Centre is an EGEE-specific concept, offering special services such as establishing a new VO. Additionally there are groups dealing with VO specific problems and support groups consisting of experts e.g. from middleware development or deployment. The whole support structure is shown in Figure 8-2.

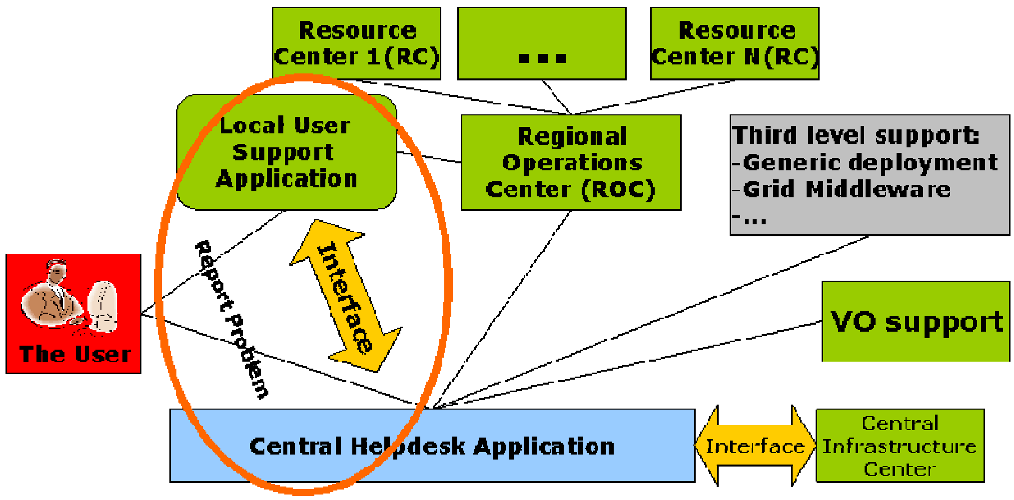


Figure 8-2 Possible schema of a federated support structure (EGEE as example)

All these support groups are interfaced to the central helpdesk application, which is used as an integration platform. The advantage of this approach is that each support group has to build only one interface via which they can reach all the other groups they might need to contact when reassigning user problems.

Such a federated structure ensures that all problems and issues of general interest are known at the central helpdesk application and can be fed to a central knowledge database that is accessible for all users from a central web portal. Together with a library of FAQs, which is also provided centrally, this constitutes a powerful tool for helping users to solve problems on their own, such as improper client-tool configuration or operation. At the central help-desk and at the ROCs users can submit their problems via a web interface or e-mail.

Technical Appendix F summarises the progress made with the implementation of the federated model and the main issues being worked on. It responds to the technical recommendations made at Den Haag, and adds new ones for future work. This work shows promising progress, and we should ensure that there is proper liaison with GGF and other major grid initiatives.

Policy recommendations for development of day-to-day grid user support

We recommend that the developments of day-to-day support systems for grids be fully supported.

It is important that user support system developments are fully documented and discussed at GGF and with other major grid projects. This applies also to the development of information systems for grid users.

8.3. Some comments on other key areas

8.3.1. Support for the underlying network infrastructure and the interaction with the Grid layer

Grid applications typically can have the following characteristics:

- very large data sets
- high end computing requirements
- remote instrumentations
- bandwidth on demand
- virtualisation of resources and data

and these characteristics lead to specific requirements for research networks.

The building blocks for the network infrastructure (federated model) for Grids are:

- campus networks
- NRENs
- GEANT
- International connectivity (US, AP, etc.)

The basic technologies of these building blocks are:

- IPV4 (state of the art today)
- Optical platforms (maturing and will be the technology of next generation research networks, because dark fibre is available and DWDM technology is ready for production network. This is a very cost efficient technology)

The services required by the research community are :

- general purpose IP (L3)
- Premium IP (L3)
- IPv6 (L3)
- VPN (L2 or L3)
- Optical VPN for specific Grid Communities like LHC (L1/L2)

Technical Appendix F discusses a typical operational environment, and the management and monitoring of the services provided. It also discusses support mechanisms for the services provided. This support for networking services must be integrated with the scheme for overall user support.

Conclusion and Recommendation

A policy for networking support for grid infrastructures should be agreed and made mandatory for all entities contributing to the operational environment. The implementation of the networking support should be interfaced to the overall user support infrastructure.

8.3.2. Software maintenance and development

The e-infrastructure would be incomplete without software. Scientific software acts as a skeletal framework for many scientific developments, as its implements incremental knowledge, approaches, algorithms and models available. Often, specific scientific codes are used by dozens or even hundreds of groups and thousands of scientists across Europe and the rest of the world – with a life span of sometimes several decades. In a way the software is in some areas – where complexity has risen to the point that researchers cannot but depend on the validity and accuracy of the code – becoming the most important carrier of scientific insights.

There are two realms of codes: the first is owned by privately or publicly held companies. and requires significant fees to use them. In a way these act as scientific publishers, incorporating the knowledge discovered by others into the software. This creates a number of problems, one of which is that one party decides where innovation can take place, which architectures are supported and optimised for. In a grid environment such features are undesirable, as they shrink the available set of resources. People also tend to have the same problematic relationship as with more classical scientific publishers: since many scientists are dependent on the codes they run for maintaining their pace of publication, financially draining the community to the point of extortion is lying around the corner. The only exit-scenario is to rebuild from scratch as a community effort, which is very time-consuming as these codes have been worked on for many man years and the internals of the software are unclear to outsiders. Even if some user communities of non-public codes would embark on such a scenario, this would still require more support for them for the next few years. For others, leveraging the buying power of the joint European users is essential.

The second realm of code is the software that is in the purely scientific domain. Many software solutions are spawned rather organically from small-scale initiatives – quite often e.g. from individual PhD-projects. Many of these codes have however outgrown their initial use and development scope, and the creators fall victim to their own success. Individual research groups that are now “responsible” for maintaining the codes – as good or as bad as possible – take considerable pride in providing such a service to the community. But they cannot be expected to pay for the whole development for years or even decades by themselves. At some point the software reaches a critical stage for their user communities: how does it continue to develop? Implementing other peoples algorithms, fine-tuning and optimising the code for new user groups and software environments and debugging – in short keeping the code up to date and usable to all –

lack the imaginative force of new discovery. Despite the importance of such codes and their widespread use, the developers therefore often lack adequate financial support or even a basic life cycle management infrastructure. It is not that people aren't seeing the importance of maintenance, optimisation and further development. On the contrary, but such activities are outside the scope of the basic funding structure and national orientation of most research funding agencies. In fact, software costs – including those of commercial codes – are systematically off-radar. That it is problematic is very clear, when we are about to enter an era where e-science is to blossom and the codes are to deliver the magic. Solutions at the EU level are needed; software crosses boundaries and local solutions are inefficient.

So there are three scenario's: the first is to leave the software as is, and just let it loose on a large scale on the e-infrastructure. This will at best just maintain inefficiency, and thereby invisibly occupy resources equivalent to significant amounts of money. If no investments are made, the science behind the software will suffer. At worst the community suffers from an accumulation of systematic programming errors, resulting in large scale scientific errors undetectable until someone does build another correct implementation or version of the software. This is obviously penny-wise and pound foolish. The second is to commercialize all the individual software packages. This may pay for some hard-needed initial changes and be economically sustainable in the future, but it is both hard to do from a legal point of view (since so far everything was built with public funds and with the contributions of many) and would create other problems. Commercialising would for instance take away the code from the open source domain – which is important for progress as it allows scientific scrutiny and the discovery of errors one the one hand and enables innovative dispute on the other. And of course it will bring about the problems described earlier as the software enters the commercial realm: reducing the amount of architectures and software environments supported and delivering scientists to the mercy of their software publisher. Third option would be to create some structural funds in order to have professional software engineers and scientists take the responsibility together for building, maintaining and consolidating scientific code. After all, not all scientists are programmers.

In both realms things are about to change because of the grid. In order to work in the new constellation with the grid middleware and new devices being brought in, significant changes will have to be made. We should leverage this, equivalent to the leveraging of the Y2K-situation that led to a significant upgrade of legacy installed base in the last years of the previous millenium.

Comment on all policy recommendations for support

The eIRG has reviewed the recommendations for grid user support in the key areas of user education, the provision of easily accessible user information, support for applications and the day to day running of the grid and networking infrastructure. The eIRG would like to stress that the planning for such support of the infrastructure should allow for the continuity of support in national structures beyond the end of major grid projects.

9. Towards a European Federated Middleware Institute

9.1. Introduction

It is clear that Grid middleware will form a fundamental component of e-Infrastructures across the European Research Area. To ensure the take-up of Grid technology in eScience it is important that Europe puts in place the necessary building blocks to support this technology. In this context, it would be highly beneficial to establish a European Federated Middleware and Application Repository, following the Open Middleware Infrastructure Institute paradigm, to provide such support.

9.2. A European middleware and application repository – "OMII-Europe"

To build reliable, robust and resilient e-Infrastructures to support the work of European science and industry requires high quality, standards-based, appropriately supported middleware software. Grid solutions, which are being deployed in many e-Infrastructure projects, rely on core Grid middleware to provide access to their resources and basic services to their users. To date, a large number of Grid technology development projects have been funded in Europe both by National Governments and by the European Commission. Many of these e-Science projects have produced software components which have been used to build prototype Grid-based e-Infrastructures. However, the plethora of current middleware initiatives lacks the coherence and direction that would pave the way towards light-weight, interoperable, plug-and-play Grid middleware environment.

The proposed institute would be a federated organization of software engineering centres utilizing their experience gained in developing robust, high-quality middleware for Grids. These software engineering centres would target to use the experience of existing National Centres in, for example, the United Kingdom (OMII UK and NeSC) and Italy (C-OMEGA), while trying to include as broad a spectrum as possible of centres with software development and maintenance expertise – spread across all Europe. The experience of EGEE and DEISA EC Research Infrastructures projects and EC Grid R&D related projects will also be a crucial factor. The proposed institute should adopt a common software engineering methodology – covering all phases from analysis through design, implementation and testing, to maintenance - and engage with other important Grid middleware development centres around the world. It is vital that the community builds on existing global standards and develop interoperable middleware solutions so that Europe can engage in the development of global e-Infrastructures to support the international collaborations of both science and industry.

Specific activities of the OMII-Europe will include:

- *Supporting the Community:* to sustain a well-managed, agile and reliable distributed Institute that delivers software to meet the evolving requirements of European science and industry.

- *Professional Software Development:* to run a highly structured, productive and stable software development process that will deliver integrated and well-supported middleware that is robust, reliable and resilient.
- *Quality Management:* to ensure appropriate quality assurance over all OMII-Europe processes, procedures and products. This will involve a definition of a set of relevant metrics providing quantitative assessments of processes and products.
- *Software Repository:* to operate an efficient process for selecting and acquiring software from a wide range of contributors, to develop and maintain the services required by the user community, to provide a convenient portal to support the download and installation of the services and their supporting environments and to support the structured development of these services through bug and problem tracking mechanisms.
- *Collaboration:* to engage in productive working relationships with other European and International fora (such as the OASIS and GGF), including National Grid Support and Operations Centres, user groups and standards organizations.
- *Outreach and Training:* to achieve close working relationships with European developers, and support for their projects so that good quality prototyping and development becomes the norm and to work closely with users to understand their needs and to help them understand and use OMII-Europe products.

Of particular importance will be continuous engagement with a wide variety of e-Infrastructure user communities to ensure that the products of OMII-Europe meet the needs of those user communities. Furthermore, this activity will be vital for us to realise the European Research Area in a meaningful way and thus meet the objectives of the European Council of Ministers. In the Lisbon declaration, making Europe the world's most advanced knowledge-based economy by 2010 is a top priority.

The opportunity to establish OMII-Europe must be seized now to ensure that Europe remains at the forefront of Grid middleware development and the deployment of e-Infrastructures.

9.3. Recommendation

The eIRG considers that Grid middleware will form a fundamental component of e-Infrastructures across the European Research Area. It is thus of key importance to endorse the principle of establishing a federated Middleware Institute to ensure the development of production quality Grid middleware leveraging EC as well as national efforts across Europe.

10. Usage policies

Chapter 7 of the Den Haag 2004 White Paper presented work towards developing a Grid Acceptable Use Policy (AUP). Its aim was that if users and applications were to abide by the simple draft AUP, it would be very likely that their use of the eInfrastructure would meet all the requirements of the many different Institutes and NREN AUPs.

The draft AUP text presented there started by defining the eligible Grid Community and went on to include a single sentence attempting to describe the foreseen general acceptable use of the Grid Infrastructure and/or Grid resources. Finally it tried to list a non-exhaustive set of possible unacceptable uses.

The final section of the chapter then explored several open issues with this simple and rather limiting approach. In particular it highlighted a number of issues which could be addressed in future versions of the White Paper, namely to include the appropriate treatment of

- for-profit activities;
- better definition of offensive and/or illegal data;
- personal or private (non-professional) use;
- proper treatment of the issues of liability and responsibility.

The subsequent e-IRG endorsement of the operation of the Joint (EGEE/LCG/OSG) Security Policy Group working on this common Acceptable Use Policy for multidisciplinary Grid infrastructures expressed its support for the current draft and encouraged the group to consolidate asap.

Following a lead on this topic from the Open Science Grid, the approach taken by the Joint Security Policy Group since Den Haag has been to make the AUP text as short and simple as possible. This not only means that there is a better chance that Grid users will actually read and understand the policy, but also the fact that the AUP no longer contains a long list of example unacceptable uses addresses many of the open issues listed above.

The aim is to have a single common User AUP, agreed to by all Grid users. Any diversity of acceptable and unacceptable use is then documented in a VO-specific way. In particular, there is no longer any explicit exclusion of for-profit or personal/private use in the general AUP nor does it attempt to define offensive or illegal data. These issues are best tackled by each VO in their own AUP. Each VO prepares its own policy and AUP which clearly expresses the aims of the VO, together with any necessary details of the applications and/or data involved. Grid infrastructure management and site/resource managers are then able to consider each VO on its merits before deciding whether or not to allow the VO to operate on the Grid, safe in the knowledge that all members of the VO will have agreed to the common User AUP at the time they registered with the VO.

The current draft JSPG text for the general user AUP is presented here.

1. *You may only perform work, or transmit or store data consistent with the activities and policies of the Virtual Organizations of which you are a member, and only on resources authorized for use by those Virtual Organizations.*

2. *You will not attempt to circumvent administrative or security controls on the use of resources. If you are informed that some aspect of your grid usage is creating a problem, you will adjust your usage and investigate ways to resolve the complaint.*

3. *You will immediately report any suspected compromise of your grid credentials or suspected misuse of grid resources to incident reporting locations specified by the Virtual Organization(s) affected and credential issuing authorities as specified in their agreements and policy statements.*

4. *You are aware that resource providers have the right to regulate access as they deem necessary for either operational or security-related reasons and that your use of the Grid is also bound by the rules and policies of the organizations through which you obtain access, e. g. your home institute, your national network and/or your internet service provider(s).*

This new draft User AUP is currently working its way through the approval processes of the EGEE, LCG and OSG projects. Once finalised and agreed it is hoped that this may also be used by other Grid infrastructures and projects.

10.1. Open Issues

Since the Luxembourg eIRG meeting, the new draft User AUP presented above has been working its way through the approval processes of the EGEE, LCG and OSG projects. Discussions with legal experts at several sites have highlighted a number of issues which still need addressing before finalising the text. These include making sure that resource providers maintain the right to control Grid access to their own resources and better treatment of data protection, criminal law, responsibilities and liabilities.

At the time of finalising this White Paper, these discussions have not yet concluded nor has a new AUP text been agreed. Once finalised and agreed to by EGEE, LCG and OSG it is hoped that this may also be used by other Grid infrastructures and projects.

10.2. Recommendation

Given that there is still no agreed text for the general AUP, the recommendation from the Den Haag White Paper is still valid:

“The eIRG notes the timely operation of an EGEE/LCG/OSG group working on a common Acceptable Usage Policy for multidisciplinary Grid infrastructures and it expresses its satisfaction and support for the current draft AUP proposed in this white paper and would like to encourage the group to consolidate it asap. It is felt that such an effort would greatly promote pan-European resource sharing for eScience.”

11. Advanced computing facilities policies

11.1. Introduction

The supercomputing arena was formed in the early nineteen eighties when the first vector computers became commercially available. Such systems were at least a factor of ten faster than other existing computers, and therefore found to be of strategic importance. Ever since that time, national policies for supercomputing have been a priority.

Supercomputing facilities need some critical mass to be affordable, and gathering this critical mass requires strategic policy activities. First country to set up a national policy was the USA, but this was soon followed by the countries in the rest of the western world and parts of Asia. Since these early days, the technology has changed drastically and the computational sciences have developed and matured. However, despite the fact that the systems have only grown larger and larger, the national focus of supercomputing policies has remained. Because of this, all countries cannot afford to keep up the continuous investments necessary today, because they fail to reach the critical mass.

Universities in the USA have access to their national systems based on the NSF's national policies. But in addition to this service infrastructure, both the USA and Japan have world leading systems that allow their scientists to do leading edge research. Furthermore, China is consistently acquiring systems belonging to the TOP500-list of world's most powerful computer installations. When trying to gain access to truly world-class computing installations, the European scientists tend to "hit their national walls" due to the lack of a European policy for advanced computing facilities.

Europe has invested a lot in the development of scientific user codes, both on the national level and European level (like the *PRISM* project in climatology). However, once crossing disciplinary boundaries, the compute intensity of those coupled codes will require new paradigms in high-end computing. For example, coupling of the research computing with sensor grids, or coupling of modelling with data acquisition are new areas where high performance computing needs to go beyond national facilities. Details of the technical rationale of the complementary nature between supercomputing and Grid are presented in section 15.1.

While bilateral co-operation agreements allow usage of advanced research equipment across borders (supercomputer use is no exception to this), there are no mechanisms for investments beyond the national scope in Europe. This is to be changed if European science and the related industrial innovation are to remain competitive with the USA, Japan, and China.

11.2. Rationale for a European policy

Why would Europe need an advanced computing policy beyond national boundaries?

Admittedly the first answer is one of policy rather than scientific ratio and it depends on the research profile Europe wishes to create and maintain for its own prosperity. This is of course a matter of choice. However, given the support for climate research, materials science, disaster prevention or management, genomics, astronomy, brain research and tissue engineering to name but a few, it is obvious that Europe has already chosen to be on the safe side in being an ICT-leading continent, where such research subjects are surely expected to contribute significantly to our economy and economic potential. If this is indeed the case, the enabling technologies in terms of infrastructures remain to be provided to realize this potential. This means that an advanced computing resources environment should be implemented on a European rather than on a national scale – using world class computing equipment.

The recommendation should be that "as long as there are significant developments in the top of the computer market, a focussed attention for these developments should be in place and special support is warranted to keep up with those developments." The key developments that are foreseen in the next ten years are presented in section 16.2.

11.3. European Advanced Computing Infrastructure

Present national large scale facilities[15] and the DEISA-project form a good starting point for the physical side of an advanced computing infrastructure in Europe. But two different visions exist for further policy:

- infrastructure, based on *disciplinary* needs (a system for climatology, one for materials science, one for genomics, etc.);
- infrastructure, based on computer *architectures* (64-bit vs. 32-bit addressing and arithmetic, low latency vs. high bandwidth, capability vs. capacity computing, etc.).

A summary of pros and cons for these visions:

11.3.1. Discipline-based approach

The main arguments for this approach are based on the perception that in-depth support, specifically directed towards the researchers in a discipline, is necessary and that community building is essential for progress in the field. On-site knowledge of the field and of the codes that are in use will play an important role in the development of the field, the efficiency of the research and the optimal usage of the equipment. Islands of knowledge can be brought together into larger and more encompassing knowledge environments. In addition, the notion of community building that will certainly result from this approach is highly valued: "this is our machine."

---

[15] See Academic Supercomputing in Europe

11.3.2. Architecture based approach

It is well known, but not always clearly communicated, that no single computer architecture will be the most efficient vehicle within one discipline (sometimes not even by far). In fact, even different codes to solve the same set of basic equations may behave orthogonally on a single architecture. The issue of user support, certainly in the context of a well developed grid environment, has to be solved but the organization of such a support structure does not need to coincide with physical computer installations. Instead, user support can be set up as a virtual organization even at the domain knowledge level, if it is organized well, and could also have some physical nucleus somewhere, if this is desired.

An approach based on architectures ensures that codes are run on the most cost effective machines for this code and puts the pressure where it should be, namely on the development of those grid components that are needed to couple all these architectures as closely as possible (one of the DEISA goals). Such an approach also contributes to all disciplines having access to the whole landscape of machinery and allows underdeveloped countries to take part in advanced research computing indiscriminately.

The practical process to support architecture based approach could be imagined to be a repeating cycle where the disciplines are “polled” about the codes and services they need in order to perform their scientific research. Based on the answers from the different disciplines, the aggregate needs for certain architectures can be compiled and used to optimize resource allocations and new acquisitions. This approach would combine the “science-centric” requirement analysis with savings based on optimizations (most cost effective architecture for individual computing task instead of limited set of architectures for each discipline) and also economies of scale (concentrating individual architectures into smaller number of centres, pooling hardware acquisitions in order to have a better bargaining position for the European computational science as a whole).

The outcome of the Amsterdam 2005e-IRG workshop clearly was in favour of the *model based on computer architectures*, not on disciplines, provided the issue of user support were well addressed in a European science Grid environment.

11.4. Advanced computer facilities and software

To be competitive on a global scale, advanced computer facilities and resources are needed on a European scale as well as on a national level,. It is far better to have, for example, five supercomputer systems in the top 12 list than to have only the #1 system on the global scale. Having five systems, for example, would also provide the structure for a network of excellence. This would help to keep and attract scientific talent from countries that so far do not pursue certain research because the necessary resources are not available. In the European research community, there are also not many applications that require the #1 very large system. Capability computing is needed, through access to suitable systems, to competitively further science. Nevertheless, local resources are still required on a national scale. Large national machines are often selected for “general purpose” jobs.

From a policy point of view, different architectures are required and should be promoted on a European level. This could be achieved by having a handful of supercomputing centres which are providing their networked high-performance computing resources both locally and to a broader scientific community. This could be achieved by using the framework of DEISA and EGEE.

Scientific software is an essential and integral part of the European e-Infrastructure. Solutions at the EU level are encouraged to handle the full lifecycle of scientific software. Local solutions are inefficient in the long run. Research funding is currently not supporting the maintenance and support of scientific software. In some cases it might even be easier to get funding for rewriting the software code than for fine-tuning, adapting and supporting an already existing code base. New and more efficient procedures have to be put in place to protect and fully employ the benefits from the already existing software code base.

12. Storage and data services

12.1. Introduction

By the end of 2005 the amount of information on the planet will increase from 3 million exabytes to more than 40 million exabytes (Source Disaster Recovery Journal Fall 2004). Generating and storing of large volumes of data has reached a critical mass. For most organizations, acceleration of data growth and the need to retain and store historical information are severely straining ICT capacities. Today's data infrastructure must offer high availability (24/7), scalability and robustness. Data storage is quickly becoming the largest expense in an enterprise ICT budget.

Parallel to this quantitative challenge, also the structuring and managing the lifecycle of the data are becoming more and more important questions as the e-Infrastructures allow the creation of larger and larger collaborations across institutional and cultural boundaries. Making sure that all the parties are working based on an up-to-date copy of the shared data is an important productivity issue. At the same time, increasing awareness of issues such as privacy and security pose further legal and regulatory challenges for the data service components.

For more detailed description of the current state of the art and trends and challenges that can be foreseen based on the situation, please refer to the section 17 in the appendix.

12.2. Access to publicly funded research data

The policy-issues related to data storage are not limited to technical issues. While establishing policies that help providing storage systems that are capable of managing the large volumes of data in a fault-tolerant manner - and enforcing the integrity and security of the stored data - are an important step forward, the production of the data itself can impose policy dependencies. On the one hand, IPR that is intended to be shared only among the partners of a collaboration have important effects on the processes related to management of the virtual organizations – in addition to the technical correctness of the security protocols and their implementations. On the other hand, in some cases the fact that the data has been produced as a result of publicly funded research activities causes somewhat opposite demand of making most of the data open to the broad public.

The *OECD Follow-up Group on Issues of Access to Publicly Funded Research Data*[16] is addressing such issues while working towards "good stewardship of public goods" and "creation of value-chains to leverage prior investments" in the digital domain. Based on this work, the OECD Committee for Scientific and Technological Policy at Ministerial

[16] http://dataaccess.ucsd.edu/

Level (29-30 January 2004) endorsed a Declaration on Access to Reseach Data from Public Funding[17], which stated in its Article 17:

> *Ministers recognised that fostering broader, open access to and wide use of research data will enhance the quality and productivity of science systems worldwide. They therefore adopted a Declaration on Access to Research Data from Public Funding, asking the OECD to take further steps towards proposing Principles and Guidelines on Access to Research Data from Public Funding, taking into account possible restrictions related to security, property rights and privacy (Annex 1).*

In the Annex one the actions the committee proposed commitment to:

> *Work towards the establishment of access regimes for digital research data from public funding in accordance with the following objectives and principles:*
>
> ***Openness:*** *balancing the interests of open access to data to increase the quality and efficiency of research and innovation with the need for restriction of access in some instances to protect social, scientific and economic interests.*
>
> ***Transparency:*** *making information on data-producing organizations, documentation on the data they produce and specifications of conditions attached to the use of these data, available and accessible internationally.*
>
> ***Legal conformity:*** *paying due attention, in the design of access regimes for digital research data, to national legal requirements concerning national security, privacy and trade secrets.*
>
> ***Formal responsibility:*** *promoting explicit, formal institutional rules on the responsibilities of the various parties involved in data-related activities pertaining to authorship, producer credits, ownership, usage restrictions, financial arrangements, ethical rules, licensing terms, and liability.*
>
> ***Professionalism:*** *building institutional rules for the management of digital research data based on the relevant professional standards and values embodied in the codes of conduct of the scientific communities involved.*
>
> ***Protection of intellectual property:*** *describing ways to obtain open access under the different legal regimes of copyright or other intellectual property law applicable to databases as well as trade secrets.*

[17] http://www.oecd.org/document/15/0,2340,en_2649_34487_25998799_1_1_1_1,00.html

__Interoperability:__ paying due attention to the relevant international standard requirements for use in multiple ways, in co-operation with other international organizations.

__Quality and security:__ describing good practices for methods, techniques and instruments employed in the collection, dissemination and accessible archiving of data to enable quality control by peer review and other means of safeguarding authenticity, originality, integrity, security and establishing liability.

__Efficiency:__ promoting further cost effectiveness within the global science system by describing good practices in data management and specialised support services.

__Accountability:__ evaluating the performance of data access regimes to maximise the support for open access among the scientific community and society at large.

Seek transparency in regulations and policies related to information, computer and communications services affecting international flows of data for research, and reducing unnecessary barriers to the international exchange of these data;

Take the necessary steps to strengthen existing instruments and - where appropriate - create within the framework of international and national law, new mechanisms and practices supporting international collaboration in access to digital research data;

Support OECD initiatives to promote the development and harmonisation of approaches by governments adhering to this Declaration aimed at maximising the accessibility of digital research data;

Consider the possible implications for other countries, including developing countries and economies in transition, when dealing with issues of access to digital research data.

On general level the above proposals are well aligned with the general goals and mandate of the e-Infrastructures reflection group, for example in the promotion of openness, professionalism, interoperability and quality & security. However, it should be noted that while the scope of the OECD recommendation is limited to data (unlike the e-IRG), it brings up the broader need to balance between openness with scientific, social and economic interests. Addressing this balancing issue in the broader scope of the e-IRG will bring in additional issues such as acceptable use policies and industrial contacts. Furthermore, sharing of computing power has potentially much higher cost than sharing of data – in the case of data various replication mechanisms can be deviced as a response to high demand, but this is not possible in case of resources such as CPU or network bandwidth.

12.3. Issues and recommendations

There is a strong need for infrastructure simplification across the borders of organizations and for investment protection over time to guaranty cost affordability of the needed storage solutions. Migration to new solutions should occur as growth occurs (just-in-time concept) and manageability must be ensured using existing (human) resources. This all leads to the need for a shared European approach to the data explosion challenge in many scientific fields, with a strategic plan to ensure flexible and efficient infrastructures and services based on open standards.

The use of open standards should enable integration and interoperability across Europe – irrespective of manufacturer. There is a strong shift in user preference from monolithic/proprietary solutions to open industry standards solutions. Such solutions are expected to enable massive scalability and to lead to price erosion over time, providing significant benefits and new opportunities to end users. Various middleware and application repository initiatives (as described elsewhere in this paper) will likely speed up this development. However, we recommend clear European guidelines for data classification, storage site classification, vendor compliance reliability, confidentiality, availability, integrity, accessibility and performance, certification and auditing.

There is also a clear need for the establishment of pan-European data availability, reliability, provenance and trust. Expensive investments and precious heritage can be at risk by technology obsolescence, security threats, affordability issues and fragility of technology. We feel that it is time for data storage to be regarded as a first class (European) citizen whereas up to now it has been regarded as just a bride maid.

We recommend therefore the establishment of a distributed shared network of European data centres, maintaining digital research data and other digital materials over their entire life-cycle and keeping them for current and future generations of users. This must include processes of digital archiving and preservation but also all the processes needed for good data creation and management, as well as the capacity to add value to the data to generate new sources of information and knowledge.

We also strongly recommend to co-ordinate the data management software development efforts across Europe and to stimulate P2P and Google-like technologies to be applied in research for data management purposes.

13. Grid and Industry in the context of the European Research Programmes

The transformation of research results into commercial products and services is a central objective of the European Research Programmes.

The emergence of Grids in research, and the potential of this new technology to lead to a new generation of ICT-infrastructures that can (r)evolutionary transform the economy poses a new challenge as to how fast and how effectively Europe can adopt these new infrastructures in its economy. Because the sharing of resources – that may be distributed on a global scale – is at the heart of Grids, the picture gets more complicated; there is also a need for the European industry to influence international standards, and for the European commercial entities to form alliances and to position themselves well in the international arena in the context of these new ICT-infrastructures.

In a (very) abstract view, there are two main trends that could be identified in the context of support by European research and competition programmes to transform research results into commercial products and services:

- A rather vertical approach supported mostly by the thematic areas of FP6 (like the IST) focuses on pilot implementations of new technologies in promising (regarding potential benefits) application fields. The main benefit of this approach is that it demonstrates that new technology "actually works" and that technology deployment can be fast, even if it is often on a restricted scale of application fields and actors involved.
- A rather horizontal approach supported mostly by the horizontal programmes of FP6 (notably the Research Infrastructures) emphasizes the provision of large-scale infrastructures to demonstrate the use of new technologies in real-world settings, to provide facilities to researchers and to industry engineers for "hands-on" experience, and to provide relevant training and operational support and advice to users. The last point aims more at lowering the barriers for the uptake of new technology by industry. This horizontal approach is of particular benefit to SMEs as the deployment and operation cost of new technology for such organizations is significantly higher than that of larger industrial entities.

The above efforts are complemented by other EU programmes that aim at increasing, in a more generic context, the competitiveness of SMEs and of European industry.

Maximum interoperability and cross-benefit from the above programmes is a continuous challenge for the European research community and funding authorities.

13.1. Recommendation

The e-IRG invites representatives of European enterprises and other commercial stakeholders to come forward and identify the expectations and needs from the business community and the contributions they expect to be able to make in return considering the long term goals of the European e-Infrastructures. e-IRG specifically also invites SME's to contribute their views and ideas.

In the same context the e-IRG wishes to be more proactive in the future in exposing its work to industry and to business.

14. Policy Roadmap

The following list of recommendations is based on the input received prior the e-IRG meeting and should be seen as a initial starting point, to be augmented based on the discussions in the Luxembourg meeting.

14.1. The role of the e-IRG in the future

It is important to reflect on the future role and scope of the e-IRG Reflection Group. A consensus should be reached about if – and to what degree – e-IRG should engage in dialog with stakeholders outside the public research infrastructure domain.

14.2. Towards National Grid Initiatives

This issue needs to further be analysed in the next White papers and discussed in future e-IRG meetings, while future Research Infrastructure projects like the second phase of EGEE and DEISA could provide support in the documentation and analysis of the feasibility and effectiveness of the EGO-NGI model. In this effort, cooperation with the NREN community should be sought, as the experience of the NREN community could be exploited analysing the commonalities and differences with the GEANT-NREN model. The following roadmap actions could be foreseen:

- Each e-IRG member state would evaluate the feasibility of establishing a NGI and planning for an open invitation for all the national projects to join in.

- A European policy for determining the common approach to the co-ordination functions of the EGO and NGIs should be established. The emphasis should be in maintaining local control of the resources, protecting the investments of the existing and potentially unknown projects (both in terms of technology and the social capital that is surrounding a specific Grid implementation). The emphasis should be in standardizing interfaces and protocols, not implementations.

14.3. Authentication and Authorisation Infrastructure

As main roadmap action the establishment of the GEANT2 AAI as a superstructure integrating all the (nation- or community-based) AA federations is advisable, and the exploration of how to make compatible current practices and systems (proxy certificates, VOMS, etc.) with the coming GEANT2 AAI.

14.4. Accounting

It is clear that work needs to continue to develop mature accounting schemes and gather experience on the grid market as a whole. Further work and studies need to be carried out in areas such as

- Gathering accounting best practices from existing consortia and computing centres and transferring them to solutions based on open standards. Also development of common vocabulary for accounting terminology as a part of the best practices work.
- Grid market pilots to identify bottlenecks and practical challenges in the economic models, for example pilots for utilisation of banking systems with Grid resource billing.
- Studies on legislation issues concerning resource trading.
- Intensifying international collaboration in defining common accounting standards and interoperability between existing and future solutions.

14.5. Legal Issues

Specific steps should be taken in order to make progress in the legal issues arena. It is advised that a dialogue between legal and technical experts takes place. An inventory of legal issues that are currently being encountered or foreseen to be relevant should be prepared taking as a starting point the Luxembourg White Paper. A dialogue could be initiated by means of a workshop, and on the basis thereof when deemed appropriate a more permanent group might be established.

14.6. Networking

Network Research is essential to meet next generation Grid requirements. Integrating evolving optical technologies with e-science applications and Grid middleware is critical. Thus e-Science networking requirements should be studied carefully, new technologies and control plane solutions should be investigated and possible integrated in the applications/middleware in order to automate the networking requests. As far as International Global Connectivity the major conclusion is that it is very much a multi-domain activity. Global connectivity will never mean global network, but rather inter-domain cooperation and interoperability. This must be taking into account for all planning that strives towards global end-to end connectivity for researchers. In this spirit it is advised to propose the following roadmap:

- Recommend joint financing of links between countries and continents, especially in relation to developing countries.
- In Europe there should be a separate test-bed or the experiments and the production must be separated in one network.

14.7. User Support

- The provision of appropriate information material should be presented in the GGF environment to enable co-operation and cross-fertilisation between major international grid projects.

- We recommend that the infrastructure planning includes the provision of resources for the support of application areas (both first application migration and also ongoing support).
- We recommend that the federated approach be pursued, and that this work be documented and presented at the GGF. This work should continue to be aware of user support activities in international grid projects such as OSG.
- The policy for networking support must be agreed and be mandatory for all entities contributing to the operational environment. The implementation of the NOC must interface to the global grid support framework as is done for the ROCs.

14.8. Towards a Middleware Institute

Grid middleware will form a fundamental component of e-Infrastructures across the European Research Area. It is thus of key importance to work towards the principle of establishing a federated Middleware Institute to ensure the development of production quality Grid middleware leveraging EC as well as national efforts across Europe.

14.9. Usage Policies

A single common User AUP agreed to by all Grid users is the way forward. Any diversity of acceptable and unacceptable use is then documented in a VO-specific way. Each VO prepares its own policy and AUP which clearly expresses the aims of the VO, together with any necessary details of the applications and/or data involved. Grid infrastructure management and site/resource managers are then able to consider each VO on its merits before deciding whether or not to allow the VO to operate on the Grid, safe in the knowledge that all members of the VO will have agreed to the common User AUP at the time they registered with the VO.

14.10. Advanced Computing Infrastructure

- A mechanism for pan-national funding for advanced computing facilities should be developed.
- The *adagium* should be that "as long as there are significant developments in the top of the computer market a focussed attention for those developments should be in place and special support is to be warranted to keep up with those developments".
- New and more efficient procedures have to be enabled to protect and fully employ the benefits from the already existing software code base.

14.11. Storage

- There is a strong need for infrastructure simplification across the borders of organizations and investment protection across time in order to guaranty cost affordability of the needed storage solutions.
- We recommend European guidelines for data classification, storage site classification, vendor compliance reliability, confidentiality, availability, integrity, accessibility and performance, certification and auditing.
- There is also a clear need for establishment of pan-European data availability, reliability, provenance and trust.
- We also strongly recommend to co-ordinate the data management software development efforts across Europe and to stimulate P2P and Google-like technologies to be applied in research for data management purposes.

14.12. Grids and Industry in the framework of EC programmes

The roadmap proposes the invitation of representatives of European enterprises and other commercial stakeholders to come forward and identify the expectations and needs from the business community and the contributions they expect to be able to make in return considering the long term goals of the European e-Infrastructures. In the same context an effort should be made in the future for the eInfrastructure community and eIRG to be more proactive in the future in exposing its work to industry and to business.

15. Technical Appendix A – Research Networking

15.1. Dark Fiber

How much is available and who owns it. What policy exists on its usage, what mechanisms are used by the researchers to get access to these networks.

15.2. Layer 1 vs. layer 2 vs. layer 3 Networks

NRENs Capabilities and Services

Is this information collected and accessible to researchers needing these services across international boundaries?

15.3. International Collaboration

The Optical network plays the most critical role in achieving this in the context of Grid Computing. How do we determine the shared use of these capacities and does a network scheduler exist? Should one be developed for all of Europe or for each individual country/domain?

15.3.1. GLIF organization

What it is and how it helps international collaboration?

The GLIF (Global Lambda Interchange Facility) community shares a common vision of building a new grid-computing paradigm, in which the central architectural element is optical networks, not computers, to support this decade's most demanding eScience applications.

There is a paradigm shift in networking through hybrid networking, combing IP and lambda networks.

- Packet switched Internet for regular many-to-many usage.
- Light paths (lambdas) for new high speed few-to-few usage.

Hybrid networking is now becoming mainstream in research and education networking. GÉANT2 will be a hybrid network. This introduces connectivity challenges like getting light paths to the desk top of the researchers and to their scientific instruments and hybrid networking functionality into the NRENs and LANs at the campuses

15.3.2. Definition of Control Plane

.An accepted definition of control plane is **"**Infrastructure and distributed intelligence that controls the establishment and maintenance of connections in the network, including protocols and mechanisms to disseminate this information; and algorithms for engineering an optimal path between end points."

More research and development will be required on this topic as the network continues to evolve to meet the high demands of the scientific research community. Below is a list of areas requiring research:

•Migrating functionality from centralized control to distributed at optical layer
–Distributed Fault management
•Self-healing opportunities at the optical layer
–Distributed Performance management
•Dynamically adjust the information to be collected to match context and near-real-time usability
–Distributed Configuration Management
•Autodiscovery
•Provisioning using signaling - GMPLS, OBS, OPS etc
•Determine what functionality makes sense from a centralized management plane vs. a distributed control plane

16. Technical Appendix B – Supercomputing

16.1. What defines a supercomputer

The dominating basics in architectures have changed dramatically over the past twenty years. From very fast single processor vector engines to massive batteries of well interconnected processors (either vector or scalar processors). Meanwhile two distinctive lines of massive computing have developed: *capability* computing and *capacity* computing. From the point of view of the researcher who has to do his scientific job, this distinction is not relevant. But in terms of things that can be done or not at all the distinction is relevant. All supercomputers from day one up to the present day have distinguished themselves as the systems to solve “latency bound” problems. Vector computers have extremely low-latency high-bandwidth connections between memory and the processor and massively parallel supercomputers have very low-latency and high-bandwidth processor interconnects (the network). As massively parallel computers are by definition modular systems the sheer size of such a system of course also determines whether a particular system will be counted as a supercomputer or as just another local piece of equipment. The important point is though, that low latency in a network physically requires very short physical distances. Remember that a 1 GHz processor ticks once in one nanosecond. Any data to be processed have to be closer than 30 cm from the heart of the processor to be there in time for the next tick according to the maximum speed of light (in vacuum). In practice signals pass glass fibre or copper wire at quite lower speeds. Any ticks for which data to be processed are not in time are a pure waste of time and money. This physical problem cannot be overcome with any technique now or in the future[18].

Grids will never take away the need for supercomputers. This is because any distributed computing using the grid will, by construction, not ever be capable of solving latency-bounds problems. The minimum latency of any communication process is physically limited by the speed of light (and in this case even more by the lower speed of light in fibre optics), and the speed of light is not big enough to guarantee, for distributed systems, the low latency communications needed for high performance. A typical MPI latency in a modern supercomputer is of the order of a few microseconds. This is boosted to a few milliseconds for WAN communications across the continent. Compared to the extremely high processor clock rates of modern CPU’s even this type of latency is killing in terms of the efficiency of the process. So real capability computing can only be achieved in compound tightly connected and physically closely packed computer systems.

---

[18] Of course by multi-threading or similar techniques the waste of cycles can be diminished, because the processor can start working for someone else, but it does not speed up the time to solution for scientific problem to be solved by the application.

The observation above is at the basis of the DEISA strategy. One of the components of this infrastructure is a tightly coupled distributed supercomputing system involving the leading IBM AIX systems in Europe. In this infrastructure, DEISA does not scatter a tightly coupled parallel application over several national supercomputers. As explained above, this is highly inefficient. DEISA uses instead job migration across national boundaries to reorganize the global workload on a European scale, so as to make room in one site for huge, demanding applications. End users get in this way much bigger partitions for a single job, and jobs from implicit DEISA users are transparently rerouted to other sites to maintain the site's job throughput. Job migration depends on network bandwidth (which can be improved), not network latency (which cannot be improved).

The grid, however, adds to the higher return on investment in such systems. Either through generalisation and the ease of the access, cross system scheduling, transparent file access across systems etc. And of course it allows connecting to the huge array of scientific devices and sensor arrays to provide real-time data input and access to distributed data – which allows for entire new applications for supercomputers. Capacity computing can and will profit from the computing power available through grids, and even more from the diversity of available systems and software environments. Whereas many scientifically very important applications can only be done on capability computers, there are just as many applications that are not latency bound. Such applications will much better be able to use the variety of simpler resources that will be available in a grid. It remains to be seen, however, if for such applications the investment of huge parallel (high bandwidth-connected but not primarily also low-latency connected) computer batteries might not be needed as well. For example for time-to-solution bound problems.

16.2. Technological developments

Are there sufficient relevant technological developments to be expected for the next ten years to set up a European advanced computing infrastructure?

Yes, there are. But instead of presenting an extensive technical exposé here on future developments, only a few highlights will be given here, which is completed by an extensive reference list at the end of the document.

Three elements contribute to the further development of real high performance computers: lower latencies and faster memories, improved architectures and technologies beyond the "silicon wall". In all these areas feasible projects are under development that will extend the time of significant changes at the very top of the high-end computer market to well beyond 2010.

Memory:

- near future: FC-RAM (Fast Cycle RAM) at least three times faster than DDR2-SDRAM
- intermediate future: MRAM (Magnetic RAM) (permanent, like SRAM)

- latency: more pipe lining, multi threading and/or
- “never leave memory”: Processing in Memory (PIM or C-RAM for Computation in RAM)
- Use FPGA's (Field-programmable gate arrays). Density 10x standard processor. Can be configured to implement memory and i/o interfaces and many important algorithms. Speedup: orders of magnitude!

Architecture:

- Connect chips by stacking them extremely closely together so that currents become coupled by capacitance: proximity bonding. This yields very high bandwidth (>20GB/s) and very low latencies (50-100 ps [!])
- Diversification of processors within one node
- Turning the architecture inside out: separately addressable (IPv6?) devices in a network centric supercomputer.

Beyond silicon:

- Yes, but not immediately: only somewhere around 2012-2014 the feature size of the switching elements will have decreased to 7-8 nm, the lower bound on a Si substrate.
- After that the nanotechnologies take over. HP works on a crossbar latch based on nanotechnology
- Opto-electronic devices: switches have been realised of 50-200 nm with switching times of 100 fs (10^{-15} sec.)
- Nano-mechanical quantum devices.

This overview is not nearly meant to serve as a comprehensive account of all possible developments in high-end computer research, but should substantiate the believe that significant developments are ahead of us, up to well beyond the 2015-2020 time frame. This would support the view that the proposals for European actions in this field will serve long time European interests and not just a temporary one. It makes from a return on investment point of view any investments in infrastructure more beneficial.

16.3. Build on DEISA?

The success of the DEISA project is proven partly from the ever growing list of supercomputer sites that wish to be involved are have joined the project. Thus it is important to recognise the values DEISA will bring to the grid and supercomputing community. On the other hand it has to be noted that on the longer term building the whole of the advanced computing arena on the DEISA project requires the DEISA concept to be compatible with the general grid trends that are evolving in the same time frame. Here we refer to the concept of “loose” versus “dense” grids. DEISA would be a dense, large granularity grid, serving the special needs and requirements of leading grand challenge applications, but otherwise compatible with the outside grid world. This fact has been recognized by the project leaders, which have recently enlarged the initial scope of the project with the definition of a precise grid architecture and roadmap for their heterogeneous supercomputing Grid. Hiding complex supercomputing environments from anonymous end users, and enabling interoperability with other Grid architectures, are some of the new objectives of the updated DEISA infrastructure.

On the other hand would building on top of the DEISA project directly upgrade the impact of the project as it would set the standards for all advanced research computing sites and give it the status of founding the high performance computing compartment of the European Science Grid.

17. Technical Appendix C - Storage and data services

Today we are experiencing a major data explosion in science and research. Supercomputers are not the only petabyte generators of data, also in experimental sciences large scientific equipment, like particle accelerators (e.g. LHC), radio telescopes (e.g. LOFAR), sensors, microarrays and high resolution scanners are highly contributing to this explosion.

- Bio-informatics queries: 500-1000 GByte in databases
- Medical imaging (fMRI): ~ 1 GByte per measurement
- Satellite world imagery: ~ 5 TByte/year
- Current particle physics: 1 PByte per year
- LOFAR (2007): 5-10 PByte per year
- LHC physics (2007): 10-30 PByte per year

The produced data need to be stored and retrieved, managed and tracked, accounted and billed. More data will be produced in 2005 than during the entire existence of humankind! By the end of 2005 the amount of information on the planet will increase from 3 million exabytes to more than 40 million exabytes (Source Disaster Recovery Journal Fall 2004). Generating and storing of large volumes of data has reached a critical mass.

- What is the impact of this data explosion on infrastructure and services?
- How to deliver, manage, store and recover the data?
- How to increase the real value of the data?
- Will the developments in storage technology be able to cope with the exponential growth of storage demands?

For most organizations, acceleration of data growth and the need to retain and store historical information are severly straining ICT capacities. Today's data infrastructure must offer high availability (24/7), scalability and robustness. Data storage is quickly becoming the largest expense in enterprise ICT budget. As ICT budgets are becoming tighter, coping with data growth is therefore a prime priority of administrators. The importance of adequate data lifecycle management is also increasing. Technology is offering a model of sharing resources. While processing power doubling every 18 months, memory size doubling every year, network speed doubles every 9 months allowing users to access their data more quickly than ever and opening the way to virtualization of ICT resources using internet, the web and of course preferably using the grid.

Several scientific disciplines, e.g. high-energy physics (LHC), and Lofar use the grid to store large amounts of data at computer/data centers classified functionally in so-called tiers. The storage requirements cannot be satisfied at a single site and there is therefore a strong need to exploit geographically distributed storage resources. The storage model is based on a hierarchical structure with the data produced and stored at the top tier, large national or intra-national centres as tier-1, national or institutional centres as tier-2 and users operating at the third tier. Many questions raise here:

- What are the major advantages and disadvantages of such approach?
- Is this approach just limited to large centralized data generators (like large equipment or large supercomputers)?
- How do we relate a storage tiering model with a processing tiering model?
- Where could efforts be combined?
- What would be the preferred funding and accounting models?

Infrastructure complexity should however be accounted for. Today many organizations have heterogeneous environments with various types of data stored creating challenges around:

- Scalability, Reliability, Security
- Disaster recovery, Business continuity
- Host, Operating system
- Legacy applications

Legislations and regulatory compliance heavily impact management of information life cycle creating challenges around:

- data lifecycle management
- keeping track with ever changing regulations
- increasing complexity
- privacy and security requirements

Europe can profit from having a shared approach to the increasing storage needs in science and research by establishing a distributed shared network of a limited number of large storage facilities, responsible for data integrity and ultimate backup and archiving. This will also yield the redundancy required for advanced data recovery.

What most end users then ask for are:

- Transparent access to distributed data archives
- Secure Backup and Archiving
- Storage interoperability regardless of platform
- Better configuration, monitoring and management tools
- Distribution of excess load throughout the WAN
- Reduced downtime
- Prediction of future storage needs from current usage patterns

- Alignment with business operations
- Service levels & Cost management
- Management of heterogeneous environment from a single interface
- Storage resource management and associated storage services, e.g. capacity planning, data migration, performance planning, etc.
- Improved TCO – with high availability, high volume processing and reduced overall cost per transaction
- On-demand services

Data grids help realise storage virtualisation masking complexities through a layer of logical abstraction and allowing capacity to be shared, partitioned, expanded and reallocated as and when required. European data grid infrastructure should be generic and able to handle different types of data of various applications, in a centralized or distributed data storage environment, depending on the data acquisition and processing model (e.g. LHC where data is mostly centralised versus life sciences where a large part of the data will be dynamically changing and stored close to the main users. Data grids should improve volume management with better capacity utilisation factors, offer greater flexibility and capacity management, lower down time and ensure higher data availability, improves the customisation of infrastructure with changing business needs, reduce management overhead, IT administration and management costs and lower the total cost of ownership.

18. Technical Appendix D- Sensor Grids

The amount of data, generated by measuring devices, has been growing very rapidly over the last decades. This growth has been caused to a large extent by the increased capability of computers to handle large volumes of data, both in terms of computing power and storage capacity. A special type of measuring device is a sensor network, where one or more types of detectors are spread over a certain region or area. The data obtained can be can scientific (e.g. geophysics, astronomy), environmental (e.g. weather, pollution) or serve a social purpose (e.g. surveillance, traffic control). The rapid growth of such networks is, next to the increased capability of computers, due to presence of high bandwidth fiber networks, that have recently been created in many European countries.

Deploying small, medium or large area sensor networks (the names refer to the surface that is covered, without being very precise) requires new techniques and concepts of various kinds. Whereas the design of a network addresses the logical and physical connectivity of the sensors, the term sensor grid is used to refer to the active control of the measuring devices, the data and computation management, and the knowledge recovery for whatever purpose the sensor network was set up for. Issues that often play a role in sensor grid design are the following:

- **Distributed sensor data access and integration.** This relates both to the heterogeneity and geographic distribution of the sensors within a sensor grid and how sensors are located, accessed and integrated within a particular study.

- **Large Data Set Storage and Management.** This relates to the sizes of data sets being collected and analysed. Online monitoring of data does not imply that all data must be stored. Decision making, for instance based on real time evaluation of the relevance of data, may lead to reduction of the volumes to be stored. Often, detailed analysis is done on data that are, at least temporarily, stored in a data warehoused for later use.

- **Distributed Data Access and Integration.** This relates to the distribution of data over various centers, due to the large data volumes involved or the special demands of the various end users. Also, in some cases the sensor data are correlated with different type of information, simultaneously detected or already stored in third-party data banks. Here, usually a variety of formats is required or encountered.

- **Open Data Analysis Computation:** This relates to the analysis applied to the data, which often requires a multitude of components such as statistical, clustering, visualisation and data classification tools. Since the end users may have widely variable requirements, it is essential that third party data analysis components and analysis workflows are integrated within the procedures of the sensor grid. Furthermore, if the analysis is to proceed over large data sets, it is essential that high performance computing resources are accessible.

In general, the special value of a sensor network and sensor grid is the real time analysis and interpretation of dynamic phenomena. When many sensors are involved, powerful 'streaming computing' is required to take care of the data. This will be important for fields as diverse as energy management, water management, precision weather forecasting (now-casting), natural disaster management, precision agriculture, earthquake monitoring, astronomy, and many others. The wide distribution of potential scientific, environmental and social fields points to the fact that the users of such sensor arrays will in general be highly multi-disciplinary and will derive from a diversity of communities. The synergy between those communities could be greatly enhanced by the use of a common e-Infrastructure. It is therefore important to include sensor grids in the e-IRG RoadMap.

The further development of sensor grids depends on various ICT hardware tools, such as fiber and wireless networking technologies, supercomputing, and developments in sensor industry (chemical, acoustical, imaging, infrared, etc.). Even more important is the innovation that occurs in software, related to various type of Grids. For instance, standardized interfaces are required, by which information can be exchanged from the simplest of sensors, e.g. cell phones, to the large computers systems that take care of the data by streaming computing. Also, controlling either a homogeneous or an inhomogeneous sensor network requires sophisticated tools, that are to some extent yet to be developed. The key ingredients of a sensor grid compared to other type of grids are the increase of the embedded processing power and the processing of streaming data flows instead of static data bases.

The data processing of sensor grids will increasingly involve comparison of the data stream with the results of computational modeling, run simultaneously with the data acquisition. Users may want to interact with these dynamic simulations that perhaps run on different computers and are guided by the real-time measurements of a (large) set of actively controlled sensors. Other important issues are authentication of the users of the sensor grid itself, the data that are generated, processed and distributed, and the computing resources. Protection is required, while maintaining low barriers for the multiple users which typically will be interested in the output of a sensor grid. So, in a sensor grid, many grid concepts come together. The requirements of interactions between different type of grids is sometimes referred to as requiring a grid of grids.

18.1. An example of a sensor grid

In the Netherlands, an example of a very large sensor grid will soon be build. This grid, which is called Lofar, is based on a wide area sensor network for astronomy, geophysics, precision agriculture and weather forecasting. When demand arises, the Lofar infrastructure will also be accessible for a wide variety of other scientific research fields as well as for the development of services and specific functions such as environmental monitoring and security. Lofar handles large data streams, gradually transforming sensor data into meaningful information for scientists of various disciplines.

The ICT activities at the moment concentrate mostly on astronomical research, since these place the highest demands on the sensor grid. In phase 1, Lofar consists geographically of a compact core area and 45 Remote Stations. Each Remote Station will be equipped with 100 high band antennas (110-240 MHz), 100 Low Band antennas (20-80 MHz), 13 three-axis vibration sensors (geophones), one to three micro barometers (for infrasound detection) and several auxiliary systems e.g. for weather monitoring and GPS time/position measurements. In the core area, with about 2 km diameter, there will be 32 substations. For the astronomical applications, there will be a total of 3200 high band and 3200 low band antennas in the core area.

First, the sensor data are digitized and processed at the station level. For the astronomical applications, the total digitized data rate at a station is 0.5 Tb/s. The (embedded) station level processing reduces this data rate to about 2 Gb/s. The geophysical sensors generate 2.4 Mb/s (geophones) and 3 kb/s (micro barometers). For these data rates, station level processing is not necessary. The bandwidth of the connection between each remote station and the central processing systems will be 10 Gbit/s. Assuming a worst case efficiency of 50%, the net sustained data rate will take about 5 Gbit/s, leaving sufficient bandwidth for monitoring en control and future expansions (both new applications and larger data rates for the existing applications).

The data of all stations are combined in a central processor. The input section of the central processor can accommodate 32 core stations and 50 remote stations at their full bandwidth, while simultaneously computational modeling can take place. The central processing power, required to perform these tasks when Lofar becomes fully operational is more than 30 Tflop/s. The data rate internally within the central processor is about 20 Tbit/s. An overview of the Lofar sensor grid is presented in Figure 18.1, below. Figure 18,2 shows the ambition of Lofar to extent on a European scale. The addition of additional sensor fields poses new challenges in terms of data transport, since these sensor data will be distributed over open networks (typically Géant).

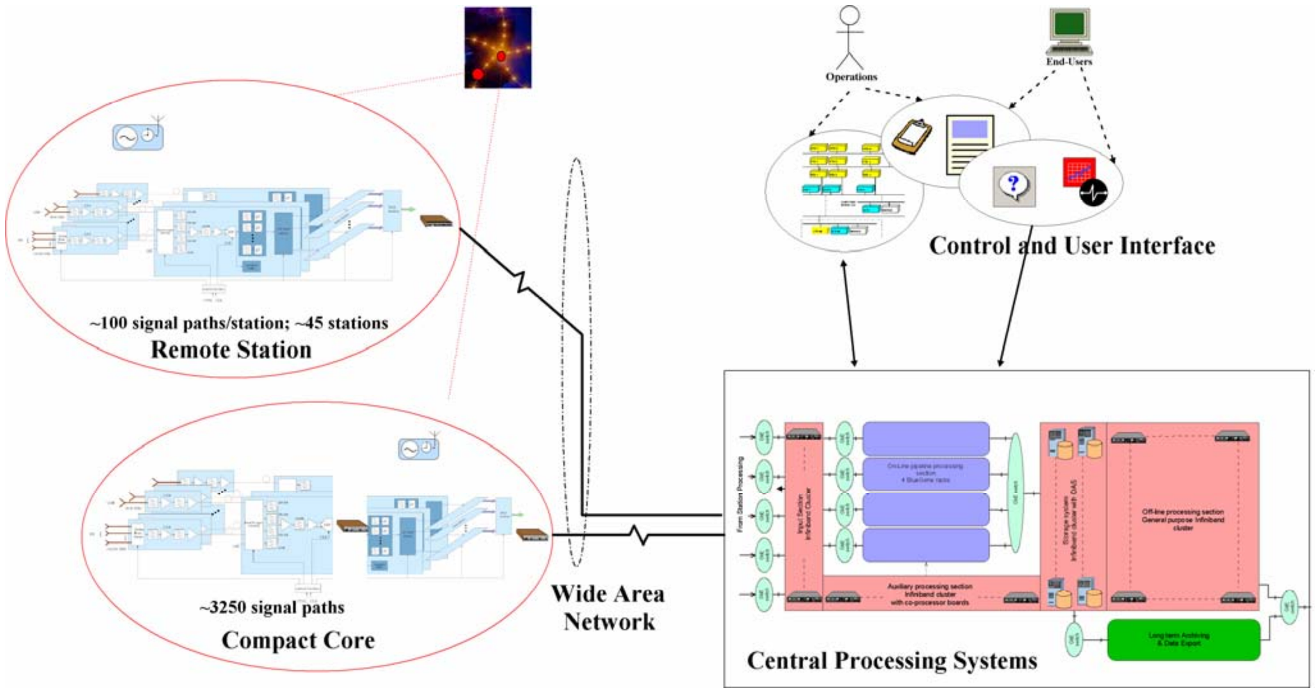


Figure 18.1 Technical overview of the Lofar sensor network

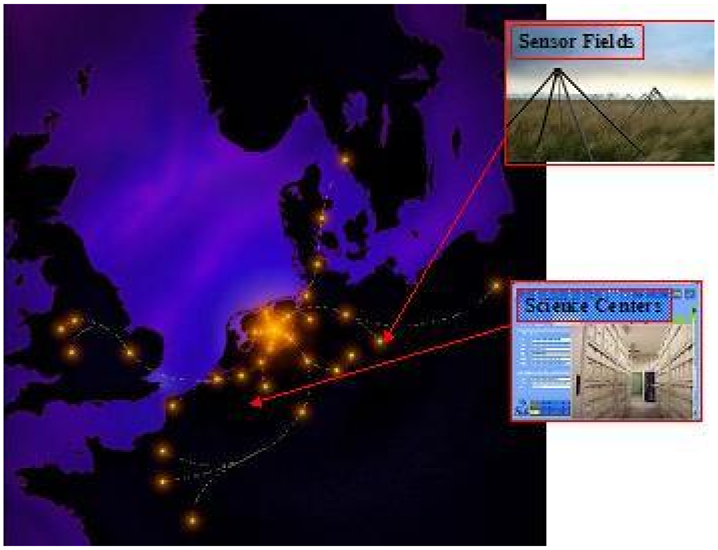


Figure 18.2: Extending LOFAR to a European Sensor Grid.

19. Technical Appendix E – Accounting

19.1. Progress since last White Paper

An update is given below of the status of the known grid accounting projects, most of which were described in detail last time

19.2. MCS

Highlights of the work are:

- A web services Resource Usage Service (RUS) that has updated the specifications of the GGF RUS-WG working group from OGSI to WS-I standards. A paper[19] on this was presented at ccGrid 2005

- The performance of this RUS was tested with insertion of a database of up to 3 million records where the scaling of insertion time is linear in this range. These results are extrapolated to be more than adequate for a Grid of the size of the NGS, possibly also will scale to the LCG size but would need further testing for this. (Reference: submitted to the UK All-Hands e-Science meeting 2005)

- We have submitted the RUS service as a possible RUS for the NGS and are awaiting evaluation of its suitability.

- Initial discussions with SGAS (SweGrid Accounting System) on collaboration.

- Our service has a security model proposed in the ccGrid paper.

19.3. SGAS

SGAS has been phased into the SweGrid production environment running NorduGrid/ARC since September 2004. At least 5 of the HPC centers involved in the SweGrid project make use of SGAS.

The current version of SGAS uses the GGF-UR taxonomy for Usage Records and has converters/translators from the Nordugrid URs and to the EGEE LCG UR formats.

Related work has taken place in NextGrid http:www.nextgrid.org Addressing both the contract and ontology aspects of accounting. An SLA use case document based on the

---

[19] *Implementing a Secure, Service Oriented Accounting System for Computational Economies*,J. D. Ainsworth, J. MacLaren, J.M. Brooke, in Proceedings of CCGrid 2005, Cardiff, UK. IEEE Press

SweGrid project was written. It is available at:
http://www.pdc.kth.se/~sandholm/docs/HPCAccountingUseCase.pdf

19.4. DGAS

DGAS (http://www.to.infn.it/grid/accounting/) was originally developed within the EU Datagrid Project (EDG) and is now being maintained and re-engineered within the EU Project Enabling Grids for E-sciencE (EGEE). The fully distributed system is designed in order for Usage Metering, Accounting and Account Balancing (through resource pricing) to be independent layers (see Fig. 19.1). A layered structure of accounting systems eases their standardization and thus enhances interoperability between systems.

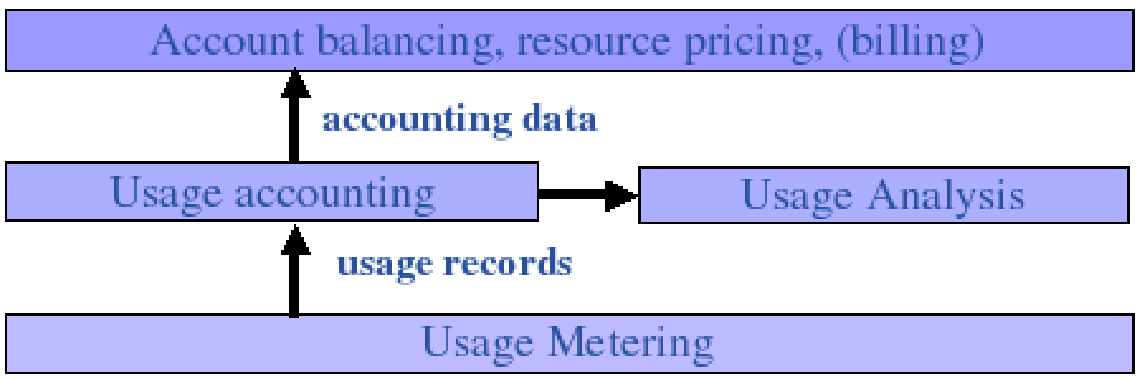


Figure 19.1 The layers of the DGAS Model

The Account Balancing layer is optional (the DGAS metering and accounting infrastructure can work without it) and may be used to implement economic models based on a market for grid resources, that may aid in balancing demand and supply as well as balancing the incoming workload. The Account Balancing provided by DGAS is intentionally generic. It may be used for different use cases, such as:

- Monitoring of overall resource consumption by users and resource contribution by owners.
- Redistribution of credits earned by a VO's resources to the VOs users (for balanced resource sharing between VOs).
- Billing/charging of users after resource usage.
- Credit/quota acquisition by users before resource usage.

The purpose of DGAS is not to define (and hence limit) the economic models/interactions between users and resource owners, but to provide the necessary means to enable them.

Some important issues addressed by DGAS

- Privacy:
    - all communications are encrypted
    - only authorised access to accounting data (users, resource admins, VO admins)
- Reliability of accounting information:
    - authorization of the accounting/payment transaction (indirectly) by the user
    - Usage records are signed using the user's certificate proxy
- Standardisation and interoperability:
    - Usage records (URs) can be transmitted from the Metering layer to the Accounting using the format of the GGF UR-WG.
- Collaboration:
    - Agreement on interfacing DGAS (provides anonymised accounting information) and APEL (publishes anonymised and aggregated accounting information).

DGAS has been deployed on a national scale on INFNgrid as well as on a smaller scale on a gLite prototype testbed.

19.5. APEL

APEL started in the LHC Computing Grid (LCG)Project and has been taken up by EGEE. A project requirement to account for resource usage within federations and at sites on a VO level has led to the development of the APEL accounting system. APEL is a log processing (or metering) application, which is used to scan batch system logs, such as PBS and LSF, to build accounting records based on the standard GGF Usage Record. It is able to piece together (by means of a join) the exact resource usage with a particular grid user from finding the necessary details stored within system and gatekeeper logs.

APEL is then used to publish the usage records into R-GMA (http://www.r-gma.org/index.html) where they are collated in a central archive at the Grid Operations Centre (GOC).

Each accounting record is unique and there is only one record per grid job. The records may be consolidated in different ways to provide high-level views of accounting data, such as the total CPU time as a function of time, Virtual Organization, country, as shown in Figure 19.2

Discussions have started on collaboration with DGAS and Open Science Grid on the exchange of Usage Records

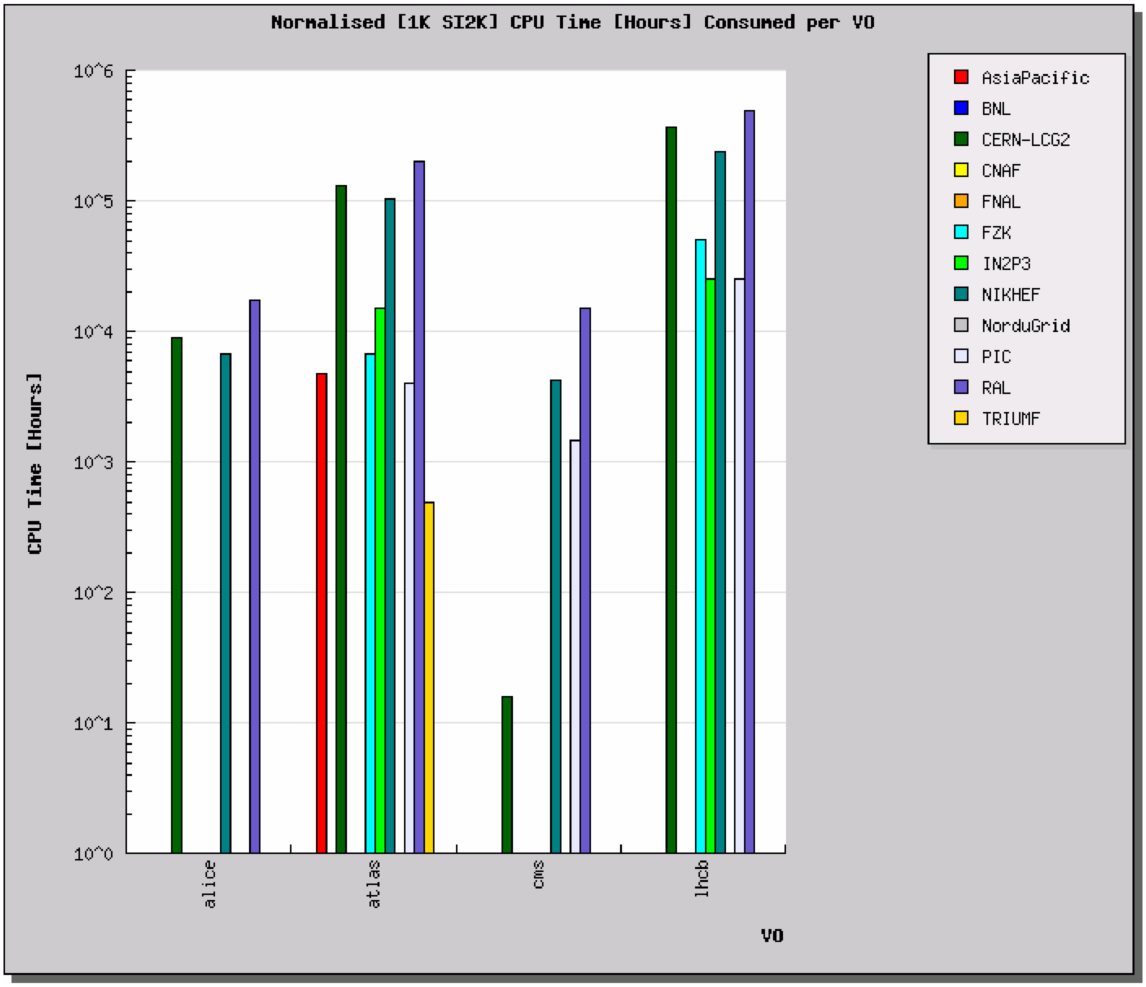


Figure 19.2: The total CPU time (Hours) provided by Tier-1 computing centres supporting LHC VOs during the latter part of 2004. Accounting data is gathered from sites running the LCG-2 middleware suite, which currently excludes a few sites (e.g. BNL, FNAL and NorduGrid).

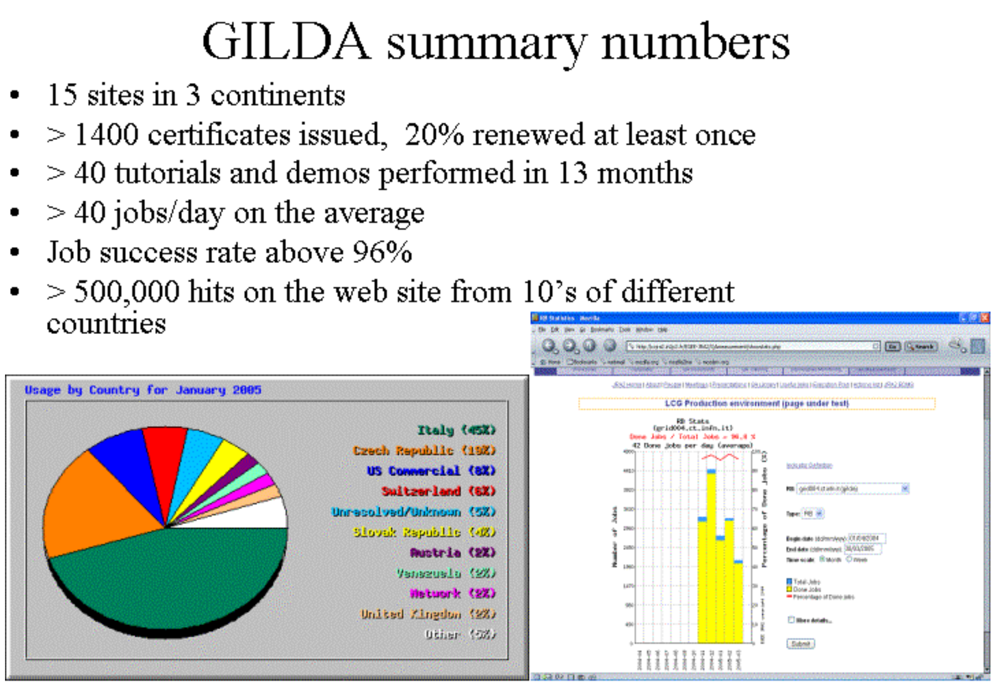
20. Technical Appendix F – User Support

20.1. Technical Annex ‘GILDA’

A very useful model used in the EGEE project has been the use of a special ‘introductory’ testbed GILDA with its own certification procedure, and sites providing the software and hardware resources for running. In addition the complexity of interfacing to grid services can be hidden behind a user portal GENIUS. In addition a team of experts provides consultancy services to the new applications getting started. Once the application has proven its basic operation it can then migrate to a full production service. This facility has been successfully used by a number of application areas including Biomedicine, Earth sciences, Computational Chemistry and Astrophysics.

It should be mentioned that similar consultancy services have been provided to the LHC experiments via the LCG/EIS team. And it is also interesting to note that members of this team have given invaluable start-up advice to the Biomedical community, again demonstrating the polyvalence of this approach. Such consultancy support can of course be distributed in the community, as long as tools exist to facilitate the overall organization.

We give below a summary of the usage of the GILDA infrastructure in early 2005



20.2. Technical Annex ‘GGUS’

Main issues being worked on in the EGEE/GGUS project are to ensure that all the support units are in place, that the workflow within these groups and the interfaces between the groups are well defined. The clear responsibilities of the groups must be known throughout the whole structure. A central co-ordination group consisting of members of all parties involved in the support processes is ensuring that the responsible persons are identified and that the structure is built up.

VO-specific support is a complicated task in an open project (like EGEE) where new VOs can enter anytime and therefore also new user support units have to be set up within these VOs. Also the new VOs have to be supported by the global support units and the workflow adapted accordingly. On the other hand, VO-specific support could be seen as just another “local user support application”; its primary function is to represent and support a group of users with a common goal, similar to a ROC whose function is to represent and support a group of users within a common region.

We look now at the advantages a federated support model can offer.

Problems are normally reported to a user’s ROC. Most ROCs are located at large universities or institutes, where a sizeable fraction of the users also work. Hence the problems will be relatively local in this model, allowing the support staff to debug by *e.g.* mining log files on the machines concerned, looking at process activities, *etc.* These options are usually not available to a central support facility.

The federated model also has the advantage that there are multiple entry points into the system; if a user’s ROC is down, she could report directly to the co-ordinating central system or to another ROC.

A minor disadvantage of the federated model is that there is always a small group of so-called “clueless users”, for whom it will be more difficult to find out “which is my ROC” than “where is the central support”.

The federated model has the *potential* to make a big impact in organizing communication between the parties involved in support, depending on how the interaction with the central support application is organized. If this application is organized along the lines of a relay system or back-end database, the communication lines between the various support teams are effectively direct.

Although based on experiences from EGEE we tried to keep the description of the federated model as generic as possible. In the following we would like to comment on the actual status of user support in EGEE. As of now the support infrastructure is mainly used by operations people to communicate, therefore an important step to test the workings and scalability of the system will be to get the real “users” onto it. It will require information, training and system adaptations to achieve this. EGEE will start a liaison with the U.S. project OSG (Open Science Grid). The plan is to make OSG look like a ROC in EGEE and make EGEE look like a Support Center in OSG. This integration should be achieved fairy straightforward in the user support area since both projects, though tackling originally the issue from different angles, came up with similar results, which consolidates our view of a federated model being best suited for user support in a grid environment.

The Den Haag White Paper gave several sub-recommendations on which we comment below:

- *Implement central helpdesk application as a relay model.* This still seems the right way to go. Experiences from EGEE have shown the importance of having a simple way of formal communication between support units, which a central relay system can provide.
- *Consider making user's IT department the Point of Entry.* Deciding this should be the responsibility of the federation, since it could interfere severely with their internal support structure to have something like this forced onto them. In the central system the specific implementation of the local support inside a federation has to be hidden. There should be only one contact point to each federation.
- *Consider a full text search model for the 'knowledge database'.* This still is considered as a very useful this to implement, judging from people's experiences of how to find information on the web.
- *Consider how to get feedback from knowledge database to the software developers.* Since software developers are also part of the overall support structure they should be informed about problems with their product. But a formal way to feedback information to the experts is still important.

20.2.1. More sub-recommendations for the federated approach

Additionally to the previous ones, which are mainly still valid, we can make a few more sub-recommendations for work in the short term.

- *Clear process definition is vital for a distributed effort like user support in a grid environment*. Only if the processes are in place will the overall support structure work. These processes must be properly documented.
- *Scalability* will become an important issue when the number of users on the grid will increase. Making the VO support groups an additional entry point for user queries should help to level to workload. How to define this 'process' is being worked on.
- *Support is done by people*, therefore it is important that all supporters are well trained in using the support infrastructure and are provided with clear and up-to-date documentation on the support processes. Only then they are able to do what is required of them.
- *Central co-ordination is important.* Even for a federated support model it is necessary to have a central body to co-ordinate the effort and have an overview over the whole structure.
- *Cooperation with other areas of support is vital.* VOs expressed their need for VO specific views of support portals offering them not only the possibility to submit

service requests, but also to access status and other information and to upload specific material needed by their members.

- *Liaise with other major Grid projects.* Connections with other Grid projects can have several benefits. Fusing the support efforts of different projects could help in building a “standard” implementation. Additionally all participants could profit from experiences made by other partners.
- *Contact GGF* to make the international community aware of this important work, and also to benefit perhaps from parallel activities in major grid projects

20.3. Technical Annex ‘Discussion of support for a typical production networking environment’

A networking operational environment is typically:

- End-to-end services over several domains have to be provided (*Campus A – NREN1 – GEANT – NREN2 – CAMPUS B* with several responsibilities and different SLAs)
- operational entities depending on the basic technology of the networks providing the e2e link have to collaborate to manage and monitor the service

The following table highlights Operational Entities and their tasks:

Network technology	Platform NOC	Hotline (Thin NOC)	IP NOC	GRID Operation Centre (GOC)
1) Optical platform	Provided by supplier of equipment	Provided by the NREN or GEANT	Provided by the NREN or GEANT	Provided by GRID Infrastructure Provider, has to have access to the underlying Operation centres
2) IP			Provided by the NREN or GEANT	Provided by GRID Infrastructure Provider, has to have access to the underlying Operation centres

The support for this operation must have the following characteristics

- the user of the GRID Infrastructure should have only one dedicated support interface for failure reports and questions, and networking issues must be routed to the GOC
- the GOC has to collaborate with the different NOCs/Hotlines for underlying resources (network, resource provider, middleware...)
- a database has to be established with information for e2e-links (network provider, technology of the involved networks, which NOCs/Hotlines are involved, SLA chains, responsibilities ...)
- support (24/7) has to be organized for the end user, resource providers and middleware providers and all necessary communication between them,

There is ongoing work on operational issues within the EGEE,and specifically on the operational interface between EGEE and GEANT/NRENs. A model

is being developed based on Premium IP and Best Effort Services provided by NRENs and GEANT. Issues of networks/services based on optical platforms have to be added to this model, and the model has to be tested for several communities. This could be the basis of a general model for the operational interface between network services and GRID Infrastructures.

21. Appendix A – Background and history of the white papers

For the development and support of the e-Infrastructures environment a series of workshops has been launched by the European Union under the aegis of the European Union Presidencies, see http://www.e-Infrastructures.org, in cooperation with the European Commission. The "e-Infrastructures" paradigm will reach its broadest scope and cross-border relevance with policy decision mechanisms that satisfy the diverse end-user communities' requirements of performance, service transparency and security, while providing economies of scale with ever-growing resources at attractive cost.

Athens, June 2003 - Launching the e-IRG

On 12th June 2003 under the auspices of the Greek presidency of the EU, the 1st workshop was held in Athens, organized by the General Secretariat for Research & Technology (GSRT), the European Commission, and the Greek Research and Technology Network (GRNET) in collaboration with the Greek National Documentation Centre (EKT). The workshop, entitled "Towards integrated Networking and Grids infrastructures for e-Science and beyond – The EU e-Infrastructures Initiative," was aimed at discussing the creation of the necessary policy decision mechanisms for the successful deployment of "e-Infrastructures" within the extended European Research Area. Among the key recommendations of the workshop was the establishment of an **e-Infrastructure Reflection Group (e-IRG)** with a membership "built from national representatives". The e-IRG "should consider and communicate clear messages on policy issues to both European Commission and existing infrastructure projects."

Rome, December 2003 - Consolidating the e-IRG and defining the scope

On 9th December 2003, the 2nd e-Infrastructure Open Workshop entitled "e-Infrastructures (Internet and Grids) – The New Foundation for Knowledge-Based Societies" took place organized by the Italian Ministry for Education, University and Research (MIUR) under the aegis of the Italian Presidency of the European Union with the High Patronage of the President of the Italian Republic, Carlo Azeglio Ciampi, and in co-operation with the European Commission.

One of the primary goals of the meeting, at which key players in the construction of the EU e-Infrastructure were present, was to review the perspectives and the technical and political issues related to the usage of the e-Infrastructure for Science and Society at the national, European and international level.

Dublin, April 2004 - The first concrete steps

On 15th April 2004, the 3rd e-Infrastructure Workshop, entitled "e-Infrastructures (Internet and Grids) – The New Foundation for Knowledge-Based Societies", took place, organized by the Irish Office of Science and Techology (OST), Science Foundation Ireland (SFI), the CosmoGrid project and Grid-Ireland, under the aegis of the Irish

Presidency of the European Union and in co-operation with the European Commission. The main results of this workshop were the following:

- Concrete next steps and actions were taken in the context of the afore-mentioned policy framework in Europe across technological and administrative domains:
    - *A concrete policy relating to grid authentication was endorsed.*
    - *The way forward for grid authorization and acceptable usage policies was described.*

The corresponding list of endorsements is given in Appendix B.

Den Haag, November 2004 – A series of recommendations

On 18th November 2004, the 4th e-Infrastructure Workshop was held as the opening event of a number of pan-European events around e-Infrastructure, grid technology and e-Science under the common title “European leadership in e-Science and grids.” The workshop was organized by the Dutch Ministry of Science, Culture and Education under the aegis of the Dutch Presidency of the European Union and in co-operation with the European Commission.

During this workshop, a series of areas was discussed where policies need to be developed including AAA (Authentication, Authorization, and Accounting), Acceptable Use Policies (AUPs) and User Support.

The list of endorsements from the Den Haag White Paper is given in Appendix B.

22. Appendix B - List of endorsements

Irish Presidency

The two main issues were endorsed by the e-IRG during the Irish presidency as follows:

"The e-IRG notes the timely operation of the EUGridPMA in conjunction with the TACAR CA Repository and it expresses its satisfaction for a European initiative that serves e-Science Grid projects.
The e-IRG endorses the principle of the EUGridPMA and TACAR. The e-IRG welcomes this development which positions Europe in the forefront of Grid and e-Science interoperability. The e-IRG strongly encourages the EUGridPMA / TACAR to continue their valuable work and recommends that they be supported by the relevant EU / national projects and agencies."

Dutch Presidency

The issues endorsed by the eIRG during the Dutch presidency are as follows:

"The e-IRG encourages work towards a common federation for academia and research institutes that ensures mutual recognition of the strength and validity of their authorization assertions."

"The e-IRG notes the timely operation of an EGEE/ LCG /OSG group working on a common Acceptable Usage Policy for multidisciplinary Grid infrastructures and it expresses its satisfaction and support for the current draft AUP proposed in the current white paper and would like to encourage the group to consolidate it as soon as possible. It is felt that such an effort would greatly promote pan-European resource sharing for e-Science."

"A forum dedicated to the co-ordination and exchange of technology and policy for disciplinary Grids should be formed. The task of the forum is to minimize duplication of efforts but still recognize and pronounce unique demands from disciplinary user communities."

"The e-IRG stresses the importance of deploying flexibly configurable and reliable end-to-end optical connections for research and education end-users (e.g. e-Science experiments). This provision should coexist with IP-routed services and build upon the European 3-tier hierarchical model consisting of the campus, NRENs and pan-European GEANT networks."

"The e-IRG gives high priority to the visibility of European infrastructures at venues such as the annual Supercomputing Conference organized in the US. The goal is to have an increased and continuous presence at booths, panels, talks, and keynote addresses. e-IRG encourages a co-ordinated European presence at the Super Computing Conference 2005. Europe should also focus on creating greater global visibility of corresponding

European venues. This could entail merging of some conferences to create critical mass and reach global impact."

Open issues for the UK presidency

Section 3:

A further amendment could be the explicit definition and use of the term "Virtual Organisation". VOs and their management is important and gets only covered at multiple places but is not discussed in detail. There remains a lot to do and to improve. Especially how VOs get better linked with the Home Organizations of their members is an open issue

Section 5

5.1 The parts of chapter 5 covering AAI are too detailed and might induce wrong expectations. About the SCHAC activity of TF-EMC2: Is it true that this group already considers "the inclusion of access entitlements and attributes supporting VO management"? What got included is a Spanish approach to encode role definitions (not in use on a larger scale yet). This is a good beginning, however standard roles are to be worked out and agreed upon first in inter-institutional and then in inter-federated environments. That will have an impact on the home institutions, since they probably have to adapt processes in order to be able to map its activities onto such standard roles. Standardization in this area would also allow VOs to make use of role information already available at Home Institutions.

5.2 Accounting seems to be very focused on billing. But there is more than just billing - e.g. statistics, credits (in an e-learning environment).

Section 6

An important aspect is the anonymity:

Anonymous access: Certain user communities need anonymous access to shared data repositories. This is typically the case if their access patterns reveal information about the nature or direction of their research work to outsiders.

Section 8

As users access the grid through membership in a (possibly short-lived) virtual organization, its creation, configuration and deletion should be made as easy as possible. In addition recommended VO policy guidelines templates should be formulated and agreed upon.

8.3.2. The two realms of code is a simplistic and incomplete view. Lots of FOSS (free and open source software) comes from outside the scientific domain. E-infrastructures use not just scientific software.

Section 9

The European Federated Middleware Institute should obtain special funding ear-marked for the implementation of missing desirable key components or functionalities of the supported grid infrastructures. Special care should be taken to widen the installable base (operating systems, versions, shared library dependencies) of existing grid software in order to spread their use to new user communities.

The European Federated Middleware Institute is mandated to run a software repository. It is supposed to make software easily available and keep track of problems and bugs. It is encouraged to also take care of security issues like vulnerabilities and provide push mechanisms to help customers to maintain not only functional, but also secure configurations. This function can be considered a grid software CERT.

Section 10

This Section drafts a generic user AUP and bullet 3 requires the user to report certain forms of abuse to specified incident reporting locations. This is fine, but the White Paper does not provide guidance on how those incident reporting locations are expected to deal with this information and how they collaborate, if at all. If that was the case, the user might be required to only report to one incident reporting location, which in turn will take further steps and inform the appropriate other affected entities in a controlled way. A policy statement on the availability of CERT services and their linkage to the existing CERT collaboration frameworks is encouraged. It should apply to entities like VOs, resource providers, home organisations and software distributors.

Section 14

The 12 paragraphs are very different in style. Some contents are found in the previous Sections. The whole chapter needs to be profoundly revised with the chapter's title in mind. So this means that the contents of earlier chapters need to be condensed so as to contribute something to a "Policy Roadmap". An introduction, and in particular, a summary and a recommendation should be added.

New sections: Add a new section on Virtual Collaboration Environments.

23. Appendix C: List of e-IRG members

Luxembourg list

eIRG board

Patrick Aerts (chair)	NCF/NWO	Netherlands
Antoine Barthel (vice-chair)	RESTENA Foundation	Luxembourg

eIRG members

Werner Kobliz	Aconet (Bundesm. für Bildung, Wissenschaft und Kultur)	Austria
Dieter Kranzlmueller	GUP/Uni Linz	Austria
Pierre Bruyere	Belnet	Belgium
Rosette Vandenbroucke	Vrije Universiteit Brussel	Belgium
Mario Campolargo	European Commission/DG Information Society	EU
Kyriakos Baxevanidis	European Commission/DG Information Society	EU
Jan Gruntorad	Cesnet	Czech Republic
Jan Windmuller	Ministry of Science Technology and Innovation	Denmark
Jaak Anton	IT advisor, Ministry of Education and Research	Estonia
Indrek Reimand	Head of Research Division	Estonia
Leif Laaksonen	CSC	Finland
Dany Vandromme	RENATER	France
Ruediger Krahl	Deutsches Zentrum fuer Luft- und Raumfahrt e.V.	Germany
Helmut Löwe	Bundesministerium für Bildung und Forschung	Germany
Fotis Karayannis	GRNET / Hellasgrid	Greece
Panayiotis Tsanakas	GRNET / Hellasgrid	Greece
Peter Hanak	National Office for Science and Technology	Hungary
Richard Hirsh	SFI	Ireland
Tom Sheedy	International Programmes, Enterprise Ireland	Ireland
Mirco Mazzucato	Istituto Nazionale di Fisica Nucleare	Italy
Francesco Beltrame	MIUR	Italy
Ilmars Slaidins	Riga Technical University	Latvia
Laimutis Telksnys	LITNET/Institute of Mathematics and Informatics	Lithuania
Theo Duhautpas	RESTENA Foundation	Luxembourg
Robert Sultana	University of Malta	Malta
Kors Bos	NIKHEF	Netherlands
Gudmund Høst	The Research Council of Norway	Norway
Rolf Larsen	Ministry of Education and Research, Dept. of Higher Education)	Norway
Norbert Meyer	Poznan Supercomputing and Networking Center	Poland
Robert Pasínski	Ministry of Scientific Research and Information Technology	Poland
Gaspar Barreira	Instituto de Prospectiva	Portugal
Dorin Carstoiu	Politehnica University of Bucharest	Romania
Branko Zebec	Ministry of Education Science and Sport	Slovenia
Sergi Girona	Barcelona Supercomputing Center	Spain

Luiz E. Ruiz	Ministry of Science and Technology	Spain
Torbjörn Wiberg	IT-Unit, Umea University	Sweden
Anders Ynnerman	SNIC	Sweden
Fiorenzo Scaroni	SWITCH	Switzerland
Anthony Hey	Engineering & Physical Sciences Research Council	UK
Alistair Dunlop	Open Middleware Infrastructure Institute	UK

Den Haag list

eIRG board

Patrick Aerts (chair)	NCF/NWO	Netherlands
Theo Duhautpas	RESTENA Foundation	Luxembourg
Brian Coghlan	Trinity College Dublin	Ireland

eIRG members

Dieter Kranzlmueller	GUP/Uni Linz	Austria
Mario Campolargo	European Commission/DG Information Society	EU
Kyriakos Baxevanidis	European Commission/DG Information Society	EU
Jan Gruntorad	Cesnet	Czech Republic
Jan Windmuller	Ministry of Science Technology and Innovation	Denmark
Mihkel Kraav	EENet	Estonia
Jari Jarvinen	CSC	Finland
Dany Vandromme	RENATER	France
Frank Schlie-Roosen	Bundesministerium fuer Bildung und Forschung	Germany
Ruediger Krahl	Deutsches Zentrum fuer Luft- und Raumfahrt e.V.	Germany
Fotis Karayannis	GRNET / Hellasgrid	Greece
Panayiotis Tsanakas	GRNET / Hellasgrid	Greece
Peter Hanak	National Office for Science and Technology	Hungary
Richard Hirsh	SFI	Ireland
Mirco Mazzucato	Istituto Nazionale di Fisica Nucleare	Italy
Francesco Beltrame	MIUR	Italy
Ilmars Slaidins	Riga Technical University	Latvia
Laimutis Telksnys	LITNET/Institute of Mathematics and Informatics	Lithuania
Robert Sultana	University of Malta	Malta
Kors Bos	NIKHEF	Netherlands
Gudmund Høst	The Research Council of Norway	Norway
Marek Niezgodka	Interdiciplinary Centre For Mathematical & Computational Modelling	Poland
Norbert Meyer	Institute of Bioorganic Chemistry PAS	Poland
Gaspar Barreira	Instituto de Prospectiva	Portugal
Dorin Carstoiu	Politehnica University of Bucharest	Romania

Branko Zebec	Ministry of Education Science and Sport	Slovenia
Julio Guzman	Ministry of Science and Technology	Spain
Anders Ynnerman	SNIC	Sweden
Anthony Hey	Engineering & Physical Sciences Research Council	UK
Alistair Dunlop	Open Middleware Infrastructure Institute	UK

List of Dublin meeting attendees

Richard Hirsh	SFI, Ireland	rhirsh@sfi.ie
Mario Campolargo	CEC	mario.campolargo@cec.eu.int
Kyriakos Baxevanidis	CEC	kyriakos.baxevanidid@cec.eu.int
Jari Jarvinen	CSC, Finland	jari.jarvinen@csc.fi
Anders Ynnerman	SNIC, Sweden	andyn@itn.liu.se
Jan Gruntorad	Cesnet, Czech Rep.	jan.gruntorad@cesnet.cz
Dany Vandromme	RENATER, France	vandromme@renater.fr
Pedro Veiga	FCCN, Portugal	pedro.veiga@fccn.pt
Laimutis Telksnys	LITNET, Lithuania	telksnys@ktl.mii.lt
Branko Zebec	Ministry of Education, Science and Sport, Slovenia	
		branko.zebec@izum.sl
Patrick Aerts	NCF/NWO, Netherlands	aerts@nwo.nl
Kors Bos	NIKHEF, Netherlands	bosk@nikhef.nl
Fotis Karayannis	GRNET, Greece	fkara@grnet.gr
Vasilis Maglaris	GRNET, Greece	maglaris@grnet.gr
Luke Drury	DIAS, Ireland	ld@cp.dias.ie
Grazynka Omarska	ICM, Poland	Gomarska@icm.edu.pl
Francesco Beltrame	MIUR, Italy	francesc@dist.unige.it
Ron Perrot	e-Science, UK	r.perrot@qub.ac.uk
Julio Guzman	Ministry of Science and Technology, Spain	julio.guzman@mcyt.es
Juan José Moreno-Navarro	Ministry of Science and Technology/Technical University of Madrid, Spain	jjmoreno@fi.upm.es
Tom Sheedy	Enterprise Ireland	tom.sheedy@enterprise-ireland.ie
Peter Hanák	National Office for Science and Technology, Hungary	
		hanak@nkth.gov.hu
Ferenc Vajda	Hungarian Academy of Sciences	vajda@sztaki.hu
Josef Lauri	University of Malta	josef.lauri@um.edu.mt
Mihkel Kraav	EENet, Estonia	mihkel@eenet.ee
Dieter Kranzlmueller	GUP, Uni Linz, Austria	kranzlmueller@gup.jku.at
Diego R. Lopez	RedIRIS, Spain	diego.lopez@rediris.es
Fabrizio Gagliardi	CERN-EGEE	f.gagliardi@cern.ch
David Kelsey	CCLRC/RAL, UK	d.p.kelsey@rl.ac.uk

White paper contributors

Luxembourg version

Editors

The overall editors of the White Paper were Fotis Karayannis, Matti Heikkurinen, Michiel Leenaras, individual sections were edited by:

2 The role of the e-IRG in shaping the ICT European Research Area: Kyriakos Baxevanidis (EU), Leif Laaksonen (CSC), including input from the European Science Grid session of the Amsterdam workshop.

3 Generic vs. disciplinary grids: Manuel Delfino (PIC), Patrick Aerts (NWO), Anders Ynnerman (SNIC)

4 Towards an European Grid Organisation and National Grid Initiatives: Fotis Karayannis (GRNET), Nikos Vogiatzis (GRNET)

5 Authentication, Authorization, Accounting policies: John Gordon (RAL), Kors Bos (NIKHEF), Diego Lopez (RedIRIS), Leif Laaksonen (CSC)

6 Legal issues in e-Infrastructures: Kees Stuurman (Van Doorne)

7 Network developments & grid requirements: Gigi Karmous-Edwards (MCNC), Dany Vandromme (RENATER), Kees Neegers (SURFNET), Vasilis Maglaris (NTUA), Fotis Karayannis (GRNET) and the GEANT consortium

8 User Support policies: Frank Harris (CERN), Torsten Antoni (FZK), Karin Schauerhammer, and Klaus Ullman (DFN), Hans Döbbeling (DANTE)

9 Towards a European Federated Middleware Institute: Tony Hey (EPSRC), Bob Jones (CERN), Ognjen Prnjat (GRNET)

10 Usage policies: Dave Kelsey (RAL)

11 Advanced computing facilities policies: Patrick Aerts (NWO), Victor Alessandrini (IDRIS)

12 Storage and data services: Anwar Osseyran (SARA), Leif Laaksonen (CSC)

13 Grid and Industry in the context of the European Research Programmes: Kyriakos Baxevanidis (EU), Michiel Leenars (NWO)

14 Policy Roadmap: Matti Heikkurinen (independent consultant)

15 Technical Appendix A – Research Networking

16 Technical Appendix B – Supercomputing

17 Technical Appendix C - Storage and data services

18 Technical Appendix D- Sensor Grids: Rosario Michael Piro (INFN)

19 Technical Appendix E – Accounting

20 Technical Appendix F – User Support

Hannelore Hammerle did review the white paper and was responsible for the Appendices.

Den Haag version

Editors

Parts of this version of the White Paper were taken from the Dublin version of this white paper. The overall editor of the White Paper was Fotis Karayannis (GRNET), individual sections were edited by:

2 The role of e-IRG: Patrick Aerts with contributions from Vincent Breton (IN2P3), Fotis Karayannis (GRNET), Kyriakos Baxevanidis (EU), Matti Heikkurinen (CERN), Risto Nieminen (CSC), Anders Ynnerman (SNIC), John Ainsworth & John Brooke (Manchester University), Antonia Ghiselli (INFN) and input from eIRG Dublin version.

3 Areas of Policy development: Patrick Aerts with contributions from Matti Heikkurinen (CERN)

4 Authentication (eIRG Dublin version)

5 Authorization policy development: David Groep (NIKHEF) with contributions from Klaus Ullmann (DFN), Karin Schauerhammer (DFN), Diego Lopez (RedIRIS), Javier Lopez (University of Malaga UMA), David Chadwick (University of Salford), Olle Mulmo (KTH), Dimitris Vergados - Konstantinos Katzourakis (University of the Aegean)

6 Accounting policy development: Kimmo Koski (CSC) with contributions from Juha Lento (CSC), Arto Teräs (CSC), Matti Heikkurinen (CERN), Gabriele von Voigt (University of Hannover)), John Ainsworth & John Brooke (Manchester University)

7 Usage Policy: Dave Kelsey (RAL) with contributions from Fotis Karayannis (GRNET), Kimmo Koski (CSC), Karin Schauerhammer & Klaus Ulmann (DFN), Arto Teräs (CSC) and Dany Vandromme (RENATER) (from Dublin version)

8 User support: Jeff Templon (NIKHEF) with contributions from Torsten Antoni, Karin Schauerhammer (DFN) and input from a discussion chaired by Flavia Donno on User Support at the LCG workshop, CERN, 2-4 Nov 2004

9 International and other related Efforts: Matti Heikkurinen (CERN) with contributions from Kyriakos Baxevanidis.

Rome version

Victor Alessandrini (CNRS-DEISA)
Kyriakos Baxevanidis (EU)
Ian Bird (CERN-LCG)
Alan Blatecky (SDSC)

Brian Coghlan (TCD-Ireland)
Fabrizio Gagliardi (CERN)
Francois Grey (CERN)
Fotis Karayannis (GRNET) – editor
Peter Kaufmann (DFN)
Dave Kelsey (UK GridPP)
Mirco Mazzucato (INFN)
Jesus Marco, (UNICAN-CSIC)
Federico Ruggieri (INFN)
Matti Veikko Johan Heikkurinen (CERN)

Dublin version

The editors for the Dublin e-IRG White paper comprised the following members:

Patrick Aerts, aerts@NWO.NL,
Kyriakos Baxevanidis, Kyriakos.Baxevanidis@CEC.EU.INT,
Kors Bos, k.bos@NIKHEF.NL,
David O'Callaghan, david.ocallaghan@CS.TCD.IE,
Brian Coghlan, coghlan@CS.TCD.IE,
Licia Florio, licia@TERENA.NL,
Antonia Ghiselli, antonia.ghiselli@CNAF.INFN.IT,
David Groep, davidg@NIKHEF.NL,
Peter Hanak, peter.hanak@NKTH.GOV.HU,
Matti Heikkurinen, Matti.Heikkurinen@CERN.CH,
Eduardo Jacob, jtpjatae@BI.EHU.ES,
Jens Jensen, J.Jensen@RL.AC.UK,
Christos Kanellopoulos, skanct@PHYSICS.AUTH.GR,
Fotis Karayannis, fkara@GRNET.GR,
Dave Kelsey, d.p.kelsey@RL.AC.UK,
Diego Lopez, diego.lopez@REDIRIS.ES,
Paolo Malfetti, p.malfetti@CINECA.IT,
Mirco Mazzucato, Mirco.Mazzucato@PD.INFN.IT,
Tom Sheedy, tom.sheedy@ENTERPRISE-IRELAND.COM,
Ferenc Vajda, vajda@SZTAKI.HU,
Dany Vandromme, vandrome@RENATER.FR,
Jules Wolfrat, wolfrat@SARA.NL,
Anders Ynnerman, andyn@ITN.LIU.SE

24. Appendix D: Abbreviations used in the White Paper

AA	Authentication & Authorization	
AAA	Authentication, Authorization, and Accounting	
AAI	Authentication Authorization Infrastructures	
AC	Attribute Certificate	
ACL	Access control list	
AUP	Acceptable Use Policy	
CA	Certificate Authority	
E2E	end-to-end	
EGEE	Enabling Grids for E-sciencE	www.eu-egee.org
e-IRG	e-Infrastuctures Reflection Group	www.e-irg.org
ESFRI	European strategy Forum on Research Infrastructures	
GGF	Global Grid Forum	www.ggf.org/
GN2	GÉANT2	
GOC	Grid Operation Centre	
HPC	High-Performance Computing	
ICT	Information and Communication Technology	
IP	Internet Protocol	
IST	Information Society Technologies	
LCG	LHC Computing Grid	www.cern.ch/lcg
LHC	Large Hadron Collider	
MoU	Memorandum of Understanding	
NOC	Network Operation Centre	
NREN	National Reseach and Education Network	
P2P	peer-to-peer	
QoS	Quality of Service	
SLA	Service Level Agreement	
SME	Small and Medium sized Enterprizes	
TEN	Trans-European Network	
VO	Virtual Organization	
VPN	Virtual Private Network	