



Design Study



EGI_DS

GRID INFRASTRUCTURES TRANSITION TOWARDS EGI

EU DELIVERABLE: D5.4

Document identifier: EGI_DS_D5-4

Date: **31 August 2009**

Workpackages: **WP5 - Establishment of EGI**

Lead Partner: **CERN**

Document status: **V1.03**

Document link:

Abstract:

This European Grid Initiative (EGI) Transition Document describes the current transition scenario from the established grid activities in Europe (such as EGEE-III [1] and related projects) to the EGI model as defined in the EGI Blueprint [2]. In parallel, project definitions responding to a call by the European Commission [3] are taking place. In this document we concentrate on high-level issues for the transition of existing grid users towards the new EGI paradigm in terms of operating the grid and supporting existing and new user communities.



Design Study



Table of Contents

EXECUTIVE SUMMARY	IV
1. INTRODUCTION	1
1.1. EDITORIAL RESPONSIBILITIES.....	1
1.2. DOCUMENT ORGANISATION.....	1
1.3. DOCUMENT AMENDMENT PROCEDURE.....	1
1.4. TERMINOLOGY.....	1
2. THE TRANSITION TO EGI	3
2.1. CURRENT STATUS.....	3
2.2. KEY REQUIREMENTS.....	3
3. THE SECOND YEAR OF EGEE III	4
3.1. MANAGEMENT.....	4
3.2. DISSEMINATION.....	4
3.3. TRAINING.....	4
3.4. APPLICATION COMMUNITY.....	4
3.5. POLICY AND STANDARDS.....	5
3.6. OPERATIONS.....	5
3.7. NETWORKING.....	5
3.8. FORMING PRODUCT TEAMS.....	5
4. MANAGING CHANGE	6
4.1. THE TASKS OF NATIONAL GRID INITIATIVES.....	6
4.1.1. <i>Selecting EGI Global Tasks</i>	6
4.2. REPLACEMENT OF REGIONAL OPERATIONS CENTRES BY NATIONAL GRID INITIATIVES.....	6
4.3. COORDINATING PROJECTS AND RESOLVING DEPENDENCIES.....	8
5. PROPOSALS FOR EGI AND RELATED PROJECTS	10
5.1. EGI-INSPIRE.....	10
5.2. SPECIALIZED SUPPORT CENTERS (SSC).....	10
5.3. MIDDLEWARE.....	11
5.3.1. <i>Transition Paths and Issues</i>	12
5.3.2. <i>Relationships with other projects or initiatives</i>	12
5.4. RELATED PROPOSALS.....	12
5.4.1. <i>Extra-European Collaboration – the CHAIN Proposal</i>	13
5.4.2. <i>Training, Dissemination, Business Outreach – the CUE Proposal</i>	13
5.4.3. <i>Dissemination channels for users, developers and policy makers – the GridTalk-II Proposal</i>	13
6. OTHER TRANSITION ISSUES	14
6.1. OPERATIONS.....	14
6.1.1. <i>Operations Coordination</i>	14
6.1.2. <i>Central and regional Grid user support and ticketing system</i>	17
6.1.3. <i>Coordination of interoperations between NGIs and other Grids</i>	17
6.1.4. <i>Operational Security Coordination</i>	18
6.2. TRAINING.....	19
6.3. USER SUPPORT.....	20
6.3.1. <i>High Energy Physics</i>	21
APPENDIX: SUMMARY OF TASKS AND EFFORT	23
BIBLIOGRAPHY	27



EXECUTIVE SUMMARY

The grid infrastructure for science in Europe was established through projects starting early 2001 with the European DataGrid (EDG) project followed by the Enabling Grids for E-Science project (EGEE), and its second and third phases (EGEE-II and EGEE-III) [1] which provide a large-scale production quality grid for scientists in Europe and beyond. The current infrastructure consists of 280 sites distributed in 48 countries, 140,000 CPU cores and 25 PB of disk space serving more than 15 science communities in 200 Virtual Organizations (VOs) with 13,000 active users. The EGEE grid (together with the Nordic DataGrid Facility in Europe and the Open Science Grid in the USA) has already achieved a high level of scalability, inter-operability and reliability in terms of services, software and operations.

The European Grid Initiative (EGI) is the next phase of the implementation of capacity computing in Europe. EGI will not provide any compute resources, but will guarantee transnational access to data and services. The EGI Blueprint [2] describes the proposal developed by the EGI Design Study (EGI_DS) [4] to establish a sustainable grid infrastructure for science in Europe in place by the end of EGEE-III in spring 2010.

After the approval of the Blueprint by a majority of country representatives, major steps towards the EGI establishment have been taken: The location of the central EGI location has been chosen to be Amsterdam and the interim EGI-Director has been elected. At the time of this writing, 30 NGIs representing their countries and two international institutions (CERN and the European Molecular Biology Laboratory, EMBL) have signed a Memorandum of Understanding (MoU) [5] and are therefore founding members of the EGI Council.

The outgoing projects and EGEE III in particular enable a smooth transition to EGI by adjusting the work plan for the final year. The major changes are:

- The scientific discipline clusters will be prepared to move towards Specialised Support Centres (SSCs).
- The operation structure will alter to reflect the move from central to distributed infrastructures.
- The software certification will be transferred to individual clusters of competence.
- The replacement of Regional Operations Centres (ROCs) by National Grid Initiatives represents a major change in the way the grid is operated and managed.

In parallel to the writing of this document, proposals responding to the EC calls FP7-INFRASTRUCTURES-2010-2 [3] are being developed. The proposals addressing the management and operation of grids in Europe, the development of software and the support of users will provide a much more detailed view of the initial phase of EGI identifying, in particular, assignments of specific tasks to NGIs and institutions. The detailed repartition of these tasks and assignment to particular calls and projects is still being discussed and are therefore not covered in this present document.

A procedure was established to finalize the assignment of EGI Global Tasks to individual NGIs. This is an important pre-requisite for the agreement on the EGI proposal. In order to ensure coherence and clear definition of responsibilities, agreements between the different projects need to be formulated.

A number of items for grid operation require special attention during the transition period. It is essential that the security coordination can migrate smoothly into the EGI era. The training effort will move from a central coordination to a distributed infrastructure.

The user support working successfully now for several scientific clusters in EGEE has to move to SSCs without disruption. This is of particular importance for High Energy Physics representing the biggest grid user community. The first operation of the largest scientific instrument, the Large Hadron Collider (LHC) at CERN takes place in the mid of the transition phase.

1. INTRODUCTION

This document outlines issues related to the transition from current Grids in Europe to the EGI.

In parallel to the writing of this document, proposals responding to the EC calls FP7-INFRASTRUCTURES-2010-2 [3] are being developed. The proposals will provide a much more detailed view of the initial phase of EGI and provide, in particular, assignments of specific tasks to NGIs and institutions. The detailed repartition of these tasks and assignment to particular calls is still being discussed and can therefore not be covered in this present document.

1.1. EDITORIAL RESPONSIBILITIES

The individual sections have been developed by different groups of people, with many contributions especially from EGEE - including members not only from the organisations involved in the EGI_DS project; the actual editorial work has been performed by WP5 with participation from all other work packages under the guidance of the EGI_DS Management Board.

We acknowledge the essential input from the editors of the EC proposals under construction.

1.2. DOCUMENT ORGANISATION

Chapter 2 summarizes the current status of the transition and enumerates the key requirements. Chapter 3 concentrates on transition issues that are being addressed in the last year workplan of the EGEE III project. Chapter 4 addresses major concerns in managing the change: the take-up of responsibility for global tasks by the National Grid Initiatives and the coordination between several projects. In Chapter 5 we give a brief overview of EGI-related projects that are being prepared in response to a call by the European Commission. In Chapter 6 we address some specific issues in the domains of operations, training and user support. The Appendix contains a list summarizing all EGI functions and the related effort.

1.3. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors.

1.4. TERMINOLOGY

This subsection provides the definitions of terms, acronyms, and abbreviations required to properly interpret this document.

Glossary

ARC	Advanced Resource Connector
CA	Certification Authority
CAO	Chief Administrative Officer
CERN	European Organization for Nuclear Research
COO	Chief Operational Officer
CPU	Central Processing Unit
CTO	Chief Technical Officer
DEISA	Distributed European Infrastructure for Supercomputing Applications
EC	European Commission
EDG	European Data Grid
EGEE	Enabling Grids for E-science
EGI	European Grid Initiative



EGI_DS	European Grid Initiative Design Study
EU	European Union
EUGridPMA	European Policy Management Authority for Grid Authentication
FTE	Full Time Equivalent
GGUS	Global Grid User Support
InSPIRE	Integrated Sustainable Pan-European Infrastructure for Researchers in Europe
LHC	Large Hadron Collider
MCB	Middleware Coordination Board
NGI	National Grid Initiative
NREN	National Research and Education Network
OGF	Open Grid Forum
PB	Peta Bytes
QA	Quality Assurance
ROC	Regional Operating Centre
SDC	Strategic Discipline Cluster
SLA	Service Level Agreement
SSC	Specialised Support Centre
UCS	User Community Services
UCO	User Coordination Officer
UFSC	User Forum Steering Committee
UMD	Universal Middleware Distribution
UNICORE	Uniform Interface to Computing Resources
VO	Virtual Organisation
WLCG	Worldwide LHC computing Grid Project

2. THE TRANSITION TO EGI

2.1. CURRENT STATUS

After the publication of the EGI Blueprint [2] in December 2008, a number of implementation milestones have already been achieved:

- Acceptance of the strategy defined in the Blueprint by the EGI_DS Policy Board;
- Choice of the location of the central EGI location, EGI.eu (previously called EGI.org) in Amsterdam;
- Acceptance of the EGI Memorandum of Understanding (MoU) and subsequent establishment of the EGI Council;
- Election of the interim EGI Director;
- Publication of calls for proposals in the 7th Framework of the European Commission targeted to EGI and related e-infrastructures;
- Establishment of editorial structure for the preparation of proposals for “EGI-PROPER”, SSCs, Middleware, and other topics related to EGI such as CHAIN;
- Establishment of the gLite Consortium;
- First drafts of proposals internally available.

At the time of this writing, 30 NGIs representing their countries and two international institutions (CERN and the European Molecular Biology Laboratory, EMBL) have signed the MoU.

In addition, the NGIs have confirmed agreement with the EGI function definitions through their responses to the EGI function questionnaire [6]. The tasks to be performed are described in the EGI_DS “Function Definition” document [7]. The summary table of individual functions has been reproduced from that document in the Appendix of the present document.

2.2. KEY REQUIREMENTS

Key requirements during the transition phase are:

- The need for utmost transparency of the changeover for the existing user communities.
- The seamless continuation of the grid operation and the provision of CPU and storage resources.
- The passing on of operations from Regional Operating Centres (ROCs) to individual National Grid Initiatives (NGIs) in each country where that operations model has been chosen.
- The continued maintenance of the middleware used by the communities and its development according to the needs of the users.
- The continued inter-operation with Grids outside Europe such as OSG [8].

3. THE SECOND YEAR OF EGEE III

In its bid to provide a world class, coherent and reliable European grid, ensuring that Europe remains at the forefront of scientific excellence, the EGEE-III project, which coordinates pan-European access to production-quality high throughput computing resources in Europe, has been working closely with the European Grid Initiative Design Study (EGI_DS) project to provide a smooth transition of the current production infrastructure to the operating model proposed for EGI. As EGEE enters its second year these changes are being phased in across the project assuming that EGI will start on the 1st May 2010 and that tasks (not resources) can be adjusted within EGEE.

3.1. MANAGEMENT

The project management team within EGEE has established strong links with the EGI_DS and the editorial boards related to the proposals that will continue and build upon the work done within EGEE. These are the 'EGI Proper', the European Middleware Initiative (EMI) and the various SSC proposals. As the internal structure of EGI becomes clearer, EGEE will start detailed transitioning planning once relevant EGI staff are in place.

The EGEE-III Technical Management Board (TMB) will continue to manage the technical process of the project focusing on the work of the product teams in monitoring their ability to deliver high quality releases to a predictable schedule and the operations automation team in regionalisation of the operations tools.

3.2. DISSEMINATION

The dissemination activity will become focused on communicating to the EGEE community, the progress of plans to transition to EGI and moving regional dissemination activity to the NGIs. This will include continuing the collaboration with EGI_DS with both projects increasing their dissemination of information relating to the transition to the EGEE user community. The EGEE dissemination activity will start relying on the emerging NGIs for undertaking these activities.

3.3. TRAINING

As envisaged in the EGI Blueprint there will be no directly funded t-infrastructure (although the formation of a support SSC that includes training is under discussion) and it is expected that NGIs will support their own training by using their own national production resources, which may include middleware other than gLite. During year II, the GILDA t-infrastructure will be merged with the production infrastructure, which in some NGIs may include non-gLite enabled production resources. This will ensure that the students' experiences during training are exactly mirrored when they use the production resources and will allow these resources to have the same professional monitoring and management that the production infrastructure benefits from. These changes will require a review of the training documentation and tutorials to reflect the new infrastructure.

3.4. APPLICATION COMMUNITY

In preparing the scientific discipline clusters (SDCs) for SSCs (Specific Support Centres) EGEE NA4 will move towards the interfaces and communication structures defined in the EGI Blueprint. Each application SSC (currently NA4 application domains) and potential support SSCs (such as NA3 training support, NA2 Business Forum, and the HEP support within SA1) will establish a co-ordinating contact point, support contact point, middleware contact point and operations contact point. These may be different people or a single person depending on the workload contained within these areas by the SSC. These representatives of the 'new' SSCs will form the interfaces into the User Forum Steering Committee (UFSC) and other emerging proto EGI bodies within EGEE – such as for middleware and operations. Each SDC will also explore how they can adopt and benefit from science gateways/portals to improve access to EGEE within their communities using existing developed technologies (e.g. from RESPECT).

3.5. POLICY AND STANDARDS

This activity is seen to continue in EGI as a central funded co-ordinated activity with input from the NGIs and this will be the model we will transition to in year II by encouraging greater contributions from the emerging NGIs.

3.6. OPERATIONS

In addition to establishing managerially equivalent bodies to those introduced in the EGI Blueprint (i.e. the operations unit) the operations structure will alter to reflect the move from central to national based infrastructures. The policy and operational procedures will be reviewed and updated as required to reflect the changes being made in the transition to EGI.

EGI will not provide any central resources for verifying the deployment of new releases. EGEE will shift from the current model of a ‘deployment testbed’ to a staged rollout of new releases onto representative sites from the infrastructure. Improvements in the certification process and the maturing of the software had already reduced the number of issues being found through the deployment testbed.

3.7. NETWORKING

This activity is seen to continue in EGI as part of the networking function within NGIs and this is the model we will transition to in year II. An ENSC (EGI Networking Support centre) will be established by defining the relevant operational tasks that need to take place as global tasks (by EGI) and the contributions that need to come from NGIs to operate this centre. During year II this will take place through the current activity participants in order to verify operational procedures and build experience.

3.8. FORMING PRODUCT TEAMS

To reflect that EGI will have no central software certification capability, EGEE is passing on this responsibility to the individual clusters of competence to form integrated product teams. Each individual product team will be completely responsible for providing certified deliverables. A team will be established (equivalent to their roles in the EGI Middleware Unit) to undertake central tasks – such as verification, identifying areas for further process automation, and the remaining integration and certification work that does not fall into a clear product team (e.g. components from external software providers) or cuts across product teams.

In preparation for EGI the gLite Collaboration will be established. The product teams within the gLite collaboration will work to establish technical and managerial ‘customer’ relationships between the EGEE project and the individual product teams. This will mimic the relationship that will exist between EGI and its software providers.

4. MANAGING CHANGE

4.1. THE TASKS OF NATIONAL GRID INITIATIVES

A National Grid Initiative (NGI) is the body that oversees the national e-infrastructure (computing, storage and other resources) within a country. In many countries, the bulk of the resources within the e-infrastructure will not be owned by the NGI, nor will their use be controlled or managed by the NGI. However, the NGI can provide the uniformity that allows users across the country to work seamlessly across shared resources at all levels: campus, regional and national.

An NGI is expected to be able to financially support the operation of its own user community and its own infrastructure.

The **EGI Global Tasks** (identified in the Appendix by O-E-*, U-E-*, M-E-* and other functions) are defined as being part of EGI.eu, in that they are central tasks (i.e. only one instance is required within the infrastructure) yet it is recognised that they do not need to be generally physically located at EGI.eu. Over the last 5 years, many of these tasks have been initially prototyped within, and are now in production use by, EGEE and its collaborators. As we enter a new phase of European e-infrastructure through EGI it makes sense that the ‘service’ being offered around each EGI Global task is the best the community has to offer. Therefore the hosting of the EGI Global tasks within EGI will be selected from a call within those NGIs that are part of the EGI project.

The **NGI International Tasks** are necessary for the functioning of any NGI, and are the responsibility of the NGI that must support the costs of this task (eligible for EC co-funding) but can delegate the execution of these tasks to another organisation. For example, NGI A wishes to delegate an NGI International Task X to NGI B. The EC contribution for task X has to go to NGI B. NGI A still needs to negotiate matching effort with NGI B.

4.1.1. Selecting EGI Global Tasks

Section 2 of the NGI Bidding Document circulated and accepted by the EGI Council reads:

“A shortlist of interested NGIs for each task will be selected by a board that have no conflict of interest and includes the EGI Director and representatives from the NGIs and the active multi-national user communities. Several boards may be needed for different sets of EGI Global Tasks to eliminate conflicts of interest.”

In line with this decision the proposal editorial team proposed that following the submission of the bids for the EGI Global tasks the selection boards is comprised of:

- EGI Interim Director (Chair)
- Two experts selected among NGI representatives from the EGI Council
- One expert from a High Usage Active User Community: HEP, LS/EBI, CompChem
- One expert from a Low Usage Active User Community: ES, AA, Fusion

Final approval or reduction of the shortlist provided by these boards will be reviewed by the EGI Council.

4.2. REPLACEMENT OF REGIONAL OPERATIONS CENTRES BY NATIONAL GRID INITIATIVES

This chapter summarizes an EGEE III document presenting the status of the Regional Operations Centres (ROCs), their organisation, management and structure, and their plans for the post EGEE era towards EGI/NGIs [9].

ROC	Number of countries	Participating Countries
ASIAN/PACIFIC	6	Taiwan, Japan, South Korea, Australia, India, Pakistan
CENTRAL EUROPEAN	8	Poland, Croatia, Austria, Czech Republic, Slovakia, Hungary, Slovenia, Belarus
CERN	Not applicable	Not applicable
FRENCH	1	France
GERMAN/SWISS	2	Germany, Switzerland
ITALIAN	1	Italy
NORTHERN EUROPEAN	9	Finland, Norway, Sweden, Denmark, The Netherlands, Belgium, Lithuania, Estonia, Latvia
RUSSIAN	1	Russia
SOUTH EASTERN EUROPEAN	8	Greece, Cyprus, Turkey, Serbia, Israel, Bulgaria, FYROM, Romania
SOUTH WESTERN EUROPEAN	2	Spain, Portugal
UK/IRELAND	2	United Kingdom, Ireland

Table 1: Countries distribution per ROC

The third phase of EGEE builds upon the achievements of the previous two phases, providing a production grid serving the needs of the European research community. The EGEE infrastructure consists of a set of services and test-beds, and the management processes and organisations that support them. The ROCs are one of the two layers of operations, support, and management, built on top of the contributing sites that provide the resources. The ROCs exist in each geographical region and support all of the sites in that region, whether those sites are formal partners of EGEE-III or not. In addition they provide front line user support functions and basic service infrastructure of the grid. In the two previous phases of EGEE the concept of the ROCs, which was the key for operational support activities, has been shown to work well. Therefore their structure was retained in EGEE III. In particular the “operator on duty” rotation is an essential part of the core grid operations. There are 11 ROCs (including one in the Asia-Pacific region) listed in Table 1 providing services to a total of 40 countries. The ROCs have been working closely together with their participating sites in order to keep production level quality in a rapidly evolving technology. To that end a series of workflows and tools have been implemented and continuously improved based on experience obtained from their use and additional user input. As EGEE III draws to completion a transition is required towards an operational model based on a set of collaborating NGIs. The challenge is to sustain the established infrastructure in the long term, through its organisational evolution towards EGI and NGIs, which will undertake the tasks today performed by ROCs. Overall, EGEE III expects that by its end operations will have been migrated to a regional model. This can serve as a basis for the next level of migration to the NGIs, adjusting to scalability and effort availability. However this migration will have to finish after EGEE

III. Progress has been made in decentralising key tasks such as the monitoring of the infrastructure functionality, but as most NGIs are still in planning phase, and with the manpower/time constraints, complete migration is not expected to be finished before the end of EGEE III. However the gradual delegation of central operational tasks to ROCs provides useful experience for the next phase, which will be the complete migration to the NGIs.

Each ROC has identified its successes and issues in the operation of its regional infrastructure. Planning for the migration to the EGI model has progressed and needs to take into account regional variations and reality on the ground.

Overall, the NGIs and EGI planning is not yet detailed enough to make a complete transition plan. The transition process will need to continue throughout EGEE-III and will require further steps beyond EGEE-III in order for devolution from ROCS to NGIs to be fully completed.

The interaction with sites and users in EGEE-III participating countries that will not be part of EGI must be addressed if global user communities and nascent regional grid activities are to be supported.

4.3. COORDINATING PROJECTS AND RESOLVING DEPENDENCIES

EU projects are generally funded as part of a call, and as such there can be an expectation that projects will need to depend on each other for maximum return on the EU’s investment. EGEE has recognised such dependencies by developing MoUs (Memorandum of Understanding) with collaborating projects. Within the EGI the situation will be more complex as illustrated in Figure 1.

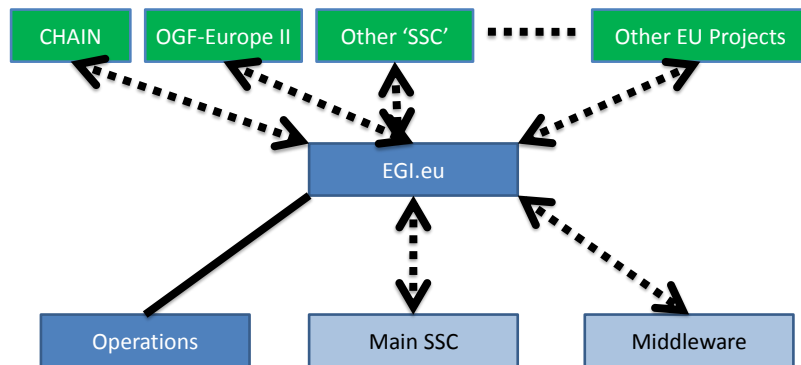


Figure 1: Project relationships

The EGI.eu and Operations activities are envisaged to be a single project. However, both the Middleware and main SSC activities are in separate proposals – yet each project will depend critically on each other. These relationships need to be governed by a protocol (such as a Service Level Agreement, Service Level Description, Memorandum of Understanding, etc.) that defines the expected interaction between collaborating projects. Many of these relationships will be bi-directional.

As part of the transition process to EGI, the protocols governing these relationships need to be provisionally defined. Formal definition of these protocols will be within the remit of the EGI Council.

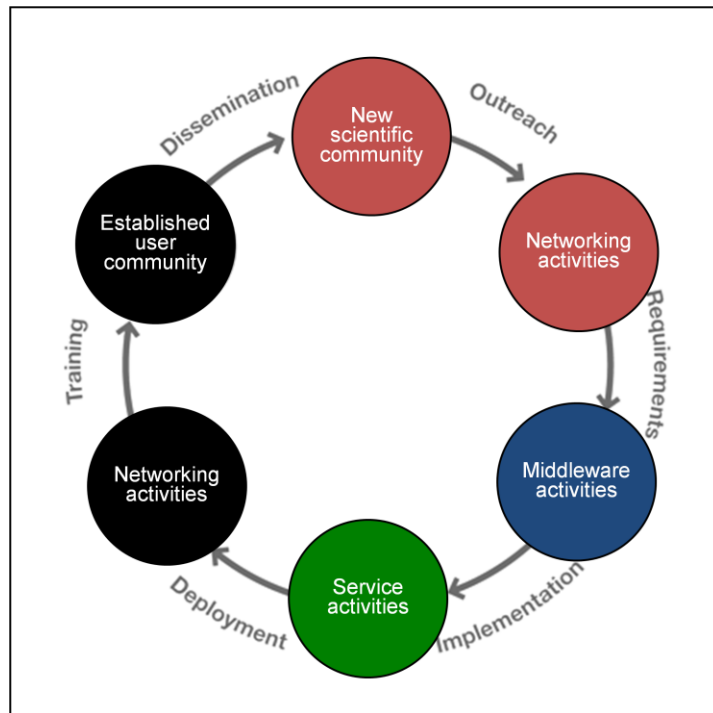


Figure 2: The virtuous cycle

A goal of the coordination of collaborating projects is to ensure that there is full support for the user communities that want to make use of the EGI infrastructure. The virtuous cycle (Figure 2) defined by the EGEE project identifies the necessary phases of the support cycle which must be enacted by the set of collaborating projects.

5. PROPOSALS FOR EGI AND RELATED PROJECTS

5.1. EGI-INSPIRE

The EGI project will be building on top of a series of national, regional, European and international initiatives and projects. The proposed acronym EGI-InSPIRE stands for “Integrated Sustainable Pan-European Infrastructure for Researchers in Europe”.

EGI-InSPIRE will follow the EGI Design Study guidelines as outlined primarily in the EGI Blueprint document and the relevant EGI functions definition deliverable.

The main objective of the EGI-InSPIRE project is to consolidate all the above efforts into a sustainable European grid e-infrastructure, with enhanced and optimised DCI services, serving an extended and increasing user base and user communities, ranging from large organised international user communities to small ad-hoc user groups and individuals. EGI-InSPIRE will offer and receive services to/from a larger set of related infrastructure, middleware, applications and support projects. The main objective can be sub-divided into the following concrete objectives:

- Migrate the existing European grid e-infrastructure and its services into a new sustainable governance model and plan for the future sustainability of the whole EGI construct. This new status will enable Europe to reconfirm its leading position in e-infrastructure research. The governance model of this newly established research infrastructure, as foreseen in the EGI Blueprint, is primarily based on a set of currently 35 NGIs and on the EGI.eu coordinating unit located in the city of Amsterdam, Netherlands. Therefore, “the majority of the (EGI) stakeholders will be NGIs” as stated in the Work Programme;
- Optimise the currently offered grid services based on the sustainable governance model and its structures, extending the geographical and user base of the e-infrastructure, raising its level of services and deploying services for heavy international user communities of EGI;
- Collaborate with middleware and application initiatives and projects, so that the latter can utilise integrated e-infrastructure services. Ensure both interoperability and interoperation with other e-infrastructures within Europe and beyond, especially with those globally required by the European research communities.

Transition towards the new operational structure has started already in 2008, and is being undertaken mainly as part of the EGEE-III and EGI_DS projects. The full implementation of EGI now requires the integration potential and assurance of the EGI-InSPIRE project such as recruiting the personnel and launching at full-speed the operation of EGI.eu, integrating the operation of NGIs possibly in a new regime, instantiating different overview boards not directly controlled by EGI or the EGI-InSPIRE project (such as the middleware or application communities ones), negotiation and signature of Service Level Agreements with NGIs, other resource or middleware providers, e-infrastructure stakeholders (such as the network or supercomputer communities) or related projects.

In parallel, EGI-InSPIRE has to implement appropriate models for its future sustainability, evaluating multiple sources of funding and possibly new business models, liaising with business and industrial stakeholders, acting in a range of different roles. As explained in the EGI Blueprint, EC funding is expected to gradually decrease over time supporting only the innovative activities and the European integration components. Routine and trivial operational tasks will increasingly be supported directly by the auspices of NGIs or by usage based subscriptions from user communities.

5.2. SPECIALIZED SUPPORT CENTERS (SSC)

All of the support and community coordination functions from the NA4 activity of EGEE-III will continue into the EGI era. A series of workshops on forming Specialized Support Centres (SSCs) to

provide both generic and domain-specific services have clarified what is expected of SSCs and how to evolve smoothly from the current model to the EGI model.

Planning for seven EGEE-III scientific communities (Astronomy & Astrophysics, Computational Chemistry, Earth Science, Fusion, Grid Observatory, High-Energy Physics, and Life Sciences) is well advanced with all of those communities participating in proposals for the FP7 Virtual Research Communities call. Two new communities (Photon Science and Humanities) will also participate in those proposals.

Because of competition between the communities, there will be several proposals around SSCs for scientific communities. One will include Astronomy & Astrophysics, Earth Science, Particle Physics and Fusion; called ROSCOE (Robust Scientific Communities for EGI), the other proposal(s) will include all of the other scientific domains.

The situation with the support functions is more complicated because of the division of responsibilities within EGI and because of the FP7 call structure. For new communities (or communities not covered by an SSC), generic support services of EGI will be provided - either directly through the NGIs or through related projects. For all types of support, the scientific SSCs are expected to play a significant role for their own user communities, complementing generic support services from EGI.

Application porting support services will be provided through a dedicated proposal in the Software Simulation and Services call.

Virtual Organization Support and Direct User Support (both operational and functional) will be provided by EGI and included in the core EGI proposal. Services targeted at new communities such as dissemination, training, and business outreach will be provided through a separate project (CUE: Creating Users of E-infrastructures) in the Virtual Research Communities call.

Overall, there will be four SSC-related proposals used to provide the services of the NA4 activity of EGEE-III. Although it is reasonably clear what types of support and which communities are targeted by each proposal, little work has been done to ensure that support procedures can be seamlessly integrated between the various projects and actors. This will come as the transition comes closer and it is clear what projects will move forward.

Organizational changes for the scientific user communities are minimal. The EGEE-III NA4 organization is very similar to that for EGI: a managing body made up of the coordinators of each SSC. In fact, the NA4 Steering Committee will start acting as the equivalent body in EGI, the User Forum Steering Committee. However, the sociological and political changes that are involved for the user communities are significant. They must be robust, independent bodies able to represent their communities within EGI and be directly responsible for reporting, providing support, etc. This transition is more difficult. The more mature communities like High-Energy Physics and Life Sciences will probably be fully-fledged SSCs at the start of EGI. Other communities will make the transition more slowly. However it is likely that all of the current communities will become fully-fledged SSCs soon after the start of EGI.

5.3. MIDDLEWARE

EMI (European Middleware Initiative) is a project proposed in support to the activities of EGI and other major European distributed computing infrastructures. It is designed to be the preferred, although not the exclusive, provider of middleware for EGI and other infrastructures. Its major goal is to maintain, standardize and evolve existing production middleware from various providers (mainly, but not exclusively the three middleware consortia ARC, gLite and UNICORE) at the agreed level of quality. The choice of middleware components to be defined by EGI is based on requests coming from the user communities through the infrastructure organizations.

The EMI project proposes to implement four major functions:

- **User support** providing expert third-level support to the first and second level activities performed by the NGIs and by EGI using the GGUS system as ticketing system

- **Development Processes** including all phases from design to release passing through implementation, building, integration, testing and certification.
- **Quality Assurance** defining and monitoring the software engineering processes with particular attention to testing, certification and release management.
- **Repositories** maintaining packages, tests, certifications and QA metrics generated by the project.

5.3.1. Transition Paths and Issues

Each of the EMI project functions described above requires a transition path from the current state:

- **User support:** the current middleware user support is not formally organized in three distinct escalation levels for all services. User support procedures have to be modified to address the level of individual Product Teams rather than a single middleware support unit incorporating existing support mechanisms for ARC, gLite, UNICORE.
- **EMI expects to use GGUS to interact with the users.**
- **Quality Assurance:** EMI is planning to adopt a very standard QA process based on existing ISO software engineering standards.
- **Repositories:** EMI will provide a single reference repository for all development phases from implementation, to integration, internal testing, certification and public releases.

5.3.2. Relationships with other projects or initiatives

EMI is expected to be a major middleware provider for EGI and other distributed computing infrastructures like PRACE. The relationship with EGI is mainly focused on providing middleware for the UMD distribution and a robust support service. In particular the following activities are foreseen:

- **UMD (Unified Middleware Distribution):** UMD is the middleware distribution defined and managed by EGI as the recommended set of services and components to be used by EGI sites. UMD is composed of many different services not developed by EGI directly, but taken from collaborating middleware providers. EMI will be the major source of components for UMD. The interaction is mainly with the EGI Middleware Unit responsible to define content and scope of the UMD releases.
- **Requirement analysis:** requirements for EMI services must come from the user communities directly or through the filter of the DCIs. The EGI Middleware Collaboration Board (MCB) will collect and analyse requirements by forming a group of experts, users, application developers, SSCs, middleware providers like EMI and EGI personnel from operations, the Middleware Unit, etc.
- **Third-Level Support:** EMI is expected to provide the third-level support collaborating with the EGI work package providing second-level support, debugging and root-cause analysis.

5.4. RELATED PROPOSALS

This chapter covers some other proposals that aim to take up functions that have been essential for the extension of grids for worldwide communities and for training, dissemination, and business outreach

5.4.1. Extra-European Collaboration – the CHAIN Proposal

A project called CHAIN (Co-ordination and Harmonisation of Advanced e-Infrastructures) is being prepared to foster collaboration with the HPC community as well as with extra-European Grids e.g. in Africa, Asia and Latin America.

It is necessary to coordinate and harmonise all the projects aiming at extending the European e-infrastructures to other regions. These projects are in their second or third phase and they are reaching the limits of what can reasonably be done in specific region-oriented projects. It is thus time to elaborate a strategy, define the instruments and deploy them for the coordination and interoperation of the European grid infrastructures with other external e-infrastructures. Furthermore it is necessary to analyse the issues related to the coordination of "Regional Extensions" to European e-infrastructures which include not only Grids but HPC and, possibly, networking.

5.4.2. Training, Dissemination, Business Outreach – the CUE Proposal

The CUE (Creating Users of E-infrastructures) proposal sets out to enable an increasing number of users and research communities from all science and engineering disciplines to access and use e-infrastructures; and remove the constraints of distance, access and usability as well as the barriers between disciplines for a more effective scientific collaboration and innovation

The CUE project proposes to build on the success of the EGEE and ICEAGE projects and its strategic aims are to:

- Reinforce the impact of e-infrastructures both in scientific and humanities research defining and delivering a stimulating programme of dissemination and educational events, both real and virtual;
- Develop novel ways of presenting e-infrastructures to prospective users and communities;
- Enlarge the number of users of the European grid infrastructure;
- Transfer grid technology with its benefits to the non-scientific public and business private sectors.

5.4.3. Dissemination channels for users, developers and policy makers – the GridTalk-II Proposal

Continued support for key dissemination channels is foreseen within the GridTalk-II proposal aiming to build on the significant achievements of the GridTalk project in bringing the success stories of Europe's e-infrastructure to policy makers in government and business, to the broader scientific community and to the general public.

GridTalk-II will work to establish close contacts with the new EGI structures such as the National Grid Initiatives, the Specialised Support Centres and the European Middleware Initiative

GridTalk-II's output will consist of four principle products:

1. The online weekly publication, International Science This Week will communicate directly with the existing and potential grid and HPC user community,
2. the GridCafe, GridCast and GridGuide suite of interactive,
3. the Real Time Monitor of GridPP, providing live views of grid activity and
4. a series of reports and materials aimed at policy makers.

6. OTHER TRANSITION ISSUES

All the tasks listed in the Appendix have specific issues related to the transition to EGI. They are defined in the function definition document [7] and will be further detailed in the proposals. We address here elected items not discussed in the previous chapters that either are of particular importance to the user community or that otherwise require particular attention.

6.1. OPERATIONS

Operations constitute central tasks essential for a reliable functioning of the grid:

Operations Coordination consisting of: the ownership of the standards that bind the NGIs together; the monitoring of the performance of the NGIs for International VOs; oversight of the global services with which the NGIs interact; and management of the issues arising from all of the above.

User Support Coordination providing maintenance and operation of the central ticketing system and of the triage of incoming problems.

Operational Security Coordination covering the development and maintenance of security policy (O-E-15) and the coordination of security and incident response (O-E-16) within the NGIs (O-N-9a).

Operational Services - the global services run by specific NGIs on behalf of the whole project must be closely integrated with the work of the NGIs that interact with them.

6.1.1. Operations Coordination

There are several levels of coordination required with different timescales.

- a) Short-term – to deal with operational exceptions as seen by VOs. In theory all issues should get resolved by the support/helpdesk processes but we see that a meeting (real or virtual) provides good information flow and guarantees appropriate escalation. Examples include the WLCG [10] Daily Operations Meeting and the EGEE Weekly Operations Meeting.
- b) Oversight of the Sites – a hierarchy of teams monitoring the sites, raising tickets, providing support. A continuous process for continuous improvement of site and service reliability.
- c) SLA Review. Consideration of how sites and NGIs are performing in the medium to long term. Takes input from monitoring and from VOs
- d) Individual service advisory groups. A meeting between service providers and stakeholders to review the performance over an extended period and propose changes to requirements, service levels, procedures, etc.
- e) Development plans for Operations Tools. Takes input from (d) but has an integrated view of how the tools interwork and share infrastructure.
- f) Liaison of ‘operations’ with other activities.

For each of the above we need to define the terms of reference as well as the membership and reporting lines.

6.1.1.1. Operation of the Grid topology and configuration repositories

Many aspects of operations rely on the availability of information from NGIs about service nodes, contact details, security contacts, certification status, sites in scheduled downtime, etc. The grid repository provides all such information. Information input is devolved to regions and sites. The current central repository (known as GOCDB in EGEE) may need to be adapted to support a two-tier distributed model. This requires the definition and implementation of common interfaces and transport mechanisms to ensure the exchange of information between different Grid domains.

By the end of EGEE III, a new version of the current repository GOCDB will have been deployed to fill this role. It will run a central database containing a sub-database per EGEE region. These will be managed by the regions.

EGI is responsible for running the central service where other services can find information and maintaining the schema of the centrally-held information. This may be the master copy for an NGI, or a cached copy if the NGI runs its own service.

An NGI is responsible for maintaining data about itself and its sites. If it runs its own repository, it is responsible for publishing the agreed schema to the central database.

Milestone¹ 1: Plan for further distribution. Participating NGIs to decide how they manage their data. Do they wish their own instance; use a regional one, or a national/regional partition of the central database, managed by the region or NGI.

Milestone: 2: Implement Plan. Beyond this milestone, changes will be an operational issue negotiable between the NGI and the global service provider.

6.1.1.2. Operation of accounting repositories for international VOs

The accounting repository is responsible for keeping records about usage of compute, storage, networking and other types of resources as required by the users, resource providers, NGIs, etc. It is the responsibility of an NGI to collect accounting data, and to keep a permanent master copy of usage records. Accounting information is needed by international VOs in order to allow VO managers to know about the amount of IT resources “consumed” by the respective users across different domains of the e-infrastructure.

By the end of EGEE III, a new version of the APEL client will have been deployed using the ActiveMQ Message Bus. The NGIs that are already publishing from their own accounting solution in EGEE III will also have ported to this new transport layer. A version of the APEL Central Repository capable of being run independently by a NGI and publishing data into the central repository will have been developed and tested but it is not assumed that this has been deployed in production. The central APEL Repository that currently holds data on LHC and EGEE VOs can act as the default repository for international VOs.

EGI is responsible for running the central service where VOs can interrogate their usage, and maintaining the schema of the centrally-held information.

An NGI is responsible for maintaining data about all usage made of its sites and for publishing data on international VOs to the appropriate repository.

Milestone 3: Plan for further distribution. Participating NGIs to decide how they manage their data, whether they wish to maintain their own instance, use a regional one, or to continue to publish accounting data directly to the central service.

Milestone: 4: Implement Plan. Beyond this milestone, changes will be an operational issue, negotiable between the NGI and the global service provider.

6.1.1.3. Operation of Grid repositories storing monitoring and performance data and other related information

Availability, status and performance information about grid services and sites are needed to check the health of the infrastructure and to verify the quality of service delivered to VOs and other NGIs. Gathering and publication of monitoring information – regarding grid functionality, grid service status, assessment of quality of the services delivered by various EGI actors (resource providers, the NGIs, etc.) – is consequently important to help the infrastructure to assess its level of service and compare it to the VO requirements. This also requires the operation of repositories and

¹ The milestones given here and in the following sections will be described with more detail and in particular with dates in the EGI proposal

supervision of the processes to populate them, the maintenance of schema for publishing of site and service status information, the ownership of the information schema used, the preparation of reports, etc.

It is assumed that all EGEE Regions and some NGIs will already be running their own Nagios monitoring infrastructure, and metric store. The data will also be published to a central metric store from which EGI-wide reporting will be possible. SAM tests will no longer be run centrally from CERN and tools that interrogated the SAM database will have been ported to use the metric store. It is also assumed that existing ROCs will continue to run monitoring for the relevant NGIs until at least Milestone 6 has been passed.

The EGI responsibilities are the publication of statistics, the maintenance of schema for central publication of site and service status information, the deployment of monitoring-related tools such as the dashboard and the alarm system, and the preparation of performance reports.

The NGI responsibilities are to monitor their sites, make the information available through a local portal and publish information to a central metric store.

Milestone 5: Plan to devolve Nagios testing from regions to countries where required.

Milestone 6: implement plan.

6.1.1.4. Operation of the Grid Operations Portals

Grid operations portals provide an entry point for various actors to support their operational needs. Different "views" are necessary according to the role of the customer (grid operators, VOs, grid site managers, Region Operations Managers, etc.). The information displayed is retrieved from several distributed sources (databases, grid information systems, etc). It provides static information about sites/VOs, and dynamic information about resources/services status and allocation. The central Operations portal is the aggregation point of regional information, which is also accessible via regional operations portals.

By the end of EGEE III a version of the CIC portal will be maintained centrally with the management distributed to regions and VOs. The first milestone will be to increase the granularity to NGIs to allow NGIs to maintain their information in the central instance. NGIs will also be able to interface a national repository of this information with the central repository.

The EGI responsibility will be to run the central portal and automate gathering of the required information

The NGI responsibility will be to maintain relevant information about their infrastructure and VOs.

Milestone 7: Delegate responsibility to NGIs where and when appropriate

6.1.1.5. Grid operation and oversight of the e-infrastructure

At the core of this task is the monitoring and support of the sites through a ticketing system that will contain issues raised as a result of SLA checking (i.e. SAM monitoring tickets or equivalent), VOs, users and system administrators. This ensures that the International VOs get the services agreed through EGI. The practical work needed includes:

- Monitoring and help desk shifts
- Triage of incoming problems, assignment of tickets to the 2nd line support units, ticket escalation to EGI.eu, ticket follow-up, suspension of sites if needed, etc.
- Certification of the sites entering the NGI grid and thus in the EGI grid, with the rules agreed with EGI, according to the site category and SLA.
- The interface with the NREN is specifically required for troubleshooting of connectivity problems, test for advancement in technologies etc.
- Maintenance of web pages for FAQ, best practices etc.
- Operation of a ticketing system integrated with the global EGI.eu ticketing system

The bulk of the work is done by NGIs using their own monitoring infrastructure for their own sites and users but there is a central role in coordination so that the international VOs receive the same level of service and cross-border issues are resolved. This is covered by O-N-3.

It is assumed that an NGI will be operating a ticketing system to coordinate its own national activities. Effort is needed to interface this national system into the EGI.eu system so that tickets can be exchanged and easily escalated. Interfacing the national help desk into the central help desk is the responsibility of the NGI, which will be supported by the central help desk staff.

6.1.2. Central and regional Grid user support and ticketing system

User support relies on a central helpdesk, which is a regional support system with central coordination (GGUS). It gives access to user documentation and support, and to a ticketing system. The central system is interfaced to a variety of other ticketing systems at the NGI level in order to allow a bi-directional exchange of tickets (for example, those opened locally can be passed to the central instance or other areas, while user and operational problem tickets can be opened centrally and subsequently routed to the NGI local support infrastructures).

Support to network end-to-end problems in the grid is also important – especially for applications requiring high-availability – as connectivity is provided by the pan-European network research backbone and by a large number of National Research and Education Networks, each providing links to sites within countries. A Network Operation Centre provides the operational interface between the grid and the relevant network players to check the end to end connectivity of grid sites (the ENOC).

The NGIs provide 1st line local/regional support to users and centres, while EGI.eu takes care of the Maintenance and Operation of the central ticketing system (GGUS like) and of the triage of incoming problems.

- a. Maintenance and Operation: run a central ticket handling system for grid and network end-to-end problems. User support relies on a central helpdesk, which is a regional support system with central coordination (GGUS). It gives access to user documentation and support, and to a problem ticketing system.
- b. Triage of tickets entering the central user support system (also known as ticket processing management in EGEE) consists of the monitoring and routing of all active tickets in the grid user support system by grid and VO experts, who are responsible of addressing the problems to the appropriate second-line specialized support units. This process combines manual as well as automated procedures.

6.1.3. Coordination of interoperations between NGIs and other Grids

Coordination is needed to foster the creation of a seamless operations model across administrative boundaries, in order to pursue pervasiveness and sustainability of the infrastructure. This is of great importance as users who want to cross grid boundaries need to know that the environments will be similar, and applications must function properly without major changes. Interoperation covers a number of aspects, such as the availability of common tests for monitoring of site status, the interconnection between helpdesks/ticketing systems, etc. “Other Grids” includes Asia-Pacific regional Grids, OSG [8], Naregi, and related infrastructure projects.

The NGIs and other national grid infrastructures have deployed different sets of middleware. The required interoperation involves research and development work to solve the problems encountered. EGEE has achieved interoperation with a large number of grid infrastructures. Besides the infrastructures at the international level, like OSG and NDGF [11], the SA1 partners have started to focus on interoperation with national grids, especially the National Grid Initiatives, for supporting the transition from EGEE to EGI. The resulting combined frameworks support more users and applications than would otherwise be possible and allow a significantly enhanced level of collaboration both between researchers and between EU e-infrastructure projects.

This role owns the definition of the operational tools interfaces, the procedures and the operational activities allowing the NGIs to interoperate. EGI aims at continuing the collaboration established with operations centres outside Europe in order to preserve the current integration of non-European sites into the production infrastructure. EGI.eu is responsible for support and coordination.

6.1.4. Operational Security Coordination

There are many important challenges to be addressed in the area of computer and network security for EGI. The collaboration of a large number of independent sites in one e-infrastructure potentially amplifies the security problems. Not only do these infrastructures contain large computing and data storage resources connected by high-speed networks, these being very attractive to potential attackers, but the connectivity and ease of use of the services means that a successful compromise of one site can threaten the e-infrastructure in general and all of the participating sites.

The aim of EGI operational security is to address the various risks involved and to maintain the availability of EGI services. In spite of all best efforts, vulnerabilities will be found in the software deployed and security incidents will happen. All of these must be promptly and efficiently handled.

Agreed common security policies are required to regulate the activities of all participants: NGIs, resource providers and user communities alike. These common policies will facilitate the use of resources by user communities in any NGI willing to support them and also encourage interoperability between EGI, the NGIs and other infrastructures elsewhere.

A common authentication trust domain is required to persistently identify all grid participants. To ensure interoperability, both at the European as well as the global scale, EGI will participate in and collaborate with the International Grid Trust Federation (IGTF), and the EUGridPMA in particular, in line with the relevant e-IRG recommendations.

The **EGI Computer Security and Incident Response Team (CSIRT)** is an activity aimed at coordinating the operational security activities in the infrastructure, in particular the response to security incidents which may affect any resource providers, inside or outside EGI.

The large number of resource providers and the infrastructure topology requires a structured approach to be adopted, based on the different levels. From bottom to top:

- Site CSIRT (hundreds, usually one per site)
- NGI CSIRT, NREN CSIRT (tens, the degree of interaction between NGI and NREN CSIRT varies)
- Main academic grid CSIRT (units)

As a result, the EGI CSIRT ensures both the coordination with peer grids (via its GRID-SEC membership), and with the NGIs and NREN CSIRTs as well as grid security monitoring, security training and dissemination and incident response improvement (e.g. security drills).

The activities and mandate of the EGI CSIRT are regulated by the EGI security policies, developed and maintained by the JSPG and adopted by EGI. Additional security procedures will be implemented and agreed by EGI via the appropriate approval process.

Each NGI appoints a "NGI Security Officer" (post O-N-9a) in order to provide the NGI CSIRT functions.

A number of NGIs are already implementing this functionality via their local NREN CSIRT. Where this is not the case, a close collaboration with the local NREN CSIRT is highly encouraged.

The resulting group of NGI Security Officers collaborate as part of the EGI CSIRT.

The EGI CSIRT is led and coordinated by the EGI Security Officer (post O-E-16), whose role and mission are defined by security policies approved by EGI and the NGIs.

The **Security Policy Group (SPG)** is responsible for the development and maintenance of security policies and for providing advice on any security policy issue. This will be led and coordinated by EGI (post O-E-15). The SPG’s primary stake-holders will be internal to EGI, i.e. the NGIs, the sites and the application communities. Building on the earlier work within the Joint (EGEE/WLCG) Security Policy Group, the SPG will continue to aim for common simple policies for interoperation across the world. Participation in the SPG by policy experts from other e-infrastructures will therefore still be encouraged.

In EGI it will be important to ensure that the grid software and deployment is as secure and free from vulnerabilities as possible. The main purpose of the **Software Vulnerabilities Group** is to eliminate existing vulnerabilities from the deployed infrastructure, primarily from the grid middleware, and prevent the introduction of new ones. The aim is to prevent grid security incidents.

EGEE-III had a Security Coordination Group, consisting of the chairs of each of the sub-groups, which met monthly by phone. It is advisable to continue this in EGI.

6.2. TRAINING

The training activity has by its nature been designed in a federated structure as seen in Figure 3 and Figure 4, allowing trainers in each country to design their training events according to local requirements.



Figure 3: Google map of EGEE-III training events in Europe

Within the EGI.eu funding effort there are very limited resources set aside for training, providing only a coordinating role for training across the NGIs, for creating/reviewing training material, and for training support services. Therefore the training will be largely conducted by the NGIs themselves, so in the second year of EGEE-III, the focus will be ensuring that the NGIs have everything they need to be completely self-sufficient with regards to training.

With respect to this some changes to the training work plan are necessary:

- For EGI facilities have to be in place following EGEE-III to ensure that training events can be run using appropriate resources as part of the production infrastructure.

- e-Learning can be considered as an aid to scalability, and coverage where full training resources are not available.
- The further development of trainers will be pushed forward, particularly in those countries where no accredited trainers currently exist.
- Countries where there are no current training partners will be approached to provide assistance in preparation for the move to EGI.
- As the EGEE partners move to the NGI model, the training required may alter, to more closely match the requirements specific to their own infrastructure, e.g., more diverse middleware implementations.
- Integrate the resources within the current GILDA infrastructure into the production infrastructure using the existing GILDA training VO that already exists on the training infrastructure. NGIs can of course enable their national training VOs on their resources. Continue to run central services (i.e. CA, BDII, GridWay, etc.) to access GILDA VO enabled production resources.

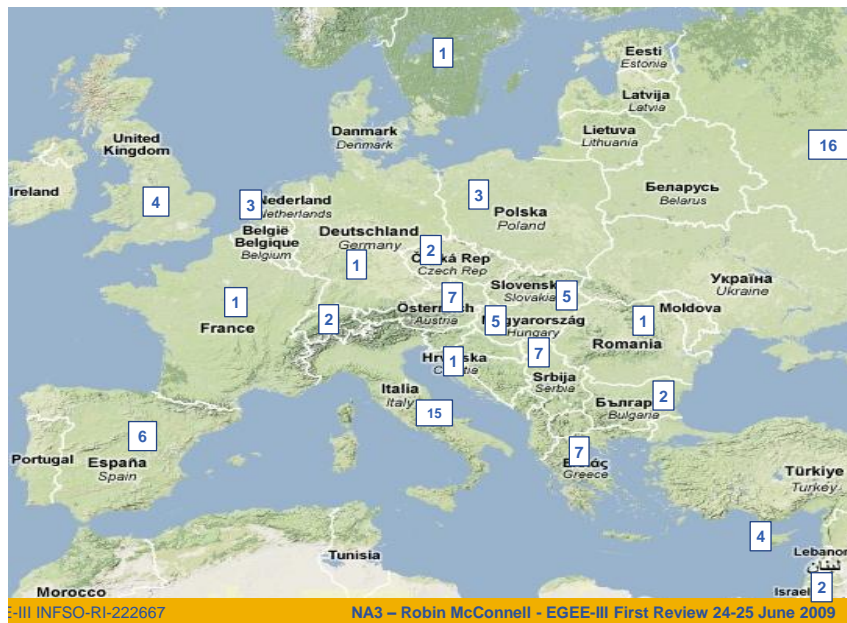


Figure 4: Location and number of accredited trainers

6.3. USER SUPPORT

In EGEE the user support was organized in seven scientific clusters within the NA4 activity:

- **Astronomy & Astrophysics:** This cluster has identified new users and ensured their training. It is notable that most of these users receive no funding and continue to use the grid because it provides a valuable service;
- **Computational Chemistry:** This cluster is using licensed software on the EGEE grid infrastructure implementing different models in production and providing the best real world experience with the benefits and problems of each model;
- **Earth Science:** The dissemination activities undertaken by this cluster have greatly increased the visibility of grid technology within the earth science community, showing directly that

grid technology can produce exciting scientific results. This recognition has allowed the Earth Science cluster to become the core of the nascent Earth Science SSC for EGI;

- **Fusion:** This cluster has ported a number of fusion applications to the EGEE grid infrastructure, showing the utility of grid computing for fusion research. Application requirements force this cluster to consider supercomputing, grid, and desktop resources, putting this cluster on the leading edge of transparently combining these different resources;
- **Grid Observatory:** This cluster has collected information about the use and performance of the various production grid services for the computer science community for analysis, creating a bridge between EGEE and the computer science community;
- **High Energy Physics:** The most important generic tools have originated in this cluster: AMGA metadata catalogue, Ganga/DIANE job management system, and the Experiment Dashboard. Although these were all developed according to the needs of the high-energy physics community, they have been useful to many other communities. A good example is the exploitation of the above tools by the Fusion community, which was also demonstrated during the 4th EGEE User Forum. Thus, the HEP cluster is a good example of cross-discipline collaboration;
- **Life Sciences:** This cluster spanning a wide range of different disciplines and applications requirements has provided the Medical Data Manager and MOTEUR among others. At the same time, users in this area have demonstrated significant scientific impact and are among the most important success stories of EGEE (e.g. the WISDOM application, the Identification of Causes for Coronary Artery Diseases etc).

The support of these communities will continue in the form of Specialized Support Centers (SSCs) as described in Chapter 5.2.

6.3.1. High Energy Physics

In this document we cannot possibly address transition issues for all communities using the grid in Europe. In this section we give the example of High Energy Physics (HEP) and in particular the Worldwide LHC Computing Grid (WLCG) [10] project because it represents the largest community of the EGEE infrastructure, both in terms of number of users and resource utilisation. In addition, tools and methods developed in the HEP community have been adopted by other grid user communities.

WLCG has been running in continuous production mode for several years and has recently carried out a Scale Testing for the Experiment Programme'09 (STEP'09) exercise. As part of this programme the LHC experiments ran over 2 million analysis jobs and generated over 6 GB/s of sustained grid traffic. The exercise was prepared with the LHC experiments during a period of about two months and executed in two intensive weeks of 24x7 operations.

It is the community that often performs (pre-) production testing of new services and / or releases – benefiting not only WLCG (and HEP) but also many other grid communities.

The complex computing models of the LHC experiments have implications on the operations procedures and grid support infrastructures. The policies and procedures established for the HEP communities – e.g. within WLCG – have frequently been adopted as part of the standard EGEE procedures.

WLCG can therefore be considered as a reference for other scientific communities. Collaborative activities between WLCG and other communities have been supported during the past years through co-funding from a variety of resources, including national grid projects and funding agencies, international projects supported through the EU and other agencies, as well as specific institutes such

as CERN, which have allowed a relatively small “grid support” team to empower a number of (very) large distributed communities.

A smooth transition from EGEE to EGI is essential for the WLCG community. A transition that satisfies this community can also be successfully applied to others. Given the schedule for the restart of the LHC in late 2009, with the data taking run continuing more or less uninterrupted until late 2010, the transition must start prior to any FP7 and/or related funding being available.

In terms of WLCG application support, there is a well-established and trusted infrastructure that should be kept largely as it is during the transition phase to the SSC infrastructure (although in terms of manpower it is already going down as people on short term contracts find employment elsewhere). A specialised support team called EIS (Experiment and Infrastructure Support) placed at CERN works daily with the experiments in order to ensure a good implementation of the grid infrastructure into their computing models. The migration of the functions of this team to the future SSC as the support infrastructure of the WLCG experiments is crucial and supported by the communities. Currently EIS is focused on gLite middleware support, offering only limited assistance to WLCG sites that use ARC middleware – this can be improved in future. The principal roles of this team are the following:

- Liaison between the WLCG community and services, operations, middleware teams and the sites;
- EIS provides grid expertise for the middleware implementation into the application computing model. This might include the development of specific application software towards the final implementation and the creation of additional components that might later be shared by other communities (as has happened in the past). In addition this team is responsible for the evaluation of new middleware functionality and to perform scalability tests based on the needs of the community;
- The EIS team manages also VO-specific services and provides user-support as required (e.g. submitting and following up on GGUS tickets on behalf of the VO);
- This team has the expertise to solve grid related issues for running applications both to the site and middleware level. The members have a deep knowledge of both grid infrastructure and WLCG computing models;
- They are responsible for application-oriented documentation and training;
- Using this knowledge they also create high level grid implementation tools;
- EIS is responsible for experiment-specific tests executed within the overall monitoring framework (SAM);
- The team also provides links to operations activities, such as creation and registration of new VOs, VO manager activities, negotiation of resources and services at the sites, etc.

Grid technology is seen as a major compute paradigm for the LHC data analysis. Nevertheless grid technology is heavily in use in other HEP activities such as the data analysis of the HERA experiments, the detector design for the future Linear Collider ILC, the theory studies and astroparticle physics like Icecube. For these tasks the major centres like DESY and INFN host a full set of grid services based upon EGEE middleware components.

APPENDIX: SUMMARY OF TASKS AND EFFORT

This chapter summarizes the information of EGI functions, owner of the function (EGI.eu, NGI, middleware consortia, SSC), and possible providers for the service (e.g. is the function provided for EGI by the personnel located in EGI headquarters or can it be delivered by the NGIs).

Table 2: Summary of EGI tasks and related estimated effort

Function			
Code	Description	Owner of the function	Effort in FTE
OPERATIONS AND SECURITY			
Operation of tools and Services			
O-E-1	Operation of the grid topology and configuration repositories	EGI.eu	1
O-E-2	Operation of accounting repositories for international VOs	EGI.eu	1
O-E-3	Operation of the grid repositories storing monitoring and performance data, and other related information	EGI.eu	2.5
O-E-4	Operation of the grid operations portals	EGI.eu	0.5
O-E-5	Grid operation and oversight of the e-infrastructure	EGI.eu	1
Support			
O-E-6	Central and regional Grid support, operation of a ticketing system and of a document repository	EGI.eu	2
O-E-7	Triage of incoming problems	EGI.eu	2
O-E-8	Gathering of requirements for user support tools and processes	EGI.eu	0.5
Other Tasks			
O-E-9	Coordination of middleware roll-out and deployment, middleware pilot and certification testbeds	EGI.eu	1
O-E-10	Coordination of resource allocation and of brokering support for VOs from NGIs	EGI.eu	0.5
O-E-11	Coordination of interoperations between NGIs and other grids. EGI.eu	EGI.eu	0.5
O-E-12	Coordination of network support	EGI.eu	0.5
O-E-13	Coordination of definition of best practices, operations procedures, operations requirements	EGI.eu	0.5
O-E-14	Operation of production grid core software services, catch-all services for international VOs, catch-all CA	EGI.eu	1
Security			
O-E-15	Coordination of security policy development and maintenance	EGI.eu	0.5
O-E-16	Coordination of security and incident response	EGI.eu	1

Development			
O-E-17	Coordination of development and maintenance of operational tools	EGI.eu	1
Operations			
O-N-1	Operation of the NGI grid topology and configuration repositories	NGIs	a)
O-N-2	Operation of the NGI accounting repository	NGIs	a)
O-N-3	Operation of repositories storing monitoring and performance data, and other related information	NGIs	a)
O-N-4	Operation of the NGI operations portal	NGIs	a)
O-N-5	NGI e-infrastructure oversight. 1 st and 2 nd line support.	NGIs	a)
O-N-9	Operations coordination at NGI level	NGIs	a)
Support			
O-N-6	Operation of the NGI ticketing system, gathering of new requirements for user support tools in the region	NGIs	a)
O-N-7	Regional helpdesk	NGIs	a)
Other Tasks			
O-N-8	Operation of production grid core software services	NGIs	a)
O-N-9	Operations coordination at the NGI level	NGIs	a)
Middleware development and support			
Middleware Unit of EGI.eu			
M-E-1	Maintain and document processes and quality criteria common to all middleware providers.	EGI.eu	1
M-E-2	Provide and support tools to enable and monitor the processes	EGI.eu	1
M-E-3	Define quality and conformance criteria for UMD components. Verify that accepted components are certified according to the agreed process and satisfy the quality and conformance criteria	EGI.eu	3
M-E-4	Maintain a repository of certified middleware components or references thereto.	EGI.eu	2
M-E-5	Follow the daily execution of the strategic plan endorsed by the MCB. Promote the EGI participation in standardisation bodies.	EGI.eu	1
Middleware maintenance			
M-C-6	Maintenance and support of security middleware components	mw consortia	10 ^{b)}
M-C-7	Maintenance and support of data management middleware components	mw consortia	15 ^{b)}

M-C-8	Maintenance and support of job management middleware components	mw consortia	24 ^{b)}
M-C-9	Maintenance and support of information management middleware components	mw consortia	10 ^{b)}
M-C-10	Other middleware maintenance related activity	mw consortia	11 ^{b)}
USER COMMUNITY SERVICES			
Representation and Networking (User Forum)			
U-N-1	User Forum Representation	NGIs / SSCs	2-5
U-N-2	General Coordination	NGIs / SSCs	2-7
U-N-3	Feedback	NGIs / SSCs	3-6
U-N-4	User Conference and other Events	NGIs / SSCs	2-5
U-N-5	Public Relations	NGIs / SSCs	3-6
Grid Planning and MCB			
U-N-6	Development of Services	NGIs / SSCs	c)
U-N-7	Coordination of Grid Planning	NGIs / SSCs	3-6
U-N-8	Technical Feedback	NGIs / SSCs	3-6
User Technical Support			
U-N-9	VO Registration and VO Database	NGIs / SSCs	2-3
U-N-10	Site Validation Tests	NGIs / SSCs	1-3
U-N-11	Core VO Service Provision	NGIs / SSCs	2-4
U-N-12	Help Desk and Direct User Support	NGIs / SSCs	2-4
Informational and High Level Tools			
U-N-13	Case Studies	NGIs / SSCs	3-6
U-N-14	Application Database	NGIs / SSCs	3-6
U-N-15	Science Gateways / Portals	NGIs / SSCs	7-15
Documentation and Training Coordination			
U-N-16	Documentation	NGIs / SSCs	2-4
U-N-17	Training Coordination	NGIs / SSCs	2.5-5
Front Desk / Services for New Communities			
U-N-18	Consulting for New Communities	NGIs / SSCs	4-8
U-N-19	Assistance for Application Porting	NGIs / SSCs	8-12
Other NGI International Tasks		NGIs	Tbd
UCS Team in EGI.eu			
U-E-1	User Forum Support	EGI.eu	1
U-E-2	Coordination of SSC activities	EGI.eu	2
U-E-3	Event Organisation	EGI.eu	1

U-E-4	Grid Planning & MCB	EGI.eu	1
U-E-5	Technical Coordination – User Technical Support	EGI.eu	1
U-E-6	Technical Coordination – Informational and and High-level tools	EGI.eu	1
U-E-7	Documentation Review Coordination	EGI.eu	1
U-E-8	Coordination of Training	EGI.eu	1
U-E-9	Services for new communities & Front Desk coordination	EGI.eu	2
External LIAISON Functions			
E-E-1	EGI.eu Dissemination team	EGI.eu	2
E-N-2	Dissemination in regional / NGI level	NGIs	0,5 per NGI
E-E-3	Industry Take-up	EGI.eu	^{d)}
E-E-3	Other External relations	EGI.eu	2
EGI management			
D-E-1	Director	EGI.eu	1
D-E-2	Assistant of Director and secretaries	EGI.eu	3
D-E-3	Chief Technical Officer, CTO	EGI.eu	1
D-E-4	Chief Operational Officer, COO	EGI.eu	1
D-E-5	User Coordination Officer, UCO	EGI.eu	1
D-E-6	Chief Administrative Officer, CAO and staff	EGI.eu	4

a) Amount of resources needed altogether for tasks O-N-1 – O-N-9 depends on the size of the NGI. Small NGI: 2-4 FTE; Medium NGI:5-10 FTE; Large NGI:14-22 FTE.

^{b)}See table in Chapter 7.1.5. for detailed break down between middlewares

^{c)} Effort not budgeted in EGI Blueprint

^{d)} EGI.eu director is assumed to take care of these functions

BIBLIOGRAPHY

- [1] **EGEE**. EGEE web page. [Online] <http://www.eu-egee.org/>.
- [2] **D5.3**. *EGI Blueprint*.
- [3] Research Infrastructures Calls: FP7-INFRASTRUCTURES-2010-2. [Online] European Commission.
http://cordis.europa.eu/fp7/dc/index.cfm?fuseaction=UserSite.CapacitiesDetailsCallPage&call_id=263#infopack.
- [4] EGI_DS. [Online] <http://web.eu-egi.eu/>.
- [5] **D5.5**. *The EGI Memorandum of Understanding (MoU)*.
- [6] **D3.4**. *Feedback on EGI Functions Definition*.
- [7] **D3.2**. *Final EGI Function Definition*.
- [8] Open Science Grid. [Online] <http://www.opensciencegrid.org/>.
- [9] *Report on the Status of ROCs and National/Regional Grid Integration*. s.l. : EGEE III, 2009. DSA1-1.3.
- [10] Worldwide LHC Computing Grid. [Online] <http://lcg.web.cern.ch/LCG/>.
- [11] Nordic DataGrid Facility. [Online] <http://www.ndgf.org/ndgfweb/home.html>.
- [12] PRACE. *PRACE*. [Online] <http://www.prace-project.eu/>.
- [13] DEISA. *DEISA*. [Online] <http://www.deisa.eu/>.

Copyright notice and limitation of liability:

Copyright © Members of the EGI_DS Collaboration, 2008.

See www.eu-EGI.eu for details on the copyright holders.

EGI_DS (“European Grid Initiative Design Study”) is a project co-funded by the European Commission as a Coordination and Support Action within the 7th Framework Programme. EGI_DS began in September 2007 and will run for 27 months.

For more information on EGI_DS, its partners and contributors please see www.eu-EGI.eu

You are permitted to copy and distribute, for non-profit purposes, verbatim copies of this document containing this copyright notice. This includes the right to copy this document in whole or in part, but without modification, into other documents if you attach the following reference to the copied elements: “Copyright © Members of the EGI_DS Collaboration 2008. See <http://web.eu-egi.eu/> for details”.

Using this document in a way and/or for purposes not foreseen in the paragraph above requires the prior written permission of the copyright holders.

The information contained in this document represents the views of the copyright holders as of the date such views are published.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE MEMBERS OF THE EGI_DS COLLABORATION, INCLUDING THE COPYRIGHT HOLDERS, OR THE EUROPEAN COMMISSION BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THE INFORMATION CONTAINED IN THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Delivery Slip

	Name	Partner/ Activity	Date	Signature
From	Jürgen Knobloch on behalf of WP5	CERN/WP5	31 August 2009	
Reviewed by	Bob Jones Panos Louridas	CERN/EGEE GRNET	31 August 2009	
Approved by	EGI Management Board		31 August 2009	

Document Log

Issue	Date	Comment	Author/Partner

Document Change Record

Issue	Item	Reason for Change