Design Study

EGI_DS

FINAL EGI FUNCTIONS DEFINITION

EU DELIVERABLE: D3.2

Document identifier:	EGI_DS_D3-2
Date:	**4/02/2009**
Workpackages:	**WP3 EGI Functions definition**
Lead Partner:	**INFN**
Document status:	
Document link:	

Abstract:

Delivery Slip

	Name	Partner/ Activity	Date	Signature
From				
Reviewed by				
Approved by				

Document Log

Issue	Date	Comment	Author/Partner

Document Change Record

Issue	Item	Reason for Change

Table of contents

1 INTRODUCTION

1.1. Purpose

The European Grid Initiative (EGI) aims at establishing a large-scale pan-European e-Infrastructure to support research projects across a wide range of scientific disciplines, enabling them to effectively access and share a variety of national resources (computing, storage, data, instruments) and to increasingly participate in global collaborations.

Co-funded by the European Commission over 27 months, the EGI Design Study (EGI_DS) was launched in September 2007 for the conceptual set-up of a new organisational model to fulfil this vision of a sustainable grid infrastructure for science in Europe. Long-term sustainability will be achieved by establishing National Grid Initiatives (NGIs) as legal organisations, which are typically supported by governments and are designed to provide a single representation at European and international level of all national players related to a national grid infrastructure, ranging from resource providers to scientific users. The study has gathered and consolidated the requirements of a wide range of research disciplines across a large number of NGIs; its purpose is to define functions and structure of an organisation designed to consolidate, operate, manage and further develop a sustainable e-Infrastructure in Europe. The overall organisational, operational, and financial framework for this future organisation (EGI.org) is detailed in the EGI Blueprint (project Deliverable D5.4) which was endorsed by the NGI representatives during their Prague meeting on 20th January 2009. While the EGI_DS Blueprint document gives the initial organisational plan for EGI.org and is targeted primarily towards NGIs and other EGI actors and stakeholders, the purpose of the present Deliverable is to provide a more detailed technical description of the EGI functions and requirements. In conjunction with the EGI Blueprint, it constitutes the conceptual platform for the EGI project designed to establish the future EGI organisation.

Organisation and content of this Deliverable are similar to Deliverable D3.1 (EGI Functions: First Definition), however each chapter has been revised to make it consistent with the final EGI Blueprint, reflecting comments received from NGI representatives and other stakeholders. The transition period is no longer addressed in this document, since it is specifically dealt with by a further Deliverable (D5.5), currently in preparation.

The specific purpose of this document is to complement the EGI Blueprint overview of the EGI functions, outlining them in more technical detail. Note that the EGI function description in this document is work in progress; feedback from the NGIs and other stakeholders, including users and VOs, is welcome and will be actively encouraged. All the received feedback will be incorporated in a further Deliverable entitled “EGI Function Definition Feedback”and to be released four months after the present one.

In the meantime, the implementation of the EGI functions will be completely detailed by the EGI project.

1.2. Editorial responsibilities

Section	Main Providers of material	Responsible editor(s)	Work Package
1		L.Perini	WP3
2		L.Perini	WP3
3	D.Cresti, S.Fantinel, , project coordinators of reported projects	A.Caltroni	WP3

4	F. Karayannis, J.Kmunicek	J.Kmunicek	WP3
5		T.Ferrari	WP3
6		T.Ferrari	WP3
7	F.Giacomini, M.Mazzucato	F.Giacomini, L.Perini	WP3
8	D.Cresti	D.Cresti, L.Perini	WP3
9	F. Karayannis, P.Öster,	L.Perini	WP3
10	L.Matyska, K.Ullmann,	A. Candiello, L.Perini	WP3

1.3. Document organisation

Following the Introduction and the Executive Summary in Chapter 2, Chapter 3 provides on overview of European grid projects to assess the functions they fulfil and to provide an estimation of the effort they deploy. Chapter 3 is a revised version of the corresponding chapter in D3.1; only the introduction and the final summary are included in the body of this document, while the rest of the material is included in the Appendix.

Chapter 4 contains a short summary of the handover for Work Packages 2 and 4 of EGI_DS that have performed the preliminary work on which the EGI function definition is based: it consists of a collection of use cases, their mapping on the function and the findings of a survey among NGIs regarding the relevance of the proposed EGI functions and regarding their progress in structuring themselves along EGI guidelines. The content of D3.1 is not repeated in this chapter, and only a summary is provided.

Chapter 5 outlines the EGI business model, specifying scope and purposes, introducing the major players and the benefits EGI will bring to them; it is a revised and condensed version of the corresponding chapter in D3.1.

From Chapter 6 onwards, this document describes the main EGI functions, including indications for their time evolution and effort estimates:

- Chapter 6 deals with Operations and Security; it is a revised version of the corresponding chapter in D3.1.
- Chapter 7 considers Middleware Maintenance Support, it is a revised version of the corresponding chapter in D3.1.

Chapter 8 deals with the support EGI will provide to Research Teams to run the application of interest to them on the grid, after becoming grid-enabled Virtual Organisations (VO); this activity was referred to as “Application Support” in D3.1, and “User Community Support (UCS)” throughout this document and the EGI Blueprint; it is a substantially revised version of the corresponding chapter in D3.1.

- Chapter 9 touches on External Liaison Functions, including Dissemination, Industry Take-up, and other issues like relations with extra-European grid infrastructures. It has been included in this document for completeness sake, albeit it is unchanged from D3.1.
- Chapter 10 outlines the Management Function of EGI; it is a revised version of the corresponding chapter in D3.1
-

Funding and related issues are outside the remit of this document; the reader is referred to the EGI Blueprint which specifically addresses these issues.

1.4. Document amendment procedure

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

1.5. Terminology

This sub-section provides the definitions of terms, acronyms, and abbreviations used in this document.

Glossary

API	Application Programming Interface
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
CSIRT	Computer Security Incident Response Team
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
eIRG	e-Infrastructure Reflection Group
ENOC	EGEE Network Operation Centre
ERA	European Research Area
ERI	European Research Infrastructure
EU	European Union
EUGridPMA	European Policy Management Authority for Grid Authentication
FTE	Full Time Equivalent
GDP	Gross Domestic Product
GGUS	Global Grid User Support
GNI	Gross National Income
gNOC	National Grid Operating Centre
GNP	Gross National Product
IGTF	International Grid Trust Federation
JRU	Joint Research Unit

JSPG	Joint Security Policy Group
LHC	Large Hadron Collider
M	Million
MC	Middleware Consortia
MCB	Middleware Coordination Board
MoU	Memorandum of Understanding
NGI	National Grid Initiative
NREN	National Research and Education Network
OCC	Operation Coordination Centre
OGF	Open Grid Forum
OMII	Open Middleware Infrastructure Institute for Europe
PB	Policy Board
QA	Quality Assurance
ROC	Regional Operating Centre
SDC	Strategic Discipline Cluster
SLA	Service Level Agreement
SSC	Specialised Support Centres
UCO	User Coordination Officer
UCS	User Community Services
UFSC	User Forum Steering Committee
UMD	Universal Middleware Distribution
UNICORE	Uniform Interface to Computing Resources
US	United States of America
VDT	Virtual Data Toolkit
VO	Virtual Organisation
W3C	The World Wide Web Consortium
WLCG	Worldwide LHC computing Grid Project
WP	Work package
WS	Workshop

2 EXECUTIVE SUMMARY

The European Grid Initiative (EGI) Blueprint (D5.4) outlines the proposal developed by the EGI Design Study (EGI_DS) to establish a sustainable grid infrastructure in Europe by the end of EGEE-III in spring 2010.

.

This Deliverable and the Blueprint have been prepared in the context of EGI_DS to describe and detail the implementation of a sustainable e-Infrastructure in Europe; both documents form the conceptual basis for the EGI project that will be prepared over the next few months by the NGIs in conjunction with other EGI stakeholders. The present Deliverable provides, however, more detailed technical descriptions, implementation scenarios and other background material that, together with the chapter "Functions of the EGI" in the Blueprint, give the complete picture of EGI function definition performed by EGI_DS to date.

This document, like the Blueprint, is based on the vision of a large pan-European distributed computing and data grid infrastructure providing such services as outlined in the EGI Vision Document (available at http://www.eu-egi.eu/vision.pdf) [1].

Thanks to the efforts of the Enabling Grids for E-SciencE project (EGEE), consolidated by its second and third phases (EGEE-II and EGEE-III) and the contribution of related European projects and of various other regional/national or scientific grid initiatives, today researchers across many disciplines in Europe and worldwide are in a position to operate in a large-scale production-quality grid ecosystem.

At present, the EGEE grid infrastructure is interconnected to other regional grid infrastructures in Europe, such as BalticGrid and SEE-Grid (South Est Europe) as well as to grid initiatives in China, India, other Asian countries, in the Mediterranean and in Latin America (EU projects EuChinaGrid, EuIndiaGrid, EuAsiaGrid, EuMedGrid, EELA, respectively). In addition, it is peered with independent international grids, such as the Open Science Grid in the US and NAREGI in Japan. It is this interconnection of grid infrastructures across the world that enables truly global collaborative research in a wide range of disciplines.

EGEE has deployed the world's largest multi-disciplinary grid infrastructure: at present, it consists of over 250 sites across Europe and more than 80,000 CPUs with over 20 Petabytes of scientific data storage, available to some 10,000 users across a wide range of disciplines.

This success, which has positioned Europe at the forefront of grid developments, is based on the high added value that grid infrastructure services offer to research teams. They enable geographically dispersed researchers working on a joint project to collaborate seamlessly by **sharing** a variety of ICT resources distributed across several compute centres, in a coordinated way that ensures both the owner's control and the efficient use of those resources.

Without EGI, each project or discipline that requires national, European or international-scale projects to be competitive such as those mentioned in the ESFRI roadmap, would need to develop their own solution for computational and data management interoperability. This would spread the cost of the infrastructure across all projects, but would also replicate the same work by different groups adding, once more, to the overall cost and resulting in many incompatible solutions.

EGI is a partnership between National Grid Initiatives (NGIs) **and** a coordinating body called EGI.org. NGIs govern EGI.org.

Each NGI is represented by a legal organisation which should:

- ensure the **operation of a national e-Infrastructure** to an agreed level of service and its integration into EGI;
- **support user communities by** providing general services to applications and promoting grid usage to new communities;
- adhere to **EGI policies** and quality criteria;
- be the **only** recognised national body in a country with a single point-of-contact, representing all research institutions and research communities together with the resource providers related to a national grid infrastructure;
- have the **capacity to sign** the statutes of EGI.org – either directly or through a legal entity representing it;
- have a **sustainable structure,** or be represented by a sustainable legal structure in order to commit to EGI.org in the long term;
- mobilise **national funding** and resources and be able to commit to EGI.org financially, i.e. to pay EGI.org membership fees and – if there is a demand for such services in the NGI – request and pay for EGI.org services;

EGI.org is seen as the "glue" providing the required pan-European coordination and enabling coherence and synergies between the NGIs for the benefit of their international user communities.

In addition to the NGIs, EGI stakeholders include:

1. Stakeholders with associate membership status: **European International Research Organisations,** like CERN, ESA, EBI etc, willing to contribute to and interested in the availability of a European e-Infrastructure and able to assist the NGIs in providing the services and resources required by their specific sectors.
2. Stakeholders with partnership status: the **Middleware Consortia** (MC), which provide the open source middleware needed to implement the European e-Infrastructure (e.g. gLite, UNICORE, ARC, and other development teams in Europe, as well as Globus, Condor etc in the US) and which have so far supplied the middleware used in current e-Infrastructures. Maintenance and development work will continue to be commissioned to those partners by EGI to guarantee a smooth transition to EGI from current grid infrastructures and to fulfil the requirements of the user communities.
3. Stakeholders represented by the NGIs include providers of computing resources (**Resource Centres – RCs**).
4. The EGI customers:

   National **Research Institutions** (RIs): Universities, research laboratories, national research organisations, etc.

   Research Teams (RTs) operating throughout Europe that come together in national and European Virtual Organisations (VOs), whose members represent the direct users of the services offered by EGI.org and NGIs to the RIs.
5. The **Funding Agencies** that fund both the Research Teams and resource providers and seek optimum return from their investments.

The EGI functions described in this document comprise the EGI management, the EGI operations, including related security provisions, the services for ensuring the maintenance, support and standardisations of the middleware used in EGI, and the User Community Services.

The top-level management layer in EGI is the EGI Council (see Figure 5 in chapter 10) The NGIs, as defined above, constitute the EGI Council. They govern EGI.org and voice their views on all EGI-related matters as voting members in the EGI Council. Other members of this body are the Associate Members, i.e. European International Research Organisations (EIRO) with formal representation in EIROFORUM or ESFRI, and non-voting representatives of extra-European partner grid infrastructures. It is expected that this representation could be reciprocated and that the EGI Council will be represented in the governing bodies of those partner grids. The upper management of EGI.org as well as the Chair of the User Forum Steering Committee will be *ex officio* Council members.

NGIs and EGI.org jointly provide a set of grid services as well as general policies and procedures which support the following tasks:

1. Electronic authentication of individual e-Infrastructure users as the people they claim to be.
2. Allocation of project or discipline collaboration members to VOs where resources are shared with rights specified by their role.
3. Allocation of computing hardware resources to those VOs and how VO members will be authorised to use them.
4. Authorisation of VOs to run computing jobs, store and retrieve data on individual computing resources (machines, data centres, facilities, etc.).
5. Distribution and scheduling of computing jobs, workflows, data retrieval and access requests to authorised computing resources.
6. Monitoring of the jobs submitted, processed, and the data stored by individuals.
7. Accounting of users and VOs in their allocations and usage of computing resources.
8. Reporting to each NGI of allocation of resources to VOs, and the use of those resources by individual users, in order to enable the NGI (and maybe via them the national funding bodies) to account for the use of funds in terms of research results produced by VOs.
9. Coordinated management of software updates, hardware upgrades while maintaining a continuous service.
10. User Community Services (UCS), gathering requirements and ensuring their implementation.

The NGI in each member state needs to support these functions to be able to interact with EGI. The technology and organisation to support these functions have been developed over the last eight years in the series of EGEE and related projects. For EGI to operate, each country needs to coordinate its

research computing, using commonly agreed procedures and interfaces to share resources, thus enabling the integration into a pan-European e-Infrastructure.

The sustainability model proposed in the Blueprint assumes that WLCG and the current major EGEE users will progressively aquire and pay the national NGI for services they need for their international collaborative activities.

The extension of the grid infrastructure usage from research to other public sectors such as health, governmental offices, civil protection etc. can constitute also an important opportunity for the NGIs and EGI.org to expand their offerings and enhance the overall productivity of society.

EGI.org will link existing NGIs and will actively support the set-up and initiation of new NGIs where none exist. The relationship between EGI.org and the NGIs is governed by the "subsidiarity principle", to the effect that tasks that are more effectively performed at either the national or regional level should be run under the NGI responsibility. EGI.org will ensure pan-European grid coordination, aiming at common solutions wherever possible.

EGI.org will provide functions to address the coordination of:

- the operation of the infrastructure (17 FTE[1]),
- the definition of common middleware interfaces and final component certification (8 FTE),
- the services for the application support and training (11 FTE),
- the external liaison functions (4 FTE)
- the management and administration (11 FTE).

It should be noted that these **51 FTE in EGI.org** represent only a small fractionof the total effort spent on grid infrastructures in Europe today.

To run an NGI as part of the EGI, **2.5 to 30 FTE** are estimated to be necessary to cover basic international tasks. The exact requirement depends on the size of the NGI, on the demands of the local user communities and on the commitment to take on international tasks. In countries with an operating grid infrastructure, most of these resources already exist.

The operation and security of the infrastructure accounts in each NGI for the majority of the above FTE, in a range estimated from **2 to 22** depending on different parameters, as detailed in the Chapter 6 of this document. The total operational effort provided by the NGI system is estimated in **225 FTE**.

Research teams are organised in user communities which will be supported by User Community Services (UCS) provided by the NGIs. This includes Specialised Support Centres (SSCs) located in a country or an institution where the respective discipline has its natural centre of gravity: the total effort provided by the NGIs for UCS is estimated in **110 FTE.**

It is essential that the underlying middleware for the European grid be maintained and further developed. This development will continue to take place in the "Middleware Consortia" or other development teams who have demonstrated to have the necessary expertise. To minimise the risk for the infrastructure to rely on unmaintained software, the costs sustained by the Middleware Consortia for the maintenance and support of the deployed components (e.g. bug fixes and small enhancements

[1] Throughout this document "FTE" is equivalent to "FTE/year"

addressing requests from operations and applications) are typically included in the costs of the EGI infrastructure and are quantified in **70 FTE.** These costs will be covered jointly by the European Commission and the Middleware Consortia under the coordination of the EGI project;;. New specific developments are to be funded by other means (for example through other EC co-funded projects). Wherever possible, EGI will foster middleware commonality and interoperability. The EGI infrastructure relies on a Unified Middleware Distribution (UMD) to provide standard and homogenous access to resources.

EGI will typically not provide any IT resources, but will enable coordinated access, interoperability and accounting between national grid infrastructures.

3. OVERVIEW OF EUROPEAN PROJECTS

The equivalent section in D3.1 provided an overview of current or recently terminated European grid projects; the content in this document has been only slightly modified by adding ETICS and updating the final table.

Only the introduction and the final table are now included in the body of this chapter; for completeness sake, all project descriptions are still provided, but have been moved to the Appendix.

The listed grid projects comprise international projects with European participation, projects aiming at extending the usage of the European infrastructure to research communities in other regions (e.g. EUMEDGRID, EUChinaGrid, EUIndiaGrid, EELA) and a sample of projects dealing with middleware, data, etc.

The objective is to illustrate that many projects implement similar functions that can be grouped in categories identified in this document; an EGI infrastructure, offering these functions, is likely to support such projects and to allow efficiency-enhancing and cost-effective synergies.

It was attempted to include a wide number of projects to provide a comprehensive overview of the European effort; however for some of the projects detailed information was not available. For each project at least an estimate of the yearly manpower or annual budget is provided.

The majority of projects (including the major ones) are presented with complete and comparable information which allows sound estimates to be formulated.

For each project a description is provided, detailing achievements and future plans; it also includes a table illustrating the effort dedicated for the three clusters of functions (excluding management tasks), expressed in FTE (calculated by dividing number of person-months by the project duration in months) and the project duration with the annual budget or PMs.

For a full description, more detailed information on the projects and for updates, please consult the EGI Knowledge Base: http://knowledge.eu-egi.eu/.

3.1. OVERVIEW OF THE EFFORT

Table 3.1 below summarises the effort of projects for which full tables were included in this chapter: EGEE-III, ETICS, BalticGrid, SEE-GRID-SCI, OMII-Europe, GridCC, BIOINFOGRID, CYCLOPS, e-NMR, Ithanet, EUChinaGRID, EUMEDGRID, EUAsiaGrid, EU-IndiaGrid, EELA-2, D4Science and ICEAGE.

All projects	**Effort in FTE**	
Middleware	Total	146.85
	Funded	80.95
Operations	Total	308.32
	Funded	184.25
User-oriented activities	Total	303.87

(includes Application support, Training, and Dissemination)	Funded	198.11
Global Effort	**Total**	**759.04**
	Funded	**463.31**

Note that for EUAsiaGrid, EU-IndiaGrid the non-EU effort is not separated, and thus is counted in the above table.

The effort included in Table 3.1 does obviously not account for all the activities in EU grid projects; only for the projects described in this document, but not accounted for in Table 3.1, the sum of the yearly budget amounts to more than 5 M Euro, which assuming a cost of 80 kEuro per FTE, accounts for more than 65 additional FTE.

The national grid projects are more difficult to treat in a homogeneous way; they do not appear to deploy any new specific function in addition to the ones of the European projects, thus they are not considered in this chapter; it should however be noted that for many countries the effort invested in these projects is considerably higher than in the European-level projects considered in this document.

4 HANDOVER FROM OTHER WORK PACKAGES

4.1 WP2 Activity Summary

WP2 started its activities before the official kick-off of the project in June 2007; it operated unfunded during the summer to prepare for the first EGI workshop that took place at the beginning of October 2007. The official start of the project and thus WP2 was on 1st September 2007. The first task of WP2 consisted in collecting a list of NGI contact points and in identifying the main players within the EGI environment. These actors were subsequently contacted to gather their requirements and use cases as part of the first consultation process of the EGI-DS project. In parallel, WP2 worked towards the establishment of the EGI Knowledge Base, which contains, among others, information on the main actors and their use cases. The bulk of the WP2 work was the collection, analysis, summarisation and presentation of the main requirements and use cases of the identified actors. Deliverable D2.1 entitled "EGI consolidated requirements and use cases" summarises the collected input as well as the main achievements of WP2 related to the above tasks and can be consulted for a more detailed analysis of the WP2 activity. A separate deliverable, D2.2, entitled "EGI Knowledge Base (first snapshot)" provides a more detailed description of the content and the technologies behind the EGI Knowledge Base. WP2 delivered successfully also a related milestone in the form of a Workshop (WS1), entitled "EGI Requirements Consolidation and Use Case Definition (NGI evolution and the road towards EGI)". A detailed summary of the WS1 is included in D2.1. During the course of the project it was agreed that requirements and use cases will continue to be prioritised as part of WP3 in order to enable further input and requirements, and that the WP2 team will continue to work under WP3 after the termination of WP2.

4.2 WP2-WP4 Ongoing Effort and Plans

An extension of a previously conducted survey evaluating the NGIs' status, functions and maturity for EGI took place in December 2008. This involved the inclusion of answers from previously performed surveys on the same topic among NGI representatives. Currently, the original survey has been enriched with several additional questions and sent to the original 54 recipients. In addition to the inclusion of originally provided answers, this survey addressed specifically the status and readiness of individual NGIs for incorporation into the EGI model as outlined in the last version of the EGI Blueprint. The analysis of the findings from the updated survey clearly indicates a significant improvement since the beginning of 2008. More than half of the NGIs reported crucial changes about their NGI structure since March 2008 as final steps towards the formal creation of a legal entity or corresponding partnership. Others reported official approval of recently established NGIs. The critical feedback concerning the guidelines mentioned in the EGI Blueprint version of December 2008 was overall positive with a 95 % awareness rate among the respondents. Although less than half of the NGIs currently comply with the guidelines, more than 80 % of the NGIs expect to do so by the end of 2009. Moreover, 84 % of the respondents consider the guidelines appropriate. This highlights the importance NGIs attribute to creating the required framework for active and successful participation in the EGI ecosystem. The next steps will focus primarily on the analysis of the provided feedback concerning all the major functions of the proposed EGI model. The feedback is expected to be provided basically on two documents: the final version of the EGI Blueprint itself (especially suggested functions dealing with applications and user-related issues) and the present Deliverable.

5 EGI ACTORS AND USE CASE

EGI is a partnership between National Grid Initiatives (NGIs) **and** a coordinating body, named EGI.org. NGIs govern EGI.org. EGI.org is seen as the "glue", providing the necessary pan-European coordination and enabling coherence and synergies between the NGIs for the benefit of their international user communities.

The EGI stakeholders include:

1. **National Grid Initiatives (NGIs):** the NGI is a legal organisation responsible for the management of the national grid e-Infrastructure and for maintaining relationships with customers (primarily the research institutes and research projects carried out by the VO constituted by research teams) as well as with a set of national partners.
2. Stakeholders with associate membership status: **European International Research Organisations** like CERN, ESA, EBI etc, willing to contribute to and interested in the availability of a European e-Infrastructures, and able to assist the NGIs in providing the services and resources required by their specific sectors.
3. Stakeholders with partnership status: the **Middleware Consortia** (MC), which provide the open source middleware needed to implement the European e-Infrastructure (e.g. gLite, UNICORE, ARC and other development teams in Europe as well as Globus, Condor etc in the US) and which have so far supplied the middleware used in current e-Infrastructures. Maintenance and development work will continue to be commissioned to those partners by EGI to guarantee a smooth transition to EGI from current grid infrastructures and to fulfil the requirements of the user communities.
4. Stakeholders represented by the NGIs include the providers of computing resources (**Resource Centres – RCs**).
5. The EGI customers:
    a. National **Research Institutions** (RIs): universities, research laboratories, national research organisations, etc.
    b. **Research Teams** (RTs) operating throughout Europe that come together in national and European Virtual Organisations (VOs) whose members are the immediate users of the services offered by EGI.org and NGIs to RIs.
6. The **Funding Agencies** that fund both the research teams and resource providers and seek optimum return from their investments.

The relationship between the various EGI players is illustrated in Figure 1.

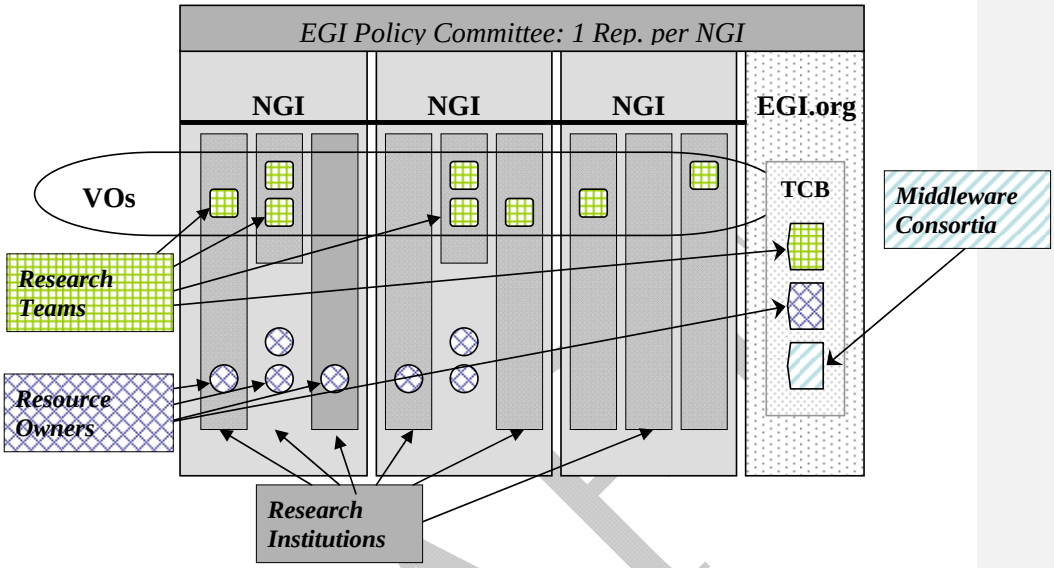

Figure 1: Actors in the EGI Business Model

Use Case

Research teams (RTs) typically collaborate on a temporary basis (at national, European or international level) within a Virtual Organisation (VO) to form a project with specific research objectives. Such a project is normally approved by peer review committees (acting at national and/or European or international level) set up by the involved research institutions or funding agencies that allocate the necessary funds, including those for IT resources.

VOs typically need to share specific IT resources, and are characterised by a particular usage model for a given set of EGI baseline services, such as authentication and authorisation services, accounting services for NGIs and VOs, services for data sharing at different levels of abstraction, services for compute sharing for different types of resources, monitoring services, etc.

Research teams belong to different research institutions (universities, laboratories, applied research institutions etc). Resource consumption may be organised in different ways through VOs. A VO may fulfil its resource requirements through resources from its constituent research institutions , through resources provided by a resource provider (either by another research institution or by a commercial partner) or in any other way which fits the VO's needs best.

An e-Research project needs to rely on a set of *software tools* which enable the *secure sharing* of all the partner organisations' "local" IT resources and data distributed over different administrative domains (**Requirement n.1**).

Such sharing may concern the CPU cycles of the commodity clusters used for data analysis, the fast interconnected parallel systems for MPI applications of computational chemists, Earth observation, biomed, weather forecast etc., the files located in distributed storage systems for image visualisation of

astronomers, medical doctors etc, the metadata located in distributed archives systems related with a large variety of applications, etc.

The sharing occurs through software services (*grid middleware*) which offer the user a uniform interface, hiding the local diversities and allowing a distinct level of authorisation according to the member's role and agreed project policies. In this way, a common pool for all kind of computing resources is created for the project, enabling the project partners to use the available distributed resources and data in the most efficient way

The reference resource centres/providers are requested by the VOs to operate the set of services which enable them to reach the above goal. Distributed accounting at institutional level is also required together with tools to monitor the activities as well as support services to the VOs to run their legacy user applications in this multi-administrative domain pool.

It should be pointed out that many VOs, especially the new ones, tend to consider their sharing requirements a special case and call for bespoke integrated vertical services (e.g. LHC experiments, ESFRI, etc). However, the use of a layer of services, shared by other VOs whenever possible, has the advantage of manpower saving and facilitates the development and, even more importantly, the seamless operation and maintenance of more fundamental baseline services essential for the sharing. This enables also partial sharing with different VOs, thus increasing the pool of available resources (see below).

It is economically much more *convenient and efficient* for funding bodies to promote, support and fund the procurement and the operation of a shared, robust, secure and certified set of baseline grid services rather than a chaotic set of tools that each VO may freely ask to adopt or develop. Such an integrated set of services can be offered and operated by EGI.org at European level and by the NGIs at national level as part of the general EGI/NGI e-Infrastructure to enable global sharing (**Economy argument**).

Past investments at European or at national level can therefore be re-used and are likely to benefit also new VOs; high-level special services that may still need to be developed will be less expensive and founded on a mature layer with a longer lifetime and wider user spectrum (**Reuse and long-term perspective argument**).

The perspective of a better global return for the money invested by funding bodies as well as of the creation of global pools based on a well-defined certified set of services constitutes the important **Requirement n.2** for EGI which is likely not to apply to all VOs.

Resource centres/providers have so far obtained the general grid middleware services they need to operate from external providers (EU projects or Middleware Consortia). In line with different VO consolidated practices, at times they are supporting more than one middleware solution. Based on the need for simplified operations, to *avoid* having to charge the *high costs* associated with the support of chaotic and very expensive multiple special environments, and having to offer a well-defined *quality of service* for multiple solutions, resource providers will typically appreciate and support the coordinated action of EGI.org and of their respective NGIs in moving towards a *progressively unified solution* for the services they will have to operate. (**Operational argument**). This constitutes the **Requirement n.3** for EGI.

6 FUNCTIONS OF EGI: OPERATIONS AND SECURITY

The Operations and Security Function includes those EGI services which are needed to ensure optimal functionality of the pan-European infrastructure and overall seamless interoperation of national and regional grids. At the same time, a common authentication trust domain is required to persistently identify all grid participants and common security policies need to be defined and enforced. In the context of a European e-Infrastructure, security policies and operational security to support and manage the activity of teams drawn from the NGIs will need to be coordinated.

The various tasks of this function need to be structured according to an agreed operational model that meets various requirements: scalability and interoperability, availability and reliability, sustainability, and autonomy of NGIs.

Many of the EGI operations and security tasks are jointly delivered by EGI.org and the NGIs, i.e. the EGI.org tasks complement those carried out by the NGIs at regional level. *NGI International Tasks* are such activities that allow national IT resources to be shared at pan-European and international level in a uniform, robust, and seamless way. Depending on the needs of the individual NGIs, the international tasks are integrated by the *NGI National Tasks* which are carried out to satisfy NGI local requirements.

In this context, common standards and/or specifications for interoperation between NGIs play a critical role in ensuring interoperability within EGI. To this end, NGIs are requested to collaborate and to jointly define specifications, policies, best practices, and in general, to share operational responsibilities. It is important to note that, at the time of writing, the devolution of operational and security activities and responsibilities is already common practice among the main grid infrastructure projects in Europe.

Operations and security tasks of EGI.org, on the one hand, and of the NGIs, on the other, along with the corresponding effort, are outlined in the following sub-chapter.

6.1 EGI.ORG TASKS

EGI Operations and Security activities are classified as follows:

1. operation of tools and services;
2. support;
3. other tasks;
4. security;
5. development.

Notation: *EGI.org and NGI tasks are numbered according to the following method: prefix O-E identifies operations services provided by EGI.org, whereas O-N identifies those provided by NGIs.. Explicit indication is given of those tasks that are deemed necessary (as opposed to optional services) and of those EGI.org activities that, technically , could be distributed to NGIs as they do not require to be located in the EGI.org site.*

6.1.1. Operation of tools and services

O-E-1 and O-N-1: Operation of the grid topology and configuration repositories (EGI.org and NGIs) – – *necessary, can be distributed*

Many aspects of operations rely on the availability of information (as applicable) 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.

O-E-2 and O-N-2: Operation of accounting repositories for international VOs (EGI.org and NGIs) — *necessary, can be distributed*

The accounting repository is designed to keep records about usage of compute, storage, networking and other types of resources as required by 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 to informVO managers about the amount of IT resources "consumed" by the respective users across different domains of the e-Infrastructure. The deployment of standard interfaces between accounting systems in different NGIs is therefore critical for the interoperable exchange of records between different domains. For each NGI, EGI.org is responsible for gathering and making publicly available accounting information (as applicable and in conformance with local laws and privacy requirements of the EGI actors). The availability of a pan-European accounting infrastructure is a key enabling component of the EGI business model.

O-E-3 and O-N-3: Operation of grid repositories storing monitoring and performance data and other related information (EGI.org and NGIs) – *necessary, can be distributed*

Availability, status and performance information about grid services and sites are needed to monitor the health of the infrastructure and to verify the Quality of Service delivered to VOs and other NGIs. The collection 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 assess its level of service and compliance with VO requirements. This entails the operation of repositories and supervision of the processes of populating them, the maintenance of schema for publishing site and service status information, the ownership of the information schema used, the preparation of reports, etc.

This task includes the gathering of network performance information for assessment of network quality and reporting purposes to ensure that the underlying network infrastructure is working properly and efficiently, and that network providers are honouring their contractual obligations.

In this context, EGI.org tasks comprise 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.

O-E-4 and O-N-4: Operation of the grid operations portals (EGI.org and NGIs) — *necessary, can be distributed*

The 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 displayed information 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.

O-E-5 and O-N-5 Grid operation and oversight of the e-Infrastructure (EGI.org and NGIs) — *necessary, can be distributed*

EGI.org operation and oversight activities cover the detection and coordination of the diagnosis of problems affecting the entire EGI e-Infrastructure during the entire lifecycle until resolution, the reporting of middleware issues to the developers, the execution of quality checks of the services provided by NGIs, and the handling of operational problems that can not be solved at the NGI level. This task coordinates the oversight of the NGI e-Infrastructures (run under the responsibility of the NGIs), which– at the NGI level – includes the monitoring of the services operated by sites, the management of tickets and their follow-up for problem resolution, 1st and 2nd line support to operations problems, the suspension of a site when deemed necessary, etc. Within EGEE, this EGI.org task is currently carried out in cooperation with the relevant regional operations centres (via rotating shifts) according to a two-level hierarchical model [COD]. This model is envisaged to evolve in such a way as to allow NGIs to autonomously run oversight activities in the region or to federate in order to share efforts.

6.1.2. Support

O-E-6, O-E-7 and O-N-6, O-N-7: Central and regional grid user support and ticketing system (EGI.org and NGIs) — *necessary, can be distributed*

User support relies on a central helpdesk, which is a regional support system under 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 to allow a bi-directional exchange of tickets (for example, tickets opened locally can be escalated 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 for network end-to-end problems in the grid is of equal importance, especially for demanding applications,,as connectivity is provided by the pan-European research network backbone, GÉANT, and by a large number of national research and education networks (NRENs), each providing links to sites within the countries. A Network Operation Centre provides the operational interface between the grid and the relevant network players to monitor the end-to-end connectivity of grid sites [ENOC].

The NGIs provide 1st line local/regional support to users and centres, while EGI.org is resposnible for the maintenance and operation of the central ticketing system (GGUS like) and for the triage of incoming problems.

a. Maintenance and operation (*can be distributed*): 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 under 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) – can be *distributed*, consists of the monitoring and routing of all active tickets in the grid user support system by grid and VO experts, who are responsible for directing the problems to the appropriate second-line specialised support units. This process combines manual as well as automated procedures.

O-E-8 Gathering of requirements for user support tools and process — *necessary, can be distributed*

Tools and the process for user support are designed to meet the requirements of customers by taking input from NGIs, VOs and resource centres. Additional requirements may arise as the current middleware stacks evolve and new user communities need to be supported. EGI.org is responsible for the coordination of this process.

6.1.3. Other tasks

O-E-9 Coordination of middleware roll-out and deployment (EGI.org, *necessary, centralised*), middleware pilot and certification testbeds (EGI.org and NGIs, *necessary, can be distributed*)

Middleware updates are required to move from certification into production as quickly as possible, whilst it needs to be ensured that such updates are actually suitable for deployment in the production grid. EGI.org coordination will be needed for strategy decision, for example to decide significant changes to processes, and to ensure that resource sites are encouraged to upgrade whenever new critical updates of supported middleware stacks are released. Being still in a phase where middleware is subject to frequent bug fixing cycles, prompt alignment of the grid services and components to the latest releases enhances functionality and improves availability of the overall infrastructure.

In addition, NGI operation of facilities for testing and certifying middleware is important for the deployment of high-quality middleware by allowing VOs and site managers to test grid components during the early development and release phase. O-E-10 Coordination of resource allocation and brokering support for VOs from NGIs (EGI.org) — *optional, centralised*

VOs can specify requirements in terms of resources to be guaranteed by the overall pan-European grid infrastructure. In this case, coordination – as required by VOs – contributes to ensuring that a suitable production infrastructure (offered grid core services and resources) is in place to meet such requirements. Tools for the automation of management and negotiation of SLAs are still to be developed. EGI.org is responsible for providing support and coordination of this process.O-E-11 Coordination of interoperations between NGIs and with other grids (EGI.org) — *necessary, centralised*

Coordination is required 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 intending to cross grid boundaries need to be assured that the environments will be similar and the applications will function adequately without major adjustments. 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” include Asia-Pacific regional grids, OSG, Naregi, and related infrastructure projects.

This role covers the definition of operational tool interfaces, procedures and operational activities to allow NGIs to interoperate. EGI aims to continue the collaboration established with operations centres outside Europe in order to maintain the current integration of non-European sites into the production infrastructure. EGI.org is responsible for support and coordination

O-E-12 Coordination of network support (EGI.org) – *necessary, centralised*

Network operation design, handling of troubles affecting international VOs, and network assessment allow EGI to establish links between grid operations and network operations. A centralised approach is proposed here in order to align this task to the other external liaison tasks run by EGI.org.

O-E-13 Definition of best practices, operations procedures, operations requirements (EGI.org and NGIs) — *necessary, can be distributed*

Interoperation relies on the definition of best practices and of general operational procedures for daily monitoring activity for sites and federations. EGI.org is responsible for the coordination of these activities.

O-E-14 and O-N-8: Operation of the production grid core software services, catch-all services for international VOs, catch-all VO (EGI.org) – *necessary and distributed*

Grid core services are components of the EGI e-Infrastructure. They are software components that typically run on server machines. “Grid service” refers to a software instance (a Web service in many cases) "that is designed to operate in a grid environment, and meets the requirements of the grid(s) in which it participates." [GLO]

In particular, core software services, provided by the Middleware Consortia, are the essential components for the overall grid functionality to be operative. Catch-all instances can be required to support small user communities. It is the responsibility of EGI.org to ensure that user communities are properly supported by the respective NGIs. Examples of gLite core software services are: the VO management service (e.g. VOMS), the file catalogue and transfer services (e.g. LFC and FTS), job management services (e.g. WMS), information services (e.g. BDII), security services, etc.

Authentication is also fundamental to get access to grid resources. Therefore, a catch-all certification authority needs to be available to any user community in EGI.

6.1.4. Security

The nature of security vulnerabilities and risks presented by grid infrastructures calls for coordination among the grid participants at various levels: at the NGI level between site managers, the NGI itself and CERTs from the national research and education networks (NRENs), between NGIs and EGI.org to adopt and enforce common policies, and between EGI and international bodies, such as EUGridPMA (the European Policy Management Authority for Grid Authentication) and IGTF (the International Grid Trust Federation).

A European-wide e-Infrastructure requires a certain degree of centralised coordination of security policies and operational security. Support and coordination of the activity of teams, drawn from the NGIs, will be the task of EGI.org.

O-E-15 Coordination of security policy development and enforcement — *necessary, centralised*

Security policy development and enforcement are needed to define an agreement on matters such as best practices, security policies, CA policies, etc. A team of security experts in NGIs will be tasked with the definition and application of standard security policies. EGI.org is responsible for support and coordination.

O-E-16 Coordination of security and incident response - *necessary, centralised*

It needs to be ensured that EGI members from NGIs follow common policies for coordinated incident response. EGI.org is responsible for coordination and support.

6.1.5. Development

O-E-17 Coordination of development and maintenance of operational tools – *necessary and centralised*

While the tools for accounting are included in the middleware, additional tools will be required to support operations. Examples are: tools for monitoring, dashboards and alarm systems, ticketing systems, portals, etc. as well as new tools to improve automation.

EGI.org is responsible for coordinating the maintenance of the set of the tools currently used in European production grids as well as for the necessary upgrades to keep them in line with the quantitative and qualitative evolution of the grid. This includes monitoring tools to measure and report on the quality of networks used by grid projects to ensure that the underlying network infrastructure is working properly and is efficiently used, and that SLA constraints with network providers are met.

It is foreseen that only coordination responsibility (necessary task) with rest with EGI.org, while a set of willing NGIs will be responsible for the development work to be co-funded by the EC.

6.2 EGI.ORG EFFORT AND TIMING

On the basis of the detailed description of the above activities, the following paragraphs summarise the list of activities carried out by EGI.org and NGIs, providing effort estimates for the first three years of EGI. For simplicity's sake, estimations are expressed in Full Time Equivalents (FTE).

6.2.1. Operation of tools and services

O-E-1. Operation of the grid topology and configuration repositories. EGI.org FTE: 1

O-E-2. Operation of accounting repositories for international VOs. EGI.org FTE: 1

O-E-3. Operation of the grid repositories storing monitoring and performance data, and other related information. EGI.org FTE: 2.5

O-E-4. Operation of the grid operations portals, EGI.org FTE: 0.5

O-E-5. Grid operation and oversight of the e-Infrastructure. EGI.org FTE: 1

6.2.2. Support

O-E-6. Maintenance and operation of central ticketing system: EGI.org FTE: 2.

O-E-7. Triage of incoming problems: assignment of tickets to the 2nd line support units, ticket escalation end ticket follow-up to ensure they get closed, EGI.org FTE: 2

O-E-8. Gathering of requirements for user support tools and processes: EGI.org FTE: 0.5

6.2.3. Other tasks

O-E-9. Coordination of middleware roll-out and deployment, middleware pilot and certification testbeds. EGI.org FTE: 1

O-E-10. Coordination of resource allocation and of brokering support for VOs from NGIs, EGI.org FTE: 0.5

O-E-11. Coordination of interoperations between NGIs and other grids. EGI.org FTE: 0.5

O-E-12. Coordination of network support, EGI.org FTE: 0.5

O-E-13. Coordination of definition of best practices, operations procedures, operations requirements, FTE: 0.5

O-E-14. Operation of production grid core software services, catch-all services for international VOs, catch-all CA: EGI.org FTE: 1

6.2.4. Security

O-E-15. Coordination of security policy development and maintenance; EGI.org FTE: 0.5
Coordination of security and incident response: EGI.org FTE: 1

6.2.5. Development

O-E-16. Coordination of development and maintenance of operational tools. EGI.org FTE: 1

Table 1: Overall effort for EGI.org operations and security critical services

ACTIVITIES	FTE
Operation of tools and services	6
Support	4.5
Other tasks	4
Security	1.5
Development	1
TOTAL	17

6.3 NGI TASKS

The list of tasks in this paragraph is intentionally not exhaustive, as only the *necessary* international tasks of an NGI were to be considered. Many of the tasks in this section are performed by the NGIs and coordinated by EGI.org. The ownership of such tasks does not prevent an NGI from devolving the

operation of the task itself to a third party, or from opting to purchase it from EGI.org. Tasks not relevant to the overall EGI operation model, or specific to national VOs, have been omitted.

It is at the discretion of the NGIs to choose the supply model that fits them best. For instance, an NGI can federate with other NGIs to share their joint effort, it can purchase a set of services from other NGIs or other partners, or put a request to EGI.org. To assist NGIs, especially during the transition phase, the possibility is foreseen for EGI.org to supply catch-all operational services – in addition to the central ones – according to demand. The number of FTE needed by EGI.org to run catch-all services is expected to be proportional to the number of NGIs requesting it.

O-N-1. Operation of the NGI grid topology and configuration repository - *necessary*

O-N-2. Operation of the NGI accounting repository - *necessary*

O-N-3. Operation of repositories storing monitoring and performance data, and other related information – *necessary*

O-N-4. Operation of the NGI operations portal – *necessary*

O-N-5. NGI e-Infrastructure oversight (monitoring of status of services operated by sites, opening of tickets and their follow-up for problem resolution), 1st and 2nd line support in case of operational problems, site suspension, reporting to EGI.org in case of middleware problems and general operational issues, etc. – *necessary*

O-N-6. Operation of the NGI ticketing system, gathering of new requirements for user support tools in the region – *necessary*

O-N-7. Regional helpdesk: support to users and site managers via a local/regional helpdesk - *necessary*

O-N-8. Operation of production grid core software services, catch-all services for international VOs, catch-all CA: running the required grid services provided by the NGI, and services required by international VOs – *optional;* availability of Certification Authority to distribute X.509 certificates to users and servers in the region - *necessary*

O-N-9. Operations coordination at NGI level - *necessary*

a) Security and incident response coordination in the region

b) Roll-out of middleware updates in the NGI

c) Resource allocation in the NGI

d) Interoperation with national and regional grids

6.4 NGI EFFORT AND TIMING

The estimation of the total NGI manpower is determined by the size of the NGI, the service level requirements to be met in the respective region, the level of participation in EGI activities, and its organisational structure (e.g. an NGI can decide to outsource tasks or take on extra tasks on behalf of other NGIs or EGI, etc.).

This document provides tentative estimations for the initial three years of EGI. NGIs are divided into three categories: “small”, “medium” and “large”. Estimations are based on the present EGEE experience, assuming that increasing automation and expertise will at least partly compensate for the increase in application variety and middleware complexity.

- Small NGI: 2-4 FTE
- Medium NGI: 5-10 FTE
- Large NGI: 14-22 FTE

Note that for countries presently involved in EGEE, during the first three years of EGI, the same amount of personnel, as currently available in EGEE, is expected to be funded to work on NGI international activities. The amount of FTE currently involved in operational activities for the EGEE III project alone is 189.9, of which 85.9 are funded by the European Commission.

Once all the NGIs that have expressed interest in EGI are properly constituted and have joined EGI, EGI is assumed to comprise 6-7 large NGIs, 12-16 medium NGIs and 16-20 small NGIs. However, during the very first year the number of NGIs could be somewhat smaller.

More details on the NGI effort estimation are provided in Appendix B.

The operations and security function is supported by manpower effort and additional hardware resources that are needed (mainly at NGI level) to host grid core software services, operational tools, testbeds and auxiliary IT services (wiki pages, agenda pages, databases, etc.). Based on the current status, it is estimated that, for some large EGEE ROCs, about 150 servers will be needed to fulfil such functions. Hardware resources required for the deployment of the NGI e-Infrastructure are funded via national funding sources (i.e. no EC co-funding is requested in this case).

6.4.1. Evolution

FTE estimates refer specifically to the overall amount of effort needed during the EGI transition phase (about three years). Increased efficiency after a few years is likely to impact on the staff requirement for the initial operational model; however, this reduction in personnel is expected to be partially matched by additional staff requirements for new activities to meet the evolving needs of new communities. As to development, a reduction in cost is forseen in about three to five years, once operational tools reach maturity. At that stage, a small fraction of funding will still be needed for maintenance of existing tools.

In three years time, some operations and security tasks of EGI.org are expected to evolve into services, of which some will be necessary and sold as a bundle, while others will be optionally subscribed to by the NGIs. Depending on the type of service, these will be charged on a per-use or flat-rate basis.

7 MIDDLEWARE DEVELOPMENT AND SUPPORT

This chapter provides the technical details that complement and complete the Middleware Section (3.2) of the Blueprint document. The concepts expressed in the latter will generally not be repeated here as the reader is assumed to have prior knowledge of the Blueprint.

7.1. MIDDLEWARE TASKS AND SERVICES

The overall goal of EGI is to establish a large-scale, production-grid infrastructure for the sharing of IT resources and data, built on national grids that interoperate seamlessly at many levels. EGI will supply reliable services to a wide range of applications, ranging from "mission critical" to prototyping and research. The approach to meet this objective consists in distributing a variety of responsibilities among the various players.

The EGI-specific technical objective, to be achieved in a coordinated effort by the EGI central organisation (EGI.org) and the National Grid Initiatives (NGIs), is to oversee, on behalf of all stakeholders, the procurement, certification, deployment, and operation of software services (i.e. the software infrastructure) and to define the organisational rules, policies and procedures that consitute the required standard access and sharing mechanisms for all sort of IT resources and data which currently are and will continue to be made available to researchers by national resource providers. At present, these providers are predominately public or semi-public resource centers of varying scope and dimensions, which will continue to be 100 percent funded at national level.

The Blueprint calls for the establishment of a Middleware function in EGI and outlines the reasons why it needs to remain in full control of the software infrastructure, which constitutes one of the key services offered to all stakeholders. The Blueprint recommends the implementation of such a function, at least during the first stage of EGI, with limited manpower in EGI.org and no further immediate mandatory contributions from the NGIs. This is to avoid the risk of disruption of the current services, used daily by thousands of researchers. The maintenance and support of the middleware components deployed in the current e-Infrastructures should continue to be co-funded at a level of 50% by the EC, without excluding contributions from other interested partners, and sustained for the remaining part by the Middleware Consortia and other development teams that have agreed to move toward a Unified Middleware Distribution (UMD) under EGI coordination. Once the EGI e-Infrastructure is well established, the maintenance and support of the widely adopted legacy services will be progressively taken over by the NGIs through the payment of service charges, while the necessary innovation and new developments, as defined by the needs of user communities and operations, will continue to be co-funded by the EC through competitive calls.

This chapter provides more details about:

1. Middleware components and Middleware Consortia.
2. Guidelines for UMD.
3. Role of the EGI.org Middleware Unit.

4. Components and services proposed for inclusion in UMD in the first stage of EGI.
5. Cost estimates for middleware maintenance.

7.1.1 MIDDLEWARE COMPONENTS AND MIDDLEWARE CONSORTIA

A variety of middleware components are currently deployed in the EU e-Infrastructures. They are the result of several years of European and international competitive efforts aimed at satisfying the needs of a large number of user communities with complementary requirements and dimensions, ranging from teams of a few individuals to very large international collaborations with thousands of researchers. They all adhere to a general service-oriented approach aimed at complying with the evolving Web services and the Open Grid Forum standards.

A large part of components are provided by three middleware distributions, namely ARC, gLite and UNICORE. Developed predominantly in Europe, all three of them are used in production in the three main EU e-Infrastructures: EGEE, DEISA and NDGF. Each provides a middleware stack suitable to meet the most fundamental needs; however, none of them represents a fully satisfactory solution for all needs. Other middleware platforms are in use in Europe (such as GridWay, pGrade, AssesGrid, GRIA, etc.) which are funded by the European Commission and by national funds; they provide higher-level services designed to complement the basic services provided by the aforementioned three major stacks. .

Recent efforts, in particular those undertaken by the OMII Europe project, have already succeeded in improving the interoperability between the three main grid platforms in use in Europe. Building on the successful developments over the last eight years, on the contributions provided mainly by other EU initiatives, as outlined above, and in conjunction with selected components originating from the US-based Globus and VDT, the ARC, gLite and UNICORE stacks provide the bulk of the services in use in the largest general-purpose EU e-Infrastructures (EGEE, DEISA and NDGF), serving thousands of researchers every day.

The ARC, gLite and UNICORE stacks thus constitute the basis for the creation of the open-source Unified Middleware Distribution (UMD) that the future European Grid Initiative (EGI) will make available to national resource providers as a key integral part of its offer and business model. The availability of certified grid services that can be easily downloaded from a common UMD repository, together with a set of common procedures, policies and rules to be established by the EGI, will enable research teams to easily access and share computational resources and data, supplied by their national resource centers and funded at national level.

The EGI Middleware function is designed to ensure the current level of quality of the deployed services in the transition period and during the initial years of consolidation of the new EGI organisation.

To fully satisfy the operational quality requirements, it is essential that, during the transition towards the new sustainable European organisation embodied by EGI, the middleware currently represented by these stacks and other identified services continues to be supported, maintained and further developed, particularly in view of emerging standards and, in some parts, completed and hardened from its current stage. The requirements will be established by

the Middleware Coordination Board to include representatives of VOs, operations and middleware development teams; this represents a logical development of current best practices in EGEE, DEISA and other national experiences.

The maintenance, development and evolution towards standards for these three EU stacks and related middleware projects are currently co-funded by national institutions or consortia and by the EC via competitive bids.

7.1.1.1 ARC

The Advanced Resource Connector (ARC) has been developed by the NorduGrid collaboration (http://www.nordugrid.org/) and associated projects since 2001. Its decentralised architecture entails high efficiency, low maintenance costs and robust performance. It is highly portable and is available for all major Linux flavours. This, in turn, allows a decentralised deployment of ARC in more than 60 sites, with over 30,000 cores. In particular, ARC is adopted by the NDGF (Nordic DataGrid Facility) to support the world's only distributed heterogeneous Tier1 centre. The next generation of ARC is currently under development; it sets out to minimise dependencies on third-party components, to improve extensibility, interoperability and to allow portability to non-Linux platforms.

The NorduGrid consortium was established in 2001 by five Nordic academic institutions and is based upon a Memorandum of Understanding (MoU) which is not legally binding. The MoU establishes the Steering Committee and the Chairperson and defines their duties. The consortium has no termination date and has no collectively owned resources. NorduGrid currently conducts consultations towards establishing an international "ARC consortium" designed to become a legal entity prior to the EGI start. NorduGrid thus guarantees middleware support, maintenance and further development of ARC beyond the scope of the current project.

7.1.1.2 GLITE

The gLite middleware is the result of a truly pan-European development effort undertaken by the EDG-EGEE project series which started in 2001 and is co-funded by the EC via competitive bids. The services offered by gLite provide the backbone of the EGEE infrastructure, the largest multi-disciplinary grid infrastructure in the world, bringing together more than 140 institutions to produce a reliable and scalable computing resource-pool available to the European and global research communities. At present, it consists of approximately 300 sites in 50 countries and gives its 10,000 users around-the-clock access to 80,000 CPU cores (largely commodity clusters with some HPC systems) and to very large (>15 PB) distributed storage systems, processing up to 300,000 jobs per day from scientific domains ranging from biomedicine to fusion science. The gLite middleware consists of an integrated set of components, compliant with open standards and covering all aspects of the grid infrastructure. Originally developed for the scientific Linux environment, extensive efforts are now underway to make it more widely available on other platforms. The gLite software is managed in all its phases, from development, over testing to release, with the tools provided by ETICS (eInfrastructure for Testing, Integration and Configuration of Software),

an EC-funded project.

The gLite community is actively pursuing the constitution of the gLite Open Consortium to maintain and evolve the gLite middleware beyond the EGEE series of projects and thus to provide a long-term sustainable roadmap for the gLite software to meet the needs of its diverse user community. The Consortium will be established as a not-for-profit entity and will be open not only to the institutions currently providing components but also to any other partner willing to contribute to the Consortium's objectives. The organisational model for the software development and related activities will be based on teams fully responsible for the individual middleware components; the coordination among the teams will rest with a Technical Coordination Board led by a Technical Coordinator.

7.1.1.3 UNICORE

The UNICORE middleware (http://www.unicore.eu/) has a traditional HPC background (for 10 years). It is used in HPC-related infrastructures, like DEISA, (serving a similar amount of CPUs as in EGEE, but concentrated on a few powerful supercomputers) and in the future PRACE (European PetaFlop/s Supercomputers), but is also implemented in non-HPC-focused NGIs, such as D-Grid and some Swiss SwiNG projects. UNICORE is characterised by its open, extensible, lean, and interoperable Web services architecture which supports many open standards, providing seamless, secure and intuitive access to grid resources. Emphasis is on workflow capabilities, security, application support and ease of installation and configuration. Since 2004, the UNICORE middleware has been open source under a BSD license and publicly available at SourceForge (http://sourceforge.net/projects/unicore/). It is developed by the open source developer community of UNICORE with a set of core partners who provide major elements of the software and are responsible for the development of the core components as well as for the release management. Institutions that have a long-term interest in the UNICORE grid technology have joined the "UNICORE Forum e.V." (http://www.unicore.eu/forum/) established in 1999; its legal status is a registered, open, non-profit association pursuant to German law. Its objective is to promote the development and distribution of UNICORE beyond the scope and duration of EU- or nationally funded projects. The UNICORE Forum e.V. currently has 32 members comprising research institutions as well as commercial organisations. The Technical Advisory Board of the UNICORE Forum e.V. is tasked with devising the roadmap and strategy of the future UNICORE development. It evaluates technical proposals, discusses technical solutions to be implemented in UNICORE and, thus drives and monitors the open source development process of UNICORE.

7.1.2 GUIDELINES FOR THE UNIFIED MIDDLEWARE DISTRIBUTION (UMD)

The Consortia mentioned in the previous section have agreed that the middleware components, tools and services they currently support have to evolve into a Unified Middleware Distribution (UMD). UMD will contain components which will satisfy the needs of the user communities and of the resource providers and conform to quality criteria defined

by the EGI.org Middleware Unit and endorsed by the Middleware Coordination Board.

The quality criteria that will determine which components will be included in UMD comprise:

- **Interoperability**: services included in a UMD release should be fully interoperable with all other UMD implementations adopted in the EGI grid infrastructure.
- **Completeness**: the set of available components and tools included in UMD, once adopted by an NGI, should allow the national infrastructure to be operated in a fully self-functional and autonomous way and should, at the same time, be completely integrated with the rest of the pan-European EGI infrastructure. The grid services included in UMD should address the needs of all current VOs; moreover, a process should be put in place to allow them to evolve according to the requirements of new scientific communities.
- **Scalability**: available services should allow the management of resources and services in an e-Infrastructure that is to cater for scientific user communities ranging in size from a few individuals to thousands of researchers. Different service implementations should be included to take into account both the need of simplicity for small user communities and scalability for the larger ones. In addition, the services should be able to cope with the anticipated growth in scale (in terms of users, services and sites operated) over a short time period.
- **Simplicity**: UMD should contain tools to download the appropriate services, to provide assistance during their configuration, and to perform as much automatic set-up as possible.
- **Extensibility:** UMD must provide interfaces (and "hooks") to allow independent development (by any interested party) of higher-level and additional services that will create a software pool from which further UMD innovation will be drawn. Gateways to other EU and non-EU (e.g. Globus) grid systems and components will be one example of services built on the extensibility interfaces.

Given the wide variety of needs, it is acceptable that different implementations of the same service or of the same interface are available at the same time in UMD, provided they are actually requested and compliant with the UMD quality criteria. However, wherever possible, a progressive specialisation of the different services will be pursued to avoid unnecessary duplication of effort.

7.1.3 ROLE OF THE EGI.ORG MIDDLEWARE UNIT

The environment, in which EU middleware development currently takes place, consists of distributed multiple teams of experts specialised in one or more services and typically organised around the three Middleware Consortia, alongside additional teams with complementary expertise belonging to other EU and international initiatives.

In order to leverage the existing clusters of competence, it is advisable to maintain this decentralised model based on autonomous teams while introducing with EGI an effective pan-

European technical and financial coordination. Decentralisation will also allow the introduction of other development teams, eventually including teams who develop components on a commercial basis.

The central technical coordination of the development teams will be supported by an EGI.org unit, called Middleware Unit (MU) and led by the Chief Technical Officer (CTO). The main objective of the MU is to ensure the availability of the required middleware services at pan-European level with the assistance of additional technical bodies, including the relevant experts appointed by the Consortia. This coordination activity requires the establishment of bodies such as a Middleware Architects Group and a Middleware Technical Management Group to follow the high-level long-term architectural issues and the day-to-day execution of the plans, respectively. The exact scope and function of these bodies, which complement the MCB, where representatives from the middleware community meet with representatives of users and operations, will be more precisely defined once this general proposal is ready for implementation to guarantee the availability and evolution of the UMD distribution and repository.

EGI.org and its technical bodies should be the single place in Europe where the requirements concerning the middleware for EGI will be planned and coordinated, in particular with respect to:

- Common baseline architecture.
- Full interoperability of existing services through standardisation.
- Validation and testing of the released services included in UMD.
- Increasing complementarities and specialisations of the included services.
- Adoption of application and operations requirements.
- Convergence and interoperability through the implementation of standard interfaces with Globus and other non-EU stacks.
- Definition of additional interfaces to allow independent development of higher level services.

It needs to be assured that the UMD software components are easily installed and configured. The goal of UMD is to make it as easy as possible for the NGI national resource providers to deploy, maintain and use the grid services that need to guarantee to the VOs teams a uniform access to their resources.

Another important objective for the EGI.org MU is to provide the necessary testing and certification of the services included in UMD to ensure seamless operation and interoperation of all the components included in UMD. This will also include provision of test suites for quality assurance and standard compliance validation of new or modified existing services. To ensure these functions in an efficient and effective way the MU is likely to rely on appropriate tools for software configuration, implementation and testing; the MU will have to make these tools available also to the middleware providers.

In addition, the MU will establish effective collaborations on an equal footing with players from outside Europe to promote the inclusion of extra-European services (like Condor and Globus) to be compliant with the same set of EGI rules.

The tasks under the responsibility of the EGI.org Middleware Unit are summarised in the below table which has been taken from the EGI Blueprint.

MW Tasks in EGI.org	FTE
Maintain and document processes and quality criteria common to all middleware providers.	1
Provide and support tools to enable and monitor the processes (such as configuration management system, bug and task tracker, wiki).	1
Define quality and conformance criteria that UMD components need to satisfy in areas such as security, performance, scalability, functionality, usability, interoperability, adherence to standards. Verify that accepted components are certified according to the agreed process and satisfy the quality and conformance criteria, specifically targeted against security vulnerabilities.	3
Maintain a repository of certified middleware components or references thereto.	2
Follow the daily execution of the strategic plan endorsed by the MCB. Promote the EGI participation in standardisation bodies.	1
Sum of Resources in EGI.org Middleware Unit	**8**

7.1.4 COMPONENTS AND SERVICES PROPOSED FOR INCLUSION IN UMD IN THE FIRST STAGE OF EGI

The main services developed by ARC, gLite and UNICORE that are extensively used on the current production infrastructures by thousands of researchers from many different scientific disciplines are summarised in the following table. The information has been originally provided by the OMII Europe project (see http://omii-europe.org/OMII-Europe/docs/DJRA20.pdf)), but has been updated since the publication of Deliverable D3.1. It is recommended that the maintenance and further development of these services continue to be supported in the EGI scenario.

OGSA capability	ARC	gLite	UNICORE
Security.AttributeAuthority	VOMS+SAML	VOMS+SAML	WS-UUDB, SAML-VOMS, UVOS
Security.Accounting	SGAS, APEL	DGAS, APEL	OGSA-RUS + UR
Data.Management.Storage	Smart-SE, dCache ARC Gridftp	StoRM, DPM	SMS
Data.Management.Transfer	FTS, GridFTP2	FTS, GridFTP	JMS, GridFTP
Data.Access.Relational			OGSA-DAI (future)
Data.Access.FlatFiles	ARC Caching	GFAL	TSI
Information.Model	GLUE, arcschema	GLUE	GLUE
Information.Discovery	OpenLDAP, WSRF	OpenLDAP	CIP
Information.Monitoring	NG-Monitor	GridICE, R-GMA	LLview, CIS, RSS
ExecMan.ExecService	Grid-Manager with A-REX (BES) or GridFTP interface	GT2 Gram, CREAM + BES	TSS, OGSA-BES
ExecMan.JobManager	ARC Client	WMS	XNJS
ExecMan.CandidateSetGen			
ExecMan.ExecPlannService			

Deliverable D3.1, “First EGI functions definition”, includes a more detailed description of the mentioned components.

7.1.5 COST ESTIMATES FOR MIDDLEWARE MAINTENANCE

An approximate evaluation of the effort needed by the three Consortia to maintain and support the existing middleware components and to adopt standards aimed at interoperability is 70 FTE. The estimation includes all phases of software preparation, from development to integration, full testing and packaging. Only the final conformance tests are under the responsibility of the EGI.org Middleware Unit.

The table below classifies the estimate of 70 FTE

	ARC	gLite	UNICORE	Total
Security	1	8	1.5	10
Data Management	5	9	1.5	15
Job Management	8	10	3	24
Information System	3	6	1	10
Other	4	5	4	11
Total	21	38	11	70

7.2 OUTLINE OF TIME EVOLUTION

The description of the Middleware function in EGI provided in this chapter and in the Blueprint refers to the first few years following the establishment of EGI. In the longer run, the middleware components should evolve into services that may be charged to customers and for which the maintenance and support may be more easily outsourced also to commercial partners.

8 USER COMMUNITY SERVICES

This section discusses the proposed organisation of the User Community Services. The first draft of these functions was referred to in D3.1 as "Application Support and Training" or "Extended Support Services". Following several rounds of feedback from the NGIs and from representatives of the user communities, and considering the intuitive purpose of these functions, the descriptions have been modified in parts and thus the entire section has been renamed **User Community Services** (UCS).

These services include activities such as:

1. **Gathering requirements** from the user communities and providing efficient channels for their **representation** vis à vis the middleware and other software providers.
2. Carrying out a review process to **integrate useful "external" software**, i.e. software packages that can help application developers use the grid infrastructure, but are not part of the core middleware distribution(s).
3. Establishing **Science Gateways** that expose common tools and services (e.g. workflow engines, web services, semantic annotation) in a transparent and user-friendly manner to user communities across various disciplines – see also http://knowledge.eu-egi.eu/index.php/Science_Gateways.
4. Establishing technical **collaborations with the large European reasearch infrastructure projects** (e.g. ESFRI) in support of customers of the European organisations.
5. Providing "umbrella" **services for collaborating projects** to streamline information management tasks and ensure some continuity of service between project cycles (e.g. maintenance of repositories, FAQs, wikis, etc.)
6. Maintaining a European Grid **Application Database** that allows applications to be "registered", enabling people to search for similar applications and contact the authors for guidance.
7. Organising European events such as **User Forum** meetings and **topical meetings** for specific user communities.
8. Providing **services for new communities**, e.g. "front desk" services, VO creation counselling, etc.
   See also http://knowledge.eu-egi.eu/index.php/UCS_Front_Desk.
9. Ensuring that user communities and grid administrators are provided with high quality **documentation** and **training services**.
   See also http://knowledge.eu-egi.eu/knowledge/index.php/Documentation_and_Training

The above activities are to be carried out mainly by the NGIs in the context of a structured network of User Community Services, under the coordination by a small team within EGI.org. Activities such as providing support to porting activities and training of users and administrators are typically delivered through NGIs, either on national level or via specific agreements with other NGIs in the context of the planned Specialised Support Centre (see below)..

For shortness sake, most of the corresponding content in D3.1 is not repeated here; where relevant (e.g. on topics such as science gateways or a possible front desk acquisition process) the reader is referred to the relevant parts of the EGI Knowledge Base at http://knowledge.eu-egi.eu.

8.1 THE EGI USER FORUM AND SPECIALISED SUPPORT CENTRES

The two main organisational entities in EGI to provide representation and services to the user communities are the **EGI User Forum** (**UF**) and the **EGI Specialised Support Centres** (**SSC**s).

The EGI User Forum was introduced in the final EGI Blueprint in response to requests by several NGIs and representatives from the user communities. It is a body specifically designed to provide representation to the various communities (e.g. bioinformatics, Earth sciences, new small communities, etc.).

The EGI User Forum (UF) is established by the communities themselves and is headed by a **Steering Committee** (**UFSC**) which interacts directly with EGI.org management and with the EGI Council via a **Chairperson**, who is a non-voting representative in the Council.

The User Coordination Officer (UCO) in EGI.org and the related UCS team are expected to interact with the UF at a central level.

Acting as the main managerial body to represent users in EGI, the UF will include representatives from any number of user groups – national, international, thematic, and "functional" (e.g. new and small user communities).

The **EGI SSCs** are also established by the user communities, as is any support centre in Europe. However, **in the context of EGI, an SSC is defined as a centre (or cluster of activities) that has a formal relationship with EGI**. The characteristics of this relationship are to be determined in the proposal phase of EGI; some initial suggestions are provided here to assist in the preparation of relevant proposals.

Each SSC is likely to have a User Forum Representative, whilst the UF Representatives (UFSC members) may not all be associated with an SSC.

The overall organisation of the EGI UF and the SSCs with their main interfaces in EGI is illustrated in Figure 2:

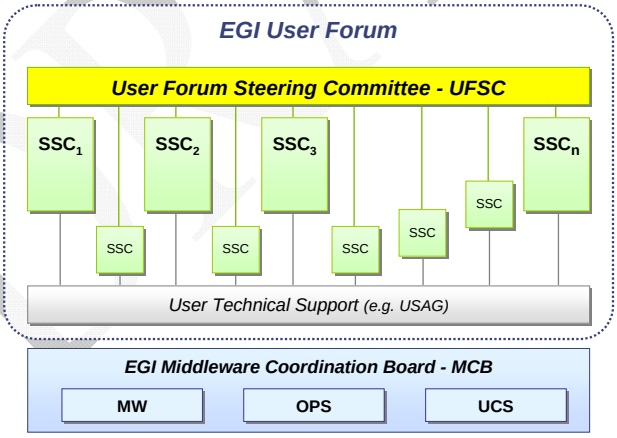

Figure 2 User Community Services in EGI

The SSCs will assist in collecting and transferring requirements and feedback from the user communities to EGI via a **User Technical Support** team covering the day-to-day technical needs in cooperation with the Operations Help-Desk team, and a **Grid Planning** team, which participates in the

EGI Middleware Coordination Board; the latter is responsible for more long-term technical planning and may establish other advisory committees to work with the EGI.org Director.

An SSC could also include **Front Desk** services, as described more in detail elsewhere. This option would be particularly recommended for an SSC dedicated to new communities. In addition, SSCs, as well as NGIs in executing their international tasks, will collaborate in several other interdisciplinary UC Services, including contributing to the **EGI Application Database**, the **External and UMD Candidate Software Review** (similar to EGEE RESPECT), the creation and maintenance of wikis, repositories, gateways, etc.

These tasks are considered proper EGI UCS tasks; the SSCs will be typically supported by EGI in executing them. However, SSCs will also be able to leverage support from their communities and collaborations from other relevant projects. The result would be a large SSC whose structure is illustrated in Figure 3:

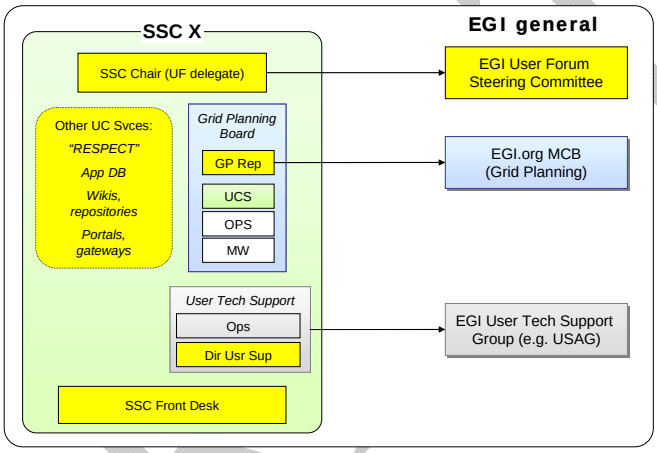

Figure 3 Example of a large SSC

There is no obligation for an NGI to support or endorse any particular SSC. SSCs will be encouraged to have a clear procedure in place to allow new members to join at a later date, or – if appropriate – to allow a community within an NGI to make partial use of its services, which would be properly acknowledged by the relevant NGI.

An SSC is expected to have European scope and visibility. In some instances, the SSCs will take over some functions from the existing EGEE strategic discipline clusters connected to specific international communities and could be highly structured themselves. While these centres are not required to structure themselves by a set template, it is assumed that some elements should be in place to facilitate communications related to EGI policy issues, such as participation in the User Forum Steering Committee, and the usage of the general technical services outlined elsewhere.

8.2 EGI SSC GUIDELINES

The set of guidelines described in this sub-section is an initial proposal aimed to provide some structure with respect to the UCS layer in an SSC. It is assumed that an SSC is governed by its user community. The guidelines below only refer to those aspects of an SSC which are relevant to its relationship with EGI.

Establishment of an SSC

During the proposal phase, the EGI partners (NGIs and experts) will make specific proposals for initial SSCs and establish the User Forum and its Steering Committee.

Once EGI is established, the EGI Council will be responsible for evaluating proposals for new SSCs in consultation with the UFSC. Details and formalities regarding this process are outside the scope of this document.

General Rights and Responsibilities

The SSCs are assumed to have a European-level existence. An SSC must have a cohesive community behind it that is able to take ownership of the SSC and to drive its evolution.

Most SSCs will be created around scientific domains. However, an SSC may also be created to meet specific “functional” requirements (e.g. a Training SSC, or an SSC for new and small communities).

The SSC will have representation in the EGI Council, Middleware, Middleware Coordination Boardgroup, and other appropriate bodies. SSCs are expected to feed their technical and non-technical requirements into EGI. The SSCs will also be able to interact with each other via the UFSC and other channels.

SSCs are expected to be "good citizens" of EGI and to follow the defined EGI policies (both security and operational policies).

The SSCs are expected to be (relatively) stable entities. The UCS layer in an SSC may be “re-adjusted” to meet the requirements of the user communities being served, and where an adjustment in personnel is requested, this will be negotiated with the UFSC and EGI Council. However, the SSC itself will evolve independently from this layer.

EGI may provide resources to the SSCs, in particular to support the central **UCS layer** of services (mainly manpower). EGI will also have a mechanism to provide seed resources for new communities, as discussed elsewhere.

SSCs will have access to resources. This includes mechanisms for making their own resources available and potentially (priority) access to centralised services, for example, help desks, central grid services, etc. This includes access to "community services" that are made available to the entire EGI user community such as operations support, middleware support, etc. SSCs will also have access to training and documentation.

SSCs will report facts and figures about their use of EGI in order to help EGI (and its funding agencies) understand the scope of the work accomplished with EGI.

SSCs are expected to operate transparently, allowing a clear view of SSCs activities and making the list of provided services available to the entire EGI community.

Typical UCS Personnel in an SSC

An SSC will have a high-level **User Forum Representative** who can nominate a deputy. For large SSCs, this individual is by default a member of the EGI UFSC.

An SSC will have a **Grid Planning Officer** who participates in EGI MCB meetings.
An SSC will have personnel for **User Technical Support** and similar tasks.
An SSC will have personnel for assisting with dissemination efforts and (web) content management.
An SSC might have a **Gateway Officer** to coordinate the development and maintenance of its Science Gateway.

Various user communities have already been approached, and several are producing hypotheses on specific SSCs. An articulated example is provided for Life Sciences in Appendix C. This case is also under review by members of ESFRI projects related to Life Sciences (in particular ELIXIR) who have graciously shared their material.

8.3 NGI INTERNATIONAL TASKS INCLUDING SSC TASKS

There is currently a projected budget of 9.9 M€ per year – equivalent to 110 FTE – for all NGI international tasks, including the SSC UCS layer. Initially approximately 50 FTE were supposed to be allocated to SSCs which provide continuity of service to the current scientific clusters in EGEE and other projects; however, it is recognised that depending on user demand, additional NGI international tasks might be clustered as SSC-related services.

At this stage, it is assumed that the international ("central") services will require the kind of effort described below. It should be kept in mind that the indicated ranges of manpower are deliberately broad to accommodate the heterogeneity of the communities' needs for UCS services. For instance, one may assume that in the short term there will be two very large SSCs (for Life Sciences and Earth Sciences) which may require 9 to 15 FTE (the latter also including activities with several collaborating projects) – hence the upper bound of the estimated effort for SSCs. Thus the global estimated ranges for manpower presuppose a rough division into "small", "medium" and "large" SSCs, as was done with other NGI international tasks.

Please note that the UCS tasks for the NGIs include only the NGI international tasks that are not otherwise allocated to SSCs. NGI national tasks are not included here; this is a departure from the global estimates at the end of Chapter 8 of D3.1, where an estimation was attempted for some common national tasks.

8.3.1 USER FORUM

Representation in EGI

The UF Representatives (UFSC members) represent their user communities at the EGI.org management level, collectively as an advisory body, and in the EGI Council, via the participation (as a non-voting representative) of the UFSC Chair.

Coordination

Coordination of all activities that concern a given user community. This includes common problems with the middleware, development of high-level services, political issues concerning SSC, funding, etc.

The coordinators, led by their UF representative, will interact "externally" with the EGI council, EGI.org administration, middleware coordinators, and (other) SSC coordinators. They must also provide information about the activities within their community or SSC to external parties and disseminate information from the external parties within their user community.

The main coordinator or UF representative is likely to be a researcher for scientific SSCs, as technical competence in the field is usually necessary to understand how grid technologies can facilitate scientific research.

The UF representatives are members of the UFSC and can count on the collaboration of the entire EGI.org UCS team.

Feedback

The UFSC should include members from all EGI communities; the process of gathering, evaluating and representing user needs is one of the primary purposes of this body.

User Conference – User Forum Events

Organisation of a large conference for users of the grid infrastructure.

This involves several parties, including the UFSC, the dissemination officers from SSCs and NGIs and other members of the UF, and is coordinated in EGI.org by the Event Organisation and User Forum Support team and the dissemination officer.

Each SSC is assumed to have a UF representative, plus often a deputy. These can be part-time assignments, but there need to be named personnel for this task.

Range of effort for 8.3.1		
SSC	**Other NGI International effort**	**Total**
8 ~ 12	4 ~ 7	12 ~ 19

8.3.2 DISSEMINATION AND EVENTS

Public Relations

Dissemination of grid activities/technologies within a particular (scientific) community. Make dissemination efforts and their results available to EGI. This typically happens through direct interactions between scientists and via the domain's conferences. Additionally, EGI should sponsor/organise meetings for specific scientific disciplines.

Direct interaction between SSCs and their user communities. Interaction with EGI to obtain funding for grid-focused meetings. Funds for organising meetings can be drawn from various sources; an SSC can also choose to "convert" its budget, currently expressed in FTE, into monetary funds. Logistical support for those meetings.

Most efficient dissemination relies on word-of-mouth promotion within a certain scientific discipline. Probably also needs general dissemination for general public or for new communities.

Events

The dissemination teams are also involved in the planning and organising of events such as the User Conference.

The UF is expected to coordinate dissemination efforts and event organisation. In large SSCs there may be some extra assistance in the region of 0.5 to 1 FTE (e.g. a Dissemination officer). It is auspicable that the central EGI.org dissemination team is given a certain degree of "central" assistance – perhaps on a rotating basis – by dissemination experts in the NGIs, for a total of 1 or 2 FTE

In addition, each NGI is expected to provide some collaboration on dissemination activities as an international task.

Range of effort for 8.3.2		
SSC	**Other NGI International effort**	**Total**
3 ~ 6	4 ~ 6	7 ~ 12

8.3.3 FRONT DESK / ACQUISITION AND SUPPORT TO NEW COMMUNITIES

The process of bringing new communities onto the infrastructure can be relatively simple or very challenging. In D3.1 a schema for a relatively large community was proposed, which is now available at http://knowledge.eu-egi.eu/index.php/UCS_Front_Desk, and is expected to evolve at that site. The following tasks are classified as front desk activities. Note the interaction of these activities with others within UCS and in Operations and Middleware, as well as with the new communities themselves.

This task is currently very labour-intensive and may require a "functional" SSC in the initial phase of EGI. However, the acquisition process of new communities is expected to become simpler as the infrastructure itself evolves.

Consulting

This service requires direct interaction with application developers to get their applications running on the grid infrastructure, and with middleware (and RESPECT software) developers to establish the best ways to use their services/APIs. It also requires interaction with other support activity personnel and with SSC leaders to identify "clients". In EGI.org, the UCS Consultant for new communities is expected to coordinate these activities.

The bulk of this activity is expected to rest with the SSCs and collaborating projects, and could require specialised effort from the hypothesised SSC for new communities.

Consulting teams typically interact with the personnel that manage informational resources, here referred to as Technical Coordination B.

Integration of Domain's Resources

The integration of a user community's computing resources with the grid infrastructure. This includes community data sources, standard computing services, and/or instruments.

This team interacts with middleware developers or with the application porting group if applicable. Direct interaction with their user community to understand what resources need to be interfaced to /made accessible from the grid. Interaction with NGIs regarding integration of hardware resources; expect EGI.org to provide contacts for NGIs.

The team must also have access to appropriate documentation and support to interface resources to the grid infrastructure. In turn, it will provide information to the community on how to access its resources via the grid. In the case of generic resources, this information should be shared with other communities.

Assistance for Application Porting

This service provides information concerning the porting of applications to the EGI infrastructure and on the integration of grid services with the application. It interfaces with application developers, either via GGUS or other fora (mailing lists, chat rooms, etc.) and with core middleware developers and

"integration software" developers (e.g. the RESPECT software) to establish how the software can be used effectively.

At the EGI.org level, this activity is considered a shared responsibility of the SSC coordination team and the Front Desk personnel, with assistance by the technical coordinators if required.

Specialised help desk personnel are involved in this effort.

It is expected that these activities will be strongly sustained by the SSCs, plus potentially by 2 or 3 general consultants drawn from the NGIs; an NGI can thus opt to devote some of its international UCS effort to this task. Alternatively, there may be a specific dedicated effort for this task in a potential Training SSC or an SSC for new communities. The effort table below assumes no such SSC, but this may be subject to change.

Range of effort for 8.3.3		
SSC	**Other NGI International effort**	**Total**
8 ~ 16	2 ~4	12 ~ 20

8.3.4 DOCUMENTATION AND TRAINING COORDINATION

As mentioned in the previous task, a dedicated Documentation and Training SSC might be proposed; thus the global UCS effort expected for this task may vary from current assumptions, which for the time being are that this entity is not in place.

Following is a brief overview of the task. For a more detailed discussion please see http://knowledge.eu-egi.eu/knowledge/index.php/Documentation_and_Training.

Documentation

Systematic review of documentation produced by entities within the project. Organisation (indexing) of the available documents along with some information about their quality and whether they are up-to-date.

Additionally, high-level documentation that treats the grid infrastructure as a coherent system must be produced and maintained. As this type of documentation is above any particular service, middleware developers cannot really be asked to provide it. Instead, EGI – either as part of the NGI international tasks or by means of a dedicated SSC - must employ technical writers to create this documentation and to keep it up-to-date. Support for multiple middleware stacks will complicate this task and is likely to require dedicated manpower for each of the supported stacks.

In EGI.org, the Documentation Review Coordinator is responsible for all these activities, with the assistance of the UCS: Technical Coordinator B and the Middleware Coordinator.

Training Coordination

Related to the high-level documentation is the creation of training courses targeted to 1) new users, 2) application developers, and 3) system administrators. Training is required by operations centres for system operators, by application developers who are developing programs to use the system and by users to allow them to access the services. Training is also required for trainers and educators regionally to assist them in disseminating experience of changes in the system which they are expected to subsequently pass on to their communities (local and in different user communities / VOs).

In EGI.org, the Training Coordinator is responsible for all these activities, with the assistance of the UCS: Technical Coordinator B and the Middleware Coordinator.

Since the EGI.org effort for this task is minimal, it is expected that the central team will require some assistance in the amount of 1 or 2 FTE.

In addition, larger SSCs may have effort dedicated for creating documentation specific to their user communities – in the region of 0.5 - 1 FTE. This does not include documentation for applications, which is outside the remit of the EGI project.

National efforts in this area are not considered here, but are assumed to be present.

Range of effort for 8.3.4		
SSC	**Other NGI International effort**	**Total**
3.5 ~ 7	1 ~ 2	4.5 ~ 9

8.3.5 TECHNICAL COORDINATION A – OPERATIONS RELATED

VO Registration and VO Database

EGI will have a central VO database with an interface for VO registration. These tools are part of the Operations portals described under Operations tasks O-E-4 / O-N-4, and under the responsibility of EGI.org Operations and UCS: Technical Coordinator A.

This service comprises such tasks as running the VO registration process, including providing support to VO managers and validating provided information; interface with VO managers for registration, with developers of other services regarding extent, format and access to registration information, and with operations to ensure comprehensive configuration information is provided. VO registration is performed in collaboration with Front Desk personnel.

Site Validation Tests

EGI will also ensure maintenance of a battery of reusable site validation tests in support of VO managers and VO members, and interfacing with Operations and Middleware personnel to assist VOs in using these tests. This service presupposes that the SAM infrastructure (or equivalent) will be available in EGI. It is also assumed that some mechanism (CVS, SVN, etc.) for versioning and maintaining code for the tests will be in place.

The creation, maintenance and availability of these tests will be ensured by a small central team consisting of Operations and UCS personnel.

The actual running of tests is the responsibility of the VOs, with some assistance in the UCS layer of the SSCs. Many actors do a certain degree of testing, but it is the SSCs to be responsible for detailed testing in close collaboration with middleware providers. It is expected that an SSC may have some dedicated effort for this task.

In EGI.org the central service is the responsibility of the Operations team and the UCS: Technical Coordinator A.

Core VO Service Provision

These services (VOMS, LFC, etc.) are already described in Operations O-E-14 and O-N-8. Their provision for all VOs running on the grid infrastructure should be guaranteed by EGI.org, but is likely to be actually run by various NGIs. The UCS: Technical Coordinator A will be responsible for working with the Operations counterpart within EGI.org to handle requests for the deployment of core services.

The NGIs must provide part of their international effort to sustain this task, which must be guaranteed by the Operations teams.

Help Desk and User Technical Support

The description of the Grid User Support and ticketing system is given in Operations tasks O-E-6, O-E-7, and O-N-6, O-N-7. These activities require a UCS element to interface directly with users of the grid infrastructure on issues of documentation and utilisation of the grid.

It is expected that the NGIs will provide this kind of service, in collaboration with their Operations counterparts.

It is also expected that large SSCs will have a User Technical Support unit which should include some (part-time) effort on the UCS side, providing help desk support focused on using community-specific software, services, data sources, etc. These activities would in any case use the common ticketing system to interact with users and other supporters.

This task is in part the responsibility of Technical Coordinator A and the personnel involved in SSC coordination.

Range of effort for 8.3.5		
SSC	**Other NGI International effort**	**Total**
4 ~ 8	3 ~ 6	7 ~ 14

8.3.6 TECHNICAL COORDINATION B – INFORMATION TOOLS

This activity is also involved with front desk activities, VO registration, and ensuring that informational tools are integrated with the Science Gateways. In addition, the task may include responsibility for designing and maintaining these gateways, in which case the relevant SSC must seek personnel with specific competences for the scientific field in question.

Case Studies

Providing written case studies of applications that have been successfully ported to the grid infrastructure. These serve as guidelines for future (similar) applications. Collaboration with application developers to port their applications to the grid to obtain case study material. Expect these studies to be made available through the EGI. The written case studies are a way to document application porting techniques and to provide a guide for future applications. Expect this to be done for each application receiving “consulting”; therefore this task interacts closely with the Front Desk teams.

This is the responsibility of the SSCs and the NGIs. At a central level, a depository or wiki should be provided, under the responsibility of Technical Coordinator B – see for instance [reference / url]

Application Database

A central database containing information about the applications running on the grid infrastructure. Serves as dissemination tool and as support resource. This should be the responsibility of EGI.org, centrally coordinated by UCS: Technical Coordinator B, possibly with some assistance by a dedicated person sought among the NGIs international tasks.

Expect end users and funding agencies to access the database, and contributions from the user community to provide information. The central team should interact with various user communities to understand what information is relevant and how to make database user-friendly and intuitive.

Any effort to populate the database is considered NGI international effort. In addition, it is hoped that there will be "external" effort from collaborating projects, in particular projects aimed at new communities.

Range of effort for 8.3.6		
SSC	**Other NGI International effort**	**Total**
4 ~ 8	3 ~ 7	7 ~ 15

8.3.7 GRID PLANNING

Development of Services

The EGI ecosystem expects the development efforts for high-level grid services and APIs (both generic and highly customised) in support of its user communities to continue. Within the EGI project there is a structured set of services for collecting and evaluating user requirements and transmitting them to development teams, for instance through the MCB and the RESPECT process, as well as via teams which assist with application porting.

However, there is no budgeted effort for actual development in the EGI project. SSCs are therefore expected to take on this responsibility, which should be properly acknowledged at the level of the EGI Council and the funding agencies.

SSCs will have access to appropriate documentation and support for interfacing new services to the core grid services; in turn, they should provide the relevant software to their user communities. If the services are generic, dissemination (and support) of services should also be offered to other communities.

At the EGI.org level, the SSC coordination team will oversee these activities to prevent duplications and to encourage sharing, in collaboration with the Grid Planning Coordinator.

Global effort for actual development is consequently 0; however, each large SSC should have a dedicated Grid Planning Officer to oversee the processes mentioned above, to participate in MCB meetings etc.

Coordination of Grid Planning

Much of the technical coordination between different disciplines currently takes place within the NA4 Steering Committee or through TMB working groups with strong NA4 participation. To avoid duplication and ensure a coherent evolution, this technical coordination must continue in the EGI era.

Expect to be able to raise issues with the infrastructure and to influence the priority for resolving them. Expect that EGI.org will provide the coordination/tracking of raised issues.

SSC coordinators should develop a consensus within their communities regarding issues and their priority. Provide funds for sponsoring these meetings and for inviting strategic stakeholders to attend.

Feedback

Providing feedback to EGI with respect to the middleware requirements, the utility of the services, operational problems, and administrative processes. SSCs will interact with their user communities and then provide collected information / experiences with the relevant coordinator in EGI.org.

This is part of the Grid Planning activities in collaboration with members of the User Forum. Each large SSC will have a UF Representative who is a member of the User Forum Steering Committee. 8 FTE are estimated for these.

Range of effort for 8.3.7		
SSC	**Other NGI International effort**	**Total**
4 ~ 8	2 ~ 4	6 ~ 12

8.3.8 SCIENCE GATEWAYS / PORTALS

For any structured scientific community, the grid is useful insofar as it provides added value to the work of that community – i.e. if the work is carried out in a manner that is easier, faster, cheaper, etc. This entails intuitive and user-friendly specific tools available to **a particular user community**. The one-size-fits-all model is not appropriate for end users who need to use their particular applications, have their particular language, and are accustomed to particular kinds of interfaces.

Hence the idea of Science Gateways, which supply these specialised services to specific communities and are built by (or in consultation with) individuals who are familiar with a specific user community.

In D3.1 an initial description of the purpose and organisation of science gateways was provided; more information can be found at http://knowledge.eu-egi.eu/index.php/Science_Gateways. It is strongly recommended that the SSCs design, build, and then maintain these tools for the benefit of their communities. These gateways can be initially simple, and evolve from currently existing portals, or they can be very sophisticated. It is in any case recommended that there be named specialised personnel (e.g. individuals with strong experience in content management and decision-making capabilities) in charge of the content and structure of these sites.

The creation of a gateway may initially take more effort than its subsequent maintenance; this effort, however, should never be 0. The table below gives estimates for the initial effort on the gateways.

Range of effort for 8.3.8		
SSC	**Other NGI International effort**	**Total**
6 ~ 12	0	6 ~ 12

8.3.9 OTHER NGI INTERNATIONAL TASKS

The estimated ranges of effort for the above tasks should not be added up without considering the discussion at the beginning of this section. However, it is expected that in line with the above guidelines for the tasks there is still a reasonable budget (perhaps in the order of 15 ~ 25 FTE) for cooperative NGI international tasks.

These coordination activities have not been speficially described in this document, as each NGI needs some flexibility in assigning its international effort, and as basic recommendations for each task have been given above.

Some cooperative effort is recommended to go toward the support for specific actions that are of interest to more than one large user community. In particular, both the Life Sciences and Earth

Sciences communities have a strong need to work in a focused manner on fostering the integration between grid and cluster computing / supercomputing.

In projects such as EGEE the regional coordinators are currently responsible for interfacing between regional support and corresponding centralised support teams. They also provide overviews of user activities within their region and act as first-line support.

Expect interfacing between users within a region and centralised support structures to continue. The regional coordinators will also continue to report about use of the grid within the region.

Tools and information to effectively make use of centralised services.

To report on activities within the region and to provide an efficient liaison between EGI.org and the regional user community.

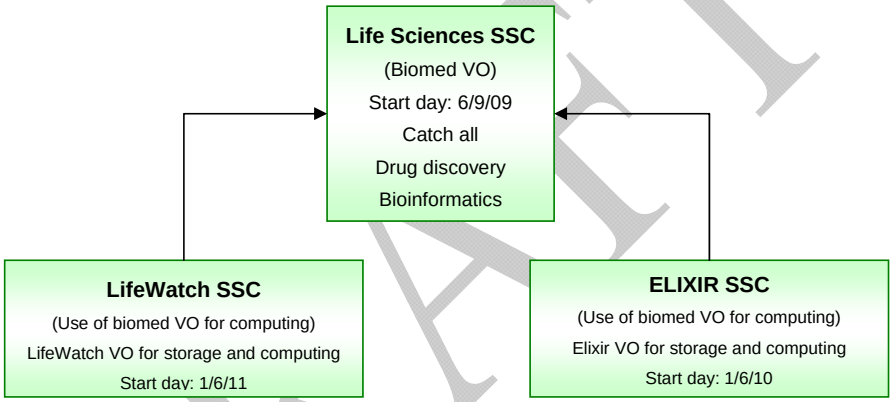

Figure 4

8.4 THE ROLE OF EGI.ORG

EGI.org will provide overall coordination for the services described above, structured as illustrated in Table 2. Aside from the activities that are carried out by senior personnel and therefore directly associated with two full-time employees, the estimated effort for the other activities are overall activity averages; event organisation, for instance, requires more than 2 FTE in certain periods and less in others and documentation-related activities are often performed in conjunction with coordination of SSC activities.

Table 2: User Community Services in EGI.org

Coordination of SSC activities A small team of coordinators to assist the User Coordination Officer (UCO) in all collaborative activities, such as (1)-(6) above, attend meetings, and work with the Grid Planning team to organise the representation of user community needs, new software etc. in EGI management.	2
Services for new and small communities & Front Desk coordination This will include a Consultant for new communities and a Front Desk Coordinator. These personnel oversee the availability of seed resources for new communities, and works with Grid Planning in analysing new trends in typology of grid users and new resources.	2
Event organisation & User Forum Support Two staff members in EGI.org will provide liaison and support for the activity of the UFSC, and coordinate the organisation of the main User Forum Events, plus others as needed in collaboration with their counterparts in the NGIs and SSCs.	2
Grid Planning & Technical Coordination A (help desk & other ops tools) One senior person to represent the UCS team in the Middleware Coordination Board and to liaise with any user committees that are established for technical representation and advisory activities with respect to the EGI Council and EGI.org management on behalf of their communities. One Technical Coordinator for all User Technical Support activities – e.g. Help Desk. Representatives of the international user communities will be members of the MCB who will steer, define priorities and provide feedback to the technical work programme of the EGI.org Director and the group of technical units in charge of the UMD component evolution and deployment.	2
Technical Coordination B (information tools) and documentation One Documentation Review Coordinator. One Technical Coordinator to perform activities related to technical information gathering, content and material creation, and support of central services such as material repository and online resources.	2
Coordination of training efforts Covering the activities in (9) related to management and coordination of training efforts in the NGIs and management of grid central services.	1
Sum of User Community Services in EGI.org	**11**

9. FUNCTION OF EGI: EXTERNAL LIAISON FUNCTIONS

9.1. TASK AND SERVICES

9.1.1. Dissemination

A small team within this function will execute the dissemination activities of the EGI.org. The team will focus on content production and coordinating activities. Technical and specific services will preferably be outsourced to third parties.

The objectives of the dissemination activities of EGI.org are:

- to ensure visibility and inform about EGI among decision makers, funding bodies, research communities, industry partners and other grid initiatives in Europe and in other parts of the world
- to inform the user communities and NGIs
- to arrange activities in collaboration with the NGIs
- to create and maintain excellent PR/media relations
- to coordinate publishing of activity and management reports
- to organise events such as EGI conferences and user forums

The dissemination activities need to be effective and well targeted. For EGI the dissemination activities at large must be executed both by EGI.org and the NGIs with a clear division of responsibilities. EGI.org will typically be in charge of tasks requiring coordination between NGIs. EGI.org will typically deal with common actions of the EGI while the NGIs are responsible for the EGI dissemination in their local and regional areas. It is important to note that in order to achieve good results, the dissemination team needs to act in close collaboration with the user-oriented, grid-operational and technical activities of EGI.

The dissemination team of EGI.org will serve as a horizontal link between the stakeholders (NGIs) and existing user communities, and has therefore a central role in maintaining the information flow to these parties. A dynamic and up-to-date website is a key element in maximising the visibility, providing support to users and stakeholders and informing about EGI. There is therefore a clear need for a professional and dedicated web editor.

The dissemination team of the EGI.org will support and coordinate the PR activities of the EGI. Press releases and Newsletters on the activity and key achievements will be published and widely distributed in order to increase visibility of the EGI. NGIs are expected to contribute by providing material to paper and electronic publications. The EGI.org will also be in charge of organising annual events and conferences, similar to e.g. the EGEE User Forum and the DEISA Symposium. These events not only increase the visibility and inform existing users, but also aim to broaden the user base. Exposure at other major events in Europe and beyond will also be coordinated and organised by the dissemination team of EGI.org, whereas NGIs are responsible for EGI representation at local and regional events. The representation may consist, for example, in a presentation, where the dissemination team would assist in identifying the right experts.

NGIs active on the international front are considered to represent themselves, but are of course free to propose coordination of any international activities with EGI.org.

In the initial phase of EGI, the core EGI.org dissemination team will be modest in size, but can be augmented by a rotation of 1-2 colleagues from the NGIs. The NGIs will be requested to provide a contact person for the dissemination activities within their own organisations.

9.1.2. Industry Take-up

It is recognised that sustainability of EGI would benefit from a persistent activity aimed at increasing participation of the private sector in the European grid infrastructure, which would bring additional competences and financial resources to the initiative.

As a publicly funded infrastructure dedicated to research, the usage policies will be determined not only by EU policies, but primarily by national law and policies. The usage policies can be expected to be comparable to those of other similar research infrastructures, such as the GÉANT network. Commercial usage is therefore limited, and usage by the business and industry sectors has primarily to be in form of research collaborations with European and national research institutes, universities and other educational institutions. The EGI.org management must develop a business model for the grid infrastructure, whose commercial potential is however forseen to be limited.

The general interest and potential use by industry can come in many forms;

- use of the EGI infrastructure in R&D (collaboration with the publicly funded research community);
- the EGI infrastructure as “state-of the-art”/”best practice” for industry;
- industry use of the EGI infrastructure for testing and learning;
- industrial projects with occasional exceptional requirements (critical computing on demand).

EGI.org is to initiate discussions with stakeholders to establish access policies for industrial research projects in the pre-competitive domain and for industrial production projects accessing innovative technologies or deploying innovative strategies. NGIs are expected to work along similar lines on a national level.

Following the recommendations of the e-Infrastructure Reflection Group (e-IRG) Task Force on Sustainable e-Infrastructures, industry has to be seen as both a potential user and a service provider. Today it is possible to identify an emerging business based on the major European grid technologies. EGI.org should welcome such initiatives and establish policies allowing emerging companies and other initiatives a fair competition in providing services for the EGI.

9.1.3. Other External Relations

External relations are defined as relations with organisations and initiatives outside EGI and of direct relevance for EGI in terms of collaboration or interoperation. Examples of such organisations and initiatives are:

- grids outside Europe
- commercial grids (e.g. cloud computing efforts)
- large-scale international research collaborations (e.g. the EIROForum organisations, ESFRI projects and WLCG)
- networking organisations (e.g. NRENs, DANTE, TERENA)
- policy and standard shaping bodies (e.g. e-IRG, ESFRI, OGF)

The EGI.org management, and specifically the Director, should be in charge of External Relations. This responsibility should primarily be focused on

- establishment of formal relations when necessary
- promotion of a common understanding on policies of grid interoperation
- influence on policy and standards shaping activities

- networking and enlargement of the EGI "sphere of influence"

The operational aspects in interoperation with other grids are handled by the EGI.org Grid Operations function.

The activity does not entail pro-active standardisation work, but handles the relations of EGI.org with organisations such as OGF, e-IRG and OASIS. EGI.org should consider membership in organisations like OGF and OASIS if deemed beneficial for EGI. The work could include coordination and reporting of participation in different standards working groups and interfacing with the technical teams involved in the actual standardisation. To maximise the outcome of the external relations activity, the EGI.org management should encourage synergies with external organisations and initiatives through the NGIs.

9.2. OUTLINE OF TIME EVOLUTION

As for other tasks, the description provided in this chapter refers to the first year of EGI. Such activities are however expected to be rather constant in time, with the very important exception of industry take-up, which is likely to start rather modestly, as a kind of feasibility study, but expexted to grow, and may in future also require some change to the EGI structure to better accommodate commercial partners.

9.3. EFFORT

Dissemination:

FTE estimation: 2 FTE for EGI.org and 0.5 for each NGI

According to the above analysis the following expertise is proposed:

- A dissemination manager – 1 FTE for EGI.org
- A web editor – 1 FTE for EGI.org
- NGI dissemination interface for EGI – 0.5 FTE for each participating NGI. As mentioned above, 1-2 of the NGI interfaces can also further staff the EGI dissemination team.

Industry Take-up:

FTE estimation: no additional manpower

It is proposed that the EGI.org Director and the management team cover these activities at the initial stage of EGI. The effort could increase substantially in the subsequent years once effective ways of collaborating with the business world are established.

Other:

FTE estimation: 2 FTE for EGI.org.

According to the above analysis the following expertise is proposed:

1) A policy and external liaison manager – 1 FTE for EGI.org
2) A standardisation liaison manager – 1 FTE for EGI.org

10. FUNCTIONS OF EGI: MANAGEMENT

10.1. EGI COUNCIL AND ITS MEMBERS

The main actors of EGI are the National Grid Initiatives (NGIs) which operate the grid infrastructures in each country and represent the requirements of their scientific communities together with resource providers and all e-Infrastructure-related institutions in a transparent way.

The **top-level management layer** in EGI is the **EGI Council**, constituted by the **NGIs** which accept the statutes. The NGIs govern EGI.org and voice their views on all EGI matters as voting members in the EGI Council. Other members of this body are the **Associate Members**, i.e. European institutions represented in the EIROFORUM or ESFRI, and **non-voting representatives** of non-European partner grid infrastructures. This representation is expected to be reciprocated, with the EGI Council being represented in the governing bodies of those partner grids.

The EGI Council may designate **committees** to work on topics specified by the Council. It may furthermore elect an **Executive**; details will be defined once the EGI.org statutes are finalised and the future EGI Council has voted on them. The **Director** and **Heads of Units** of EGI.org as well as the Chair of the User Forum Steering Committee will be *ex officio* Council members.

10.2. EGI.ORG AND ITS MANAGEMENT

The EGI.org full-time **Director** provides the organisational interface to the EGI Council, to funding and policy bodies (EC etc.) and to several EGI committees on the one hand, and to the heads of the EGI.org units on the other. For all internal and external activities, the EGI.org Director has an assistant. The EGI.org Director will be supported by a secretariat and by dedicated staff to prepare policy developments, representation on European level, and to support the EGI Council.

In EGI.org four permanent ***units*** are identified: the *Administration Unit* headed by the Central Administration Officer (CAO), the *Operations Unit* headed by the Central Operational Officer (COO), the *Middleware Unit* headed by the Central Technical Officer (CTO) and the *User Community Services* headed by the User Coordination Officer (UCO). The administration also includes staff to cover public relations, human resources, administrative and legal services.

Projects may, based on EGI.org's findings, be embedded in these units or they may be organised as a separate project-oriented unit within EGI.org, but need to be always embedded in the organisation's structure.

The following graph summarises the characteritics of the EGI.org **management structure**:

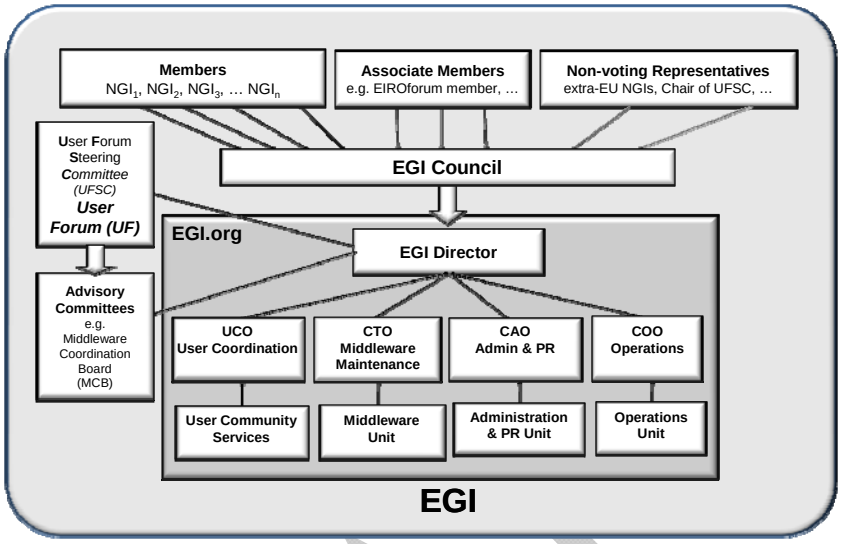

Figure 5 EGI Management Structure

The following table quantifies the management-related positions mentioned above:

Position	**FTE**
Director *Assistant to the Director* *Secretaries*	1 1 2
CTO	1
COO	1
UCO	1
CAO *Admin. Staff* *Legal expert*	1 2 1
Total (positions paid by membership fees)	**11**

10.3. EGI USER FORUM

The **user communities** will have representation and support mechanisms via the EGI **User Forum** (UF). The UF will organise an annual general meeting of all user communities to facilitate information exchanges at all levels. More details on User Communities Support and its structure are provided in Chapter 8.

User communities are represented in the **User Forum Steering Committee** (UFSC) through the respective SSC or – if there is no SSC - through their larger international grid-based projects. The **Chairperson** of the UFSC is an *ex officio* member of the EGI Council. The UFSC advises both the Council and the EGI.org Director on all matters regarding the involvement of users of the EGI e-Infrastructure.

At the management level, the SSCs will assist in collecting and transferring the requirements and feedback from the user communities to EGI through the Grid Planning team (see Chapter 8); the team participates in the EGI Middleware Coordination Board which is responsible for more long-term technical planning; it may also establish other advisory committees to work with the EGI.org Director.

10.4. EGI MIDDLEWARE COORDINATION BOARD (MCB)

The Middleware Coordination Board (MCB) is the EGI body that sets technical priorities and takes all decisions concerning the **maintenance**, **support** and **evolution** of the **middleware** deployed on the EGI e-Infrastructure; more details about the Middleware Support are provided in Chapter 7. The MCB is composed of representatives of the following areas, appointed in agreement with the EGI.org management:

- the main **middleware** developers of the components in use in the EGI e-Infrastructure (i.e. the three European Middleware Consortia);
- the **operations** function representing all operational requirements of EGI.org, NGIs and resource providers;
- the **User Community** Services (UCS) teams on behalf of the Specialised Support Centers, representing the various user communities organised in thematic disciplines.

11. BIBLIOGRAPHY

[1] *EGI BluePrint, http://www.eu-egi.eu/vision.pdf.*
[2] *Grid Operations Cookbook; EGEE-II Project Deliverable DSA1-5, Oct 2007.*
[3] *Draft Definition of the EGI Organization, EGI_DS Deliverable, Mar 2008*
[4] *The EGEE Network Operation Centre (http://technical.eu-egee.org/index.php?id=144).*
[5] *The EUgridPMA: coordinating Grid authentication in e-Science (https://www.eugridpma.org/)*
[6] *GGUS SUpport Model, EGEE and LCG Document N. 9100, Jul 2006.*
[7] *Treadwell, J.; Open Grid Services Architecture Glossary of Terms, GFD-I.120.*
[8] *The International Grid Trust Federation (http://www.gridpma.org/http://www.gridpma.org/)*
[9] *Grid Security Incident Handling and Response Guide, Open Science Grid, Nov 2004*
[10] OMII Europe, http://omii-europe.org/OMII-Europe/docs/DJRA20.pdf
[11] *D2.1. EGI consolidated requirements and use cases.*
[12] The SA1 Operational Procedures Manual: https://twiki.cern.ch/twiki/bin/view/EGEE/EGEEROperationalProcedures#1_SA1_Operational_Procedures_Man
[13] GGUS monthly reports (https://gus.fzk.de/pages/metrics/download_metrics_reports.php).

APPENDIX A: DESCRIPTION OF EUROPEAN PROJECTS SUMMARISED IN CHAPTER 3

A.1 INFRASTRUCTURE PROJECTS

A.1.1 EGEE

Building on investments from member states into national resources and thanks to the EC-funded EGEE project (Enabling Grids for E-sciencE), Europe has developed a scientific grid infrastructure for a wide range of research communities in and across many member states .

More than 250 sites in 48 countries contribute to the EGEE infrastructure which, at present, provides round-the-clock access to over 80,000 CPUs to communities across Europe, in areas such as Archeology, Astronomy, Astrophysics, Civil Protection, Computational Chemistry, Earth Sciences, Finance, Fusion, Geophysics, High Energy Physics, Life Sciences, Multimedia, Material Sciences; the infrastructure is serving over 10000 registered users spread across some 90 Virtual Organisations. In 2007, about 25PB of data were stored in disk and tape/MSS storage.

Peaks of 3.5 Mjob per month have recently been observed on the EGEE infrastructure, which corresponds to 115 kjobs per day. During 2007 20578 kSI2kyears of CPUs were used. The high-energy physics (HEP) community currently accounts for two thirds of the use of these computing resources, with the rest being used by researchers in other fields (see above). It is expected that the HEP community alone will increase the usage by a factor of 5 during the next year. Massive data transfer rates of up to 1.5 GB/s have already been reached.

EGEE III Effort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA1: Middleware Engineering SA3: Integration, Testing and Certification TNA5.3: Monitor EGEE contributions to standardisation activities	Total	52.8
		Funded	26.4
Operations	SA1: Grid Operations SA2: Networking Support	Total	189.9
		Funded	94.9
User oriented activities	NA3: User Training and Induction NA4: User community support and expansion NA2: Dissemination, Communication and Outreach	Total	121.7
		Funded	60.9
Global Effort		**Total**	**364.4**
		Funded	**182.2**

Project duration: 24 months
Project's home page: http://www.eu-egee.org/

A.1.2 BalticGrid-II

The BalticGrid Second Phase (BalticGrid-II) project is designed to increase the impact, adoption and reach, and to further improve the support of services and users of the recently created e-Infrastructure in the Baltic States.

This will be achieved by an extension of the BalticGrid infrastructure to Belarus; interoperation of the gLite-based infrastructure with UNICORE- and ARC-based Grid resources in the region; identifying and addressing the specific needs of new scientific communities such as nano-science and engineering sciences; and by establishing new grid services for linguistic research, Baltic Sea environmental research, data mining tools for communication modelling and bioinformatics.

The e-Infrastructure, based on the successful BalticGrid project, will be fully interoperable with the pan-European e-Infrastructures established by EGEE, EGEE-associated projects, and the planned EGI, with the goal of establishing a sustainable e-Infrastructure in the Baltic region.

The e-Infrastructure of 26 clusters deployed in five countries during the first phase of BalticGrid is expected to expand, both in capacity and capability of its computing resources.

The BG-II consortium is composed of 13 leading institutions in seven countries: 7 in Estonia, Latvia and Lithuania, 2 in Belarus, 2 in Poland, and one in Sweden and Switzerland, respectively.

The overall vision is to support and encourage scientists and services used in the Baltic region to conveniently access critical computing resources available to them both within Europe and beyond, and thus to faciliate effective research collaborations.

BalticGrid-II Effort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA1: Enhanced Application Services on Sustainable e-Infrastructure NA4: Policy and Standards Development	Total	4.25
		Funded	4.25
Operations	SA1: Grid Operations SA2: Network Resource Provisioning	Total	16.25
		Funded	16.25
User oriented activities	NA2: Education, Training, Dissemination and Outreach NA3: Application Identification and Collaboration SA3: Application Integration and Support	Total	16.87
		Funded	16.87

Project duration: 24 months

Yearly effort: PM 448; Annual budget: €1,499,000

Project's home page: http://www.balticgrid.org/

A.1.3 SEE-GRID-SCI

The South-East European (SEE) e-Infrastructure initiatives are committed to ensuring equal participation of a less-resourced region like South-East Europe in European e-Infrastructure trends. The SEEREN initiative has deployed a regional network, interconnected with the pan-European GÉANT backbone, and has established a regional grid infrastructure through the SEE-GRID initiative.

SEE-GRID-SCI leverages the SEE e-Infrastructure to enable new scientific collaborations among user communities.

SEE-GRID-SCI stimulates widespread e-Infrastructure uptake by new user groups across the region, fostering collaboration and providing advanced resources to a wider range of researchers, with an emphasis on strategic groups in seismology, meteorology and environmental protection. The initiative thus aims to have a catalytic and structuring effect on target user communities that currently do not directly benefit from the available infrastructures.

In parallel, in response to user demand, it aims to expand the regional e-Infrastructure by increasing the computing and storage resources and involving new partner countries in the region.

Finally, SEE-GRID-SCI endeavours to help consolidate national grid initiatives in the region, paving the way for them to be part of a longer-term sustainable grid infrastructure in Europe.

SEE-GRID-SCI Effort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA1 Development of application-level services	Total	3.6
		Funded	2.7
Operations	SA1 Infrastructure Operations	Total	9.6
		Funded	9.6
User oriented activities	NA4: User communities support NA3: Dissemination and Training	Total	12
		Funded	12

Project duration: 24 months

Yearly effort: PM 302; Annual budget: €1,014,443

Project's home page: http://www.see-grid.eu/

The following table summarises the effort for the infrastructure projects (EGEE-III, BalticGrid, SEE-GRID-SCI).

All Infrastructure Projects	**Effort in FTE**	
Middleware	Total	60.65
	Funded	33.35
Operations	Total	215.75
	Funded	120.75
User oriented activities	Total	150.57

(includes Application support, Training, and Dissemination)	Funded	89.77
Global Effort	**Total**	**426.97**
	Funded	**243.87**

A.2 DEVELOPMENT PROJECTS

A.2.1 ETICS

The objective of the ETICS 2 project (e-Infrastructure for Testing, Integration and Configuration of Software) is to offer a software build, test and quality assurance validation service across different infrastructures and to promote the widespread adoption of grid-based software engineering technologies by existing and new infrastructures. The approach consists in consolidating and expanding the availability, flexibility and efficiency of the existing ETICS services across those infrastructures, capturing commonalities and promoting open standards on software build, testing and quality assurance.

During the first phase of the ETICS project (2006-2007), a number of major challenges in the adoption of common build and test services across several projects were identified through the close collaboration with many projects and via dissemination events. The challenges can be summarised as follows:

- The lack of skilled personnel able to design and implement efficient validation test suites for complex grid scenarios associated to the complexity of the deployment and management of grid software.
- The lack of widely adopted validation procedures and metrics as well as a relative lack of trust at technological level between users and providers.
- The diversity of utilisation scenarios, the need for specialised validation methods and tools and the necessity of supporting emerging technologies and standards (e.g. IPv6).
- The lack of grid-aware, implementation-independent test design and workflow management tools.
- The complexity of setting up and managing secure, multi-platform validation testbeds.
- The dispersion of resources across multiple software repositories, third-party testbeds and private resources, which are not accessible through common mechanisms, making it difficult to share information and protect existing investments in legacy systems.

ETICS 2 addresses the identified challenges by providing a common software configuration model, multi-platform and language-independent build and test tools, an open repository of packages and reports produced during builds and test runs, extensible tools to collect software metrics, generate reports and monitor the overall quality and standard compliance of distributed software and a standard-based certification model (Grid-QCM). The focus is on maximising automation of the software development process from build to release, minimising the time and effort required to perform complex tests in realistic grid and distributed environments.

In addition, ETICS extends the availability of distributed build and test services to multiple infrastructures, allowing the automatic deployment of complex multi-node tests using the major European and international middleware implementations such as gLite, UNICORE and Condor.

ETICS 2 Effort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA1 - Testbed Management Technologies JRA2 - Test Management Tools	Total	4.6
		Funded	2.8
Operations	SA1 - Service Management SA2 - Infrastructures Support	Total	10.3
		Funded	7.7
User oriented activities	NA2 - Dissemination, training and certification	Total	2.0
		Funded	2.0

Project duration: 24 months
Yearly effort: PM 218.25; Annual budget: €1,336,000.00
Project's home page: http://www.eticsproject.eu/etics

A.2.2 OMII-EUROPE

The EU-funded Open Middleware Infrastructure Institute for Europe (OMII-Europe) has delivered a collection of re-engineered components that enable interoperability between key grid middleware platforms.

Components are selected for their potential in the field of interoperability: similar functionalities, availability and maturity of standards, open nature of the standard, etc.

In line with a service-oriented approach, the focus is on individual components rather than on full middleware distributions to prove that interoperability can be achieved even among completely different grid middleware architectures.

The final objective is to make the quality-assured re-engineered components available in a common repository to be re-introduced in their original middleware releases.

OMII-Europe started in May 2006 with 16 established partners from Europe, the USA and China.

OMII-Europe's primary focus is on the gLite, Globus and UNICORE platforms. Specific services were identified for re-engineering, such as job execution (BES/JSDL), data integration (OGSA-DAI), VO management (VOMS), accounting (RUS) and portal capability (GridSphere).

During the first year emphasis was on relationship-building among all internal and external partners, on participation in OGF and other working groups and on the design and prototyping of the components with the aim of delivering alpha versions by the end of the project year.

The second year saw the beginning of QA tests, ramp-up of training events, continued cooperation with partner projects and participation in standardisation events and the bulk of development leading to the delivery of final versions of all components by the end of the project.

OMII-Europe sets out to promote the definition and the implementation of open standards in all fields of grid computing. The project established the concept that standards are fundamental for the future of

grid middlewares and proved that interoperability can be achieved even between very different architectures.

OMII-Europe Effort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA1: Re-engineering of services JRA2: Identification of new services JRA3: Infrastructure integration JRA4: Benchmarking	Total	38.2
		Funded	17.3
Operations	SA1: Repository SA2: Quality Assurance SA3: Support	Total	14.9
		Funded	6.2
User oriented activities	NA2: Outreach and inreach NA3: Training	Total	3.4
		Funded	2.2

Project duration: 24 months
Yearly effort: PM 678; Annual budget: €3,174,191
Project's home page: http://www.omii-europe.org

A.2.3 GRIDCC / DORII

While remote control and data collection was part of the initial grid concept, most recent grid developments have been focusing on the sharing of distributed computational and storage resources. In this context, compute-intensive applications only have to use these grid elements in order to access an unlimited amount of computational power and disk storage. However, scientific and technical facilities provide concrete use cases where a strong interaction between the instrumentation and the computational grid is required.

The EU-funded GRIDCC project, launched in September 2004, provides a validated technology that can be deployed on top of existing grid middleware, exploiting grid opportunities for the secure operation and monitoring of remote instrumentation. EGEE gLite is the natural reference grid middleware for GRIDCC and the EGEE e-infrastructure is the natural framework to deploy and integrate this instrument grid technology on.

The goal of GRIDCC was to build a geographically distributed system able to provide access to and control of distributed complex instrumentation, ranging over a large number of diverse environments, from a set of sensors used by geophysical stations monitoring the state of the Earth to a network of small power generators supplying the European power grid. These applications rely on real-time and highly interactive operations of grid computing resources. To achieve this goal the project has pursued three main objectives:

- To develop generic grid middleware, based on existing building blocks (Grid Services), which will enable the remote control and monitoring of distributed instrumentation.
- To incorporate this new middleware into a few significant applications to validate the software both in terms of functionality and quality of service. These applications include, among others,

European Power Grid, Meteorology, Remote Operation of an Accelerator Facility, High Energy Physics Experiment.

- To widely disseminate the new software technology and the results of the application evaluations on the test beds, and to encourage a wide range of stakeholders to evaluate and adopt this grid-oriented approach to real-time control and monitoring of remote instrumentation.

GridCC Effort Table

	Project Activities	**Effort in FTE**	
Middleware	WP1: System Architecture WP2: Real-time and Interactive web services WP3: Grid-Enabled Instrumentation WP4: Brokering access to existing Grid resources WP5: Cooperative Environment (user-oriented?)	Total	21.7
		Funded	10.8
Operations	N / A	Total	0
		Funded	0
User oriented activities	WP6: Integration and Pilot Applications WP7: Information dissemination and exploitation	Total	15.7
		Funded	7.8

Project duration: 36 months
Yearly effort: PM 449; Annual budget: €1,763,000
Project's home page: http://www.gridcc.org/cms/

A.2.4 INTERACTIVE EUROPEAN GRID

The objective of the Interactive European Grid project is the deployment of an advanced grid-empowered infrastructure in the European Research Area specifically oriented to support the execution of demanding interactive applications. The Interactive European Grid, whilst interoperable with EGEE, will focus on the support for remote interactive collaboration and the reinforcement of the global framework for operation of virtual organisations for research projects in areas like biomedicine, astronomy, environment, physics, robotics, archaeology that are likely to benefit from being grid-enhanced. The initiative exploits the expertise generated by the EU CrossGrid project to provide researchers with interactive and simultaneous access to large distributed facilities through a friendly interface with powerful visualisation.

Project duration: 24 months
Annual budget: €1,318,500
Project's home page: http://www.interactive-grid.eu/

A.3 FIELD-SPECIFIC PROJECTS

A.3.1 BIOINFOGRID

Since the completion of the Genome Project, given the vast number of identified sequences, the problems associated with computational resources needed to process biological data have increased dramatically. Moreover, the amount of data continues to increase at a high speed as new technologies of high-throughput expression analysis create a continuous flow of information to be processed and interpreted. Furthermore, comparative genomics and genetic variation studies which employ modern analysis methods to identify genes in diseases create additional data challenges.

The BioinfoGRID project has successfully demonstrated the potential of grid computing for addressing the computational challenges the bioinformatics community is faced with. More specifically, BioinfoGRID has evaluated applications in the fields of Genomics, Proteomics, Transcriptomics and Molecular Dynamics, showing that data calculation times can be significantly reduced by distributing the computation across thousands of computers over the EGEE grid infrastructure. It is now possible to walk through the sequencing of the Human Genome and to study complex multigenic diseases, analysing in parallel thousands of molecular components.
However, the BioinfoGRID project has also identified limitations that still exist in terms of friendliness, completeness, robustness and standards compliance of the existing tools for the biological data access and management as well for the grid jobs submission, monitoring and bookkeeping. The project has also highlighted the need for continued dissemination activities to raise grid awareness within the bioinformatics community.

BIOINFOGRID Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N / A	Total	0
		Funded	0
Operations	N / A	Total	0
		Funded	0
User oriented activities	WP1: Genomics applications in grid WP2: Proteomics Applications in grid WP3: Transcriptomics Applications in grid WP4: Database and Functional Genomics Applications WP5: Molecular Dynamics Applications WP6: Coordination of technical aspects and relation with grid infrastructure projects, user training, application support and resources integration WP7: Dissemination and Outreach	Total	12.2
		Funded	9.3

Project duration: 24 months
Yearly effort: PM 146; Annual budget: €527,104
Project's home page: http://www.bioinfogrid.eu/

A.3.2 CYCLOPS

CYCLOPS brings together two important communities: GMES (Global Monitoring for Environment and Security) and grid, focusing on the operative sector and needs of European Civil Protection (CP). The main objectives of CYCLOPS are:

1) To disseminate EGEE results to the CP community, assessing the suitability of the EGEE infrastructure for CP applications. A variety of activities will focus on dissemination and outreach, training, workshops, possibly in conjunction with EGEE events, and on promoting a close collaboration between the two communities.

2) To provide the EGEE community with knowledge and requirements that characterise CP services. These requirements will also be used to assess the possibility for the development of an advanced grid platform to enable real time and near-real time services and to implement a security infrastructure very similar to defence systems standards.

3) To evaluate the possibility to utilise current EGEE services for CP applications, developing research strategies to enhance the EGEE platform.

4) To develop research strategies to enhance the EGEE platform, especially for Earth sciences resources.

CYCLOPS will contribute to EU policy developments, establishing liaisons and synergies with other existing projects and initiatives dealing with GMES, grid and complementary sectors, including PREVIEW, Risk EOS, RISK-AWARE, BOSS4GMES, EGEE Networking Activities and Application Support, e-IRG and INSPIRE. In fact, Consortium partners are involved in all these projects and initiatives.

Furthermore, CYCLOPS aims to address the OGF standardisation needs as far as the Earth and Space Science community, GMES and gLite are concerned.

In this context, the project actively contributes to the OGC (Open Geospatial Consortium) OGF initiative.

CYCLOPS Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N / A	Total	0
		Funded	0
Operations	N / A	Total	0
		Funded	0
User oriented activities	WP2: Coordination with EGEE activities WP3: Civil Protection System analysis WP4: research and Innovation Strategies definition WP5: Dissemination & Exploitation	Total	5.1
		Funded	5.1

Project duration: 24 months

Yearly effort: PM 61; Annual budget: €412,500

Project's home page: http://www.cyclops-project.eu/

A.3.3 E-NMR

e-NMR aims at deploying and unifying the NMR computational infrastructure in system biology, a project funded under the 7th Framework Programme of the European Union (Contract no. 213010 - e-NMR).

NMR plays an important role in life sciences (biomolecular NMR), and structural biology in particular, at both European and international level. Its overall objective is to optimise and extend the use of the NMR research infrastructure of EU-NMR through the implementation of an e-Infrastructure in order to provide the biomolecular NMR user community with a platform designed to integrate and streamline the computational approaches necessary for NMR data analysis and structural modelling (e-NMR). Access to the e-NMR infrastructure will be provided through a portal integrating NMR software and grid technology.

e-NMR Effort Table

	Project Activities	**Effort in FTE**	
Middleware	WP3: Design and development of the e-NMR Grid platform	Total	5.1
		Funded	5.1
Operations	WP2: e-NMR Grid deployment and operation	Total	1.4
		Funded	1.4
User oriented activities	WP1: Monitoring, Standardisation and Outreach	Total	1.8
		Funded	1.8

Project duration: 36 months
Yearly effort: PM 100; Annual budget: €922,217
Project's home page: http://www.e-nmr.eu/

A.3.4 ITHANET

Ithanet is a Euro-mediterranean network of research centres conducting molecular and clinical research into thalassaemia and related haemoglobinopathies. Participants of Ithanet include all major European research institutions active in haemoglobinopathy research and a number of collaborating partner institutions from non-EU Mediterranean and Black Sea countries.

The main objective of Ithanet's co-ordination action is to enhance the scientific potential of this research community by using e-Infrastructures.. It focuses a set of powerful e-infrastructure tools on the needs of researchers, clinicians, patients and the public, giving them the ability to carry out collaborative research, to pool resources, to exchange data and to disseminate research results efficiently and cost-effectively.

Using e-Infrastructure tools to consolidate and strengthen a research community with a specific geographic distribution and research topic, Ithanet strives to create new opportunities for high-impact collaborative research in the European Research Area.

Ithanet Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N/A	Total	0.0
		Funded	0.0
Operations	WP2: e-Infrastructure (collaboration tools)	Total	1.05
		Funded	1.05
User oriented activities	WP3: Tools for clinical research WP4: Tools for molecular research WP5: Training and knowledge transfer WP6: Portal WP7: Dissemination	Total	3.3
		Funded	3.3

Project duration: 24 months
Yearly effort: PM 52; Annual budget: €603,650
Project's home page: http://www.ithanet.eu/

A.3.5 DEGREE

A major challenge for the DEGREE (Dissemination and Exploitation of GRids in Earth science) project is to build a bridge linking the Earth sciences (ES) and grid communities throughout Europe, with particular focus on the EGEE-II project. An ES applications panel with a range of candidate applications suitable for porting to grid will ensure key ES requirements for porting and deployment on the grid middleware are identified, communicated and discussed within the grid community. At the same time, the DEGREE SSA will ensure the ES community is informed and updated on grid-related developments and potential benefits.

The results will provide feedback to the grid community and dissemination in the ES community will increase awareness of and involvement in grid developments.

To ensure ES requirements are taken into account in the next grid generation, DEGREE will initiate collaborations on various levels: at short, medium and long term via EU horizontal initiatives, specific collaborations with grid projects and participation in the e Infrastructure Reflection Group (e-IRG).

Objectives:

- Disseminate, promote uptake of grid in the wider ES community
- Reduce the gap between ES users and grid technology
- Inform ES users of grid benefits and its capability of tackling new and complex problems.

Project duration: 24 months
Annual budget: €670,000
Project's Home page: http://www.eu-degree.eu/

A.3.6 EUROVO-DCA

The concept of a Virtual Observatory (VO) consists in providing transparent access to the world's ever-expanding astronomical data through a standard interface which allows scientists to discover, access, analyse, and combine nature and lab data from heterogeneous data collections in a user-friendly way. A VO is a collection of interoperating data archives and software tools which utilise e-Infrastructures to form a scientific research environment in which astronomical collaborative research can be conducted.. Euro-VO is the European implementation of this idea that will produce a unified data and service resource (a data and service grid) able to perform complex data discovery and manipulation tasks across the whole range of astronomical research topics.

The Euro-VO Data Centre Alliance project will co-ordinate the national and European Agencies' Virtual Observatory initiatives, supporting the implementation of the Virtual Observatory framework by the European Data Centres to populate the Virtual Observatory with data produced by the European astronomy infrastructures.

Project duration: 28 months

Yearly effort: PM 72.9; Annual budget: €702,857

Project's home page: http://www.euro-vo.org/pub/index.html

A.4 INTERNATIONAL COOPERATION PROJECTS

A.4.1 EUCHINAGRID

Co-funded by the European Commission, the FP6 EUChinaGRID project officially started on 1st January 2006 with the aim to support the interconnection of the existing European and Chinese grid infrastructures and to enable their interoperability, thus creating a network of research collaboration between Europe and China.

EUChinaGRID provided specific support actions to foster the integration and interoperability of the grid infrastructures in Europe (EGEE) and China (CNGrid) for the benefit of e-Science applications and worldwide grid initiatives, in line with the support for the intercontinental extension of the European Research Area (ERA).

The project studied and supported the extension of a pilot intercontinental infrastructure using EGEE-supported applications and promoted the migration of new applications onto the grid infrastructures in Europe and China; this was achieved by training new user communities and supporting the adoption of grid tools and services for scientific applications. A set of existing Euro-Chinese collaborations in research with demanding computational needs was selected as pilot applications to validate the infrastructure.

During the 27 months of duration, the project achieved several goals.

The pilot infrastructure includes 12 sites, 5 of which are in China (4 in Beijing and 1 in Shandong). All relevant grid services were implemented and are maintained to facilitate the access of users and Virtual Organisations (VO) through the web portal (www.euchinagrid.eu). Some of these core Grid services are hosted in China.

Emphasis was on designing a fully interoperable e-Infrastructure, both horizontally (i.e. between European and Chinese middleware) and vertically (i.e. between grid middleware and the different versions of the IP protocol). Efforts towards both objectives led to promising results; furthermore, EUChinaGRID findings in this field raised interest amongst middleware developers in EGEE and

ETICS communities leading to joint activities such as a code checker for IPv6 compliance, implemented in the ETICS building system.

A Gateway between gLite and GOS was built and extensively tested and improved. The Gateway allows exchanging jobs between the two infrastructures addressing differences related to the Job Description Languages and the Security mechanisms.

Application deployment has equally achieved significant impact in several science fields:

- High-energy experiments (ATLAS and CMS) at the CERN Large Hadron Collider (LHC) can run their applications on the pilot infrastructure.
- Astroparticle experiment ARGO-YBJ, a joint collaboration between Chinese and Italian researchers, is currently collecting data on Cosmic Ray showers in the YangBaJing laboratory in Tibet; a complete system has been deployed to perform the data transfer from YangBaJing to IHEP (Beijing) and INFN-CNAF (Bologna) sites, using the EUChinaGRID Grid infrastructure deployed on the 2.5 Gbps link provided by the ORIENT project.
- EUChinaGRID also supports biological applications in the field of simulation and discovery of new proteins. The work in this field, carried out in the laboratories of the Biology Department of University of Roma Tre (UROM3), Jagiellonian University – Medical College (JU-MC) and Peking University (PKU), resulted in the first ab-initio protein structure prediction processes ever deployed in a grid environment. The parallel approaches adopted by UROM3 and JU-MC have been compared on a large sample of candidates (2x104), while the predicted protein structures are being experimentally verified by the PKU group.

EUChinaGRID undertook an intense dissemination activity with two website versions in English and Chinese and more than 300 Chinese researchers, engineers and students took part in the advanced knowledge tutorials held in China. A specific dissemination action was targeted towards the community of middleware developers to raise their awareness about IPv6 compliance and interoperability issues and to identify actions and best practices to overcome these problems. This included the delivery of focused workshops and tutorials that involved over 150 developers, and the development of a dedicated IPv6 website (http://www.euchinagrid.org/IPv6/index.html)and finally, the collaboration with related projects such as 6DISS.

EUChinaGrid Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N / A	Total	0
		Funded	0
Operations	WP2: Network planning and interoperability study WP3: Pilot infrastructure operational support	Total	20.08 (10 Non-EU)
		Funded	14.01 (6 Non-EU)
User oriented activities	WP5: Applications WP5: Dissemination	Total	37.64 (13 Non-EU)
		Funded	21.81 (5 Non-EU)

Project duration: 27 months

Yearly effort: PM 693; Annual budget: €577,777

Project's home page: http://www.euchinagrid.org/

A.4.2 EUMEDGRID

Funded by the EC and coordinated by INFN, the FP6 project EUMEDGRID aimed to support the development of a grid e-Infrastructure in the Mediterranean region and to promote the porting of new applications on the grid platform, thus allowing Mediterranean scientist to collaborate more closely with their European colleagues.

EUMEDGRID has raised grid awareness and disseminated competences across the Mediterranean and, at the same time, identified new research groups to be involved in the project, assisting them to exploit the enormous potential of the grid for improving their own applications.

The implementation and coordination of a grid infrastructure at a national (or even broader) level can be regarded, especially in the beneficiary countries, as an opportunity to optimise the usage of existing limited storage and computing resources and to enhance their accessibility for all research groups.

The EUMEDGRID project was conceived with this objective in mind and has succeeded in establishing a pilot grid infrastructure for research in the Mediterranean region which is interoperable and compatible with EGEE and related initiatives. Emphasis was on improving both the technological level and the expertise of networking and computing professionals across the Mediterranean, thus fostering the deployment of an effective Mediterranean grid infrastructure to support eScience. In fact, the two main objectives of the project were as follows: firstly, creating a human network in e-Science across the Mediterranean and secondly, addressing technical issues and supporting the implementation of a pilot grid infrastructure and applications in the area.

EUMEDGRID lasted for 26 months and made a considerable step forward during the second project year with a series of considerable achievements.

Cooperation among all the participants was demonstrated by the enthusiastic participation in joint workshops and meetings organised during the duration of the project and by the successful promotion of the creation of National Certification Authorities and National Grid Initiatives. Impressive results were also obtained in events fostering knowledge dissemination on grid technology and services. They attracted over 700 participants ranging from system administrators, researchers to end users. Feedback was gathered through dedicated questionnaires.

Promotion of National Grid Initiatives carried out in all non-EGEE partner countries registered a good level of success with programmes already operational in Algeria, Egypt, Morocco, Tunisia and Turkey and well advanced plans, with clear commitments, in Cyprus, Jordan, Syria and the Palestinian Territories. The project was very active in promoting the creation of national Certification Authorities designed to issue digital certificates to ensure secure grid access. The process is completed in Morocco, the first African country to become member of EUGridPMA, and is well advanced in the other countries, with a temporary catch-all CA in place to meet the needs of EUMEDGRID users.

A pilot grid infrastructure, composed to date of 25 sites in 13 countries, was set up during the project's duration.

Applications selected to run on the EUMEDGRID e-Infrastructure span several fields of interests: High-energy physics, biology and biomedical, hydrology, archaeology, seismology and vulcanology. New communities and applications of regional interest were also identified by means of a survey

based on web questionnaires[2]. Efforts to port the first applications on the EUMEDGRID e-Infrastructure began in the 1st quarter of 2006 with CODESA and ArchaeoGrid, a hydrological and an archaeological application, respectively, both pertinent to the Mediterranean region.

Various other applications were deployed during a dedicated event in Cairo, namely the first "EUMEDGRID School for Application Porting" (EGSAP-1 http://www.EUMEDGRID.org/egsap-1/) on 17-28 April 2007. Conceived as a full-immersion experience for selected new communities of regional interest, the school was deemed of paramount importance for the up-take of new applications on the regional pilot infrastructure. EGSAP-1 was consequently one of the largest dissemination efforts during the project, involving new communities in the project activities and supplying the required know-how to exploit the e-Infrastructure and deploy their own applications.

All selected applications were ported to the EUMEDGRID infrastructure. Moreover, these applications were also ported to the GENIUS web portal.

EUMEDGRID is however not confined to scientific issues, although the opportunity to port applications of regional importance, such as the hydro-geological and medical ones, on the pilot infrastructure is really exciting. Fostering grid awareness and the growth of new competences within scientific communities in Europe's neighbouring countries is a concrete initiative towards bridging the digital divide and towards a peaceful and effective collaboration among all partners.

At social level, e-Infrastructures can help mitigate phenomena such as digital divide and, possibly, revert brain drain, allowing brilliant minds in the region to make valuable contributions to cutting-edge European scientific activities which would effectively enlarge the European Research Area (ERA). Research and education networks and grids are fundamental infrastructures that allow non-EU researcher to excel in their home laboratories without the need to migrate to more advanced countries.

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

EUMEDGRID Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N / A	Total	0
		Funded	0
Operations	WP3: Pilot infrastructure operational support	Total	22.33 (15.09 Non-EU)
		Funded	12.1 (4.76 Non-EU)
User oriented activities	WP4: Application support WP2: Requirement capture and analysis WP5: Dissemination and Outreach	Total	26.57 (10.56 Non-EU)
		Funded	20.84 (8.36 Non-EU)

Project duration: 26 months

Yearly effort: PM 587; Annual budget: €759,231

[2] https://secure.um.edu.mt/EUMEDGRID/questionnaire/wp2/, https://secure.um.edu.mt/EUMEDGRID/questionnaire/wp4/

Project's home page: http://www.eumedgrid.org/

A.4.3 EUASIAGRID

The EUAsiaGrid proposal contributes to the objectives of the EU Research Infrastructures FP7 Programme by "promoting international interoperation between similar infrastructures with the aim of reinforcing the global relevance and impact of European e-Infrastructures".
The project's main goal is to pave the way towards an Asian e-Science grid infrastructure, in synergy with other European grid initiatives in Asia, namely EGEE-III via its Asia Federation, and the EUChinaGRID and EU-IndiaGRID projects and their follow-on efforts.
Exploiting existing global grid technologies, with particular focus on the European experience with the gLite middleware and applications running on top of it, the project plans to encourage federating approaches across scientific disciplines and communities.
EUAsiaGrid acts as a support action, aiming to define and implement a policy to promote the gLite middleware developed within the EU EGEE project across Asian countries.
Emphasis will be on dissemination, training, support for scientific applications and result monitoring. The use of the grid e-Science infrastructure is not only promoted on a geographical level, but also targeted towards new communities likely to benefit from this infrastructure, such as social sciences, disaster mitigation, building on the knowledge of more experienced fields, like high-energy physics and bioinformatics. The project intends to interact with standardisation bodies and other projects to help make the results sustainable over time.

EUAsiaGrid Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N / A	Total	0
		Funded	0
Operations	N / A	Total	0
		Funded	0
User oriented activities	WP2: Requirement capture and coordination policy definition WP3: Support of scientific applications WP4: Dissemination WP5: Training	Total	15.0
		Funded	13.1

Project duration: 24 months
Yearly effort: PM 180; Annual budget: €727,075
Project's home page: http://www.euasiagrid.org/

A.4.4 EU-INDIAGRID

EU-IndiaGrid is a European project that has established and currently maintains e-Infrastructure ties with the Indian generalised grid infrastructure. Among the partners of the project are the Indian NREN

(ERNET) and the Indian NGI (GARUDA). EU-IndiaGrid is formally endorsed by the Indian Government, as a letter sent to Ms Reding (EC) by the Indian Government Principal Scientific Advisor, Dr Chidambaram, is testimony of.

In addition to extensive dissemination and training activities, EU-IndiaGrid has set up a testbed running applications from several scientific communities, and has reported on its interoperation efforts in the context of many collaborative and standardisation/interoperability events. The efforts have established some specific requirements for interoperation between the European gLite middleware and the Indian middleware.

The project intends to collaborate with standardisation bodies and projects such as (a possible continuation of) OMII-Europe by implementing these requirements (either directly or in an effort mediated by EGI), and aims to consolidate the current EU-India Grid relationship.

EU-IndiaGrid Effort Table

	Project Activities	**Effort in FTE**	
Middleware	N / A	Total	0
		Funded	0
Operations	WP3: Network Planning Support WP4: Pilot grid infrastructure operational support	Total	4.7
		Funded	3.1
User oriented activities	WP5: Applications WP2: Building an eScience Network Community WP6: Dissemination & Networking Events	Total	8.9
		Funded	6.0

Project duration: 24 months

Yearly effort: PM 163; Annual budget: €640,410

Project's home page: http://www.euindiagrid.org/

A.4.5 EELA-2

EELA-2 aims at building a high-capacity, production-quality, scalable grid facility, providing round-the-clock, worldwide access to distributed computing, storage and network resources needed by the wide spectrum of applications from European-Latin American scientific collaborations, with special focus on:

- offering a complete set of versatile services fulfilling applications requirements;
- ensuring the long-term sustainability of the e-Infrastructure beyond the term of the project.

Such an ambitious project would not be possible without the prior existence of a consolidated e-Infrastructure, set up with the original intention to build a sustainable grid platform. This was the objective of the EELA project (www.eu-eela.org/first-phase.php) that provided its users with a stable, well supported Grid infrastructure based on 16 Resource Centres (RCs) with over 730 CPU cores and 60 Terabytes of storage space, thus proving that the deployment of a European-Latin American e-Infrastructure was not only technically viable but also demand-driven.

The EELA-2 vision is two-fold:

- consolidate and expand the current EELA e-Infrastructure built on the GÉANT2/European and RedCLARA/LA National Research & Education Networks (NRENs), to become an e-Infrastructure facility that provides a full set of enhanced services to all types of applications from a wide range of scientific areas in Europe and Latin America;
- establish the conditions for a sustanable e-Infrastructure, beyond the project duration.

EELA-2 Effort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA1: Development of Services for Applications and Infrastructure	Total	7.5
		Funded	5
Operations	SA1: Grid Infrastructure Service Activity SA2: Network Resource Provision	Total	31.5
		Funded	18.3
User oriented activities	NA3: Application Support NA2: Dissemination and Training	Total	17.0
		Funded	8.7

Project duration: 24 months
Yearly effort: PM 672; Annual budget: €1,284,160
Project's home page: http://www.eu-eela.eu/

A.5 DATA MANAGEMENT PROJECTS

A.5.1 D4SCIENCE

Co-funded by the European Commission's Seventh Framework Programme for Research and Technological Development, D4Science is one of the main European e-Infrastructure projects. The project started in January 2008, has a duration of 2 years and involves 11 partners.

D4Science aims to continue the efforts that the GÉANT, EGEE, and DILIGENT projects have initiated towards establishing networking, grid-based, and data-centric e-Infrastructures that accelerate multidisciplinary research by eventually overcoming barriers to heterogeneity, sustainability and scalability.

The main objective of D4Science is to lay the foundations for next-generation collaboration and knowledge-management environments by deploying an infrastructure that allows members of dynamic Virtual Research Environments (VREs) to create on-demand transient digital libraries based on shared computing, storage, multi-type content and application resources. Knowledge sharing and collaboration in a secure, coordinated, dynamic and cost-effective manner are to be the two major facilities offered by the combination of hardware, network, software and content elements that constitute the D4science infrastructure. Whilst the infrastructure is designed to support many different research and industrial applications, two specific communities have been selected to validate the project: the Environmental Monitoring and Fisheries and Aquaculture Resources Management management communities.

The objectives of the project will be achieved through the synergetic operation of Networking, Service and Joint Research Activities. The overall objective of the Networking Activities (NA) is to serve the needs of the communities. The experience gathered in connection with these large communities will facilitate a future extension of the e-Infrastructure capabilities to other scientific communities. This will be achieved by disseminating the project outcomes, training of the various players, and exploiting and collecting feedback of the D4Science e-Infrastructure through the implementation of the communities VREs.

The Service Activities (SA) aim at providing and maintaining a stable, reliable and usable e-Infrastructure to these (and possible other) D4Science user communities.

Finally, the Joint Research Activities (JRA) address the technical requirements raised by the Environmental Monitoring and Fisheries and Aquaculture Resources Management communities against the gCube framework.

D4ScienceEffort Table

	Project Activities	**Effort in FTE**	
Middleware	JRA4: gCube Development	Total	6.1
		Funded	4.6
Operations	SA1: Infrastructure Operation SA2: Community Specific Operations SA3: Software Integration, Testing and Distribution	Total	9.4
		Funded	9.4
User oriented activities	JRA1: Overall Planning and Development Coordination JRA2: Environmental Monitoring Community-specific Software Development JRA3: Fishery Resources Management Community-specific Software Development NA3: Communication and Dissemination NA4: Training NA5 Communities VREs Definition, Validation and Exploitation	Total	15.25
		Funded	10.75

Project duration: 24 months

Yearly effort: PM 200; Annual budget: €1,575,000

A.5.2 DRIVER

DRIVER is building the testbed for a future knowledge infrastructure of the European Research Area. Designed to be complementary to GEANT, the underlying network infrastructure for computing resources, data storage and data transport, DRIVER will deliver the content resources, i.e. any form of scientific output, including scientific/technical reports, working papers, pre-prints, articles and original research data.

The objective in a second phase is to establish the successful interoperation of both data network and knowledge repositories as integral parts of the e-Infrastructure for research and education in Europe.

The knowledge infrastructure testbed, delivered by DRIVER, will be based on nationally organised digital repository infrastructures, similar to GN2 and the NRENs. The successful DARE network in the Netherlands, recently presented to the public by the project partner SURF, will serve as a model to DRIVER.

DRIVER with its testbed is not intended to build a specific digital repository system with pre-defined services, based on a specific technology and needs of specific communities.

The testbed will in its inception focus on the infrastructure aspect, i.e. open, clearly defined interfaces to the content network, which allow any qualified service providers to build services on top of it. Like the data network GÉANT, DRIVER's knowledge infrastructure offers mainly a well structured, reliable and trustworthy basis. DRIVER opens up knowledge to the communities; it does however not prescribe how to use the knowledge.

Project duration: 18 months

Yearly effort: PM 244.7

A.6 POLICY AND PUBLIC RELATIONS PROJECTS

A.6.1 BELIEF

BELIEF's aim is to create a platform where e-Infrastructure stakeholders can collaborate, reach out to new communities and exchange knowledge, thus helping to ensure that e-Infrastructures are both developed and used effectively worldwide. It will be a one-stop shop for information on e-Infrastructure documentation and activities for both research and industry and will thus aid the knowledge transfer between them.

Project duration: 24 months

Annual budget: €604,226.5

Project's home page: http://www.beliefproject.org/

A.6.2 E-IRGSP

The e-IRGSP project provides a number of services to support the work of the e-Infrastructure Reflection Group (e-IRG), such as a secretariat (in The Hague, The Netherlands), a knowledge base and policy and editorial support. e-IRG comprises official government delegates from the 25 EU member states, as well as from associated countries.

Project duration: 24 months

Yearly effort: PM 22.5; Annual budget: €183,042

Project's home page: http://e-irg.eu/

A.7 OTHER PROJECTS

A.7.1 ICEAGE

At European level, e-Infrastructure has been identified as a key element for the creation of the European Research Area (ERA) to stimulate industry, improve the lives of citizens, accelerate research and gain international competitive advantage. This presupposes a diverse, knowledgeable and creative community able to effectively exploit e-Infrastructure.

With the support of the European Union, the ICEAGE project aimed at encouraging and supporting the incorporation of education in distributed computing in academic courses throughout the ERA. Built on EGEE, ICEAGE has enabled students and educators to obtain and develop grid education via sustained, large-scale, multi-purpose e-Infrastructures. ICEAGE differs from EGEE in that its primary goals are educational and therefore embraces a wide variety of approaches to e-Infrastructure.

ICEAGE has catalysed the necessary infrastructure and skills by establishing a worldwide initiative to inspire innovative and effective grid education. Grid education implies the use of education in the grid, but also the use of the grid in education. In the context of ICEAGE, the term "grid" is indeed used in a broad sense to include computing and communications technology, working practices, and policies that underpin e-Infrastructure.

ICEAGE Effort Table

	Project Activities	**Effort in FTE**	
Middleware	t-Infrastructure – development and provision (with several middleware co-existent)	Total	3
		Funded	2
Operations	t-Infrastructure operation (during Grid Schools)	Total	2
		Funded	1
User oriented activities	WP1 - Extend and Advance Grid Education – Grid Education Policy Development WP2 - Advanced Grid Education Support, Outreach, Induction & Training Services WP3 - Educational events and Summer Schools WP4 - t-Infrastructure – development and provision	Total	13
		Funded	9

Project duration: 24 months
Yearly effort: PM 216; Annual budget: €600,000
Project's home page: http://www.iceage-eu.org/

A.7.2 ISSEG

ISSeG aims to contribute to the consolidation of the European grid infrastructure in the field of computer security, by creating and disseminating practical expertise on the deployment of Integrated Site Security (ISS); this is to complement efforts undertaken within the Enabling Grids for E-sciencE (EGEE) projects Grid Security. ISS is a concept where all Site Security components (technical, administrative, educational) are developed in a coordinated fashion. The ISSeG vision is that Grid

Security, which focuses on inter-site security, middleware, and authentication, needs to be complemented by a comprehensive ISS strategy at every centre. The ISSeG consortium comprises three large scientific centres, namely CERN, CCLRC and FZK, which are all involved in EGEE.

The project objectives will be achieved by the creation and capture of raw expertise through full-scale ISS deployment at CERN and FZK, and by dissemination through the provision of applicable recommendations and methodologies for further ISS deployments.

Project duration: 24 months

Yearly effort: PM 102.5; Annual budget: €655,000

Project's home page: http://www.isseg.eu/

A.7.3 RINGRID

RINGrid provides an architecture which integrates scientific instruments in the e-Infrastructure and promotes a vision towards next-generation remote instrumentation systems. It encompasses the current state-of-the-art and near-future technology, delivers a conceptual design of missing architectural pieces to achieve such vision and assumes a grid environment and high-speed network interconnections.

Project duration: 18 months

Yearly effort: PM 123; Annual budget: €666,110

Project's home page: http://www.ringrid.eu/

APPENDIX B:

Formatted: English (U.K.)

B.1 NGI EFFORT ESTIMATION

The purpose of this section is to further elaborate on NGI effort requirements for Operations and Security activities in EGI. In particular, this document identifies the metrics to be adopted to estimate the effort needed for international tasks by NGIs of different sizes and with different levels of complexity and involvement in EGI. Various metrics can be adopted to estimate the size and complexity of an NGI, as shown in **Error! Reference source not found.**

Table 3: examples of metric for NGI size estimation

Category	Metric	Description
Size metrics	Number of production sites	Number of certified grid sites that are part of the NGI e-Infrastructure, exluding other sites just configured for testing purposes. The effort needed by an NGI to support site managers in the country and to monitor the infrastructure is somehow related to this quantity. This information is typically provided by the Grid Information services.
	Total computing power	Amount of computing resources made available to international VOs. The power of a computing device typically depends on the type of application it runs and can not be quantified univocally. It can be approximated by the total number of CPUs or cores provided and the related computing power, which can be estimated via a given benchmark of choice (e.g. SpecInt2000). Information about this metric is typically provided by the Grid Information services.
	Total disk and tape space	Amount of storage space made available to international VOs. Information about this metric is typically provided by the Grid Information services.
Operations metrics	Overall NGI availability/reliability	Overall level of service offered to the supported VOs, intended to be a composed function of the availability/reliability on various services, such as the grid technical services operated by the sites, the grid central services operated by the NGI, the grid operational tools, etc.
	Number of trouble tickets per NGI	The amount of tickets reaching final status in the EGI central ticketing system can be computed for every NGI to estimate the workload sustained in supporting users and site managers. This number should include both central and regional tickets.

Usage metrics	Number of users per NGI	The number of users served by a given NGI can be approximately estimated by gathering the number of supported VOs, and by calculating the distribution of users per country, according to the accredited Grid CA that released the corresponding personal certificate.
	Accounting records	The aggregated amount of (normalised) wall clock time/CPU time consumed by applications running in the region.

Operations metrics such as availability and reliability are of great importance, as they can have a significant impact on the amount of effort needed to operate a given NGI e-Infrastructure. For example, according to the requirements of the supported VOs, some NGIs may be interested in operating a highly reliable grid, while others may just provide a best-effort service. As to the former, the enforcement of high-level services requires manpower to run shifts, for grid oversight activities, etc.

To address this variety of needs, NGIs may be periodically requested to declare a *target* level for their services, according to which the effort requirements are estimated. During the transition period, discrepancies between the *target* and the actual service level delivered may be acceptable, in order to give all NGIs the chance to develop the required level of maturity.

B.2 METRIC ANALYSIS

B.2.1 Operations metrics: Number of trouble tickets per NGI

Tickets for an NGI obviously imply a management workload, but they can also be an indication of a poor level of service, and thus are not necessarily proportional to the size of the infrastructure operated, as it is currently the case in EGEE. A study has been performed to estimate the current distribution of tickets per country. Information about the number of *central* tickets (i.e. those tickets managed via the central helpdesk system in EGEE) was gathered from the GGUS monthly reports [REP][3].

Error! Reference source not found. shows the number of GGUS tickets for each EGEE ROC from January to September 2008. As can be seen, the number of tickets is not always related to the number of sites operated by a given ROC (e.g. this is the case of Italy and South East Europe). This can be attributed to a number of factors, such as the middleware used and the middleware update frequency in the region. Secondly, the reports used in this study only count the tickets managed via the central ticketing system (i.e. tickets limited to the regional scope are not included), so the numbers considered in this study may be partly incomplete.

[3] For Central Europe information of number of tickets was missing for two months during the time interval considered.

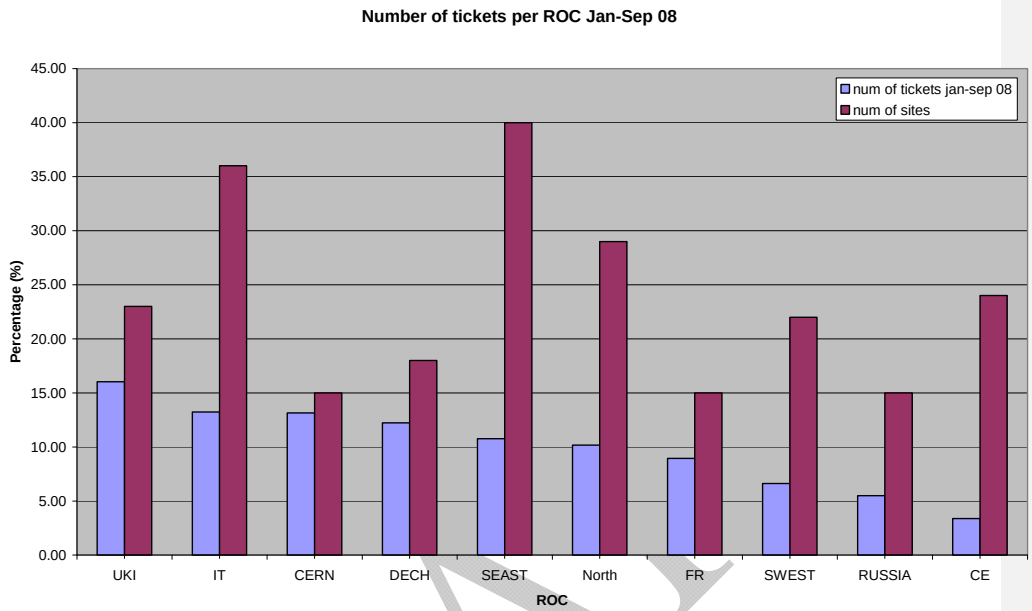

Figure 6: number of EGEE GGUS tickets per ROC (Jan-Sep 08), and comparison with the number of sites operated in each region

B.2.2 Usage metrics: Number of Grid users per NGI[4][5]

The distribution of grid users among countries involved in the EGEE, SEE-Grid and Baltic-Grid infrastructures was estimated during January 2009. The purpose of this analysis was two-fold: firstly, to get an idea of what type of procedure can be put in place to calculate the density of users per countryand secondly, to verify the level of correlation between the numbers of users, the amount of tickets handled in a country and its size.

25 production VOMS servers have been queried to get a list of registered VOs and the related users with certificates released by an accredited Grid CA. The number of VOs registered at the time of writing is 139, varying considerably in size.

For each valid VO in the list, a query was issued to extract the list of the member certificates registered. This list of users was subsequently reprocessed to:

[4] We are grateful to Andrea Ceccanti from INFN CNAF for his effort in performing this analysis, the results of which are herein described.

[5] We are grateful to Andrea Ceccanti from INFN CNAF for his effort in performing this analysis, the results of which are herein described.

- remove those VOs enabled for testing, demonstration activities, or past projects, and as such not used for production activities;
- remove certificates from non-accredited CAs (expired CAs, CAs for testing and demonstration activities, etc.);
- remove double entries in case of multiple certificates held by individual users;
- avoid double counting in case of a user certificate that is part of two or more VOs;
- remove the Kerberized CA of Fermilab (extensively used by the CDF VO), where a single user typically holds three or more different X.509 certificates;

As at the time of writing most of the accredited Grid CAs can be easily associated with a given European country, the distribution of users among countries was determined from the country associated to the respective certificate. The three notable exceptions are the CERN CA (arbitrarily assigned to Switzerland even if it is a European International Research Organisation, and many certified users hold both a certificate from the home institute and a certificate from CERN), and Baltic-Grid and SEE-Grid, which both include users from multiple countries.

CAs with less than 10 registered users were not considered, and only the top 30 VOs were included in the following diagrams.

Table 4: Grid users per country and per CA in descending order

CA ShortName	Country	Users	%
CERN Trusted Certification Authority	CH	902	14.42%
GridKa-CA	DE	897	14.34%
DOEGrids CA 1	USA	789	12.62%
INFN CA	IT	718	11.48%
GRID-FR	FR	497	7.95%
Baltic Grid Certification Authority	BalticGrid	398	6.36%
UK e-Science CA	UK	351	5.61%
IRISGridCA	ES	224	3.58%
NIKHEF medium-security certification auth	NL	214	3.42%
HellasGrid CA 2006	GR	199	3.18%
Russian Data-Intensive Grid CA	RU	144	2.30%
Academia Sinica Grid Comp. Certification Authority & Mercury	TW	231	3.69%
Grid Canada Certificate Authority	CA	116	1.85%
KEK GRID Certificate Authority	JP	93	1.49%
Polish Grid CA	PL	78	1.25%
gridca-cn	CN	63	1.01%
CyGridCA	CY	46	0.74%
IUCC	IL	45	0.72%
AEGIS-CA	RS	43	0.69%
SEE-GRID CA	SeeGrid	37	0.59%
TR-Grid CA	TR	36	0.58%
BEGrid CA	BE	33	0.53%

LIP Certification Authority	PT	31	0.50%
AustrianGrid CA	AT	24	0.38%
CESNET CA	CZ	23	0.37%
KISTI Grid Certificate Authority	KR	22	0.35%

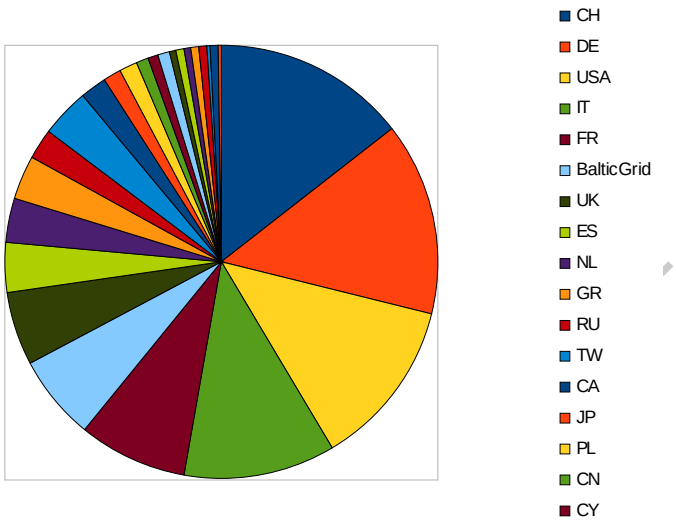

Figure 7: Users distribution per country/project

B.2.3 Size metrics: Number of production sites and Total computing power

The current size of the 36 NGIs operated in the framework of the main infrastructural grid projects in Europe (BalticGrid, EGEE, and SEE-Grid) has been estimated to verify to what extent size metrics can be composed to estimate their size today.

The current status of several NGIs shows that size cannot be just based on singe metrics such as the number of sites, or the amount of resources provided. For example, as shown in Figure 3[6], some countries provide a large fraction of CPU power, which is however concentrated on a relatively small number of sites (e.g. France, UK, etc.), while other countries feature a large number of sites and a small average amount of computing power per site (e.g. Italy, Greece, Russia, etc.). Consequently, it was attempted to estimate the NGI size according to a different formula.

[6] In this study, the numbers of usable/installed specInt2000 and of Tera Bytes were not considered, given the relatively large number of sites that – at the time of writing – did not publish and include correct information into the Grid information system.

The size of an NGI has been estimated by considering the overall amount of CPU offered and of the number of sites operated[7]. The *gstat* tool, which gathers information from the production grid information services, was used in this exercise as a source of input. Data was collected at different times during the second half of October 2008. The reference parameters were the *total* number of CPUs (C_{tot}) and sites (S_{tot}) and the respective number of CPUs C_{NGI}(NGI) and sites S_{NGI}(NGI) per NGI currently available in the EGEE, BalticGrid and SEE-Grid infrastructures. Only production sites have been considered. The size of an NGI was calculated according to this formula:

$$Size(NGI) = \frac{1}{2} [C_{NGI}(NGI) / C_{tot} * 100 + S_{NGI}(NGI) / S_{tot} * 100]$$

Note that at the time of writing, the number of sites belonging to NDGF – a site distributed across DK, FI, NO and SE – is equal to seven. Figure 2 illustrates the size distribution in Europe, estimated as the percentage of the overall size of EGEE, plus SEE-Grid and BalticGrid.

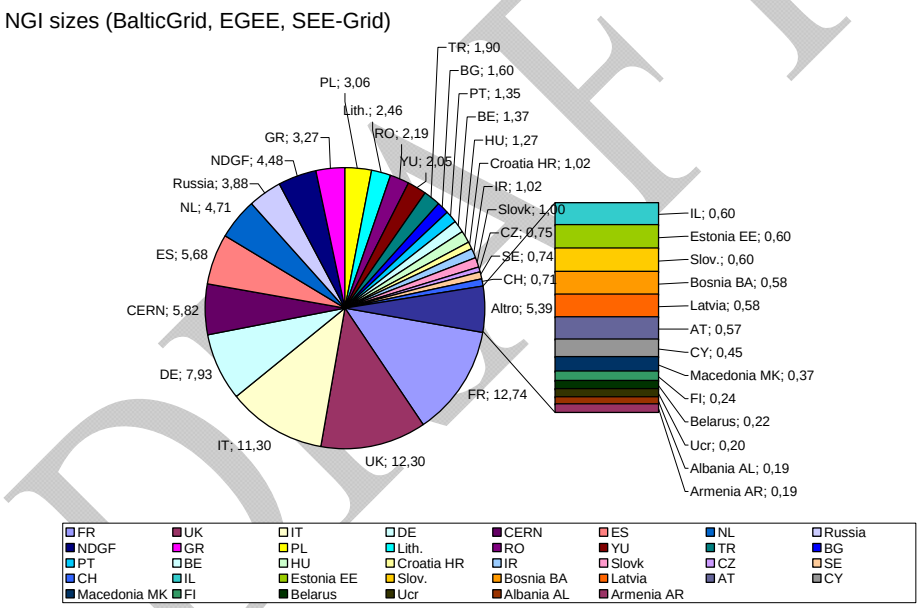

Figure 8: NGI size distribution (size %)

[7] The published number of CPUs often provides partial information as, given the widespread usage of multicore servers, the number of cores in production is a complementary information that needs to be taken into account to accurately quantify the amount of computing resources offered. To this extent, work is being undertaken in WLCG to review the current usage of the GLUE information schema to publish information about the installed capacity and to improve its usage. As to the number of sites, sites in temporary scheduled downtime and error status were included in the sum.

B.3 NGI EFFORT DISTRIBUTION: SIMULATION

By applying the NGI size estimation formula as illustrated above, a linear function to assign effort according to the NGI size is proposed; wherever possible, the result is compared to the current level of funding of NGIs in EGEE.

B.3.1 Current funding in EGEE-III

The following table summarises the level of Operations and Security funding per NGI provided to countries in EGEE III. For each NGI this amount has been computed from the overall amount of funding of the project, by calculating the proportion between FTE assigned to the activity SA1 and the number of FTE provided to the NGI for other project activities. Data were extracted from the EGEE Consortium Agreement v1.8. The funding level per NGI is compared to the number of CPUs offered and the number of sites operated (in this case excluding resources from BalticGrid and SEE-Grid).

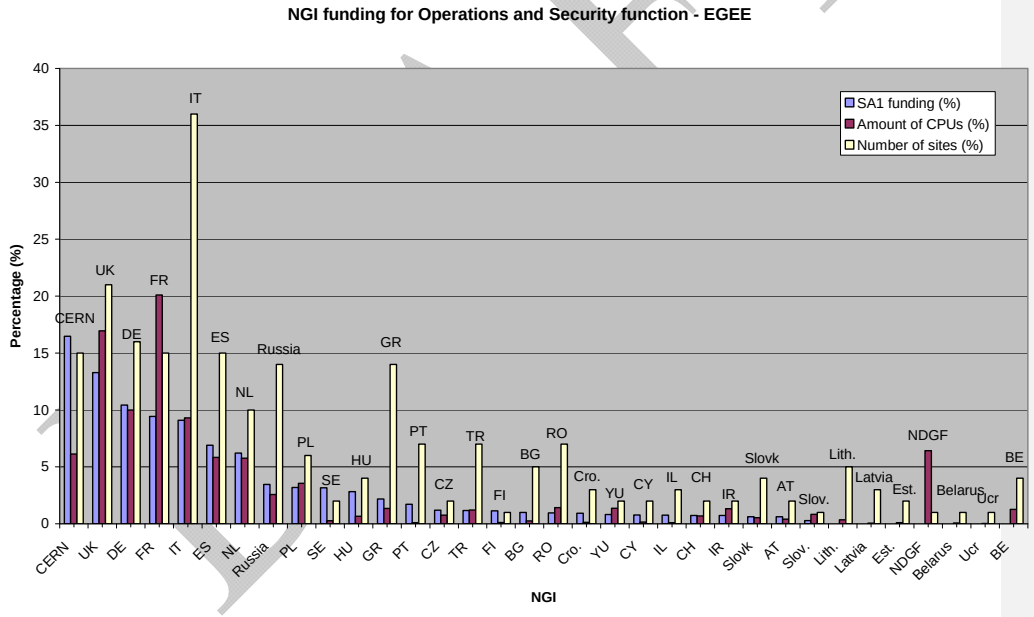

Figure 9: Operations and Security funding in EGEE-III

The following table provides detailed information on the current number of FTE funded by the EC for every NGI/EIRO in EGEE III.

Table 5: NGI/EIRO distribution of funded FTE in EGEE III for operation and security activities

NGI	TOT FTE
AT	1.54
Belarus	0
Belgium	0
Bulgaria	2.5
CERN	17.5
CH	1
Cipro	1.96
Croatia	1.96
CZ	2.41
DE	16.33
ES	13.2
Est.	0
FI	1
FR	18.75
GR	5.46
HU	1.58
Ireland	1.5
Israel	2.16
IT	19.5
Latvia	0
Lith.	0.00
NL	8.5
NO	0
PL	6.33
PT	4.16
Romania	2.37
Russia	17.66
SE	5
Serbia	2.29
Slovakia	1.37
Slovenia	0.66
Turkey	2.75
Ucr.	0
UK	15.5

B.3.2 Effort distribution

Based on the NGI size function previously illustrated (given by the combination of number of sites and amount of computing power provided), a linear function was defined to compute the NGI effort as a function of its size.

In this function the minimum number of assigned FTE was defined to be 2 FTE (for small NGIs), with a maximum of 22 (for larger NGIs). In particular, the number was calculated by multiplying the size with a coefficient (2 in our case) and by rounding up the result to the minimum larger integer. The effort was capped to 22 FTE or fixed to 2 FTE in case of out-of-range results. The resulting FTE distribution is illustrated in Figure 4.

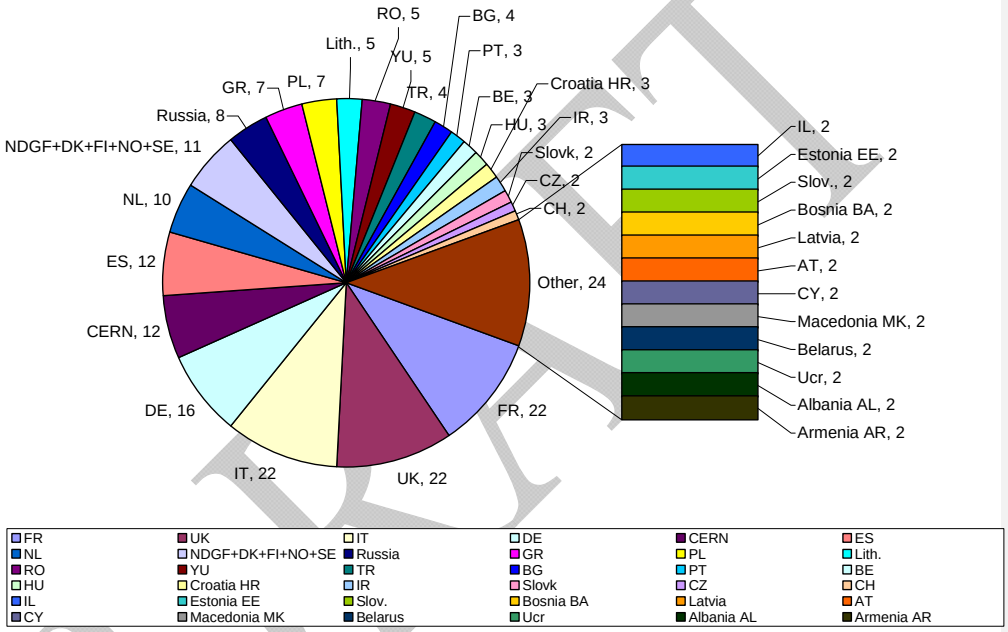

Figure 10: Simulated distribution of FTE among NGIs

B.3.3 Conclusions

The simulated FTE distribution provides results very similar to the current EGEE effort distribution, and to the distribution obtained from keys where manpower is assigned proportionally to the Gross Domestic Product of a country. EGI_DS plans to review the effort distribution function illustrated above to take into account additional metrics as discussed, and possibly new metrics according to the operational tools available in the future.

APPENDIX C:

C.1 LIFE SCIENCES SSC

The Life Sciences (LS) community is very large and heterogeneous and although at the beginning of EGI there will probably be just one SSC for this entire community, it is very likely that, as grid usage in the field evolves, more specialised SSCs may spin off from the original one.

There are several thematic clusters of LS grid users: Bioinformatics, Medical Imaging, Drug discovery and Health. The needs of these communities are fairly homogeneous with the exception of certain security requirements (anonymisation of data) for the Health group.

All LS communities have expressed a need to access **cluster grids and supercomputers**, and are thus very interested in the integration of these resources on the grid.

At the same time, there are about 8 ESFRI design studies related to life sciences which are establishing their own infrastructures – in particular sustained by the ELIXIR infrastructure – and need to be properly interfaced with EGI:

INSTRUCT	Integrated Structural Biology Infrastructure	www.instruct-fp7.eu
Infrafrontier	Infrastructure for Phenomefrontier and Archivefrontier	www.infrafrontier.eu
EATRIS	The European Advanced Translational Research Infrastructure	www.eatris.eu/
BBMRI	European Biobanking And Biomolecular Resources	www.biobanks.eu
ECRIN	Infrastructures For Clinical Trials And Biotherapy	www.ecrin.org
ELIXIR	Upgrade Of European Bioinformatics Infrastructure	www.elixir-europe.org

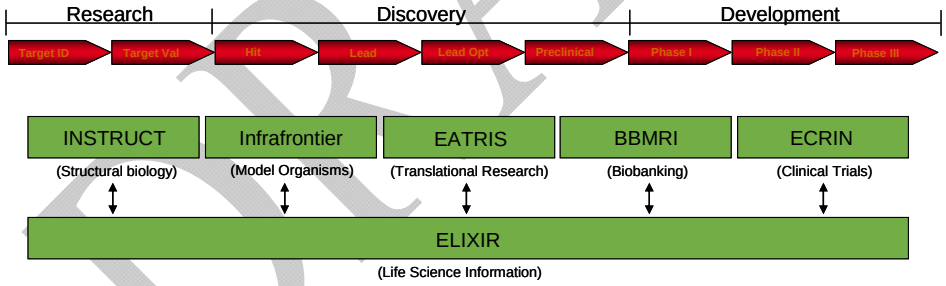

Figure 11: General schema of the planned ESFRI infrastructures. Image courtesy of A. Lyall (Elixir).

In addition to the needs mentioned above, the LS community expects more user-friendly interfaces and is prepared to dedicate specific effort to the design of a **Life Sciences Gateway**, which will serve not only to facilitate access and use of the grid by its growing communities, but also as a template for gateways in other disciplines.

Furthermore, the importance of international **standards** (OGF, Web Services) cannot be overestimated, as the community requires integration of resources into European infrastructures and initiatives, such as the ESFRIs and the Virtual Physiological Human [ref].

A possible scenario for this SSC in relation to the ESFRIs could be the following:

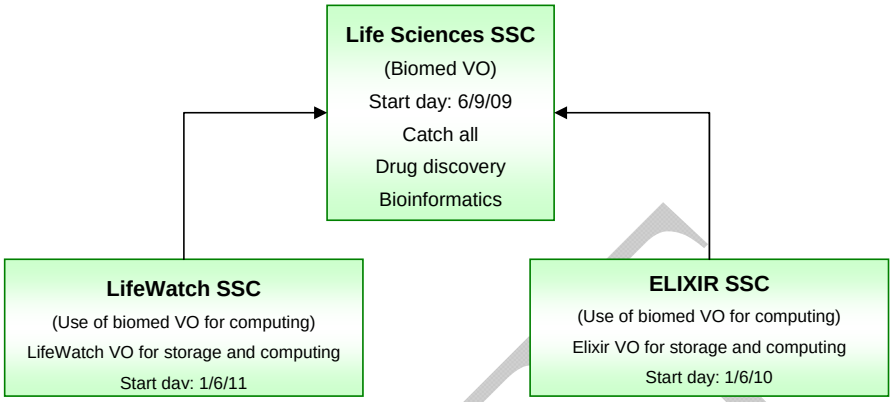

Figure 12: The LS SSC and two ESFRI SSCs

VO Structure

In terms of usage typology, the community can be divided into three categories:

- Research teams coming together for a limited time (at the national, European or international level) to constitute a project which pursues specific research objectives. This use case is very well discussed in the documents. These users are keen to have their own VO for the duration of their project and therefore also their own scientific gateway.
- Research teams involved in an ongoing scientific activity that require regular access to limited resources (equivalent to 1-2 CPU years) and from time to time, access to more advanced resources (equivalent to 10-20 CPU years). These users are the customers of the biomed VO that they join permanently. The biomed VO should have its own scientific gateway.
- The ESFRI infrastructures will comprise communities with well-defined long-term scientific topics and therefore specific requirements. Each ESFRI should have its own international VO, its own scientific gateways and be managed by the ESFRI collaboration board.

UCS Layer

The Life Sciences SSC proposes to structure its UCS layer services as follows:

1. 1 **User Forum Representative** and member of the UFSC; if possible, this individual should be assisted by at least a part-time representative for each thematic subcluster, for a total of 1.5 or 2 additional FTE. These individuals are also involved in coordination.
2. 1 **Grid Planning Officer** to participate in the EGI MCB meetings. Possibly assisted by the VO managers (not counted here).
3. 2 FTE for **User Technical Support** and other support tasks.

The personnel above will also collaborate on collecting, evaluating and representing the requirements of the LS community to the relevant bodies, namely the EGI Council, MCB, etc. The UF representative and GPO are also responsible for recommending specific actions (e.g. the creation of projects) for the development of new services.

4. A team of 2 FTE to design and build the **Life Sciences Gateway.**
5. 1 web **content manager** to work with the Gateway team and other information-related tasks.
6. 1 FTE for coordination of documentation and training specific to the LS community.
7. A team of 2 FTE for consultation with new communities, application porting support, and for maintainance and monitoring of the Application Database for Life Sciences.

The global UCS layer for the LS SSC is thus initially estimated to range between 9 and 14 FTE.