Co-ordination & Harmonisation of Advanced e-Infrastructures
Research Infrastructures – Support Action
Grant Agreement n. 260011
www.chain-project.eu

SEVENTH FRAMEWORK PROGRAMME
Capacities specific program - Research Infrastructures
FP7–INFRASTRUCTURES–2010-2

Project acronym: *CHAIN*
Project full title: Co-ordination & Harmonisation of Advanced e-Infrastructures
Grant agreement no.: 260011
Start date of project: 1 December 2011 Duration: 24 months

Deliverable D4.3
"Organisational study for interoperations of EGI with external grid infrastructures"

Version 1.0

Due on: 30th September 2012 Submitted on: 21/11/2012

Organisation name of lead beneficiary for this deliverable: CESNET
Dissemination Level: Public

Abstract:

This deliverable provides an overview of the current organization and operation models of EGI, with emphasis on to the interoperation with other (non-European) grid infrastructures. This is followed by a study of possible organizational structures for the interoperations and concludes with a proposal of an organizational model as suggested and already tested within the CHAIN project. The organizational model is proposed to also serve as a basis for future EGI plans on collaboration with external grids.

Document log

Issue	Date	Comment	Author
V0.1	30/06/2012	ToC	Ludek Matyska (CESNET)
V0.2	31/10/2012	Main body, missing Executive summary and Conclusion	Ludek Matyska (CESNET)
V0.3	03/11/2012	Modifications and new text	Federico Ruggieri (INFN)
V0.4	08/11/2012	Minor changes	Rafael Mayo (CIEMAT)
V0.5	08/11/2012	Few corrections	S. Qian (PKU)
V0.6	12/11/2012	New text and modifications	F. Ruggieri (INFN)
V0.7	12/11/2012	Few corrections	Margaret Ngwira (UbuntuNet)
V0.8	15/11/2012	New text and Executive Summary	F. Ruggieri (INFN)
V0.9	19/11/2012	Extended Section 4.3	L. Matyska (CESNET)
V0.9a	19/11/2012	Corrections and additions	O. Prnjat (GRNET)
V1.0	20/11/2012	Merge of versions	L. Matyska (CESNET)

Table of contents

1. Executive Summary

An analysis of the possible organisational models has been conducted over the last 23 months with several meetings that began with the brainstorming during the kick-off meeting in Rome (December 2010) and continued with another brainstorming at the EGI premises in Amsterdam the following January 2011. Refinements and updates of the discussed models and options were done regularly at the EGI conferences every semester, the last being held in Prague (EGI TF2012) last September. These meetings provided a platform for discussion on and establishment of the appropriate organisational model documented here and successfully supported and tested during the CHAIN project.

The document starts discussing the EGI organisational model and the current approach towards non-European Grid sites and NGIs.

After having analysed the state of the art and the available options two main models are then analysed: federated and peer-to-peer, with their respective advantages and weak points. The federated model explores the possibility to extend the existing EGI federated approach to other countries and NGIs providing clear advantages such as the strong collaboration to which any newcomer can adhere. The weak points are related to the potential high number of members that such an organisation can have with the risk of involving hundreds of delegates, logistics behind any meeting will be similar to the organisation of a large conference. The model, moreover, imposes the structure that fits well the European situation (with coordinating roles of the European Commission and Parliament), but may not be suitable to other regions and parts of the world.

The peer-to-peer model has some very strong points, such as that the model does not impose a common technical solution deployed globally at all levels; the level of integration (level of actual collaboration) can be adapted on a per-region (per-CHAIN-partner) basis without endangering collaboration elsewhere. Also, more formal structures can be introduced later, when their need arises - an example is a (weak) coordination of this peer-to-peer collaboration through one partner, with eventual rotation of this responsibility among partners (using as an example the rotating presidency of the EU). Naturally, there are also some weak points and risks related to the loose coupling of partners with the peer-to-peer collaboration that may slow down progress in areas where a common agreement is necessary and it may result in weaker interoperability (but this is a standard price to be paid for competitiveness among different technical solutions).

This organizational model better reflects the actual situation, where some regions already developed or are developing their own approach towards the distributed computing infrastructure and EGI already has MoUs in place for the actual collaboration. It does not contradict the federated organizational model more suitable for a single region.

The technical aspects still open are also discussed in the document and a proposal for an evolution of the current application level interoperability, based on Science Gateway, is also exposed.

2. Introduction

2.1. Purpose

This document presents and analyses different organizational models for interoperations of the EGI with external infrastructures. It starts with a brief description of the current EGI organizational and operational models that serve as a basis for further discussion. The major alternatives for possible organizational models are presented together with their strong and weak points; the peer-to-peer organizational model for the global collaboration is selected as the best framework. The organizational model, proposed by CHAIN project, is presented and discussed in more details in the 5th chapter, reflecting experience gained previously within the project (and described e.g. in the Deliverable D4.2).

2.2. Terminology glossary

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

EGI	European Grid Initiative
EGI-InSPIRE	European Grid Initiative-Integrated Sustained Pan-European Infrastructure
CNGrid	China National Grid
Gisela	Grid Initiatives for e-Science virtual communities in Europe and Latin America
NREN	National Research and Education Network
CLARA	Cooperacion Latino-Americana de Redes Avanzadas; Coordinator of Latin America NRENs
ROC	Regional Operation Centre
EELA	E-science grid facility for Europe and Latin America
EELA-2	E-science grid facility for Europe and Latin America 2
EUAsiaGrid	https://www.euasiagrid.org/
EUChinaGRID	Interconnection & Interoperability of Grids between Europe & China
EUIndiaGrid	Sustaining and enabling interoperability between Europe and India
ASREN	Arab States Research and Education Network Organisation
UbuntuNet	http://www.ubuntunet.net/
APGI	Asia/Pacific Grid Initiative
APAN	Asia Pacific Academic Network
TEIN	Trans-Eurasia[SQ1] Information Network
ASGC	Academia Sinica Grid Computing centre
SG	Science Gateway

A complete glossary of term used in conjunction with the grid infrastructure and EGI collaboration is available on the EGI InSPIRE project web pages
http://www.egi.eu/about/glossary/

3. EGI organizational and operational model

European Grid Initiative (EGI) is a collaboration and community of independent resource providers working with the specific research communities to create and maintain (operate) the European grid infrastructure, a distributed computing and storage facility. Its organization and operation models are briefly described in the following sections; more details can be found on the EGI web pages (http://www.egi.eu.) and in many EGI-InSPIRE project deliverables (http://www.egi.eu/about/egi-inspire/d-and-ms/index.html). EGI Collaboration is a result of a series of related projects and activities (series of EGEE projects, SEE Grid, EGI Design Study etc.), that each contributed to the current EGI status.

3.1. EGI organization and governance

EGI can be defined as a federation of nationally and domain specific resource providers coordinated by a dedicated unique organization: EGI.eu in Amsterdam. The EGI organization is based on a basic two-tier model: National resource providers are coordinated at the national level by individual National Grid Initiatives (NGI), each having a representative at the EGI (European) level. NGIs thus serve at the national level the same role as the EGI.eu at the European one. However, as we will see later, the NGIs can differ in the extent and scope of additional services that are not provided through EGI.eu. Also, international resource providers, such as those directly owned by the European International Research Organizations (EIRO) like CERN or EMBL, can be represented through the international organization they belong to.

NGI and international organization representatives together form the EGI Council, the supreme governing body of the EGI Collaboration. A strict rule of one NGI per country (or one representative per EIRO) is kept in order to leave any national discussion to the national level and be able to focus primary on the international collaboration and issues. EGI Council uses a weighted scheme, where the number of votes per a member relates to the membership fee that is linked with the GDP. The EIRO members can select their own membership fee (and therefore the number of votes they have), but a total number of votes owned by all EIRO members together is limited to be always less than the total number of votes owned by NGIs. The full membership in the EGI Council is restricted to the EU member states and also countries in accession to the EU or with specific relationship with EU (e.g. Switzerland). Apart from full members, EGI Council offers affiliated membership to the NGIs from countries that are outside the defining set. EGI Council can withhold the membership (or put it on hold, i.e. withhold the voting rights) in case the particular country does not pay the assigned membership fee.

EGI Council founded EGI.eu—a Dutch Foundation (or Stichting)—whose primary goal is to coordinate and manage EGI on behalf of its members. This is provided through a set of services, complementing and/or adding value to the activities developed and performed by NGIs and EIROs.

As the supreme governing body, EGI Council is responsible for defining the strategy of the EGI Collaboration and it also governs the EGI.eu, defining e.g. the policies of this organization. The EGI Council is represented by its chair who is elected for a two year term. As the EGI Council meets just few times per year, the day-to-day management of the EGI.eu is delegated to the EGI.eu director whose activities are supervised by the EGI Executive Board (EB). EB has 7 members, six are directly elected by the EGI Council and the seventh is the EGI Council chair who also chairs the EB.

EGI.eu focuses on services with added value in the know-how domain, not competing as a resource provider. The major four areas within the responsibility of EGI.eu are:

- Strategy and policy

EGI.eu focuses on the strategic planning and support for the execution of the agreed strategy, on the policy development and also on liaison with external partners. The work is supported by the EGI.eu Strategy and Policy Team (SPT), which also organizes meetings and workshops on strategic issues and it also coordinates and/or supports internal policy groups. Currently, the following policy groups are active within the EGI Collaboration:

- Technology Coordination Board (TCB)
- Operational Tools Advisory Group (OTAG)
- Operations Management Board (OMB)
- User Community Board (UCB)
- Security Policy Group (SPG)
- Software Vulnerability Group (SVG)
- Security Coordination Group (SCG)
- Computer Security Incident Response Team (CSIRT)

- Communication, events and outreach
  This service aims to widely publicise results of EGI and its individual members and partners. The communication team prepares documents and news for different target audiences: user communities (existing and new), journalists, general public, collaborating projects, but also decision makers, funding agencies and governmental representatives. EGI.eu uses wide communication means (website, wikis, social media channels, …) and also participation at events. The activity also covers organization of the major EGI community events:
  - The Community Forum in Spring (March/April);
  - The Technical Forum in Autumn (September).

  This activity also supports knowledge transfer between user communities, helping especially new users to become acquainted with the expertise and tools accumulated and developed by EGI, its predecessors and also collaborating activities.
- Operations, management, and technology provision
  This is in more detail discussed in the next section
- Unspecified activities (the Virtual Team framework)
  The Virtual Team framework is a pool of distributed competences within the EGI Collaboration (i.e. the competences mostly available in individual NGIs or other EGI Council members). It serves as a mechanism to support interaction and collaboration on non-technical matters, focusing specifically toward better work with existing research communities, outreach to new ones, new approaches to policies, strategy preparation etc. Its primary goal is to complement the operational activities within EGI Collaboration by more far sighted or more risky activities to keep the EGI Council participants involved also outside the strict (production quality) operations. There are no permanent Virtual Teams, each VT has a clearly defined goal and time of duration. Each NGI can propose new VT and if the proposal gets enough attraction among other NGIs (usually at least three NGIs should be involved), it gets some (human) resources and starts its work. This model encourages activity among NGIs and minimizes a risk that some important non-operational issue is left over due to insufficient human resources at the EGI.eu level.

The responsibility of the organization (of national resource providers and eventually other national participants) at the national level lies fully with the NGIs. This means, apart from the "one NGI per a country" rule no other requirements are imposed on the NGIs. Individual countries selected organizational models that best suits them—in some countries the NGI is affiliated with the NREN, in other countries they are part of some other organization and in some NGIs are independent entities. Also, while EGI.eu does not own any grid resources, NGIs in some countries are moderate or even important resource providers themselves, meaning they also execute direct operational responsibilities on their resources, not only

coordinating resources provided by other bodies. The model is flexible enough to accommodate practically any organizational model each individual NGI deploys. It is also able to deal with NGIs at different maturity level, without endangering the overall EGI Collaboration. However, fully developed (mature) NGIs usually deploy at the national level all the services EGI.eu is running, plus some country specific (there is a natural difference between the internal organization between small, medium, and large countries) and eventually also the roles and responsibilities of service providers. Additionally, the actual training of technical staff and users is performed at the national level.

3.2. EGI operations

EGI uses Platform—*a composition of IT infrastructure and IT services that together enable independent solution providers to build other technologies or processes, or both, on top if it*— as primary architectural layers for its operations. Two distinct types of platform are used:

- The **EGI Infrastructure Platform** that provides a set of core infrastructure services, namely
  - Operations Portal—provision of status information
  - Monitoring Portal—through the Service Availability Monitoring continuously tests the performance of the infrastructure and its services
  - Accounting Portal—provision of a homogeneous view of the usage data and resource usage
  - Helpdesk—support to users and operations staff
  - Configuration Database (GOCDB)—provision of general information about resource centres
  - Metric Portal—metrics collection and visualization
- The **EGI Collaboration Platform** that extends the infrastructure platform and includes services that are not critical for the primary EGI operations and are domain neutral (common across user communities). These include
  - Training marketplace—training coordination among countries, communities and projects
  - Application database—provision of information about computing tools
  - Requirements tracker—a ticketing system, to keep track of requirements
  - Science gateways—a community-specific set of tools, applications and data integrated in a web portal
  - Workflows and workflow systems—helping organize computations in a reusable way

Recently, EGI started introducing cloud interfaces to some of the available resources as a result of increased demand for more flexible resource provisioning. A **Cloud Infrastructure Platform** is being developed to provide the capabilities required by the future EGI cloud service.

Through the concept of platforms and their services, EGI is becoming able to provide its operations tools in a reusable way to address typical resource and technology independent needs of other grid infrastructures. The platform tools are available in a variety of configuration options, ranging from a simple provisioning as open-source software for local independent deployment through running these tools in a Software as a Service (SaaS) fashion to opening access to the deployed production instances within the EGI infrastructure. Use of these tools by external grid infrastructures (in any mentioned fashion) allows easy integration with existing EGI production services, making interoperation an easy achievable goal.

At the lower layer, EGI operations rely on the UMD middleware, provided by EMI (European Middleware Initiative) and other independent external providers. The UMD middleware is based on a set of standards that define interfaces necessary for a proper interoperability between grid infrastructures.

4. Organizational models for external collaboration

EGI is a federation of primary resource providers explicitly coordinated by a dedicated body—the EGI.eu organization. In this chapter, we discuss strength and shortcomings of this model when applied to the global collaboration and will also present a different approach to the EGI external collaboration. In this context, by "external collaboration" we mean any process that allows include into the decision process and integrate at the operational level any new resource provider not covered (usually by the country) through any already existing partner of the EGI Collaboration, i.e. EGI Council member. The main purpose is to allow resource providers and international organizations outside of Europe to somehow "join" EGI and to allow users to seamlessly use all and any available resource. The ideal state is to have neither organizational nor technical barriers for the truly global grid infrastructure.

4.1. Federated model

The federated organizational model used by EGI is open to new members of the collaboration. While the current EGI Council status restricts the membership to an explicitly defined group, there is no restriction to the affiliated membership. The federated model itself does not have any limits on the number of partners (members), so it is conceivable to speculate about simple extension of the EGI federated model (including change in the current status) through a modification of the full membership to cover wider areas than the currently EGI does.
Without going into details of the necessary EGI Council status, let's analyse the resulting structure.
There are clear advantages in going through a simple extension of the EGI federated model:

- There is already a strong collaboration to which any newcomer can adhere.
- The process can be adapted to the maturity (preparedness) of each new partner, e.g. through several membership levels.
- As the EGI operations share a common operational infrastructure (or at least its basic elements), there are clear rules that each new member must fulfil at the technical (operational) level. This guarantees seamless and direct interoperability.
- The coordinating organization is responsible for defining a common view of the whole collaboration, providing clear inputs into strategic decision made at the Council level.
- Through extended membership, the EGI Collaboration will be able to collect higher sum from membership fees without raising them for existing members. This will help to financially stabilize the EGI.eu and global operations it is responsible for.

However, this organizational model does have some weak points:

- While in theory there is no limit to the number of members in the federation, there are practical limits, implied by the transaction costs of too large federations. Even if the membership is restricted to countries only, the "future EGI" could become similar to UNO with some 200+ members.
- This organizational model risks becoming too cumbersome. Regardless of the voting schema it is almost impossible to secure sufficient quorum for a meeting of hundreds of delegates, logistics behind any meeting would be a nightmare, etc. Also, any discussion would need to split the whole constituency into smaller groups to make sure all the members are properly allowed to present their views.
- The model imposes the structure that fits well the European situation (with coordinating roles of the European Commission and Parliament), but may not be suitable to other regions and parts of the world.

- The model also imposes specific technical solution and operation. However, this will remove differences between regions, with a risk losing some of the innovative potential of different (even competing) approaches.
- Some regions (like Latin America) are already well internally organized, but this model does not directly reflect nor appreciate such an organization, although the use of the EGI-InSPIRE groupings model (internal organization of the Collaboration Board that mostly overlaps with the EGI Council) can be applied here.
- Really large countries (e.g. China and India, but also USA) fit rather poorly into this model, without very extensive reconstruction of the voting scheme.
- EGI considers adopting the ERIC (European Research Infrastructure Consortium) model, which means additional restrictions on the consortium membership.
- Last but not least, this model would require finding a different name—European in the EGI would become too restrictive and will need to be replaced by something like GGI (Global Grid Initiative). Also, in such a case the ERIC would become a no-option, as it is not properly suited for a global continents spanning organization. Another negative aspect is additional unwanted reduction of competitiveness.

The potential drawbacks of the federated model make this model unacceptable as a final solution to the problem of the organization of the collaboration of EGI at global level.
On the other hand, this organizational model is already used as an interim solution for single non-European countries. They are becoming accepted as (temporary) affiliated members to help them get access to advanced technologies and accelerate their own development. However, this process is not easy and cannot go too much beyond the current size of EGI Collaboration.

4.2. Peer-to-peer organization

As the federated organization used by EGI works rather well at the continental (world region) level, we can use it as a building block for a proposal of a different organization of the global collaboration. This organizational model is based on the assumption that federated EGI-like collaborations can be successfully applied at the regional level (or at the level of large countries) and these federations can mutually collaborate at a peer-to-peer level. The number of partners in such an organization is naturally kept to acceptably low number—the CHAIN project already works with 8 regions and large countries (Mediterranean and Arab countries, sub Saharan Africa, Middle Asia, Asia-Pacific, India, China, and Latin and Middle Americas) to which we can add the developed countries and regions (USA, Canada, Australia and New Zealand, Russia) to cover the whole world.
This model has some very strong points:

- Such number of partners can collaborate on a peer-to-peer basis without need for a dedicated coordinating organization.
- The model does not impose any prescribed (or globally agreed on) internal organization per each partner.
- There is no need for a common technical solution deployed globally at all levels. Agreement on standard interfaces is more natural and easier done within a smaller number of partners.
- The model supports competitiveness, as individual partners (regions) can develop and keep their own technical solutions, interacting through agreed standard interfaces only.
- The model can be implemented piecewise, starting with bilateral agreements (between EGI and individual regions and/or between regions themselves), evolving gradually into a more complex peer-to-peer collaboration. The level of integration (level of actual collaboration) can be adapted on a per region (per partner) basis without endangering collaboration

elsewhere. Also, more formal structures can be introduced later, when their need arises—as an example is a (weak) coordination of this peer-to-peer collaboration through one partner, with eventual rotation of this responsibility among partners (using as an example the rotating presidency of the EU).

- Each partner of such a global collaboration can proceed at its own pace, emphasising different internal organizational and technical aspects. No one is rushed nor delayed because of a decision made elsewhere.
- Different levels of (technical and organizational) integration can be agreed between individual partners, leading to different levels of interoperability available to users.

Naturally, there are also some weak points and risks:

- The loose coupling of partners with the peer-to-peer collaboration may slow down progress in areas where a common agreement is necessary (but this can be at least partially overcome through different levels of collaboration).
- As this model does not impose any agreement in internal technical solutions (e.g., different middleware can be easily used), it may result in weaker interoperability (but this is a standard price to be paid for competitiveness among different technical solutions).
- Especially during the bootstrapping phase it leads to more complex administration, as individual agreements (e.g. MoUs) must be prepared, signed and followed.
- The different agreements between partners (and their different maturity) can lead to more confusing environment for users. This can be mitigated through an overlay imposed by the international research communities (VRCs) that can further reduce differences between regions through application of additional middleware layer (e.g. using Scientific Gateways to hide most of the complexity and differences between technically different infrastructures).
- There is a potential security risk that some regional organizations will include as their members countries under UNO restrictions, complicating the agreements within the peer-to-peer collaboration (however, at least at the technical level this can be resolved practically without additional effort, compared to the federated model).

This organizational model better reflects the actual situation, where some regions already developed or are developing their own approach towards the distributed computing infrastructure and EGI already has MoUs in place for the actual collaboration. It does not contradict the federated organizational model more suitable for a single region.
In addition to the previous considerations, the changes in the technology and in the user requirements need a co-ordination and harmonisation activity to smooth out all the unnecessary incompatibilities and provide exchange of views on the evolution of the eInfrastructures and their services. Also the changes in operational aspects of external infrastructures have to be followed and the interoperation supported. These have been the roles of CHAIN in these last two years, promoting discussions, keeping track in updates in operational structures and approaches, considering alternative models – all of these focusing on inter-regional issues. The central point has been to bring at the same table the regional infrastructure operators, as well as VRCs and regional eInfrastructures providing discussion agendas and proposing practical solutions.
The peer-to-peer organizational model is used as a generic framework for the actual proposal in the next chapter, however the current activities of EGI and other regional infrastructures, investigating possible migrations towards alternative computing models (such as cloud solutions), can potentially impact the model. . The possible way of monitoring and addressing this evolution is also discussed in the following chapters.

4.3. Technical considerations

The organizational models are not studied in isolation, but they are considered as the key tools towards the actual goal: provide world-wide infrastructure that is capable to run user jobs and provide all necessary information about them regardless of the location of actually used resources. The key aspect is the interoperability at the level of the infrastructure platform (ability to accept and run jobs and to make results available) and at the collaborative platform level (information and user support systems). The interoperability must be supported also by the adopted organizational model; otherwise users will not be able to accept such an organization.

The simplest solution is provided within the narrowly taken federated organization model, with its strict rules on the deployed middleware (resources are not allowed to run other than approved middleware). The interoperability here is taken care of at the entry of the middleware into the system: only components already interoperable with the infrastructure are accepted and endorsed. This is the model currently used within the EGI, where the interoperability is defined through the UMD. The advantage is the full integration; disadvantage is rather slow adoption of new solutions, especially if they have potentially disruptive effect on existing production infrastructure.

On the other hand, the peer-to-peer model implicitly expects different solutions and systems deployed by individual peers. In this case, the interoperability should be achieved through common interfaces, either defined as existing standards or agreed on among the peers. While this approach reduces the extent of integration of individual infrastructures, it gives much higher freedom to individual partners to select their own “best” solution without leaving them outside the club.

EGI Collaboration in its recent Deliverable D4.6 “EGI Operations Architecture (Infrastructure platform and collaboration platform integration)” already started defining a way how services from the EGI operations portfolio (mentioned above in section 3.2) can be adopted and extended to support integrated operations with external Research Infrastructures. While the document is primary focused on other European Research Infrastructures, its conclusions and recommendations are valid also for the international collaboration that is subject of this document.

The following services are identified as the most crucial to support interoperation between infrastructures:

- Topology information
  Proper operation of a grid infrastructure relies on knowledge of topology information, i.e. general mostly static information about sites (Resource Centers) and their services and service endpoints. Within EGI this information is kept in the Grid Configuration Database (GOCDB) and includes information both about sites themselves (including relevant personal contact information) and about services (type/name, service endpoint and eventual endpoint location). This database also keeps information about resources available at site, their scheduled intervention plan and status.
  This kind of service is essential to provide an authoritative list of resources and services available together with their location and status information and also contacts in case of any issues.
- Accounting information
  Currently, OGF UR Working Group (UR-WG) is working on a revision of the standard for compute accounting. There is also a proposal from EMI for the storage accounting. Using these standards will allow to exchange accounting information between different grid infrastructures in a defined and clean way. While the accounting information per se is not necessary for the actual operations, it is usually required by virtual organizations for their own bookkeeping purposes. A distributed nature of the EGI accounting infrastructure

allows for easy combination with any other infrastructure that uses the standard usage records. The summaries can be taken care of through the EGI Accounting infrastructure (based on APEL), provided the other research infrastructure uses the standard UR and confirms to rather simple rules and requirements presented in the EGI D4.6 Deliverable. Using the current infrastructure, it is also possible to exchange the EGI summaries with external portals to allow the data visualization be done locally.

- Monitoring
  While accounting information is used for a post processing analysis, monitoring takes care of service status in a real-time and notifies site administrators if anything goes wrong. This service is primary important within the infrastructure, but its importance spans infrastructures when they serve international VOs and the VO operators need a concise view of the whole VO-related virtual infrastructure. While there are no official standards for general resource monitoring, there are well established open-source projects like Nagios that could serve as the basic building blocks for the interoperable monitoring infrastructure.
- Information discovery
  With a plethora of services running in and on any large grid infrastructure, an ability to search, filter and understand the available information is necessary to actually fulfill any requests users or administrators may have. The GLUE specification can serve as a de facto standard for a conceptual information model for grid entities. Using natural language and UML Class Diagrams, this model is made independent from concrete implementation-specific data models. This makes GLUE a primary proposal for the interoperability of grid entities descriptions needed for the service discovery spanning different grid infrastructures. In parallel to GLUE, the LDAPv3 (RFC 4350) protocol and search syntax provides another standard to be used to query information form the information discovery services.
  Within EGI, the Information discovery service comprises the following five aspects:
    - Service Discovery—find the service
    - Service Selection—narrow the search according to some highly dynamic data
    - Monitoring—state of the service
    - Oversight—aggregated view of services and their properties
    - Diagnostics—tracing of problems

  Interoperability must deal all of these aspects, however not with all of them at the same level. The Service Discovery and Service Selection should work over the different infrastructures if a single query is to be sufficient to find and select a particular service. Sharing the Diagnostics aspects helps to discover and mitigate any problem that arises with the available services, but is not essential to actually run jobs on the grids. Also, the Oversight aspect may not be shared, making some views and presentation more complicated or even impossible to achieve, but again it is not absolutely necessary for simple inter-operations among different grids.
- User support (helpdesk integration)
  End user support is a crucial for grid infrastructure users' satisfaction, as they provide means to help users overcome either imagined (e.g. caused by misunderstanding derived from insufficient training and experience) or actual problems with the infrastructure. It is even more important when users are dealing with several grid infrastructures at once. In such cases even small differences can lead to huge user's dissatisfaction as a helpdesk of any one particular grid infrastructure may not been able to help them with problems coming from use of a different infrastructure. Therefore, integration of user support (helpdesks) is crucial to success of interoperable grid infrastructures.
  EGI uses a distributed support infrastructure consisting of topical and regional helpdesk systems integrated through the central GGUS system. This way a formalized

communication between users encountering a problem with the infrastructure and those able to resolve it is established, without users having to understand (and track) where the problem they are experiencing originated. The linking of different helpdesk systems through the GGUS enables passing the incident reports transparently to users. Also, connecting different tools and tracking systems into the GGUS "hub" allows experts from different areas to use the system without giving up the resolution environments they are most familiar with.
To facilitate use of such a distributed system by other infrastructures, EGI provides the xGUS helpdesk template to allow them to build their own user support infrastructure based on the EGI model. xGUS based support systems allows easy integration with the central EGI GGUS one, as both share the same internal model and technology. This fits well with the peer-to-peer organizational model where the EGI GGUS can serve as a backend (hub) where individual (and otherwise) independent xGUS instances can plug in and use it for support information exchange. This setup also allows sharing access to technology experts of different middleware solutions used and shared by (some of) interconnected grid infrastructures.

Technical solutions are of course subject to review and update; this is mandatory in a highly dynamic environment that includes scientific collaborations, software developers, laboratories and universities. The interoperation at the middleware level is thus a moving target that evolves with the needs of the communities and the technology innovation. On the other hand disruptive technologies couldn't be introduced affecting the service availability and thus soft migration paths are necessary. The interoperability at application level is less dependent on such changes if the layers underneath are based on standards and shielding the user from the different technologies used. The Science Gateway (SG) approach, promoted by the CHAIN project, operates as a transversal solution across middlewares and applications facilitating the interoperability at application level and leaving sufficient freedom of implementation to the different middleware solutions.
The SG, however, does not solve all the interoperation problems and specifically some work is still needed to approach the following topics:

- Error management and ticketing system – if an application fails on one of the sites of an eInfrastructure, a preliminary analysis of the problem could point to a misconfiguration of that site and the corresponding (to the eInfrastructure) ticketing system should be used to report the failure;
- Generalise the SG approach to Clouds or even simple computing clusters, thus opening the door to other resources which are not accessible via Grid;
- Authentication and Authorisation still require harmonisation processes in order to proceed towards the single sign-on and, at the same time, retain the strong security characteristics needed by a worldwide infrastructure;
- Migration to Cloud based eInfrastructures has also to be closely monitored to avoid, whenever possible, interoperability issues similar to those currently faced by Grid infrastructures;
- Last but not least data infrastructures have not yet been fully addressed at intercontinental level, although some partial and/or vertical (domain specific) solutions exist (e.g. HEP, Earth Observation).

A preliminary proposal for interoperability could thus be limited to an extension of the SG paradigm to other types of eInfrastructures (e.g. Clouds, Clusters) with some new features related to Data Exchange (similar to Globus OnLine) as it is shown schematically in the following figure.

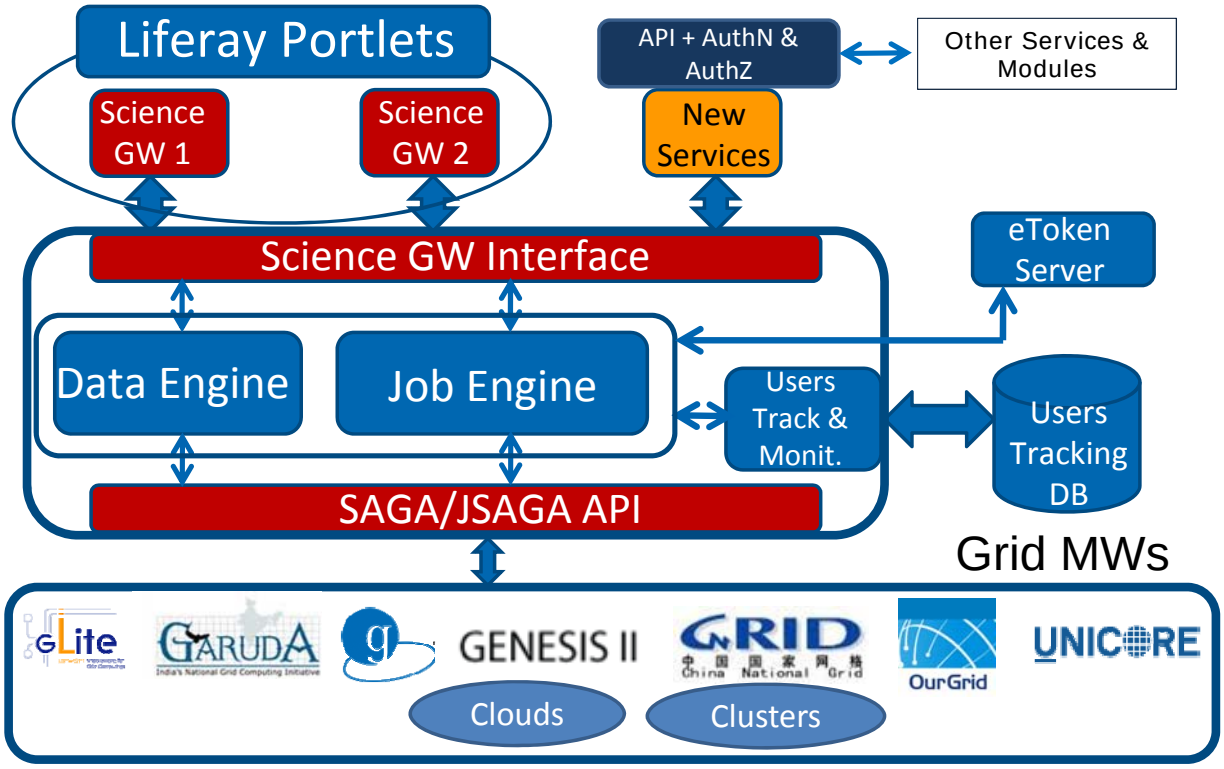

Figure 1 - Science Gateway Architecture

The number and quality of the still open issues requires, however, further work of harmonisation that should continue beyond the CHAIN's time limit.
The deliverable D4.4 will present the road-map foreseen by the CHAIN project, with some of the conclusions derived by the current document.

5. The proposed organizational model

The analysis of the possible organisational models has been conducted over the last 23 months with several meetings that began with the brainstorming during the kick-off meeting in Rome (December 2010) and continued with another brainstorming at the EGI premises in Amsterdam the following January 2011. Refinements and updates of the discussed models and options were done regularly at the EGI conferences every semester, the last being held in Prague (EGI TF2012) last September. These meetings provided a platform for discussion on and establishment of the appropriate organisational model documented here and successfully supported and tested during the CHAIN project.

After having analysed the state of the art and the available options, the CHAIN Deliverable D4.2 presented the preliminary organizational model suitable for the provision of the interoperability of EGI with external grid infrastructures. It has been based on a simplified version of the peer-to-peer model presented in the previous chapter, using Science Gateways as the bridge between different infrastructures.

This model has been successfully applied by several regional eInfrastructures (e.g. EUMEDGRID, GISELA) and projects (e.g. DECIDE, INDICATE) and CHAIN has demonstrated by means of a SG the possibility to run applications across several middlewares (i.e. EMI, Globus, GOS, OurGrid, etc.) and regional eInfrastructures including Latin America (GISELA), Africa and Arab Countries (EUMEDGRID, SAGrid), Asia Pacific, China (CNGrid), India (GARUDA), United States (GENESIS II).

Science Gateways, however, do not pretend to solve operational issues related to the different eInfrastructure involved. A program failure in one of the sites accessed should be followed by a clear procedure that allows the final users to be supported.

The organisational model should thus include a support system that is capable of managing the failures and dispatching the corresponding requests of technical support to the correct ticketing system of the infrastructure concerned.

Moreover, other topics emerged concerning technical and organisational issues such as Clouds, data services, Authentication & Authorisation, etc.

The CHAIN project thus proposes to maintain a concertation table between EGI and the regional infrastructures with regular meetings taking place at the EGI events (Community Forums and/or Technical Forums) and supported by new EC Co-funded projects.

The new CHAIN-REDS project, recently approved by the EC, could take care of such follow-up discussions and provide a valuable frame for the just mentioned follow-up activities.

6. Conclusions

The CHAIN Project after an accurate investigation of the state of the art of regional e-Infrastructures in Africa, Asia and Latin America has restricted the possible interoperation with EGI to two models: the federated one, which provides the highest harmonisation potential at the cost of an unacceptable overload in the organisational structure; and the peer-to-peer model which provides a looser coordination but with the benefit of a lighter organisational structure.

The evolution of technologies, the current foreseen shift of paradigm towards Clouds, the development of federations of identity providers and the growing need for open access to data, however, add new objectives to the coordination and harmonisation efforts worldwide.

This document thus proposes to continue the discussions between EGI and the other regional eInfrastructures to monitor the evolution towards Clouds and provide suggestions in the early stage to avoid possible interoperations and interoperability issues.