EGI-InSPIRE

USER FEEDBACK AND RECOMMENDATIONS

EU MILESTONE: MS305

Document identifier:	EGI-InSPIRE-WP3-MS305-v6
Date:	**07/01/2011**
Activity:	**NA3**
Lead Partner:	**EGI.eu**
Document Status:	**FINAL**
Dissemination Level:	**PUBLIC**
Document Link:	https://documents.egi.eu/document/211

Abstract

One of the prime responsibilities of the User Community Coordination activity of the EGI-InSPIRE project is to capture the needs, requirements and recommendations from user communities of the European Grid Infrastructure. This input has to be normalised, prioritised then made available to technology providers inside and outside of the project. This document describes the channels, tools and bodies that the User Community Coordination activity has established to run this process, and also includes the first set of requirements and recommendations that have been collected in the first six months of the project. The milestone will be reviewed annually, to guarantee that the requirement capturing and collection process evolves together with the user communities and that other stakeholders of the EGI collaboration are informed about the emerging needs of users.



I. COPYRIGHT NOTICE

II. DELIVERY SLIP

	Name	Partner/Activity	Date
From	Gergely Sipos	EGI.eu/NA3	30/09/2010
Reviewed by	**Moderator:** Michel Drescher **Reviewer:** Maria Girone	EGI.eu/SA2 CERN/SA3	05/11/2010
Approved by	**AMB & PMB**		14/12/2010

III. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
1	30/09/2010	First draft with TOC.	Karolis Eigelis / EGI.eu
2	30/09/2010	First draft of content relating requirements gathering process including diagrams added.	SB, GS, KE, NF / EGI.eu
2.5	02/11/2010	Draft for internal review, review by SB, KE	GS / EGI.eu
3	05/11/2010	Draft for external review	GS / EGI.eu
4	22/11/2010	Updated draft based on feedback of reviewers (MD, MG, NF, SB)	GS / EGI.eu
6	03/12/2010	Requirement gathering process refined based on feedback from UCB	GS / EGI.eu

IV. APPLICATION AREA



This document is a formal deliverable for the European Commission, applicable to all members of the EGI-InSPIRE project, beneficiaries and Joint Research Unit members, as well as its collaborating projects.

V. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors. The procedures documented in the EGI-InSPIRE "Document Management Procedure" will be followed: https://wiki.egi.eu/wiki/Procedures

This document is the first version of four milestone documents to be produced by the EGI-InSPIRE project: MS305, MS308, MS311 and MS314. Updates of the current milestone will take place at months 18 (MS308), 30 (MS311) and 42 (MS314). This report will be also complemented by information made available from the publicly available Annual Report on EGI's user communities (month 11, 23, 35 and 47).

VI. TERMINOLOGY

A complete project glossary is provided at the following page: http://www.egi.eu/about/glossary/.

VII. PROJECT SUMMARY

To support science and innovation, a lasting operational model for e-Science is needed – both for coordinating the infrastructure and for delivering integrated services that cross national borders.

The EGI-InSPIRE project will support the transition from a project-based system to a sustainable pan-European e-Infrastructure, by supporting 'grids' of high-performance computing (HPC) and high-throughput computing (HTC) resources. EGI-InSPIRE will also be ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit user communities within the European Research Area.

EGI-InSPIRE will collect user requirements and provide support for the current and potential new user communities, for example within the ESFRI projects. Additional support will also be given to the current heavy users of the infrastructure, such as high energy physics, computational chemistry and life sciences, as they move their critical services and tools from a centralised support model to one driven by their own individual communities.

The objectives of the project are:

1. The continued operation and expansion of today's production infrastructure by transitioning to a governance model and operational infrastructure that can be increasingly sustained outside of specific project funding.
2. The continued support of researchers within Europe and their international collaborators that are using the current production infrastructure.
3. The support for current heavy users of the infrastructure in earth science, astronomy and astrophysics, fusion, computational chemistry and materials science technology, life sciences and high energy physics as they move to sustainable support models for their own communities.
4. Interfaces that expand access to new user communities including new potential heavy users of the infrastructure from the ESFRI projects.
5. Mechanisms to integrate existing infrastructure providers in Europe and around the world into the production infrastructure, so as to provide transparent access to all authorised users.
6. Establish processes and procedures to allow the integration of new DCI technologies (e.g. clouds, volunteer desktop grids) and heterogeneous resources (e.g. HTC and HPC) into a seamless production infrastructure as they mature and demonstrate value to the EGI community.

The EGI community is a federation of independent national and community resource providers, whose resources support specific research communities and international collaborators both within Europe and worldwide. EGI.eu, coordinator of EGI-InSPIRE, brings together partner institutions established within the community to provide a set of essential human and technical services that enable secure integrated access to distributed resources on behalf of the community.

The production infrastructure supports Virtual Research Communities (VRCs) − structured international user communities − that are grouped into specific research domains. VRCs are formally represented within EGI at both a technical and strategic level.

VIII. EXECUTIVE SUMMARY

The NA3 (WP3) work package of the EGI-InSPIRE project is responsible for defining and running support processes for user communities of the European Grid Infrastructure (EGI). The evolution of the European Grid Infrastructure is driven by the users. Therefore capturing and communicating feedback from users to the infrastructure as well as technology operators and providers is a key goal for NA3 and for the project as a whole.

This document describes the channels, processes and bodies that the User Community Coordination activity has established and runs in order to collect, capture, process, and report requirements and recommendations. Input from users will be collected by NA3 through various electronic channels and face-to-face communication mechanisms. The User Community Board (UCB), the group of representatives of user communities, then prioritises the collected input with the help of the EGI.eu User Community Support Team (UCST) and the User Services Advisory Group (USAG). The prioritised, weighted requirements prepared by these groups provide direction for the evolution of middleware services as well as the user-oriented technical services provided by EGI.

This milestone also includes the first set of requirements and recommendations that have been collected through these processes during the first six months of the project.

TABLE OF CONTENTS

1 INTRODUCTION

The EGI collaboration operates a grid infrastructure that spans across approximately 60 countries with more than 300 sites. The consortium operates a large number of middleware components and user-oriented software tools on the nodes. The user communities of EGI are heterogeneous, both in terms of their scientific use case, their grid use case and their experience with grids. Capturing, processing and responding to users' requirements and needs in such an environment in a timely manner is a challenge. The proposed process therefore ensures that:

- Threshold for user communities to provide feedback is as low as possible.
- Useful feedback is separated from harmful input and noise.
- No community's input is neglected.
- No input is lost or distorted during the processing phase.
- Input provided by the communities in different form and with different levels of details will be normalised. Normalised requirements will be compared, merged, summarized if and where necessary.
- The prioritisation mechanism (especially across user communities) is fair, i.e.
    - Input from large communities with more users are treated with higher priority than those coming from smaller groups;

  AND
    - Input from communities with the potential for growth is treated with higher priority than those coming from groups with no, or with minor potential to grow.
- Prioritised requirements are communicated to the right stakeholders of the EGI ecosystem.
- The requirement gathering process works efficiently, i.e. the input is processed and communicated to technology providers within a designated timeframe, contributing to timely implementation on the infrastructure.

This document describes the channels, processes and bodies that the NA3 work package has implemented to support a requirement gathering process with the above features and describes how these processes fit into the big picture of requirements management in EGI. The document ensures that:

- Existing and emerging user communities can learn about the process itself and can see what channels are available for them to communicate feedback about EGI services.
- EGI technology providers and operators can learn how the collection and processing of requirements is implemented within NA3, and thus they can adapt their own internal processes accordingly in order to shorten the implementation cycle for end users.
- Stakeholders of EGI can have a snapshot of the requirements that have been collected during the first six months of the project.

Finally the document describes future plans and both the channels and tools that will be established by NA3 in order to further develop the requirements gathering and requirements processing activities.

2 BACKGROUND

EGI builds on the work and achievements of a number of earlier initiatives most notably, the EGEE sequence of projects. The legacy from EGEE is a number of processes and a body of knowledge. In addition to the EGEE legacy, a number of other projects and bodies have articulated requirements on behalf of the community. These include the following:

- The EGI Design Study: http://web.eu-egi.eu/
- EEF: The European e-Infrastructure Forum: http://www.einfrastructure-forum.eu/
- The ESFRI preparatory phase projects: http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri

Documents and other outputs from these sources have been studied in order to contribute to the initial phase of the EGI requirements gathering process.

The EEF final report[1] describes the following key requirements that cut across all disciplines to emerge from their analysis of the ESFRI Roadmap projects:

1. Single sign-on
2. Virtual Organisations
3. Persistent storage
4. User Support
5. Training and consultancy
6. Web-service interfaces
7. Workflows
8. Global scope
9. Integration with cloud systems and volunteer desk-top systems

One of the key changes that has taken place in this area with respect to the transition from EGEE to EGI is the move from Regional Operational Centres (ROCs) to NGIs. The other change is the introduction of Virtual Research Communities (VRCs) to coordinate the needs and requirements of discipline-specific clusters of Virtual Organisations (VOs) in the form of communities of practice. Instead of having direct channels from one or more user communities and the technology providers, EGI mediates between the user communities and the technology providers. The third mayor change is that the UMD middleware (Unified Middleware Distribution) used on the production infrastructure includes independent bags of services with some common/shared service.

[1] https://documents.egi.eu/public/RetrieveFile?docid=12&version=1&filename=EEF_report_FINAL.pdf

3 CAPTURING USER FEEDBACK AND RECOMMENDATIONS

3.1 The process

The high-level process of capturing input and feedback from users is shown in Figure 1 and it aims to collect feedback from the two most important groups of users of the European Grid Infrastructure:

- *End users*: the researchers, application developers, etc. who use but do not operate software services in the European Grid Infrastructure

AND

- *Service operators*: people who operate grid sites and other software services for end users. Most of the services these persons provide are developed outside of the NA3 activity, even outside of the EGI-InSPIRE project and are provided by external technology developer projects. Some of the services are developed inside NA3, particularly the technical services for users in the TNA3.4 task (EGI-InSPIRE task name "Technical services").

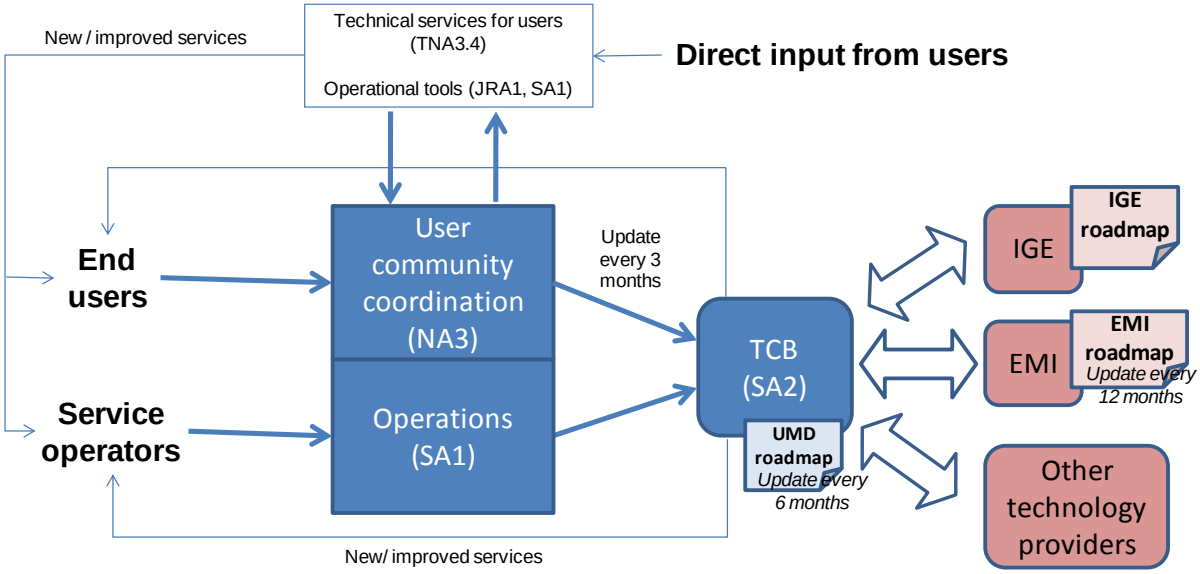

Figure 1. Feedback and recommendations gathering process.

Requirements and recommendations from service operators are captured by the Operations activity (SA1) of EGI-InSPIRE. The requirements and recommendations collected by SA1 are complemented by end user feedback which was captured, collected, normalised and prioritised by NA3. The Technology Coordination Board (TCB) is the central point where all the prioritised requirements from the different sources come together, where possible conflicts between the needs are resolved, and where external technology providers join the process. The requirements endorsed by the TCB are described in the UMD roadmap, then picked up and implemented by technology providers outside of EGI-InSPIRE. A smaller set of requirements, that affect tools that are internal to EGI-InSPIRE, particularly the tools developed within the TNA3.4 activity do not have to cross the TCB and can be endorsed internally by NA3.

Note that the various components of the workflow described in Figure 1 work with different frequencies. The goal of EGI is to fine-tune this process in such a way that the quality of the output

(quality of new software on the infrastructure) is maximised, the new functions meet the users' needs and that the turn-around time is minimised, i.e. new requirements become implemented as quickly as possible. On top of that, EGI aims to run the requirement process as visibly as possible, enabling the different stakeholders to easily join in and benefit from the process.

The SA1 and SA2 related parts of the process (i.e. requirements gathering and processing by Operations and the TCB) are discussed elsewhere [R4]. The focus of this document is on the end user-related part and this is presented in a more detailed way in Figure 2.

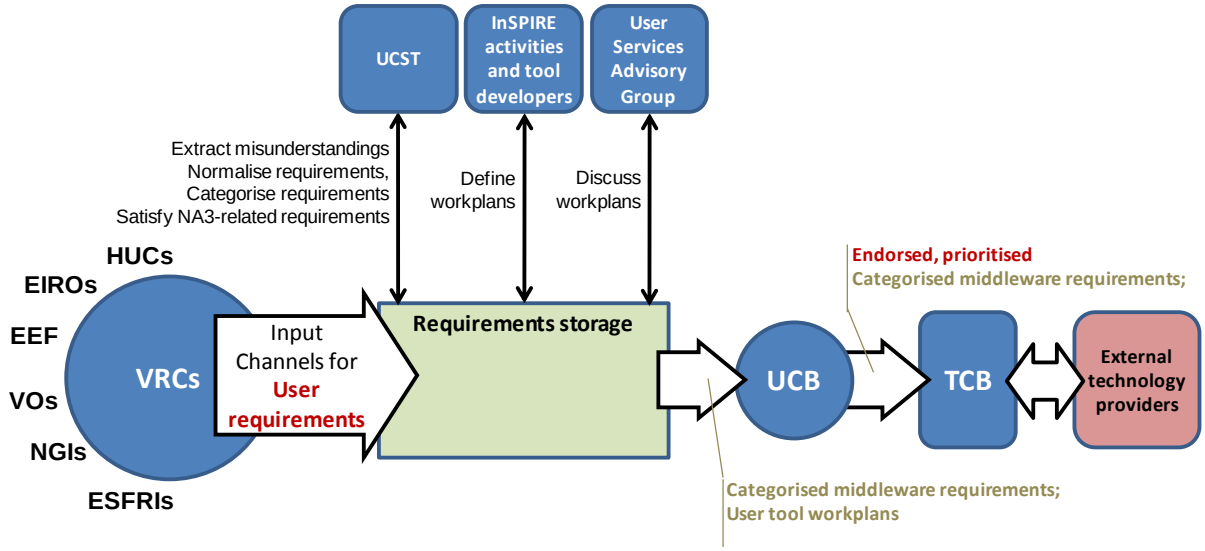

Figure 2. Collecting and processing requirements and recommendations from user communities.

The main driver for the end user requirement gathering and processing workflow is the User Community Support Team (UCST) of EGI.eu in Amsterdam. This team defines and operates the input channels through which end user communities can feed EGI with requests and requirements for new features, and/or changes. These channels are detailed in Section 3.2. The channels have been designed to capture feedback from the existing, formally recognised communities, as well as from the emerging, formally not yet recognised groups.

The most important source of end user requirements are the EGI VRCs, the sustainable EGI user communities that are capable of some or all of the following:

- port new applications to the infrastructure
- extend the infrastructure with domain-specific software according to their members' needs
- expand their user base without depending (heavily) on user support services of EGI.

Besides the EGI VRCs, the other formally recognised end user entities are the mature Heavy User Communities (HUCs) and the established Virtual Organisations (VOs):

- Heavy User Communities (HUCs) are mature scientific collaborations that have been using the EGEE grid and EGI routinely for many years and thus have become more structured and advanced in terms of grid usage. HUCs' activities are carried out within the SA3 work package of EGI-InSPIRE in the first three years of the project, during which it is expect that VRCs will be establish for these communities. HUCs include the High Energy Physics, Earth Sciences, Life Sciences and the Astronomy and Astrophysics communities. More information about the services and capabilities of HUCs can be found in [R6]. HUC representatives can

[2] ESFRI: European Strategy Forum on Research Infrastructures – http://ec.europa.eu/research/esfri/

Week	Activity	Material
1	Capture snap shot of requirements	Various inputs (see Section 3.2.1)
3	Requirements “normalised”	Internal report
5	Prioritisation by UCST	Preliminary analysis added
7	Priorities for user tools discussed by USAG	External input added by USAG members
9	Issues raised are discussed with communities, prioritisation revised if needed.	
11	Prioritised requirements reviewed by UCB	
12	UCB endorsed prioritised requirements published to technology development community	

3.2 The practice

EGI is reaching out to many communities in the form of VRCs and inviting and encouraging them to articulate their requirements. These VRCs will each have their own preferences for capturing and prioritising their requirements. EGI will be flexible and supportive to a multitude of mechanisms and methods for receiving these and will then normalise and prioritise them for further processing.

The user requirements’ gathering is a complex and non-trivial process due to several reasons. Coordination of efforts between Virtual Organizations, Virtual Research Communities, National Grid Initiatives and EGI.eu is crucial, in order to assess the requirements from the communities they embrace. This is one of the main goals of the EGI.eu UCST; *to bring together a pro-active user support service across Europe under the EGI umbrella*.

Though the requirements from these communities are diverse, reflecting different fields of study and approaches to their scientific endeavors, a set of general requirements can be envisaged. It is thus critical to assess specific and general requirements such that proper prioritisation can be performed. Another important aspect is related to the heterogeneity of people operating and using the e-Infrastructure. Their IT background, their vocabulary to describe computing problems is different, thus strategies to guarantee a successful dialog between both parties are required. The various channels ensure that users with different experience, different levels of experience, different forms of representation are all reached and heard.

The NA3 activity defined a page [R5] within the EGI Wiki to provide a complete and up to date picture of the requirement gathering process run by the activity. The page describes the goal of the process, the channels, the methods and bodies used by the process, and the actual output which has been achieved so far. The remaining parts of this document is a replication of this information, providing a snapshot of the requirement gathering activity.

3.2.1 Requirements gathering channels

The activity already opened several channels towards users in order to gather requirements from them. These channels – document based requirement capturing, online forms, surveys – are described in the following subsections. Other tools to congregate user requirements are under study, such as blogs, social networking and request trackers. The next version of the milestone will include sections on these tools too.

3.2.1.1 Documents

A first tentative methodology to gather the requirements was done by filtering the contents of documents written by different communities. This task revealed the need of EGI UCST to provide a user friendly service to collect the necessities, such that no ambiguities exist and the requirement is properly understood by both parties. The documents that the UCST already processed are (documents can be found at [R5]):

- SHARE Roadmap by the Healthgrid community
- Production Grid Infrastructure - Use Case Collection v1 by OGF Production Grid Infrastructure WG
- ESFRI project requirements for Pan-European e-infrastructure resources and facilities by European E-Infrastructure Forum
- Requests for Enhancements of gLite to support bio-NMR applications by the Deploying and unifying the NMR e-Infrastructure in System Biology (e-NMR) project

The documents that the UCST recently received and have not processed yet are:

-
- gLite requirements collected by the Project Technical Forum of the EGEE project in 2005
- Documents released by e-IRG and the ICEAGE project on training and education

The activity recently invited all the HUCs to provide/prepare similar requirement description documents for the UCST in whatever form they prefer.

The requirements and recommendations that have been identified so far from these documents can be found in Appendix A and online in the requirement gathering page of EGI Wiki [R5].

3.2.1.2 Direct input from VOs and HUCs

Prior to EGI, in EGEE the user communities were represented by VOs and by application clusters, currently called Heavy User Communities. Because VRCs have yet to be established, VOs and HUCs remain the most valuable entities for EGI to communicate with users. Until VOs and HUCs become integrated parts of VRCs, the UCST has to capture requirements and recommendations from users through VOs and HUCs as well.

Recently, (during the month of October) the UCST sent out a questionnaire to the VO managers and to the HUC coordinators and asked them to collect input from their members on two topics on which EGI.eu required input from user communities:

- How would the transition of gLite based EGI sites from gLite version 3.1 to 3.2 affect the users?
- What bug fixes and other enhancements would users like to see in the first EMI middleware expected to be released in April 2011?

The deadline for VOs to answer was in the second half of October. Because of the low response rate (23 VO out of the 223 - 10.3%, which represents 2,046 out of the 12,409 users - 16.4%.) and

because of the received answers EGI together with EMI agreed to extend the support for the WN, UI and LFC components of the gLite 3.1 middleware until the end of April 2011. This information has been circulated to HUCs, NGIs and VOs.

The input received from VOs and from two HUCs (High Energy Physics and Life Sciences) for the EMI middleware enhancement question can be found in Appendix B. This will be discussed at the inaugural UCB at the end of November.

Besides collecting valuable feedback from users, VO surveys also help the UCST identify nonresponsive VOs, VOs with non-functioning managers. These findings can drive the development of new communication mechanisms, and can initiate the suspension or removal of inactive VOs from the infrastructure. As the first step, an email message has been sent out to all the HUCs, NGIs, VOs and NA3 members about the details of the requirement gathering process, including the channels to communicate any additional requirements to EGI.

3.2.1.3 Direct input from NGIs

The main actors of EGI are the National Grid Initiatives (NGIs). These ensure the operation of the grid infrastructures in each country as well as a transparent representation of the requirements of all their scientific communities together with resource providers and all e-Infrastructure-related institutions. NGIs operate User Support Teams (USTs) that contribute to the distributed user-support services of EGI [R1]. One of the responsibilities of NGI USTs is collecting and providing feedback about existing user services; collecting and providing requirements about missing grid and support functionalities and tools. NGI USTs do this on behalf of the user communities they locally support, and on behalf of their own members. EGI.eu is proactive in the process and performs the following tasks:

- Identification of confirmed NGI UST contacts: The first step of requirement gathering through USTs was the identification of user support contacts. NGIs have been asked to delegate one main contact from their UST who – among other things – can communicate local feedback and requirements to the EGI.eu User Community Support Team (UCST). The process of collecting confirmed NGI UST contacts began in June and by now most of the NGIs (about 80%) delegated user support contacts. The whole process is described in the MS301 milestone document [R2].
- Surveys, questionnaires: the EGI.eu UCST surveys NGIs to collect input on well defined, user related topics. The EGI.eu UCST aims to use a single tool to survey communities with a pre-defined frequency. Whether this is possible and is in line with related activities inside and outside of the project is still an open question.

The first NGI survey has been recently sent (end of September, beginning of October) to all the confirmed NGI UST contacts. The goals of this first questionnaire were: (1) to collect information about the status and services of NGI user support teams, and the services and support tools these teams expect from EGI; (2) identify the interests of NGIs in the field of user support and; (3) initiate discussions with them, collect input from them for the improvement of user support services and tools. EGI.eu UCST made all the surveys and the received responses are visible to the EGI community through the NA3 Requirement gathering Wiki page [R3]. The responses received so far from NGIs to the user support survey are also available at [R3] and in Appendix C of this document.

3.2.1.4 Events

Representatives of the EGI.eu UCST attend user-centric events that are organised by grid and scientific communities (DECIDE workshop, NEERI 2010, NDGF Strategy Workshop). These events

provide excellent opportunities for the UCST to identify emerging user requirements and/or to further discuss needs that are already known but are not detailed enough yet to be ratified by the UCB. While so far the primary focus of the UCST at such events has been the identification and engagement with VRC candidates, it is expected that once VRCs are established these events will foster the requirements gathering process by UCST.

The UCST is planning a series of user requirement gathering workshops to be held during throughout 2011. Some of these will be co-located with events such as the EGI Technical and User Forums, others will be held as standalone events. The meetings will be open to all, however, specialists whose areas of interest relate to the specific themes will be targeted and encouraged to attend. Workshops will last for one day and will be broken down into a number of sessions. Typically, a couple of these will be parallel breakout sessions during which more detailed topics may be examined and then reported back to the full workshop. It is anticipated that these events will deliver the following opportunities:

- An opportunity for researchers to make informed statements about their requirements relating to infrastructure technology and provision,
- a refinement of development plans for specific applications and services prior to submission to UCB for ratification,
- an arena for academics to present their latest achievement including describing how infrastructure has enables, enhanced or hindered this work,
- an opportunity for developers to showcase their latest technology,
- an opportunity for application and middleware strategists to outline potential new directions to an informed and interested audience,
- an opportunity for new collaborative partnerships to emerge,
- a forum for constructive discussion about and planning about all of the above points.

3.2.1.5 Online form

A second approach is also being devised based on an online form which is integrated into the EGI web site. Using the form any user community, can possibly through a use case scenario, give evidence of the needs from the infrastructure (see Figure 3 below). It is assumed that only those communities who are not yet formally engaged with EGI (as a VRC or as a VO) will use this form. Even though we expect low number of responses, the mechanism makes it possible for any community to provide feedback to EGI. Given the extremely low threshold to submit input through this channel, the UCST must be very careful with the received data. It should make sure that the entered data comes from credible community with potential of becoming a heavy user of the infrastructure. The form is connected to a Google spreadsheet, so data entered can be seen and accessed only by a few selected members of the EGI.eu UCST.

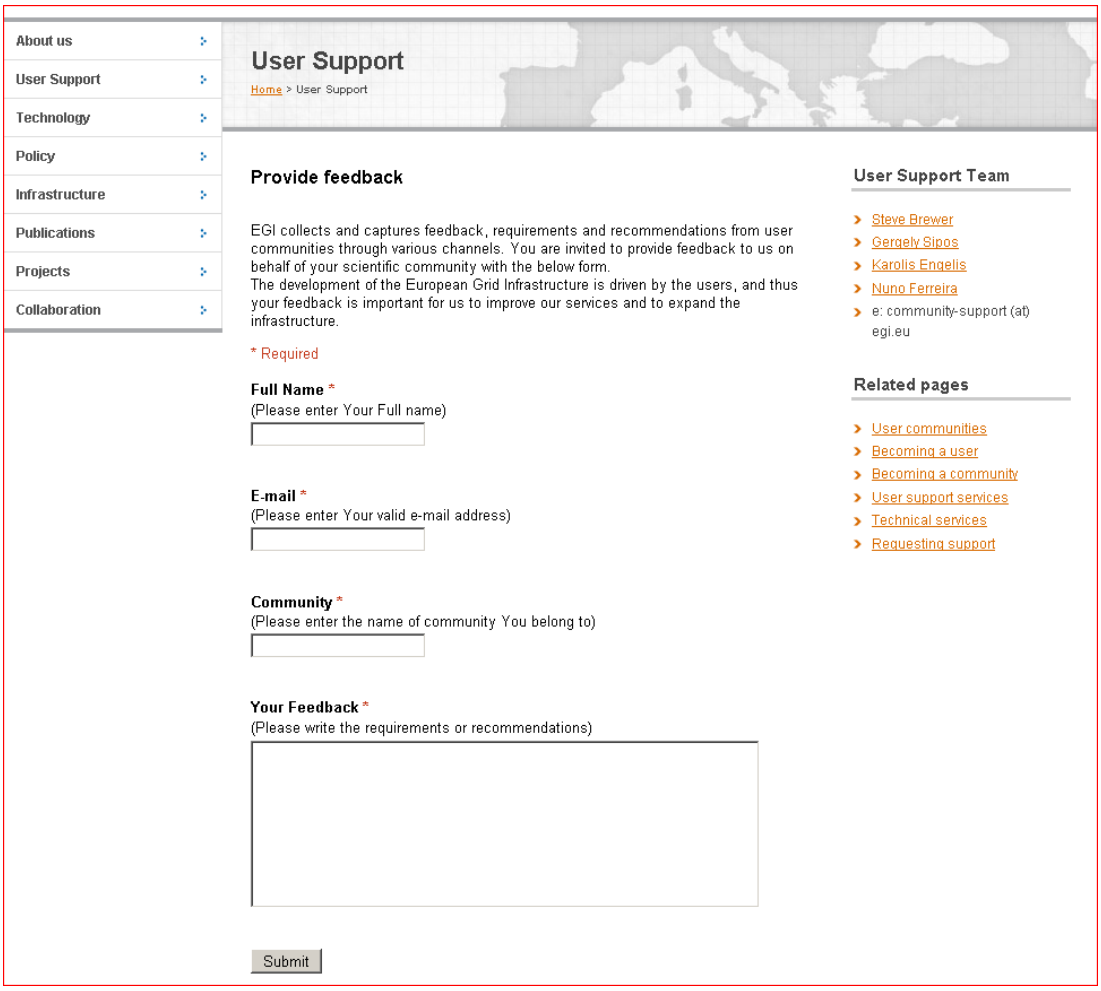

Figure 3. Online form to capture feedbacks and requirements from formally not recognised communities.

4 TIMETABLE FOR IMPLEMENTATION AND PROCESS EXECUTION

The mechanisms described in this document will eventually become a regular cyclical process in the fullness of time. A significant part of the reason for this is to clearly decouple the requirements gathering process from the middleware and user tool development cycles. This is important as it both prevents this function from being purely reactive to the needs of the development teams and also presents a much more even process to the user community. It is proposed that this cyclical process is based on iterations with an updated prioritised requirements report being made available to the development community. In this context the term development community covers both the Technology Coordination Board and the related external providers such as EMI and IGE but also the partners related to NA3's Technical Services development. In order to get to this situation a number of milestones will need to be reached:

Project month	**Event**	**Outcome** **Expected outcome**	**Date**
PM4	First NGI survey	Report on UST service status; Suggestions for improvements of user support activities and services	October 2010
PM4	First VO and HUC survey	Feedback on gLite transition; Expected enhancements for first EMI middleware release	October 2010
PM5	First TCB meeting	Agreement on requirement gathering process between NA3 and other activities and projects	October 25th 2010
PM6	First VRC established	Confirmed channels to communities	November 2010
PM6	First USAG meeting	Technical discussion of the workplans of user facing tools, driven by the requirements captured from users	November 2010
PM7	First UCB meeting	Endorsement of collected requirements; Identification of missing input; Feedback on requirement gathering process	November 2010

5 CONCLUSION AND FUTURE PLANS

The NA3 activity has established channels to the user communities that are currently represented in EGI in the form of VOs and HUCs and through NGIs. The first set of requirements has been collected through these channels and is available through the EGI Wiki [R3] and in this report. The UCST supported by the USAG will process and discuss these requirements further before ratification by the UCB, and presentation to the TCB.

While the UCB, TCB and USAG work with pre-defined frequencies the requirement gathering process from the users' point of view is continuous. Communication channels for existing and for new communities towards the UCST are always open, requirements can be communicated to the team in various forms.

The team also works on improving and extending the process:

- Open additional input channels towards users, e.g. a blog with discussion forum to generate interaction among users on topics defined by the blog posts.
- Improve existing channels by e.g. using online voting and surveying tools to reach VOs, NGIs and VRCs and to evaluate responses.
- Setup a tool that ease the maintenance of requirements and could be used to define, categorise, search and handle the users' needs.
- Define topics on which users can provide feedback on. For example, ask users about (1) new features that the EGI middleware should implement; (2) new features that the application database should implement; etc.

The surveys that have run recently also revealed a few issues that the UCST will deal with, but these are outside the scope of this document:

- Most of the VOs cannot be reached through their managers.
- The email addresses of some of the VO managers are invalid, contact with such VOs is thus lost.
- Obtaining information about the user support activities of NGIs is hard - some are not responsive for questionnaires, others give very short answers.
- Lack of funding for user communities is still an obstacle for some to establish VRC.

6 REFERENCES

R 1	User Community Support Processes (D3.1 deliverable): https://documents.egi.eu/document/106
R 2	User Support Contacts (MS301 milestone): https://documents.egi.eu/document/60
R 3	Surveys for user communities (section in EGI Wiki): https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities
R 4	EGI-InSPIRE, D5.1 UMD Roadmap: https://documents.egi.eu/document/100
R 5	Requirements gathering (page in EGI Wiki): https://wiki.egi.eu/wiki/Requirements_gathering_details
R 6	Capabilities offered by Heavy User Communities (D6.1 deliverable): https://documents.egi.eu/document/154

7 APPENDIX A –REQUIREMENTS CAPTURED FROM DOCUMENTS

These requirements have been captured by the UCST from documents received from HUCs, VRC candidates and other contacts. The UCST invites any scientific community to submit materials that describe their e-Infrastructure requirements. All these materials and captured requirements are available at [R5].

7.1 From SHARE roadmap

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**HealthGrid** (Innovative medicine community)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	SHARE the journey. A European Healthgrid Roadmap
Link	http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Flexible, secure and scalable IT infrastructure

Summary of Requirement(s)

At the request of the European Commission, the European Federation of Pharmaceutical Industries and Associations (EFPIA) has identified the main barriers to innovation in Life Sciences research in Europe with the objective of establishing a European technology platform for innovative medicines.

Several key bottlenecks in the R&D process were described, where scientific and technological advances would be of direct benefit to the pharmaceutical industry by improving efficacy of tests and containing costs. The knowledge management area was identified as key to leveraging the potential of new technologies such as genomics and proteomics and to analyze the huge quantity and diversity of information in an integrated way.

The scientific requirements identified to accomplish the knowledge management goal are not specific to grids. Nevertheless, they are relevant to Health Grids if they want become relevant infrastructures for biopharmaceutical research and development. This community thus pinpoints the need for technology which

supports a distributed/federated, service oriented, and ontology driven architecture, providing a collaboration medium, facilitating effective computation and capable of generating, organizing and managing knowledge.

From the point of view of EGI.eu UCST, among the non-specific requirements there's one that falls under our umbrella of responsibility. The IT infrastructure should be flexible, secure (covering all aspects of data protection encountered in a biomedical context), and scalable.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**HealthGrid** (Epidemiology community)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s) **Name** **Link**	 SHARE the journey. A European Healthgrid Roadmap http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Data security

Summary of Requirement(s)

The epidemiology community is strongly focused in ICT-driven research in two main areas, namely patient-customized research and population-level research. Both of these scenarios share in general the problem of access to distributed, critically sensitive and heterogeneous data, resulting in overall costly computing processes.

Users from this community have several requirements, some should be addressed internally by them since they are directly related to their applications nevertheless, EGI.eu UCST could spot some relevant requirements which are under our umbrella. Epidemiology users ought to be able to take as granted:

- Security mechanisms are sufficient to protect their data. Other than being sensitive to security issues, they should not need to know anything in detail about encryption, secure transfer, delegation or other technical issues.

Ensuring that the research is done in a secure and legally-compliant framework is thus a must.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	**HealthGrid** (Epidemiology community)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s) **Name** **Link**	 SHARE the journey. A European Healthgrid Roadmap http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Quality of service

Summary of Requirement(s)

The epidemiology community is strongly focused in ICT-driven research in two main areas, namely patient-customized research and population-level research. Both of these scenarios share in general the problem of access to distributed, critically sensitive and heterogeneous data, resulting in overall costly computing processes.

Users from this community have several requirements, some should be addressed internally by them since they are directly related to their applications nevertheless, EGI.eu UCST could spot some relevant requirements which are under our umbrella. Epidemiology users ought to be able to take as granted:

- The services are reliable, efficient and permanent. They may not understand, or want a detailed explanation of why a service is down, or why a job is taking so long. They are expecting a quality of service similar to any other utility.

Availability of efficient infrastructures and usage policies should be enforced. Applications will require resources and reliable infrastructure to work on under a clear Quality of Service (QoS) promise.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**HealthGrid** (Epidemiology community)
Requirement Type	☐ Support ☐ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s) **Name** **Link**	 SHARE the journey. A European Healthgrid Roadmap http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Community internal requirements

Summary of Requirement(s)

The epidemiology community is strongly focused in ICT-driven research in two main areas, namely patient-customized research and population-level research. Both of these scenarios share in general the problem of access to distributed, critically sensitive and heterogeneous data, resulting in overall costly computing processes.

Users from this community have several requirements; some should be addressed internally by them since they are directly related to their applications and how they are implemented. For completeness, EGI.eu UCST reports those requirements below, since they are strongly coupled with the applications needs from this community, and EGI could also provide technical expertise in the decision making process regarding this issue.

Requirements in a broad sense can be summarized as follows:

- Effective semantic annotation of data. Data is poorly coded and interoperability of coding is not trivial. Extracting knowledge from medical data, however, is a main objective.
- Effective integration of distributed and heterogeneous data. Integrating distributed resources requires exchange protocols, secure mechanisms, patient de- and re-identification, and automatic data analysis services.
- User-friendliness of applications and services. The tools should be available through protocols and interfaces similar to those used in the users' normal research. Not only must the applications be as compliant as possible with current systems and interfaces, but so must the technologies.
- Reliability, scalability and pervasiveness. All services must be robust and trustful and should be scaled without reducing performance.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	**HealthGrid** (Virtual Physilogical Human – VPH - community)
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 SHARE the journey. A European Healthgrid Roadmap http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Collaboration between community and grid experts

Summary of Requirement(s)

The Virtual Physiological Human (VPH) research project is promoting a methodological and technological framework that once established will enable the investigation of the human body as a single complex system. It is well understood by the VPH research community that grid technology is required to pursue effectively this ambitious goal. In an attempt to bridge the gap between the VPH community and the grid community, a group of experts exchanged views on the relevance of grids for VPH. The main conclusions were summarized in four topics, namely: research requirements; requirements specific to grid computing; requirements specific to grid data and knowledge management; and requirements relevant to grid technology adoption and application deployment.

Regarding research requirements:

"The vast scope and integrative approach of the VPH project can only successfully be addressed using the resource sharing mechanisms provided by a grid infrastructure. However, analysis of the present situation shows there are barriers to such deployment. We propose to overcome this situation by deploying on the existing infrastructures some grid services that could be of extreme usefulness for the VPH community; this should attract VPH researchers to the large scale infrastructures, and should help grid developers to become more aware of the special needs of this emerging scientific community. The collaboration between the VPH and grid communities should enlarge the computing and storage resources as well as the services made available to the VPH community and foster the identification of new scientific research areas to which a grid environment can be appropriate."

In short, this community will be responsible to deploy grid services for their community, and feedback from grid experts on the best practices to accomplish the goal, should be considered as a request.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	**HealthGrid** (Virtual Physilogical Human – VPH - community)
Requirement Type	☐ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☒ Other : Don't know how to categorize this one
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 SHARE the journey. A European Healthgrid Roadmap http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Easy access to grid

Summary of Requirement(s)

The Virtual Physiological Human (VPH) research project is promoting a methodological and technological framework that once established will enable the investigation of the human body as a single complex system. It is well understood by the VPH research community that grid technology is required to pursue effectively this ambitious goal. In an attempt to bridge the gap between the VPH community and the grid community, a group of experts exchanged views on the relevance of grids for VPH. The main conclusions were summarized in four topics, namely: research requirements; requirements specific to grid computing; requirements specific to grid data and knowledge management; and requirements relevant to grid technology adoption and application deployment. Research requirements, as understood by EGI.eu UCST, belong to the scope of VPH developers.

Regarding requirements specific to grid computing:

"It is essential to take an approach that integrates all resources beyond the desktop into a cohesive infrastructure, accessible by all VPH researchers as necessary. This means allowing researchers access to resources in a uniform manner, from their local departmental cluster to the biggest HPC machines available on a national or EU basis, and including everything in between. […] The multi-scale nature of the VPH project demands that access to such resources be provided in as seamless a way as possible, and where appropriate, mechanisms be developed to allow the automatic migrating of simulations between different scales or different platforms (and by implication, between resources appropriate to run the simulation at a particular scale)."

In summary, users should have an easy access to their applications ported into the grid, such that everyone can benefit from it. The heterogeneity of the grid should also not impose problems to researchers, whether the research is done at a small or big scale.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	**HealthGrid** (Virtual Physilogical Human – VPH - community)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s) **Name** **Link**	 SHARE the journey. A European Healthgrid Roadmap http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Data: persistency, availability, security

Summary of Requirement(s)

The Virtual Physiological Human (VPH) research project is promoting a methodological and technological framework that once established will enable the investigation of the human body as a single complex system. It is well understood by the VPH research community that grid technology is required to pursue effectively this ambitious goal. In an attempt to bridge the gap between the VPH community and the grid community, a group of experts exchanged views on the relevance of grids for VPH. The main conclusions were summarized in four topics, namely: research requirements; requirements specific to grid computing; requirements specific to grid data and knowledge management; and requirements relevant to grid technology adoption and application deployment.

Regarding requirements specific to grid data and knowledge management:
"In many grid contexts the data are transient in nature; they are produced by the simulation runs, but after being analyzed, can be stored off line or even trashed. Persistent data collections must be provided to the VPH community. Large Scale Infrastructures should make available storage services designed to ease the upload and download of large binary objects, and their replication computationally near by the execution nodes. The management of storage and execution resources should be designed to have inherent security and knowledge management features. Security is vital because of the sensitive nature of the clinical or genetic data VPH sometime involves. Current technologies are insufficient to protect the privacy of the data outside the health network barriers, according to legal regulations and ethical principles request. Technologies are also not scalable when dealing with fine-grain authorization and delegation methods in current practice could be not sufficient for medical applications. Finally, the accumulative nature of VPH imposes that everything is organized under solid knowledge management models, which make possible to keep organized and usable even very large information spaces."
In summary, all aspects related to data, namely, persistency of storage, security and availability, are a concern to this community.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	**HealthGrid** (Virtual Physilogical Human – VPH - community)
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s)	
Name	SHARE the journey. A European Healthgrid Roadmap
Link	http://eu-share.org/roadmap/SHARE_roadmap_long.pdf
Requirement *Unique name identifier*	Cross community interaction

Summary of Requirement(s)

The Virtual Physiological Human (VPH) research project is promoting a methodological and technological framework that once established will enable the investigation of the human body as a single complex system. It is well understood by the VPH research community that grid technology is required to pursue effectively this ambitious goal. In an attempt to bridge the gap between the VPH community and the grid community, a group of experts exchanged views on the relevance of grids for VPH. The main conclusions were summarized in four topics, namely: research requirements; requirements specific to grid computing; requirements specific to grid data and knowledge management; and requirements relevant to grid technology adoption and application deployment.

Regarding requirements relevant to grid technology adoption and application deployment (closely coupled with the research requirements):

The VPH community is very heterogeneous and "largely ignores the large scale infrastructures, avoiding deploying large scale collections, and excluding the use of massive computational resources as an opportunity to solve some of its modeling problems. On the other hand, with a few notable exceptions, the HPC infrastructures and the other grid stakeholders are so far failing to provide the services needed to handle the VPH community computing needs. What is required is the encouragement of cross community interaction, in order to build meaningful dialogue between grid developers and VPH researchers. Providing higher-level tools that allow VPH researchers to interact with the resources that they need to achieve their scientific objectives in a uniform manner, abstracting where necessary the underlying difficulties of dealing with grid middleware, will help engage with researchers […]".

In summary, VPH stresses the urgent need for collaboration between grid experts and VPH researchers. This community also recommends the need to "identify a few VPH CPU intensive applications which would benefit immediately of the existing grid infrastructures. The deployment of these applications will allow identifying the missing services on the existing infrastructures and will rise up the grid awareness in the VPH community."

7.2 From EEF report

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☐ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☒ Other : Worldwide collaboration/interoperation
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Global collaboration

Summary of Requirement

All ESFRI sectors identified the need to collaborate with parties beyond Europe's borders. The European E-infrastructure Forum members can leverage their existing international contacts for the benefit of the ESFRI projects. For ELIXIR the USA (NCBI) and Japan (DDBJ) are identified key partners. The global requirements that have been articulated to date are within the activity footprint of the GÉANT global reach and relationships are in place to assist in organizing network requirements beyond Europe. Availability of services should not be a problem but for specific locations access capacity to those locations, including capacity available for elements of the service portfolio, would need to be confirmed and, if appropriate, addressed.

From a grid and HPC point of view, there are already a number of interoperation points addressed via the Infrastructure Policy Group which meets at the Open Grid Forum where EGEE/EGI, DEISA, TeraGrid, OSG and NAREGI meet regularly to simplify their interaction.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	ESFRI project requirements for Pan-European e-infrastructure resources and facilities. https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Integration and interoperability

Summary of Requirement

The majority of the ESFRI projects consulted are unconcerned as to where resources come from, be they grids, clouds or supercomputers, and are primarily interested in easy-to-use, yet powerful and secure, data management facilities. External commercially operated clouds, which allow the user to create 'virtual computers' including the applications and operating systems of their choice, can be a very good for-purchase solution for a user who needs additional resources on demand. However, the tasks of some other ESFRI projects require computation only possible on sophisticated, high-performance computational resources such as supercomputers —currently not provided by clouds. Similarly, there are important policies questions to be addressed concerning large-scale data management and archive on commercial cloud services.

Each computing paradigm has its advantages and drawbacks, and, in the end, a custom fit solution for each user community will surely work the best. The work of standards bodies, such as Open Grid Forum, aims to make it easier to bring clouds, grids and supercomputer installations together by defining interfaces designed to simplify and promote their interoperability. Interoperability work undertaken so far among EEF partners (EGEE and DEISA), as well as EGEE with volunteer grids (such as Enabling Desktop Grids for e-Science, EDGeS) and cloud systems (via the RESERVOIR project) has been driven by the needs of users, such as the fusion and life science communities. These distributed computing solutions will continue to complement grid computing in the future and, depending on the evolution of the needs of the research communities, will become part of the European e-infrastructure ecosystem offering.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Long-term data storage and accessibility

Summary of Requirement

The ability to provide long-term (measured in decades rather than years) storage and accessibility was identified by several sectors though terminology differs. Persistent Identifiers (PIDs) and metadata are key issues for the user communities. There are existing services for registering, storing and resolving digital object identifiers (DOI) such as the handle system being offered by several consortia. Effective access to persistent data from the European e-infrastructures implies:

- Guarantees of quality of service and access for long-term storage will be required for the centers offering persistent data. To provide access implies that such centers are connected to the network (GEANT) and to ensure suitable quality of service (i.e. availability/reliability) they should be integrated into operations monitoring schemes such as those deployed by EGEE/EGI.
- The middleware deployed by European e-infrastructures will have to be modified to support access to persistent data using PIDs at these sites. Provenance of data allowing the origins of data to be recorded and traced and its movement between databases has also been mentioned by several ESFRI projects.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Consistent identity management and Single Sign On

Summary of Requirement

All the ESFRI projects consulted identified consistent identity management and single sign-on as a fundamental requirement. A unified single sign-on service has to ensure an individual's identity can be used across Network, HPC and grid services. All the e-infrastructures in EEF have existing Authentication and Authorization Infrastructures (AAI) which are similar but not identical and are separately managed. Harmonizing policies for Authentication, Authorization and potentially Accounting and Auditing will simplify access and usage to the e-infrastructures. The issue for EEF to offer what is requested by the ESFRI projects is to make these existing AAI systems interoperate so that a users identity can be established once and accepted by all the e-infrastructures.
All the e-infrastructures have dedicated security structures, procedures and measures in place intended to ensure the secure operation of the infrastructures. There are already examples of the security incidence response groups in the e-infrastructures co-operating and this will be generalized to ensure an effective and timely response to security threats can exist across the whole of the e-infrastructure ecosystem.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Training and consultancy

Summary of Requirement

All ESFRI projects have expressed the need for training, education or external expertise in their use of e-infrastructures. The EEF members all offer varying levels of existing support in this area.

GEANT provides network performance analysis expertise and has an E-Learning portal. The ICT requirements that projects have described are very broad-ranging as far as “Networks” are concerned. In technology and capacity terms, there is nothing that cannot be accommodated by current and predicated technology departments. The requirements range from multiple access of databases from a diverse population of users, through to much more concentrated flows between key project locations. The portfolio of services that is available across the GÉANT service area, including high performance IP configurable Point to Point connections and, where appropriate, dedicated wavelength capacity, is capable of meeting the needs that have been articulated. For more complex requirements, a design, and the implementation of networking needs will require cooperative effort between the research network community and the project participants. Performance, particularly where demanding applications are being supported, will need monitoring and fine-tuning. This is a non-issue and the techniques of addressing it are established. It needs to be stated, that overall performance in terms of complex systems could be challenging, as it involves interactions between different systems under separate management control. Network tools to help debug such problems are available. As part of customer support, GÉANT is prepared to analyze and diagnose performance problems.

DEISA provides general training for new users of the distributed HPC infrastructure, and has made two specific training courses for the fusion and Virtual Physical Human (VPH) communities. PRACE is offering training in all

HPC related topics. The material from workshops, summer and winter schools is made available online. EGEE has an extensive training program and material repository which is used at a wide range of events for end-users, application developers, site managers and even the trainers themselves.

Experience shows that training programs have the most impact when tailored to the specific needs of the target user community. The EEF members are willing to contribute trainers and material to training programs and events that could be organized by the ESFRI projects either individually or on a sector basis.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	User support

Summary of Requirement

To effectively use the European e-infrastructures, users should have quick responses to questions and high-quality documentation. All the large-scale European e-infrastructures of today offer specific user support facilities.

EGEE offers user support through the central Global Grid User Support (GGUS) portal via a web form or e-mail, or at their Regional Operations' Centre (ROC) or their VO which will be continued within EGI. This central helpdesk keeps track of all service requests and assigns them to the appropriate Support Units.

DEISA has a similar support structure accessible via the central DEISA Helpdesk. PRACE will offer a broad range of support ranging from first-level helpdesk to user- and community- oriented application optimization and scaling. GEANT's regional model relies on user support provided by the NRENs as well as a centralized end-to-end co-ordination capability.

The EEF members will work to make these existing user support structures cooperate and requests can be issued to appropriate support groups across the different technologies and geographical regions.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Consistent support for Virtual Organizations

Summary of Requirement

All the ESFRI projects consulted identified the ability to control access to resources, data and applications on a community level as being necessary for at least some subset of their user communities and foreseen use cases. The HPC and grid infrastructures currently offer support for virtual organizations to differing levels of granularity and with differing semantics so a goal would be to offer consistent support for Virtual Organizations by harmonizing current features.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s) **Name** **Link**	ESFRI project requirements for Pan-European e-infrastructure resources and facilities. https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Web-services interfaces

Summary of Requirement

A common theme from all the ESFRI projects consulted was the identification of web-services as a standardized manner of packaging e-infrastructure services. Many ESFRI projects highlighted the importance of well-defined interfaces for web-services, a registration facility and the ability to discover (search for) new web-services and a consistent view to managing the life-cycle of web-services. The consequence of these findings is that, where appropriate, e-infrastructures should offer web-service interfaces for their relevant services wherever possible and allow the user communities to build on these to produce their own customized web-services.

The RESPECT program (Recommended External Software for EGEE CommuniTies) publicizes and provides access to proven and useful grid software and services that work well in concert with the EGEE-produced gLite open source middleware. Third-party software packages (including commercially licensed ones) can also be integrated into the EGEE grid environment. Support for similar repositories of community specific and third-party packages that can run across the whole of European e-infrastructure will be required.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	**28 ESFRI projects** (CLARIN, ESS, DARIAH, SHARE, CESSDA, ELIXIR, BBMRI, ECRIN, EMBRC, ERINHA, Euro-BioImaging, Infrafrontier, Instruct, EU Openscreen, EISCAT-3D, EMSO, EPOS, EUFAR-COPAL, EURO-ARGO, EUSAAR-I3, EARLINET-ASOS, IAGOS, ICOS, LifeWatch, European XFEL, CTA, FAIR, SKA); **2 VO projects** (eNMR, neuGRID)
Requirement Type	☐ Support ☐ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	ESFRI project requirements for Pan-European e-infrastructure resources and facilities.
Link	https://documents.egi.eu/document/12
Requirement *Unique name identifier*	Workflows

Summary of Requirement

The need for workflows was identified by all ESFRI research sectors. There are many workflow tools or frameworks employed by the user communities and this diversity is certain to remain. As a result, the EEF members will work to provide an environment which can support a range of workflow tools and environments. The implication of supporting such cross infrastructure workflows seamlessly is that the AAI, virtual organization and data management inter-operation aspects mentioned in previous reported requirements must be in place.

7.3 From OGF Production Grid Infrastructure Use Case Collection

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Project VPH(Virtual Physiological Human)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	VPH(Virtual Physiological Human) (Use Case)

Summary of Requirement

- Storage resources for the interaction of the jobs are necessary and should be advertised.
- Furthermore NAMD is needed for the simulations. It would be, thus, nice to be able to ask an information service regarding software availability.
- File transfer services such as GridFTP.
- Advertising of services and endpoints is also necessary.
- The ability to submit multiprocessor jobs (MPI) is also needed.
- Authentication and authorization services for user authentication and file transfers

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Different research domains as multimedia, finance, archeology, civil protection, astronomy, astrophysics, computational chemistry, earth sciences, fusion, and computer science.
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Efficient MPI-based Parallel Job Executions (Use Case)

Summary of Requirement

Although the OGSA-BES and JSDL specifications provide functionalities for creating and managing jobs on remote computational resources, they don't provide full support for job parallelism activities. A lot of profiles are available to fill the gap but they are often a complex mechanism that does not encourage the user to use it. Also, these profiles or standard extensions do not allow using the massively parallel systems in a very efficient manner that could be done by better specifying hardware setups (e.g. network topologies).
A better solution would be to extend the OGSA-BES and JSDL specifications by integrating the profiled items or, create a new specification which covers all these aspects.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Trust Anchor, Security Engineer of a Site, User Domain Manager, Grants access rights to the User, RA (Registration Authority), CA (Certificate Authority), CSIRT (Computer Security Incident Response Team)
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Enforce Security of a Production Service Grid (Use Case)

Summary of Requirement

– Robust standardized Information System describing Grid entities to be monitored for security,
– Robust standardized Security design: Credentials definition, Setup, Process...
– Robust standardized Monitoring permitting to quickly detect security issues,
– Robust standardized Logging and Accounting holding persistent records permitting easy Security audits.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Application Gridification Team, DG Application Validation Team, UserDomain Manager, Bridge
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☒ Other : EDGI
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Prepare Desktop Grid Version of Application (Use Case)

Summary of Requirement

– Standardized Information System describing Grid entities able to accepts Activities,
– Standardized Security design permitting to associate specific Resources to Credentials
– Standardized Application Repository.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Any User
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☒ Other : EDGI
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Marshal Activities from Service Grid to Desktop Grid (Use Case)

Summary of Requirement

– Robust standardized Information System describing Grid entities able to accepts Activities,
– Robust standardized Security design permitting to associate specific Resources to Credentials,
– Robust standardized Application Repository permitting Users to pick an Application version which is adequate for execution inside a Desktop Grid.
– Robust standardized Logging holding persistent records permitting Users to audit job history (for example in case of error).
– Robust standardized Accounting holding persistent records permitting Users to analyze resource usage.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	many different disciplines: physics (HEP particularly), chemistry, computer science, computer engineering, systems engineering, biochemistry, biology, economics, public health science, and chemical engineering—just to name a few.
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	Production Grid Infrastructure - Use Case Collection v1 http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	High Throughput or High Capacity Computing (Use Case)

Summary of Requirement

This use case relies on the ability schedule, start, stop, monitor and manage jobs of various types as well as staging data to and from jobs and accessing intermediate job data.
At the moment, the existing BES, JSDL, and JSDL extension specifications can be used to support most of functionality required for the execution of, and management of jobs as well as staging data to and from jobs.
One area not directly covered by the existing specifications, is the ability to access the intermediate data of a job in progress. However, there are several existing specifications that could be adopted to provide part of this functionality. In particular, the RNS specification can be used to model the directory structure of the job's working directory and provide an interface to access and manipulate directory entries; the ByteIO specification(s) can be used to model data files and provide the interface to manipulate them; and the BES specification could be extended to provide a mechanism to access the root of a job's working directory – for example, by profiling BES to define the interface for the BES Activity port type to include this functionality or by adding a new port type to the BES specification.
Additional profiling or extensions of JSDL and BES may be required for some functionality or to increase interoperability between grid implementations.
From a security perspective, existing OGSA security profiles and specifications cover many, but not all of the use case.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	wide range of users from many disciplines
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Mid-Range Computing (Use Case)

Summary of Requirement

This use case relies on the ability to create and manage jobs on remote resources. In this case, there seems to be very few functions that aren't handled by existing BES, JSDL, and JSDL related specifications.

Some amount of profiling may be desirable to achieve/encourage interoperability such as specifying mandatory supported data staging protocols and/or data staging protocol discovery for BES implementations.

From a security perspective, existing OGSA security profiles and specifications cover most of the use case.

Functional requirements:

Resource access: Users of mid-range computing applications typically submit jobs directly, may make use of workflow tools or may use gateways to submit jobs.

Authentication, authorization, and accounting (AAA): These users require standard Unix and batch AAA as implemented at individual resource provider sites and following the policies those sites adhere to.

Fault tolerance: Some mid-range computing applications have built-in periodic checkpointing features to guard against loss of work in the event of a system crash. Some mid-range jobs may employ workflow restart capabilities for fault tolerance.

Scheduling: Scheduling the job consists of at least three logical phases: determining where the job can run based on resource and account requirements, selecting a resource on which to execute the job, and preparing the execution environment for executing the job (e.g., staging data, getting the binaries in place, etc.) Determining where the job may run may require an information system that describes execution environments, their capabilities, software installed, etc. The ability to start both single and "vector" batches of jobs should be supported.

Advance scheduling: Some mid-range computing applications may require special scheduling capabilities including pre-stage and post-stage data job scheduling or sequential job dependencies.

Workflow: Some mid-range computing applications have workflow components. Most of these workflows are managed by the users using homegrown tools or special workflow tools.

Data storage: The data sets produced by mid-range computing applications can range from small to quite large. The storage and archival resources available must be both sufficiently large to store such data sets as well as sufficiently fast to handle them in a timely manner.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	users with deadlines for work
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Special Quality-of-Service (QOS) Computing (Use Case)

Summary of Requirement

This use case relies heavily on the ability to make reservations of resources (be they compute, or data), and at the proper time, execute jobs on those resources. Further, for the urgent computing aspect of this use case, the ability to manage jobs already running (either to kill, suspend, checkpoint, etc.) is needed.
t the moment, the existing BES, JSDL, and JSDL extension specifications can be used to support the execution of, and management of jobs. Since most of the functionality for execution in this scenario is implementation and not interface (i.e., the interface to BES doesn't have to change to take into account urgent computing needs), few if any extensions are necessary. However, extensions would be needed to support the reservation aspect of this use case. For this, a reservation port type (interface) would be needed that could then be folded in to an appropriate JSDL extension to support the reservation. The back end BES implementation would of course need to be able to handle this JSDL extension and the reservation itself (though its not necessarily the case that the BES implementation has to manage the reservation, merely be aware of and work with it).
From a security perspective, existing OGSA security profiles and specifications cover most of the use case.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	The consumers of the GROMACS technology
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	GROMACS-based Molecular Dynamics in Bio-molecular Systems (Use Case)

Summary of Requirement

Support for MPI tasks, like those executed by GROMACS, in Grid environments in general is rather poor: neither information systems, nor execution services offer adequate functionalities. Many Grid solutions also suffer from inferior data handling. PGI is expected to extend existing standard specifications such that they accommodate GROMACS use case.

Attractiveness of Grids for GROMACS users lies primarily in potential availability of several resources to choose among, all accessible in a uniform manner (e.g., SSO), and all offering identical (or near identical) execution environment, such that users don't have to adapt their workloads for each individual resource.

Therefore, PGI is expected to deliver common standards for:

- authorization and related security aspects,
- publishing information relevant for GROMACS services discovery, such as e.g.:
  - presence of MPI support as such
  - details of the software environment
  - hardware configuration
- execution service interface capable of
  - supporting data-intensive tasks
  - supporting MPI tasks
- job description language, capable of specifying GROMACS workflows, including:
  - required input and output data locations,
  - required application software,
  - required resource usage (CPU, memory, disk space, connectivity) taking into account MPI aspects

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	many bio-informatics communities that actually seek to perform computational bioscience without the need to understand the complexity of the underlying escience infrastructures or computational paradigms.
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	Production Grid Infrastructure - Use Case Collection v1 http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Multi-Grid Drug Discovery Workflow (Use Case)

Summary of Requirement

Production Grid-driven realistic reference model based on open
Standard: Although we used several standards in this drug discovery use case (OGSA-BES, JSDL, GLUE2, GridFTP, security profiles, etc.) their usage as a whole ecosystem, so to say, was rather unclear. This mainly includes computing, data,
security standards as well as information flow aspects and standards. A reference model or greater realistic architecture would be important.

Secure Interaction: Connections between Grid technologies should be only based on SSL/TLS (i.e. https) connections avoiding the
need for dedicated technology protocols as seen previously with the Grid Security Infrastructure (GSI) (i.e. httpg). The
concept of end-user rights delegation should be not strictly bounded together with the connection mechanism (as done
within GSI).

Grid Application Improvements: Grid application job descriptions satisfied basic needs in this use case but were not satisfactory enough to describe
an application in this multi-Grid setup. Some improvements covering but are not limited to application types classification (e.g. parallel, etc.), application type refinements (e.g. pre-installed, submitted, etc.), revised application executable definition, application software statements, application family extension (e.g. LIBRARY), application software requirements, application output joins, etc.

Application Execution Adjacencies: In this workflow, we had several challenges in the execution environment itself. Thus we need better support for scientific application executions with standard-based information aspects on the lowest possible level (i.e. resource management system level), including common environment variables, execution modules as well as module characteristics, etc.

High Performance Computing Extensions to open standards: While runs using CREAM-BES on EGEE had been relatively ok, submission with UNICORE-BES to DEISA
lacked important HPC specific information. Therefore, we seek to submit and execute applications more efficiently
than currently with GLUE2, JSDL, or OGSA-BES. For instance, required aspects are support for network topologies (torus, global tree, Ethernet, etc.), shape reservations (x X y X z), network information enhancements, available shape characteristics, high message support, task/core mapping definitions, available task/core mappings exposure and re-use, etc.

Sequence Support for Computational Jobs: An analysis of lessons learned obtained from the WISDOM use case leads to specific missing features encountered
during the production Grid interoperability with respect to the support of automatically started pre- and postprocessing
functionalities within JSDL using different application execution modes. AMBER, for instance,
consists of a set of applications (~80 executables), where some of them are used to transform input data in a suitable
format for production runs and/or transform outputs in several other formats necessary for further analysis. Of course, these transformation and short running pre- processing steps should be executed in a serial mode, while the actual corresponding molecular dynamic simulation is executed in a parallel mode. Pre-job sequences (pre-processing, compilation), Post-job sequences (post-processing).

Manual Data-staging support: A careful analysis of many lessons learned from this production cross-Grid application use case that takes
advantage of EGEE and DEISA resources, revealed that in many cases the scientists require a more flexible mechanism during data-staging processes in order to better coordinate distributed data and computation. This is
true, irrespective of whether data is transported to where the computational resource resides, or if computation is
decomposed and job submissions are performed towards the location of data or even a hybrid of both methods is
adopted. One example was the careful manual data input selection (aka manual data-staging) from the outcome of
the EGEE workflow step in order to use only good results for the time-constrained workflow step in DEISA.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	The research group of this use case is geographically distributed in Japan. They have the demand for using heterogeneous architecture and large scale computing resources for running their applications effectively. Each application communicates with other applications using GridMPI. Therefore, the customers have a demand for using a high-bandwidth/low-latency dedicated network between geographically distributed computing resources.
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 Production Grid Infrastructure - Use Case Collection v1 http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	RISM-FMO Coupled Simulation (Use Case)

Summary of Requirement

- Advanced resource reservation
- Inform advanced reservation (scheduling) table for meta-scheduler and work load manager
- GridMPI support
- Inform allocated worker nodes list (for each GridMPI activity)

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	End-users, Scientific application developers, Resource providers
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	Data intensive computation job processing on cluster-based Grids (Use Case)

Summary of Requirement

he use case is currently largely addressed by the ARC middleware, and to a large extent – by the gLite middleware as well. However, submission of this kind of tasks across different middlewares is very problematic due to absence of common interfaces. On application level in the infrastructures like WLCG this lack of interoperability and desired functionality is solved in the very unorthodox manner: all HEP experiments deploy agent-based (a.k.a. pilot) jobs, which represent dummy payloads without data and with a trivial “transformation” which has a task of fetching the actual payload from an application-specific service (job pool). While being a very convenient and flexible framework for applications, this approach creates multiple problems for resource owners, and does not help application developers at all; neither does it promote standardization.
By delivering a common execution service interface and job description complete with well-defined and flexible data staging capabilities, increased user friendliness and better fault tolerance, PGI will reduce the burden on application developers, which in turn will attract more user communities to the Grid.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Potentially, all European academic6 researchers engaged in e-Science; Potentially, all European providers of research computing services; European Grid technology providers, primarily through EMI, IGE and EDGI projects;
Requirement Type	☒ Support ☒ Infrastructure ☒ Operations & Tools ☒ User applications & Tools ☐ Other : ……
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Production Grid Infrastructure - Use Case Collection v1
Link	http://www.ogf.org/Public_Comment_Docs/Documents/2010-10/PGI_use_cases_v3.pdf
Requirement *Unique name identifier*	European Grid Infrastructure (Use Case)

Summary of Requirement

As EGI encompasses all possible services used by national Grid infrastructures across Europe, it requires standardization on all levels: computing interfaces, data management, information, security etc. At the moment, arguably the only commonly implemented standard is SRM 2.2. Production Grid Infrastructure is expected to deliver common profiles for all other relevant services.

7.4 From "gLite enhancements requests" deliverable of the eNMR project

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	VO : enmr.eu
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s) **Name** **Link**	 Request for Enhancements of gLite to support bio-NMR applications http://www.enmr.eu/files/Deliverable3.3.pdf
Requirement *Unique name identifier*	Quality of grid service

Summary of Requirement

The e-NMR community performs periodic surveys to their user community, in order to gather feedback which is critical to improve their services. Moreover, it tries to pinpoint the bottlenecks per application ported to the Grid by the e-NMR software managers. One of the survey questions was:

> "Did you find any limitation on gLite current performances?"

The replies gathered clearly indicated that "the performance relies very much on the reliability of the GRID (disappearing or failing jobs cause severe delays and reduction of performance)". This requirement can be summarized like this: reliable grid middleware; quality of Service, especially when submitting a huge number of jobs.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	VO : enmr.eu
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 Request for Enhancements of gLite to support bio-NMR applications http://www.enmr.eu/files/Deliverable3.3.pdf
Requirement *Unique name identifier*	Data availability

Summary of Requirement

An anonymous on-line survey (targeting both academia/researchers as well as a number of biotech and pharmaceutical companies) showed that “easy access to data and databases; easy to use, standardized API and fine grained access policies” is a critical requirement.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	VO : enmr.eu
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s)	
Name	Request for Enhancements of gLite to support bio-NMR applications
Link	http://www.enmr.eu/files/Deliverable3.3.pdf
Requirement *Unique name identifier*	Authentication and authorization

Summary of Requirement

Since e-NMR platform relies heavily on the use of web portals, to hide the grid complexity to their end-users, all aspects concerning authentication/authorization for usage in portals are critical to this community.

An overview of the requirements regarding this subject:
Common ways of authentication and authorization: Standardized Authentication and authorization mechanisms for usage in portals as well as in direct Grid access. A globally accepted, trustworthy Grid user identity or 'Passport' with an infrastructure providing it is part of this.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	VO : enmr.eu
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☒ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Request for Enhancements of gLite to support bio-NMR applications
Link	http://www.enmr.eu/files/Deliverable3.3.pdf
Requirement *Unique name identifier*	Encryption and protection of data

Summary of Requirement

From the surveys done by this community, encryption and protection of data on the grid storage elements is also an important requirement, especially for industry users. Notice however that the e-NMR project is aware of some solutions provided by EGI to cope with this requirement, such as the Encrypted Data Storage system based on Hydra keystores.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	VO : enmr.eu
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☒ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	Updated Request for Enhancements of gLite to support bio-NMR applications
Link	http://www.enmr.eu/files/Deliverable3.7.pdf
Requirement *Unique name identifier*	MPI support

Summary of Requirement

e-NMR hosts some applications which could potentially benefit from more advanced gLite capabilities like MPI support. The consortium is closely following the activity of the MPI working group set up in EGEE and planned to continue its effort in the context of EMI, so that future new developments can be quickly incorporated in the e-NMR grid as soon as ready for deployment.

7.5 From documents about training and education

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	EGI.eu UCST
Requirement Type	☐ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☒ Other : Digital Library
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	"Distributed Computing Education, Part 5: Intellectual property Rights" Elizabeth van der Meer, Malcolm Atkinson and David Fergusson,, IEEE Distributed Systems Online, (2008),
Link	http://www2.computer.org/portal/web/computingnow/dso/1208/education
Requirement *Unique name identifier*	License & Copyright for The Digital Library

Summary of Requirement

• Licenses produced by the Creative Commons should be recommended for digital repositories containing e-science education and training materials.
• Where appropriate, depositors should be permitted to upload their own usage licenses, but when this is done, the repository must make all reasonable efforts to explain this usage license to users downloading the material.
• A separate deposition license or agreement is also required and must be clearly displayed whenever a depositor attempts to upload new materials; depositors should have to explicitly agree to this license before they can upload materials.
• Any digital repository of e-science education and training materials should make it easy for users downloading materials to cite them. This can be done by including appropriate metadata containing, for example, the details of the author and copyright holders. This is particularly important if the materials themselves don't identify the author, which is often the case.
• All metadata should be based on international standards such as Dublin Core to facilitate querying, harvesting, federation, and so on.
• Repositories should let consumers search materials by license type or with a license type filter in place to ensure that materials available under a particular license are easy to find. This should reduce the danger of, for example, a consumer accidentally modifying and distributing materials that were originally distributed under a license prohibiting this.

• Repositories should provide depositors with an FAQ or short explanation of common problems that might arise when depositing materials. For example, it should point out that the depositor's employer might hold the copyright for the materials, and it should remind depositors to check the licenses of any material that they might have reused in creating their material.
• The repository must include a simple mechanism to remove or block materials and derivative works in case of disputes.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	EGI.eu
Requirement Type	☐ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☒ Other : Training and Education
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s)	
Name	"Strategies and Policies to Support and Advance Education in e-Science" "Curricula Development for e-Science: Meeting the Challenges" "Distributed Computing Education, Part 1: A Special Case?" "Distributed Computing Education, Part 4: Training Infrastructure"
Link	https://www.humyo.com/FRnSjhN/04-Project%20Activity/WP3/User%20support%20services/Training/DavidFergusson/Atkinson_y2009.pdf?a=s5zXUYQkdlo https://www.humyo.com/FRnSjhN/04-Project%20Activity/WP3/User%20support%20services/Training/DavidFergusson/Atkinson_y2009a.pdf?a=Xs7L0CSFapA http://csdl2.computer.org/comp/mags/ds/2008/06/mds2008060002.pdf http://csdl2.computer.org/comp/mags/ds/2008/10/mds2008100002.pdf
Requirement *Unique name identifier*	Training and Education multiple Requirements

Summary of Requirement

- establish a committee of leading educators across disciplines to expedite the creation of curricula goals and principal topics, launched and supported by major conferences highlighting educational priorities and opportunities in the field.
- continue meetings in international contexts, such as the Curricula Development Workshop in Brussels and community group meetings held at OGF 22 and 23, to develop an understanding of educational goals and curricula.
- support e-Science centers in identifying role models to interact with students, and encourage universities to nurture grassroots groups to advance e-Science education.
- continue to build federated repositories of shared experiences and practice in e-Science education.
- teaching students underlying processes, concepts, and critical thinking,
- practical expression of these processes and concepts through examples relevant to each

student's discipline

- collaboration between computer scientists and application scientists to achieve a successful outcome.

Teaching environments must provide:

1. materials to let educators rapidly take up this task (such as textbooks and repositories of teaching materials),
2. education for the educators to help them bootstrap in this area,
3. seed teaching infrastructures to be used initially for teaching and to become templates for organizations in developing their own infrastructures, and
4. a supportive policy framework that encourages cooperation and sharing (for example, in managing intellectual property rights or sharing teaching infrastructures).

T-Infrastructure must provide:

- Convenient provision of authentication and authorization is crucial. This can be achieved by using automatic provision via Web interface to users, batch provision to educators for courses and training events, and provision of certificates via prerequisite self-paced exercises (summer schools).
- The provision of directly accessible training material should be prioritized. We provided training material via trainer-maintained wiki pages on GILDA and through the ICEAGE and EGEE digital libraries. Trainer materials should be strictly related to and tested on the tinfrastructure (limiting the problem of adapting a generic procedure to a specific context).
- Developers of training infrastructure should use virtual machines for support systems administration training and for provision of reconfigured middleware as a downloadable, installable context for users to experience.
- Web portals ameliorate interface complexity, as we identified specifically in the cases of GENIUS

8 APPENDIX B – REQUIREMENTS FOR MIDDLEWARE ENHANCEMENT FROM HUCS AND VOS

The question asked to VOs and HUCs was the following:

> *The European Middleware Initiative project (EMI) is collecting requirements to define the development plans for the ARC, gLite and UNICORE middleware stacks for the first year of EMI. (until April 2011) EMI year 1 developments focus on service enhancements (security, consolidation of services and libraries, usability) instead of adding completely new functionalities. The EGI user community is expected to provide input to this plan. Please list and explain what enhancements your VO/HUC would like to see in the new middleware release.*

The received answers follow below and are available at [R5]:

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
Heavy User Community – Life Sciences

Requirement Type

☐ Support
☐ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☒ Other : UMD stack

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name HUCs Survey
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
UMD stack

Summary of Requirement

The list below is prioritized from most to less critical.

- Coherent/homogeneous set of APIs for all middleware services (WS and java), that can be installed on a client without a complete UI (ie an non-SL host)
- No single-point-of-failure services (VOMS and LFC currently are critical and single point of failures), i.e. redundancy of critical services is needed.
- LFC administration tools: tooling for SE decommissioning or maintenance procedures, including automatic replication/migration of files if possible and automatic notification of file owners in all cases.
- DMS workload management: avoid overloading / time-out of data access that cause many job failures (these should be postponed and retried until success)
- Management of many small files and improved replica selection.
- Middleware support for pilot jobs
- single sign-on through Shibbolet or other federated identity mechanism to avoid exposing users to X509 certificates

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Tristan Glatard on behalf of biomed VO and LS community
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☒ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	Savannah: https://savannah.cern.ch/bugs/?func=detailitem&item_id=74976 GGUS: https://gus.fzk.de/ws/ticket_info.php?ticket=63103
Requirement *Unique name identifier*	use of GlueSEStatus BDII attribute

Summary of Requirement

Dear gLite developers,

Downtime notifications are quite difficult to process by VOs due to the lack of automated tools to handle it. In case of SEs these downtimes can be critical since they have a direct impact on the availability of files, therefore on the VO production.

Would that be a possibility that SEs publish their status (downtime, decommissioning, etc) using the GlueSEStatus BDII attribute? That would allow VOs to efficiently process this information with lcg-infosites.

This issue is already discussed there:
https://gus.fzk.de/ws/ticket_info.p...
I would be happy to elaborate on that if needed.

Best regards,
Tristan, for the biomed VO

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Tristan Glatard on behalf of biomed VO and LS community
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☒ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	Savannah: https://savannah.cern.ch/bugs/?func=detailitem&item_id=73846 GGUS: https://gus.fzk.de/ws/ticket_info.php?ticket=62886
Requirement *Unique name identifier*	dCache's grid-mapfile is cleared when VOMS is not online

Summary of Requirement

Hello,

It seems that the grid-mapfile of dCache is cleared from all VO entries when the VO's VOMS server is not accessible. In GGUS ticket #62886, the site admin suggests that this should be fixed in dCache by keeping the old grid map file when the VOMS server is not responding rather than removing all the entries.

I guess that this is also a valid comment to check for other SE implementations. Where should this request go?

Thanks in advance,
Tristan, for the biomed technical teams

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Tristan Glatard on behalf of biomed VO and LS community
Requirement Type	☐ Support ☐ Infrastructure ☒ Operations & Tools ☐ User applications & Tools ☐ Other : .....
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical
Requirement Source(s)	
Name	
Link	GGUS: https://gus.fzk.de/ws/ticket_info.php?ticket=63425 https://gus.fzk.de/ws/ticket_info.php?ticket=62854
Requirement *Unique name identifier*	establishing a strong site decommissioning procedure

Summary of Requirement

Here is a new pair of tickets conveying a clear requirement for us: establishing a strong site decommissioning procedure.
https://gus.fzk.de/ws/ticket_info.php?ticket=62854
https://gus.fzk.de/ws/ticket_info.php?ticket=63425

This incident costed some significant manpower and we still lost 12,355 files, not to mention the impact on user expectations.

As already mentioned during the OLA session at the TF we as the biomed VO are willing to give technical feedback on the decommissioning procedure.

Best regards,
Tristan, for the biomed VO

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
Heavy User Community – WLCG

Requirement Type

- ☐ Support
- ☐ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☒ Other : UMD stack

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name HUCs Survey

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
UMD stack

Summary of Requirement

The timescale of the first EMI release is not convenient for the LHC running schedule in 2011. It should not be expected that WLCG will pick up a release on this timescale - but I don't want to include a long discussion on the accelerator cycle here.

I also understand from discussions with Markus and Maria A that there are supposed to be direct channels between WLCG and EMI for discussing these requirements where I am sure more detailed info can or perhaps has been provided.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: CALICE

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
WMS stability and performance

Summary of Requirement

The WMS is unstable and slow.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: cdf

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other …….:

Priority
Self-assessment from report
☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Demand for WMS and enhancements at EMI

Summary of Requirement

CDF uses the glite WMS for its job submission on the European Grid. We are interested in every enhancement of this component, such as implementing site feedback and automatic resubmission of jobs that are waiting too long at a site's queue. We are already in touch with the WMS developers about this, and we are trying to test some new features in our job submission environment.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: desktopgrid.vo.edges-grid.eu

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other …….:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Demand on LCG-CE, stability and proper development

Summary of Requirement

Requests for UMD enhancements:

1. Prefer to see better tested, more stable releases with less bugs in each new release over new features or services.
2. Development should focus on stabilising services and fix known bugs first before considering enhancements.
3. We would like to see proper support in yaim for publishing different information about multiple sub-clusters. This is to complete this development: http://goc.grid.sinica.edu.tw/gocwiki/How_to_publish_different_memory_limits_for_different_queues_on_the_same_CE
4. YAIM update and config overwrites custom settings, the same custom configuration must be set manually after every service update

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/…*	VO: enmr.eu

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name	VOs Survery
Link	https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement *Unique name identifier*	SSO and other enhancements

Summary of Requirement

most important issue is SSO via Shibbolet or other federated identity mechanism, in order to avoid the users of bio-NMR community of bothering with X509 personal certificates.

More detailed request for enhancements are described in two project deliverables publicly available at http://www.enmr.eu/files/Deliverable3.3.pdf and http://www.enmr.eu/files/Deliverable3.7.pdf

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: fusion

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Stabillity of UMD services and improvements

Summary of Requirement

The most important thing is The general stability of The services in order to avoid as many aborted jobs as possible;
Avoid WMS high workload when submitting great number of simulations;
Obtain a quicklier refresh time of The status of a job. Sometimes it takes too long to get an updated status of a job which makes users think things are not going properly.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: gridmosi.ici.ro

Requirement Type

☒ Support
☐ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Comparison of middlewares , middleware transition affects on UMD

Summary of Requirement

So far we used gLite middleware only, we don't have much experience with other solutions like Unicore or Arc or Globus.
We would like to switch to Unicore for example, and we will like to have more documentation regarding the transition of site from gLite to all the other middleware stacks.
A functional comparison per component, with gLite components as a reference, and the other middleware stacks having a similar functionality in the mirror, would be very useful.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: HONE

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other .....

Priority
Self-assessment from report
☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Demand for WMS , stability of CREAMCE, gLite reliability at EMI

Summary of Requirement

HONE presently relies on submitting via the WMS and therefore depends on that service.
HONE has adopted the workflows to existing gLite functionality. For us improvements in reliability are of higher importance then feature enhancements.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: ILC

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
WMS and error handling enhancements

Summary of Requirement

* The WMS is unstable and slow.
* Error messages and Error status needs to be improved for the middleware commands.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: ildg

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Improvements on Integration of infrastructures

Summary of Requirement

Improved integration of unicore based infrastructures like DEISA and glite based infrastructures like EGEE/EGI

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: lofar

Requirement Type

☐ Support
☒ Infrastructure
☒ Operations & Tools
☐ User applications & Tools
☐ Other ……. :

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
gLite components enhancements and improvements of VO management

Summary of Requirement

* Something that does not work is the option to claim a whole node via the middleware. The option of using the GLUE attribute “LogicalCPUs” does work for a number that is smaller than the total number of CPUs (cores) per node, but the job will remain in the queue forever when it asks for the total number of CPUs (cores) per node. It will never be executed. This is quite problematic because LOFAR wants to use this broken feature.

* Enhanced information feedback by middleware error messages. At the moment, many of the error messages leave the user clueless about the real problem.

* Enhanced WMS throughput. In the current WMS implementation, each request takes 6 seconds to be processed.

* Enhanced SRM performance. The current SRM implementations are not able to handle more than two requests per second.

* JDL does not offer GLUE attribute for available disk space on worker node. The user has NO means to select worker nodes that have the required amount of local disk space. This is a basic functionality expected by an ordinary user, and should not be missing.
* LFC reliability is currently not acceptable, i.e. the LFC can not be trusted. Failing registrations of

files are not communicated to the end user by the middleware, which means that files are lost in practice (file exists, but reference is lost). Vice versa, failing file transfers are not communicated to the end user by the middleware, which means that non-existing files are registred in the LFC (reference exists, but file is lost). In case the LFC registration or the file transfer fails, the complete action should be rolled back by the middleware and an error message (failed) sent to the user.

In addition (this is however not a middleware issue) it is not clear to the LOFAR VO manager what tasks are expected from a VO manager and what options/services are available to the VO manager. Is there a document providing guidelines on this? For example, how to retrieve and distribute information about (un)scheduled downtimes of relevant sites, what information to provide to new users about getting GGUS support, when should issues be communicated to EGI and when to the involved NGIs (lofar is a pan-european organisation that involves multiple NGIs), etcetera.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: lsgrid

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
gLite components enhancements

Summary of Requirement

* Enhanced information feedback by middleware error messages. At the moment, many of the error messages leave the user clueless about the real problem;

* Enhanced WMS throughput. In the current WMS implementation, each request takes 6 seconds to be processed.

* Enhanced SRM performance. The current SRM implementations are not able to handle more than two requests per second.

* JDL does not offer GLUE attribute for available disk space on worker node. The user has NO means to select worker nodes that have the required amount of local disk space. This is a basic functionality expected by an ordinary user, and should not be missing.

* LFC reliability is currently not acceptable, i.e. the LFC can not be trusted. Failing registrations of files are not communicated to the end user by the middleware, which means that files are lost in practice (file exists, but reference is lost). Vice versa, failing file transfers are not communicated to the end user by the middleware, which means that non-existing files are registred in the LFC (reference exists, but file is lost). In case the LFC registration or the file transfer fails, the complete action should be rolled back by the middleware and an error message (failed) sent to the user.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: nordugrid.org

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other …….:

Priority
Self-assessment from report
☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Proper gLite services support and improvements and development at EMI

Summary of Requirement

Our VO welcomes the fact that ARC components will become officially available for the EGI community, and we hope that many more European sites will deploy it. We also expect adequate operational support from EGI to sites and user communities that chose ARC components. We expect that EMI will make possible to include ARC services into the European Grid infrastructure in the manner identical to gLite and UNICORE ones, such that users will be able to enjoy the larger overall resource. Specifically, we anticipate that it will be possible for our VO to negotiate access to gLite- and UNICORE-powered services, and that it will be possible to access gLite- and UNICORE-powered services using ARC client tools familiar to our users. Such services include e.g. computing, storage, athorisation, monitoring, accounting etc.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/...
VO: pheno

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other .....
:

Priority
Self-assessment from report
☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Proper gLite services support and improvements and development at EMI

Summary of Requirement

no problem as long as glite command-line tools do not arbitrarily change their interface (command name, option flag names, etc)

Better quality control with the goal of a coherent user interface of command line tools. At the moment there is no coherent set of option flags. For example, depending on the tool, we have sometimes '-all', sometimes '--all'.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: vlemed

Requirement Type

☒ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Data management and middleware improvements

Summary of Requirement

Our VO is only making use of gLite.
The vlemed VO would be interested in a EGI / EMI led discussion for the improvement of data management facilities.
Some other improvements that would be really welcome:
- finer grained access controls
- enforcement of access control from the catalogue to the physical files and v.v.
- data lifetime management
- stable, documented and supported API to middleware (python, java) with proper error handling

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…

VO: vo.complex-systems.eu

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☒ User applications & Tools
☐ Other …….:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier

Better support for MPI applications and other type of jobs

Summary of Requirement

Better support for MPI jobs and the MPI-START mechanism; support for hybrid jobs (i.e. mixed MPI,openmp); Better support for perusal mechanism; Better DAG and Parametric jobs manipulation; Provide a VO software administration framework

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: vo.iscpif.fr

Requirement Type

- ☐ Support
- ☒ Infrastructure
- ☐ Operations & Tools
- ☐ User applications & Tools
- ☐ Other …….:

Priority
Self-assessment from report
☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery

Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Demand on efficient virtualization system and data system

Summary of Requirement

We would like to have an efficient virtualization system on each worker node (kvm) to run user defined virtual machines (VM). Users have software with complex dependencies which may be hard to package for SL5 or for any other predefined OS distribution. We would like to enable the user to provide his own home brewed specific VM at job submission time to ease the migration of existing applications to the grid and to improve reliability.

We would like to use IRODS for datamanagement in our VO.

EGI.eu UCST requirement identification

Document ID
EGI.eu UCST internal Doc. ID

Requestor(s)
VRC(s)/VO(s)/ESFRI Project(s)/…
VO: VOCE

Requirement Type

☐ Support
☒ Infrastructure
☐ Operations & Tools
☐ User applications & Tools
☐ Other …….
:

Priority
Self-assessment from report

☐ Unknown ☐ Low ☐ Medium ☐ Important ☒ Critical

Requirement Source(s)

Name VOs Survery
Link https://wiki.egi.eu/wiki/Requirements_gathering_details#Surveys_for_user_communities

Requirement
Unique name identifier
Proper MPI support

Summary of Requirement

There is a intensive request concerning support of MPI. The primary deployment requirements for sites supporting VOCE VO are as follows:

- Implemented MPI support (according to MPI Working Group recommendations)
- Support of Open MPI
- Support of MPI start scripts

In parallel, a request regarding MPI deployment at sites supporting VOCE VO to guarantee ssh-like interconnection of worker nodes assigned to a specific computational jobs is following:

- Passwordless SSH login among worker nodes

9 APPENDIX C – INPUT FROM NGIS ON USER SUPPORT ACTIVITIES AND SERVICES

The question asked to NGI User Support Teams was the following:

Could the performance of your User Support Team be improved by additional services, resources and tools? If yes, what kind of services, resources and tools are these?

The received answers follow below and are available at [R5]:

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	MetaCentrum, CZ NGI;
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 NGI UST questionnaire EGI.eu humyo document
Requirement *Unique name identifier*	High-level middleware frameworks

Summary of Requirement

Specific user communities can advantageously benefit from high-level middleware frameworks as those available under RESPECT program. However, this requires tight connection with the targeted community, provision of corresponding consultancy and subsequently efforts invested into customization and modification of available tools according to end users needs and requirements.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	GRID-Ireland, NGI_IE;
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 NGI UST questionnaire EGI.eu humyo document
Requirement *Unique name identifier*	Comparison of portals and frameworks

Summary of Requirement

The GRID-Ireland NGI mentions that their User Support Team will benefit hugely from “an evaluation, comparison and contrast table of various portals and frameworks”, since this “would ease the burden in performing costly local evaluations”.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	GRID-Ireland, NGI_IE;
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☒ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☒ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 NGI UST questionnaire EGI.eu humyo document
Requirement *Unique name identifier*	Divulgate available user tools and support

Summary of Requirement

GRID-Ireland NGI requests sessions at EGI User and Technical Forums, about tools for user and application support.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	ISRAGRID, IL NGI;
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 NGI UST questionnaire EGI.eu humyo document
Requirement *Unique name identifier*	User knowledge base

Summary of Requirement

This NGI stresses the need of having a good online wiki where problems and solutions could be easily accessed by everyone. They also pinpoint the need of proper management of questions and answers from the users.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	ISRAGRID, IL NGI;
Requirement Type	☐ Support ☒ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☒ Unknown ☐ Low ☐ Medium ☐ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 NGI UST questionnaire EGI.eu humyo document
Requirement *Unique name identifier*	Improve grid interface

Summary of Requirement

This NGI recognizes the need to have better interfaces to the grid system, since current “interface is hard to use, sensitive and gives cryptic errors”.

Note:

The requirement is largely based on our interpretation from the input given by ISRAGRID, since the information provided is not very clear.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	RU NGI;
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☐ Other :
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s) **Name** **Link**	 NGI UST questionnaire EGI.eu humyo document
Requirement *Unique name identifier*	Localised User Support Tools

Summary of Requirement

We strongly need localization of supporting tools.

EGI.eu UCST requirement identification

Document ID *EGI.eu UCST internal Doc. ID*	
Requestor(s) *VRC(s)/VO(s)/ESFRI Project(s)/...*	Nordic NGIs
Requirement Type	☒ Support ☐ Infrastructure ☐ Operations & Tools ☐ User applications & Tools ☒ Other : EGI-InSPIRE Technical Services
Priority *Self-assessment from report*	☐ Unknown ☐ Low ☐ Medium ☒ Important ☐ Critical
Requirement Source(s) **Name** **Link**	
Requirement *Unique name identifier*	Localization of User support tools and services

Summary of Requirement

The user technical services should be localized for NGIs, to allow the customized provenance of these services in the Nordic countries. Similarly to Google services, or xGUS.