**EGI-InSPIRE**

**Establishing the**

**Operational Tool**

**product teams**

**EU DELIVERABLE: MS702**

|  |  |
| --- | --- |
| Document identifier: |  |
| Date: |  |
| Activity: | **JRA1** |
| Lead Partner: | **INFN** |
| Document Status: | **Draft** |
| Dissemination Level: | **PUBLIC** |
| Document Link: | https://documents.egi.eu/document/xx |

|  |
| --- |
| AbstractThis document presents the operational tool product teams working for the WP7 (JRA1) activity. Details on their development and testing infrastructure is also be given after a brief introduction to the developed tool. The common release procedure and the relation with external bodies in order to get new requirements are described.  |

Copyright notice:

Copyright © Members of the EGI-InSPIRE Collaboration, 2010. See [www.egi.eu](file:///C%3A%5CUsers%5Ccesini%5CAppData%5CLocal%5CTemp%5Cwww.egi.eu) for details of the EGI-InSPIRE project and the collaboration.

EGI-InSPIRE (“European Grid Initiative: Integrated Sustainable Pan-European Infrastructure for Researchers in Europe”) is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 7th Framework Programme. EGI-InSPIRE began in May 2010 and will run for 4 years.

This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc/3.0/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA. The work must be attributed by attaching the following reference to the copied elements: “Copyright © Members of the EGI-InSPIRE Collaboration, 2010. See [www.egi.eu](file:///C%3A%5CUsers%5Ccesini%5CAppData%5CLocal%5CTemp%5Cwww.egi.eu) for details of the EGI-InSPIRE project and the collaboration”.

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

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

**Delivery Slip**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Name** | **Partner/Activity** | **Date** |
| **From** |  |  |  |
| **Reviewed by** | **Moderator:** **Reviewers:**  |  |  |
| **Approved by** | **AMB & PMB** |  |  |

**Document Log**

|  |  |  |  |
| --- | --- | --- | --- |
| **Issue** | **Date** | **Comment** | **Author/Partner** |
| 1.0 | 29/06/2010 | First draft |  |
| 2.0 |  |  |  |
| 3.0 |  |  |  |
| 4.0 |  |  |  |

**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 the 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 the ESFRI projects. 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 − structured international user communities − that are grouped into specific research domains. VRCs are formally represented within EGI at both a technical and strategic level.

**Table of contents**

1. Introduction 5

1.1. Purpose 5

1.2. Application area 5

1.3. References 5

1.4. Document amendment procedure 6

1.5. Terminology 6

2. JRA1 activity Overview 7

3. The product teams development infrastructure 7

3.1. Operation Portal 7

3.2. EGI Helpdesk 8

3.3. Grid Configuration Database 8

3.4. Accounting Repository 9

3.4.1. Accounting for different resource types 9

3.5. Accounting Portal 9

3.6. Service Availability Monitoring 9

3.7. MESSAGE BUS CONFIGURATION 10

3.8. Metrics portal 10

4. Interaction with external bodies 11

5. Release procedure and documentation 11

6. Conclusions 12

# Introduction

## Purpose

This document lists the tools that will be developed by the WP7 activity and describes the distributed development infrastructure that will be used. A description of the release procedure and the interaction with external bodies needed to get feedback and new requirements for the development will also be given.

## 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.

## References

**Table 1: Table of references**

|  |  |
| --- | --- |
| R 1 | http://www.symfony-project.org/ |
| R 2 | https://cic.gridops.org/index.php?section=roc&page=generaldoc |
| R 3 | http://maven.apache.org/ |
| R 4 | https://forge.in2p3.fr/projects/show/opsportaluser |
| R 5 | https://cvs.in2p3.fr/operations-portal/package/ |
| R 6 | http://www.remedy.com/ |
| R 7 | https://gus.fzk.de/pages/docu.php |
| R 8 | https://savannah.cern.ch/projects/esc |
| R 9 | https://gus.fzk.de/pages/owl.php |
| R 10 | http://goc.grid.sinica.edu.tw/gocwiki/GOCDB4\_Architecture |
| R 11 | http://goc.grid.sinica.edu.tw/gocwiki/GOCDB\_Technical\_Documentation |
| R 12 | https://savannah.cern.ch/projects/gocdb/ |
| R 13 | https://fedorahosted.org/koji/ |
| R 14 | http://goc.grid.sinica.edu.tw/gocwiki/ApelHome |
| R 15 | https://edms.cern.ch/document/726137 |
| R 16 | http://www3.egee.cesga.es/gridsite/accounting/CESGA/links/acct\_ibergrid06\_final14.pdf |
| R 17 | http://bestpractical.com/rt/ |
| R 18 | http://subversion.apache.org/ |
| R 19 | http://www.nagios.org/ |
| R 20 | https://edms.cern.ch/document/1060571 |
| R 21 | https://twiki.cern.ch/twiki/bin/view/EGEE/ExternalROCNagios |
| R 22 | http://www.atlassian.com/software/jira/ |
| R 23 | https://tomtools.cern.ch/jira/browse/SAM |
| R 24  | https://koji.afroditi.hellasgrid.gr/koji/ |
| R 25 | http://rpm.hellasgrid.gr/mash/ |
| R 26 | http://www.sysadmin.hep.ac.uk/rpms/egee-SA1/ |
| R 27 | http://www.sysadmin.hep.ac.uk/svn/grid-monitoring/ |
| R 28 | http://activemq.apache.org/ |
| R 29 | https://tomtools.cern.ch/jira/browse/MIG |
| R 30 | http://www3.egee.cesga.es/gridsite/accounting/CESGA/links/paper\_metrics\_iber2010.pdf |
| R 31 | https://rt.egi.eu/rt/index.html |

## 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>

## Terminology

A complete project glossary is provided in the EGI-InSPIRE glossary:

 <http://www.egi.eu/results/glossary/>.

# JRA1 activity Overview

The JRA1 activity is responsible for the development and continuous evolution of the operational tools used by the EGI production Grid. Future developments will follow the evolution of the infrastructure that is expected to encompass additional resource types, such as desktop grids, cloud and high capacity computing resources. The work started during the EGEE-III project to allow for a fully distributed or regionalized model of the operational tools should be completed during the early stages of the activity. The evolution of the accounting system towards an economic and business model is also an objective of the activity which is composed by five tasks:

* TJRA1.1 – Activity management: is the coordination of the development work, including the definition of the release roadmaps, the representation of the activity within the projects bodies and the reporting on the status of the activity. Runs for 4 years.
* TJRA1.2 – Maintenance and development of the deployed operational tools: is the underlying bug fixing and development work needed by the following operational tools:
	+ The Operation Portal
	+ The EGI Helpdesk
	+ The Grid Configuration Database (GOCDB)
	+ The Accounting Repository
	+ The Accounting Portal
	+ The Service Availability Monitoring
	+ The Metrics Portal

A general description of the tools will be given in section 3 where the development infrastructures for each tool will be presented. This task runs for 4 years.

* TJRA1.3 – National Deployment Models: is the development needed to support the deployment at a regional (or NGIs) level. This activity, started during the EGEE-III project, should be completed by the end of the first year of the project.
* TJRA1.4 – Accounting for different resource types: is the work needed to evolve the EGI accounting system in order to encompass the different resource types that will be included into the production infrastructure and to support an economic model needed for the self sustainability of the EGI Grid. This is a 3 years task starting from the second one.
* TJRA1.5 – Integrated Operation Portal: the development needed for a restructuring of the Operation Portal and its harmonization with other portal frameworks such as the GOCDB. This task spans over the first three years of the project.

The relationship between the JRA1 activity and other project bodies is described in section 4.

# The product teams development infrastructure

In this section the development infrastructure used by each product team within the activity is described. A brief introduction to the developed tools is also given.

## Operation Portal

The Operation Portal is an integration platform, allowing for strong interaction among existing tools with similar scope and filling up gaps wherever functionality was lacking. It also implements numerous work flows derived from procedures for several of its features out of requirements expressed by end users or administrators of Virtual Organizations (VO), Regional Operations Centres (ROC) or Resources Centres. The information on display is retrieved from several different distributed static and dynamic sources – databases, Grid Information System, Web Services, etc. – and gathered onto the portal. Criss-crossing this information enables the users to display high level views where static and dynamic data yield representative views of the EGI grid. This resulted in numerous tools that turned precious to sites like user tracking or monitoring failures alarm notification.
Complementary to this informative goal, the portal also fosters communication between different actors, through channels like EGEE broadcast, and Downtime Notification Mechanisms and putting in place procedures to address their interactions needs.
The release of the first production regionalized package is from June 8th 2010 and this release includes the dashboard. This package will be constantly improved during the coming year:
•   To add the other features integrated step by step into the Central Instance.
•   To ease the upgrade or the set-up of the package.

The development infrastructure is located at IN2P3, in Lyon and the development team is composed by three people (68 PMs over 4 years), two responsible for the Symfony [1] developments and one for the web service integration and the global architecture. Details on the architecture and implementation can be found on the portal at [2]. Maven [3] is used to build the web service and Symfony scripts for the remaining code. The internal bug/task tracking system is hosted on the IN2P3 Forge system [4]. The communication channels include the cic-information<at>in2p3.fr mailing list and a request form on the Portal. In case of bugs a ticket can be opened either on GGUS or directly on the internal bug tracking system. Two completely independent infrastructures are in place for development and testing, but can be synchronized via SVN. Once the development is done and validated the code is propagated to the test infrastructure where the new features are first checked by the developer in charge and after by one of the other developers. Distribution of rpms is done via SVN [5].

## EGI Helpdesk

The EGI Helpdesk (GGUS) is the primary way to request support about any production infrastructure related problem. Trouble tickets are created upon request and are used to track the problem until a solution is reached. The Helpdesk system is not foreseen to be deployed at the NGI level, but regionalized views are currently under test.

Developers and development infrastructure are located at KIT in Karlsruhe. Four developers (47 PMs over 4 years) compose the team, one is responsible for the frontend, one for the Remedy [6] integration, one for the failover and system administration, one for the regional view development.

Details on the architecture and implementation can be found on the main GGUS documentation page [7].

SVN is used for code management, feature requests and bug tracking is done through a savannah project hosted at CERN [8]. The GGUS mailing list (ggus-info<at>cern.ch) can be used to contact the developers – it will be moved soon to an EGI address.

The request tracker contains a “shopping list” that is an open forum where wishes’ and requirements are recorded and discussed by an advisory group. The decisions made by this group are then published on the Ongoing Worklist [9] which is the official plan where everybody can get information about what is going on GGUS.

New releases are timely announced via various communication channels and the new features are moved to the GGUS training system where are intensively tested by the GGUS developers and by the regional helpdesk developers if the interfaces to other helpdesk systems are involved by the new code. Automatic test scripts for basic functionalities are also available. Manual testing is always done after each release for key features, such as the “Alarm Ticket” chain.

## Grid Configuration Database

The Grid Configuration Database (GOCDB) contains general information about the sites participating to the production Grid: the installed m/w services, the administrators and security contacts, the scheduled and unscheduled downtimes are examples of the information stored into the GOCDB. These information are available through a series of web pages accessible via X509 certificates. The project will provide an instance of the GOCDB service that could be deployed nationally and information federated into a central instance.

Implementation details and architecture are available at [10] and [11]. The development and testing infrastructures are located at RAL-STFC, Didcot, UK. Two developers (for a total of 1 FTE) are responsible for the development of the tool. The bug/task tracking system is a savannah project hosted at CERN [12], KOJI [13] is used as build system. Communication channels to the team are: the savannah project, the GGUS support unit and the mailing list gocdb-admins<at>mailtalk.ac.uk.

Source code and RPMs are distributed via SVN.

## Accounting Repository

The EGI accounting repository stores information about the usage sites resources within the EGI production infrastructure. It receives records from accounting probes running in the sites and exports accounting information to consumers (i.e. the accounting portal for visualization).

Currently the accounting repository runs as a central instance only, but development are foreseen to have a distributable version that can be installed at a regional or NGI level.

The accounting repository is developed by the same team developing the GOCDB at STFC (54 PMs over 4 years for both components).

Details on the architecture and implementation are available at [14] and [15].

Communication channels are the GGUS helpdesk and the support mailing list APEL-SUPPORT<at>JISCMAIL.AC.UK.

### Accounting for different resource types

The work needed for task TJRA1.4 (accounting for different resource types, see section 2) will be spread over four partners: LUH (18 PMs), INFN (26 PMs), SFTC (27 PMs), CESGA (18PMs).

The development infrastructure and the tools that will be used to carry on the task are still not well defined since the task will start on the second year of the project, but of course the work will be done in strong connection with the already existing accounting development infrastructures inside the EGI (section 3.4) and EMI projects.

## Accounting Portal

The Accounting Portal is a graphical frontend for the production infrastructure accounting data. Accounting statistics are available for different kinds of users: i.e. end users, VO managers, site administrators. Implementation and architecture details can be found at [16].

The central accounting portal will be extended to support regional deployments of NGIs accounting systems and national deployment of the accounting repository. The Accounting Portal will be distributed to NGIs that wish to install it, but also new regional views on the central instance will be created for NGIs not running their own instance of the portal.

The developers and development infrastructure are located at CESGA.

The development group is composed by three developers that can be contacted via GGUS tickets or via the grid-admin<at>cerga.es mailing list. Internal CESGA RT [17] helpdesk is used as bug/task tracking system. Subversion [18] is used to manage the code.

Development version is available for preview and testing on a dedicated instance. The testing procedures are in general composed by two step:

* tests performed by the development group
* final review performed by the requestor of the new functionality

## Service Availability Monitoring

The Service Availability Monitor (SAM) is the system that will be used to monitor the resources within the production infrastructure. It includes the following components:

* The test framework: based on the NAGIOS system [19] set up and customized by the NAGIOS Configurator (NCG)
* DataBase components: The Aggregated Topology Provider (ATP), the Metric Description DataBase (MDDB) and the Metrics Result DataBase (MRDB)
* Message bus to publish the monitoring results
* A visualization tool: MyEGI

All the components are developed using the same infrastructure and process, but the development team is distributed across 2 partners: SRCE (15 PMs) is responsible for the NCG component, CERN is responsible for the DB components and MyEGI (18 PMs), the message bus configuration is done in conjunction with a team located at AUTH (see section 3.8).

The SAM system is a distributable package and its regionalization was almost completed during EGEEIII [20]. The validation process of the regional instances is still ongoing and progress is tracked on this wiki page [21].

Communication channels for the whole group include a GGUS support unit and the following mailing lists hosted at CERN: egee3-operations-automation-discuss<at>cern.ch, regional-nagios-admins<at>cern.ch.

JIRA [22] [23], hosted at CERN, is used as bug/task tracking system and KOJI [13] [24], which is provided by AUTH, as build system.

YUM repositories are created by the KOJI building system at AUTH [25] and synced off-site to Manchester [26] both sites can act as distribution and the code is accessible through an SVN repository [27].

Each release is bundled in a JIRA version which has a single task tracking the validation of that release through a dedicated testbed where testing is done to check that the bugs are fixed.

## MESSAGE BUS CONFIGURATION

Within the activity a group located at AUTH (12PMs) will work to ease the integration of the tools with the broker network that EGI is running. This broker network is based on ActiveMQ/FUSE [28] and even if its development or maintenance is not responsibility of the JRA1 activity, work will be carried on to provide configuration and tools that will maintain a message bus infrastructure that fit each application's requirements. Some tools, as SAM for example, are already integrated with the ActiveMQ broker network and their message bus configuration will be kept updated as the EGI production broker will evolve.

The AUTH team is using the same tools and providers/hosts for them as SAM:

* it has JIRA (at CERN) [22] [29] for bug/task tracking system and KOJI (at AUTH) [25] as the building system
* the YUM repositories at AUTH and Manchester [25][26]
* the sources for our developments are in Manchester's SVN repository [27]

Production message bus network maintenance and operation communication is handled at the following mailing list hosted at AUTH: prod-broker-network<at>grid.auth.gr. Development related communication are handled within JIRA.

## Metrics portal

The metrics portal automatically collects a set of metrics that can help to measure project performance and keep track of its evolution displaying all of them in a single place. The portal collects information from different sources like GOCDB, GGUS, GridView, etc. using various connectors and stores it into a local database. Information are then available through web pages that can be exploited to create statistical reports in few seconds. No regional model is foreseen, the portal will run on a central installation only.

Architecture and implementation details can be found at [30].

The development group and the testbeds are located at CESGA (33 PMs for both Accounting Portal and Metrics Portal). The communication channel for the group is the mailing list grid-admin<at>cesga.es.

Internal CESGA RT is used as bug/task tracking system, subversion for managing the code.

Testing procedures include internal review done by the developers of the new functionalities and a preview instance made available to the functionality requestors.

# Interaction with external bodies

Interaction with external bodies is fundamental in order to get input, feedback and new requirements for the developed tools. Two advisory groups are foreseen by the project description of work: the Operational Tools Advisory Group (OTAG) and the User Services Advisory Group (USAG):

* Being composed by representatives from the NGIs, from the m/w developers and from the JRA1 activity, the OTAG will be the main supervisory group for the development progress and the place where technical discussion about the evolution of the tools will take place
* The USAG has representatives from the EGI.eu user communities and will focus on the requirements for the complete set of services run by the project but could also impact on the operational tools - in example end user requirements could be addressed to the GGUS Helpdesk

In order to create the proper schedule for the development it will be important that the outcome of the advisory groups work will be a single prioritized table. The prioritization is an important step and possible conflicts should be resolved when multiple requirements from different groups will impact on the same tool. This is not expected to happen frequently and will be analyzed case by case by the management of the activity together with the advisory groups.

The development progress of new features, requested and approved by the advisory groups, will be tracked using the project tracking system [31].

# Release procedure and documentation

Each tool will be released as a standalone package. The release procedure, presented in Figure 1, will apply to each tool.

Figure - Operational Tools Release Procedure

When development of a new release is completed (T0) a first announce is done to the EGI community using the broadcast tool of the Operation Portal. The announce has to contain the release notes, the documentation links, a detailed test plan, an indication of the expected release date and all the information needed by the conformance criteria set by the SA2 activity for the software providers of the project. A JRA1 internal phase of testing will take place until all the tests described into the test plan are passed – tests will include a documentation review. At this time (T1) a second release announce will be broadcasted containing the actual release date (TR), the release notes, the documentation links and a document describing the testing phase results. The release can be an immediate installation on the production instance for the centralized tools or will follow a deployment process, with a possible initial testing phase on a selected number of instance (StagedRollout), according to SA1 needs.

Each tool will maintain its own documentation on the web, links to these web pages will be maintained and reviewed before each release on the project wiki in order to have a single access point to all the tools documentation.

# Conclusions

In this milestone the development and testing infrastructure for the operational tools under the responsibility of the JRA1 activity was described. The infrastructure is distributed across the partners participating to the activity and is composed by different (in terms of hardware resources, bug/task tracking systems, communication channels and deployment facilities) sub-infrastructures, one for each tool. However the requirements workflow and the release process that will be used during the whole project lifetime will be common to all the tools. The development progress will be tracked using a project level facility for all the tools and a single access point for the tools documentation will be created.