



EGI-InSPIRE

Operational Tools Regionalization Work Plan

EU MILESTONE: MS703

Document identifier:	EGI-MS703-107-V7
Date:	06/10/2010
Activity:	JRA1
Lead Partner:	CERN
Document Status:	DRAFT
Dissemination Level:	PUBLIC
Document Link:	https://documents.egi.eu/document/107

Comment [J1]: Obtain from the EGI document server: documents.egi.eu

Abstract

This document assesses the work plan on the regionalization of the operational tools, identifying the upcoming releases and associated functionality.



Copyright notice:

Copyright © Members of the EGI-InSPIRE Collaboration, 2010. See www.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 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 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	David Collados	CERN/SA1	19/07/2010
Reviewed by	Moderator: Reviewers:		
Approved by	AMB & PMB		

Comment [J2]: To be completed by the Project Office on submission to the AMB and PMB.

Comment [J3]: To be completed by the Project Office on submission to the EC or when published on the website.

Document Log

Issue	Date	Comment	Author/Partner
1.0	19/07/2010	First draft	D. Collados/CERN
2.0	16/08/2010	Added comments for GOCDB, APEL and Operations Portal. Improved chapter 4.	D. Horat/CERN
3.0	02/09/2010	Added GGUS block	D. Collados/CERN
7.0	06/09/2010	Added executive summary, conclusions and minor changes	D. Horat/CERN
8.0	06/10/2010	Considered reviews on EGI document 165	D. Collados/CERN
9.0	07/10/2010	Added accounting repo regionalization plan	D.Collados/CERN

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.....	ERROR! BOOKMARK NOT DEFINED.
1.5. TERMINOLOGY.....	6
2. EXECUTIVE SUMMARY	7
3. OPERATIONAL TOOLS REGIONALIZATION STATUS IN EGI-INSPIRE	8
4. ROADMAPS IN THE REGIONALIZATION PROCESS	9
4.1. ACCOUNTING PORTAL.....	15
4.2. METRICS PORTAL	16
4.3. MONITORING INFRASTRUCTURE	15
5. CONCLUSIONS	17

1. INTRODUCTION

1.1. PURPOSE

This document is a report on the work plan of the regionalization process of the operational tools that needs to be addressed by the WP7. The upcoming functionality is presented here.

1.2. 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.

1.3. REFERENCES

Table 1: Table of references

R 1	Milestone MS401: Operational Tools Regionalization Status (https://documents.egi.eu/document/48)
R 2	The CIC Operations Portal (https://cic.gridops.org)
R 3	<i>Global Grid User Support (GGUS) Plan, EGEE-III Deliverable DSA1.1, August 2008</i> (https://edms.cern.ch/document/931935)
R 4	Mathieu, G.; Casson, J.; <i>GOCDDB4, a New Architecture for the European Grid Infrastructure</i> ; to appear in the Proc. of the Int. Symposium on Grid Computing 2010, Mar 2010, Academia Sinica, Taipei, Taiwan
R 5	Jiang, M.; Del Cano Novales, C.; Mathieu, G. et al; <i>An APEL Tool Based CPU Usage Accounting Infrastructure for Large Scale Computing Grids</i> , in the Proc. of ISGC 2010, Mar 2010, Academia Sinica, Taipei, Taiwan
R 6	The Accounting Portal (http://www3.egee.cesga.es)
R 7	The Service Availability Monitoring Framework (http://www.youtube.com/watch?v=PADq2x8q0kw)
R 8	The Metrics Portal (http://www3.egee.cesga.es/gridsite/accounting/CESGA/links/EGEE-III-SA1-Metrics_Portal_Roadmap_and_Requirements-v3.1.xls)
R 9	The Regional Dashboard (https://operations-portal.in2p3.fr/index.php)
R 10	The Messaging Infrastructure (https://twiki.cern.ch/twiki/bin/view/LCG/MessagingSystemforGrid)
R 11	The Apache ActiveMQ messaging system (http://activemq.apache.org)
R 12	Nagios (http://www.nagios.org)
R 13	EGI-InSPIRE Document Management Procedure (https://wiki.egi.eu/wiki/Procedures)
R 14	EGI-InSPIRE Glossary (http://www.egi.eu/results/glossary/)
R 15	SVN repository for Operations Portal Regional Package (https://cvs.in2p3.fr/operations-portal/package/)
R 16	Validation of Operations Regional Package

	https://cvs.in2p3.fr/operations-portal/package/validation.pdf?revision=HEAD
R 17	EGI InSPIRE Milestone MS701: Operations Portal Work plan https://documents.egi.eu/document/39
R 18	GGUS Portal Interface Standards https://gus.fzk.de/pages/ggus-docs/interfaces/docu_ggus_interfaces.php
R 19	GOCDB binaries repository http://www.sysadmin.hep.ac.uk/rpms/egee-SA1/centos5-devel/x86_64/repoview/gocdb.html
R 20	GOCDB SVN repository (http://www.sysadmin.hep.ac.uk/svn/grid-monitoring/trunk/gocdb/)
R 21	GOCDB wiki page http://goc.grid.sinica.edu.tw/gocwiki/GOCDB_Regional_Module_Technical_Documentation
R 22	GOCDB regionalization status http://goc.grid.sinica.edu.tw/gocwiki/GOCDB4_Regionalisation
R 23	GOCDB development plans (http://goc.grid.sinica.edu.tw/gocwiki/GOCDB4_development)
R 24	APEL wiki page (http://goc.grid.sinica.edu.tw/gocwiki/ApelHome)

1.4. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors. The procedures documented in the EGI-InSPIRE “Document Management Procedure” [R13] will be followed.

1.5. TERMINOLOGY

A complete project glossary is provided in the EGI-InSPIRE glossary [R14]



2. EXECUTIVE SUMMARY

The EGI-InSPIRE project continues the transition to a sustainable pan-European e-Infrastructure started in EGEE-III, integrating and interoperating individual national grid infrastructures and technologies throughout Europe. Among other objectives, the operational tools will be extended by the project to support a national operational deployment model, and to integrate new Distributed Computing Infrastructures and their associated accounting information. For this, EGI attempts to have a strong regionalization model in the area of operational tools. The large effort devoted to this activity during the second year of the EGEE-III project, providing a regionalized instance of each of the tools or a regionalized view on the central instances is a very good starting point.

This document describes the regionalization activity of the different operational tools, focusing on the status and roadmap of the respective efforts.



Error! No text of specified style in document.

Document Identifier

Date: 06/10/2010

3. OPERATIONAL TOOLS REGIONALIZATION STATUS IN EGI-INSPIRE

The status of deployment of regionalized tools is fully described in the ‘Operational Tools Regionalization Status’ EGI’s milestone [R1], covering the following tools: the *Operations Portal [R2]*, *GGUS [R3]*, *GOCDDB [R4]*, the *Accounting Repository [R5]*, the *Accounting Portal [R6]*, the *Service Availability Monitoring Framework [R7]*, the *Metrics Portal [R8]* and the *Regional Dashboard [R9]*.

4. ROADMAPS IN THE REGIONALIZATION PROCESS

4.1. OPERATIONS PORTAL AND REGIONAL DASHBOARD

The Operational Portal and Regional Dashboard is a set of tools and interfaces dedicated to the 1st Line Support, Regional or Central Operators on Duty daily work.

STATUS OF REGIONALIZATION

The development of a portal regional package started in EGEE-III: the goal of this work is to distribute the different features and modules present into the Central Operations Portal into a regional package for NGIs to cope with NGI needs.

The first release of the first production package was on June 8th 2010 and it included the first release of the dashboard module. The different parts of the package and the documentation are distributed via a SVN repository [R15].

The regional package has been deployed in production in 2 NGI:

- NGI_CZ
- NGI_GRNET

2 NGIs are currently deploying it:

- NGI_IBERGRID – the validation process is completed but they want to extend the test period
- NGI_BY

ARCHITECTURE AND SYNCHRONIZATION

The application is composed of:

- a web service named Lavoisier configured to handle Nagios notifications, store and provide data cache from GOCDB and GGUS and to generate metrics reports about the use of the dashboard (alarms raised, tickets handled ...);
- a PHP web application to provide a user interface based on Symfony framework;
- a database generated automatically with the web application during the installation.

The regional instance is linked with the central instance of Lavoisier; and creation, update, delete of records are synchronized so as not to disrupt global oversight operations. Synchronisation is achieved through REST and SOAP and records are synchronized every 10 minutes by using php scripts. Any problem detected during the synchronisation is reported in a mail sent to the webmasters.

The central and the regional instances (Instance 1 and 2 on the schema below) have been built on the same model in order to behave in the same way and to be easily interoperable. A validation procedure [R16] has been established to validate the regional instances when a given region expresses this need. All regions or NGIs are able to opt for either the central or the regional instance. We will continue to offer as now regional views for the future features – when it makes sense - to the NGIs.

As you can see on the schema, the architecture is exactly the same on a NGI and central instance.

The distributed components are the same, just the configuration changes:

- with PHP code in the application that permits to distinguish a regional instance from a central one
- on the Lavoisier side to filter information and to keep only data relevant to the NGI.

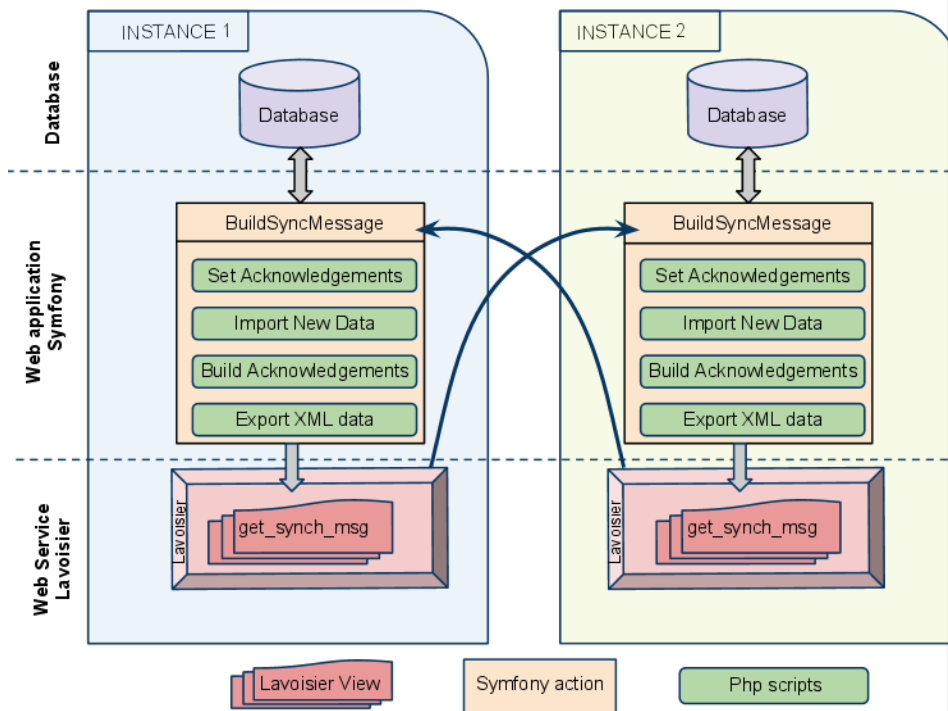


Figure 1 - Overview of the architecture and synchronization

On the way a regional instance interacts to a helpdesk, we foresee two scenarios. This is represented in Figure 2, below. The current and standard model corresponds to scenario (A), where a regional portal interfaces directly a GGUS helpdesk. An evolution of this model is represented in scenario (B), where tickets are not created and updated in GGUS but directly in a regional helpdesk.

To put in place a system based on model (B), some technical points must be solved first:

The first one is the work-flow behind the ticket. If we are directly connected to the regional helpdesk, it is not clear how we can transfer the problem to Central Operator in the dashboard if the ticket is not registered in GGUS (when a ticket expired it should be automatically visible to the Central Operator, so it should be transferred automatically to GGUS).

On the other hand, access to the regional helpdesk must be standard, so they can be interfaced in a common way.

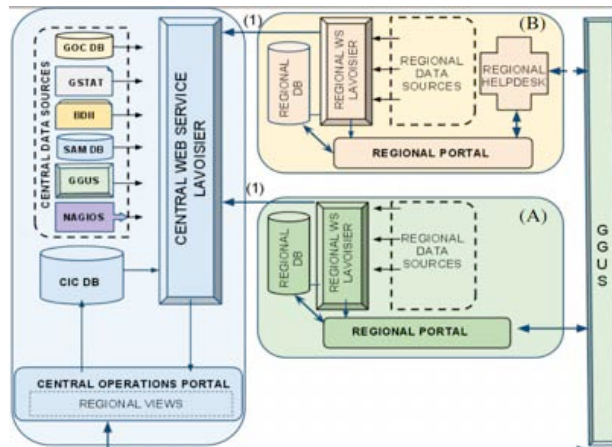


Figure 2. Interaction to helpdesks

REGIONALIZATION ROADMAP

This regional package is currently using the same central data sources as the central instance but this package is designed to integrate the regional version of GOCDB and regional helpdesks if they are deployed. This regional package will be constantly improved during the coming year (2010/2011) to add some of the other features that will be integrated step by step into the Central Instance. These features will be added one by one into the central instance, and if relevant to a regional instance we will provide an update to the package.

For each central feature released, we will provide it in the regional package in the following months, if applicable.

For the central instance the different timelines are summarized in the table below.

For each new feature a study will be carried out to understand if the deployment of such features at a regional level is relevant.

Features	Release Date
Migration to Symphony of the VO ID Card	August 2010
Migration to Symphony of the Broadcast Tool	September 2010
Migration to Symphony of the Downtime notifications system.	December 2010
Enhancements of Lavoisier Programmatic Interface	December 2010
Integration of VO specific tests within the Dashboard	February 2011
Release of Lavoisier 2.0	March 2011

For more details about the complete work plan, please read MS701, Operations Portal Work plan [R17].

4.2. GGUS

The GGUS portal is the central helpdesk of the EGI infrastructure. It acts as an integration platform between other helpdesk systems (regional and topical) that are in production in EGI. The set-up with a central tool and various different satellite systems has been in place since the beginning of EGEE. All project-wide support units are present in the GGUS system, regional or specialised topical support units can be reached via GGUS through the interfaces their respective ticket systems have with GGUS. This enables to track all trouble tickets and service requests of general interest in the GGUS system, independent of the tool in which the ticket is worked on. Since infrastructure is already regionalized GGUS was not part of the regionalization effort that started in EGEE for the operational tools which was steered by the OAT.

One of the major changes to the user support infrastructure caused by the transition from EGEE to EGI is the move from ROCs to NGIs. This means that instead of 12 regional ticket systems (one in each ROC) now every NGI has to set up the respective user support infrastructure, and interface it to the central GGUS system. This of course increases the number of interfaces that have to be maintained in GGUS. In order to cope with this it is necessary to strictly standardise the interfaces to new NGI helpdesk systems. Only web services and messaging interfaces are being accepted now. The interface standards for these two methods are described in the documentation collection of the GGUS portal [R 18].

In EGI we have identified three options how NGIs can connect to the central user support infrastructure:

- NGIs can set up a regional ticket system and interface GGUS via web services or messaging. This option is currently used by five NGIs.
- NGIs can use the GGUS system directly. In this case there is only one support unit in GGUS for the whole NGI. This option is currently used by 11 NGIs.
- NGIs can use a customisable version of GGUS called xGUS. This was developed to give NGIs a simple way to set up a ticket system without having to host it themselves.

STATUS OF xGUS

The design phase of a customisable version of GGUS started in 2009 during EGEE-III, resulting in resources being assigned to developing such a system in the proposal for EGI-InSPIRE under JRA1.

The xGUS system has the same basic functionality of GGUS and is based on the same workflow engine. xGUS consists of a web front end that can be customised with a NGI specific banner and a backend that includes all basic support portal functionalities such as a ticket database, an email engine, user administration, a news module and basic portal administration. This means that fields like 'Type Of Problem', 'Affected Site', 'Affected VO', 'Responsible Unit', and link lists can be administered directly via the portal.

The template comes with synchronization to GGUS, so tickets coming from or going to GGUS are duplicated and synchronized automatically.

xGUS is hosted central by the GGUS team, relieving the NGIs from having to maintain a ticket system locally, so it is not meant to be deployed locally. Currently there is a first instance of xGUS in place for NGI-DE, the German NGI. This tool is now used in production as the regional ticket system for NGI-DE. A second instance for NGI-AEGIS, the Serbian NGI, is being finalised at the moment. Discussions with other NGIs on the use of xGUS have started.

The xGUS system will undergo the same release procedures as the GGUS system, as the two systems are closely related.

4.3. GOCDB

GOCDB stands for **Grid Operations Centre DataBase**, and stores and gives access to information about regions and countries, resources, users, roles and contacts. It also links this information together logically.

GOCDB STATUS

As of the end of July 2010 GOCDB is in a transition phase between GOCDB3 and GOCDB4.

- GOCDB4 central portal is online since November 2009. Data appearing there are synchronised from GOCDB3 data
- GOCDB3 Programmatic Interface has been decommissioned on the 15th of July 2010. As a result, all client tools connecting to GOCDB are all contacting GOCDB4 and there is no more dependency on GOCDB3
- the GOCDB4 input system which is due to replace GOCDB3 web portal is under finalisation and testing.

STATUS OF REGIONALIZATION

The last released version of GOCDB regional module dates from March 2010 (gocdb4.0d3). It is not a fully functional tool but all the basic components are included so that initial deployment can be tested using that version. It is available from the GOCDB repository [R 19].

The latest development version of the module has not been packaged yet, but the code is available from the SVN repository [R 20]

This development version is a fully functional system, although there are minor developments and polishing issues before this can be labelled "production". The estimated timeline for this is September 2010.

The documentation is available on the GOCD wiki page [R 21]

GOCDB regionalization status is available online at [R 22]

Important note: GOCDB regional modules can't be operated in production until GOCDB3 has been fully decommissioned.

REGIONALIZATION ROADMAP

Main GOCDB development items linked to the tool regionalization are as follows, with the expected release date:

Features	Release Date
Provide a production quality package for GOCDB regional module	September 2010
Optimize data access for the regional module to improve overall performances	October 2010
Finalise the publication interface from regional to central GOCDB	October 2010
Consolidate and polish the web admin interface for GOCDB regional module	October 2010



Propose a basic definition fo a publication interface from regional non-GOCDB system to the central GOCDB	November 2010
Provide a MySQL version of the GOCDB regional module	April 2011

GOCDB development plans are published on the development wiki page [R 23]

4.4. ACCOUNTING REPOSITORY

The Accounting Service, comprised of data collection and reporting services, is a large centralized database which collects and aggregates CPU usage information from sites across the Grid through different types of sensors.

APEL parses batch system and Grid gatekeeper logs at each Grid site to generate CPU usage records and publishes them into a centralized repository at a GOC (Grid Operations Centre) using a transport mechanism. The records are processed and summarized and can then be accessed through the Accounting Portal [R 6]

APEL SERVER STATUS

As of the end of July 2010 the APEL server is in a transition phase between the R-GMA based system and its ActiveMQ based replacement. The ActiveMQ based APEL client has been released to production within gLite3.2 in early June 2010. About 12 sites are using this version at the time of writing. The server accepts publication from both systems in parallel. The respective Usage Records are then processed and archived in exactly the same way.

STATUS ON REGIONALIZATION

The APEL regionalized architecture was presented as a poster in EGEE09 conference and is available from the APEL home wiki page [R 24]. This architecture is still being followed. Of the three types of region identified (an NGI or group of NGIs is a region), region A (direct publishing) is now in place. Work has started on the other two types with the development of a summary publishing client (required for both type B (distributed APEL) and C(own accounting solution)) and summary ingestion in the server and a redesign of the repository to make it more suitable for distribution.

REGIONALIZATION ROADMAP

The milestones in the roadmap are described in this table:

Features	Release Date
Delivery of new glite 3.2 node type glite-APEL for publishing cpu accounting.	June 2010
Delivery of example summary client for integration by Region Cs, eg OSG, DGAS, SGAS	November 2010
Roll-out of new glite-APEL in production (started June 2010)	December 2010
Retire R-GMA production service	December 2010
Ingestion of summary records by central repository. This will allow existing	December 2010

RegionCs to publish by ActiveMQ and stop direct database insertion. Also any new Region Cs to start publishing.	
Delivery of prototype regionalised repository for testing by a few Region Bs.	February 2011

4.5. ACCOUNTING PORTAL

The Accounting Portal [R 6] is the interface to the accounting repository.

The prototype version of the regional accounting portal is already available. Currently several NGIs have expressed their interest in deploying a regional instance of the regional portal.

During the following months the installation scripts will be improved making easier to deploy the portal. Additional improvements will be incorporated based on the feedback from the candidate NGIs. New functionalities incorporated into the central accounting portal will be ported to the regional accounting portal.

In the future the regional accounting portal will use the new accounting infrastructure based on Messaging [R10]. This new accounting infrastructure allows greater flexibility to configure regional accounting repositories. Several modifications are foreseen in the regional accounting portal to support ActiveMQ [R11].

Besides the regional accounting portal a country view is offered in the central accounting portal that allows NGIs to review the usage in their region. This view will be continuously improved based on NGI feedback.

The roadmap for the accounting portal is the following.

Features	Release date
Improved installation support. GOCDBPI-V4 support	December 2010
New improvements based on NGI feedback	March 2011
ActiveMQ support	June 2011

4.6. MONITORING INFRASTRUCTURE

The monitoring infrastructure supervises the production and pre-production grid sites. It provides sets of probes which are submitted at regular intervals, databases to store test results and a messaging infrastructure to distribute the output of the tests. In effect, it provides monitoring of grid services from a user perspective.

The monitoring infrastructure based on Nagios [R12] and Messaging is already completely regionalized, as described in the EGI-InSPIRE milestone MS401 [R1]



Error! No text of specified style in document.

Document Identifier

Date: 06/10/2010

4.7. METRICS PORTAL

The Metrics Portal [R 8] is a tool to obtain statistics about grid operations. There are currently no regionalization plans for the Metrics Portal.



5. CONCLUSIONS

JRA1 has the primary goal of coordinating the operational tools development carried out by various groups. The guidelines are being defined with the fundamental support of the OTAG advisory group.

The operations portal and regional dashboard team have already released a first version of the regional dashboard and have provided a schedule of future releases.

GGUS offers with xGUS a simplified regional helpdesk instance for NGIs. These instances are operated centrally but can be customized by the regions. Also NGIs can run their own systems connected to GGUS via standardized interfaces.

GOCDB is currently in a transition phase between GOCDB v3 and GOCDB v4.

APEL repository regionalization roadmap has not been addressed yet.

The initial version of the regional accounting portal is already available. Currently several NGIs have expressed their interest in deploying a regional instance of the regional portal.

The monitoring infrastructure based on Nagios [R12] and Messaging is already completely regionalized, as described in the EGI-InSPIRE milestone MS401 [R1]

There are currently no regionalization plans for the Metrics Portal.