EGI-InSPIRE

OPERATIONAL LEVEL AGREEMENTS WITHIN THE EGI PRODUCTION INFRASTRUCTURE

EU MILESTONE: MS411

Document identifier:	EGI-MS411-v1.0
Date:	**20/07/2011**
Activity:	**SA1**
Lead Partner:	**AUTH/GRNET**
Document Status:	**DRAFT**
Dissemination Level:	**PUBLIC**
Document Link:	https://documents.egi.eu/document/524

Abstract

This document presents the status of the EGI Operational Level Agreements framework. It illustrates the changes introduced to the Resource Centre Operational Level Agreement during the course of the first year of the project and defines plans for future extensions that are needed to support the EGI service level management activities.

I. DELIVERY SLIP

	Name	Partner/Activity	Date
From	Dimitris Zilaskos	AUTH-GRNET/TSA1.8	28/06/2011
Reviewed by	**Moderator:** **Reviewers:** <<To be completed by project office on submission to AMB/PMB>>		
Approved by	**AMB & PMB** <<To be completed by project office on submission to EC>>		

II. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
1	1 June 2011	First draft	Dimitris ZIlaskos/AUTH-GRNET
2	28 June 2011	Final draft	Dimitris ZIlaskos/AUTH-GRNET
3	07 July 2011	First internal revision	T. Ferrari, P. Solagna

III. 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.

IV. 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

V. TERMINOLOGY

A complete project glossary is provided at the following page: http://www.egi.eu/about/glossary/. Additional Operations specific terms are provided at the Operations glossary page: https://wiki.egi.eu/wiki/Glossary

VI. 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.

VII. EXECUTIVE SUMMARY

Operational Level Agreements (OLAs) are defined in the Information Technology Infrastructure Library [ITIL] and describe how IT groups work together to meet IT service level requirements

The purpose of the OLA is to optimize the delivery of IT services to customers and users. It is an internal agreement that defines how two different organizations will work together to support the delivery of a set of IT services to customers and users.

Within EGI, the OLAs serve various purposes:

- to ensure mutual understanding of the principles of cooperation between parties of the operations community;

- to define the responsibilities of each party;

- to define the services to be delivered and the corresponding level parameters according to the needs of the users and/or the operators of the infrastructure as applicable;

- to support the EGI service level management procedures needed to monitor the delivered service quality.

During the project year 1 (PY1) of EGI-InSPIRE a task force was constituted to drive the evolution from the former EGEE Service Level Description [SLD] to Site-NGI OLA and eventually to the EGI Resource Centre OLA [RCO], which was finally approved in May 2011. The Resource Centre OLA defines the agreement between a Resource Centre (RC) and the respective Resource infrastructure Provider (RP) on the commitments needed to ensure an available and reliable grid infrastructure.

The new OLA was adapted to define the change management process, to adopt the new terminology defined in the Operations Architecture [ARCH] to define the conditions to be met to be an EGI certified RC, and to define the suspension policy that is applicable to underperforming RCs.

As the OLA is closely coupled to the operational tools – in particular the Availability Calculation Engine and the Operations Portal – that are needed to measure the Availability and Reliability performance indicators, effort was also spent to define the future tool extensions and their implications on the tool development roadmaps.

The RC OLA was consolidated during PY1. During PY2 new OLAs will be defined to address Global Services offered by EGI.eu and the respective partners (EGI.eu OLA) and the Local Services offered by the RPs (RP OLA). In addition, the procedure for Service Level Management will be further automated via the enhancement of the Service Availability Monitoring (SAM) framework.

TABLE OF CONTENTS

1   INTRODUCTION

This Milestone presents the status of the EGI Operational Level Agreement (OLA)[1] framework, the related Service Level Management procedure and suspension policy, and elaborates on the future extensions to the current framework that are needed to complete the description of the operational services, the respective Service Level Targets [ITIL] and the duties of the service providers.

The OLA is the mechanism adopted in EGI to integrate resource providers into the pan-European production infrastructure while ensuring interoperation of operational services, Quality of Service, and to enforce a common set of policies and procedures.

EGI operations are based on a distributed service-oriented model comprehending Global and Local Services [ARCH], where different stakeholders play the role of service provider.

Crucial to EGI is the maximization of the Quality of Service experienced by the end-user, which is affected by the overall quality offered by the providers of operational services. It is therefore important that the providers commit to a minimum set of requirements to jointly offer a predictable, secure and highly available service infrastructure.

The first OLA prototype was consolidated in the framework of the EGEE project series, to define the minimum set of services and the corresponding Service Level Targets to be exchanged between Resource Centres and ROCs.

EGI will extend this prototype framework to comprehend various types of OLAs (see Figure 1).
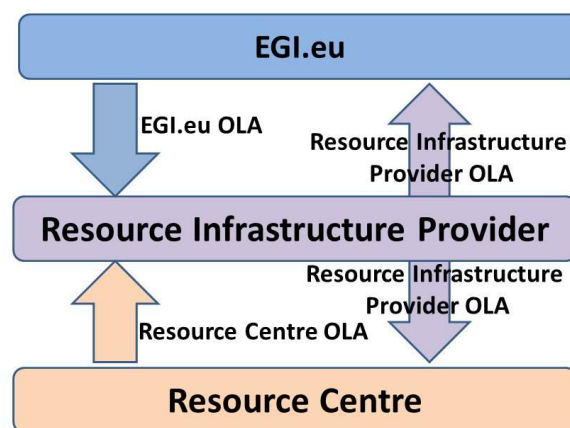
Figure 1. The EGI OLA framework and the relevant service providers. The arrows in the diagram indicate the partners involved in the respective OLAs.

The **Resource Centre OLA** (RC OLA) defines the minimum set of required *operational procedures and policies* to be endorsed and enforced locally, and the minimum set of *Local Services* to be offered by the RC, and the corresponding guaranteed service levels (the EGI Resource Centre OLA v 1.0 was approved in May 2011) [RCO].

The **Resource Infrastructure Provider OLA** (RP OLA) defines the minimum set of necessary *operational procedures and policies* to be endorsed and enforced locally by Resource infrastructure

---

[1] https://wiki.egi.eu/wiki/Glossary#Operational_Level_Agreement

Providers (RPs) and the corresponding Operations Centre [GLO], and the minimum set of necessary *Local Services* provided by the Operations Centre and the respective guaranteed service levels. The Resource infrastructure Provider OLA is currently under development.

The **EGI.eu OLA** defines the set of EGI *Global Services* that EGI.eu offers in collaboration with the EGI partners to the Resource infrastructure Providers, and the respective minimum guaranteed service levels.

The RC OLA, the RP OLA and EGI.eu OLA all together constitute the EGI OLA framework (Figure 1).

EGI.eu, certified Resource Centres and the accredited RPs will be requested to adhere to the relevant parts of such framework.

2 RESOURCE CENTRE OLA

The enhancement of the framework for the monitoring of the service levels delivered, and the progressive improvement of the quality of the infrastructure, are part of the EGI-InSPIRE project objective 1: "the continued operation and expansion of today's production infrastructure" [DOW].

The Resource Centre OLA (RC OLA) defines the minimum set of operational services to be delivered by a certified Resource Centre, the obligations, and a set of metrics with the corresponding minimum service level targets required. The RP can extend the EGI OLA framework according to the local needs by ensuring that the additional requirements do not introduce any conflict with what stated in the EGI RC OLA.

The acceptance of the RC OLA is now a requirement for being part of the certified infrastructure. New RCs are requested to endorse it as defined by the Resource Centre Registration and Certification procedure [PROC09]. The Resource Centre Operations Manager is the new role introduced to define the person responsible of accepting the RC OLA and of ensuring its enforcement by the RC.

While in EGEE times signatures were requested and collected during the agreement negotiation, in EGI it is the responsibility of the RP to record acceptance according to the local needs.

The enhancements introduced concern various technical areas.

- **Terminology**: it was adapted and harmonized. Sites are now referred as *Resource Centres*; the EGEE Regional Operations Centre (ROC) entity is now more generically defined as *Operations Centre* [GLO]. The infrastructure operated by the Operations Centre is defined *Resource Infrastructure*, and the legal entity responsible for it is the *Resource infrastructure Provider*. *Service Level Availability* is replaced by *Operational Level Agreement* in compliance with ITIL.
- **Responsibilities**: the *Resource Centre Operations Manager* role was introduced to define the Resource Centre contact responsible of accepting the OLA and of making sure it is endorsed by the Resource Centre. The OLA clarifies that the accuracy of GOCDB information is a joint responsibility of the Resource Centre Operations Manager and of the Resource Provider Operations Manager.
- **Deployed middleware**: any supported grid technology that complies with the UMD specification [UMD] can be deployed. The minimum set of capabilities to be provided by a Resource Centre was relaxed in compliance with the current availability computation algorithm. This change was necessary as more flexibility in the type of services deployed by RCs needs to be taken into account, in particular to foster the integration of novel resources (data repositories, desktop Grids, virtualized resources, etc.).
- **Deployed resources**: all constraints concerning the minimum amount of resource capacity to be provided were removed.
- **Service Level Targets**: the maximum response time to GGUS tickets was raised from 4 to 8 operating hours as this is more adequate for all types of RCs. VOs can optionally deploy other types of tickets (such as ALARM tickets already in use by the WLCG community) in case of severe issues that require immediate action.
- **Suspension policy**: a RC that does not provide the minimum requested availability for three consecutive months is eligible for suspension. The minimum availability to be provided was raised from 50% to 70%.

- *Supported VOs*: DTEAM and OPS need to be mandatorily supported by the Resource Centre for troubleshooting and monitoring purposes to reflect the decision of the OMB.
- *Change management*: the OLA amendment procedure is now defined.

The evolution of the RC OLA following the directions above has been driven by the works of a dedicated task force[2] during PQ2 and PQ3. The task force has been in charge of the definition of a roadmap for the evolution of the RC OLA and of the overall framework. The proposed roadmap was developed taking into account the input collected from NGIs through a questionnaire [QST].

The results of the questionnaire were presented at the EGI Technical Forum that took place in Amsterdam on September 2010 [ETF]**.** There the decision was taken that an OLA Task Force should be formed to lead the evolution of the OLAs.

---

[2] The NGIs that contributed are: France, Greece (coordination), Italy, NDGF, Poland, Serbia and Sweden.

3 TOOL REQUIREMENTS

Tools development efforts affect the implementation of the OLA roadmap, as the OLA Service Level Targets must match what the tools are able to provide for quality assessment and reporting.

Participants of the TSA1.8 and JRA1 tasks, EGI operations, and the ACE team at CERN worked together to define a set of OLA use cases requiring tool enhancements. These are documented at [USE].

Following more brainstorming and refining of the requirements, a final version of a proposal to JRA1 regarding the developments in Operation tools was presented and approved in the December 2010 OMB.

They key requirements have been about getting a flexible grouping of services and the possibility to add new custom profiles for the generation of availability reports. The purpose of these is:

- to define custom availability calculations based on arbitrary combinations of RC services as deemed necessary;
- to group operational tools and other RP-level services in order to produce RP availability reports;
- to group together service instances that are physically hosted by different RCs and/or Resource Infrastructures, and produce the corresponding availability reports. This grouping is useful for services that can be deployed in cluster mode and geographically distributed.

The implementation of these requirements requires extensions to both GOCDB and to POEM - the SAM component in charge of the management of monitoring profiles [POEM].

4 SERVICE LEVEL MANAGEMENT

EGI Service Level Management is the process responsible for ensuring that EGI *Operational Level Agreements and Underpinning Contracts are appropriate for the agreed Service Level Targets*, in conformance with ITIL.

The EGI process relies on the distribution of monthly Availability and Reliability reports that provide information about the performance of the individual RCs, as well as of the EGI Resource Infrastructures. Both EGI participants and the integrated infrastructures are concerned by this process as all EGI certified RCs are bound to the same minimum set of Service Level Targets and to the acceptance of the RC OLA.

4.1 Targets

Two service targets are constantly monitored: Availability and Reliability.

Availability is defined to be the percentage of time that the service/RC was up and running[3]; the minimum monthly value must equal or exceed 70%. Reliability is also the percentage of up time of a service/RC but it is computed excluding periods of scheduled intervention. While Availability measures the level of functionality delivered by a set of capabilities, Reliability estimates the quality of problem/incident management of a RC.

During PY1 the algorithm for the computation of Availability and Reliability did not change. However, some enhancements were introduced.

- The number of physical CPUs used to estimate the relative weight of a RC Availability/Reliability when computing the overall infrastructure target was replaced by the installed capacity estimated in HEP-SPEC 06 (10 HEP-SPEC 06 = 4 KSI2000) [HS06]. The weight of a Resource Centre Availability and Reliability is now estimated more accurately as HEP-SPEC 06 is a more precise estimation of the capacity installed as Resource Centres typically provide heterogeneous CPUs with different power.

- From March 2011 a snapshot of the RC topology is taken at the end of each reporting period in order to accurately retrieve topology information in case of Availability/Reliability amendments are needed. In particular, the list of RCs where in status "certified" on the last day of the month is recorded. By doing so, when re-computing statistics in case of problems after the month is expired (say on day X of the subsequent month), RCs that were uncertified before the end of the reporting month but became uncertified before day X, are excluded from re-computations. Before March 2011, these RCs were included in the reports. As this deficiency contributed to degrade the overall NGI performance monthly statistics during the NGI creation processes, the problem was fixed.

- From May 2011 additional service types were introduced in the computation of the Compute Element Availability: CREAM and ARC-CE. This change was needed as lcg-CE is close to end-of-support and it is being progressively replaced by alternative implementations.

---

[3] For details about the Availability and Reliability computation algorithm see https://twiki.cern.ch/twiki/pub/LCG/GridView/Gridview_Service_Availability_Computation.pdf

4.2 Process

Availability and Reliability reports are generated usually within the first week of each new month and the overall distribution process is managed under the responsibility of task TSA1.8. After a first sanity check for any obvious issues, reports are placed in the EGI document server for public access. Easy access to all reports is provided by the EGI wiki [PERF].

The Operations Management Board is notified when a new monthly report is available. A ticket is then opened in GGUS and assigned to the Central On Duty (COD) operations team. COD is responsible of:

- following up with RCs that have missed the monthly Service Level Targets. Tickets are opened to the relevant Operations Centres, which hold the responsibility of contacting the RCs in their domain to collect explanations.

- requesting the suspension of RCs that have been underperforming for three consecutive calendar months (see the details about the Suspension Policy for underperforming RCs in Section 5). COD has the right to suspend in case of an unresponsive Operations Centre. RCs eligible for suspension can be exempted upon request of the respective Operations Centre if there are documented reasons to believe that the technical problems causing low performance were solved.

The full Service Level Management procedure is documented at [PERF].

The procedure to be followed by RC administrators and Operations Centres to request amendments in the monitoring results and in the Availability/Reliability reports is currently under discussion.
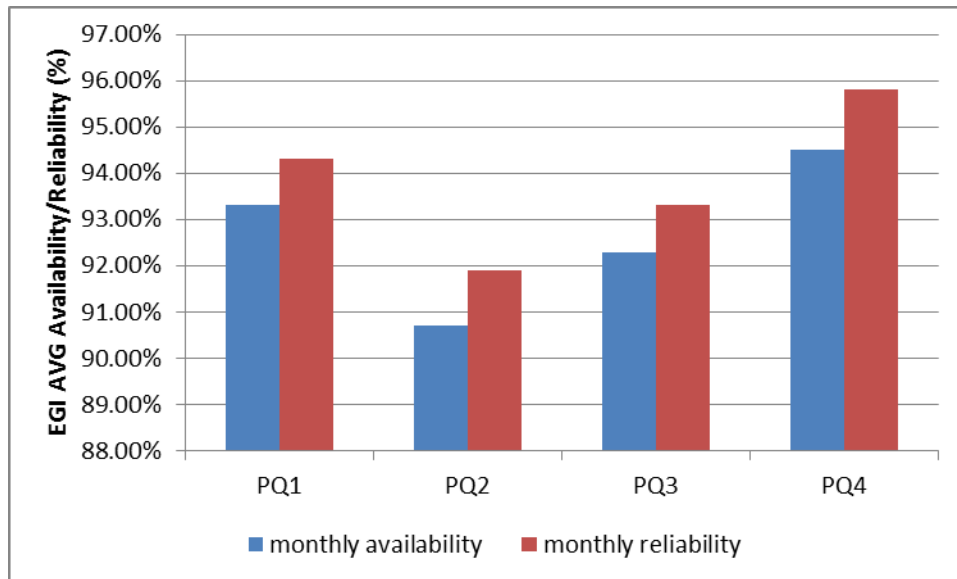
4.3 EGI Availability and Reliability

The overall EGI monthly Availability scored during PY1 of EGI-InSPIRE is 92.73%, and Reliability amounts to 93.85%. These values respectively compare to 91.88% and 92.73%, which are the values scored during the last year of EGEE-III[4]. Six are the RCs that were suspended during PY1 of EGI-InSPIRE.

Figure 2 shows the Availability and Reliability trend from May 2009 (start of PY2 of EGEE-III) to April 2011 (end of EGI-InSPIRE PY1). The objective of EGI-InSPIRE during the first year was to preserve the EGEE service level while migrating from EGEE Regional Operations Centres to independent NGIs. The stop of operations of Central Europe and South East Europe generated a large number of medium-small NGIs at different levels of maturity, and this is reflected in a deep of EGI Availability and Reliability during PQ2 and PQ3 (Figure 2).
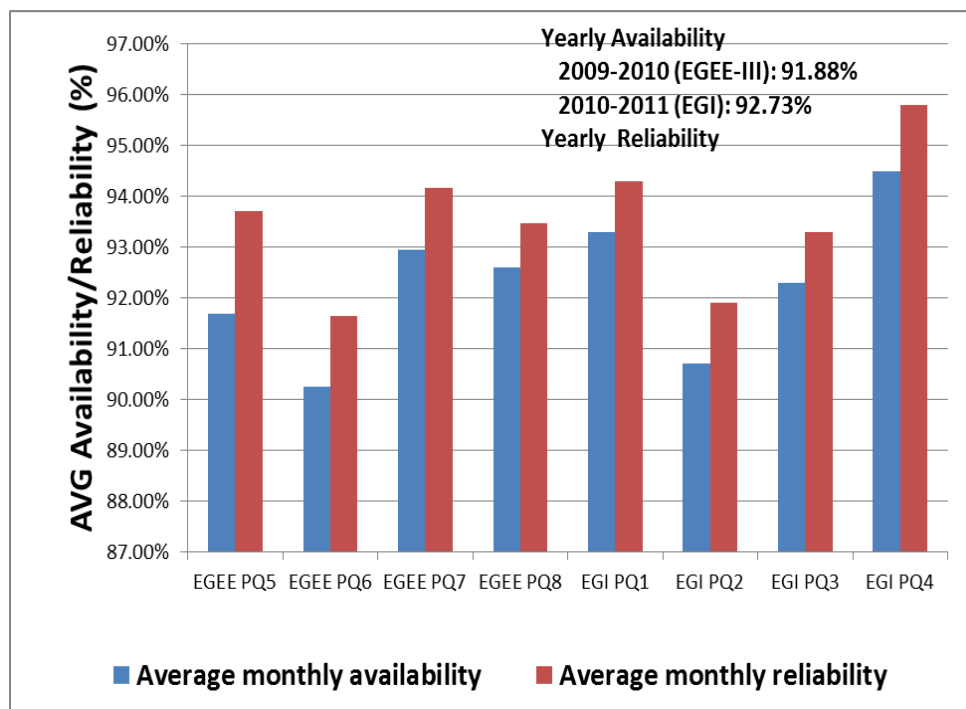
Both the EGI monthly Availability and Reliability have slightly improved in comparison with the corresponding figures recorded during PY2 of EGEE-III, this is shown in Figure 3.

---

[4] The complete set of EGEE-III Availability and Reliability statistics is available at https://edms.cern.ch/document/963325.

Figure 2. EGI monthly Availability and Reliability from EGI-InSPIRE PQ1 to PQ4. Availability and Reliability are computed as the mathematical average across the weighted RC monthly Availability and Reliability of the certified RCs in EGI.


Figure 3. Evolution of the overall infrastructure monthly Availability and Reliability from May 2009 to April 2011.

Status of the monthly Availability and Reliability performance can now be monitored during the course of the month through the Operations Portal. For each RC, the running monthly Availability

and Reliability are reported in order to assist RC administrators to proactively take counter measures in case of low performance.

A snapshot of the relevant Operations Portal Dashboard feature is provided in Figure 4.
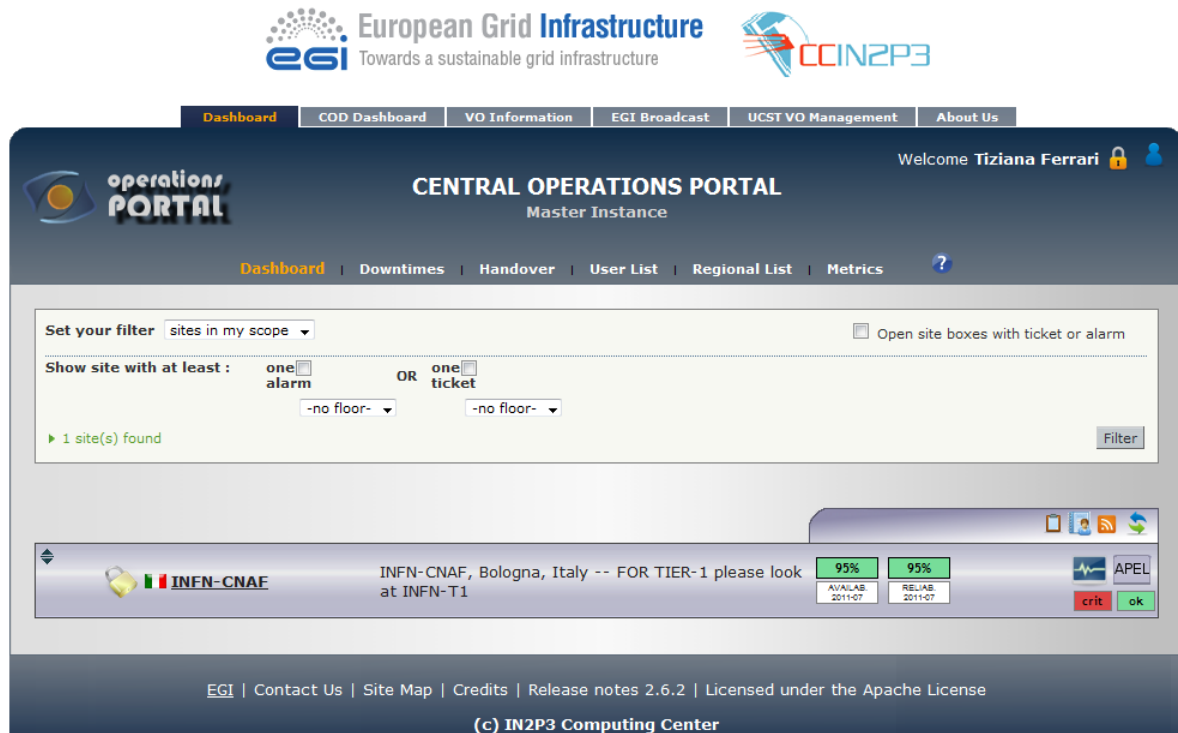
Figure 4. Monitoring of the progress of the monthly Availability and Reliability through the Operations Portal Dashboard (example provided for the INFN-CNAF production RC).

5   SUSPENSION POLICY

Infrastructure performance has been progressively improving during 2010 and 2011 as detailed in Section 4.3. As a consequence of this, the Operations Management Board decided to update the EGEE-III suspension policy to be stricter. It originally required an under-performing RC to be suspended when not delivering a minimum availability of 50% for three consecutive months. The minimum Availability was increased in May 2011 from 50% to 70% in order to keep it consistent with the minimum monthly Availability requested as Service Level Target by the RC OLA.

The new policy was endorsed after 6 months of internal assessment of the impact on the production infrastructure [POL].

COD is the operations entity responsible of monitoring the service level delivered by RCs. The entities responsible of performing suspension are the Operations Centre of the affected infrastructure, or alternatively, COD in case of an unresponsive Operations Centre.

6 CONCLUSIONS AND FUTURE WORK

OLAs are needed in EGI in order to provide a set of high-quality services to end-users by defining the service quality levels and the responsibilities of the individual operational units.

As the infrastructure matures and new technologies are adapted, OLAs will continue to evolve in order to match the changing environment and the increased expectations the user communities have. The feedback from RCs and RPs is of particular importance in order to have OLAs that reflect the actual expectations from the infrastructure. At the same time this evolution process allows awareness for the responsibilities to be increased among the involved players, which combined with the monthly tracking of Availability and Reliability allows the performance and the usability of the infrastructure to increase as EGI-InSPIRE progresses.

As the RC OLA has been finalised, focus for the next months will be on the RP OLA. Originally started as an NGI OLA, this OLA will include the responsibilities of the RPs towards their RCs, users and the EGI ecosystem.

A first version of this OLA was drafted in April 2011 and is available at [RPO]. The OLA Task Force has been collecting feedback from NGIs via the mailing list and a more complete version of the RP OLA is planned to be produced within summer 2011.

The Availability/Reliability reporting framework will be progressively extended to cover the central operational tools provided by EGI.eu in collaboration with various partners. To do so, the VIRTUALOPS GOCDB structure that includes the central operational tools will be renamed to EGI.EU, updated with the tool endpoints hosted in the egi.eu domain, and all local instances will be migrated to the respective NGI/ROC structures. EGI.EU will be used for the generation of a new profile that takes into account all EGI.eu central tools. This is a pre-requisite for the implementation of EGI.EU Availability and Reliability reports.

In parallel to this, the development of tool probes will be completed in collaboration with the Product Teams in JRA1. These are necessary for the collection of the monitoring results needed for the computation of Service Level reports.

7 REFERENCES

ARCH	EGI Operations architecture, EGI-InSPIRE Deliverable D4.1, Feb 2011 (https://documents.egi.eu/document/218)
ETF	OLA Workshop, Amsterdam, EGI Technical Forum Sep 2010 (https://www.egi.eu/indico/conferenceTimeTable.py?confId=48#20100915)
DOW	EGI-InSPIRE Description of Work (https://documents.egi.eu/document/10)
GLO	EGI Glossary (https://wiki.egi.eu/wiki/Glossary)
HS06	HEP-SPEC 06 Definition (http://go.egi.eu/hep-spec06)
ITIL	Information Technology Infrastructure Library web RC http://www.itil-officialRC.com/home/home.asp
PERF	https://wiki.egi.eu/wiki/Availability_and_reliability_monthly_statistics
POEM	Profile Management Database (https://tomtools.cern.ch/confluence/display/SAM/POEM)
POL	Evaluation of the impact of increasing availability suspension threshold https://wiki.egi.eu/wiki/Availability_and_reliability_threshold_change_impact
PROC09	Resource Centre Registration and Certification Procedure (https://wiki.egi.eu/wiki/PROC09)
QST	OLA NGI Questionnaire (https://documents.egi.eu/document/58)
RCO	EGI Resource Centre Operational Level Agreement, May 2011 (https://documents.egi.eu/document/31)
RPO	RP OLA draft https://documents.egi.eu/document/463
SLD	EGEE III Service Level Description (https://edms.cern.ch/document/860386)
UMD	Unified Middleware Distribution Roadmap (http://www.egi.eu/deployed-technology/technology-roadmap/)
USE	Availability and Reliability use cases, meeting minutes (https://documents.egi.eu/document/466)