

**EGI-Engage**

**First release of the new Accounting Portal deployed in production**

D3.6

|  |  |
| --- | --- |
| **Date** | 31 March 2016 |
| **Activity** | WP3 |
| **Lead Partner** | CSIC / CESGA |
| **Document Status** | DRAFT |
| **Document Link** | <https://documents.egi.eu/document/2790> |

Abstract

This deliverable describes the first release of the EGI Accounting Portal during EGI-Engage including the developments made during the first year of the EGI-Engage project. The EGI Accounting Portal receives data from APEL and ultimately from sites participating in the EGI and WLCG infrastructures as well as from sites belonging to other Grid organisations that are collaborating with EGI. This is crossed with metadata from other sources to offer an integrated view of accounting data on the EGI Infrastructure.

**COPYRIGHT NOTICE**



This work by Parties of the EGI-Engage Consortium is licensed under a Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). The EGI-Engage project is co-funded by the European Union Horizon 2020 programme under grant number 654142.

**DELIVERY SLIP**

|  |  |  |  |
| --- | --- | --- | --- |
|  | ***Name*** | ***Partner/Activity*** | ***Date*** |
| **From:** | Iván Díaz Álvarez | CSIC / CESGA | 28/03/2016 |
| **Moderated by:** |  |  |  |
| **Reviewed by** |  |  |  |
| **Approved by:** |  |  |  |

**DOCUMENT LOG**

|  |  |  |  |
| --- | --- | --- | --- |
| ***Issue*** | ***Date*** | ***Comment*** | ***Author/Partner*** |
| **v0.1** | 03/28/16 | Document creation | I. Díaz / CESGA |
| **...** |  |  |  |
| **...** |  |  |  |
| **v.n** |  |  |  |

**TERMINOLOGY**

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

**Contents**

[1 Introduction 5](#_Toc447203799)

[2 Service architecture 5](#_Toc447203800)

[High-Level Service architecture 6](#_Toc447203801)

[3 Integration and dependencies 7](#_Toc447203802)

[APEL SSM 7](#_Toc447203803)

[Metadata Gathering 7](#_Toc447203804)

[4 Release notes 8](#_Toc447203805)

[Requirements covered in the release 8](#_Toc447203806)

[5 Feedback on satisfaction 9](#_Toc447203807)

[6 Future plans 9](#_Toc447203808)

**Executive summary**

This deliverable describes the first EGI-Engage release of the EGI Accounting Portal including the developments made during the first year of the project. The EGI Accounting Portal receives data from APEL and ultimately from sites participating in the EGI and WLCG infrastructures as well as from sites belonging to other Grid organisations that are collaborating with EGI. This is crossed with metadata from other sources to offer a integrated view of accounting data on the EGI Infrastructure.   
The Accounting Portal has a MVC architecture with a backend loading data and metadata into a MySQL database which is accessed with a Model layer. There are several views that expose this as a web application.

The dependencies are in two groups, the main data gathered from APEL using SSM and Metadata gathered from several sources that is used to categorize and make sense of the accounting data.   
This release is the first with a new implementation based on the Django Python framework, and the Dojo and Bootstrap Javascript frameworks that both update the interface to support the latest web technologies and exposes new functionality and flexibility to the end user.   
The next features to be implemented are the rest of the accounting reports, the Storage view, new geographical graphing capabilities and the ability to compare data from two separate periods.

# Introduction

The following table provides a summary of the tool covered in this release.

|  |  |
| --- | --- |
| **Tool name** | *Accounting Portal* |
| **Tool url** | [*https://accounting-devel-next.egi.cesga.es/*](https://accounting-devel-next.egi.cesga.es/) |
| **Tool wiki page** | [*https://wiki.egi.eu/wiki/Accounting\_Portal*](https://wiki.egi.eu/wiki/Accounting_Portal) |
| **Description** | *EGI Core Service – The Accounting Portal provides data accounting views for users, VO Managers, NGI operations and the general public.* |
| **Value proposition** | *This new release of the Accounting Portal uses a new architecture and user interface that updates it to use new Web technologies and exposes more flexibility to the end user.* |
| **Customer of the tool** | *EGI, WLCG, Site and VO Admins, Infrastructure Users, others.* |
| **User of the service** | *Any stakeholder interested on the accounting of jobs in the EGI infrastructure.* |
| **User Documentation** | <https://documents.egi.eu/document/2789> |
| **Technical Documentation** | <https://documents.egi.eu/document/2545> |
| **Product team** | *CESGA / CSIC* |
| **License** | *Apache License, Version 2.0* |
| **Source code** | [*https://github.com/cesga-egi/accounting*](https://github.com/cesga-egi/accounting) |

# Service architecture

The service architecture provides an overview of the key (logical) service components and their dependencies to help better understand the structure and logical as well as technical setup of the main data gathered from APEL using SSM and Metadata gathered from several sources that is used to categorize and make sense of the accounting data.he service.

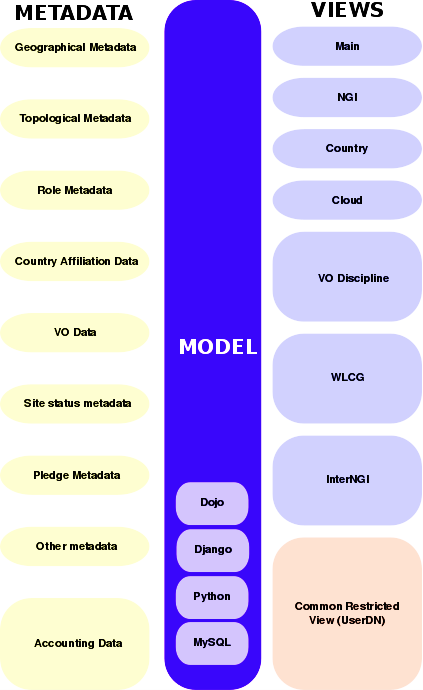
## High-Level Service architecture

The Accounting Portal is a web application which has as its primary function to provide users with customized accounting reports, containing tables and graphs, as web pages. It also offers RESTful web services to allow external entities to gather accounting data.

This chapter details the basic architecture of the Portal, which consists on:

1. A backend, which aggregates both data and metadata in a MySQL database, using the APEL SSM messaging system1 to interact with the Accounting Repository and several scripts, which periodically gather the data, and metadata described below.
2. A Model represented by database schemas both external and internal which define database tables for several types of accounting (grid, cloud, storage, multicore, user statistics etc.) and metadata (topology, geographical data, site status, nodes, VO users and admins, site admins etc.) and a series of parametrized queries,
3. A set of views that expose the data to the user. These views contain a form to set the parameters and metric of the report, a number of tables showing the data parametrized by two selectable dimensions and filtered by several parameters, a line graph showing the table data, and pie charts showing the percentage distribution on each dimension. It is planned that this part of the portal will evolve with interactive graphs, responsive in real time, reactive and only exposing advanced controls on user demand.

A graphical representation of these components is depicted on Fig. 1.

**

1 – Accounting Portal Architecture

# Integration and dependencies

The dependencies of the Portal are divided in two categories, the main data gathered from APEL using SSM and Metadata gathered from several sources that is used to categorize and make sense of the accounting data.

## APEL SSM

The Accounting Portal has to refresh its database periodically with data from the Accounting Repository to assure that information published is up-to-date. Secure Stomp Messenger (SSM), a queue messaging system based on ActiveMQ, is used for synchronization purpose and also for the communication between sites and the Accounting Repository. The SSM system is composed by:

1. A SSM loader for each accounting source (multicore, cloud, storage, etc.). This daemon waits for messages arriving on a queue and authenticates them with a DN and certificate. If a message is deemed valid, it is saved to a spool directory for further processing.
2. A DB loader, this daemon monitors the spool directory and, if there are messages, these are imported in the DB in order. This import at present does not delete the previous data in the tables; it only overwrites it. Manual intervention is needed for stale data.

The accounting data is sent several times per day by APEL in chunks of 1000 registers. This obviates the need for the portal to do pull requests.

## Metadata Gathering

Metadata is a category of data that complements the raw accounting data and allows the portal to organize, categorize and impart new meaning to it. This metadata includes:

1. **Geographical Metadata**: Country and NGI affiliation of sites. Generally, this follows current borders, but there are important exceptions. This is gathered from GOCDB using its XML-based API.
2. **Topological Metadata**: Sites are presented in trees, there are Country and NGI trees that correspond to geographical classifications, but there are also trees based on topological classifications like Tier1 and Tier2 sites, OSG sites and uncategorised sites. Inside Tier2 sites, the federation they belong to is also important and can trigger special code in some cases. Gathered from several sources, including OSG and WLCG databases.
3. **Role Metadata**: VO members and managers, and the site admins records. This metadata controls the access to restricted views. Information is gathered from GOCDB and individual VOMS servers constructing a list of individual VOMSes and querying them with the VOMS API.
4. **Country affiliation data**: Each user record contains a user identifier that has his/her user name, institution and sometimes country. Scripts in the backend map each user with a country based on the institution which issues their certificate. This data is used in anonymized statistics per country on: how much resources from other countries are used by given country and the distribution of its resources used by other countries.
5. **VO Data**: To make possible VO selection in the user interface, the portal stores lists of VOs. They are also used to filter incorrect VO names, provide access to VO managers, and arrange accounting by VO discipline (such as “High Energy Physics”, “Biomedicine”, “Earth Sciences”, etc.). Information is gathered from the Operations portal using its XML based APIs.
6. **Site status metadata**: Sites must be filtered to exclude those that are not in production (due to being closed or being in test mode). There must be also metadata to aggregate the accounting history of sites whose name has been changed. There are requirements to extend this functionality to NGIs. Information is gathered from GOCDB using its XML tables and internal tables compiled as part of EGI PROC 152.
7. **Pledge metadata**: The WLCG reports have to contain only those sites where MoUs or other pledges between VOs and sites are honoured, so the validity date and pledged hours are needed. Information is gathered from WLCG using the REBUS service.
8. **Other metadata:** There are also other metadata like local privileges, SpecInt calculations, publication status, VO activities and more. Some of these metadata is calculated internally using other types of metadata and published for other EGI operational tools, like VO activity data and Site UserDN publishing

# Release notes

## Requirements covered in the release

This release is the first with a new implementation based on the Django Python framework, and the Dojo and Bootstrap JavaScript frameworks that both update the interface to support the latest web technologies and exposes new functionality and flexibility to the end user.

* **Accounting Portal navigation tree** - <https://rt.egi.eu/rt/Ticket/Display.html?id=10156> – This is included as the new VO Discipline navigation tree that allows the user to see the three-level discipline hierarchy, view all subdisciplines and VOs for any discipline in any level and jump seamlessly between them.
* **Report generator in Portal -** <https://rt.egi.eu/rt/Ticket/Display.html>[?id=9733](https://rt.egi.eu/rt/Ticket/Display.html?id=9733). This is covered by the VO discipline functionality in the previous requirement.
* **Front page** - <https://rt.egi.eu/rt/Ticket/Display.html?id=8823> - Implemented, it uses last week instead of last 24h since we don't receive data so fresh currently.
* **Export data function** - <https://rt.egi.eu/rt/Ticket/Display.html?id=10159> - This is supported as JSON for the time being. XML and CSV will be included later.
* **Unofficial VO list in the Portal** - <https://rt.egi.eu/rt/Ticket/Display.html?id=9704> - There is a curated list of EGI VOs - disciplines now.
* **EGI Scientific Discipline Classification in the accounting portal**- <https://rt.egi.eu/rt/Ticket/Display.html?id=9583> - As with #10156 above.
* **Simplify access to some basic functionality. Avoid the use of complex forms for common statistics and get accounting information for some common queries** - <https://rt.egi.eu/rt/Ticket/Display.html?id=9075> – Implemented with an accordion hiding advanced form options by default.
* **RESTFUL interface to XML output (or JSON, csv) -**  <https://rt.egi.eu/rt/Ticket/Display.html?id=9052> - We currently use JSON.
* **Accept Storage Records by messaging**. <https://rt.egi.eu/rt/Ticket/Display.html?id=9051> Done.
* **Improve graphs visualization (New)** <https://rt.egi.eu/rt/Ticket/Display.html?id=8827> Done.
* **Portal to accept all data by messaging** - <https://rt.egi.eu/rt/Ticket/Display.html?id=9048> We currently only use SSM, a messaging service.

# Feedback on satisfaction

*Who was involved in testing and what the outcome of the review was*

# Future plans

The next features to be implemented:

* **More Accounting reports –** Specialised report pages will be done, in some cases utilizing the geographical capabilities below.
* **Storage view –** A view will be added for the storage accounting records, with appropriate metrics and automatic unit scaling.
* **New geographical graphing capabilities –** The graphing engine based in Dojo allows to represent numerical data using a geographical vectorial map. Maps will be provided for European countries and the World that will allow a new way to visualize the accounting data.
* **Comparing data from two separate periods –** The normal accounting view already has a start and end date selector that defaults to the last year. This will include a second time period that will serve as comparison base to compare any metric and grouping and derive trends.
* **GPGPU Accounting –** Accelerated computing records will be supported shortly after they are made available from APEL.