**EGI-InSPIRE**

EGI Federated Cloud Blueprint v1

**EU MILESTONE: MS520**

|  |  |
| --- | --- |
| Document identifier: | EGI-MS520-1773-v1bis-20130529-MD.docx |
| Date: | 29/05/2013 |
| Activity: | **SA2** |
| Lead Partner: | **OerC** |
| Document Status: | **DRAFT** |
| Dissemination Level: | **PUBLIC** |
| Document Link: | <https://documents.egi.eu/document/1773>  |

|  |
| --- |
| Abstract<< The abstract should provide a brief neutral overview of the document and its contents and main conclusions. Once complete the abstract should be copied into the abstract field on the document server.>><<document handling and production procedure is provided in <https://documents.egi.eu/document/33> >> |

1. 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 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, and 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.

1. Delivery Slip

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Name** | **Partner/Activity** | **Date** |
| **From** | <<The lead author/editor>> |  |  |
| **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>> |  |  |

1. Document Log

|  |  |  |  |
| --- | --- | --- | --- |
| **Issue** | **Date** | **Comment** | **Author/Partner** |
| 1 |  | First draft |  |
| 2 |  |  |  |
| 3 |  |  |  |

1. 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. 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>

1. Terminology

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

<<The authors should check if the acronyms are covered by the glossary page and if the definition is still correct; all the amendments should be communicated to glossary@egi.eu>>

1. 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.

1. EXECUTIVE SUMMARY

<< The text should provide a summary of the full report so that the reader can ‘in a page’ understand the problem it has been written to cover. This includes an overview of the background material and motivation for the report, a summary of the analysis, and the report’s main conclusions.>>

TABLE OF CONTENTS

1 Introduction 7

2 Federation model 8

3 Management Interfaces 10

3.1 VM management interface: OCCI 10

3.2 Data management interface: CDMI (I Livenson) 11

3.2.1 CDMI Objects 12

3.2.2 Detection of capabilities 12

3.2.3 Export protocol 12

3.3 Virtual Organisation Management & AAI: VOMS 12

3.4 Image management interface: vmcatcher/caster 13

3.4.1 Image distribution Design 13

3.4.2 Image transfer Specifications 15

3.4.3 Implementations of image list based image transfer. 16

4 EGI Core Services for Cloud 19

4.1 Information discovery: BDII 19

4.1.1 Technical implementation of the federated cloud information system 19

4.2 Central service registry: GOCDB 20

4.3 Monitoring: SAM 20

4.4 Accounting: (A. Packer) 21

4.5 4.5 Image metadata publishing & repository (K Koumantaros) 23

5 5. Federating Cloud Resources to EGI (M Turilli) 25

5.1 5.1 Overview: 6 steps process (M Turilli) 25

5.2 5.2 OpenNebula (ON RP team) 25

5.3 5.3 OpenStack (OS RP team) 27

5.4 5.4 StratusLab (StratusLab RP team) 29

5.5 5.5 WNoDeS (E. Ronchetti) 30

5.6 5.6 Synnefo (K Koumantaros) 30

6 6. Conclusions (D Wallom) 31

7 Conclusion 32

8 References 33

# Introduction

The current high throughput model of grid computing has proven extremely powerful for a number of different communities, though it has also limited in a number of cases the uptake of e-infrastructure. EGI therefore chose to investigate how to broaden the communities and application design models that would be able to take advantage of EGI core infrastructure and the level of dependency its production infrastructure is able to provide.

It was clear that the utilisation of virtualization and Infrastructure as a Service (IaaS) cloud computing was a clear candidate to enable this transformation though with a number of different open source providers already in use across a number of different resource providers it was clear that the previous approach of a mandated single software stack was not going to be possible. Instead following on from a number of different activities already on-going in the EC including SIENA, an approach that required the utilisation where possible of open standards where available and where not methods that have broad acceptance in the e-infrastructure community were essential.

The Task Force as originally configured has an 18-month mandate from September 2011, which is subdivided into 3 succinct six-month blocks;

1) Setup - Identify resource and technology providers and draft the model

2) Consolidation - Engage exemplar user communities and start configuration of test-bed

3) Integration - Establish test-bed fully with early adopter user communities and document

Overall goals for the activity were:

* Write a blueprint document[[1]](#footnote-1) for EGI Resource Providers that wish to securely federate[[2]](#footnote-2) and share their virtualised environments as part of the EGI production infrastructure;
* Deploy a test bed[[3]](#footnote-3) to evaluate the integration of virtualised resources within the existing EGI production infrastructure for monitoring[[4]](#footnote-4), accounting[[5]](#footnote-5) and information services[[6]](#footnote-6);
* Investigate and catalogue the requirements[[7]](#footnote-7) for community facing services based on or deployed through virtualised resources;
* Provide feedback[[8]](#footnote-8) to relevant technology providers on their implementations and any changes needed for deployment into the production infrastructure;
* Identify and work with user communities[[9]](#footnote-9) willing to be early adopters of the test bed infrastructure to help prioritise its future development;
* Identify issues[[10]](#footnote-10) that need to be addressed by other areas of EGI (e.g. policy, operations, support & dissemination).

# Federation model

The federation of IaaS Cloud resources in EGI is built upon extensive autonomy of Resource Providers in terms of ownership of exposed resources. The current federation model in EGI for exposing Grid resources requires Resource Providers to deploy and operate a specific set of Grid Middleware components before they were integrated into the EGI federated Grid production infrastructure. In contrast to that, the federation model for distributed IaaS Cloud resources allows a lightweight aggregation of local Cloud resources into the EGI Cloud Infrastructure Platform. At the heart of the federation are thus locally deployed Cloud Management stacks. In compliance with the Cloud computing model, the EGI CLIP does not mandate deploying any particular or specific Cloud Management stack; it is the responsibility of the Resource Providers to research, identify and deploy the solution that fits best their individual needs for as long as the offered services implement the required interfaces and domain languages. These interfaces and domain languages, and the interoperability of their implementation with other solutions are the focus of the federation.

Consequently, the EGI Cloud Infrastructure Platform is modeled around the concept of an *abstract* Cloud Management stack subsystem that is integrated with components of the EGI Core Infrastructure Platform (see Figure 1).

Figure 1: Architecture of the EGI Cloud Infrastructure Platform

This architecture allows EGI to define the CLIP as a relatively thin layer of federation and interoperability around local deployments and integrations of Cloud Management stacks.

This architecture defines interaction ports with a number of services from the EGI Core Infrastructure Platform, and the EGI Collaboration Platform. At the same time, it defines the required external interfaces and corresponding interaction ports. All these ports will have to be realised by local Cloud Management stack deployments.

The main interaction points of Resource Providers must take care of thus are:

* Integrate with the EGI Core AAI
* Integrate with the EGI Core Accounting system
* Integrate with the EGI Core Monitoring system
* Provide a standardised Cloud Computing interface (OCCI)
* Provide a standardised Cloud Storage interface (CDMI)
* Provide a standardised interface to an Information Service

Additionally, by means of using the Appliance Repository and the VM Marketplace from the EGI Collaboration Platform EGI Cloud Infrastructure Platform is providing VM image sharing and re-use across EGI Research Communities.

At the time of writing the Cloud federation platform is maintaining its own, separate Information discovery system. Even though it is using the GLUE2 schema, some extensions and tweaks are not compatible with the canonical GLUE 2 specification. Hence Cloud Resource providers maintain local LDAP endpoints (usually deployed as a resource BDII) aggregated into a Cloud Platform Information Discovery service, which in turn allows access to the data using LDAP v3.

Figure 2 provides an overview of the current realisations of the abstract Cloud Management stack subsystem in the EGI Cloud federation. It illustrates that each existing realisation inherits the obligation to implement the interaction points from the generalised parent Cloud Management stack. At the same time, the EGI Federated Clouds Task (funded through the EGI-InSPIRE project) gives Resource Providers a platform to share their implementation solutions for a commonly deployed specific Cloud Management stack (e.g. OpenNebula and OpenStack). Section 5 is dedicated to the documentation of the steps necessary to integrate a local deployment of a given Cloud Management stack into the EGI Cloud federation.

Figure 2: Current realisations of the abstract Cloud Management stack component

Through this collaboration, Resource Providers gradually develop and mature deployment and configuration profiles around common Cloud Management stacks as illustrated in Figure 2. This blueprint document is the first iteration of the resource provider oriented documentation.

# Management Interfaces

To federate a cloud system there are several functions for which we must have a common interface defined. These are each described below and in all define the method by which a ‘user’ of the service would be able to interact.

## VM management interface: OCCI

**The Open Cloud Computing Interface** (OCCI) is a RESTful Protocol and API designed to facilitate interoperable access to cloud-based resources across multiple resource providers and heterogeneous environments. The formal specification is maintained and still being extended by OGF’s OCCI-WG, for details see <http://occi-wg.org/>. Intended deployment is depicted by Figure 3.

Figure 3Deployment of OCCI in a provider's infrastructure


OCCI’s specification consists of three basic elements, each covered in a separate specification document:

**OCCI Core** describes the formal definition of the the OCCI Core Model [1]. **OCCI HTTP Rendering** defines how to interact with the OCCI Core Model using the RESTful OCCI API [2]. The document denes how the OCCI Core Model can be communicated and thus serialised using the HTTP protocol. **OCCI Infrastructure** contains the definition of the OCCI Infrastructure extension for the IaaS domain [3]. The document defines additional resource types, their attributes and the actions that can be taken on each resource type. Detailed description of the abovementioned elements of the specification is outside the scope of this document. Simplified description is as follows.

OCCI Core defines base types **Resource**, **Link**, **Action** and **Mixin**. Resource represents all OCCI objects that can be manipulated and used in any conceivable way. In general, it represents provider’s resources such as images (Storage Resource), networks (Network Resource), virtual machines (Compute Resource) or available services. Link represents a base association between two Resource instances, it indicates a generic connection between a *source* and a *target*. The most common real-world examples are Network Interface and Storage Link connecting Storage and Network Resource to a Compute Resource. Action defines an invocable operation tied to a specific Resource instance or a collection of Resource instances. In general, Action is designed to perform complex high-level operations changing the state of the chosen Resource such as virtual machine reboot or migration. The concept of mixins is used to facilitate extensibility and provide a way to define provider-specific features.

In the Federated Cloud environment, OCCI is deployed as a variety of platform-specific implementations. An ongoing EGI-InSPIRE mini-project *TSA 4.4 Providing OCCI support for arbitrary Cloud Management Frameworks* aims to provide a common implementation to further improve interoperability.

[1] R. Nyren, A. Edmonds, A. Papaspyrou, and T. Metsch, “Open Cloud Computing Interface - Core," GFD-P-R.183, April 2011. [Online]. Available: http://ogf.org/documents/GFD.183.pdf

[2] T. Metsch and A. Edmonds, “Open Cloud Computing Interface - HTTP Rendering," GFD-P-R.185, April 2011. [Online]. Available: http://ogf.org/documents/GFD.185.pdf

[3] “Open Cloud Computing Interface - Infrastructure," GFD-P-R.184, April 2011. [Online]. Available: http://ogf.org/documents/GFD.184.pdf

## Data management interface: CDMI (I Livenson)

SNIA Cloud Data Management Interface (CDMI) defines an open standard for operations on storage objects in a RESTful way. Semantically the interface is very close to AWS S3 and MS Azure Blob, but is more open and flexible for implementation.

Figure 4: Cloud storage reference model (courtesy to SNIAcloud.com)

Figure 4 shows the conceptual model of a cloud storage. CDMI offers clients a way for operating both on a storage management system and single data items. The exact level of support depends on the concrete implementation and is exposed to the client as part of the protocol.

The design of the protocol is aiming both at flexibility and efficiency.  Certain heavyweight operations, e.g. blob download, can be performed also with a pure HTTP client to make use of the existing ecosystem of tools. CDMI is built around the concept of Objects, which vary in supported operations and metadata schema. Each Object has an ID, which is unique across all CDMI deployments.

### CDMI Objects

There are 4 objects most relevant in the context of EGI Federated cloud TF:

* **Data object**: Abstraction for a file with rich metadata.
* **Container: A**bstraction for a folder. Export to non-HTTP protocols is performed on the container level. Container might have other containers inside of them.
* **Capability:** Exposes information about a feature set of a certain object.
* **Domain:** Deployment specific information.

### Detection of capabilities

CDMI supports partial implementation of the standards by defining optional features and parameters. In order to discover what functionality is supported by a specific implementation, CDMI client can issue a GET request to a fixed url: /cdmi\_capabilities.

### Export protocol

Attachment of the storage items to a VM can often be performed more efficiently using protocols like NFS or iSCSI. CDMI supports exposing of this information via container metadata. A client can make use of this information to attach a storage item to a VM over an OCCI protocol.

More information about the CDMI standard can be found at <http://cdmi.sniacloud.com/>. An ongoing EGI-InSPIRE mini-project *TSA4.5 CDMI Support in Cloud Management Frameworks* aims to provide an implementation of CDMI, which integrates with OCCI-based infrastructure and supports Fedcloud use-cases.

## Virtual Organisation Management & AAI: VOMS

Within the task force, it has been decided to generally allow VOMS proxy certificates of VOs ‘fedcloud.egi.eu’ and ‘ops’, the latter mainly for Nagios monitoring. There are modules available for each cloud middleware that have been developed by TF members. Configuring these modules into a provider’s cloud installation will allow members of these VOs to access the cloud.

Figure 5 shows the main components involved. The user retrieves a VOMS attribute certificate from the VOMS server of the desired VO and thus creates a local VOMS proxy certificate. The VOMS proxy certificate is use in subsequent calls to the OCCI endpoints of OpenNebula or OpenStack using the rOCCI client tool.

The rOCCI client directly talks to OpenNebula endpoints, which map the certificate and VO information to local users. Local users need to have been created in advance, which is triggered by regular synchronizations of the OpenNebula installation with Perun. Perun manages the VOs that participate in the FCTF testbed.

In order to access an OpenStack OCCI endpoint, rOCCI client needs to retrieve a Keystone token from OpenStack Keystone first. The retrieval is transparent to the user and automated in the workflow of accessing the OpenStack OCCI endpoint. It is triggered by the OCCI endpoint rejecting invalid requests and sending back an HTTP header referencing the Keystone URL for authentication. Users are generated on-the-fly in Keystone, it does not need regular synchronization with Perun.

Figure 5: Model of the Federated Cloud authentication architecture

Generic information about how to configure VOMS support for OpenStack Keystone can be found at <http://keystone-voms.readthedocs.org/en/latest/>. Information specific to FCTF is located at <https://wiki.egi.eu/wiki/Federated_AAI_Configuration#OpenStack>.

For OpenNebula, the information can be found here: [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:\_Federated\_AAI:OpenNebula](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3A_Federated_AAI%3AOpenNebula).

Stratus lab provides multiple authentication mechanisms at once. They are documented here: <http://stratuslab.eu//documentation/2012/10/07/docs-syadmin-auth.html>.

## VM Image management

In a distributed, federated Cloud infrastructure, users will often face the situation of efficiently managing and distributing their VM Images across supporting Cloud resource providers. The VM Image management subsystem provides the user with an interface into the EGI Cloud federation to notify supporting resource providers of the existence of a new or updated VM Image. Sites then examine the provided information, and pending their decision pool the new or updated VM Image locally for instantiation.

This concept introduces a number of capabilities into the EGI Cloud Infrastructure Platform:

* **VM Image lifecycle management** – Apply best practices of Software Lifecycle Management at scale across EGI
* **Automated VM Image distribution** – Publish VM images (or updates to existing images) once, while they are automatically distributed to the Cloud resource providers that support the publishing research community with Cloud resources.
* **Asynchronous distribution mechanism** – Publishing images and pooling these locally are intrinsically decoupled, allowing federated Resource Providers to apply local, specific processes transparently before VM images are available for local instantiation, for example:
* **Provider-specific VM image endorsement policies** – Not all federated Cloud resource providers will be able to enforce strict perimeter protection in their Cloud infrastructure as risk management to contain potential security incidents related to VM images and instances. Sites may implement a specific VM Image inspection and assessment policy prior to pooling the image for immediate instantiation.

Two command-line tools provide the principal functionality of this subsystem; “vmcaster” to publish VM image lists and “vmcatcher” to subscribe to changes to these lists, respectively.

Figure 6: Main components and actors of the VM Image management subsystem



Research Communities ultimately create and update VM Images (or delegate this functionality). The Images themselves are stored in Appliance repositories that are provided and managed elsewhere, typically by the Research Community itself.

A representative of the Research Community then generates a VM Image list (or updates an existing one) and publishes it on an authenticated host, which is typically using a host certificate signed by a CA included in the EGI Trust Anchor profile.

Federated Clouds Resource Provider then subscribe to changes in VM Image lists by regularly downloading the list from the authenticated host, and comparing it against local copies. New and updated VM Images are downloaded from the appliance repository referenced in the VM Image list into a local staging cache and, where required, made available for further examination and assessment.

Ultimately, Cloud resource Providers will make VM Images available for immediate instantiation by the Research Community.

# EGI Core Services for Cloud

## Information discovery: BDII

Users and service managers need tools to retrieve information about the whole infrastructure and filter the huge amount of data to retrieve the subset of the infrastructure that fulfill their requirements. To achieve this target the information about the services in the infrastructure must be structured in a uniform schema and published by a common set of services usable both by automatic tools and human users.

The current standard deployed in EGI for the implementation of the common information system is the Berkeley Database Information Index (BDII). It is a software based on a LDAP server, and it is deployed in an hierarchical structure, distributed over the whole infrastructure.

The information system is structure in three levels: the grid or cloud services publish their information (e.g. specific capabilities, total and available capacity or VO supported by the service) using a standard Open Grid Forum format, the GLUE schema2. The information published by the services are collected by a Site-BDII, a service deployed in almost every site in EGI. The Site-BDIIs are queried by the Top-BDIIs - the higher level of the hierarchy- which contain the information of all the sites in the infrastructure and their services. NGIs usually provide an authoritative instance of Top-BDII, but every Top-BDII, if properly configured, should contain the same set of information. While Site-BDII is configured by site managers to pull the data from the resource information providers of the services, Top-BDII automatically retrieves the list of the production sites from GOCDB and pulls the information form their Site-BDIIs.

Users and tools can use the Top-BDII to look for the services that provide the capabilities and the resources to run their activities. A typical example of Top-BDII query is retrieving the list of services that support a specific VO.

### Technical implementation of the federated cloud information system

Currently the Federated Cloud information system is built starting from the resource provider level. Every resource provider is requested to deploy a LDAP server publishing the information about their services structured used the GLUE schema. The best technical choice is to go for OpenLDAP, which is available in almost all the \*nix machines in the world. On top of that, OpenLDAP is the server used by the gLite BDIIs, therefore it would be easy to use the same configuration files set-up used for the GRIS (Grid Resource Information Service) or the GIIS (Grid Index Information Service).

Cloud services are not yet implementing information providers, therefore the information are published directly by a site-level information provider, comparable to a site-bdii in the structure of the information published. This solution is considered acceptable as the number of cloud services deployed by a single resource provider are usually not as many as the services in grid sites.

The LDIF file to be loaded in the local LDAP server is generated by a prototype custom script, and the information published are the following:

* Cloud computing resources
* Service endpoint
* Capabilities provided by the service, such as: virtual machine management or snapshot taking. The labels that identify the capabilities are agreed within the taskforce.
* Interface, the type of interface – e.g. webservice or webportal – and the interface name and version, for example OCCI 1.2.0
* User authentication and authorization profiles supported by the service, e.g. X.509 certificates
* Virtual machines images made available by the cloud provider
* Operating system, and other environment configuration details
* Maximum number of cores – and physical memory – allocable in a single virtual machine

Clearly, multiple virtual machine types can be associated to a single cloud service. There are no limitations in the number of services published by a resource centre either.Currently the information published is modelled using the latest GLUE2.0 schema definition. An extension of the GLUE schema is under development to address the specific requirements of Cloud resources and add information about storage and network services.

The Federated Cloud taskforce deploys a central Top-BDII that pulls automatically the information from the local LDAP servers of the resource providers. This service can be used as a single entry point to query for all the resource centres supported by the test-bed, by users or other automatic tools.

Currently the central federated cloud Top-BDII is reachable at this address: [ldap://test03.egi.cesga.es:2170](file://localhost/ldap/%3A%3Atest03.egi.cesga.es/2170)

An example of a possible query is:

ldapsearch -x -H ldap://test03.egi.cesga.es:2170 -b o=glue '(objectClass=GLUE2Endpoint)' | perl -p00e 's/\r?\n //g' | grep -E 'GLUE2EndpointURL|GLUE2EndpointInterfaceName|GLUE2EndpointInterfaceVersion|dn\:' | awk '{printf("%s%s", $0, (NR%4 ? " === " : "\n"))}' | awk '{print ""$2" "$5" "$8" "$11}' | awk -F "GLUE2DomainID=" '{print $2}' | awk -F "," '{print $1 " "$3}' | awk '{print $1" "$4" "$3" "$5}' | sort

This query provides the interface names exposed by all the resource providers, the version of the implemented interface, and the endpoint that can be used to contact the service through the specific interface.

From a technical standpoint the resource centre LDAP server must answer to the port 2170, in order to be automatically polled by the Top-BDII

## Central service registry: GOCDB

EGI’s central service catalogue is used to manage the production infrastructure topology. The service is provided using the GOCDB tool that is developed and deployed within EGI. To allow Resource Providers to expose Cloud resources to the production infrastructure, a number of new service types were added to GODCB:

* eu.egi.cloud.accounting
* eu.egi.cloud.information.bdii
* eu.egi.cloud.storage-management.cdmi
* eu.egi.cloud.vm-management.occi
* eu.egi.cloud.vm-metadata.marketplace

For the time-being (i.e. until EGI is integrating federated Cloud resources into production), all registered Cloud resources are maintained in  test-bed mode.

## Monitoring: SAM

The SAM (Service Availability Monitoring) system is a framework consisting of:

Nagios monitoring system (https://[www.nagios.org](http://www.nagios.org/)),

custom databases for topology, probes description and storing results of tests

web interface MyWLCG/MyEGI (<https://tomtools.cern.ch/confluence/display/SAM/MyWLCG>)

Probes used to perform check of services are provided by service developers. SAM team only maintains probes which test internal SAM functions (e.g. communication with messaging system, database synchronization, etc). More information on SAM can be found [here](https://wiki.egi.eu/wiki/SAM).

Current set of probes used for monitoring cloud resources consists of:

OCCI probe - creates an instance of a given image by using OCCI and checks its status

BDII probe - basic LDAP check tries to connects to cloud BDII

Accounting probe - checks if the cloud resource is publishing data to Accounting repository

TCP checks - basic TCP checks used for CDMI services.

Central SAM instance is deployed for monitoring cloud resources (<https://cloudmon.egi.eu/nagios>). Once the set of probes is finalized probes will be included into official SAM release. Adding probes to official SAM will follow procedure “Adding new probes to SAM” (<https://wiki.egi.eu/wiki/PROC07>).

## Accounting: (A. Packer)

<<Describe the key components of the accounting subsystem, with a focus on those components that need to be installed at the Resource Provider’s infrastructure. Link through to other docs where possible for other APEL components>>

To account for resource usage across the resource providers the following have been defined:

* The elements to be accounted for;
* Mechanisms for gathering and publishing accounting data to a central accounting repository;
* How accounting data will be displayed by the EGI Accounting Portal.

A Cloud Accounting Usage Record defines the data elements which resource providers should send to the central Cloud Accounting repository.  These elements are as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Key** | **Value** | **Description** | **Mandatory** |
| VMUUID | string | Virtual Machine's Universally Unique IDentifier | Yes |
| SiteName | string | Sitename, e.g. GOCDB Sitename | Yes |
| MachineName | string | VM Id |  |
| LocalUserId | string | Local username |  |
| LocalGroupId | string | Local groupname |  |
| GlobalUserName | string | User's X509 DN |  |
| FQAN | string | User's VOMS attributes |  |
| Status | string | Completion status - started, completed, suspended |  |
| StartTime | int | Must be set if Status = Started (epoch time) |  |
| EndTime | int | Must be set if Status = completed (epoch time) |  |
| SuspendDuration | int | Set when Status = suspended (seconds) |  |
| WallDuration | int | Wallclock - actual time used (seconds) |  |
| CpuDuration | int | CPU time consumed (seconds) |  |
| CpuCount | int | Number of CPUs allocated |  |
| NetworkType | string | Description |  |
| NetworkInbound | int | GB received |  |
| NetworkOutbound | int | GB sent |  |
| Memory | int | Memory allocated to the VM (MB) |  |
| Disk | int | Disk allocated to the VM (GB) |  |
| StorageRecordId | string | Link to associated storage record |  |
| ImageId | string | Image ID |  |
| CloudType | string | e.g. OpenNebula, Openstack |  |

Scripts have been provided for OpenNebula and Openstack implementations to retrieve the accounting data required in this format ready to be sent to the Cloud Accounting Repository.  These scripts are available from:

* OpenNebula –<https://github.com/EGI-FCTF/opennebula-cloudacc>
* Openstack –<https://github.com/EGI-FCTF/osssm>

The APEL SSM (Secure Stomp Messenger) package is provided by STFC for resource providers to send their messages to the central accounting repository.  It is written in Python and uses the STOMP protocol, the messages contain cloud accounting records as defined above in the following format (example data added) where %% is the record delimiter:

APEL-cloud-message: v0.2

VMUUID: https://cloud.cesga.es:3202/compute/47f74797-e9c9-46d7-b28d-5f87209239eb 2013-02-25 17:37:27+00:00

SiteName: CESGA

MachineName: one-2421

LocalUserId: 19

LocalGroupId: 101

GlobalUserName: NULL

FQAN: NULL

Status: completed

StartTime: 1361813847

EndTime: 1361813870

SuspendDuration: NULL

WallDuration: NULL

CpuDuration: NULL

CpuCount: 1

NetworkType: NULL

NetworkInbound: 0

NetworkOutbound: 0

Memory: 1000

Disk: NULL

StorageRecordId: NULL

ImageId: NULL

CloudType: OpenNebula

%%

...another cloud record...

%%

...

%%

The OpenNebula and Openstack scripts produce the messages to be sent using the SSM package in the correct format.

The SSM package can be downloaded from here: <http://apel.github.io/apel/>

Detail about configuring SSM and publishing records may be found here:

[https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario4#Publishing\_Records](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3AScenario4#Publishing_Records)

SSM utilizes the network of EGI message brokers and is run on both the Cloud Accounting server at STFC and on a client at the Resource Provider site.  The SSM running on the Cloud Accounting server receives any messages sent from the Resource Provider SSMs and they are stored in an “incoming” filesystem.

A Record loader package also runs on the Cloud Accounting server and checks the received messages and inserts the records contained in the message into the MySQL database.

A Cloud Accounting Summary Usage Record has also been defined and the Summaries created on a daily basis from all the accounting records received from the Resource Providers is sent to the EGI Accounting Portal.  The EGI Accounting Portal also runs SSM to receive these summaries and the Record loader package to load them in a MySQL database storing the cloud accounting summaries.

The EGI Accounting Portal provides a web page displaying different views of the Cloud Accounting data received from the Resource Providers:

<http://accounting.egi.eu/cloud.php>

## 4.5 Image metadata publishing & repository (K Koumantaros)

<<Describe the key components and concepts of the VM Marketplace, the Appliance repositories. Highlight the various roles and responsibilities of the involved stakeholders.>>

The task force uses the appliance repository and marketplace developed by stratuslab as repositories for images and their metadata.   IaaS providers endorse images that are suitable for their infrastructure by signing their metadata and uploading them on the marketplace (<https://marketpalce.egi.eu>)  and make the image available either to EGI appliance repository (<https://appliance-repo.egi.eu>) or their local appliance repository.  The user is then able to browse the metadata for suitable images to instantiate in one of the federated IaaS.

# 5. Federating Cloud Resources to EGI (M Turilli)

The model of federation chosen in the EGI is one where all resources available to the user are equal. therefore this is therefore suitable for resource consumers who do not have their own cloud infrastructure available to them. Within this section we detail the concepts behind the method of federation chosen and how those technology providers within the task force have been able to adapt or integrate their technologies. We therefore have for each technology provider a section which details work done for integration and any specific configuration a deployer of this technology needs to do to connect to EGI. We also describe technology premise that the activity itself started with.

Add here the criteria for being integrated (and counted) into the testbed.

Mention here also the certification process (that is one step further) to be officially counted in the prod infra figures

<<The key idea of this section is specificity. While the previous sections describe the architectural and design concepts behind the federated Cloud infrastructure, this section lays out the concrete steps a Resource Provider would have to undertake in order to technically integrate their Cloud resources with the EGI Cloud federation Section 5.1 will provide the 6 principal steps as a sort of overview.

The following subsections will list the specific actions pertinent to the described CMF. So for example, if a Resource Provider would want to deploy WNoDeS, section 5.5 would be the right section to read on how to do that.>>

## 5.1 Overview: 6 steps process (M Turilli)

The initial plan for federation of cloud resources within EGI was based on 6 different functional requirements;

1)

<<see above>>

## 5.2 OpenNebula (ON RP team)

(section 5.1 is still empty, I’m not sure what to focus on; here is my attempt to provide a somewhat concise step-by-step guide -- Boris)

A new Resource Provider using OpenNebula or OpenNebula-based CMF has to take the following steps to join the EGI Cloud Federation. There is only one prerequisite and that is fully functional OpenNebula installation capable of deploying, sustaining and shutting down virtual machines. There are no requirements for the underlying architecture, Resource Providers in question may choose the virtualization platform, network and storage configuration according to their preferences and needs. It is highly recommended to install OpenNebula v3.8.x where x denotes the latest security update and coordinate any future upgrades with the task force to avoid infrastructure fragmentation. Resource providers installing OpenNebula from scratch should follow its step-by-step installation and configuration guides available online [1].

The actual integration with the EGI Cloud Federation consists of the following steps:

1. Additional OpenNebula configuration
2. rOCCI-server installation and configuration
3. Integration with user management service -- Perun
4. Integration with accounting service -- APEL
5. Integration with image management service -- vmcaster/vmcatcher
6. Integration with information system -- LDAP/BDII
7. Registration of deployed services in GOCDB

Each of the above-mentioned steps is a requirement for every Resource Provider wishing to join the EGI Cloud Federation. Resource Providers are welcome to deploy and offer additional services such as object storage (CDMI) but this is not a requirement at this time. Detailed description of the listed steps is as follows.

**Additional OpenNebula configuration**

Integration with EGI Cloud Federation requires the use of X.509 authentication mechanism in communication with OpenNebula. Resource Providers are encouraged to follow the step-by-step configuration guide provided by OpenNebula developers available online [2]. There is no need to change authentication driver for the *oneadmin* user or create any user accounts manually at this time.

**rOCCI-server installation and configuration**

The EGI Cloud Federation uses OCCI as its VM management protocol. It is necessary to install a fully compliant OCCI 1.1 server on top of RP’s existing OpenNebula installation. OpenNebula’s OCCI implementation is *not* compliant with the OCCI 1.1 specification. This functionality is provided by the rOCCI-server project. Detailed installation and configuration instructions are available online in the FedCloud wiki [3].

**Integration with Perun**

The current rOCCI-server implementation doesn’t handle user management and identity propagation hence an integration with a third-party service is necessary. The Perun management server developed and maintained by CESNET is used to provide user management capabilities for OpenNebula Resource Providers [4]. It uses locally installed scripts (fully under the control of the Resource Provider in question) to propagate changes in the user pool to all registered Resource Providers. They are required to install and configure (if need be) these scripts and report back to EGI Cloud Federation for registration in Perun. Installation and configuration details are available online in the EGI-FCTF repository on GitHub [5].

**Integration with APEL**

One of the required integration points is accounting. The EGI Cloud Federation employs APEL framework with extended accounting records. Every Resource Provider is required to install the APEL SSM client and OpenNebula accounting script. As with the previous cases, installation and configuration details are available online on GitHub and in the FCTF wiki [6][7].

**Integration with vmcaster/vmcatcher**

Resource Providers are required to integrate their OpenNebula with an image management service used within the federation. As with the previous cases, installation and configuration details are available online in the FCTF wiki [9]. This service ensures that all images are trusted and up-to-date for all Resource Providers across the federation.

**Integration with TopBDII**

Details about services offered by the Resource Provider in question are advertised to the rest of the EGI Cloud Federation using an LDAP server -- BDII. Resource Providers are encouraged to follow instructions available online in the FCTF wiki [8].

**Registration in GOCDB**

(I’m sorry, I have no idea how the process of registration in GOCDB goes or what is required … generally speaking, it should be the same for all Resource Providers regardless of their CMF, right?)[10]

[1] [http://opennebula.org/documentation:archives:rel3.8](http://opennebula.org/documentation%3Aarchives%3Arel3.8)

[2] [http://opennebula.org/documentation:archives:rel3.8:x509\_auth](http://opennebula.org/documentation%3Aarchives%3Arel3.8%3Ax509_auth)

[3] [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:\_Federated\_AAI:OpenNebula](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3A_Federated_AAI%3AOpenNebula)

[4] <http://perun.metacentrum.cz/web/>

[5] <https://github.com/EGI-FCTF/fctf-perun>

[6] <https://github.com/EGI-FCTF/opennebula-cloudacc>

[7] (Is there a step-by-step installation and configuration guide for APEL SSM?)

[8] <https://wiki.egi.eu/wiki/Fedclouds_BDII_instructions>

[9] [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario8:Configuration#VMcatcher](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3AScenario8%3AConfiguration#VMcatcher)

[10] (Is there a step-by-step registration guide for a FedCloud site in GOCDB?)

## 5.3 OpenStack (OS RP team)

<<see above>>

Mention glancepush and glancpush-vmcaster

This section describes steps necessary for new Resource Provider (RP) using Openstack middleware to join EGI Cloud Federation. It is strongly recommended using the last Openstack version. Specifically, the VOMS-enable authentication will require Grizzly version of Keystone. The installation and configuration of Openstack are available at [1].

The actual integration with the EGI Cloud Federation consists of the following steps:

a)      VOMS-enable Keystone installation and configuration

b)      OCCI installation and configuration

c)       Integration with accounting service APEL

d)      Integration with vmcaster/vmcatcher

e)      Integration with information system LDAP/BDII

f)       Registration of deployed services in GOCDB

Each of the above-mentioned steps is a requirement for every Resource Provider wishing to join the EGI Cloud Federation. Resource Providers are welcome to deploy and offer additional services such as object storage (CDMI) but this is not a requirement at this time. Detailed description of the listed steps is as follows.

**a)      VOMS-enable Keystone installation and configuration**

The installation and configuration of VOMS-enable Keystone is available at [2]. That will enable X.509 authentication mechanism and allows users with valid VOMS proxy certificate to log in. The actual VO for EGI Cloud Federation fedcloud.egi.eu should be enabled in the configuration. There is an option for automatic creating new users for trusted VO on the fly.

**b)      OCCI installation and configuration**

The steps of installation and configuration of OCCI is available at [3]. The installation and configuration should be done on the machine with Nova server. The OOCI implementation is not perfect; occasionally Nova server needs to be restarted for refreshing OCCI configuration (especially when new images are added).

**c)       Integration with accounting service APEL**

Like RP with OpenNebula, the client for accounting service APEL must be installed and configured. The details of installation and configuration of APEL for Openstack is available at [4], [5].

**d)      Integration with vmcaster/vmcatcher**

Resource Providers are required to integrate their Openstack with an image management service used within the federation. Installation and configuration details are available online in the FCTF wiki [6]. This service ensures that all images are trusted and up-to-date for all Resource Providers across the federation. In addition to vmcaster/vmcatcher, glancepush-vmcatcher [7] uses vmcatcher's event handler to signal glancepush that a new image was updated in vmcatcher's cache and glancepush will check and publish images from vmcatcher cache to glance service in Openstack.

**e)      Integration with information system LDAP/BDII**

Integration with BDII for RP with Openstack is identical as in the OpenNebula case. The instructions are available online in the FCTF wiki [8].

[1] http://docs.openstack.org/install/

[2] http://keystone-voms.readthedocs.org/en/latest/index.html

[3] https://github.com/tmetsch/occi-os

[4] https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario4

[5] https://github.com/EGI-FCTF/osssm/wiki

[6] https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario8:Configuration#VMcatcher

[7] https://github.com/EGI-FCTF/glancepush

[8] https://wiki.egi.eu/wiki/Fedclouds\_BDII\_instructions

## 5.4 StratusLab (StratusLab RP team)

A StratusLab cloud based production release offers computing, storage, and networking services as well as the Marketplace, a high-level service to facilitate sharing of machine images.

Actually StratusLab integrates and uses OpenNebula as the Virtual Machine Manager.

In order to technically integrate a StratusLab Cloud with the EGI Cloud Federation, the following steps should be implemented:

Authentication:

X509/VOMS authentication is done by creating users in the OpenNebula VMM service with the X509 driver. This driver should be enabled in the OpenNebula configuration file.

As the VOMS validation is passing through Apache2, grid\_site and then rOCCI-server. Apache2 and grid\_site should be properly configured. In our case, install gridsite packages, then load gridsite module in apache configuration file.

Compute: Nothing to do, rOCCI-server is well integrated with OpneNebula, and StratusLab actually integrates OpenNebula as it’s Virtual Machine Manager.

Network: StratusLab networking service permits the allocation of “public”, “private” and “local” network. When creating VM templates one of these networks should be specified.

Public IP are visible from outside and inside the cloud.

Local IP are visible from the Cloud (VMs running in the Cloud), access external services through NAT, this type of IP addresses could be useful for MPI jobs.

Private IP are visible only fron the host where they are running, go to outside via NAT.

Storage: StratusLab developed it’s own storage service solution based on a disk approach. One of its functionnality is to cache machine images, making deployment of VM very fast.

This service is well integrated with the other StratusLab services, and doesn’t need any additional configuration for rOCCI-server.

Marketplace: We are using StratusLab Marketplace to instantiate VM in the StratusLab Cloud. The StratusLab Marketplace is like a registry of images metadata.

Unlike the other RPs, in the metadata we don’t specify network nor storage elements. Instead in the VM templates it’s mandatory to specify identifier image url  from the marketplace in the SOURCE filed. (E.g. SOURCE=http://marketplace.egi.eu/metadata/HGSxEvjFP0TUo1mMcT-M63Y-2KF/airaj@lal.in2p3.fr/2013-02-12T14:11:55Z).

NB. In StratusLab, OpenNebula is used only as VMM service, and in the near future, it’ll be dropped. Integration will be done directly with libvirt.

More precisely, this configuration will works till the StratusLab 13.05 release, OpenNebula as VMM will be dropped in the  StratusLab 13.08 release, in August 2013.

## 5.5 WNoDeS (E. Ronchetti)

<<see above>>

## 5.6 Synnefo (K Koumantaros)

Synnefo (<http://www.synnefo.org/>) is open source cloud software used to create massively scalable IaaS clouds. It uses Google Ganeti for the low level VM management and also talks to the outside world through the OpenStack APIs with extensions for advanced operations.  Synnefo in conjunction with Google GANETI (<https://code.google.com/p/ganeti/>) is the software thats empowers GRNETs ~Okeanos service (<https://okeanos.grnet.gr>) that has currently supports 2100 users

with 2941 VMs and 10119 Virtual cores.   ~Okeanos is only partialy integrated with the Federated cloud infrastructure using snf-occi (<http://www.synnefo.org/docs/snf-occi/latest/index.html>), an implementation of the OCCI specification on top of synnefo’s API kamaki.   Development for the rest of the modules required is currently foreseen for the near future but due to lack of manpower and parallel developments of the synnefo API there is no estimate for the date of delivery for each module.

# 6. Conclusions (D Wallom)

# Conclusion

# References

|  |  |
| --- | --- |
| R 1 |  |
| R 2 |  |
| R 3 |  |
| R 4 |  |
| R 5 |  |

1. [https://wiki.egi.eu/wiki/Fedcloud-tf:Blueprint](https://wiki.egi.eu/wiki/Fedcloud-tf%3ABlueprint) [↑](#footnote-ref-1)
2. [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:\_Federated\_AAI](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3A_Federated_AAI) [↑](#footnote-ref-2)
3. [https://wiki.egi.eu/wiki/Fedcloud-tf:Testbed](https://wiki.egi.eu/wiki/Fedcloud-tf%3ATestbed) [↑](#footnote-ref-3)
4. [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario5](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3AScenario5) [↑](#footnote-ref-4)
5. [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario4](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3AScenario4) [↑](#footnote-ref-5)
6. [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario3](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3AScenario3) [↑](#footnote-ref-6)
7. [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:\_Outreach#Requirements](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3A_Outreach#Requirements) [↑](#footnote-ref-7)
8. [https://wiki.egi.eu/wiki/Fedcloud-tf:Blueprint:Solutions\_Intentory](https://wiki.egi.eu/wiki/Fedcloud-tf%3ABlueprint%3ASolutions_Intentory) [↑](#footnote-ref-8)
9. [https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:\_Outreach](https://wiki.egi.eu/wiki/Fedcloud-tf%3AWorkGroups%3A_Outreach) [↑](#footnote-ref-9)
10. [https://wiki.egi.eu/wiki/Fedcloud-tf:Blueprint:Security\_and\_Policy](https://wiki.egi.eu/wiki/Fedcloud-tf%3ABlueprint%3ASecurity_and_Policy) [↑](#footnote-ref-10)