H2020-EINFRA-2015-1

VI-SEEM

VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean

Deliverable D3.3

Infrastructure overview, assessment and refinement plan

Author(s): VI-SEEM Consortium

Status –Version: e

Date: March 23rd, 2017

Distribution - Type: Public

Abstract: Deliverable D3.3 – "Infrastructure overview, assessment and refinement plan", provides an overview of the VI-SEEM infrastructure resources and services offered in the middle of the project, as well as the plan for new services deployment until the end of the project.

The VI-SEEM Consortium consists of:

GRNET	Coordinating Contractor	Greece
CYI	Contractor	Cyprus
IICT-BAS	Contractor	Bulgaria
IPB	Contractor	Serbia
NIIF	Contractor	Hungary
UVT	Contractor	Romania
UPT	Contractor	Albania
UNI BL	Contractor	Bosnia-Herzegovina
UKIM	Contractor	FYR of Macedonia
UOM	Contractor	Montenegro
RENAM	Contractor	Moldova (Republic of)
IIAP-NAS-RA	Contractor	Armenia
GRENA	Contractor	Georgia
BA	Contractor	Egypt
IUCC	Contractor	Israel
SESAME	Contractor	Jordan

The VI-SEEM project is funded by the European Commission under the Horizon 2020 e-Infrastructures grant agreement no. 675121.

Document Revision History

Date	Issue	Author / Editor / Contributor	Summary of main changes
January 31st, 2017	a	Dusan Vudragovic	Initial version of ToC.
February 6th, 2017	b	Dusan Vudragovic	Final version of ToC.
March 15th, 2017	c	Ioannis Liabotis, Anastas Misev, Andreas Athenodorou, Todor Gurov, Dimitar Dimitrov, Mihajlo Savic, Andreas Panteli, Antun Balaz, Dusan Vudragovic, Vladimir Slavnic, Petar Jovanovic, Silviu Panica, Neki Frasheri, Nicolai Iliuha, Ramaz Kvatadze, Tamas Maray, Youssef Eldakar	Integration of partners' contributions.
March 21st, 2017	d	Dusan Vudragovic	First draft.
March 30th, 2017	e	Dusan Vudragovic, Ognjen Prnjat, Ioannis Liabotis	Final editing.

Preface

In the last decade, a number of initiatives were crucial for enabling high-quality research - by providing e-Infrastructure resources, application support and training - in both South East Europe (SEE) and Eastern Mediterranean (EM). They helped reduce the digital divide and brain drain in Europe, by ensuring access to regional e-Infrastructures to new member states, states on path to ascension, and states in European Neighborhood Policy area – in total 14 countries in SEE and 6 in EM.

This VI-SEEM project brings together these e-Infrastructures to build capacity and better utilize synergies, for an improved service provision within a unified Virtual Research Environment (VRE) for the inter-disciplinary scientific user communities in the combined SEE and EM regions (SEEM). The overall objective is to provide user-friendly integrated e-Infrastructure platform for regional cross-border Scientific Communities in Climatology, Life Sciences, and Cultural Heritage for the SEEM region; by linking compute, data, and visualization resources, as well as services, models, software and tools. This VRE aspires to provide the scientists and researchers with the support in full lifecycle of collaborative research: accessing and sharing relevant research data, using it with provided codes and tools to carry out new experiments and simulations on large-scale e-Infrastructures, and producing new knowledge and data - which can be stored and shared in the same VRE. Climatology and Life Science communities are directly relevant for Societal Challenges.

The driving ambition of this proposal is to maintain leadership in enabling e-Infrastructure based research and innovation in the region for the 3 strategic regional user communities: supporting multidisciplinary solutions, advancing their research, and bridging the development gap with the rest of Europe. The VI-SEEM consortium brings together e-Infrastructure operators and Scientific Communities in a common endeavor.

The overall objective is to provide user-friendly integrated e-Infrastructure platform for Scientific Communities in Climatology, Life Sciences, and Cultural Heritage for the SEEM region; by linking compute, data, and visualization resources, as well as services, software and tools.

The detailed objectives of the VI-SEEM project are:

1. Provide scientists with access to state of the art e-Infrastructure - computing, storage and connectivity resources - available in the region; and promote additional resources across the region.
2. Integrate the underlying e-Infrastructure layers with generic/standardised as well as domain-specific services for the region. The latter are leveraging on existing tools (including visualization) with additional features being co-developed and co-operated by the Scientific Communities and the e-Infrastructure providers, thus proving integrated VRE environments.
3. Promote capacity building in the region and foster interdisciplinary approaches.
4. Provide functions allowing for data management for the selected Scientific Communities, engage the full data management lifecycle, link data across the region, provide data interoperability across disciplines.
5. Provide adequate user support and training programmes for the user communities in the SEEM region.

6. Bring high level expertise in e-Infrastructure utilization to enable research activities of international standing in the selected fields of Climatology, Life Sciences and Cultural Heritage.

The VI-SEEM project kicked-off in October 2015 and is planned to be completed by September 2018. It is coordinated by GRNET with 15 contractors from Cyprus, Bulgaria, Serbia, Hungary, Romania, Albania, Bosnia-Herzegovina, FYR of Macedonia, Montenegro, Moldova (Republic of), Armenia, Georgia, Egypt, Israel, Jordan. The total budget is 3.300.000 €. The project is funded by the European Commission's Horizon 2020 Programme for Excellence in Science, e-Infrastructure.

The project plans to issue the following deliverables:

Del. no.	Deliverable name	Nature	Security	Planned Delivery
D1.1	Project management information system and "grant agreement" relationships	R	CO	M01
D1.2	3-Monthly progress report	R	CO	M03n *
D1.3a	First period progress reports	R	CO	M18
D1.3b	Final period progress reports	R	CO	M36
D2.1	Internal and external communication platform, docs repository and mailing lists	DEC	PU	M02
D2.2	Promotional package	DEC	PU	M04
D2.3	Dissemination and marketing plan	R	PU	M05
D2.4	Training plan	R	PU	M06
D2.5	Promotional package with updates	R	PU	M16
D2.6	1st Dissemination, training and marketing report	DEC	PU	M18
D2.7	2nd Dissemination, training and marketing report	R	PU	M35
D3.1	Infrastructure and services deployment plan	R	PU	M04
D3.2	Service registry, operational and service level monitoring	R	PU	M12
D3.3	Infrastructure overview, assessment and refinement plan	R	PU	M18
D3.4	VRE AAI Model and compatibility with other eInfrastructures	R	PU	M27
D3.5	Final infrastructure overview and assessment report	R	PU	M36
D4.1	Data sources and services deployment plan	R	PU	M06
D4.2	Description of the initial deployed data services	R	PU	M11
D4.3	Description of the final data platform available to VRE users	R	PU	M23
D4.4	Final report on data, services, availability and usage	R	PU	M35

D5.1	Detailed technical implementation plan for VRE services and tools	R	PU	M04
D5.2	Data management plans	R	PU	M06
D5.3	User-oriented documentation and training material for VRE services	R	PU	M13
D5.4	Report on integrated services and the VRE platform	R	PU	M14
D5.5	Final report on integrated services and the VRE platform	R	PU	M36
D6.1	Framework for VRE resource and service provision	R	PU	M09
D6.2	1st Report of open calls and integration support	R	PU	M20
D6.3	Sustainability and business model	R	PU	M24
D6.4	2nd Report of open calls and integration support	R	PU	M36

Legend: R = Document, report, DEC = Websites, patent fillings, videos, etc., PU = Public, CO = Confidential, only for members of the consortium (including the Commission Services).

* *n=1,2,3,...12*

Table of contents

References

[1] Project VI-SEEM-675121 - Annex I - Description of the action

[2] VI-SEEM D2.6 – 1st Dissemination, training and marketing report

[3] OASIS Security Assertion Markup Language (SAML)
https://docs.oasis-open.org/security/saml/v2.0/saml-metadata-2.0-os.pdf

[4] VI-SEEM D3.1 – Infrastructure and services deployment plan

[5] TOP500 Supercomputer Sites
https://www.top500.org/

[6] European Grid Infrastructure
https://www.egi.eu/

[7] VI-SEEM Virtual Organization VOMS-admin interface
https://voms.ipb.ac.rs:8443/voms/vo.vi-seem.eu

[8] gLite User Guide
https://edms.cern.ch/document/722398/

[9] VI-SEEM D6.1 – Framework for VRE resource and service provision

[10] EUDAT - Research Data Services, Expertise & Technology Solutions
https://eudat.eu/

[11] ARGO Service Monitoring system
http://argoeu.github.io/

[12] Official Nagios webpage
https://www.nagios.org/

[13] Python Flask framework
http://flask.pocoo.org/

[14] React library
https://facebook.github.io/react/

[15] osTicket – Support Ticket System
http://osticket.com/

[16] Official GitLab webpage
http://www.gitlab.com/

[17] FitSM – standards for lightweight IT service management
http://fitsm.itemo.org/

[18] VI-SEEM D3.2 – Service registry, operational and service level monitoring

List of Figures

List of Tables

Glossary

AAI	Authentication and Authorization Infrastructure
AMQP	Advanced Message Queuing Protocol
API	Application Programming Interface
BDII	Berkeley Database Information Index
CA	Certification Authority
CMDB	Configuration Management Database
CPU	Central Processing Unit
EGI	European Grid Infrastructure
EM	Eastern Mediterranean
EUDAT	European Data Infrastructure
GLUE	Grid Laboratory Uniform Environment
GOCDB	Grid Operations Centre Database
GOCDB	Grid Operations Centre Database
GPU	Graphics Processing Unit
HPC	High Performance Computing
HTTP	HyperText Transfer Protocol Secure
IGTF	Interoperable Global Trust Federation
ITIL	Information Technology Infrastructure Library
ITSM	IT Service Management
JSON	JavaScript Object Notation
LB	Logging and Bookkeeping
LDAP	Lightweight Directory Access Protocol
LFC	LCG File Catalogue
OASIS	Organization for the Advancement of Structured Information Standards
OLA	Operational Level Agreement
PKI	Public Key Infrastructures
PX	MyProxy software
REST	Representational State Transfer
SAML	Security Assertion Markup Language
SC	Service Catalogue
SEE	South East European
SEEM	South East Europe and Eastern Mediterranean
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
SP	Service Portfolio
TTS	Token Translation Service
UI	User Interface
VI-SEEM	VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean
VM	Virtual Machine
VO	Virtual Organization

VOMS	Virtual Organization Membership Service
VRE	Virtual Research Environment
WMS	Workload Management System
XML	Extensible Markup Language

Executive summary

What is the focus of this Deliverable?

The focus of this deliverable is to provide an overview of integrated computing (HPC, Grid, Cloud) and storage services available through the VI-SEEM infrastructure in the middle of the project. The deliverable also reports on the usage of the underlying resources, developed and deployed services that enable uniform management and access to offered resources, as well as a plan for deployment of new services until the end of the project.

What is next in the process to deliver the VI-SEEM results?

The contents of this deliverable form a basis for other workpackages so as to make efficient use of the infrastructure service offering. The service registry includes both common infrastructure services (WP3 for e-Infrastructure and WP4 for data services) that are used by all three supported scientific communities, as well as application-specific services (deployed within the WP5). Deployed Authentication and Authorization Infrastructure provides user access management used by WP3, WP4, and WP5 service components.

In particular, the contents of this deliverable will be used in the following VI-SEEM activities:

- All WP3 activities,
- WP4.1 – Data services design,
- WP4.3 – Data collection and provisioning,
- WP4.4 – Data analysis,
- WP5.1 – Refinement of service requirements and technical assessment for integration,
- WP5.3 – Development VRE platform,
- WP5.4 – Overall integration of services.

What are the deliverable contents?

The deliverable contents include a detailed overview of the infrastructure available to the VRE, and deployed infrastructure services that cover the needs of application services and users. In particular, it gives detailed information on:

- Integrated computing and storage resources – infrastructure services, as well as usage of these resources through development access, and calls for production use of resources and services by user communities (Section 2).
- Deployed operational and resource management tools that facilitate interoperability of the infrastructure, and efficient management of the available computing and storage resources (Section 3).
- Deployed service portfolio management system that supports the service portfolio management process and service discovery (Section 4).
- Deployed Authentication and Authorization Infrastructure (AAI) – VI-SEEM Login service – responsible for management of user access to the resources and services (Section 5).

The deliverable reports the status of the infrastructure and its usage at project month M18. Each section also includes plan for the deployments of infrastructure services until the end of the project.

Conclusions and recommendations

Based on a comprehensive plan for the deployment of the infrastructure and services presented in D3.1 – Infrastructure and services deployment plan, the project so far integrated available computing (HPC, Grid, Cloud) and storage resources, and deployed infrastructure services that fulfill the needs of application services and users. The main conclusions from this process and recommendations for the next period are presented in this deliverable, and can be summarized as follows:

1. Up to this point, VI-SEEM integrated 22,816 CPU-cores, 759,296 GPU-cores, 18,300 Xeon Phi-cores, 3,112 Grid CPU-cores, 14,152 Cloud VM-cores, and 18 PB of storage space. Current level of commitment of these resources to the project are as follows: 12.2% of CPU-cores, 12.9% of GPU-cores, 10% of Xeon Phi-cores, 10.3% of Grid CPU-cores, 4.5% of Cloud VM-cores, and 7.6% of storage space. Most of these resources were integrated by project month M4 (January 2016), while the rest was integrated by project month M9 (June 2016). Although percentage of dedicated resources is slightly below the defined key performance indicator value (in terms of Cloud and storage resources), it is based on real resource demands and will be increased with the 2nd and 3rd call for production use of the resources, as required.
2. Provided infrastructure resources are utilized during: 1st integration phase from project month M5 to M12, 2nd integration phase from M5 to M18, and at present time by accepted applications from the 1st call for production use of resources, which is still ongoing. In total, 1.5 million CPU-hours and 79 million GPU-hours were consumed, mainly by climatology and life science communities. Cloud resources are mainly consumed by digital cultural heritage and life science communities, and in overall 39 Cloud VMs were used. Major use of resources is expected in the ongoing 1st call, where 13.7 million CPU-hours, 3.5 million GPU-hours, and 1.0 million Xeon Phi-hours are allocated.
3. In total, 5 (out of 6 planned) operational and resource management tools were deployed for efficient management of the available computing and storage resources. Monitoring system is partially deployed, and will be fully integrated by the project month M20 (May 2017). All available operational and resource management tools are fully integrated with project's Authentication and Authorization Infrastructure (AAI).
4. VI-SEEM service portfolio management system was deployed to support the service portfolio management process within VI-SEEM. The tool could be used by other infrastructures as well, and therefore is listed in the VI-SEEM Innovation Register [2]. It is fully integrated with project's AAI.
5. VI-SEEM AAI – VI-SEEM Login service – is based on OASIS SAML [3] standard to implement AAI services, and has been established as a central identification service able to identify, authorize and authenticate VI-SEEM users to further allow their access to the offered services. Integration process is straightforward, documented, tested, and implemented in several use-cases.

1. Introduction

Infrastructure and service deployment within the VI-SEEM project is performed both horizontally and vertically through four major tasks of WP3:

- Task 3.1 – Service registry and service level definition,
- Task 3.2 – Implementation of resources and service provisioning,
- Task 3.3 – Operations and resource management of the e-Infrastructure scientific communities,
- Task 3.4 – Authentication, authorization and access management.

The horizontal deployment is done through the provision of a unified underlying e-Infrastructure layer on top of the existing network links among the countries that participate in the project. This layer consists of heterogeneous (HPC, Grid, Cloud, Storage) and geographically distributed resource centers. The vertical deployment covers development of a service catalog/portfolio for the efficient management of project’s services, as well as Authentication and Authorization Infrastructure to enable a uniform management and resource access provisioning. Additionally, a bidirectional task supports site managers to operate integrated services on top of underlying infrastructures, to operate common user services, and support users in their day-to-day usage of the infrastructure.

This deliverable gives an overview of integrated computing (HPC, Grid, Cloud) and storage resources, usage of these resources, developed and deployed services that enable a uniform management and access to these resources, as well as a plan for deployment of new services until the end of the project.

2. Infrastructure resources

The infrastructure of the VI-SEEM project consists of resources of various types - HPC resources - clusters and supercomputers with different hardware architectures, Grid sites, Clouds with possibility to launch virtual machines (VMs) for services and distributed computing, and storage resources with possibility for short and long term storage. The heterogeneous nature of the infrastructure presents management challenges to the project's operational team, but is also an advantage for the users because of its ability to support different types of applications, or different segments of the same application. Modern, state-of-the-art technologies for computing, virtualization and storage, are made available to the developers and end-users. Amount of the resources collected and presented in this section is inline with expected size of the infrastructure given in project's description of the action – Annex I [1].

Particularly, key performance indicators related to e-Infrastructure, and defined in project description, defines size of the infrastructure as 21,500 CPU-cores, 325,000 GPU-cores, and 18,500 Xeon Phi-cores of HPC, 2,900 Grid-cores, 10,500 Cloud VM-cores, and 11 PB of storage space. It also defines dedication to the project as 5-15% of HPC, 10-15% of Grid, and 5% of Cloud, and 10% of storage resources. Up to this point, VI-SEEM integrated 22,816 CPU-cores, 759,296 GPU-cores, 18,300 Xeon Phi-core, 3,112 Grid CPU-cores, 14,152 Cloud VM-cores, and 18 PB of storage space. Dedication to the project needs is the following: 12.2% of CPU-cores, 12.9% of GPU-cores, 10% of Xeon Phi-core, 10.3% of Grid CPU-cores, 4.5% of Cloud VM-cores, and 7.6% of storage space.

Although percentages of dedicated resources does not perfectly match the defined key performance indicators (in terms of Cloud and storage resources), it is based on real resource demands and will be increased with the 2nd and 3rd call for production use of the resources, as required.

2.1. HPC resources

The HPC resources of the project consist of clusters with low-latency interconnection and supercomputers listed in Table 1. Detail description of particular resource was presented in D3.1 – Infrastructure and services deployment plan [4]. In order to keep these descriptions dynamically up-to-date, resource wiki pages have been created as well. In Table 1, links to corresponding wiki pages are listed. Two of these resources, Avitohol from Bulgaria and Leo from Hungary, were at the November 2016 Top500 list [5] of supercomputers (at 388th and 402nd place), while another one, ARIS from Greece, was present at the June 2015 list (at 468th place). Most of the systems are based on CPUs with x86_64 instruction set, some of them equipped with accelerators, but there is a BlueGene/P system, as well as one based on the Cell processor (PS3 cluster IMAN1-Booster/King).

Resource	Country Code	Wiki URL
ARIS	GR	https://wiki.vi-seem.eu/index.php/HPC_Resources/ARIS
Cy-Tera	CY	https://wiki.vi-seem.eu/index.php/HPC_Resources/Cy-Tera
Avitohol	BG	https://wiki.vi-seem.eu/index.php/HPC_Resources/Avitohol
PARADOX	RS	https://wiki.vi-seem.eu/index.php/HPC_Resources/PARADOX
NIIFI SC	HU	https://wiki.vi-seem.eu/index.php/HPC_Resources/NIIFI_SC
Leo	HU	https://wiki.vi-seem.eu/index.php/HPC_Resources/Leo
InfraGRID	RO	https://wiki.vi-seem.eu/index.php/HPC_Resources/InfraGRID
ICAM BlueGene/P	RO	https://wiki.vi-seem.eu/index.php/HPC_Resources/ICAM_BlueGene/P
UPT-HPC	AL	https://wiki.vi-seem.eu/index.php/HPC_Resources/UPT-HPC
MK-03-FINKI	MK	https://wiki.vi-seem.eu/index.php/HPC_Resources/MK-03-FINKI
Armcluster	AR	https://wiki.vi-seem.eu/index.php/HPC_Resources/Armcluster
BA-HPC	EG	https://wiki.vi-seem.eu/index.php/HPC_Resources/BA-HPC
Gamma	JO	https://wiki.vi-seem.eu/index.php/HPC_Resources/Gamma
Zaina	JO	https://wiki.vi-seem.eu/index.php/HPC_Resources/Zaina

Table 1: Wiki pages with detail description of HPC resources.

Summary information about the available resources and their dedication to the project is shown in Table 2. Most of these resources were integrated by project month 4 (January 2016), while the rest of them were integrated by project month 9 (June 2016). In total, 24.6 million CPU-hours, 372.0 million GPU-hours, 16.2 million Xeon Phi-hours, and 5.3 million IBM Cell-hours per year are made available to the scientific communities.

Resource	Country Code	Total			Dedicated (hours per year)		
		CPU-cores	GPU-cores	Phi-cores	CPU-hours	GPU-hours	Phi-hours
ARIS	GR	8,520	253,440	2,196	7,038,000	4,625,280	120,231
Cy-Tera	CY	1,392	16,128	-	1,223,597	16,953,754	-
Avitohol	BG	2,400	-	18,300	2,110,000	-	16,030,800
PARADOX	RS	1,696	108,544	-	1,000,000	47,542,272	-
NIIFI SC	HU	768	-	-	421,882	-	-
Leo	HU	1,344	628,992	-	588,672	275,498,496	-
InfraGRID	RO	456	3,136		798,912	5,494,272	-
ICAM	RO	4,096	-	-	7,176,192	-	-
UPT-HPC	AL	144	-	-	126,144	-	-
FINKI	MK	768	-	-	336,384	-	-
Armcluster	AR	128	-	-	112,128	-	-
BA-HPC	EG	1,968	-	-	3,447,936	-	-
Gamma	JO	8	2,496	-	70,080	21,864,960	-
Zaina	JO	56	-	-	147,168	-	-
	Total	23,744	1,012,73	20,496	24,597,095	371,979,034	16,151,031

Table 2: Available HPC resources and their dedication to the project.

2.2. Cloud resources

List of available Cloud resources is given in Table 3 together with a link to the corresponding wiki page with technical description of particular Cloud center.

Resource	Country Code	Description URL
~Okeanos	GR	https://wiki.vi-seem.eu/index.php/Cloud_Resources/Okeanos
CyI Cloud	CY	https://wiki.vi-seem.eu/index.php/Cloud_Resources/CyI_Cloud_Facility
Avitohol	BG	https://wiki.vi-seem.eu/index.php/Cloud_Resources/Avitohol
InfraGRID Cloud	RO	https://wiki.vi-seem.eu/index.php/Cloud_Resources/InfraGRID_Cloud
UPT-Cloud	AL	https://wiki.vi-seem.eu/index.php/Cloud_Resources/UPT-Cloud
ETFBL-CC01	BA	https://wiki.vi-seem.eu/index.php/Cloud_Resources/ETFBL-CC01
MK-04-FINKI_CLOUD	MK	https://wiki.vi-seem.eu/index.php/Cloud_Resources/MK-04-
MD-RSC	MD	https://wiki.vi-seem.eu/index.php/Cloud_Resources/MD-Cloud
IIAP Cloud	AR	https://wiki.vi-seem.eu/index.php/Cloud_Resources/IIAP_Cloud
IUCC InfinityCloud	IL	https://wiki.vi-seem.eu/index.php/Cloud_Resources/IUCC_InfinityCloud

Table 3: Wiki pages with detail description of Cloud resources.

Summary information about Cloud resources and their dedication to the project is given in Table 4. Cloud resources are integrated sequentially starting from the beginning of the project, and from project month 10 (July 2016) all 468 dedicated virtual machines are made available to the scientific communities.

Resource	Country Code	Total VMs	Dedicated VMs	Dedicated VM-hours
~Okeanos	GR	10,000	200	1,752,000
CyI Cloud	CY	176	18	157,680
Avitohol	BG	2,400	120	1,051,200
InfraGRID Cloud	RO	400	46	402,960
UPT-Cloud	AL	12	6	52,560
ETFBL-CC01	BA	60	13	113,880
MK-04-FINKI_CLOUD	MK	436	24	210,240
MD-Cloud	MD	12	3	26,280
IIAP Cloud	AR	96	10	87,600
IUCC InfinityCloud	IL	560	28	245,280
	Total	**14,152**	**468**	**4,099,680**

Table 4: Available Cloud resources and their dedication to the project.

2.3. Grid resources

The Grid resources within the project are provided mostly from smaller clusters that have already installed Grid middleware and also share their resources within the European Grid Infrastructure [6]. Some of the clusters that are not certified at the level of the EGI had to be incorporated by using modified BDII services. Links to wiki pages with detail description of Grid resources are given in Table 5.

Resource	Country Code	Description URL
Hellas Grid	GR	https://wiki.vi-seem.eu/index.php/Grid_Resources/Hellas_Grid
BG01-IPP	BG	https://wiki.vi-seem.eu/index.php/Grid_Resources/BG01-IPP
AEGIS01-IPB-SCL	RS	https://wiki.vi-seem.eu/index.php/Grid_Resources:AEGIS01-IPB-SCL
MK-03-FINKI	MK	https://wiki.vi-seem.eu/index.php/Grid_Resources/MK-03-FINKI
MREN01CIS	ME	https://wiki.vi-seem.eu/index.php/Grid_Resources/MREN01CIS
MD-GRID	MD	https://wiki.vi-seem.eu/index.php/Grid_Resources/MD-GRID
ArmCluster	AR	https://wiki.vi-seem.eu/index.php/Grid_Resources/ArmCluster
GE-01-GRENA	GE	https://wiki.vi-seem.eu/index.php/Grid_Resources/GE-01-GRENA

Table 5: Wiki pages with detail description of Grid resources.

Summary information about the available Grid resources and their dedication to the project is shown in Table 6. These resources were integrated by project month 5 (February 2016). In total, 299 dedicated Grid CPU-cores are provided to scientific communities.

Resource	Country Code	Total CPU-cores	Dedicated CPU-cores	Dedicated CPU-hours
Hellas Grid	GR	864	43	370,000
BG01-IPP	BG	640	120	1,051,200
AEGIS01-IPB-SCL	RS	704	50	438,000
MK-03-FINKI	MK	768	38	336,384
MREN01CIS	ME	32	16	140,160
MD-GRID	MD	40	12	105,120
GE-01-GRENA	GE	64	20	175,200
Total		**3,112**	**299**	**2,616,064**

Table 6: Available Grid resources and their dedication to the project.

Project dedicated Virtual Organization (vo.vi-seem.eu) has been deployed at the IPB, and its web interface, VOMS-admin is available at [7]. Beside of the virtual organization management service, a set of Grid core services (VOMS, BDII, LFC, WMS, LB, PX) are deployed at IPB, IICT-BAS, and GRNET. Functional description of Grid core services from a user's point of view is given in gLite User Guide [8].

2.4. Storage resources

The storage resources could be considered in relation to the respective type of resource, e.g., the HPC storage should be available to HPC jobs and the Grid storage should be available to Grid jobs. Such storages are described on wiki pages listed in Table 7. The storage resources, described here, are those that fall outside of such types of usage and are provided for long-term or other special types of data storage. Table 7 lists links to detailed description of particular storage resources integrated into VI-SEEM infrastructure.

Resource	Country Code	Description URL
ARIS	GR	https://wiki.vi-seem.eu/index.php/Storage_Resources/ARIS
Cyprus	CY	https://wiki.vi-seem.eu/index.php/Storage_Resources/ONYX,_Cy-Tera
Avitohol	BG	https://wiki.vi-seem.eu/index.php/Storage_Resources/Avitohol
PARADOX	RS	https://wiki.vi-seem.eu/index.php/Storage_Resources/PARADOX
NIIFI HSM	HU	https://wiki.vi-seem.eu/index.php/Storage_Resources/NIIFI_HSM_Service
NIIFI iSCSI	HU	https://wiki.vi-seem.eu/index.php/Storage_Resources/NIIFI_iSCSI_Service
UVT HPC GPFS	RO	https://wiki.vi-seem.eu/index.php/Storage_Resources/UVT_HPC_GPFS
ETFBL-CC01	BA	https://wiki.vi-seem.eu/index.php/Storage_Resources/ETFBL-CS01
MK-04-FINKI	MK	https://wiki.vi-seem.eu/index.php/Storage_Resources/MK-04-FINKI_CLOUD
RENAMstor	MD	https://wiki.vi-seem.eu/index.php/Storage_Resources/RENAMstor
IIAP Storage	AR	https://wiki.vi-seem.eu/index.php/Storage_Resources/IIAP
BA-IA	EG	https://wiki.vi-seem.eu/index.php/Storage_Resources/BA-IA
IUCC Storage	IL	https://wiki.vi-seem.eu/index.php/Storage_Resources/IUCC-InfinityCloud-

Table 7: Wiki pages with detail description of Storage resources.

Summary of provided storage space is given in Table 8. Storage resources are integrated sequentially starting from the beginning of the project, and from project month 10 (July 2016) full capacity, 330 TB of disk and 510 TB of tape storage space, is made available to the scientific communities.

Resource	Country Code	Total		Dedicated	
		Disk [TB]	Tape [TB]	Disk [TB]	Tape [TB]
ARIS	GR	1,000	3,000	50	210
Cyprus	CY	500	-	100	-
Avitohol	BG	96	-	5	-
PARADOX	RS	96	-	5	-
NIIFI HSM	HU	87	6,800	3	300
NIIFI iSCSI	HU	1,000	-	50	-
UVT HPC GPFS	RO	50	-	5	-
ETFBL-CC01	BA	1		0.5	-
MK-04-FINKI	MK	36	-	2	-
RENAMstor	MD	7	-	1.75	-
IIAP Storage	AR	8	-	3	-

BA-IA	EG	5,200	-	100	
IUCC Storage	IL	40	-	5	-
	Total	**8,121**	**9,800**	**330.2**	**510**

Table 8: Available Storage resources and their dedication to the project.

2.5. Usage of resources

Provided infrastructure resources and services overall are so far mainly used through the development access, as well as through calls for production use of resources and services. The VI-SEEM development access facilitates the development and integration of services by the selected collaborating user communities. In this process, applications are given access to the infrastructure and necessary computational resources for a six-month period, during which application developers are expected to develop and integrate relevant services. The calls for production use of resources and services target specific communities and research groups that have already began development of their projects. These calls are intended for mature projects, which require significant resources and services to realize their workplans. Therefore, significant utilization of VI-SEEM resources comes from the calls for production use of resources and services, and an order-of-magnitude smaller utilization comes from the development access. Main categories of access provided to the VI-SEEM resources and services are defined in D6.1 – Framework for VRE resource and service provision [9].

The VI-SEEM development access is divided into three integration phases, and three separate calls for production use of resources and services will be issued during the project lifetime. Up to now, two integration phases and one call have been issued. Thus, the provided infrastructure resources are utilized during:

- 1st integration phase from project month M5 to M12,
- 2nd integration phase from M5 to M18,
- 1st call for production use of resource from M18 to M30.

List of applications supported during the 1st and 2nd integration phases of development access is shown in Table 9 and Table 10. Table 9 lists applications whose deployment required HPC infrastructure. For each application, dedicated HPC resource, and consumed amount of resource needed for development purposes are given. In total, 1.5 million CPU-hours and 93 thousand of GPU-hours (in CPU hours of the GPU hosts) are consumed mainly by climatology and life science communities.

Application Acronym	**Country Code**	**VRC**	**Integration Phase**	**HPC Resource**	**Resources used**	
					CPU-hours	**GPU-hours (in CPU hours of the GPU hosts)**
EMAC	CY	CR	1st	Cy-Tera	58,480	1,032
WRF-ARW	CY	CR	1st	Cy-Tera	30,500	-
VINE	GE	CR	1st	ARIS	1,163	-
WRF-Chem	GR	CR	1st	ARIS	452,200	-
DREAMCLIMATE	RS	CR	1st	PARADOX	223,000	-
ACIQLife	BG	CR	1st	Avitohol	657,856	-

MD-Sim	GR	LS	1st	ARIS	-	79,000
PSOMI	ME	LS	1st	FINKI	1,400	-
SIPD	GR	CR	2nd	ARIS	20,000	13,000
TVRegCM	BG	CR	2nd	Avitohol	87,104	-
DRS-ACS	MK	CR	2nd	FINKI	500	-
CNCADD	MK	LS	2nd	FINKI	500	-
				Total	**1,532,703**	**93,032**

Table 9: Used HPC resources during the development access.

During the development access Cloud resources are mainly consumed by digital cultural heritage and life science communities. In total, 36 virtual machines are used for these purposes, and Table 10 gives usage per application at particular Cloud center.

Application Acronym	Country Code	Virtual Community	Integration Phase	Cloud Resource	Resources used
					Used VMs
BVL	RO	DCH	1st	CyI-Cloud	4
BVL	RO	DCH	1st	InfraGRID	10
MD-Sim	GR	LS	1st	~OKEANOS	4
Manuscript	IL	DCH	2nd	CyI-Cloud	4
Dioptra	CY	DCH	2nd	CyI-Cloud	4
3DINV	GR	DCH	2nd	~OKEANOS	4
AutoGR	GR	DCH	2nd	~OKEANOS	4
DICOMNetwork	MD	LS	2nd	FINKI_CLOUD	1
SEMaCD	RS	LS	2nd	PARADOX	1
				Total	**36**

Table 10: Used Cloud resources during the development access.

Overall, 21 applications are accepted in the 1st call for production use of resource and services. Majority (20 out of 21) of the applications require access to HPC and Cloud resources. One of them (VINE application from Georgia) requires Grid infrastructure resources, and has been granted 50,000 Grid CPU-hours. One application (AUTH_WRF371M_EUR0.44 application from Greece) is granted access to the VI-SEEM repository service only, and therefore is missing from Table 11 and Table 12. All applications requested, and were provided with access to storage resources directly, via access to HPC/Grid resource, or indirectly, via deployed data services (simple storage service, repository service, archival service, etc.).

Access to HPC resources is provided to 14 applications, and access to Cloud resources to 6 applications. Table 11 lists applications from the 1st call for production use of resources and services, and their distribution over available VI-SEEM HPC resources. For each application, amount of allocated resources is given. In total, 13.7 million CPU-hours, 3.5 million GPU-hours, and 1.0 million Xeon Phi-hours are allocated.

Application Acronym	Country Code	Virtual Community	HPC Resource	Resources allocated		
				CPU-hours	GPU-hours	Phi-hours
VINE	GE	CR	ARIS	38,000	-	-
MECCA_ECHAM_GPU	CY	CR	Cy-Tera	200,000	40,000	-
TVRegCM	BG	CR	Avitohol	960,000	-	-
ACIQLife	BG	CR	Avitohol	650,000	-	1,000,000
CH-CBIR	BA	DCH	Leo	-	3,000,000	-
MULTIDRUG	BG	LS	Avitohol	500,000	-	-
D3R	GR	LS	ARIS	-	1,500	-
MolSurf	BG	LS	Leo	450,000	550,000	-
DRS-ACS	MK	CR	NIIFI SC	500,000	-	-
DRS-ACS	MK	CR	InfraGRID	500,000	-	-
CNCADD	MK	LS	BA-HPC	1,500,000	-	-
DREAMCLIMATE	RS	CR	PARADOX	1,000,000	-	-
DACEM1	GR	CR	ARIS	3,000,000	-	-
WRF-ARW	CY	CR	Cy-Tera	422,000	-	-
DACEM2	GR	CR	ARIS	4,000,000	-	-
			Total	**13,720,000**	**3,591,500**	**1,000,000**

Table 11: Allocated HPC resources in the 1st call for production use of resource and services.

Table 12 lists applications from the 1st call for production use of resources and services, and their distribution over available VI-SEEM Cloud resources. In overall, 39 virtual machines are allocated for this purpose.

Application Acronym	Country Code	Virtual Community	Cloud Resource	Resources allocated
				Used VMs
PSOMI	ME	LS	MK-04-FINKI_CLOUD	1
DIOPTRA	CY	DCH	CyI-Cloud	3
Continuous_LST	IL	CR	IUCC InfinityCloud	5
DATANMNH	BG	DCH	Avitohol	10
MC4CH	BG	DCH	MK-04-FINKI_CLOUD	10
BVL	RO	DCH	InfraGRID	10
			Total	**39**

Table 12: Allocated Cloud resources in the 1st call for production use of resource and services.

3. Operations and resource management

Efficient management of the available computing and storage resources, as well as interoperability of the infrastructure is achieved by a set of operational tools. Static technical information, such as name, geographical location, contact and downtime information, list of service-endpoints provided by a particular resource center within the infrastructure etc., is manually entered and made available through VI-SEEM GOCDB database. Based on this information, project monitoring system is able to automatically trigger execution of monitoring service probes, and to enable efficient access to results of the probes via customized monitoring web portal. Using standard metrics VI-SEEM accounting system accumulates and reports utilization of the different types of resources. User support and service-related problems are resolved mainly through the helpdesk system, but via technical mailing list as well. VI-SEEM source code repository contains all codes developed within the project, while the technical wiki collects technical documentation, know-hows, best practices, guidelines, etc.

Deployed VI-SEEM operational and resource management tools, responsible partners, and frontend URLs are listed in Table 13. Technical details on deployment of particular tool are given in this section.

System	Responsible Partner	Country Code	URL
GOCDB	UKIM	MK	https://gocdb.vi-seem.eu/portal
Monitoring	GRNET/UNI BL	GR/BA	https://mon.vi-seem.eu/
Accounting	IICT-BAS	BG	https://accounting.vi-seem.eu/
Helpdesk	UNI BL/UOM	BA/ME	https://support.vi-seem.eu/
Source code repository	UNI BL/UOM	BA/ME	https://code.vi-seem.eu/
Technical wiki	CYI	CY	https://wiki.vi-seem.eu/

Table 13: List of deployed operational and resource management tools.

3.1. GOCDB

To maintain compatibility with similar e-Infrastructure projects, such as EGI-Engage [6] or EUDAT2020 [10], the GOCDB was chosen as the configuration management database to be used within the VI-SEEM project. GOCDB is a Configuration Management Database (CMDB) for recording and managing assets in federated environments. It defines a number of topology objects including admin-domains, sites, services, service-groups, service-endpoints, service-downtimes, users and roles. The tool provides a web portal for editing information and a REST style programmatic interface (PI) for querying data in XML. Relationships between different objects are defined using a well constrained relational schema that closely resembles a subset of the GLUE 22 information model. A comprehensive role-based permissions model controls user permissions. Remarkably, the project can self-manage their users: users make requests for roles over target objects and users that already hold the necessary role(s) can accept or reject those role requests.

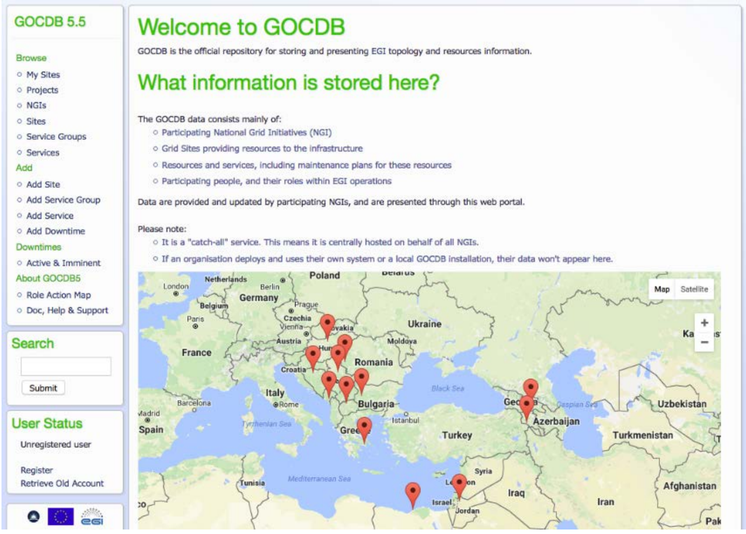

Figure 1: Front page of the VI-SEEM GOCDB system.

An instance of the GOCDB was deployed by the partner UKIM for the needs of the project. The instance is customized to reflect the actual needs and objects of the VI-SEEM project infrastructure and services. VI-SEEM is defined at the level of project and at the level of NGI in the deployed instance. This allows the sites to be declared as part of the NGI, which in this case is the project itself. Project partners that are resource providers are defined as GOCDB sites. Since all the partners are providers, there are currently 16 sites in the database, one per each partner. Each partner is responsible for maintaining the information about its site.

GOCDB user guide is documented at https://wiki.vi-seem.eu/index.php/VI-SEEM_GOCDB, and usage of its programmatic interface is described at https://wiki.vi-seem.eu/index.php/VI-SEEM_GOCDB_PI_Technical_Documentation.

Following the requirements for the project, there are 15 defined service types, used to group the different services. Currently, the GOCDB database contains 73 services, out of which 69 are in productions phase. There are 21 GOCDB registered users in the database.

3.2. Monitoring

VI-SEEM project has a very complex structure, from infrastructure that spans several paradigms from classical supercomputers to grid and cloud computing, services that can be either VRE-specific or shared among them, to applications that can and will cross the borders of the mentioned domains. Possessing a high quality monitoring system is a must in such circumstances. As the monitored infrastructure is complex and heterogeneous, the monitoring system was required to be scalable and modular. Drawing upon previous experience in similar environments, the ARGO Service Monitoring system [11] was chosen for the basis. One of the most important characteristics of ARGO, from the project standpoint, is its modularity that enables to effectively adapt it to ever-changing needs of project infrastructure users and administrators.

ARGO architecture is based on layered approach with lowest layer, the Monitoring Engine, being based on Nagios [12]. This layer is used for service endpoints monitoring and uses probes and add-ons in order to provide required raw monitoring data. Availability and Reliability calculations are performed by the Analytics Engine layer that consumes monitoring data through the use of Connectors. Web API is the top layer and it enables for simpler integration with external tools and systems (including reports, alarms, etc). This architecture is decoupled in nature and there is a considerable flexibility in extending or adapting each of the components. This layered architecture of ARGO system is illustrated in Figure 2.

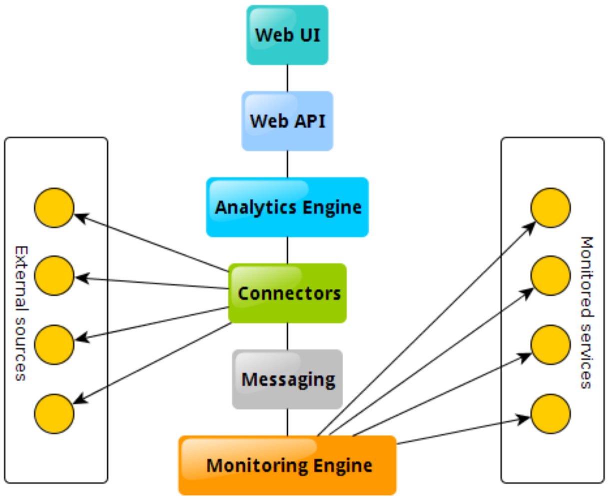

Figure 2: Modular architecture of ARGO service monitoring system.

Within the VI-SEEM project, ARGO instance is provided by GRNET in a centralized configuration which was deemed an appropriate solution for current project needs. Web UI is hosted at ETFBL-CC01 cloud computing infrastructure at https://mon.vi-seem.eu/.

Due to specific project needs, a custom Web UI (illustrated in Figure 3) was developed to enable for simpler and more efficient access to data by both project personnel as well as end-users. The Web UI was designed as single page web application and uses Python Flask framework [13] based application as the back-end and React [14] Javascript based front-end with custom made components. The Web UI provides three main sections:

- **Dashboard** – overview page presenting summary monitoring information for each monitored community. It contains tabular data on latest available service statuses, availability and reliability figures as well as graphical representation for past three months.
- **Current Service Status** – provides data on current status of monitored services as well as a timeline with service status data for services currently in non-OK status for the past three days.
- **Availability and Reliability** – provides “drill-down” approach to detailed data. It enables users to either browse starting with the latest data available or to filter the

data by time period and granularity level for calculated values. It also provides uniform UI for accessing both services’ and service components’ monitoring data.

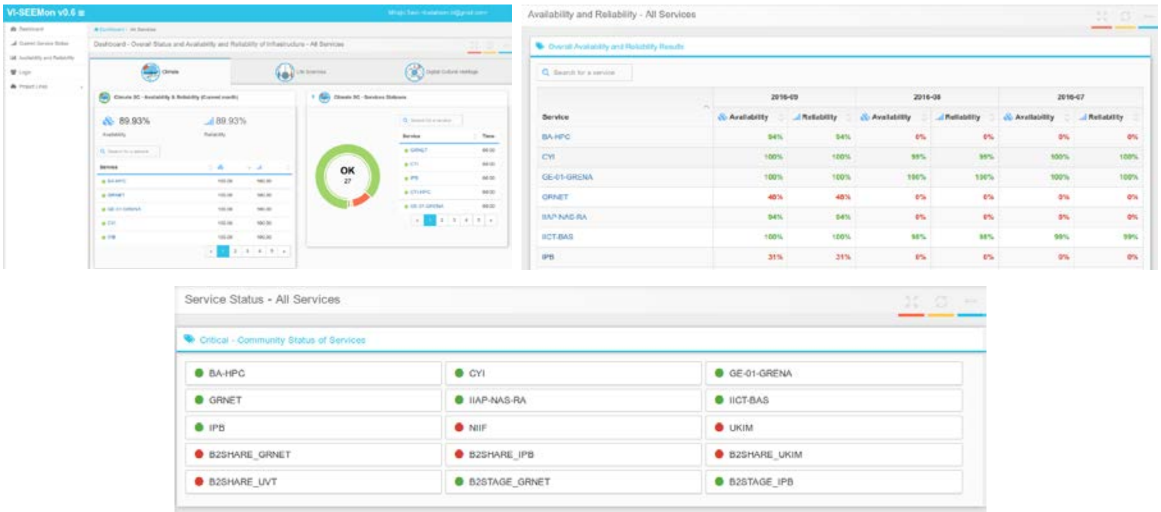

Figure 3: Main sections of the monitoring Web UI.

Overall monitoring system architecture is illustrated in Figure 4. It has been integrated with project wide AAI and provides both unauthenticated, but limited, access in Guest mode as well as authenticated access for registered users. Integration has been documented in the project wiki pages (https://wiki.vi-seem.eu/index.php/VI-SEEM_AAI_mod_auth_mellon_Integration). The system pairs the data from GOCDB, ARGO and AAI and provides full access to otherwise restricted data for registered users having elevated access level in GOCDB (such as system administrators, etc.).

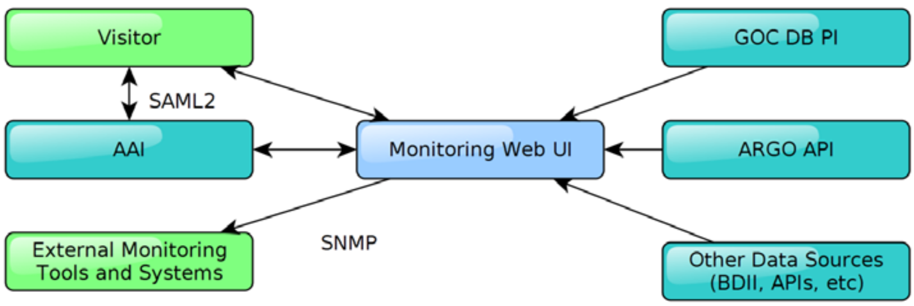

Figure 4: Overall monitoring system architecture.

As none of the systems on the internet exist in isolation, it is important to provide standardized access to data for external tools and existing monitoring systems. Although there is a plethora of “standards”, there are a very small number of actual standards defined by relevant standard bodies. We have chosen to implement access to a subset of monitoring data via SNMP protocol as it is by far the most widely supported standard in network monitoring systems. This work is currently under development.

3.3. Accounting

The accounting system accumulates and reports utilization of the different types of resources in the VI-SEEM infrastructure using standard metrics. The information that is relevant to the VI-SEEM resources usage is gathered and organized in one central database. Accounting reports are available via web interface at https://accounting.vi-seem.eu/.

Architecture of the accounting system is illustrated in Figure 5. Each resource manager node that holds relevant information publishes them through an accounting client. Within the accounting framework this is a client-type service. The accounting system provides an API and publisher clients for this API written in Python. If needed each resource administrator can extend or write a new publisher based on their specific requirements and policies. For ease of deployment and simplicity the dependencies of the Python clients are kept to a minimum.

The accounting publishers communicate with the server side via REST API. They can upload whole logs with accounting data or use API directly. The server-side is implemented in Python using Flask web framework and offers both API and web UI. Each resource type has a standardized accounting format that is described under the API section in the web UI.

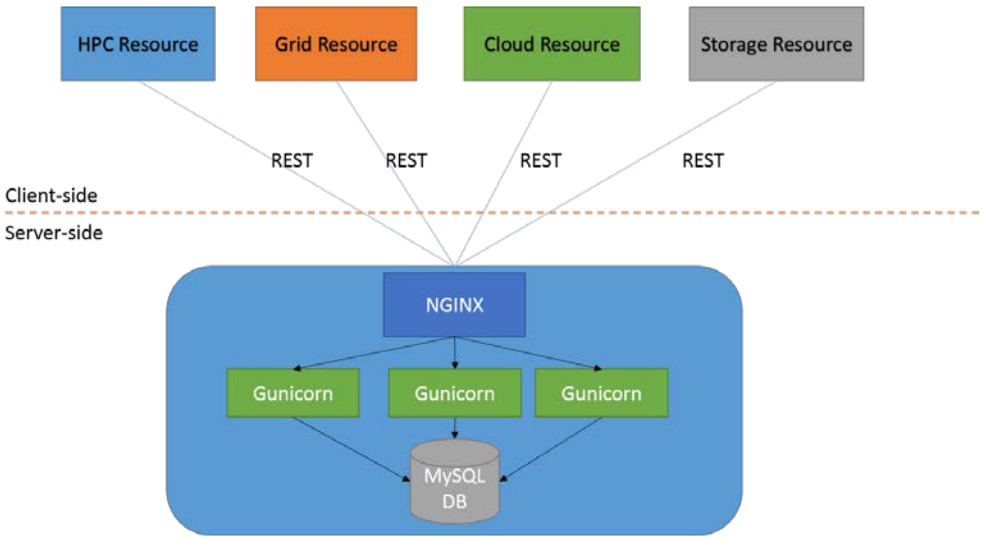

Figure 5: Architecture of the accounting system.

The access to the accounting reports is available via web interface (shown in Figure 6) that supports the VI-SEEM central login via SAML. The user can access the information grouped in table by date, year, country, resource name, virtual research community and application in rows and columns. Each of the resource types will have separate view – compute data for HPC and Grid computing data, cloud data for cloud accounting and storage data for storage accounting. Each generated report is presented in table and simple charts grouped by the user choice.

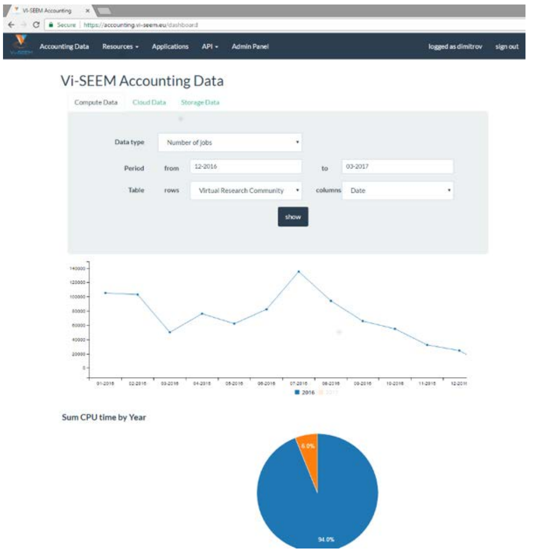

Figure 6: Dashboard view of the accounting system.

The hardware requirements for the client side are Python (2.6+, 3.3+). The server-side is installed on machine with CentOS7 with installed virtualenv for virtual environment for Python, Gunicorn as light and fast Python WSGI application server, Nginx as a reverse-proxy in-front providing a high-performance handling of the user requests and MySQL as database server.

3.4. Helpdesk

After performing analysis of several available open source helpdesk/issue-tracking system, we have chosen the osTicket [15] as the solution for the project needs. This is an application written in PHP that utilizes MySQL (or compatible) database for data storage. Interaction with the system is possible through web interface, JSON/XML based HTTP API and email.

VI-SEEM helpdesk system is hosted at ETFBL-CC01 cloud computing infrastructure and is available at https://support.vi-seem.eu/ and its front page is shown in Figure 7. Its use is documented for both end users and agents at the project wiki pages (https://wiki.vi-seem.eu/index.php/Support_Ticket_System).

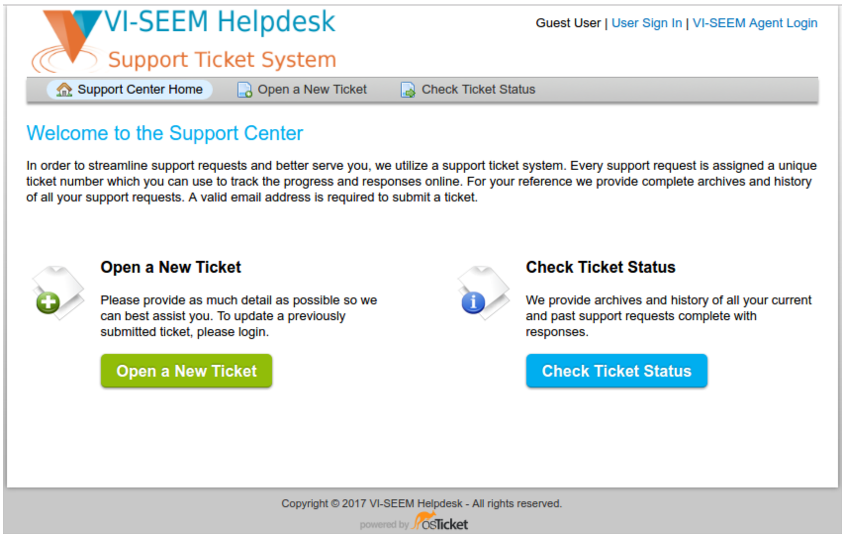

Figure 7: Front page of the Helpdesk system.

There are two types of users within the helpdesk:

- End user - is a user from one of the three VRE communities who is using services.
- Agent - is a member of institution which provides support for the end users.

End user can use VI-SEEM helpdesk as a guest user or as a registered user. In both case one has to use a real email account to be able to communicate with support staff.

Basic workflow of using the helpdesk system is as follows:

1. End user has some problem (ticket status: open),
2. End user opens a ticket (ticket status: open),
3. System automatically sends confirmation mail (ticket status: open),
4. Agent or supervisor assigns ticket to a specific agent (ticket status: assigned),
5. Agent works on the problem (ticket status: in progress),
6. Agent submits solution to the ticket (ticket status: resolved),
7. Agent closes the ticket (ticket status: closed).

Basic workflow is a straight line, however, it is possible that ticket has to be reassigned, reopened or that agent need additional information about issue. In those cases, ticket life cycle has loops. During the ticket life, end user can check status of the ticket at any point in time. Agent's view of one ticket's history is illustrated in Figure 8.

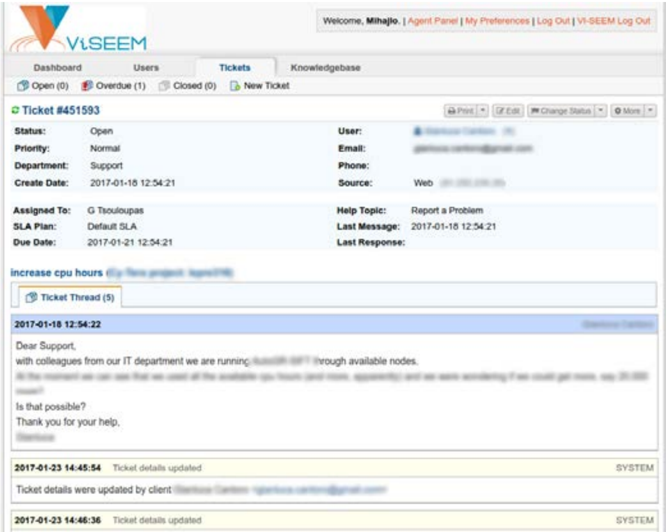

Figure 8: Agent's view of a ticket history.

OsTicket supports fine grained access control and elaborate categorizations, enabling quite complex workflows envisioned for the project. When it comes to end-user facing side, it allows for many possibilities from allowing guest users to open tickets and follow their status to acting as a fully closed invitation-only system.

When it comes to internal support organization, the system allows for both horizontal and vertical segmentation with possible domain border crossing entities when needed. Currently, help topics are organized as follows:

- Feedback – low priority topic for generic feedback,
- Report a Problem; Report a Problem / Access Issue – higher priority for generic problem reporting,
- WP3 – Infrastructure; Service Registry – generic WP3 issues,
- WP4 – Data Platforms – generic WP4 issues,
- WP5 – Applications; Domain Specific Services – issues related to whether to concrete applications or services dependent on,
- Issues with resources – this is a set of topics, one for each resource center – created tickets are assigned to designated support person at the resource center.

When it comes to integration with project infrastructure, there are three main domains of interest: AAI, service registry and separate support tools already present within the project infrastructure.

Integration with project AAI was implemented by using the SAML 2.0 protocol and mod_auth_mellon Apache module. The manner the integration was executed (by using external web server authentication support in osTicket) and separating the access URLs for local and AAI users enable for operation of both kind of accounts in parallel. Support agents' accounts are integrated with the project infrastructure by using static mapping (originating from GOCDB). This procedure was fully documented in the project wiki pages (https://wiki.vi-seem.eu/index.php/VI-SEEM_AAI_mod_auth_mellon_Integration).

Integration with service registry/catalogue was used to create topics for "Issues with resources" described previously. Current implementation is static in nature and was aimed at reducing of the potentially overwhelming number of support topics for end-users.

Integration with other related systems (monitoring systems) is currently under development.

Primary instance of the helpdesk is hosted at University of Banja Luka Faculty of Electrical Engineering. It is planned to install the service in an active-passive fail-over configuration with secondary instance hosted at the University of Montenegro.

3.5. Source code repository

VI-SEEM source code repository (illustrated in Figure 9) is one of the critical services of the project as it is to host all the code of applications and services developed within the project. As the project has quite complex structure with many partners and several work packages closely related, it was decided that distributed version control system was the best fit for project needs. Chosen solution had to offer support for distributed, non-linear development with strong support for data integrity and traceability of versions.

After initial testing of several available source code repository systems, we have chosen GitLab [16] for the project needs. This system was created as a web management tool for projects that use git for version control system and mostly mirrors concepts and features available on a highly popular hosting site github.org. It is built primarily as a Ruby application although it integrates other various open source tools and libraries. Aside from source code versioning control, GitLab also provides issue tracking and Wiki. Although there are other services within the project that cover those areas, code development can be very specific and is often tightly coupled with code related issue tracking and some form of highly specific knowledge base. GitLab also has support for file browser, code review, merging, testing as well as a simple interface than enables easy import of existing projects from other hosting solutions like GitHub, Bitbucket, Google Code or any other repository with available git access. Support for migrating from SVN based operations to git is provided by svn2git utility.

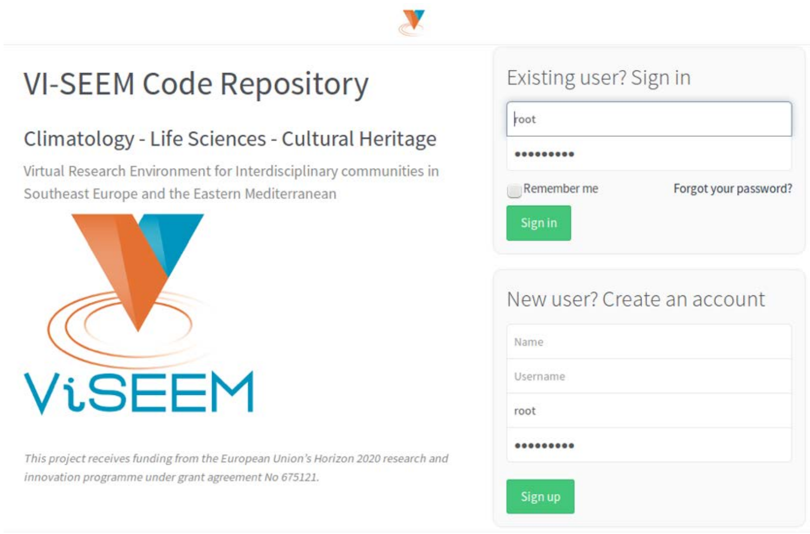

Figure 9: VI-SEEM source code repository front page.

Integration with project infrastructure and services is based mostly on allowing users to access the system by using project Authentication and Authorization Infrastructure and publish the projects’ metadata to VI-SEEM Data Discovery Service. Prior to full integration with project AAI, users were allowed to self-register, but valid email address from the set of predefined project partners' institutional email domains was required.

The system has been fully integrated in the project AAI and has full support for mapping local existing users to externally authenticated users (such as AAI ones) by linking the user accounts with the same registered email addresses. The procedure of integration with AAI has been fully documented at the project Wiki pages (https://wiki.vi-seem.eu/index.php/VI-SEEM_AAI_GitLab_Integration).

In order to provide metadata on non-private projects available at the source code repository, we have developed a metadata publisher (gitlab2ckan). This tool collects list of projects, branches, tags and description tags from GitLab instance and publishes it to VI-SEEM Data Discovery Service by using CKAN API.

The use of the system was also documented with examples covering common use cases and this documentation is also available at the project Wiki pages (https://wiki.vi-seem.eu/index.php/Support_Ticket_System).

The primary system instance is hosted at ETFBL-CC01 cloud computing infrastructure and is available at https://code.vi-seem.eu/. Fail-over service in an active-passive fail-over configuration with secondary instance hosted at the University of Montenegro has also been implemented. Access is possible from either web interface or from command line by utilizing git tools.

3.6. Technical wiki

VI-SEEM technical wiki (http://wiki.vi-seem.eu/), which is hosted at the Cyprus Institute and is based on the MediaWiki platform, has been integrated with the VI-SEEM Authentication and Authorization Infrastructure (AAI). This integration provides single sign-on capabilities to the VI-SEEM wiki pages and therefore wiki users can now be authenticated through the central Vi-SEEM AAI without having special credentials for the wiki pages.

To perform the integration Shibboleth Apache module has been installed as well as the Shibboleth MediaWiki extension. The Apache module was needed in order to enable the web server hosting the wiki pages to exchange authentication and authorization data using the Security Assertion Markup Language (SAML) standard. In particular, the authentication and authorization data are exchanged between the identity provider, i.e. the VI-SEEM AAI and the service provider, i.e. the Vi-SEEM wiki web server. The Shibboleth MediaWiki extension is used to redirect the users to the VI-SEEM AAI, in order to be authenticated, and maps the SAML attributes of the authenticated user to wiki attributes.

The procedure for a user to login to the VI-SEEM wiki page is fairly simple. While the user is located at the main page of the wiki, a “VI-SEEM Login” link is visible in the top right corner of the page as shown in Figure 10.

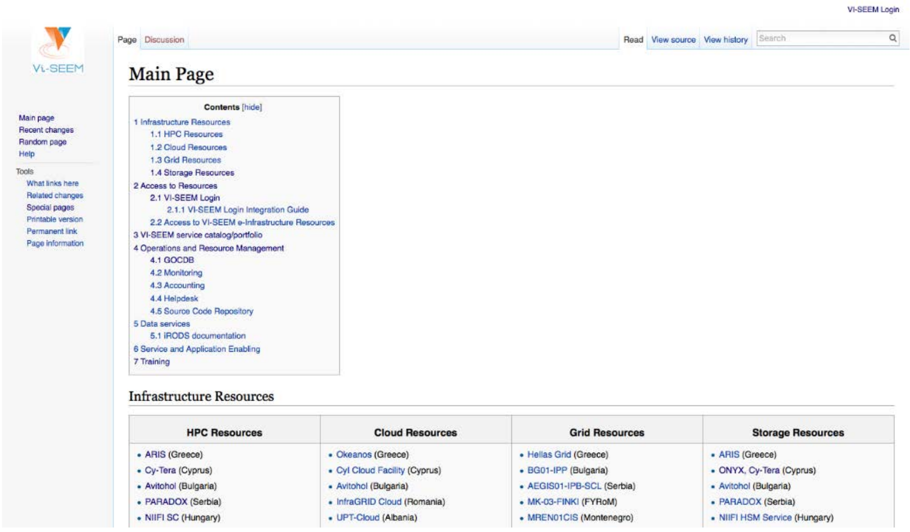

Figure 10: Main page of VI-SEEM technical wiki.

Once the user clicks on the "VI-SEEM Login" link, he will be redirected to the VI-SEEM AAI site, where the list of academic institutions as well as social networks is listed. After authentication account selection, and if the authentication has been successful, he will be redirected back to the VI-SEEM wiki page logged in with the selected account. The wiki username of each user is the email address associated with the account used for the authentication. In order to terminate the session, the user has to click on the "Logout" link located on the top right corner of the page.

VI-SEEM technical wiki has 269 pages, 173 uploaded files, approximately 95,000 of page views, and 42 registered users.

4. Service catalog/portfolio and management

The VI-SEEM service portfolio management system has been developed to support the service portfolio management process within VI-SEEM as well as being usable for other infrastructures if required. The main requirements for the creation of this tool have been collected from the service management process design within VI-SEEM work packages WP3 (infrastructure services), WP4 (storage services) and WP5 (application level services). The service management system has been designed to be compatible with the FitSM [17] service portfolio management. Requirements gathered in the context of EUDAT2020 project have also been considered for compatibility and completeness. Following is a list of the main functionalities of the service portfolio management system.

- There is a set of different roles that have access to the functionalities of the tool. These are:
    - The **potential customers** or **end users** of the services that are listed in the **service catalog**. Such users are able to see the list and details of the services that are currently into production or beta stages and are for offer to them. Such users are also able to order or use the services via the service catalog, interact with the helpdesk or the dedicated support channel for that service and see features and use cases of each service.
    - The **service managers** within the VI-SEEM environment. Such users are able to see all the details for the services that are in the **service portfolio** which contains a superset of the service and the information found in the service catalog.
    - The **service owners** that are the persons responsible for each service listed in the service portfolio. Members with this role have the full responsibility of the content that is provided within the service catalog and portfolio for the services under their responsibility.
- The service catalog contains only information about services that are to be provided to the potential customers and end users of the services. A subset of all the service information is provided for a subset of the services that are registered in the service portfolio.
- Each service can have multiple versions in the service portfolio. Each version can have a different readiness status i.e. "concept", "under development", "beta", "production", "retired", etc.
- The service owners and the service managers are able to state which combinations of service and service versions should appear in the service catalogue.
- The service can be either customer facing or internal to the organization.
- The dependencies between services are depicted in the service portfolio in order to facilitate the SLA management process better.
- The components that are required for deploying each service are detailed in the service portfolio providing information needed for the operations staff to deploy such services.
- The service portfolio is accessible via a RESTful API to accommodate 3rd party application building and different views of the service catalog depending on the needs of the organization.
- The system is fully integrated with the VI-SEEM AAI infrastructure.

- The service portfolio/catalog has web UI to accommodate the main service catalog view for the VI-SEEM project as well as to facilitate adding and editing of information from service managers and service owners. The web UI has three main components:
  - The service catalog view, with open access.
  - The service portfolio view, with restricted access using the VI-SEEM AAI.
  - The WRITE UI, with restricted access for editing the service portfolio/catalog.

The design of the system was detailed in the D3.1 – Infrastructure and services deployment plan. The technical details of the initial version of the system were detailed in the D3.2 – Service registry, operational and service level monitoring [18].

4.1. Newly developed elements

Since the system has been developed and put into the production, several improvements were made to increase its usability. The main improvement is the development of the custom write UI, that replaces the original write UI build by customizing the Django admin interface. The new write UI (illustrated in Figure 11) offers a more intuitive interface to add and modify service and their elements.

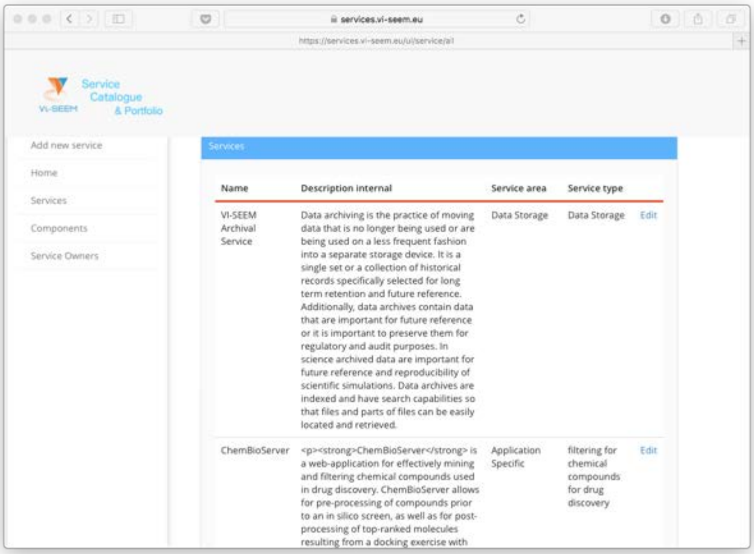

Figure 11: New WRITE UI of Service catalog/portfolio.

For each service, the write UI enables entering and editing of all necessary information (illustrated in Figure 12), logically grouped in 5 categories:

- Service, containing basic information, name, description (both internal and external), service area and type, procedure for requesting the service, funding, value for the customers, risks and competition, contact information and information about the service owner,

- Versions, enabling the system to track different (past, present and future) versions of the service,
- Internal dependencies,
- External dependencies,
- Users/customers roles.

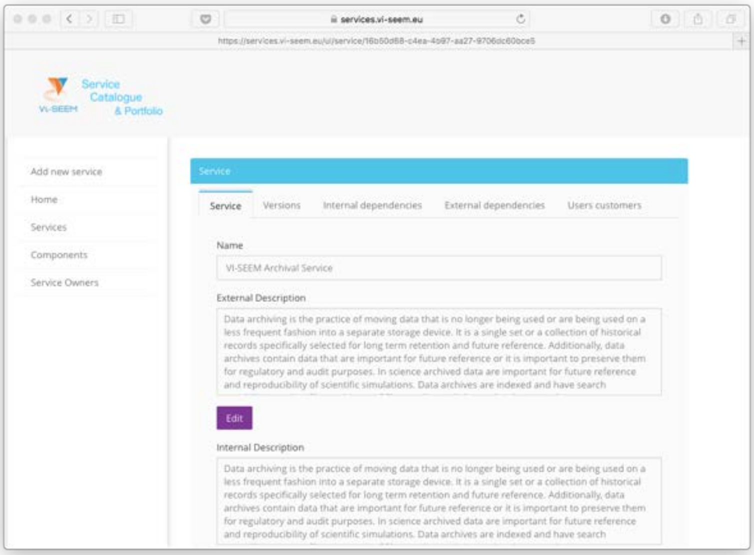

Figure 12: Editing the service information in the service catalog/portfolio.

Each service can be based on one or more service components. The components can be entered and edited through the write UI by specifying their basic info (name and description), their implementations and implementation details, as it is show in Figure 13.

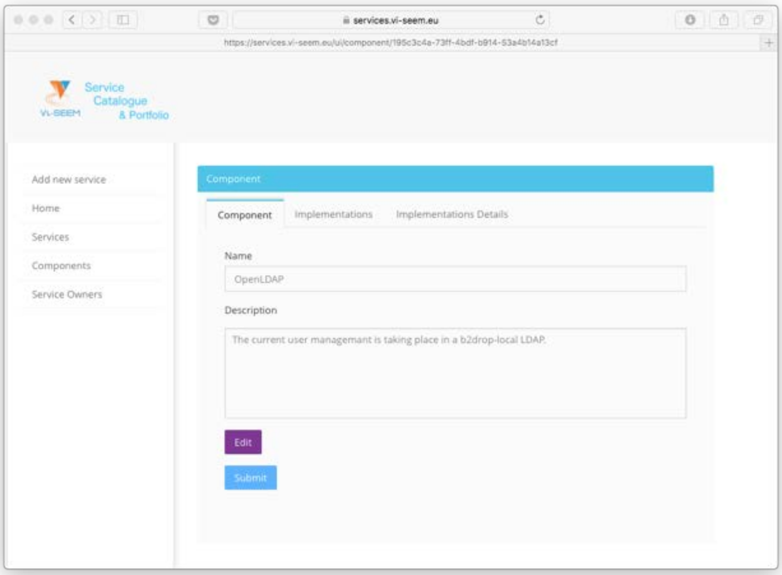

Figure 13: Service component editing of the service catalog/portfolio.

According to FitSM, each service should have a service owner, responsible for the operation of the service. The details editable by the system for the service owners include their names and contacts (email and phone), as well as their respective institutions. Service owner view of the service catalog/portfolio is shown in Figure 14. Currently, there are 17 service owners in the system, one per each service in the service portfolio.

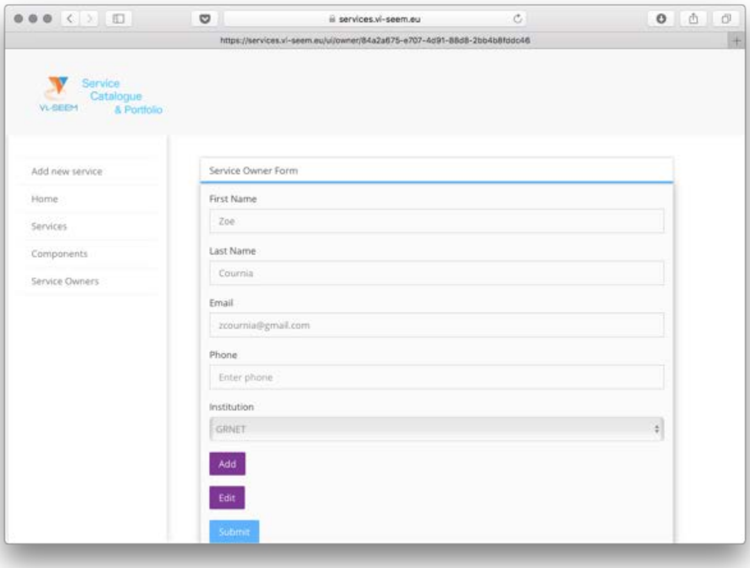

Figure 14: Service owner view of the service catalog/portfolio.

4.2. Usage statistics

Currently, the VI-SEEM service portfolio contains 21 services. Out of them, 16 are presented to the end-users in the service catalog, one is in deployment and 4 are internal services. The services in the service catalogue are grouped in the following 5 service areas.

- Computing resources related services, providing the foundation for the other, higher level services, such as:
  - VI-SEEM Cloud,
  - VI-SEEM Grid,
  - VI-SEEM HPC.
- Data storage, containing services that are related to all the aspects of the storage of data relevant to the project, such as:
  - VI-SEEM Archival Service,
  - VI-SEEM Simple Storage Service,
  - VI-SEEM Repository.
- Application specific services, addressing the specific requirements of the virtual research communities involved in the project, such as:
  - ChemBioServer,

    - VI-SEEM Regional Community Datasets,
    - VI-SEEM Live Access Server,
    - AFMM,
    - VI-SEEM Scientific Application Environment,
    - VI-SEEM Workflow, Software and Tools Repository,
    - NANO-Crystal,
    - VI-SEEM Clowder.
- Authentication and authorization services, such as VI-SEEM Login,
- Service provisioning, such as VI-SEEM Service Portfolio Management System.

The internal services listed in the service portfolio are:

- VI-SEEM GOCDB instance, a configuration management database,
- VI-SEEM Helpdesk,
- VI-SEEM Monitoring,
- VI-SEEM PID Handle services.

We expect that during the lifetime of the project, and even after its completion, this list will continue to grow, providing the necessary support to the research communities in the SEE and EM region.

5. Authentication and Authorization Infrastructure

VI-SEEM AAI is based on OASIS SAML standard to implement AAI (Authorization, Authentication and Identification) services, and has established a central identification service - VI-SEEM Login service – able to identify, authorize and authenticate VI-SEEM users to further allow their access to the offered services.

The following service have already been integrated with the VI-SEEM Login service:

- VI-SEEM Repository,
- VI-SEEM Code Repository,
- VI-SEEM Monitoring,
- VI-SEEM Simple Storage Service,
- VI-SEEM Simple Storage Service instance at ETFBL,
- VI-SEEM Service Portfolio Management System,
- VI-SEEM Technical Wiki

5.1. VI-SEEM Login

The VI-SEEM Login service enables research communities to access VI-SEEM e-Infrastructure resources in a user-friendly and secure way. More specifically, VI-SEEM Login allows researchers whose home organizations participate in one of the eduGAIN federations to access the VI-SEEM infrastructure and services using the same credentials they are using at their home organization. Furthermore, VI-SEEM Login supports user authentication with social identities, enabling even those users who do not have a federated account at home organization to be able to seamlessly access the VI-SEEM services without compromising the security of the VI-SEEM infrastructure. The current list of supported Identity Providers is available at

https://aai.vi-seem.eu/proxy/module.php/core/authenticate.php?as=sso.

VI-SEEM Login serves as a central hub between federated Identity Providers (IdP) and Service Providers (SP). More specifically, it acts as a Service Provider towards the external Identity Providers and as an Identity Provider towards the VI-SEEM Service Providers. Common technical services, such as Identity Provider discovery and user registration, are provided centrally by VI-SEEM Login, and do not have to be implemented by each individual service. VI-SEEM Login is also responsible for aggregating user attributes originating from various authoritative sources (IdPs and attribute authorities) and delivering them to the connected Service Providers in a harmonized and transparent way. Service Providers can use the received attributes for authorization purposes.

The core of VI-SEEM Login service is the IdP/SP proxy component, which acts as a bridge between the VI-SEEM services and external authentication sources and identity providers. This separation between the internal services and the external authentication sources/identity providers allows the service developers to focus on the service features and abstract away the complexity of multiple IdPs, Federations, Attribute Authorities and different authentication and authorization technologies. This complexity is “outsourced” and handled centrally by the proxy. Services need to establish trust with just one entity, the IdP/SP proxy. Typically, services will have one static configuration for the IdP/SP proxy. Such configuration, based on one IdP, removes the requirement from the services to operate their own IdP Discovery Service, which is a common requirement for services supporting federated access. Furthermore, all internal services will get consistent and harmonized user identifiers and attributes, regardless of the home organization or the research community that the user belongs to. Finally, this separation allows the management of the internal services to be independent of the management cycles at the home organizations IdPs. IdPs establish trust with one entity, the operator of the IdP/SP proxy and they are not impacted by the change operations of each individual service.

The above architecture has been instantiated and deployed on ~okeanos cloud implementing a high-availability and load balancing schema. The deployment consists of a Load balancer (nginx), 2 IdP/SP Proxy servers connected to a redundant Data Base backend consisting of a Cache server and a Data Base Master server.

SimpleSAMLphp caches user sessions in Memcached, an in-memory key-value store for small chunks of arbitrary data. COmanage maintains VI-SEEM user profile information in PostgreSQL DB cluster. Data is synced between the master (read/write) and the hot standby slave (read-only queries). Sessions are distributed and replicated among different Memcached servers, enabling load-balancing and fail-over. Finally, user requests are load balanced among multiple SimpleSAMLphp web front-ends that use the back-end matrix of Memcached servers.

Currently the functionality implemented is the following:

- REFEDS Research & Scholarship SP available via eduGAIN:
- Centralised IdP Discovery
- User Enrolment / account linking
- User Consent
- Support for OIDC/OAuth2 IdPs:
    - Google, Facebook, LinkedIn, ORCID
- Support for SAML services

VI-SEEM Login integration guides have been developed and are available in the VI-SEEM wiki. In the following months, more VI-SEEM services registered in the VI-SEEM service catalog will be integrated with the VI-SEEM Login service. According to agreed deployment plan, application level services will be the ones to be integrated first.

6. Conclusions

Based on a comprehensive plan for the deployment of the infrastructure and services presented in D3.1 – Infrastructure and services deployment plan, the project so far integrated available computing (HPC, Grid, Cloud) and storage resources, and deployed infrastructure services that fulfill the needs of application services and users. The main conclusions from this process and recommendations for the next period are presented in this deliverable, and can be summarized as follows:

1. Up to this point, VI-SEEM integrated 22,816 CPU-cores, 759,296 GPU-cores, 18,300 Xeon Phi-cores, 3,112 Grid CPU-cores, 14,152 Cloud VM-cores, and 18 PB of storage space. Current level of commitment of these resources to the project are as follows: 12.2% of CPU-cores, 12.9% of GPU-cores, 10% of Xeon Phi-cores, 10.3% of Grid CPU-cores, 4.5% of Cloud VM-cores, and 7.6% of storage space. Most of these resources were integrated by project month M4 (January 2016), while the rest was integrated by project month M9 (June 2016). Although percentage of dedicated resources is slightly below the defined key performance indicator value (in terms of Cloud and storage resources), it is based on real resource demands and will be increased with the 2nd and 3rd call for production use of the resources, as required.
2. Provided infrastructure resources are utilized during: 1st integration phase from project month M5 to M12, 2nd integration phase from M5 to M18, and at present time by accepted applications from the 1st call for production use of resources, which is still ongoing. In total, 1.5 million CPU-hours and 79 million GPU-hours were consumed, mainly by climatology and life science communities. Cloud resources are mainly consumed by digital cultural heritage and life science communities, and in overall 39 Cloud VMs were used. Major use of resources is expected in the ongoing 1st call, where 13.7 million CPU-hours, 3.5 million GPU-hours, and 1.0 million Xeon Phi-hours are allocated.
3. In total, 5 (out of 6 planned) operational and resource management tools were deployed for efficient management of the available computing and storage resources. Monitoring system is partially deployed, and will be fully integrated by the project month M20 (May 2017). All available operational and resource management tools are fully integrated with project’s Authentication and Authorization Infrastructure (AAI).
4. VI-SEEM service portfolio management system was deployed to support the service portfolio management process within VI-SEEM. The tool could be used by other infrastructures as well, and therefore is listed in the VI-SEEM Innovation Register [2]. It is fully integrated with project’s AAI.
5. VI-SEEM AAI – VI-SEEM Login service – is based on OASIS SAML [3] standard to implement AAI services, and has been established as a central identification service able to identify, authorize and authenticate VI-SEEM users to further allow their access to the offered services. Integration process is straightforward, documented, tested, and implemented in several use-cases.