H2020-EINFRA-2015-1

VI-SEEM

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

Deliverable D3.1

Infrastructure and services deployment plan

Author(s): Dusan Vudragovic (ed), VI-SEEM consortium

Status –Version: Final

Date: January 29, 2016

Distribution - Type: Public

Abstract: Deliverable D3.1 – "Infrastructure and services deployment plan", provides an overview of the infrastructure available to the VRE, and the services deployment plan to cover the needs of the applications services and users.

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
December 1st, 2015	a	Dusan Vudragovic	Initial version of ToC.
December 11th, 2015	b	Dusan Vudragovic	Final version of ToC.
December 20th, 2015	c	Ioannis Liabotis, Constantinos Lazarou, Stelios Erotokritou, Andreas Panteli, Emanouil Atanassov, Sofiya Ivanovska, Stiliyan Stoyanov, Antun Balaz, Dusan Vudragovic, Vladimir Slavnic, Ivan Marton, Kristof Bajnok, Tamas Frank, Silviu Panica, Neki Frasheri, Dorian Minarolli, Mihajlo Savic, Boro Jokimovski, Anastas Misev, Luka Filipovic, Nicolai Iliuha, Hrachya Astsatryan, Ramaz Kvatadze, Youssef Eldakar, Zivan Yoash, Salman Matalgah	Integration of partners' contributions.
January 29th, 2016	d	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] VOMS-Admin page of vo.vi-seem.eu Virtual Organization
https://voms.ipb.ac.rs:8443/voms/vo.vi-seem.eu

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

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

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

[6] VI-SEEM Source Code Repository
https://code.vi-seem.eu/

[7] VI-SEEM Wiki page
http://wiki.vi-seem.eu/

[8] Information Technology Infrastructure Library
https://en.wikipedia.org/wiki/IT_service_management#Information_Technology_Infrastructure_Library_.28ITIL.29

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

[10] OTRS IT service management
http://www.otrs.com/

[11] Citsmart IT service management
http://www.citsmart.com.br/

[12] iTop IT service management
http://www.combodo.com/?lang=en

[13] Geant service eduGAIN
http://services.geant.net/edugain/

List of Figures

List of Tables

Glossary

AAI	Authentication and Authorization Infrastructure
AMQP	Advanced Message Queuing Protocol
API	Application Programming Interface
CA	Certification Authority
CMDB	Configuration Management Database
EGI	European Grid Infrastructure
EM	Eastern Mediterranean
EUDAT	European Data Infrastructure
IGTF	Interoperable Global Trust Federation
ITIL	Information Technology Infrastructure Library
ITSM	IT Service Management
LDAP	Lightweight Directory Access Protocol
OLA	Operational Level Agreement
PKI	Public Key Infrastructures
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
VO	Virtual Organization
VRE	Virtual Research Environment

Executive summary

What is the focus of this Deliverable?

The focus of this deliverable is to provide detailed execution plan for integration of the available computing (HPC, Grid, Cloud) and storage resources, as well as a plan for development and deployment of the operational services that will enable uniform management and access provision to the VRE resources to be used by the supported scientific communities.

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

The contents of this deliverable will form a basis for the other work packages so as to make efficient use of the e-Infrastructure services. The service registry, whose initial design and plan for development is presented in this deliverable, includes both the common services (WP3 for e-Infrastructure and WP4 for data services) that are used by all three scientific communities, as well as the application-specific services that WP5 will deploy. Authentication and Authorization Infrastructure proposed in the document will enable user access management deployed by the infrastructure and WP3, WP4, and WP5 service components.

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

- WP4.1 – Data services design
- WP4.3 – Data collection and provisioning
- WP4.4 – Data analysis
- WP5.1 – Refinement of service requirements & tech 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 services available to the VRE, and the services deployment plan to cover the needs of the applications and users. In particular, it gives detailed information on:

- Available computing and storage resources, and plan for their integration into the regional interoperable VI-SEEM infrastructure (Section 2).
- Plan for development and deployment of operational tools that will enable resource monitoring and management, and facilitate interoperability of the infrastructure (Section 3).
- The initial plan for development and deployment of service registry, a formal portfolio and catalogue of services offered by the project, which will enable service portfolio management and service discovery (Section 4).
- Plan for development and deployment of Authentication and Authorization Infrastructure (AAI) responsible for management of user access to the resources and services (Section 5).
- The extensive appendices give overview of the available computing (HPC, Grid, Cloud) and storage resources.

Conclusions and recommendations

This deliverable presents a comprehensive plan for the deployment of available HPC, Cloud, Grid and storage resources, and development of services and management solutions for operation of this infrastructure. The main conclusions of this deliverable can be summarized as follows:

1. Integration of available computing and storage resources started from the first month of the project. Most of the resources are made available in PM04 (January 2016), while the rest of the resources will be integrated gradually, by PM10 (July 2016). The infrastructure will deliver 18.8 CPU, 371.6 GPU, 16.0 Xeon Phi, 5.3 IBM Cell, and 1.6 Grid CPU millions of hours per year, as well as 468 virtual machine cores to the scientific communities. Partner responsible for coordinating the integration is IICT-BAS.
2. Set of operational tools for resource monitoring and management will be deployed, and user support will be provided. These tools will be in production by PM09 (Jun 2016), and partner IPB is responsible for this task.
3. Service management and discovery will be provided via service registry – a formal portfolio of services offered by the project. The registry will be deployed by PM09, and partner UKIM is responsible for all related activities. Beta version of the service will be deployed by PM06.
4. Authentication and authorization infrastructure will be deployed within the WP3 by PM09. It will allow users to be authenticated using a single identity, and provide service for uniform managing of access to the resources regardless of their type and location. Beta version of the service will be deployed by PM06. Partner responsible for this task is NIIFI.
5. The specific operational solutions were chosen so as to insure compatibility with related project such as AARC, EUDAT2020, EGI-Engage, etc.

The deliverable confirms the strong commitment by the resource-provisioning partners (all project partners) to share substantial computing and storage resources, collaborating with all the countries in SEEM region, as well as to share the responsibility for common operation of these resources, supported by a number of operational/management tools.

1 Introduction

Infrastructure and service deployment within the VI-SEEM project will be performed horizontally as well as vertically through four major tasks. The horizontal deployment will be 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) and geographically distributed resource centers. The vertical deployment will cover development of a service portfolio for the efficient management of the project's services, as well as authentication and authorization infrastructure to enable a uniform management and resource access provisioning. Additionally, a bidirectional task will support site managers for operating the integrated services on top of the underlying infrastructures, operate common user services, and support the users in their day-to-day usage of the infrastructure.
In this deliverable we focus on detailing the task features and specifying the distribution of work between partners in this respect.

2 Infrastructure resources and availability plan

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, storage resources with possibility for short and long term storage. The heterogeneous nature of the infrastructure presents a challenge to the project's operational team, but is also an advantage for the users because of its ability to support different types of applications. Modern, state-of-the-art technologies for computing, virtualization and storage, are made available to the developers and end-users, in most cases early during the project's lifetime. The amount of resources planned meet or exceed the initial commitments and in our experience will be adequate to support the ramp-up of applications' usage. 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].

2.1 Available HPC resources

The HPC resources of the project consist of clusters with low-latency interconnection or supercomputers. Currently two of them, Avitohol from Bulgaria and Leo from Hungary, are at the November top500 list of supercomputers (at 389th 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 are BlueGene/P systems, as well as one based on the Cell processor (PS3 cluster IMAN1-Booster/King). The summary information about the available resources and their dedication to the project is shown in Table 1.

Resource	Country	Total			Dedicated (hours per year)			PM
		CPU-cores	GPU-cores	Phi-cores	CPU-hours	GPU-hours	Phi-hours	
ARIS	Greece	8,520	-	-	3,000,000	-	-	PM01
Cy-Tera	Cyprus	1,392	16,128	-	1,829,088	21,192,192	-	PM09
Avitohol	Bulgaria	2,400	-	18,300	2,102,400	-	16,030,800	PM04
PARADOX	Serbia	1,696	108,544	-	742,848	47,542,272	-	PM01
NIIFI SC	Hungary	768	-	-	421,882	-	-	PM04
Leo	Hungary	1,344	628,992	-	588,672	275,498,496	-	PM04
InfraGRID	Romania	456	3,136		350,400	5,494,272	-	PM04
ICAM	Romania	4,096	-	-	7,176,192	-	-	PM04
UPT-HPC	Albania	144	-	-	126,144	-	-	PM04
FINKI	FYROM	768	-	-	336,384	-	-	PM04
Armcluster	Armenia	128	-	-	112,128	-	-	PM04
BA-HPC	Egypt	1,040	-	-	1,822,080	-	-	PM04
Gamma	Jordan	8	2,496	-	70,080	21,864,960	-	PM08
Zaina	Jordan	56	-	-	147,168	-	-	PM06
Total		**22,816**	**759,296**	**18,300**	**18,825,466**	**371,592,192**	**16,030,800**	**PM09**

Table 1 – Available HPC resources and their dedication to the project.

At the beginning of the lifetime of applications the available resources will be more than enough, although some strain on the resources can be expected towards the end of the project. However, since resource deployment is dependent on some unpredictable factors like national funding, we can expect that in 2 years' time some partners will be able to bring in additional resources and cater for the increased use.

Dynamic of the integration of available HPC resource into the VI-SEEM infrastructure is shown in Figure 1. Most of these resources will be integrated by project month 4 (January 2016), while the rest of them will be integrated by June 2016. In total 18.8 CPU, 371.6 GPU, 16.0 Xeon Phi, and 5.3 IBM Cell millions of hours per year will be made available to the scientific communities.

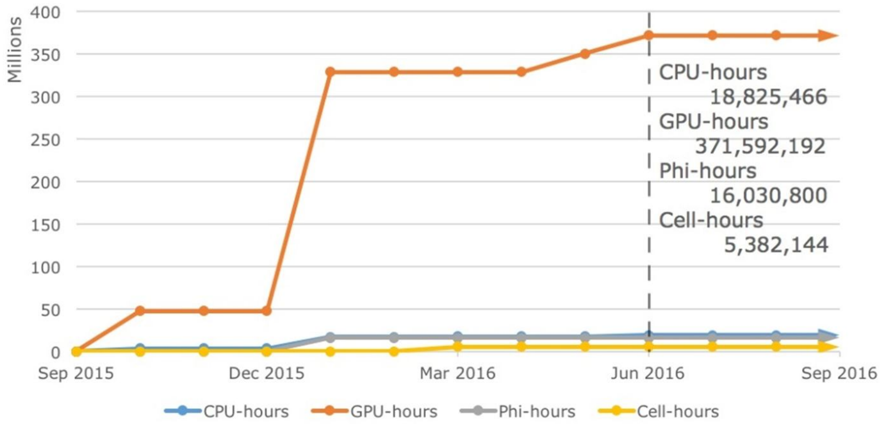


Figure 1 – Plan for integration of available HPC resources.

2.2 Available Cloud resources

The cloud resources provided to the project will be used in two ways. Those clouds that provide the ability to launch VMs with public IPs will provide the possibility to deploy VRE services, for their main or backup/fail-over instance, as was envisioned in the project proposal. Some VMs that possess only private IPs will be used for distributed data processing where necessary. However, at this point it is expected that most of the data processing will be done on the HPC resources, which have better features in this respect. The summary information about available cloud resources is presented in Table 2. Overall it transpires that the available cloud resources will be sufficient for the expected usage.

Resource	Country	Total VMs	Dedicated VMs	Dedicated VM-hours	PM
~Okeanos	Greece	10,000	200	1,752,000	PM01
CyI Cloud	Cyprus	176	18	157,680	PM06
Avitohol	Bulgaria	2,400	120	1,051,200	PM10
InfraGRID Cloud	Romania	400	46	402,960	PM04
UPT-Cloud	Albania	12	6	52,560	PM10

ETFBL-CC01	Bosnia	60	13	113,880	PM06
MK-04-FINKI_CLOUD	FYROM	436	24	210,240	PM04
MD-Cloud	Moldova	12	3	26,280	PM04
IIAP Cloud	Armenia	96	10	87,600	PM06
IUCC InfinityCloud	Israel	560	28	245,280	PM04
	Total	**14,152**	**468**	**4,099,680**	**PM10**

Table 2 – Available Cloud resources and their dedication to the project.

Dynamic of integration of available Cloud resource into the VI-SEEM infrastructure is shown in Figure 2. Most of these resources will be integrated by project month 4 (January 2016), while the rest of them will be integrated by July 2016. In total 468 Virtual Machines, i.e. 4 million of VM-hours will be made available to the scientific communities.

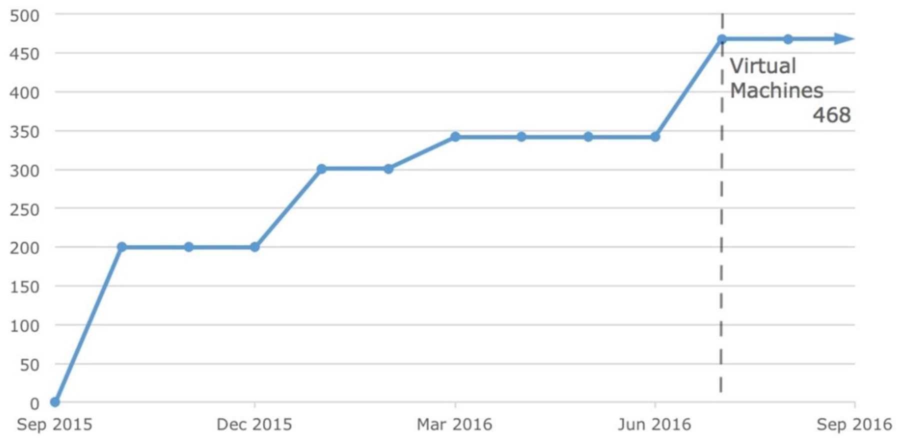


Figure 2 – Plan for integration of available Cloud resources.

2.3 Available Grid resources

The Grid resources, available for the project, will be provided mostly from smaller clusters that have already installed Grid middleware and are part of the European Grid Infrastructure. Some of the clusters that are not certified at the level of the EGI will have to be incorporated through appropriate means, e.g., by using modified BDII services. The provisioning of these clusters for the VI-SEEM VRE is not part of EGI activity and as such is distinct effort.

One cluster that was originally proposed as a Grid cluster was deemed more appropriate to be used as an HPC resource and has been moved to the section above. The Grid resources are summarized in Table 3.

Resource	Country	Total CPU-cores	Dedicated CPU-cores	Dedicated CPU-hours	PM
Hellas Grid	Greece	864	43	370,000	PM04

BG01-IPP	Bulgaria	640	120	1,051,200	PM04-PM16
AEGIS01-IPB-SCL	Serbia	704	35	308,352	PM01
MK-03-FINKI	FYROM	768	38	336,384	PM04
MREN01CIS	Montenegro	32	16	140,160	PM05
MD-GRID	Moldova	40	12	105,120	PM04
GE-01-GRENA	Georgia	128	20	175,200	PM04
	Total	**3,176**	**190**	**1,660,896**	**PM05**

Table 3 – Available Grid resources and their dedication to the project.

Project dedicated Virtual Organization (vo.vi-seem.eu) has been already deployed at the IPB, and its web interface, VOMS-admin is available at [2]. Beside of the virtual organization management service, a set of Grid core services will be deployed at IPB, IICT-BAS, and GRNET. Plan for deployment of these services is specified in Table 4. Functional description of Grid core services from a user's point of view is given in gLite User Guide (https://edms.cern.ch/ui/file/722398/1.4/).

Service	Responsible partner	In production
VOMS	IPB	PM03
BDII	IPB, GRNET	PM03, PM06
LFC	IPB	PM03
WMS	IPB, GRNET, IICT-BAS	PM03, PM06, PM06
LB	IPB, GRNET, IICT-BAS	PM03, PM06, PM06
PX	IPB	PM03

Table 4 – Plan for deployment of Grid core services.

Dynamic of integration of available Grid resource into the VI-SEEM infrastructure is shown in Figure 3. The most of Grid sites will be integrated by project month 5 (February 2016), while the rest of them will be integrated by February 2016. In total 1.6 millions of CPU-hours per year will be made available to the scientific communities.

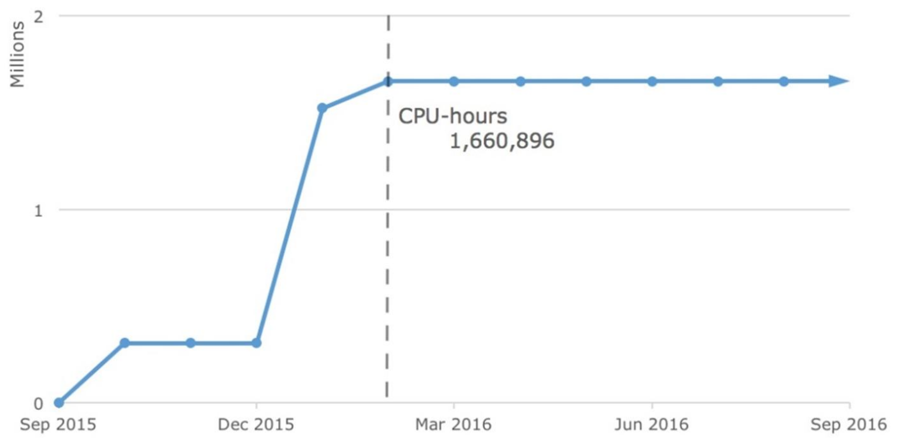


Figure 3 – Plan for integration of available Grid resources.

2.4 Available Storage resources

The storage resources that were described above, together with the dedications to the VI-SEEM project, are 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. 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 5 summarizes the storage resources of this type, planned to be provided by partners.

Resource	Country	Total		Dedicated		PM
		Disk [TB]	Tape [TB]	Disk [TB]	Tape [TB]	
ARIS	Greece	1,000	3,000	50	210	PM04, PM10
Cyprus	Cy-Tera	500	-	100	-	PM04
Avitohol	Bulgaria	96	-	5	-	PM04
PARADOX	Serbia	96	-	5	-	PM01
NIIFI HSM	Hungary	87	6,800	3	300	PM04
NIIFI iSCSI	Hungary	1,000	-	50	-	PM04
UVT HPC GPFS	Romania	50	-	5	-	PM04
ETFBL-CC01	Bosnia	1		0.5		PM06
MK-04-FINKI	FYROM	36	-	2	-	PM04
RENAMstor	Moldova	4	-	1	-	PM04
IIAP Storage	Armenia	8	-	3	-	PM04
BA-IA	Egypt	5,200	-	100		PM04
IUCC Storage	Israel	40	-	5	-	PM04
Total		**8,118**	**9,800**	**329.5**	**510**	**PM10**

Table 5 – Available Storage resources and their dedication to the project.

Dynamic of integration of available storage resource into the VI-SEEM infrastructure is shown in Figure 4. The most of these resources will be integrated by project month 4 (January 2016), while the rest of them will be integrated by July 2016. In total 330 TB of disk and 510 TB of tape storage will be provided.

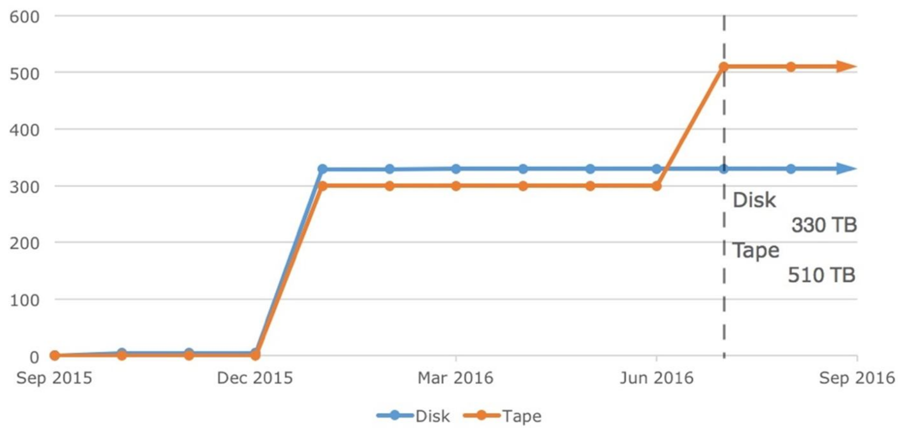


Figure 4 – Plan for integration of available Storage resources.

3 Operations and resource management

In order to achieve efficient management of the available computing and storage resources, and facilitate interoperability of the infrastructure, a set of operational tools will be deployed. Monitoring will be based on ARGO Service Monitoring system. This system allows integration with GOCDB that will contain all technical details relevant for configuration of monitoring system. Integration will automatically trigger monitoring of all services once they are inserted in the GOCDB. Accounting system will provide accurate measurements of the utilization of the different types of resources in the VI-SEEM infrastructure, and results will be provided through a web interface. Source code repository described in this section will contain all codes developed within the project. User support and service-related problems will be resolved mainly through the Trouble ticketing system, but via technical mailing list as well. Technical documentation, know-hows, best practices, guidelines, as well as training material will be provided via project's wiki system.

Deployment of operational tools is distributed among VI-SEEM partners and will be finalized in PM09 as shown in Table 6.

System	Responsible partner	Country	In production
Monitoring	GRNET	Greece	PM06
GOCDB	UKIM	FYR of Macedonia	PM09
Accounting	IICT-BAS	Bulgaria	PM09
Source code repository	UoBL, UoM	Bosnia and Herzegovina	PM03, PM06
Trouble ticketing system	UoBL, UoM	Bosnia and Herzegovina	PM06, PM09
Technical wiki	CYI	Cyprus	PM03

Table 6 – Plan for deployment of operational and resource management tools.

3.1 Monitoring

VI-SEEM project has very complex structure, from infrastructure that spans several concepts 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 domains mentioned. It is of paramount importance to have a stable and fully functional monitoring component in order to provide satisfactory environment for all involved. Based on rich experience from recent projects of similar nature that also involved grid, HPC and cloud infrastructure monitoring, it was decided to primarily base the monitoring on the ARGO Service Monitoring system [3] and adapt it to cover specific project needs.

ARGO Service Monitoring represents a flexible and scalable framework suitable for monitoring of status, availability and reliability of services. Modular nature enables for integration with external system which is an important characteristic as the system will have to cooperate with various other project systems in order to provide efficient environment for both project personnel and end-users.

Architecture (shown in Figure 5) is based on layered approach with lowest layer, the Monitoring Engine, being based on Nagios [4] for service endpoints monitoring by utilizing custom probes and a set of optional add-ons. Availability and Reliability layer sits above in hierarchy and is comprised of so called Connectors that connect Monitoring Engine to Analytics Engine on top of which is Web API. This approach provides for flexible configuration that covers monitoring metrics, notifications and alarms, status reports, support for SLAs/OLAs as well as integration with external Service Catalogs and Configuration Management Databases. While ARGO supports distributed model with dispersed monitoring engines, it was decided that centralized model best suits the project at this stage, with possible migration to more complex topology if the need arises.

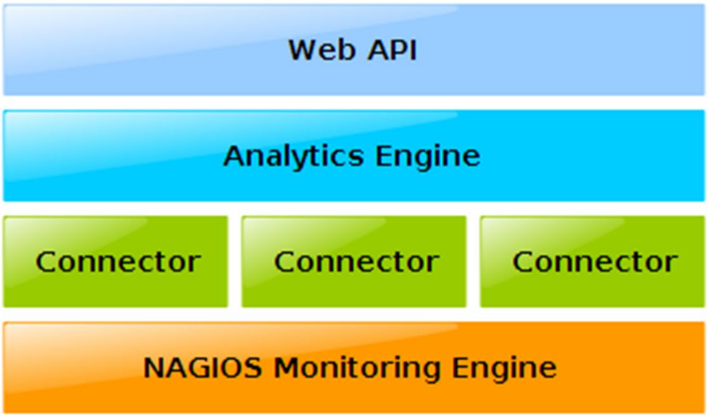


Figure 5 – Architecture of ARGO system.

ARGO service within VI-SEEM project will be provided by GRNET with initial deployment and testing in PM04 and PM05 and production use in PM06. Development of custom components will be carried during and after this period as new requirements are met.

When discussing the effort within the project on adapting and extending existing ARGO components, we can identify four main domains ordered by priority: custom probes, Web UI, Reporting component and Alternative interfaces.

Design and implementation of the probes should be carried out by service developers and has to follow guidelines specified in EUDAT Task 6.1.5 – A/R Monitoring, section Guidelines for developing monitoring probes. These guidelines cover in detail development, packaging, integration, testing and deployment of newly developed probes. All source codes for developed probes will be available at project source code repository.

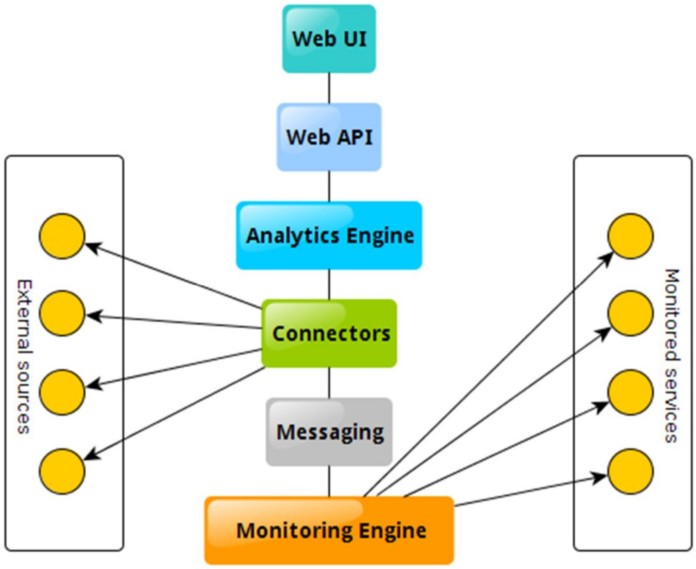


Figure 6 – ARGO Modular Architecture.

Modular architecture of ARGO system is shown in Figure 6. Custom Web UI will have to be developed in order to enable VREs and other actors to use the monitoring system in a straight-forward and fully efficient manner. This component should provide support for multi-tenancy, federated identities as well as integration with other project-specific services and components. Significant problem that needs to be solved is defining the scope and depth of raw and processed data available to different classes of users that exposes the needed data but without overwhelming the users or leaking unintended information.

Reporting component will provide customized reports for status, availability and reliability for use by both VREs as well as operations personnel. Although the underlying data might be the same, the views suitable for different user classes dictate different reports as well as interfaces.

Alternative interfaces aim to provide simpler integration of monitoring data generated by the project into existing monitoring systems at partner institutions. One such approach could be exporting a subset of data through widely supported SNMP standard.

3.2 GOCDB

The proposed configuration management database to be used within the VI-SEEM project will be GOCDB. GOCDB is a Configuration Management Database (CMDB) for recording and managing assets in e-infrastructure projects. It defines a number of topology objects including admin-domains, sites, services, service-groups, service-endpoints, service-downtimes, users and roles. GOCDB is the central CMDB for the EGI-Engage and EUDAT2020 projects. 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. Importantly, 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.

Core objects can also be extended using an extensibility mechanism that allows custom key-value pairs to be added to those objects. These objects can then be flexibly filtered when selecting/querying data.

An authentication abstraction layer has been integrated to allow different authentication mechanisms to be supported using a pluggable 'AuthenticationProvider' interface. Requests are authenticated using extensible 'Authentication' tokens. Implementations are provided for x509, SAML2 and username/password. These methods of authentication are described in detail in the Section 5.

An instance of the GOCDB will be deployed by the partner UKIM for the needs of the project. The instance will be customized to reflect the actual needs and objects of the VI-SEEM project infrastructure and services. It will also be integrated with the project's Authentication and Authorization Infrastructure (AAI).

3.3 Accounting

The purpose of the accounting system is to provide accurate measurements of the utilization of the different types of resources in the VI-SEEM infrastructure, using commonly accepted metrics. All the accounting information, relevant to the VI-SEEM use of the provided infrastructure resources will be gathered in one central database, using accounting publishers appropriate for the respective types of resources. Overall architecture of the accounting system is shown in Figure 7.

An accounting publisher package will be installed on the controlling nodes where information about the usage can be collected, for each type of resources (HPC, Grid, Cloud Iaas and storage). From the point of view of the accounting framework this will be a client-type service.

The accounting publisher packages will be written in Python and will have multiple independent modules to collect the resource usage data. In most cases this means to parse log files of the corresponding type. For ease of deployment the dependencies of the packages will be kept to a minimum.

The accounting publisher will format data into messages and transfer them via AMQP server. The server-side software will receive the messages from the AMQP server and decode the incoming formatted data, translate it into SQL queries for insertion in the central relational database. The server-side package will be written again in Python. For the development the previous experience of the project team in the gathering of regional accounting will be useful and some of the software packages will be built upon previous sources.

With regards to the metrics that are to be collected, we follow established practices in HPC and distributed infrastructures to determine exactly what data should be available.

The Uniform Accounting Record format will be used for gathering accounting information about computational jobs, regardless whether they are from the Grid or HPC.

Care will be taken to properly account for the Multi-CPU jobs, e.g., to attach 1000 CPU-hours usage for a job that uses 1000 CPUs for one hour - something which was a problem with some other Grid accounting software.

About the storage we should be able to determine how much is the storage that is used by the project at each of the sites, with acceptable granularity.

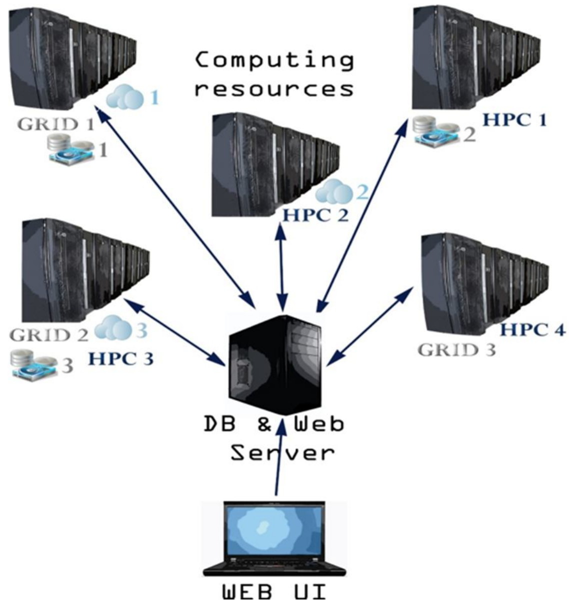


Figure 7 – Overall architecture of the accounting system.

Access to the accounting

Access to the accounting data will be provided through a web interface – the accounting portal. It will be build using HTML, CSS, JavaScript, JQuery, Ajax and PHP. The interface will support the different roles that are necessary. The roles that are envisaged at this stage are as follows:

- User: any user should be able to query the system and obtain information about their own use, as well as some general information about all the utilization.
- Team leader: leader of application team should be able to obtain information about the usage by their team.
- System administrators: system administrators should be able to obtain information about the VI-SEEM usage of their own systems. In this way they should be also able to compare with their internal information, if they have such systems deployed.
- Project staff: leaders of work packages, etc. should be able to obtain information about the overall utilization of VI-SEEM infrastructure, with sufficient granularity.

The accounting is the third pillar of the AAA (triple-A, Authentication, Authorization and Accounting), and as such will be built to use the Authentication and Authorization Infrastructure accepted for the project, so as to avoid the creation of any separate infrastructures. Through authentication users will legitimate themselves to gain access to the system. Following successful login, a user will obtain authorization for doing tasks, appropriate to their roles.

There will be variety of options to choose from to elaborate the search and get appropriate results. The user web-interface will provide settings for specifying start and end of time period, job type – Grid, HPC or Cloud, to which virtual organization it belongs, from which region, how many CPU hours it took and etc. The accounting application will process such requests and generate the needed data in tables and/or simple charts grouped by and sorted according to user's choice.

For example, a user can request to see comparison for how many CPU hours he used, grouped by HPC, Grid and Cloud in descending order for the last quarter. Project manager would be interested to see how different virtual organizations in the project utilize computing resources throughout the year, by monthly and quarterly periods. The system will provide data to show how the commitments of the partners are being fulfilled.

Hardware and software requirements

On the client side the main requirement is to have Python version 2 (2.6+) or Python version 3 (3.3+) with pika library for messaging service and access to the Internet.

The server must support hardware virtualization technology with a minimum of 8 logical cores and 16 GBs of RAM memory. Operating system of the server will be CentOS or Red Hat Enterprise Linux. This service will be operated at IICT-BAS. The message broker will also be installed there, using open source software to support the AMQP. The database is expected to be mysql, but no special mysql features will be used.

3.4 Helpdesk

After performing analysis of several available open source helpdesk / trouble ticket system, we have chosen the osTicket (osticket.com) 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.

Selected system supports fine grained access control and elaborate categorizations, enabling 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. As the project develops and evolves, it is possible to fine tune options available to different class of both end-users as well as support staff (agents).

When it comes to internal support organization, the system allows for both horizontal and vertical segmentation with possible domain border crossing entities when needed. This is achieved by segmenting along departments and support teams, as well as defining "Level 1 Support" responsible for initial interactions with users. For example, as the project is segmented in various work packages (departments) it is natural that similar structure will be present in helpdesk organization, but creating support teams that cross various departments enables for providing effective support for both application developers as well as end-users that need to troubleshoot issues that demand coordinated effort by work packages 3, 4 and 5.

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

Integration with project AAI will be established through the SAML 2.0 protocol. In the initial stages of the project, at least, it is planned to allow for a more liberal approach to end-user access enabling guest users to initiate support requests. Support accounts will be integrated with project AAI infrastructure. The integration will be implemented either by using either SimpleSAMLphp or mod_auth_mellon.

Integration with service registry/catalogue will be performed by creating an external application that will perform synchronization of internal helpdesk database and external data sources. This will mostly manifest itself to end-users as a hierarchy of offered help topics the system provides. This organization hides the internal organization of helpdesk from end-users and simplifies their interactions. Level of acceptable automation is to be determined during the use of the system as the system can provide great flexibility, yet one must retain relative simplicity for the end-users.

Integration with other related systems already existing within the project infrastructure or at project partners' premises will be developed as the needs arise. There is support for ticket filtering enabling to automatic routing of new tickets thus providing certain level of automation and initial ticket handling according to the defined rules. This functionality can be adapted to better integrate into existing workflows if needed.

The system supports somewhat simplistic but effective and easy to use reporting module with all needed data present in the open database and available for further analysis if such requirements become apparent during the project duration.

Since different work packages and VRE communities will probably have different requirements regarding the information important to issues being resolved, it is worth noting that the system supports creating custom forms for data entry tied to different help topics. This has positive effect on both end-users, as they have less cluttered interface and have less data to fill, and support staff, as they have to process only the relevant data.

Pilot installation of the helpdesk will be installed and tested in PM04 and will be available at http://support.vi-seem.eu/ in PM05 when testing finishes and integrated with other relevant project services and is ready for production use by project participants in PM06/PM07.

Primary instance of the helpdesk will be hosted at University of Banjaluka 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

Source code repository 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 [5] for the project needs. As can be devised from the name, 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 various other open source tools and libraries. There are several ways to deploy GitLab in local infrastructure, but for the project we chose Omnibus method as it fits best with project needs.

Aside from source code versioning control, GitLab also provides for 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 aptly named svn2git utility.

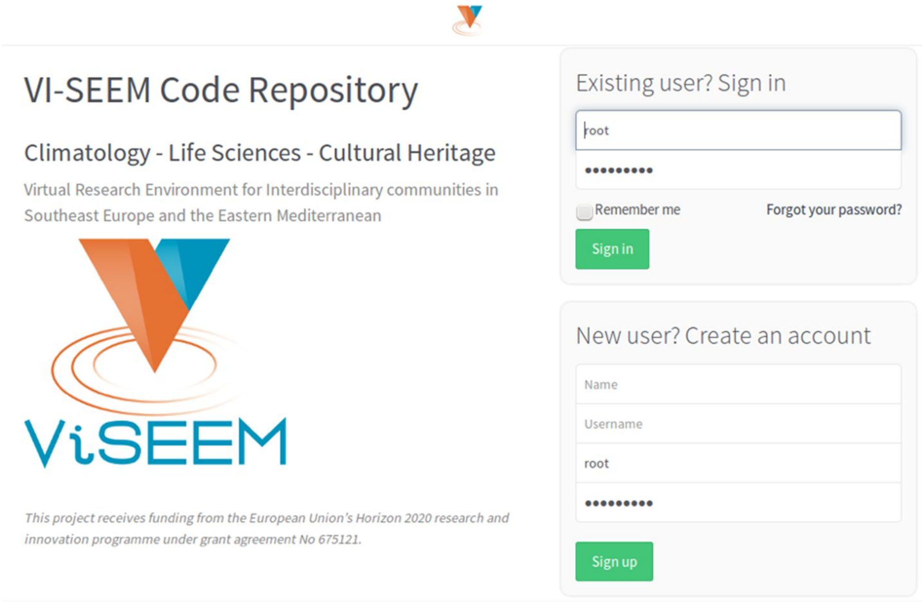


Figure 8 – VI-SEEM source code repository front page.

Integrating with project infrastructure and services is based mostly on allowing users to access the system by using project Authentication and Authorization Infrastructure. GitLab is quite flexible when it comes to user authentication and provides for a plethora of supported options via the OmniAuth package. This potentially allows users to log in with their existing Google, Facebook, Twitter, GitHub or Bitbucket credentials, but more importantly for the project needs, by utilizing Shibboleth or SAML 2.0. Prior to full integration with project AAI, user are allowed to self-register, but valid email address from the set of predefined project partners' institutional email domains is required.

The system was initially deployed in Q1 of the project and was put in production after it passed testing. It is fully operational and is available at the address [6]. Access is possible from either web interface (Figure 8) or from command line by utilizing git tools.

Primary instance of Source Code Repository System will be hosted at University of Banjaluka 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.

3.6 Technical wiki

The technical wiki of the VI-SEEM project is based on the MediaWiki platform. MediaWiki is a free software open source wiki package, created originally for use on Wikipedia. It's designed to be run on a large server farm for a website that gets millions of hits per day. MediaWiki is an extremely powerful, scalable software and a feature-rich wiki implementation that uses PHP to process and display data stored in a database, such as MySQL.

Pages use MediaWiki's wikitext format, so that users without knowledge of XHTML or CSS can edit them easily. When a user submits an edit to a page, MediaWiki writes it to the database, but without deleting the previous versions of the page, thus allowing easy reverts in case of vandalism or spamming. MediaWiki can manage image and multimedia files, too, which are stored in the filesystem.

VI-SEEM wiki is hosted at CyI partner and listens on [7]. The content of the wiki is publicly accessible for read only, while modifying requires a logged in user account. User accounts can be created only by VI-SEEM wiki administrators. Wiki system will be integrated with the project AAI when it becomes available.

4 Service Management and Service Catalogue / Portfolio

IT Service Management (ITSM) refers to the activities that an organization performs to plan, deliver, operate and control IT services that are offered to customers. Such activities facilitate the offering of quality IT services which provide value to customers, meeting their needs. Several approaches to ITSM exist and are described in the literature, while the most prominent is the Information Technology Infrastructure Library (ITIL) [8]. The current version of ITIL is being used by a large number of IT organizations for efficiently delivering services to their users and is a full and relatively complex ITSM framework. For new organizations that are also distributed in nature, i.e. services are being offered by a federated structure of different organizations the FitSM [9] is considered to be more appropriate as it is lightweight and it is compatible with ITIL being a first step for all organizations that later want to implement ITIL.

FitSM provide a set of process specific requirements for a service management system. Such requirements are listed in Table 6.

#	REQUIREMENT
PR1	Service Portfolio management (SPM)
PR2	Service Level Management (SLM)
PR3	Service Reporting Management (SRM)
PR4	Service Availability & Continuity Management (SACM)
PR5	Capacity Management (CAPM)
PR6	Information Security Management (SM)
PR7	Customer Relationship Management (CRM)
PR8	Supplier Relationship Management (SUPPM)
PR9	Incident & Service Request Management (ISRM)
PR10	Problem Management (PM)
PR11	Configuration Management (CONFM)
PR12	Change Management (CHM)
PR13	Release & Deployment Management (RDM)
PR14	Continual Service Improvement Management (CSI)

Figure 9 – FitSM requirements for a service management system.

Task 3.1 of WP3 deals in detail with service management requirements PR1 and PR2 where a formal service portfolio and service catalogue as well as the appropriate processes and responsibilities will be implemented and maintained.

The Service Portfolio (SP) is an internal list that details all the services offered by a *service provider* including those in preparation, live and discontinued. The service portfolio includes meta-information about *services* such as their *value* proposition, target *customer* base, *service* descriptions, technical specifications, cost and price, *risks* to the provider, service level packages offered etc.

The Service Catalogue (SC) is a customer facing list of services that are in production and provide value to the customers of the service provider. The SC, among others, provides also information on service options including the various SLAs available for each service.

At a high level the service catalogue is a subset of the service portfolio, both in terms of the number of services that they contain and also in terms of the number of fields or attributes each one holds.

FitSM defines a set of activities for the maintenance of the service portfolio and the corresponding catalogue. For the service portfolio such activities are:

- Initial Process Setup
  - Define a way to document the service portfolio;
  - Define a way to describe / specify a specific service;
  - Set up an initial service portfolio (including service specifications) covering at least all live services provided to customers, as far as they are in the scope of the service management system;
  - Create a map of the bodies / parties (organizations, federation members) involved in delivering services.
- Ongoing process execution
  - Manage and maintain the service portfolio;
  - Manage the design and transition of new or changed services;
  - Manage the organizational structure involved in delivering services.

For the service catalogue such activities are:

- Initial Process Setup
  - Define the structure and format of the service catalogue, and create an initial service catalogue based on the service portfolio;
  - Define a basic/default SLA valid for all services provided to customers, where no specific/individual SLA are in place;
  - Define templates for individual SLAs, OLAs and UAs;
  - Identify the most critical supporting service components, and agree OLAs and UAs with those contributing to delivering services to customers;
  - Agree individual SLAs with customers for the most important/critical services.
- Ongoing process execution
  - Maintain the service catalogue;
  - Manage SLAs;
  - Manage OLAs and UAs.

In the duration of the project WP3 will implement the above processes that are required for the Service Portfolio and Service Catalogue management. The following describes the structure of the service portfolio/catalogue and the definition of roles and responsibilities within the SPM/SLM process. This approach takes into consideration similar approaches on service portfolio management and service catalogue that are being performed in projects such as EUDAT2020 and PRACE-4IP.

Service Portfolio/Catalogue Template

Following a high level description of the VI-SEEM service portfolio/catalogue, fields that it will contain are grouped in the relevant categories shown in Table 7. In Table 7 fields that start with a * sign are part of the service catalogue as well as the service portfolio.

Field Name	Comment
Service Information/Description	
* Name	The name of the service
* Description	A description of the service
Service Status	(New, In development, In testing, Production, Deprecated)
Service Owner	The person responsible for all aspects of the service (i.e. development, and operations).
Contact Information (Internal)	The process to communicate any issue with the service internally in the organization or the federation
* Contact Information (External)	The process for the end users to communicate any issue with the service.
* Request Procedure	A description of the procedure available to potential users in order to request usage of the service.
* Service Category	Core, Operational, Storage, Computation, Application
Service Type	Customer Facing, Internal, Supporting
* Users of the service	A description of who are the potential users of the service
* Use cases and user requirements	A set of use cases and requirements this service fulfills
* Features	The list of service features
Service Components	The components necessary to deploy the service (can be distinguished in necessary or optional)
Service Dependencies	The dependencies of the service in other services provided by the organization or the federation
* Service Options	The different service packages
* Service Level Agreements	The SLA associated with every service option
Service Prerequisites	
* Usage Policy	A link to the usage policy for the service
* User Documentation	A link to the user documentation for the service
Operational Documentation	A link to the operational documentation for the service
Monitoring	A link to the monitoring portal for the service
Accounting	A link to the accounting information for the service
Business Case Information	
Value for customers	The value that the service provides to customers
Risks	The risks associated with the service
Competitors	A list of competitors for the service

Table 7 – VI-SEEM service portfolio/catalogue categories.

The following step in the process is the creation of a map with the bodies or the persons within the VI-SEEM federation that are responsible for delivering the services. Table 8 defines that at a high level.

Role	Responsibilities	Actor
Service Portfolio/Catalogue Process Owner	To control the SPM and SLM processes, maintain the catalogue and portfolio and report to senior management	T3.1 leader
Service Technical Coordinator / Architect	Has the overall view of services being developed or operated in the organization from the technical point of view	WP3, WP4 and WP5 leaders
Customer Relationship Manager	Gathers requests for new features from feature / service requestor, Initiates a new service / service change to the service portfolio, identifies services that need decommissioning	WP5 leader, SC leaders
Service Portfolio Approval Committee	Review and approves new services or changes to services	Technical Board
Service Owner	Has the overall responsibility for one specific service which is part of the service portfolio, Acts as the primary contact point for all (process-independent) concerns in the context of that specific service	A person assigned by the Service architect, usually a person working for the relevant to the service WP
Service Design Team	The team that is responsible for the design, implementation and maintenance of a service	

Table 8 – VI-SEEM service portfolio/catalogue responsibilities.

T3.1 will work towards fine-tuning and implementing the above processes. A tool will be developed to implement the service portfolio and catalogue based on the details presented bellow as described in the following sections.

Further to that specific SLA and OLA templates will be developed within T3.1 for ensuring the provision of high quality services. The work for the design of SLA and OLA templates as well as their implementation timeline, will be defined when the initial set of services is identified i.e. PM08 of the project.

4.1 Overview of available technologies

In order to select the necessary tools to support the IT service management, extensive analysis of the available tools was performed. The focus was on the open source tools that had active communities and/or operating support. The analysis showed that there are many such tools, but most of them are built having ITIL as guiding principles. There are almost no tools (or none that we are aware of) that are built using FitSM as the basics. Three of the tools were taken a closer look at, by making test deployments.

OTRS::ITIL is the most renowned IT service management tool. It is built as an extension to the OTRS tool, one of the oldest and most successful tools for helpdesk and process management. It is implemented in perl, with a community of over 5000 active members. Its primary goal was to be a helpdesk management tool, growing up to become a full ITSM solution. It fully supports ITIL recommendations. It has been deployed at more than 150.000 institutions in many different sectors. More about OTRS::ITIL can be found at [10].

Citsmart [11] is IT service management software, built to maintain the efficiency of the processes for delivering IT services and promote the improvement of business by increasing the quality of the services. It is built as a web based platform, using Java. It is certified to be ITIL compliant for the following processes:

1. Incident Management
2. Request Fulfillment
3. Knowledge Management
4. Service Catalog Management
5. Service Level Management
6. Change Management
7. Release & Deployment Management
8. Problem Management
9. Service Asset & Configuration Management

The software is built by a Brazilian company Central IT as a tool that is deployed mostly in the governmental institutions to support ITSM. The contact was established with the company itself through several Skype meetings. Regarding to their representatives, future versions of Citsmart might support FitSM and federated environment, such the one in the VI-SEEM and similar projects.

iTop [12] tool is built by the Combodo company, but it comes in a community version also, available through SourceForge. It is developed in PHP and offers very high degree of customization, both with and without code modification. It is also based on ITIL. If offers support for configuration management, service management, incident and problem management, SLA management, change management etc. It offers REST/JSON API for external access, and support for various authentication mechanisms. It is a very robust tool, which on the other hand makes it too complex to use for smaller enterprises. It does not support federated environments.

Although there are many ITSM tools, most of them are built having ITIL as the guiding principles. None of the tools support federated environment out of the box, which was the first and most important reason to lean toward the development of own solutions for the service catalog and portfolio management. The other reason was that there were already tools in place, with sufficient human expertise, to cover some of the elements that are covered by the analyzed tools (GOCDB, ARGO, …).

4.2 Design and plan for implementation

The high level architecture is presented in the Figure 10. The implementation will be done in Python, using Django framework. The code will be available through the project's code repository. MYSQL will be used as the data store for the solution. System will be accessible via RESTful interface, using JSON API requests from external applications and/or VI-SEEM site. There will also be a front-end, developed using AngularJS framework giving access to the project catalog/portfolio.

In the initial phase, the data entry will be done using the Django admin interface. The system will support the project-wide AAI methods. The implementation timeline for the service catalog/portfolio systems will be:

1. Final design (PM06)
2. Database implementation (PM07)
3. Beta version of the RESTful interface (M08)
4. Production version of the RESTful interface (M10)
5. Beta version of the UI (M10)

6. Production version of the UI (M12)

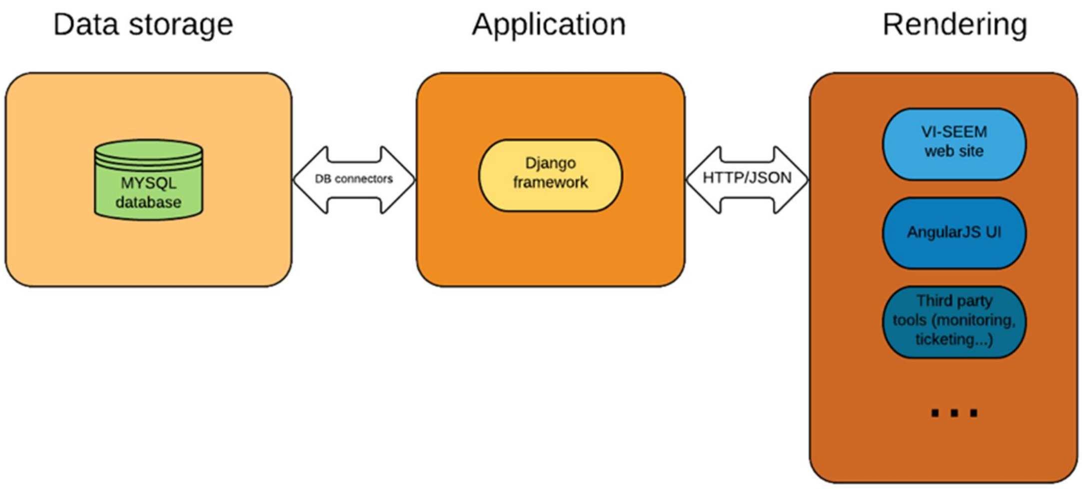


Figure 10 – High level design of the service portfolio catalogue tool.

5 Authentication and Authorization Infrastructure

The goal of WP3 is to develop an authentication and authorization infrastructure for VI-SEEM communities that is:

- secure, in order to protect valuable resources;
- distributed, in order to provide a manageable environment for communities with various types and sizes;
- easy to use for the researchers, by facilitating them to access the resources with the credentials they are used of and provide single sign-on where it is possible;
- standards-based, so that the achievements of the work package will be maintainable and interoperable.

Initial version of the project AAI will be developed by PM06, it will be tested during PM07-PM08, and ready to be used in production by PM09.

5.1 Overview of available technologies

5.1.1 SAML

Security Assertion Markup Language is an XML-based, OASIS data format for exchanging authentication and authorization data between parties, in particular, between an identity provider (IdP) and a service provider (SP). Products supporting SAML version 2 are deployed extensively both at governments, higher education and commercial enterprises worldwide.
Even though the standard was designed to work for all kinds of applications, in practice the *Web Browser SSO Profile* what is mostly deployed, thus the support for non web-based applications is problematic.

5.1.2 Identity Federations and eduGAIN

SAML was designed to facilitate authentication and authorization inside and between administrative domains. However, in order to be able to rely on SAML protocol messages between different administrative domains, trust relationships between the entities must be set up.
While it is technically possible to maintain bilateral trust links, it is exponentially proportional to the number of the entities. To address this, a set of entities may trust a third party to do the necessary entity verification and vetting procedures and maintain the SAML2 metadata, which contains verified and trustworthy information about the entities. In the Higher Education and Research world, the National Research and Education Network (NREN) provider is one of the natural choices for such a third party, therefore national Identity Federations running SAML for research become quite common by now.
However, there are more and more research projects that are not limited to a single country – therefore the grounds for an international collaboration had to be developed. By 2015, 40 national identity federations worldwide signed up for eduGAIN [13], a 'confederation' (a federation of federations) now being operated by Géant.

It is worth noting that even though eduGAIN is principally used as a trust framework for SAML2 entities, there is a growing interest for making this framework suitable for other trust technologies such as Moonshot.

5.1.3 HEXAA

In Higher Education Identity Federations, usually the identities are provided by traditional organizations, such as universities, research institutes, libraries etc. This means that joining such an organization as an individual is normally a formal and well-defined process.

However, some research communities have provided use cases that were stretching the limits of the identity federations. They need information that is relevant to the person’s role within the community and not to the organizational identity, therefore this information should not be managed and provided by the home organization or by the Identity Provider. All the same, the communities still rely on the federations for authentication as a minimum – therefore the trend is to create collaborative environments for identity federations. These environments can be implemented by building external attribute providers that are integrated to existing federations.

HEXAA (Higher Education eXternal Attribute Authority) was designed to be a collaboration platform that implements SAML2 Attribute Authority interface for the Service Providers, thus they can use standard SAML2 Attribute Query protocol to fetch additional information. By doing this, the applications that have already implemented federated authorization can be integrated to HEXAA without any change.

5.1.4 X.509

X.509 is an ITU-T standard for Public Key Infrastructures (PKI). The use of asymmetric keys enables a relying party to authenticate a user with a certificate from a trusted third party (a Certification Authority). The X.509 certificate contains the public key counterpart of the user’s private key as well as information about the subject (user), the issuer name, validity period, various other attributes (extensions) and a digital signature.

IGTF (Interoperable Global Trust Federation) is a global community that provides a trust framework for grids relying on X.509 authentication. It consists of regional PMAs (Policy Management Authorities) that are responsible for accrediting Certification Authorities who are eligible for authenticating Grid users. PMAs may specify the minimum operational, identity vetting, security and documentation requirements a CA must fulfill in order to get accredited as a Grid CA. The list of accredited CAs are released regularly as a bundle. By using this bundle, every Grid user might be securely authenticated by every relying party.

5.1.5 Token Translation Service

Token Translation Service (TTS) is a general term for translating an authentication token (such as a SAML assertion) to another means of authentication such as an X.509 certificate, an SSH key or an LDAP entry. They are used extensively to provide a federated gateway for services that are not designed for federated access.

AARC project has a task that evaluates different kinds of TTS. Services include:

- *GÉANT Trusted Certificate Service* (TCS), that enables obtaining grid-enabled X.509 certificates in a self-service manner near real-time after federated authentication and authorization;
- *LDAP Façade*: a software that provisions users to LDAP after SAML login, so that LDAP-based services may be integrated directly.
- *CILogon*: a software that is capable of issuing grid-enabled X.509 certificates after SAML or OpenID authentication

5.1.6 Moonshot

Moonshot is a unifying technology for extending the benefits of federated identity to a broad range of non-web services, including cloud infrastructures, high performance computing & grid infrastructures, and other commonly deployed services including mail, file store, remote access, and instant messaging. Moonshot builds on deployed, proven technology, including:

- Strong authentication as used by eduroam (EAP/RADIUS);
- Strong authorization as used by many national federations (SAML); and
- Strong service/application integration as used by many major applications (GSSAPI or Microsoft SSPI).

Moonshot is an implementation of the IETF's Application Bridging for Federated Access Beyond web (abfab) Working Group's set of open standards. It uses trust router components to route trust between administrative domains, which is an orthogonal federation concept to both EduGAIN and IGTF.

At of the time of writing, the definition and specification of the Moonshot technology is finished, however there are some concerns that may delay deployments for production services:

- For some services it is necessary to run the most up to date (or 'bleeding edge') version, while some patches still need to be integrated to the mainline code (OpenSSH server, FreeRadius)
- A Moonshot client (identity selector and moonshot libraries) must be installed and configured at the end users' machines, while the client is supported on Windows and Linux only (no support for users with Mac OS X)
- A global trust infrastructure such as eduGAIN and IGTF still needs to be established

For the above reasons it is not currently advisable to deploy services that rely on Moonshot *only*, however this technology should be made available as an alternative where it is technically feasible.

5.2 Design and plan for implementation

5.2.1 Architecture overview

The proposed solution for VI-SEEM AAI contains the following components:

- a VI-SEEM proxy that manages connections between:
    - eduGAIN SAML IdPs,
    - SAML IdPs that are not in eduGAIN yet,
    - SAML services,

  - authorisation backend (HEXAA);
- HEXAA;
- Home for the homeless IdP;
- Shibboleth SP or SimpleSAMLphp SP for web-based services;
- Token Translation SP for non-web services.

The proposed architecture is illustrated in Figure 11.

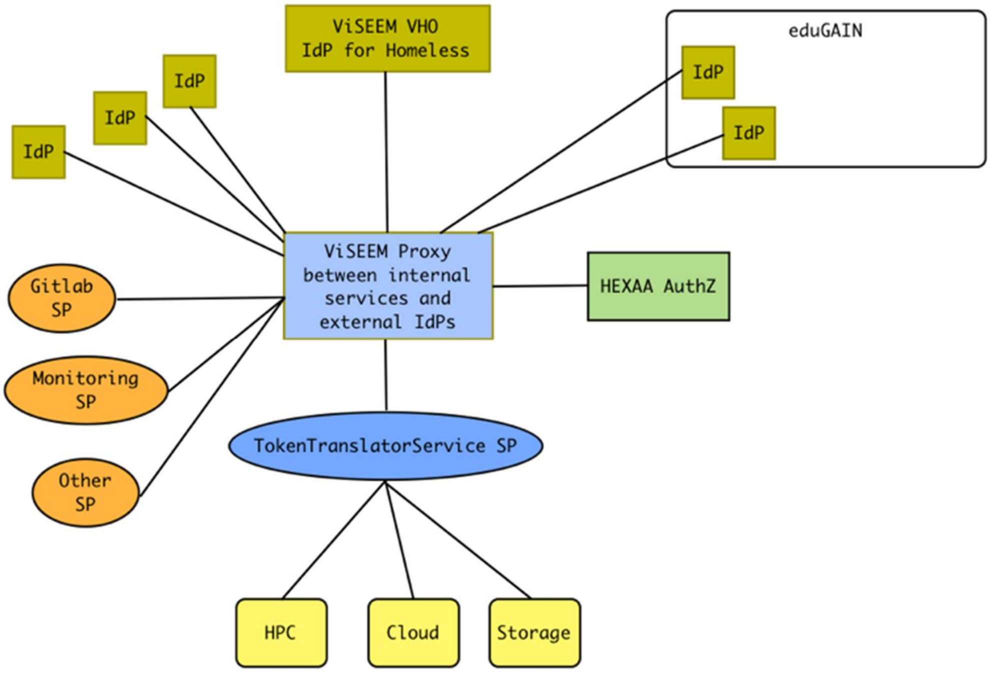


Figure 11 – Project AAI architecture.

5.2.2 Federating with eduGAIN (VI-SEEM Proxy)

EduGAIN aims to be a global trust framework, however it can only take identity federations in as members. Furthermore, every participating federation must have a written policy, which can be a burden for many developing research communities. Some of the countries participating in VI-SEEM are in eduGAIN, but others are in the process of joining, while some countries don't even have a federation.

To support both eduGAIN-enabled and other participating organizations, a mixed federation strategy is advisable. It contains a proxy that is operated by VI-SEEM, which connects the following components together:

- eduGAIN IdPs,
- VI-SEEM IdPs and VHO,
- VI-SEEM web-based services (such as GitLab and other operations tools),
- VI-SEEM Token Translation Service for enabling federated non-web access,
- authorization backend for Virtual Organizations (HEXAA).

It is obvious that the Proxy should be a Service Provider for eduGAIN, however it is still not decided whether it should provide identities as well. This would enable users outside eduGAIN to use normal eduGAIN services, which means that certain eduGAIN services could be used without any modifications. However, this would mean that the proxy should be sponsored by one of the participating national identity federations, and the policy under which the proxy would operate and issue (proxied) identities is not yet formalized.

It is also yet to be decided how the proxy should cater for non-eduGAIN IdPs, thus what are the base requirements for a member organization to push its identity provider to VI-SEEM.

The consequence of the proxy model is that VI-SEEM services should require a harmonized set of attributes, because for the IdPs they are not distinguishable, so the IdP needs to send the union of the required attributes on each login. While this is not privacy respecting in theory, in practice most of the services require the following attributes to be present:

- permanent identifier (such as eduPersonPrincipalName),
- email address,
- name (either as displayName or a surname + givenName),
- affiliation information (eduPersonScopedAffiliation or eduPersonAffiliation).

The above attributes are covered by REFEDs Research & Scholarship entity category, and as long as all VI-SEEM services are eligible for using R&S (what means they are *'Service Providers that are operated for the purpose of supporting research and scholarship interaction, collaboration or management, at least in part'*), the Proxy might have the R&S entity category assigned.

5.2.3 Home for the homeless (VHO)

While the goal is to use organisation-asserted identities where it is possible, there are a number of researchers in the VI-SEEM project who do not have an existing SAML identity provider to authenticate with. In order to let those individuals to use VI-SEEM services, a Virtual Home Organisation (VHO) must be set up.

VHO is an IdP for sponsored identities. A user might be sponsored either

- by an individual: every user with an active organisational identity may create sponsored identities;
- by a community (group): authorised personnel of groups managed in HEXAA might sponsor identities. Groups have more flexibility with sponsoring: custom expiration policy, invitation text, shared administration.

The invited user can choose whether to register a new set of credentials (username and password) or to use his or her existing social identity (Facebook, LinkedIn, Google+) for authentication. Using a social identity might have some benefits, such as no management cost for password management (recover lost passwords, etc), but some drawbacks as well.

5.2.4 Distributed authorization (HEXAA)

HEXAA is a SAML2 Attribute Authority (AA) that has three main functions:

- distributed management of community members and roles;
- allow service administrators to delegate sum permissions to certain communities;
- allow individuals to supply attributes that are not present in their organizational identity such as ORCID, SSH keys or X.509 certificates.

Technically, HEXAA contains of three components:

- SAML2 SP for managing groups and permissions;
- SAML2 AA for supplying information to services;
- an API to provide additional automatized functionality such as user provisioning.

HEXAA supplies authorization information in eduPersonEntitlement attributes. In order to adapt distributed authorization in services, the service provider must authorize the users based on eduPersonEntitlement attribute values.
Because HEXAA is also connected to the Proxy, all attributes including eduPersonEntitlement is delivered to SAML2 Service Providers as one assertion, because the Proxy does the SAML2 Attribute Request and passes the received attributes on to the services.

5.2.5 Non-web resources

The majority of the resources in VI-SEEM project are not entirely web-based, such as

- HPC resources,
- Grid resources,
- Storage resources.

Direct SAML integration of such services would be hard, as they might require SAML ECP or Moonshot technologies deployed, which are not common for production deployments. Instead, one or more Token Translation Service should be deployed and configured for allowing federated access to those web-based services. While all of the above services can be integrated by using LDAP as a basic protocol, it is recommended to have a centralized LDAP directory structure (that may be implemented by a distributed deployment of LDAP servers) that is capable of provisioning users after SAML login and place the necessary credentials (such as SSH keys, X.509 certificates, certificate DN's) as well as authorizing information (such as group membership, expiration time) there. By relying on the native LDAP integration code, which is mostly in place, it is possible to use the services in the manner the users are already used of.

On the other hand, *cloud resources* may be exceptions to the above, because more and more cloud software is able to do SAML authentication and authorization directly. There are known production-ready solutions for at least OpenStack and OpenNebula, and because direct SAML integration has several benefits (in terms of provisioning, UI integration etc), it is recommended to use SAML for clouds, where it is technically possible.

6 Conclusions

This deliverable presents a comprehensive plan for the deployment of available HPC, Cloud, Grid and storage resources, and development of services and management solutions for operation of this infrastructure. The main conclusions of this deliverable can be summarized as follows:

1. Integration of available computing and storage resources started from the first month of the project. Most of the resources are made available in PM04 (January 2016), while the rest of the resources will be integrated gradually, by PM10 (July 2016). The infrastructure will deliver 18.8 CPU, 371.6 GPU, 16.0 Xeon Phi, 5.3 IBM Cell, and 1.6 Grid CPU millions of hours per year, as well as 468 virtual machine cores to the scientific communities. Partner responsible for coordinating the integration is IICT-BAS.
2. Set of operational tools for resource monitoring and management will be deployed, and user support will be provided. These tools will be in production by PM09 (Jun 2016), and partner IPB is responsible for this task.
3. Service management and discovery will be provided via service registry – a formal portfolio of services offered by the project. The registry will be deployed by PM09, and partner UKIM is responsible for all related activities. Beta version of the service will be deployed by PM06.
4. Authentication and authorization infrastructure will be deployed within the WP3 by PM09. It will allow users to be authenticated using a single identity, and provide service for uniform managing of access to the resources regardless of their type and location. Beta version of the service will be deployed by PM06. Partner responsible for this task is NIIFI.
5. The specific operational solutions were chosen so as to insure compatibility with related project such as AARC, EUDAT2020, EGI-Engage, etc.

The deliverable confirms the strong commitment by the resource-provisioning partners (all project partners) to share substantial computing and storage resources, collaborating with all countries in SEEM region; as well as to share the responsibility for common operation of these resources, supported by a number of operational/management tools.

Appendix A Technical specification of HPC resources

Appendix A.1 ARIS

ARIS (Advanced Research Infromation System) is an HPC cluster based on IBM's NeXtScale platform, incorporating the Intel ® Xeon ® E5 v2 processors, (Ivy Bridge) and has a theoretical peak performance (Rpeak) of 190,85 TFlops and a sustained performance (Rmax) of 179,73 TFlops on the Linpack benchmark. With a total of 426 compute nodes that incorporate 2, 10 core CPUs (Ivy Bridge - Intel Xeon E5-2680v2, 2.8 GHZ), it will offer more than 8500 processor cores (CPU cores) interconnected through FDR Infiniband network, a technology offering very low latency and high bandwidth. Each compute node offers 64 Gbyte of RAM. In addition, the system offers about 1 Petabyte (quadrillion bytes) of storage, based on the IBM General Parallel File System (GPFS). The system software allows developing and running scientific applications and provides several pre-installed compilers, scientific libraries and popular scientific application suites. Other technical information is provided in the table below.

Administrative Data	
Name	ARIS
Short Description	Greek Tier-1 HPC System
Owner	GRNET S.A.
Country	Greece
Computational Power	
Number of servers	426
Server specification	IBM NeXtScale nx360 M4
CPU per server	2
RAM per server	64 GB
Total number of CPU-cores	8,520
Max number of parallel processes	8,520
Interconnect type	FDR-14 Infiniband
Interconnect latency	2.5 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	IBM GPFS
Total storage (TB)	1 TB
Accelerators type	-
Number of cores	-
Accelerators per server	-
Servers equipped with accelerators	-
Peak performance CPU (Tflops)	190.85
Peak performance accelerators (Tflops)	-
Peak performance (Tflops)	190.85
Real performance (Tflops)	179.73
Operating system	RedHat Enterprize Linux
Version	6.6

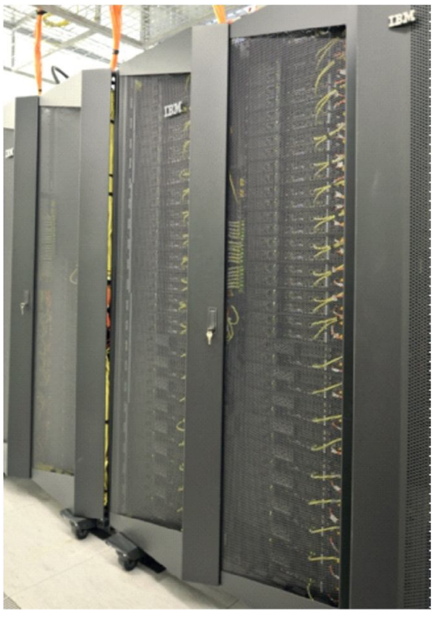
Batch system/scheduler	SLURM
Development tools	intel,pgi,gnu, intelmpi, openmpi, gdb,gdb-ia,pgdbg,ddd, VTune,Scalasca,mpiP,gprof,pgprof
Libraries	ACML, ATLAS, BOOST, ElmerFEM, ELPA, FFTW, GSL, libFLAME, Libint, Libxc, METIS, MKL, OPENBLAS, PARMETIS, SCALAPACK, Voro++, GLPK, JasPer, NETCDF, HDF5, UDUNITS, MED
Applications	ABinit, ABySS, BigDFT, CP2K, Desmond, GAMESS US, GROMACS, LAMMPS, MDynaMix, MPQC, NAMD, NWChem, Octopus, OpenMD, PLUMED, Quantum ESPRESSO, WRF, WRF-CHEM, Code Saturne, OpenFOAM
Dedication to VI-SEEM	
CPU (percent)	5%
Storage (percent)	5%
Accelerators (percent)	-
CPU (core-hours per year)	3,000,000
Storage in TB	50
Accelerators (hours per year)	-
Integration	
System operational since	Jul 15
Available to the project from	PM01
Expected date system to be phased	N/A
Interfaces	SSH
Photo	

Appendix A.2 Cy-Tera

Cy-Tera (shown in Figure) is a hybrid CPU/GPU HPC cluster composed of 116 iDataPlex dx360 M3 nodes and a theoretical peak performance of 30.5 TFlops. 98 of these are twelve core compute nodes and 18 of these are GPU nodes with dual NVidia M2070 GPUs. Each node has 48 GB of memory and 4xDR Infiniband network for MPI and I/O to the 300 TB GPFS filesystem. Other technical information is provided in the table below.

Administrative Data	
Name	Cy-Tera
Short Description	Hybrid CPU/GPU HPC cluster
Owner	The Cyprus Institute
Country	Cyprus
Computational Power	
Number of servers	116
Server specification	iDataPlex dx360 M3
CPU per server	12
RAM per server	48 GB
Total number of CPU-cores	1,392
Max number of parallel processes	1,392
Interconnect type	QDR Infiniband
Interconnect latency	1.2 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	GPFS
Total storage (TB)	260 TB
Accelerators type	NVidia M2070 GPUs
Number of cores	16,128
Accelerators per server	2
Servers equipped with accelerators	18
Peak performance CPU (Tflops)	12
Peak performance accelerators (Tflops)	18
Peak performance (Tflops)	30
Real performance (Tflops)	30
Operating system	CentOS
Version	6.6
Batch system/scheduler	Slurm
Development tools	http://cytera.cyi.ac.cy/index.php/resources/software.html
Libraries	http://cytera.cyi.ac.cy/index.php/resources/software.html
Applications	http://cytera.cyi.ac.cy/index.php/resources/software.html
Dedication to VI-SEEM	
CPU (percent)	15%
Storage (percent)	15%
Accelerators (percent)	15%
CPU (core-hours per year)	1,829,088
Storage in TB	20 TB
Accelerators (hours per year)	21,192,192
Integration	
System operational since	May 2012

Available to the project from	PM09
Expected date system to be phased	N/A
Interfaces	SSH
Photo	

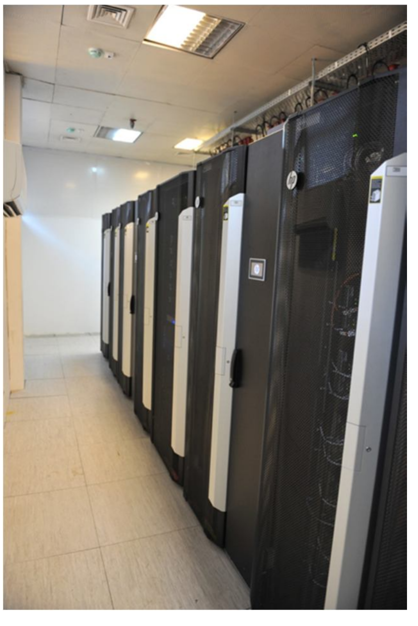


Appendix A.3 Avitohol

The Supercomputer System Avitohol at IICT-BAS consists of 150 HP Cluster Platform SL250S GEN8 servers, each one equipped with 2 Intel Xeon E5-2650 v2 8C 2600 GHz CPUs and two Intel Xeon Phi 7120P co-processors. Six management nodes control the cluster, with 4 of them dedicated to the provision of access to the storage system through Fibre Channel. The storage system is HP MSA 2040 SAN with a total of 96 TB of raw disk storage capacity. All the servers are interconnected with fully non-blocking FDR Infiniband, using a fat-tree topology. The HP CMU is used for fabric management, together with the torque/Moab combination for local resource management. Most of the computing capacity of the system comes from the Intel Xeon Phi 7120P co-processors, which use the Multiple Integrated Core (MIC) technology. For optimum use of these resources Intel compilers and the Intel MKL are deployed. Since this supercomputer is relatively new, it is in the process of deploying the software and libraries and streamlining the computational environment. Other technical information is provided in the table below.

Administrative Data	
Name	Avitohol
Short Description	Bulgarian multifunctional high performance computing cluster
Owner	IICT-BAS
Country	Bulgaria
Computational Power	
Number of servers	150
Server specification	HP ProLiant SL250s Gen8
CPU per server	2
CPU type	Intel Xeon E5-2650v2 8C 2.6GHz
RAM per server	64 GB
Total number of CPU-cores	2,400
Max number of parallel processes	4,800
Interconnect type	FDR InfiniBand
Interconnect latency	1.1 µs
Interconnect bandwidth	56 Gbps
Local filesystem type	Lustre
Total storage (TB)	96
Accelerators type	Intel Xeon Phi 7120P
Number of cores	61
Accelerators per server	2
Servers equipped with accelerators	150
Peak performance CPU (Tflops)	50
Peak performance accelerators (Tflops)	362
Peak performance (Tflops)	412
Real performance (Tflops)	264
Operating system	Red Hat Enterprise Linux for HPC Compute Node
Version	6.7 (Santiago)
Batch system/scheduler	Torque/Moab

Development tools	Intel Compilers (C/C++, FORTRAN), GNU Compilers, OpenMPI, CUDA, TotalView, Scalasca, TAU, gprof, gdb, pgdbg, Program Database Toolkit
Libraries	Intel MKL, HDF5, FFTW, NetCDF, GSL, LAPACK Boost, BLAS
Applications	Gromacs, NAMD, Desmond, VMD
Dedication to VI-SEEM	
CPU (percent)	10%
Storage (percent)	5%
Accelerators (percent)	10%
CPU (core-hours per year)	2,102,400
Storage in TB	5
Accelerators (hours per year)	16,030,800
Integration	
System operational since	Jun 2015
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	SSH, gridFTP
Photo	



Appendix A.4 PARADOX

Fourth major upgrade of PARADOX installation (Paradox IV) became operational during September 2013. This upgrade consists of 106 working nodes and 3 service nodes. Working nodes (HP ProLiant SL250s Gen8, 2U height) are configured with two Intel Xeon E5-2670 8-core Sandy Bridge processors, at a frequency of 2.6 GHz and 32 GB of RAM (2 GB per CPU-core). The total number of new processor-cores in the cluster is 1696. Each working node contains an additional GP-GPU card (NVIDIA Tesla M2090) with 6 GB of RAM. With a total of 106 NVIDIA Tesla M2090 graphics cards, PARADOX is a premier computer resource in the wider region, which provides access to a large production GPU cluster and new technology. The peak computing power of PARADOX is 105 TFlops. Other technical information is provided in the table below.

One service node (HP DL380p Gen8), equipped with an uplink of 10 Gbps, is dedicated to cluster management and user access (gateway machine). All cluster nodes are interconnected via Infiniband QDR technology, through a non-blocking 144-port Mellanox QDR Infiniband switch. The communication speed of all nodes is 40 Gbps in both directions, which is a qualitative step forward over the previous (Gigabit Ethernet) PARADOX installation. The administration of the cluster is enabled by an independent network connection through the iLO (Integrated Lights-Out) interface integrated on motherboards of all nodes.

PARADOX cluster is installed in four water-cooled racks. The cooling system consists of 4 cooling modules (one within each rack), which are connected via a system of pipes with a large industrial chiller and configured so as to minimize power consumption.

Administrative Data	
Name	PARADOX
Short Description	Serbian supercomputing cluster
Owner	Institute of Physics Belgrade (IPB)
Country	Serbia
Computational Power	
Number of servers	106
Server specification	HP ProLiant SL250s Gen8
CPU per server	2
RAM per server	32 GB
Total number of CPU-cores	1696
Max number of parallel processes	With Hyper-Threading enabled 3392
Interconnect type	QDR Infiniband
Interconnect latency	1.15 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	Lustre
Total storage (TB)	96
Accelerators type	NVIDIA TESLA M2090
Number of cores	512
Accelerators per server	1
Servers equipped with accelerators	106
Peak performance CPU (Tflops)	35.3
Peak performance accelerators (Tflops)	70.5
Peak performance (Tflops)	105.8

Real performance (Tflops)	N/A
Operating system	Scientific Linux
Version	6.4 (Carbon)
Batch system/scheduler	Torque/Maui
Development tools	Intel Compilers (C/C++, FORTRAN), Portland Group Compilers (Fortran/C/C++ with accelaterator support for CUDA, OpenACC, OpenCL), GNU Compilers, OpenMPI, CUDA, TotalView, Scalasca, TAU, gprof, gdb, pgdbg, Program Database Toolkit, ANTLR3 C
Libraries	Intel MKL, HDF5, FFTW, NetCDF, GSL, LAPACK, Boost, BLAS
Applications	Gromacs, NAMD, Desmond, VMD, AgroTagger
Dedication to VI-SEEM	
CPU (percent)	5%
Storage (percent)	10%
Accelerators (percent)	5%
CPU (core-hours per year)	742,848
Storage in TB	10
Accelerators (hours per year)	47,542,272
Integration	
System operational since	Sep 2013
Available to the project from	PM01
Expected date system to be phased	N/A
Interfaces	SSH, gridFTP
Photo	

Appendix A.5 NIIFI SC

This is one of the smaller systems NIIF is operating since 2012, having 64 Opteron 12-core CPUs. This is one of the two machines operated at NIIF HQ, Budapest. The system is integrated to PRACE European HPC network. Other technical information is provided in the table below.

Administrative Data	
Name	NIIFI SC
Short Description	Hungarian HPC cluster Budapest site
Owner	NIIF Institute
Country	Hungary
Computational Power	
Number of servers	32
Server specification	HP CP4000BL
CPU per server	2
RAM per server	66 GB
Total number of CPU-cores	768
Max number of parallel processes	1536
Interconnect type	QDR Infiniband
Interconnect latency	2.5 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	IBRIX
Total storage (TB)	50
Accelerators type	N/A
Number of cores	N/A
Accelerators per server	N/A
Servers equipped with accelerators	N/A
Peak performance CPU (Tflops)	6
Peak performance accelerators (Tflops)	N/A
Peak performance (Tflops)	6
Real performance (Tflops)	5
Operating system	Red Hat Enterprise Linux
Version	6
Batch system/scheduler	SLURM
Development tools	Intel SDK
Libraries	Intel MKL, etc.
Applications	Misc.
Dedication to VI-SEEM	
CPU (percent)	5%
Storage (percent)	0%
Accelerators (percent)	
CPU (core-hours per year)	421,882
Storage in TB	
Accelerators (hours per year)	
Integration	
System operational since	2013
Available to the project from	PM04

Expected date system to be phased out	2018
Interfaces	SSH, gsiSSH
Photo	

Appendix A.6 Leo

The HPC is the newest machines named after Leo Szilard, a Physicist and inventor of the nuclear reactor. The Top500 qualified heavily accelerated machine has 252 Nvidia GPUs to enable running highly accelerated codes to help the 168 Sandy Bridge 8-core CPUs. The system is located at Debrecen, the second largest city in Hungary, and shown in Figure. The machine is integrated to PRACE European HPC network. Other technical information is provided in the table below.

Administrative Data	
Name	Leo
Short Description	Hungarian HPC cluster
Owner	NIIF Institute
Country	Hungary
Computational Power	
Number of servers	84
Server specification	HP SL250S
CPU per server	2
RAM per server	125 GB
Total number of CPU-cores	1,344
Max number of parallel processes	2,688
Interconnect type	FDR Infiniband
Interconnect latency	2.5 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	Lustre
Total storage (TB)	585
Accelerators type	Nvidia K20x and K40x misc.
Number of cores	2688
Accelerators per server	3
Servers equipped with accelerators	84
Peak performance CPU (Tflops)	
Peak performance accelerators (Tflops)	
Peak performance (Tflops)	248
Real performance (Tflops)	208
Operating system	Red Hat Enterprise Linux
Version	6.5
Batch system/scheduler	SLURM
Development tools	Intel SDK
Libraries	Intel MKL, etc.
Applications	Maple, Matlab, etc.
Dedication to VI-SEEM	
CPU (percent)	10%
Storage (percent)	0%
Accelerators (percent)	10%
CPU (core-hours per year)	588,672
Storage in TB	0
Accelerators (hours per year)	125,269,402
Integration	

System operational since	2015
Available to the project from	PM04
Expected date system to be phased out	N/A
Interfaces	SSH
Photo	

Appendix A.7 InfraGRID

InfraGRID Cluster consists of 50 compute nodes powered by IBM BladeCenter-H technology. The entire solution is built up from 4 blade center chassis, each with 14 HS21 blade server. Each blade server has a dual quad core Intel Xeon E5504 CPU (clocked at 2.00Ghz) and 10GB RAM memory. The connectivity is delivered by: (a) Infiniband for interconnect and storage, (b) fiber channel for dedicated storage and (c) Gigabit for service networking. This cluster was initially setup in 2009. In 2011 the InfraGRID Cluster received an update by adding a new IBM BladeCenter-H chassis powered with 7 dual CPU/GPU HS22 blade servers. The CPU is powered by Intel XEON technology with a clock speed at 3.46Ghz and 32GB RAM memory. The GPU cards are NVidia Tesla M2070Q (448 GPU cores and 6GB GDDR5 RAM memory). All hardware details are available in the table below.

The cluster is managed by two service nodes, one dedicated to user access (also called head node) and one dedicated exclusively for service actions and cluster management. The storage is shared using GPFS file system over Infiniband using two dedicated NSD nodes. The service management is conducted using IBM BladeCenter the advanced management module (AMM) that is built in into both blade center chassis and blade servers (IMM). This feature allows remote administration and monitoring of any of the installed hardware. Infragrid Cluster is cooled with air. The displacement of the cooling units is in-row with a well delimited hot-cold area. Currently three APC InRow cooling units are installed and working in a cluster configuration to obtain cooling unit high-availability behavior.

Administrative Data	
Name	InfraGRID Cluster
Short Description	UVT HPC Center – InfraGRID Cluster
Owner	Universitatea de Vest din Timisoara (UVT)
Country	Romania
Computational Power	
Number of servers	57
Server specification	IBM BladeCenter HS21
CPU per server	2
RAM per server	10 GB
Total number of CPU-cores	400
Max number of parallel processes	800
Interconnect type	QDR Infiniband
Interconnect latency	2.5 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	GPFS
Total storage (TB)	50
Accelerators type	NVIDIA TESLA M2070Q
Number of cores	448
Accelerators per server	1
Servers equipped with accelerators	7
Peak performance CPU (Tflops)	3.5
Peak performance accelerators (Tflops)	3.5
Peak performance (Tflops)	2.11

Real performance (Tflops)	3.5
Operating system	Red Hat Enterprise Linux
Version	6
Batch system/scheduler	LoadLeveler
Development tools	Intel Compilers (C/C++, FORTRAN), GNU Compilers, OpenMPI, CUDA, gdb, pgdbg
Libraries	Intel MKL, HDF5, FFTW, NetCDF, GSL, LAPACK, Boost, BLAS
Applications	Misc.
Dedication to VI-SEEM	
CPU (percent)	10%
Storage (percent)	0%
Accelerators (percent)	20%
CPU (core-hours per year)	798,912
Storage in TB	0
Accelerators (hours per year)	5,494,272
Integration	
System operational since	Sep 2009
Available to the project from	PM04
Expected date system to be phased out	N/A
Interfaces	SSH
Photo	

Appendix A.8 ICAM BlueGene/P

ICAM BlueGene/P is a supercomputer designed for highly scalable applications. Our infrastructure is based on one IBM BlueGene/P rack with 1,024 physical CPUs (4,096 cores) and more than 1TB of RAM memory. The computing power is backed by a 19TB (RAID 5) storage system delivered using a 10Gbps Ethernet connection. The full specifications of the supercomputer are described in the table below.

ICAM BlueGene/P is managed by two special nodes, one dedicated to users (the head node) and one dedicated to management and monitoring activities. The storage is exported to the supercomputer by two NSD nodes connected directly to the storage system. The storage space is exported using IBM GPFS solution. ICAM BlueGene/P is cooled by air using two dedicated cooling units. The cooling model splits the cooling zone into two areas: hot zone above the technical floor and cold zone under it. The two cooling units are working in cluster mode to ensure high-availability.

Administrative Data	
Name	ICAM BlueGene/P
Short Description	UVT HPC Center – ICAM BlueGene/P Supercomputer
Owner	Universitatea de Vest din Timisoara (UVT)
Country	Romania
Computational Power	
Number of servers	1,024
Server specification	PowerPC 850
CPU per server	1
RAM per server	4 GB
Total number of CPU-cores	4,096
Max number of parallel processes	4,096
Interconnect type	IBM 3D Torus Interconnect
Interconnect latency	<1 μs
Interconnect bandwidth	40 Gbps
Local filesystem type	GPFS
Total storage (TB)	19
Accelerators type	N/A
Number of cores	N/A
Accelerators per server	N/A
Servers equipped with accelerators	N/A
Peak performance CPU (Tflops)	13.1
Peak performance accelerators (Tflops)	N/A
Peak performance (Tflops)	13.1
Real performance (Tflops)	11.2
Operating system	SLES
Version	10
Batch system/scheduler	LoadLeveler
Development tools	IBM XL (C, Fortran), OpenMPI
Libraries	Misc.
Applications	Misc.
Dedication to VI-SEEM	

CPU (percent)	20%
Storage (percent)	0%
Accelerators (percent)	0%
CPU (core-hours per year)	7,176,192
Storage in TB	0
Accelerators (hours per year)	N/A
Integration	
System operational since	Sep 2013
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	SSH
Photo	

Appendix A.9 UPT-HPC

HPC resources of UPT were implemented from an initiative of Albania Ministry of Infrastructure in collaboration with RP of China. It includes three blocks of 8 blade servers and two separate servers interconnected through two Ethernet switch of 1 Gbps. In total the system has 208 cores. Only two of blade blocks are active, the third one is not in use due to technical problems, leaving only 144 active cores. Interconnection is done using two Ethernet switches of 1 Gbps. Two local Ethernet segments are configured, the internal one used for MPI exchanges, the other external for Internet access. Only two separate servers may be accessed from outside using SSH, serving as frontal to access the system. Currently, only minimal software for parallel processing is installed: GCC and MPICH-2. Other open source software may be installed if requested. Other technical information is provided in the table below.

Administrative Data	
Name	UPT-HPC
Short Description	Polytechnic University of Tirana HPC Site
Owner	Polytechnic University of Tirana
Country	Albania
Computational Power	
Number of servers	18
Server specification	Blade servers SUGON
CPU per server	2 x Intel Quad-core E5506 2.13 GHz
RAM per server	4 GB DDR3 1.333GHz
Total number of CPU-cores	144
Max number of parallel processes	144
Interconnect type	Ethernet
Interconnect latency	
Interconnect bandwidth	1 Gbps
Local filesystem type	NFS
Total storage (TB)	18 x 146 GB 10K RPM
Accelerators type	N/A
Number of cores	N/A
Accelerators per server	N/A
Servers equipped with accelerators	N/A
Peak performance CPU (Tflops)	
Peak performance accelerators (Tflops)	
Peak performance (Tflops)	
Real performance (Tflops)	
Operating system	Scientific Linux
Version	6.5
Batch system/scheduler	Torque/Hydra
Development tools	GNU Compiler, MPICH-2
Libraries	
Applications	
Dedication to VI-SEEM	
CPU (percent)	10%
Storage (percent)	10%

Accelerators (percent)	N/A
CPU (core-hours per year)	126,144
Storage in TB	1.5
Accelerators (hours per year)	N/A
Integration	
System operational since	2013
Available to the project from	PM04
System to be phased out	N/A
Interfaces	SSH
Photo	

Appendix A.10 MK-03-FINKI

MK-03-FINKI Cluster consists of 84 compute nodes powered by HP BladeSystem c7000 technology. The entire solution is built up from 2 BladeSystem chassis, each with 32 HP BL2x220c G7 blade servers. Each blade server has a dual six core Intel Xeon L5640 CPU (clocked at 2.267Ghz) and 24GB RAM memory. The connectivity is delivered by: (a) QDR Infiniband for interconnect and storage and (b) Gigabit for service networking. This cluster was initially setup in 2012. All hardware details are available in the table below.

The cluster is managed by six service nodes, one dedicated to user access (also called head node), one dedicated exclusively for service actions and cluster management and four dedicated for the storage. The storage is shared using Lustre file system over Infiniband using two dedicated nodes for MGS/MDS, and two dedicated for OSS, all in HA mode. The service management is conducted using HP BladeSystem Onboard administration. This feature allows remote administration and monitoring of any of the installed hardware. MK-03-FINKI is cooled using HP Modular Cooling system. The displacement of the cooling unit is between the self-cooled racks.

Administrative Data	
Name	MK-03-FINKI
Short Description	Macedonia national HPC system
Owner	Ss. Cyril and Methodius University in Skopje
Country	Macedonia
Computational Power	
Number of servers	64
Server specification	Blade servers HP BL2x220c
CPU per server	2 x Intel Six-core L5640 2.267 GHz
RAM per server	24 GB
Total number of CPU-cores	768
Max number of parallel processes	1536
Interconnect type	QDR Infiniband
Interconnect latency	2 µs
Interconnect bandwidth	40 Gbps
Local filesystem type	Lustre, NFS
Total storage (TB)	4TB, 5TB
Accelerators type	N/A
Number of cores	N/A
Accelerators per server	N/A
Servers equipped with accelerators	N/A
Peak performance CPU (Tflops)	6.9
Peak performance accelerators (Tflops)	N/A
Real performance CPU (Tflops)	5.9
Real performance accelerators (Tflops)	N/A
Operating system	Scientific Linux
Version	5
Batch system/scheduler	Torque/MAUI
Development tools	GNU Compiler, MPICH-2
Libraries	
Applications	

Dedication to VI-SEEM	
CPU (percent)	5%
Storage (percent)	10%
Accelerators (percent)	N/A
CPU (core-hours per year)	336,384
Storage in TB	1+1
Accelerators (hours per year)	N/A
Integration	
System operational since	Feb 2012
Available to the project from	PM04
System to be phased out	N/A
Interfaces	SSH
Photo	
	

Appendix A.11 Armcluster

In 2004, in the Institute for Informatics and Automation Problems of NAS RA (IIAP) the first high Performance computing cluster (Armenian Cluster - Armcluster) in Armenia had been developed, which consists of 128 Xeon 3.06 GHz (64 nodes) processors. Myrinet high bandwidth and Gigabit networks interconnect the nodes of the cluster. The Myrinet network is used for computation and Gigabit for task distribution and management. The cluster achieved 523.4 GFlops performance by HPL (High Performance Linpack) test. Many intellectual software packages to support the advance in the field of modeling and analysis of quantum systems, signal and image processing, theory of radiation transfer, calculation of time constants for bimolecular chemical reactions, a system of mathematically proved methods, fast algorithms and programs for solving of certain classes of problems in linear algebra, calculus, algebraic reconditibility, test-checkable design of the built-in control circuits have been developed. In addition, user friendly tools and scientific gateways have been developed for Armcluster. Other technical information is provided in the table below.

Administrative Data	
Name	Armcluster
Short Description	Armenian Cluster
Owner	Institute for Informatics and Automation Problems of NAS RA
Country	Armenia
Computational Power	
Number of servers	64
Server specification	Intel Xeon 3.06.GHz
CPU per server	2 x Intel Xeon 3.06.GHz
RAM per server	2 GB
Total number of CPU-cores	128
Max number of parallel processes	128
Interconnect type	Myrinet2000
Interconnect latency	3.2 us
Interconnect bandwidth	4 Gbps
Local filesystem type	NFS
Total storage (TB)	40 GB
Accelerators type	N/A
Number of cores	N/A
Accelerators per server	N/A
Servers equipped with accelerators	N/A
Peak performance CPU (Tflops)	783.36
Peak performance accelerators (Tflops)	N/A
Peak performance (Tflops)	N/A
Real performance (Tflops)	N/A
Operating system	Scientific Linux
Version	6.5
Batch system/scheduler	Torque/Maui
Development tools	GNU Compiler, MPICH-2
Libraries	Blas, ScalaPack, MKL
Applications	GROMACS, NAMD, WRF, SCALAPACK

Dedication to VI-SEEM	
CPU (percent)	10%
Storage (percent)	10%
Accelerators (percent)	N/A
CPU (core-hours per year)	112,128
Storage in TB	2
Accelerators (hours per year)	N/A
Integration	
System operational since	2005
Available to the project from	PM04
System to be phased out	N/A
Interfaces	SSH
Photo	

Appendix A.12 BA-HPC

The Bibliotheca Alexandrina (BA) has been operating a High-Performance Computing (HPC) cluster since August 2009. The goal of this initiative is to provide the computational resources needed for modern scientific research in the various domains as a merit-based service to researchers locally and regionally. The cluster consists of 130 compute nodes, providing a total of 1,040 CPU cores, each with access to 1 GB of RAM. Storage for input and output data is provided by a Lustre file system hosted on storage hardware with a total raw capacity of 36 TB. The cluster is wired with 10-Gbps DDR InfiniBand. The BA-HPC participated in the LinkSCEEM-2 project and continues to participate in joint calls with the Cy-Tera cluster operated by the Cyprus Institute. The majority of usage on the system comes from projects by researchers at Egyptian universities. In the VI-SEEM project, the BA is dedicating 20 percent of the system, i.e., approximately 1.8 million core hours yearly, for hosting projects that will be granted access to HPC resources through VI-SEEM. In addition, on the BA large-scale storage cluster, 100 TB are being dedicated to the VI-SEEM project. Other technical details are provided in the table below.

Administrative Data	
Name	BA-HPC
Short Description	The Bibliotheca Alexandrina High-Performance Computing Cluster
Owner	Bibliotheca Alexandrina
Country	Egypt
Computational Power	
Number of servers	130
Server specification	Sun Blade X6250 Server Module
CPU per server	2
RAM per server	64 GB
Total number of CPU-cores	1,040
Max number of parallel processes	1,040
Interconnect type	DDR Infiniband
Interconnect latency	3.3 μs
Interconnect bandwidth	10 Gbps
Local filesystem type	Lustre
Total storage (TB)	36
Accelerators type	N/A
Number of cores	N/A
Accelerators per server	N/A
Servers equipped with accelerators	N/A
Peak performance CPU (Tflops)	11.8
Peak performance accelerators (Tflops)	N/A
Peak performance (Tflops)	11.8
Real performance (Tflops)	9.1
Operating system	CentOS
Version	6.7
Batch system/scheduler	Open Grid Scheduler
Development tools	GCC
Libraries	Managed via EasyBuild

Applications	Managed via EasyBuild
Dedication to VI-SEEM	
CPU (percent)	20%
Storage (percent)	20%
Accelerators (percent)	N/A
CPU (core-hours per year)	1,822,080
Storage in TB	7.2
Accelerators (hours per year)	N/A
Integration	
System operational since	Aug 2009
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	SSH
Photo	
	

Appendix A.13 Gamma

IMAN1 (Jordan) as supercomputing center has diversity of HPC cluster architectures (Gama, Zaina, Booster/King) which are made available for VI-SEEM project. One of them is Gamma. Gamma is Intel Xeon based hybrid CPU/GPU computing node equipped with NVIDIA Tesla K20 GPU card. This computing node is used for data visualization and pattern / object detection. Mainly, this computing node is used for academic and research purposes. Technical details are given in the table below.

Administrative Data	
Name	Gama
Short Description	IMAN1-01
Owner	SESAME
Country	Jordan
Computational Power	
Number of servers	1
Server specification	IBM M4 dx360
CPU per server	2
RAM per server	32 GB
Total number of CPU-cores	4
Max number of parallel processes	8
Total storage (TB)	600 GB
Accelerators type	NVIDIA TESLA K20
Number of cores	2496 CUDA Cores
Accelerators per server	1
Servers equipped with accelerators	1
Operating system	Scientific Linux
Version	6.4 (Carbon)
Batch system/scheduler	N/A (to be installed)
Dedication to VI-SEEM	
CPU (percent)	100%
Storage (percent)	100%
Accelerators (percent)	100%
CPU (core-hours per year)	70,080
Storage in TB	.6
Accelerators (hours per year)	21,864,960
Integration	
System operational since	2013
Available to the project from	M08
Expected date system to be phased	N/A
Interfaces	SSH, VNC
Photo	

Appendix A.14 Zaina

Zaina is Intel Xeon based computing cluster with 1 Gbit Ethernet interconnect. This cluster is used for code development, code porting and Synchrotron Radiation application purposes. It compounds of Two Dell PowerEdge R710 and Five HP ProLiant DL140 G3 servers. Technical details are given in the table below.

Administrative Data	
Name	Zaina
Short Description	IMAN1-02
Owner	SESAME and JAEC
Country	Jordan
Computational Power	
Number of servers	7
Server specification	Two Dell PowerEdge R710 and five HP ProLiant DL140 G3
CPU per server	2
RAM per server	Dell (16 GB) HP (6 GB)
Total number of CPU-cores	4
Max number of parallel processes	Dell (16) HP (8)
Total storage (TB)	1 TB NFS share
Accelerators type	N/A
Number of cores	
Accelerators per server	
Servers equipped with accelerators	
Operating system	Scientific Linux
Version	6.4 (Carbon)
Batch system/scheduler	N/A (to be installed)
Dedication to VI-SEEM	
CPU (percent)	30%
Storage (percent)	30%
Accelerators (percent)	
CPU (core-hours per year)	147,168
Storage in TB	1 TB NFS share
Accelerators (hours per year)	
Integration	
System operational since	2014
Available to the project from	M06
Expected date system to be phased	N/A
Interfaces	SSH, VNC
Photo	

Appendix A.15 IMAN1-Booster/King

Both of booster (IMAN1-03-PS3) and King (IMAN1-04-PS3) Clusters are HPC cluster based on the IBM CELL Processor 8 C SPEs + 1 PPE. Booster cluster is a small scale cluster consisted of five PlayStations (PS3) used to run pilot projects and to make prove of concept of code porting on the IBM CELL Processor computing architecture. However, the successful jobs on booster cluster are moved to the large-scale cluster King (IMAN1-04-PS3), which is consisted of 250 PlayStation, more technical information is provided in the table below.

Administrative Data	
Name	IMAN1-Booster/King
Short Description	Booster (IMAN1-03-PS3), King (IMAN1-04-PS3)
Owner	SESAME
Country	Jordan
Computational Power	
Number PlayStation	Booster (QTY=5), King (QTY=250)
Server/ PlayStation specification	PlaySation3
CPU per server	1 PPE (Power Processing Element), 64-bit PowerPC
RAM per server	256MB XDR DRAM
Total storage (TB)	1 TB NFS share
Accelerators type	SPE (Synergistic Processing Element)
Number of cores	8 SPEs (Synergistic Processing Element)
Accelerators per server	8
Servers equipped with accelerators	Every PlayStation
Operating system	Yellow Dog Linux
Version	6.2
Batch system/scheduler	Open MPI, (to be installed)
Dedication to VI-SEEM	
CPU (percent)	50%
Storage (percent)	30%
Accelerators (percent)	30%
CPU (core-hours per year)	672,768
Storage in TB	1 TB NFS share
Accelerators (hours per year)	5,382,144
Integration	
System operational since	2011
Available to the project from	PM06
Expected date system to be phased out	N/A
Interfaces	SSH
Photo	


ROW
ROW 2

Appendix B Technical specification of Cloud resources

Appendix B.1 Okeanos

GRNET has developed its own IaaS cloud solution which is providing virtualized computing resources for free to the Greek Universities and public Research Centres that already utilize GRNET's network infrastructure and services. The cloud service called ~okeanos is operated using an in-house developed software stack called Synnefo. The software builds on top of existing proven open source software (e.g. Google Ganeti), enhanced and expanded in order to provide a robust and complete IaaS cloud solution. ~okeanos is currently used on daily basis by thousands of users, university students, researches and members of the academia.

GRNET operates ~okeanos via the company-owned data center which currently comprises 22 racks, 400+ servers and 4 Petabytes of storage. The servers were ~okeanos is hosted have the following characteristics:

- CPU: dual AMD Opteron 6172, 2100MHz, 12cores
- Memory: 192GB (24x8GB)
- Disk: 8x600GB SAS
- Network: 4x1G + IPMI (iLO)
- Size: 2RU

This infrastructure is currently being extended with additional hardware and different architectural configurations (SSDs, GPUs etc), in order to satisfy the increasing and diversifying demand for computing resources. Technical details are given in the table below.

Administrative Data	
System Name	~Okeanos
Short Description	GRNET's cloud service, for the Greek Research and Academic Community.
Owner	GRNET S.A.
Country	Greece
Computational Power	
Number of servers	400
Server specification	HP DL385G7
CPU per server	2
RAM per server	192
Total number of CPU cores	5520
Max number of VMs	10,000
Smallest VM CPU cores	1
Smallest VM RAM	1
Largest VM CPU cores	8
Largest VM RAM	8
Total storage	300 TB

Cloud environment	
Cloud environment type	synnefo.org
Interfaces	Standard Openstack
User access	Acdemic Login, Classic Accounts (API Based, Web UI)
Dedication to VI-SEEM	
Dedication % of max VMs	2%
Dedication storage %	2%
Dedication in VM hours	1,752,000
Dedication storage in TB	100
Integration	
Cloud system operational since	2010
Expected date system to be available	PM01
Expected date system to be phased	N/A
Interface to be provided	~Okeanos
Photo	
	

Appendix B.2 CyI Cloud Facility

The CyI Cloud Facility is a cloud infrastructure pilot based on OpenStack and hosted on two different systems – upon 16 IBM x3550 and upon 6 Dell PowerEdge servers with 32GB and 196GB memory per server respectively. With 176 CPU cores and 60TB storage the system will soon be available for users to use its services. Other technical information is provided in the table below.

Administrative Data	
System Name	CyI Cloud Facility
Short Description	Cloud infastructure pilot based on OpenStack
Owner	CyI
Country	Cyprus
Computational Power	
Number of servers	16+6
Server specification	16 IBM x3550, 6 Dell PowerEdge
CPU per server	8
RAM per server	32GB, 96GB
Total number of CPU cores	176
Max number of VMs	176
Smallest VM CPU cores	1
Smallest VM RAM	1 GB
Largest VM CPU cores	8
Largest VM RAM	96 GB
Total storage	40TB
Cloud environment	
Cloud environment type	OpenStack
Interfaces	-
User access	Web/SSH
Dedication to VI-SEEM	
Dedication % of max VMs	10%
Dedication storage %	10%
Dedication in VM hours	157,680
Dedication storage in TB	10%
Integration	
Cloud system operational since	N/A
Available to the project from	PM06
System to be phased out	N/A
Photo	

Appendix B.3 Avitohol

The Bulgarian supercomputer cluster Avitohol will have also ability to run VMs, using Openstack. The restriction of no-public IPs for the VMs may be lifted in the future, if there is substantial need for such kind of virtual machines for the services. Detail technical information is provided in the table below.

Administrative Data	
System Name	Avitohol
Short Description	Bulgarian multifunctional high performance
Owner	IICT-BAS
Country	Bulgaria
Computational Power	
Number of servers	150
Server specification	HP ProLiant SL250s Gen8
CPU per server	2
RAM per server	64 GB
Total number of CPU cores	2,400
Max number of VMs	2,400
Smallest VM CPU cores	1
Smallest VM RAM	4 GB
Largest VM CPU cores	16
Largest VM RAM	-
Total storage	96 TB
Cloud environment	
Cloud environment type	OpenStack
Interfaces	Standard Openstack
User access	N/A
Dedication to VI-SEEM	
Dedication % of max VMs	5%
Dedication storage %	5%
Dedication in VM hours	1,051,200
Dedication storage in TB	5 TB
Integration	
Cloud system operational since	N/A
Available to the project from	PM10
System to be phased out	N/A
Dedication storage in TB	10%
Interface to be provided	Standard Openstack
Restrictions on VMs	No public IP

Appendix B.4 InfraGRID Cloud

InfraGRID Cloud offers cloud infrastructure as a service (IaaS). The cloud infrastructure is powered by OpenStack technology. The table below describes the resources offered by InfraGRID Cloud.

The virtual image repository consists of several basic images supported by OpenStack. InfraGRID Cloud uses KVM as a virtualization technology so, in general, any HVM (hardware virtual machine) compliant image can be supported by InfraGRID Cloud.

Administrative Data	
Name	InfraGRID Cloud
Short Description	UVT HPC Center – InfraGRID Cloud Solution
Owner	Universitatea de Vest din Timisoara (UVT)
Country	Romania
Computational Power	
Number of servers	50
Server specification	IBM BladeCenter HS21
CPU per server	2
RAM per server	10
Total number of CPU-cores	400
Max number of VMs	400
Smallest VM CPU Cores	1
Smallest VM RAM	0.5
Largest VM CPU Cores	8
Largest VM RAM	8
Total storage (TB)	50
Cloud environment type	OpenStack
Dedication to VI-SEEM	
Dedication % of max VMs	12%
Storage (percent)	10%
Dedication in VM hours	402,960
Storage in TB	5
Integration	
System operational since	January 2015
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	SSH, OpenStack Standard API, OpenStack CLI

Appendix B.5 UPT-Cloud

Faculty of Information Technology of Polytechnic University of Tirana has part of the grid system created in framework of former FP7 SEE-GRID projects supported by the national programme for R&D. The plans are to use this infrastructure to set-up and integrate in the future virtualized cloud infrastructure resources. The plan includes installation of OpenStack cloud management platform supporting virtualization technologies such as Xen or KVM. The idea is to support two functionalities: (a) increased flexibility by running different application on different operating systems thus widening the support for different research communities; and (b) to open the possibility for virtual resource management for performance and energy optimizations.

The cluster consists of 8 node cluster. Each node is an HP ProLiant DL320 Server with dual-core Intel Xeon 3040, 1.86GHz, 2GB RAM + 80GB HDD and two Gigabit Ethernet cards. Total number of cores is 16. We plan to upgrade memory to 4-8 GB per node and add an additional number of 6 nodes reaching in total 14 nodes or 28 cores. We plan to setup and run the OpenStack cloud platform on Ubuntu 14.0.4 server using KVM virtualization hypervisor. Two nodes will be used for controller and storage/network service while 6 nodes as compute nodes. It will run a minimum of 12VM with (1VCPU + 1GB RAM + 30GBHDD) per VM and later the number and capacity of VMs can be increased depending on the upgrade of hardware. For the project the total available amount of cloud resource will be 6VM (1VCPU + 1GB RAM + 30GB HDD) per VM starting from month 10 of the project. In total, it will be dedicated: 6VM-cores, 52,560 VM-hours. Other technical information is provided in the table below.

Administrative Data	
Name	UPT-Cloud
Short Description	Polytechnic University of Tirana Cloud Site
Owner	Polytechnic University of Tirana
Country	Albania
Computational Power	
Number of servers	8
Server specification	HP ProLiant DL320 Generation 5
CPU per server	2
RAM per server	2
Total number of CPU-cores	16
Total number of VM cores	12
Interconnect type	Ethernet
Interconnect bandwidth	1 Gbps
Total storage (GB)	640
Operating system	Ubuntu Linux/Scientific Linux/Windows
Dedication to VI-SEEM	
CPU (percent)	50%
Storage (percent)	50%
VM CPU (VM core-hours per year)	52,560
Storage in GB	320
Integration	
System operational since	N/A
Available to the project from	PM10

Expected date system to be phased	N/A
Interfaces	Web-based, SSH
Photo	
	

Appendix B.6 ETFBL-CC01

Cloud Computing infrastructure at ETFBL-CC01 is to be used for project needs in areas of general services and applications as well as hosting services to be used for Source Code Repository, Helpdesk/Trouble Ticket System and as a Monitoring System development environment. This installation is also closely related to ETFBL-CS01 which will provide additional storage for those needs. Detail technical information is provided in the table below.

Administrative Data	
System Name	ETFBL-CC01
Short Description	University of Banja Luka Faculty of Electrical Engineering Computing Cluster
Owner	University of Banja Luka Faculty of Electrical Engineering
Country	Bosnia and Herzegovina
Computational Power	
Number of servers	3
Server specification	HP DL180
CPU per server	2 (4 cores per CPU)
RAM per server	8 GB
Total number of CPU cores	40
Max number of VMs	60
Smallest VM CPU cores	1
Smallest VM RAM	256 MB
Largest VM CPU cores	8
Largest VM RAM	4 GB
Total storage	3 TB
Cloud environment	
Cloud environment type	OpenStack
Interfaces	Standard Openstack
User access	N/A
Dedication to VI-SEEM	
Dedication % of max VMs	33%
Dedication storage %	33%
Dedication in VM hours	115,632
Dedication storage in TB	1 TB
Integration	
Cloud system operational since	N/A
Available to the project from	PM06
System to be phased out	N/A
Interface to be provided	Standard Openstack
Restrictions on VMs	No public IP

Appendix B.7 MK-04-FINKI_CLOUD

MK-04-FINKI_CLOUD is a cloud infrastructure based on OpenNebula and is hosted on 15 HP blade servers, each with 24GB RAM and 12 HT CPU cores, totaling in 432 vCPU cores and 17TB of storage. The system is in production from 2014 as a National and EGI FedCloud integrated cloud system. Other technical information is provided in the table below.

Administrative Data	
Name	MK-04-FINKI_CLOUD
Short Description	Macedonia national research Cloud
Owner	UKIM
Country	Macedonia
Computational Power	
Number of servers	18
Server specification	HP BL2x220c
CPU per server	12
RAM per server	24
Total number of CPU-cores	216
Max number of VMs	432
Smallest VM CPU Cores	1
Smallest VM RAM	0.5
Largest VM CPU Cores	12
Largest VM RAM	12
Total storage (TB)	18
Cloud environment type	OpenNebula
Dedication to VI-SEEM	
Dedication % of max VMs	10%
Storage (percent)	10%
Dedication in VM hours	210,240
Storage in TB	2
Integration	
System operational since	May 2014
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	occi, sunstone

Appendix B.8 MD-Cloud

RENAM Scientific Cloud is deployed on 2 servers using OpenStack and/or Synnefo Cloud middleware. OpenNebula was deployed during participation in the regional project "Experimental Deployment of an Integrated Grid and Cloud Enabled Environment in BSEC Countries on the Base of g-Eclipse (BSEC gEclipseGrid)". Deployed installation includes 1 server with the following parameters: 8 cores, 16 Gb RAM, 1,5 Tb Storage. OpenStack was deployed during participation in the GN3plus project service activity SA2 "Testbed as a Service" – GEANT Testbed Service (GTS, http://services.geant.net/gts). Installation includes 1 server: 8 cores, 16 Gb RAM, 1,5 Tb Storage. Technical details are given in the table below.

Administrative Data	
Country	Moldova
Name	MD-Cloud
Short Description	MD Scientific Cloud
Owner	RENAM
Computational Power	
Number of servers	2
Server specification	Intel Callahan S5000VCL (Intel Xeon 1,6 GHz)
CPU per server	2
RAM per server	16
Total number of CPU-cores	16
Max number of VMs	12
Smallest VM CPU cores	1
Smallest VM RAM	1 Gb
Largest VM CPU cores	4
Largest VM RAM	4 Gb
Total Storage	2 TB
Cloud environment	
Cloud environment type	Openstack/Synnefo
Interfaces	WEB, SSH
Dedication to VI-SEEM	
Dedication % of max VMs	25%
Dedication storage %	25%
Dedication in VM hours if smallest VM	26,280
Dedication storage	0.5 TB
Integration	
System operational since	N/A
Available to the project from	PM04
Expected date system to be phased	N/A
Restrictions on VMs	1 public IP4
Photo	

Appendix B.9 IIAP Cloud

IIAP Cloud offers cloud infrastructure as a service (IaaS) using the OpenStack middleware. The cloud infrastructure consists of 2 servers and 96 VM-cores. The virtual image repository consists of customized Oss for earth science and life science communities. Technical details are given in the table below.

Administrative Data	
Country	Armenia
Name	IIAP Cloud
Short Description	Armenian Scientific Cloud
Owner	IIAP
Computational Power	
Number of servers	2
Server specification	Intel Xeon E5-2680 v3 2.5GHz
CPU per server	4
RAM per server	128
Total number of CPU-cores	96
Max number of VMs	96
Smallest VM CPU cores	1
Smallest VM RAM	1 GB
Largest VM CPU cores	4
Largest VM RAM	16 GB
Total Storage	2 TB
Cloud environment	
Cloud environment type	Openstack
Interfaces	WEB, SSH
Dedication to VI-SEEM	
Dedication % of max VMs	10%
Dedication storage %	10%
Dedication in VM hours if smallest VM	84,096
Dedication storage	3 TB
Integration	
System operational since	N/A
Available to the project from	N/A
Expected date system to be phased	N/A
Restrictions on VMs	No public IP

Appendix B.10 IUCC InfinityCloud

The IUCC InfinityCloud Facility is a cloud infrastructure pilot based on OpenStack and hosted on 75 IBM x3450 servers with 64GB memory per server. With 560 CPU-cores and 40TB storage, the system will soon be available for users to use its services. Other technical information provided in the table below.

Administrative Data	
Country	Israel
Name	InfinityCloud
Short Description	Infiniband Cloud
Owner	IUCC
Computational Power	
Number of servers	70
Server specification	IBM System x3450 (Intel Xeon CPU E5472@ 3.00GHz)
CPU per server	8
RAM per server	64
Total number of CPU-cores	560
Max number of VMs	560
Smallest VM CPU cores	1
Smallest VM RAM	1 Gb
Largest VM CPU cores	8
Largest VM RAM	16 Gb
Total Storage	40 TB
Cloud environment	
Cloud environment type	Openstack
Interfaces	WEB, SSH
Dedication to VI-SEEM	
Dedication % of max VMs	5%
Dedication storage %	8%
Dedication in VM hours if smallest VM	245,280
Dedication storage	5 TB
Integration	
System operational since	N/A
Available to the project from	PM04
Expected date system to be phased	N/A
Restrictions on VMs	50 public IP4
Photo	


icsl1199

Appendix C Technical specification of Grid resources

Appendix C.1 Hellas Grid

The main scope of HellasGrid is the provision of High Throughput Computing services to Greek academic and research community. The HellasGrid infrastructure is composed by six (6) clusters of computational and storage resources located at Athens (HG-01-GRNET, HG-02-IASA, HG-06-EKT), Thessaloniki (HG-03-AUTH), Patras (HG-04-CTI-CEID) and Heraklion (HG-05-FORTH). Technical information is provided in the table below.

Administrative Data	
Country	Greece
Name	Hellas Grid
Short Description	GREEK NGI Infrastructure
Owner	GRNET S.A.
Computational Power	
Number of servers	100
Server specification	HP
CPU per server	2
RAM per server	64
Total number of CPU-cores	864
Max number of Jobs	1728
Local temp storage per server	100GB
Max storage available to a grid job	500GB
Grid Storage	
Type of grid storage	lustre
Total grid storage	2TB
Dedication to VI-SEEM	
Dedication % of max jobs	5%
Max number of VI-SEEM jobs	2,000
Dedication in CPU core hours per year	370,000
Dedication storage in TB from the Grid	0.5
Integration	
System operational since	N/A
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	gLite flavour of EMI

Appendix C.2 BG01-IPP

The Grid resources from Bulgaria will be provided initially from the older HPCG cluster – BG01-IPP. Technical details of BG01-IPP cluster are given in the table below. Since this cluster is towards the end of its usable lifetime, the Grid services will be later migrated to the new system – Avitohol.

From point of view of users, there will be no change in the way the services are used and the dedicated cores will be kept.

Administrative Data	
Country	Bulgaria
Name	BG01-IPP
Short Description	HP Cluster Platform Express 7000 enclosures
Owner	IICT-BAS
Computational Power	
Number of servers	36
Server specification	HP ProLiant BL280c
CPU per server	16
RAM per server	8
Total number of CPU-cores	640
Max number of Jobs	640
Local temp storage per server	40 GB
Max storage available to a grid job	2 TB
Grid Storage	
Type of grid storage	Dcache
Total grid storage	5 TB
Dedication to VI-SEEM	
Dedication % of max jobs	5%
Max number of VI-SEEM jobs	32
Dedication in CPU core hours per year	280,320
Dedication storage in TB from the Grid	1 TB
Integration	
System operational since	2010
Available to the project from	PM04
Expected date system to be phased	PM16
Interfaces	gLite flavour of EMI
Photo	

Appendix C.3 AEGIS01-IPB-SCL

First Grid site in Serbia AEGIS01-IPB-SCL is setup in Scientific Computing Laboratory, Institute of Physics in Belgrade. This Grid site is a set of 89 worker nodes (2 x quad core Xeon E5345 on 2.33 GHz with 8GB of RAM) and 15 service nodes (Xeon based nodes) fully dedicated to the GRID community. OS is Scientific Linux with EMI-gLite middleware. Technical specification is provided in the table below.

The configuration consists of CE, DPM storage on 3 nodes (27TB), sBDII and MON. VI-SEEM Grid core services are also part of AEGIS01-IPB-SCL with two WMS nodes, top level BDII, LFC, VOMS and UI machine.

Administrative Data	
Country	Serbia
Name	AEGIS01-IPB-SCL
Short Description	First and largest Grid site in Serbia, AEGIS01-IPB-SCL, located at the Scientific Computing Laboratory, Institute of Physics Belgrade
Owner	Institute of Physics Belgrade (IPB)
Computational Power	
Number of servers	88
Server specification	Supermicro SuperServer 6015T
CPU per server	2
RAM per server	8
Total number of CPU-cores	704
Max number of Jobs	704
Local temp storage per server	128 GB
Max storage available to a grid job	7.7 TB
Grid Storage	
Type of grid storage	DPM
Total grid storage	41 TB
Dedication to VI-SEEM	
Dedication % of max jobs	5%
Max number of VI-SEEM jobs	704
Dedication in CPU core hours per year	308,352
Dedication storage in TB from the Grid	5 TB
Integration	
System operational since	2007
Available to the project from	PM01
Expected date system to be phased	N/A
Interfaces	gLite flavour of EMI

Appendix C.4 MK-03-FINKI

MK-03-FINKI is a grid site provided by the UKIM. This Grid site is a set of 64 worker nodes (2 x six core Xeon L5640 on 2.267 GHz with 24GB of RAM) and 6 service nodes. OS is Scientific Linux with EMI-gLite middleware. The configuration consists of CE, DPM storage, sBDII and MON. Technical specification is provided in the table below.

Administrative Data	
Country	Macedonia
Name	MK-03-FINKI
Short Description	Macedonia largest grid site
Owner	UKIM
Computational Power	
Number of servers	64
Server specification	HP BLC2x220c
CPU per server	2
RAM per server	24
Total number of CPU-cores	768
Max number of Jobs	768
Local temp storage per server	250 GB
Max storage available to a grid job	4 TB
Grid Storage	
Type of grid storage	DPM
Total grid storage	2 TB
Dedication to VI-SEEM	
Dedication % of max jobs	5%
Max number of VI-SEEM jobs	
Dedication in CPU core hours per year	336,384
Dedication storage in TB from the Grid	2 TB
Integration	
System operational since	2012
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	gLite flavour of EMI

Appendix C.5 MREN01CIS

MREN01CIS is first academic cluster for scientific computing in Montenegro, established in 2006. It was used in several FP6 (SEE-GRID-2) and FP7 (SEE-GRID-SCI, HP SEE, EGI Inspire) projects. MREN01CIS, as a heterogeneous cluster, consist of 32 CPU cores and service nodes (CE, SE, WMS/LB, BDII, Argus, and Nagios). Cluster is situated in Data center of University of Montenegro.

MREN CA is operational from 2008 and supports certification activities for Montenegrin academic institutions. Other technical details are provided in the table below.

Administrative Data	
Name	MREN01CIS
Short Description	Montenegrin academic cluster
Owner	UoM
Country	Montenegro
Computational Power	
Number of servers	16
Server specification	Mixed infrastructure
CPU per server	2
RAM per server	2 GB
Total number of CPU-cores	32
Max number of parallel processes	32
Total storage (TB)	50 GB
Max storage available to job	10 GB
Grid Storage	
Type of grid storage	DPM
Total grid storage	2 TB
Dedication to VI-SEEM	
CPU (percent)	50%
Storage (percent)	50%
Storage in TB	1 TB
Integration	
System operational since	2006
Available to the project from	PM05
Expected date system to be phased	N/A
Interfaces	EMI
Photo	

Appendix C.6 MD-GRID

RENAM Association as the National scientific-educational network of Moldova (NREN) is coordinating Grid activities in Moldova since 2007 when it was established MD-Grid - National Grid Initiative (NGI). First GRID-Site started operation in 2008. MD-Grid consist of three GRID-Sites located in the Research and Educational Network Association of Moldova (RENAM), in the Institute of Mathematics and Computer Science of Academy of Science of Moldova (IMI ASM) and the State University of Moldova (USM). Grid site parameters:

- Total nodes: 74 CPU-cores, 69 GB of RAM, 6 TB HDD;
- Computing nodes: 40 CPU-cores, 40 GB RAM, 2 TB HDD.

Technical specification of MD-GRID site is given in the table below.

Administrative Data	
Country	Moldova
Name	MD-GRID
Short Description	MD-GRID Infrastructure
Owner	RENAM, IMI, USM
Computational Power	
Number of servers	10 (Worker Nodes)
Server specification	HP ProLiant DL140 G3, Intel Callahan S5000VCL 2.3 GHz
CPU per server	4
RAM per server	4
Total number of CPU-cores	40
Max number of Jobs	40
Local temp storage per server	16 GB
Max storage available to a grid job	0,5 TB
Grid Storage	
Type of grid storage	SE, DPM
Total grid storage	2 TB
Dedication to VI-SEEM	
Dedication % of max jobs	30%
Max number of VI-SEEM jobs	12
Dedication in CPU core hours per year	105,120
Dedication storage in TB from the Grid	0.5 TB
Integration	
System operational since	2008
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	SSH, gridFTP
Photo	

Appendix C.7 ArmCluster

IIAP coordinates the ArmNGI (Armenian National Grid Initiative), which is a national Armenian effort to establish a nationwide grid environment for computational science and research. The goal of the ArmNGI is to pursue a variety of scientific users in utilizing the Grid for their applications. All these applications rely on a wide range of diverse computer science technologies composed from standard grid middleware and sophisticated high-level extensions. Now the computational resources (about 500 cores) of Armenian Grid infrastructure distributed among our leading research (National Academy of Sciences, Yerevan Physics Institute) and academic (Yerevan State University, State Engineering University of Armenia) organizations are located in the cities of Yerevan and Ashtarak. Being the most powerful computational resource in the field of science and education in Armenia, the Armcluster constitutes the core of the Armenian grid infrastructure.

Appendix C.8 GE-01-GRENA

In 2009 Georgian Research and Educational Networking Association GRENA established the first Grid facility in Georgia GE-01-GRENA, which is included in the European Grid Infrastructure. Last upgrade of the system took place during 2013-14, when 2 working nodes (DELL PowerEdge R720) and storage system (DELL PowerEdge R720 XD) were installed. In 2013 Tbilisi State University and GRENA established Certification Authority. GE-01-GRENA site is located at GRENA data center and its elements are interconnected with Gigabit Ethernet. Data center is connected to the GRENA backbone with three 1 Gbps capacity links. During 2016 these links will be upgraded to 10 Gbps. The Data center is equipped with redundant power supply system (2 generators and 2 UPS) and cooling system. The technical information is provided in the table below.

Administrative Data	
Name	GE-01-GRENA
Short Description	Georgian Grid infrastructure
Owner	Georgian Research and Educational Networking Association
Country	Georgia
Computational Power	
Number of servers	2
Server specification	DELL PowerEdge R720
CPU per server	32
RAM per server	32
Total number of CPU-cores	64
Max number of parallel processes	20
Total storage (TB)	7.7
Accelerators type	NVIDIA TESLA M2090
Number of cores	512
Accelerators per server	1
Servers equipped with accelerators	2
Operating system	Scientific Linux
Version	6.4 (Carbon)
Batch system/scheduler	Torque/Maui
Dedication to VI-SEEM	
CPU (percent)	30%
Storage (percent)	30%
Accelerators (percent)	30%
CPU (core-hours per year)	175,200
Storage in TB	2
Accelerators (hours per year)	8,760
Integration	
System operational since	2010
Available to the project from	PM04
Expected date system to be phased	N/A
Interfaces	SSH, gridFTP
Photo	

Appendix D Technical specification of Storage resources

Appendix D.1 ARIS

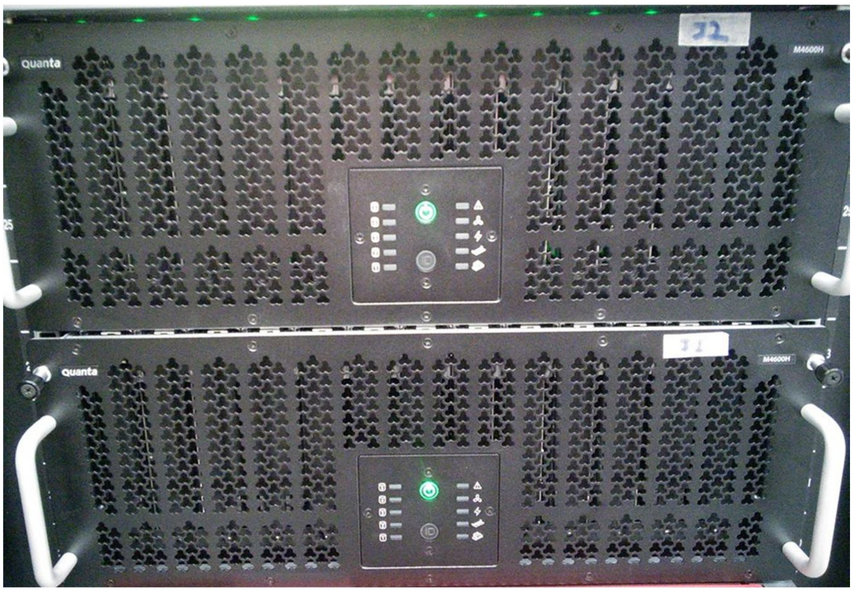
The Storage resources for specialized data storage and preservation will be provided from the ARIS supercomputer system, as shown in the table below.

Administrative Data	
System Name	ARIS
Short Description	Greek Tier-1 HPC System
Owner	GRNET S.A.
Country	Greece
Storage system	
Description of storage systems	IBM Tape Storage, IBM GPFS
Total storage	1 PB, 3 PB
Dedication to VI-SEEM	
Dedication % of available storage	7%
Dedication storage in TB	50 TB, 210 TB
Integration	
System operational since	01-Jul-15
Available to the project from	PM4, PM10
Expected date to be phased out	N/A
Interface to be provided	gridFTP

Appendix D.2 ONYX, Cy-Tera

ONYX is an internally implemented storage system using commodity hardware with IBM storage servers and Quantas JBOD systems running upon BGFS and ZFS. With a storage capacity of 500TB this is an expandable storage system, able to service the storage requirements of several users since November 2014. Other technical information is provided in the table below.

Administrative Data	
System Name	ONYX
Short Description	Internally implemented storage system with storage servers and JBOD system and drives running upon BGFS and ZFS.
Owner	CyI
Country	Cyprus
Storage system	
Description of storage systems	Quantas JBOD system with BGFS and ZFS
Total storage	500 TB
Dedication to VI-SEEM	
Dedication % of available storage	20%
Dedication storage in TB	100 TB
Integration	
System operational since	Nov 2014
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	SSH
Photo	



Appendix D.3 Avitohol

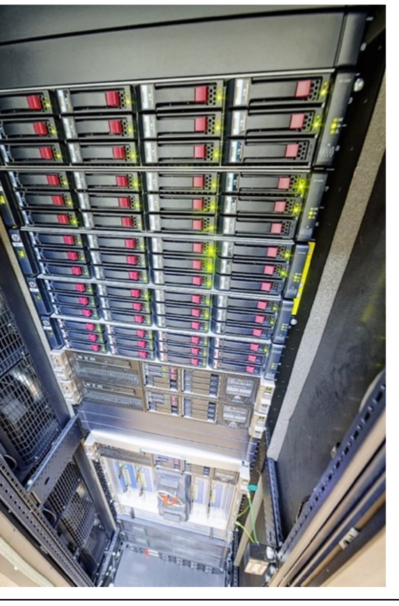
The Storage resources for specialized data storage and preservation will be provided from the Avitohol supercomputer system, as shown in the table below.

Administrative Data	
System Name	Avitohol
Short Description	Bulgarian multifunctional high performance computing cluster.
Owner	IICT-BAS
Country	Bulgaria
Storage system	
Description of storage systems	HP MSA 2040 Storage
Total storage	96 TB
Dedication to VI-SEEM	
Dedication % of available storage	5%
Dedication storage in TB	5 TB
Integration	
System operational since	N/A
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	gridFTP

Appendix D.4 PARADOX

PARADOX provides a data storage system, which consists of two service nodes (HP DL380p Gen8) and 5 additional disk enclosures. One disk enclosure is configured with 12 SAS drives of 300 GB (3.6 TB in total), while the other four disk enclosures are configured each with 12 SATA drives of 2 TB (96 TB in total), so that the cluster provides around 100 TB of storage space. Storage space is distributed via a Lustre high performance parallel file system that uses Infiniband technology, and is available both on working and service nodes. Also, the storage can be used externally via gridFTP protocol. Technical details are provided in the table below.

Administrative Data	
System Name	PARADOX
Short Description	PARADOX cluster storage system
Owner	Institute of Physics Belgrade (IPB)
Country	Serbia
Storage system	
Description of storage systems	PARADOX HP MSA 2040 disk enclosures
Total storage	96 TB
Dedication to VI-SEEM	
Dedication % of available storage	5%
Dedication storage in TB	5 TB
Integration	
System operational since	2013
Available to the project from	PM01
Expected date to be phased out	N/A
Interface to be provided	SFTP, gridFTP
Photo	



Appendix D.5 NIIFI HSM Service

Both disk and tape storage resources for specialized data storage and preservation will be provided from the NIIF HSM system, as shown in the table below. In total, 3 TB of disk and 300 TB of tape storage will be dedicated to the VI-SEEM community.

Administrative Data	
System Name	NIIF HSM service
Short Description	Hungarian Hierarchical Storage Management service
Owner	NIIF Institute
Country	Hungary
Storage system	
Description of storage systems	IBM Storwize V7000 Unified
Total storage	86,73 TB of disk, and 6,8 PB of tape
Dedication to VI-SEEM	
Dedication % of available storage	4%
Dedication storage in TB	3 TB + 300 TB
Integration	
System operational since	2015
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	SFTP, gridFTP

Appendix D.6 NIIFI iSCSI Service

Storage resources for specialized data storage and preservation will be provided from the NIIF iSCSI system, as shown in the table below. In total, 50 TB of disk storage will be dedicated to the VI-SEEM community.

Administrative Data	
System Name	NIIF iSCSI storage service
Short Description	NIIF iSCSI storage service
Owner	NIIF Institute
Country	Hungary
Storage system	
Description of storage systems	Fujitsu DX90S2
Total storage	1 PB
Dedication to VI-SEEM	
Dedication % of available storage	5%
Dedication storage in TB	50 TB
Integration	
System operational since	2013
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	SFTP, gridFTP

Appendix D.7 UVT HPC GPFS

UVT HPC GPFS provides a data storage system which consists of two storage nodes (NSDs) and three different data storage enclosures and systems (IBM DS3400, DS400 and IBM NetApp 7.x family). All the storage systems are exported over fiber channel (dual 4 Gbps and 8 Gbps) to the storage nodes. All the clients (services, nodes etc.) have access to the storage space using the GPFS protocol over Infiniband QDR (40 Gbps) connection (configured in fail-over mode). The total available space is about 50TB from which 10%, 5TB, is dedicated to the project. Details are provided in the table below.

Administrative Data	
Name	UVT HPC GPFS
Short Description	UVT HPC Center – GPFS Storage
Owner	Universitatea de Vest din Timisoara (UVT)
Country	Romania
Storage system	
Description of storage systems	IBM DS3400, DS4000 and NetApp 7.x
Total storage	50 TB
Dedication to VI-SEEM	
Dedication % of available storage	10%
Storage in TB	5
Integration	
System operational since	February 2013
Available to the project from	PM04
Expected date to be phased out	N/A
Interfaces	SFTO, FTP, HTTP, NFS

Appendix D.8 ETFBL-CS01

ETFBL-CS01 will provide 1 TB of storage accessible via Web user interface as well as DAV protocol for the project needs. The aim is to use the storage as a simple way to store and share data between interested parties in a straight-forward manner, easing data delivery to third parties. Details are provided in the table below.

Administrative Data	
Name	ETFBL-CC01
Short Description	University of Banja Luka Faculty of Electrical Engineering Computing Cluster
Owner	University of Banja Luka Faculty of Electrical Engineering
Country	Bosnia and Herzegovina
Storage system	
Description of storage systems	ownCloud storage
Total storage	1 TB
Dedication to VI-SEEM	
Dedication % of available storage	50%
Storage in TB	0.5
Integration	
System operational since	N/A
Available to the project from	PM06
Expected date to be phased out	N/A
Interfaces	ownCloud

Appendix D.9 MK-04-FINKI_CLOUD

MK-04-FINKI_CLOUD provides a data storage system, which consists of four service nodes (HP DL380 Gen8) and 4 additional disk enclosures. The disk enclosure are configured with 60 SAS drives of 600 GB (36 TB in total). Storage space is distributed via a Lustre high performance parallel file system that uses Infiniband technology, and is available both on working and service nodes. Also, the storage is available on the cloud system for disk and block storages. Technical details are provided in the table below.

Administrative Data	
System Name	MK-04-FINKI_CLOUD
Short Description	HPGCC cluster storage system
Owner	UKIM
Country	Macedonia
Storage system	
Description of storage systems	SAN system HP P2000
Total storage	36 TB
Dedication to VI-SEEM	
Dedication % of available storage	5%
Dedication storage in TB	2 TB
Integration	
System operational since	2012
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	gridFTP, cdmi

Appendix D.10 RENAMstor

MD-Grid NGI that is coordinating by RENAM provides a data storage system on FreeNAS 9.3, which consists of 1 Supermicro service node with 5x1Tb HDD in ZFS Z1 raid, so that the service provides 4 TB of storage space, as shown in the table below.

Administrative Data	
System Name	RENAMstor
Short Description	MD Scientific Storage
Owner	RENAM
Country	Moldova
Storage system	
Description of storage systems	FreeNAS 9.3
Total storage	4 TB
Dedication to VI-SEEM	
Dedication % of available storage	25%
Dedication storage in TB	1 TB
Integration	
System operational since	N/A
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	FTP

Appendix D.11 IIAP

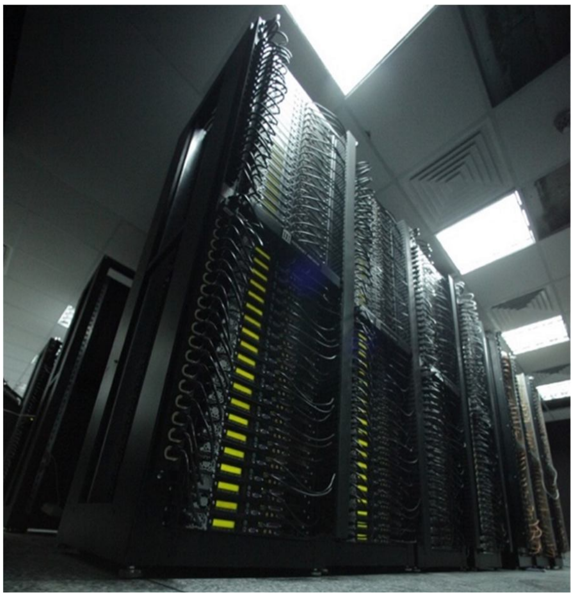
IIAP provides a data storage system on QNAP Turbo NAS server (total 8TB disk space). The storage accessible via Web user interface as well as sFTP protocol for the project needs. Details are provided in the table below.

Administrative Data	
System Name	IIAP Storage
Short Description	AM Scientific Storage
Owner	IIAP
Country	Armenia
Storage system	
Description of storage systems	Storage QNAP TS-809U-RP
Total storage	8 TB
Dedication to VI-SEEM	
Dedication % of available storage	10%
Dedication storage in TB	3 TB
Integration	
System operational since	2014
Available to the project from	PM04
Expected date to be phased out	N/A
Interface to be provided	SFTP

Appendix D.12 BA-IA

The BA has been operating a large-scale storage cluster since 2002. This is commonly known as the Bibliotheca Alexandrina Internet Archive, initially established in a collaborative effort with the San Francisco-based Internet Archive for hosting a comprehensive archive of webpages that allows users of the archive to navigate the web of the past, going back all the way to 1996. Today, while this large-scale storage cluster continues to serve the purpose of hosting the BA web archive, it has also become home to other collections, most notably material digitized at the BA. The cluster is constructed out of commodity hardware, runs all open-source software, and has seen multiple upgrades over the years, growing from the initial 100 TB in 2002 to the current total capacity of 5.2 PB. Technical details are provided in the table below.

Administrative Data	
Name	BA-IA
Short Description	The Bibliotheca Alexandrina Internet Archive
Owner	Bibliotheca Alexandrina
Country	Egypt
Storage system	
Description of storage systems	Internet Archive "Petabox"
Total storage	5.2 PB
Dedication to VI-SEEM	
Dedication % of available storage	2%
Storage in TB	100
Integration	
System operational since	2014 (latest batch of hardware installed)
Available to the project from	PM04
Expected date to be phased out	N/A
Interfaces	SFTP access
Photo	



Appendix D.13 IUCC-InfinityCloud-Storage

The Storage resources for specialized data storage and preservation will be provided from the IUCC InfinityCloud system, as shown in the table below.

Administrative Data	
Name	IUCC-InfinityCloud-Storage
Short Description	IUCC Openstack Cloud Cluster
Owner	IUCC
Country	Israel
Storage system	
Description of storage systems	InfinityCloud storage
Total storage	40 TB
Dedication to VI-SEEM	
Dedication % of available storage	8%
Storage in TB	5
Integration	
System operational since	N/A
Available to the project from	PM04
Expected date to be phased out	N/A
Interfaces	ownCloud