



EGI-Engage

Second intermediate report

M 1.3

Date	14 October 2016
Activity	NA1
Lead Partner	EGI.eu
Document Status	Final
Document Link	https://documents.egi.eu/document/2832

Abstract

This report describes the EGI-Engage activities from 1st March 2016 to 31st August 2016, issues encountered, plans for the next intermediate period (PM19-PM24), metrics and the dissemination and exploitation plan. Effort consumption during the reference period is presented to analyse alignment and deviations with the envisaged work plan.



This material by Parties of the EGI-Engage Consortium is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

The EGI-Engage project is co-funded by the European Union (EU) Horizon 2020 program under Grant number 654142 <http://go.egi.eu/eng>

COPYRIGHT NOTICE

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

DELIVERY SLIP

	<i>Name</i>	<i>Partner/Activity</i>	<i>Date</i>
From:	Małgorzata Krakowian	EGl.eu/WP1	
Moderated by:	Tiziana Ferrari	EGl.eu/WP1	
Reviewed by	AMB and PMB		
Approved by:			29 Sep 2016

DOCUMENT LOG

<i>Issue</i>	<i>Date</i>	<i>Comment</i>	<i>Author/Partner</i>
v.ToC	15 Jul 2016	First version of Table of Content	Małgorzata Krakowian, EGI.eu
v.1.0	27 Sep 2016	First complete document for integration and review	Małgorzata Krakowian, EGI.eu
v.2.0	14 Oct 2016	Updated effort tables, typo corrections	Celine Bitoune, EGI.eu

TERMINOLOGY

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

Contents

1	Executive Summary	7
2	Strategy, Policy and Communications.....	11
2.1	Summary.....	11
2.2	Main Achievements.....	12
2.2.1	Communication and Dissemination	12
2.2.2	Strategy, Business Development and Exploitation.....	15
2.2.3	SME/Industry Engagement and Big Data Value Chain.....	18
2.3	Issues and Treatment	23
2.4	Plans for next period	23
2.4.1	Communication and Dissemination	23
2.4.2	Strategy, Business Development and Exploitation.....	23
2.4.3	SME/Industry Engagement and Big Data Value Chain.....	24
3	E-Infrastructure Commons	26
3.1	Summary.....	26
3.2	Main Achievements.....	27
3.2.1	Authentication and Authorisation Infrastructure	27
3.2.2	Service Registry and Marketplace	28
3.2.3	Accounting.....	28
3.2.4	Operations Tools.....	29
3.2.5	Resource Allocation – e-GRANT	31
3.3	Issues and Treatment	32
3.4	Plans for next period	32
3.4.1	Authentication and Authorisation Infrastructure	32
3.4.2	Service Registry and Marketplace	32
3.4.3	Accounting.....	32
3.4.4	Operations Tools.....	33
3.4.5	Resource Allocation – e-GRANT	34
4	Platforms for the Data Commons	35
4.1	Summary.....	35
4.2	Main Achievements.....	36

4.2.1	Federated Open Data	36
4.2.2	Federated Cloud	36
4.2.3	e-Infrastructures Integration	38
4.2.4	Accelerated Computing	38
4.3	Issues and Treatment	40
4.4	Plans for next period	40
4.4.1	Federated Open Data	40
4.4.2	Federated Cloud	41
4.4.3	e-Infrastructures Integration	41
4.4.4	Accelerated Computing	42
5	Operations.....	43
5.1	Summary.....	43
5.2	Main Achievements.....	44
5.2.1	Operations Coordination	44
5.2.2	Development of Security Operations.....	47
5.2.3	Integration, Deployment of Grid and Cloud Platforms.....	49
5.3	Issues and Treatment	52
5.4	Plans for next period	53
5.4.1	Operations Coordination	53
5.4.2	Development of Security Operations.....	53
5.4.3	Integration, Deployment of Grid and Cloud Platforms.....	54
6	Knowledge Commons	56
6.1	Summary.....	56
6.2	Main Achievements.....	57
6.2.1	Training.....	57
6.2.2	Technical User Support.....	58
6.2.3	ELIXIR.....	60
6.2.4	BBMRI.....	61
6.2.5	MoBrain.....	61
6.2.6	DARIAH	62
6.2.7	LifeWatch	63
6.2.8	EISCAT_3D	64

6.2.9	EPOS	64
6.2.10	Disaster Mitigation	66
6.3	Issues and Treatment	66
6.4	Plans for next period	66
6.4.1	Training.....	67
6.4.2	Technical user support.....	67
6.4.3	ELIXIR.....	68
6.4.4	BBMRI.....	68
6.4.5	MoBrain.....	68
6.4.6	DARIAH	68
6.4.7	LifeWatch	69
6.4.8	EISCAT_3D	69
6.4.9	EPOS	69
6.4.10	Disaster Mitigation	70
7	Consortium Management	71
7.1	Summary.....	71
7.2	Main Achievements.....	71
7.2.1	Administrative and Financial Management	71
7.2.2	Technical Management	71
7.2.3	Quality management	73
7.3	Issues and Treatment	73
7.3.1	Milestones and Deliverables.....	74
7.3.2	Consumption of Effort	75
7.3.3	Overall Financial Status.....	78
7.4	Issues and Treatment	81
7.5	Plans for period PM19-PM24	81
8	Project Metrics	82
8.1	Overall metrics.....	82
8.2	Activity metrics	83
8.2.1	NA1 – Project Management.....	83
8.2.2	NA2 – Strategy, Policy and Communication.....	83
8.2.3	JRA1 – E-Infrastructure Commons	84

8.2.4	JRA2 – Platforms for the Data Commons	84
8.2.5	SA1 – Operations	86
8.2.6	SA2 – Knowledge Commons	87
	Dissemination report	89

1 Executive Summary

Dissemination. EGI successfully contributed to the creation and implementation of a new event concept meant to create a forum where e-Infrastructures and Research Infrastructure experts can meet to discuss requirements, disseminate outputs of their work and discuss new collaboration. The event is called “Digital Infrastructures for Research”, and the first edition took place in Krakow (28-30 September) with the involvement of EUDAT, GEANT, OpenAIRE and RDA Europe. EGI successfully mobilized its user community to contribute to the event. A completely new web site was brought live at the end of September with a focus on services, research user communities and business.

Service catalogue. The new service catalogue of EGI was released in September 2016 in the form of an online register that allows customer interaction to obtain access, and of a brochure for general dissemination purposes. This was a major milestone in EGI given the large legacy of production services and platform, which will enable better value communications towards funding agencies and user groups. We have also evolved the EGI service portfolio, which is structured in two sections (EGI service portfolio, EGI internal service portfolio) with the idea to create a third section (EGI community service portfolio). We worked together with other e-Infrastructures to define a common way to describe a joint service catalogue for research and we agreed a first draft structure. The service catalogue was part of a broader activity aiming at implementing an Integrated Management System (IMS) for the EGI Foundation via the FitSM standard.

Strategy. The EGI strategy was implemented and evolved¹. The themes addressed in this period were: long-term business operations of NGIs, evolution of EGI in the context of the European Cloud Initiative policy, service portfolio strategy and collaboration with commercial cloud providers. We also devoted relevant effort to develop a focused strategy for the Earth Observation sector with engagement actions with EC DG Growth and the Copernicus programme.

We also have continued to maintain the FitSM standard and also to organise trainings to increase expertise across the community with 3 courses held in May and another 5 planned until the end of the year (Sept - Krakow, Nov - Amsterdam).

Business engagement. We consolidated the program with promotional material and a dedicated section on the website. We improved the definition of the process for the onboarding of private organisations. We also established new relationships with actors from the private sector, both as potential suppliers and users for EGI. We have progressed in creating and implementing pilots, especially in the Earth Observation sector where reuse of research data creates innovation potential. We signed a new MoU with CloudSME² and completed negotiation with the one for Terradue³. We also defined a pilot experiment with the Sinergise SME⁴ and EOProc⁵ to

¹ <https://documents.egi.eu/document/2417>

² <https://www.egi.eu/news/egi-joins-forces-with-cloudsme-ug/>

³ <https://www.egi.eu/news/egi-terradue-a-better-cloud-service-for-science/>

⁴ <http://sinergise.com/>

demonstrate how EGI can offer access to Sentinel data and enable exploitation and creation of added value services by external actors. In terms of market analysis, we conducted a study in selected sectors with special focus on agri-food. As part of our business engagement activities, the project authored the strategy document “Exploiting Earth Observation Data at European Scale” to promote the EGI strategy towards EO data with the EC (especially with the Directorate General Growth). We also contributed to the definition of a European Marketplace for Earth Observation services (MAEOS). An activity has started involving major commercial cloud providers to discuss technical interoperability solutions at IaaS level to enable the EGI Cloud Federation with commercial public clouds to compute and access data seamlessly. The SMEs involved in collaborative activities are: Terrasigna, Sinergise, EOProc, Terradue, and CloudEO. We also established collaboration with the Big Data Europe project⁶ and are discussing how EGI can offer infrastructure services to serve their big data platform supporting societal challenges.

Policy. Since September 2016, EGI is contributing to the Open Science Policy Platform (OSPP)⁷ with the support of the project. OSPP was established by the Directorate-General for Research and Innovation to create a Commission Expert Group advising on the development and implementation of open science policy in Europe.

Federation services. The project is rolling into operations the new AAI service “CheckIn” that dramatically advances the accessibility of EGI services. CheckIN provides different Levels of Assurance, allowing, user authentication via OAuth2 and OpenID Connect, including social credentials (Google, Facebook, LinkedIn and ORCID); querying Attribute Authorities via SAML2 Attribute Queries and REST APIs; the generation of long-lived, non-reassignable, non-targeted, opaque and globally unique user identifiers based on the attributes received from the identity provider; user enrolment and account linking via the integration with the COmanage Registry. The EGI marketplace has been prototyped focusing on the testing of the on-line platform and its suitable to render EGI service categories and attributes. The ELIXIR AAI was integrated with the EGI one.

The accounting repository has been enhanced with new features for Cloud, CPU and Storage accounting. For Cloud accounting, changes required to support version 0.4 of the Cloud Accounting Usage Record were implemented and to support the accounting of long running VMs. Work has been done to introduce a new metric in the repository, the Dataset accounting. The first prototype of the new accounting portal was also released. Version 4.0 of the Operations Portal was successfully rolled to production, as well as new versions of ARGO and GOCDB. e-GRANT worked on a first prototype of a pay-for-use platform that integrates two EGI processes, pay-for-use and resource allocation, in a unique system. The technical roadmap of the marketplace, long-tail of science tools and e-GRANT was also reviewed and updated aiming at removing duplication where present and improving the user experience by enforcing the realization of a unique consistent service access workflow to all services of the external portfolio.

⁵ <http://www.eoproc.de/>

⁶ <https://www.big-data-europe.eu/>

⁷ <http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>

HTC and Cloud Compute. Federated cloud technical components (OOI, Keystone-VOMS, accounting, and Cloud and GPGPU information publishing) have been advanced, the Application Database features for VM Image management are now in advanced prototype status. These are a breakthrough development to increase the usability of the federated cloud for what concerns the capability of launching and managing standalone or complex VM structures. The technical integration of gCube/D4Science is now complete and an SLA is now being negotiated between EGI and D4Science. Distributed access to GPGPU computing via the CREAM CE is complete, now supporting different batch systems: HTCondor, Slurm, SGE and LSF.

Data management. The open data hub prototype features have been tested with research communities (LifeWatch and Human Brain Project) in preparation to the first prototype release due in November 2016.

- **Competence Centres and support to RIs and research collaborations.** EGI-EUDAT integration task continues to develop the selected use cases to support the use of services of both infrastructures. ICOS is the most relevant use case for this collaboration, and current work is focused in accessing selected EUDAT services from EGI Federated Cloud resources. A total of 8 Service Level Agreements for research collaborations are now finalized. Platforms capable of accessing distributed GPGPU (DISIVIS, POWERFIT and AMBER) were rolled to production. The BBMRI CC reached an integrated BiobankCloud-EGI FedCloud release, and initiated deployment at 3 affiliated biobanks. The EPOS Competence Centre continued working on the AAI, Computing and Data federation use cases. In response to the first project review, technical support activities were increased towards an additional set of Research Infrastructures and research projects, these include ICOS, EuroBioImaging, Euro-Argo, ENES, EMSO, as well as with 8-10 research project and institutional communities. In this consultancy work the topics were dominated by various ways of exploiting the EGI Federated Cloud.

Operations. The Cloud Middleware Distribution is completed; the software release team is preparing a dry run to test it before the first release. This will allow the application of existing software quality assurance procedures to cloud middleware. Security policy updates and cloud-specific security monitoring instruments are now in operations. The Operational Level Agreements framework was updated to allow the federation of community-specific RI data centres. A fully central deployment allows having faster development of the monitoring infrastructure, a more reliable service and to reduce the effort required for the NGIs to monitor their sites, maintaining unaltered the effort cost of the core activity. Collaboration with other e-Infrastructures (GEANT, EUDAT, PRACE, XSEDE and others) on security trust and development of best practice continues to take place under the auspices of the recently formed WISE community⁸. During the reporting period the ARGO monitoring system have been fully centralised.

In the area of development of new trust framework and security policies, several trust and policy issues were pushed through to approval by IGTF at EUGridPMA and TAGPMA meetings. This

⁸ <https://wise-community.org>

included the accreditation by IGTF of a new "CILogon-like" CA for Europe (RCauth.eu) for use in the new EGI-Engage AAI service, an update to the IGTF IOTA profile, work on private key protection models for centralized credential stores and new video-supported identity vetting models.

A detailed activity plan for period PM13-PM24 was produced taking into account the feedback from the first project review.

Consumption of effort per WP is in-line with the envisaged plan.

2 Strategy, Policy and Communications

2.1 Summary

In the area of communication and event organisation, we organised and ran the spring EGI conference in Amsterdam around the theme 'Opening science in Europe and in the World'. This conference attracted more than 200 participants who participated in some 20 sessions over 3 days. We also developed a new concept for a conference called “Digital Infrastructures for Research” jointly organised with the following e-Infrastructures: EUDAT, GEANT, OpenAIRE and RDA Europe. We also actively maintained the communications channels with more than 40 news items and blog posts on EGI sites, and two editions of the Inspired newsletter with a total of 17 articles. In terms of publications, we finalised a new brochure to promote the business engagement program to private organisations, we defined the structure and content for a dedicated brochure to promote the new EGI service catalogue and we authored a new case study document related to a scientific result supported by EGI and published in Nature. We also contributed to a joint work led by the Unit e-Infrastructures from DG CONNECT for a joint booklet to promote services co-funded by the unit.

In the area of strategic planning, we supported the EGI Council with policy briefs and workshop organisation to further identify steps to implement and evolve the EGI strategy⁹. The themes addressed in this period were: long-term business operations of NGIs, evolution of EGI in the context of the European Cloud Initiative policy, service portfolio strategy and collaboration with commercial cloud providers. We also devoted relevant effort to develop a focused strategy for the Earth Observation sector with engagement actions with EC DG Growth and the Copernicus programme. In the area of governance, we have launched the call for applications to become a member of the EGI Strategy and Innovation Board (SIB), the external advisory board for EGI. We also defined the structure, key benefits and initial set of indicators for the EGI impact report to be developed along the dimensions of research, innovation, collaboration and skills.

In the area of service strategy, we have contributed to the definition of the Integrated Management System (IMS) for the EGI Foundation and to develop the definitions for the Service Portfolio Management (SPM) and Continual Improvement (CI) with a newly added Business Development and Stakeholder (BDS) processes. We have also evolved the EGI service portfolio, which is structured in two sections (EGI service portfolio, EGI internal service portfolio) with the idea to create a third section (EGI community service portfolio). We worked together with other e-Infrastructures to define a common way to describe a joint service catalogue for research and we agreed a first draft structure. A second dedicated session on the topic will be held in Krakow at the DI4R conference in September. We also have continued to maintain the FitSM standard and also to organise trainings to increase expertise across the community with 3 courses held in May and another 5 planned until the end of the year (Sept - Krakow, Nov - Amsterdam).

⁹ <https://documents.egi.eu/document/2417>

In the area of cross-border procurement of e-Infrastructure services, we have organised a workshop with the key research infrastructures involved in the activity to continue the assessment of the opportunities and barriers, and to identify recommendations.

Concerning the activity on enabling the pay-for-use business model, the work focused on finalising the provider agreement framework and developing the last features needed within the e-GRANT tool towards moving to production.

In the area of business engagement, we consolidated the program with promotional material and a dedicated section on the website. We have improved the definition of the process for the onboarding of private organisations. We also established new relationships with actors from the private sector, both as potential suppliers and users for EGI. We have progressed in creating and implementing pilots (especially in the EO sector). We signed a new MoU with CloudSME and completed negotiation with the one for Terradue. In terms of market analysis, we conducted a study in selected sectors with special focus on agri-food.

2.2 Main Achievements

2.2.1 Communication and Dissemination

Events

This section provides a list of events organized or being organized in reporting period:

EGI Conference 2016, in Amsterdam, 6-8 April 2016

<https://indico.egi.eu/indico/event/2875/>

The event was organised by the EGI-Engage project, this time with no local organiser, which took place at the Amsterdam Science Park. The conference's programme was organised around the theme of 'Opening science in Europe and in the World' and placed research communities, NGIs and EIROs and collaborating e-Infrastructures under the spotlight as key players for the delivery of digital services for borderless science. We welcomed 222 participants who participated in 19 sessions over 3 days. The conference hosted an INDIGO-DataCloud project meeting and the first edition of the 'Design your e-Infrastructure' workshop – an event organised by the EGI User Community and Support Team (UCST) to accelerate the onboarding of new user communities.

Digital Infrastructures for Research 2016 (DI4R 2016)

<http://www.digitalinfrastructures.eu/>

This event was initially planned to be the EGI Community Forum for 2016 and co-hosted by Cyfronet, the Polish representative in the EGI Council. In order to strengthen the collaboration in the e-Infrastructures domain, the concept was expanded towards a joint conference around the theme "Serving the user base" co-organised by EGI in collaboration with the following e-infrastructures: EUDAT, GÉANT, OpenAIRE and RDA Europe. As of the writing of this report, all logistical arrangements are in place and the programme is finalised and online. Keynote speakers have been identified and secured and brochures were prepared for sponsors and exhibitors. The

website of the event is up and running. The different co-organisers contributed to the organisation with different items. EGI's contribution (entirely supported by EGI-Engage) consisted of:

- Liaising with the local host to negotiate terms with the organisers steering committee
- Liaising with the local host to prepare content for the Exhibition and Sponsors brochure
- Hosting the Programme Committee face-to-face meeting
- Leading the event's communication activities
 - Drafting all news items and announcements
 - Sourcing quotes and content for the website
 - Keeping the event's newsfeed populated
- Contributing to the Programme Committee's activities

The DI4R event represents a shift in the events' strategy of EGI for the future and it is expected to have a major impact in facilitating and strengthening cross-infrastructure collaboration activities and in improving the outreach to research communities with joint offers. Taking this into account, we outlined a set of success factors and ways to gauge them. They are related to the following aspects:

- EGI retains its voice and is visible on the programme
- The EGI community participates in the joint event
- Cross-pollination: EGI collaborates more and expands its network
- Efficiency of logistics

Active communication channels

Newsfeed and newsletter

The EGI newsfeed has been updated with 23 news items (as of 6 August), of which about one quarter give visibility to EGI's partners and/or users. Issues 23 and 24 of Inspired, the EGI newsletter, have been published. In total, the newsletters presented 17 articles and features of which 9 were sourced in the community (i.e., they were not written by EGI Foundation staff). This trend continues from previous project periods and gives us confidence that our steps towards more inclusive communications channels, at the disposal of the community, have been successful.

Blog

As of the end of July 2016, 19 blog posts were published as a mix of announcements and opinion pieces. The blog's content is now self-sourced, i.e.: most items are published on the request of community members that actively request their announcement / pieces to be placed in our blog. This reinforces the idea that we are being successful in becoming a platform for discussion and community building.

Publications

The NA2.1 worked closely with the other WP2 tasks to develop a number of publications. They were:

1. EGI: Open for Business: a leaflet with the basic ideas behind the EGI Business Engagement Programme, designed to convey an attractive message that can be used a conversation starter. PDF of a previous version with an old name¹⁰.
2. Impact report and Impact brochure: this long-term (PY2) project started as a leaflet to be prepared for events and evolved to the idea of developing an Impact report based in four dimensions – research, skills, innovation, and collaboration. More information below.
3. EGI Service Catalogue: developed the concept of a brochure for the EGI Service Catalogue, based on the service descriptions prepared for the new website. The brochure will be used at events and as hand-outs to potential users to show the range of services and basic features, including use cases and testimonials.
4. e-Infrastructures: Making Europe the best place for research and innovation: this booklet, published by the European Commission's e-Infrastructures Unit, describes how e-Infrastructures contribute to innovation and research. EGI contributed to the project by suggesting the overarching concept and proof-reading the final copy (see document¹¹).

Champions programme

The Champions programme was designed and implemented during the first months of EGI-Engage, however it was not as effective as hoped: despite the effort to advertise the programme through the channels available, only four applications were received. We took remediation measures described in the EGI-Engage periodic report for year 1¹², nevertheless none of them has led to an improvement. Therefore, we decided to refocus the Champions Programme towards the goal of onboarding new user groups – which is different from the original approach, but has more chances of success and will lead to more engagement and more users. The first activity was the 'Design your e-infrastructure' workshop in Amsterdam¹³ which will be followed by a second edition co-located with the DI4R 2016 in Kraków, embracing more e-Infrastructures participating in joint onboarding of new user groups¹⁴.

Website redesign

As reported and justified in the EGI-Engage PY1 report and milestone M2.3 The website is reviewed¹⁵, we worked on the development of a new website for EGI. The new website will be about EGI, its services, solutions and supported users as a whole. The main aims of the website are:

- Showcase EGI services and solutions
- Establish EGI as a leading data and computing e-infrastructure for research
- Increase service awareness
- Give visibility to the stakeholders

¹⁰ http://www.egi.eu/%20news-and-media/publications/EGI_Business_Engagement_Programme.pdf

¹¹ http://www.egi.eu/export/sites/egi/news-and-media/publications/booklet_infrastructures.pdf

¹² <https://documents.egi.eu/document/2794>

¹³ <https://indico.egi.eu/indico/event/2895/>

¹⁴ <https://indico.egi.eu/indico/event/3025/>

¹⁵ <https://documents.egi.eu/public/ShowDocument?docid=2765>

During the first half of PY2, we worked together with a website designer to develop a customised Wordpress template. We also started to add in content. As of August 2016, we aim to have the new website launched in October 2016, at the latest.

Case studies

We published one case study¹⁶ during this period, based on a paper produced thanks to the EGI Federated Cloud and published in Nature. This is below the target. The main reason is that we decided to refocus effort on the development of the new website during this period.

2.2.2 Strategy, Business Development and Exploitation

Strategic Planning and Evaluation (EGI Foundation)

In May 2015, the EGI Council adopted the new strategy¹⁷ and an initial summary report on the implementation was provided with deliverable D2.9¹⁸. During this period, this task supported the detailed evaluation of the strategy implementation by updating the live document with progress tracking for each objective identified in the strategy¹⁹.

The task also supported the decision making of the EGI Council with the following activities:

- Helped with the organisation of two Council workshops: 6 April 2016 focusing on the long-term operations and funding of NGIs; and 28-29 June 2016 focusing on the long-term business operations of the EGI federation and the potential role within the European Cloud Initiative
- Authored the policy brief “Collaborating with Commercial Cloud Providers” to support the Council in evaluating what kind of engagement should be pursued with private actors of the Cloud sector²⁰.
- Authored the policy brief “EGI service portfolio” to propose the adoption of a new structure and definitions for the EGI service portfolio²¹.
- Authored the policy brief “EGI Strategy and Innovation Board (SIB) - Profiles”²².

In the area of strategy for the Earth Observation sector, given the high relevance and needs for actions in exploiting EO data, during the last period we have intensified our actions:

- Authored the strategy document “Exploiting Earth Observation Data at European Scale” to promote the EGI strategy towards EO data with the EC (especially with the Directorate General Growth)²³

¹⁶ http://www.egi.eu/case-studies/medical/genetics_salmonella_infections.html

¹⁷ <http://go.egi.eu/strategy2020>

¹⁸ <https://documents.egi.eu/document/2669>

¹⁹ <http://go.egi.eu/strategy-implementation>

²⁰ <https://documents.egi.eu/document/2852>

²¹ <https://documents.egi.eu/document/2630>

²² <https://documents.egi.eu/document/2884>

²³ <https://documents.egi.eu/document/2851>

- Authored a poster for the Copernicus Value Chain workshop (26 April 2016²⁴) to communicate the EGI strategy and solutions on “Exploiting Earth Observation Data at European Scale: EGI Data Hub for Sentinel Data”²⁵
- Organised meetings with the EC Directorate General Growth to better understand the policy needs from the EC and explain how EGI could support those needs
- Prepared and submitted an abstract for session on exploitation of Copernicus Data at the DI4R conference that was accepted
- Contributed to two EC surveys on exploitation of Copernicus and EC Space strategy
- Contributed to one survey on the definition of an European Marketplace for Earth Observation services (MAEOS)

Governance Evolution (EGI Foundation)

During this period, we prepared and launched the call for applications to become a member of the EGI Strategy and Innovation Board (SIB)²⁶. To develop this call, we used the profile descriptions that were accepted by the EGI Council through the dedicated policy brief²⁷.

Impact Assessment (EGI Foundation)

We drafted an initial structure for the report to capture the impact of EGI. This is structured along four dimensions: 1) research: to capture the impact of EGI in enabling better and more efficient research; 2) innovation: to capture the impact of EGI in improving processes, products/services, technology or business models that benefit the researchers; 3) collaboration: to capture the impact of EGI in facilitating and supporting collaborations; 4) skills: to capture the impact of EGI in developing new skills. We have an initial draft of benefits and related indicators and we plan to develop a full report by the end of PY2.

Service & Solution Strategy (EGI Foundation)

Service management

During this period, we progressed in the definition of a more mature Service Portfolio Management (SPM) process as part of the EGI Integrated Management System (IMS). We updated the service portfolio entry template and split the previous portfolio in two:

- EGI service portfolio: collection of services provided from the EGI federation externally
- EGI internal service portfolio: collection of services provided to the EGI federation in order to enable them to work together

We also started working on the concept of an EGI Community service portfolio capturing the collection of services that are offered by the research communities and that may be of general interest for being promoting to the wider EGI user base. This would require EGI to establish agreements with the providers to ensure clear service levels and capacity to be available. During

²⁴ <http://www.copernicus.eu/value-chain-workshop>

²⁵ <https://documents.egi.eu/document/2857>

²⁶ <https://documents.egi.eu/document/2131>

²⁷ <https://documents.egi.eu/document/2884>

this period, we also analysed the expected outcomes from the Indigo-DataCloud project to identify potential new service components that could support the creation of new services within EGI.

In the context of collaborating with other e-Infrastructures to develop a joint service catalogue of services for research, we have collaborated with EUDAT, OpenAIRE, the BlueBRIDGE project, the THOR project and GEANT to define a common way to describe services and to share best practices around service portfolio management. We organised and held a workshop at the EGI Conference (6 April 2016²⁸), and we planned a follow-up session at the DI4R conference²⁹. An initial output of this work is a draft definition of a scheme to describe a joint service catalogue³⁰.

EGI IT Service Management (ITSM)

During this period, consistent effort was devoted to complete the transition from the mediawiki-based Service Management System towards the confluence-based Integration Management System (IMS). This work is performed within the initiative to achieve ISO 9001 and ISO 20000 certification for the EGI Foundation. The additional requirements to comply with the standards were identified and work has been scheduled with the various process owners to improve the definition of the various processes. Compared to FitSM, two processes were introduced: Business Development & Stakeholder (BDS) and Budgeting, Accounting for Services (BA). All activities are supervised by the Services and Solutions Board (SSB) who met three times during the period under review. An internal audit was conducted in May (M2.4) with another planned for the end of Sept in preparation for the external audits later in the year as part of the certification process.

FitSM

With regards to trainings, we successfully delivered 3 courses in May 2016 (Foundation, both Advanced Levels) with a total of 19 total participants. We planned a Foundation training that will be co-located at the DI4R conference in Krakow (27 Sept) along with 4 more courses in November³¹ that cover all levels including our first Expert level course. We disseminated FitSM at the INFN Computing Commission Workshop (17 May 2018, Elba, Italy). Following this, INFN (the Council representative for Italy) expressed the interest to organise in-house courses (Foundation + 1 Advanced) tentatively scheduled for October.

With regards to the evolution of the FitSM standard, we attended the monthly calls from the FitSM Working Group managed by ITEMO³² chaired the FitSM Subgroup on marketing, and supported other subgroups in areas such as training material and exams; community contributions; external relationships; website and internal tools. In addition to supporting the overall future and evolution of FitSM, participation also allows the EGI Foundation trainers direct

²⁸ <https://indico.egi.eu/indico/event/2875/session/15/?slotId=0#20160406>

²⁹ <http://www.digitalinfrastructures.eu/content/joint-service-catalogue-research>

³⁰ <https://indico.egi.eu/indico/event/2875/session/15/contribution/38/material/0/0.pdf>

³¹ <http://go.egi.eu/fit-sm-trainings>

³² <http://www.itemo.org/>

access and influence to implement participant feedback from each of the training. Finally, we collaborated in preparing and submitting a poster to the DI4R conference³³.

Cross-border procurement of e-Infrastructure services

CERN organised a Cross Border Joint Procurement session at the EGI Conference in Amsterdam on 7 April 2016³⁴. The contributors to this activity (BBMRI, CSCI, RBI, INGV and EGI.eu) participated and a report from the session was created with a summary of the work performed, assessment of the opportunities for the future and recommendations³⁵. A follow-up session for the DI4R conference was prepared and submitted (see Cloud procurement³⁶).

Pay-for-use Implementation

This activity is led by the EGI Foundation and is collectively supported by a number of providers (IICT-BAS, CSIC, GRNET, INFN, TUBITAK). Technical developments regarding the e-GRANT tool is provided by Cyfronet though funded under JRA. During this period, the work focused on the actions that are needed to move to pay-for-use capabilities to production, particularly:

- Created a “Letter of Intent” template for any EGI provider to formally state their ability and willingness to provide services on a pay-for-use basis as well as the access policy/limitations (e.g. research/innovation; commercial). This letter is provided by the institutions legal/administration department and required to be visible in e-GRANT; so far, we received 4 signed letters from: CESGA (Spain); CSC (Finland); CSIC (Spain), and 100%IT (UK)
- Developed terms and conditions for the e-GRANT tool as well provided the necessary requirements to finalize the remaining features in order to move into production (Sept 2016).
- Drafted an internal P4U Provider Agreement outlining process, procedures and general rules/conditions for provider participation.
- Drafted the documentation for both the customer and provider documentation

2.2.3 SME/Industry Engagement and Big Data Value Chain

(EGI Foundation)

During this period, we worked on promoting and maturing the business engagement program. With regards to the promotional aspects, we created and populated a new dedicated section on the website to host information for the private sector to engage with EGI³⁷. Given the current major re-design of the website, this content will be later ported to the new website expected in the last quarter of 2016. We also created a brochure in collaboration with NA2.1 to disseminate the key aspects of the business engagement program. The brochure was distributed for the first

³³ <http://fitsm.itemo.org/informational-material>

³⁴ <https://indico.egi.eu/indico/event/2875/session/27/#20160407>

³⁵ <https://documents.egi.eu/document/2883>

³⁶ <http://www.digitalinfrastructures.eu/programme>

³⁷ <http://go.egi.eu/for-business>

time at the I4MS conference and we got the opportunity to initiate a discussion with the I4MS coordinator to define a strategy to be involved with digital innovation hubs³⁸.

From the process viewpoint, we developed a questionnaire to collect initial information from interested private partners that help us evaluate the business case for cooperation³⁹. We also defined procedures to on-board new private customers that were integrated in the BDS and Customer Relationship (CRM) processes within the IMS. We also co-created a session about SME engagement with other e-Infrastructures scheduled for the DI4R conference⁴⁰.

With regards to establishing new relationships in the sector of private organisations as suppliers of services, we had conversations with Microsoft Research to increase knowledge on EGI services for them and on Azure services for EGI. Microsoft offered the opportunity to access grants for resources that EGI could use for testing. We discussed also the opportunity to extend the rOCCI server implementation to expose a Microsoft Azure endpoint using the standard interface. We also had conversations with the scientific computing department of Amazon and a number of areas for collaborations were identified. We had discussions with T-Systems to explore how the Open Telekom Cloud can be integrated in the EGI Federated Cloud. Under the existing collaboration with UberCloud, we achieved the publishing of EGI Cloud services in the UberCloud marketplace and performed an initial analysis of potential companies to run pilot.

Considering private partners as users of EGI services, we started a discussion with Funboostr (startup) and with Terrasigna⁴¹. We have also defined a pilot experiment with the Sinergise SME⁴² and EOProc⁴³ to demonstrate how EGI can offer access to Sentinel data and enable exploitation and creation of added value services by external actors. In the context of the TEISS stimulus project funded by ESA, some of the effort was supported by EGI-Engage as well; through the activity we achieved integration of the CYFRONET cloud provider (part of the EGI Federated Cloud) in the SlipStream platform connected to the CloudEO store (German SME); the work was presented at the EGI conference in Amsterdam; we also wrote the final report covering business models and technical architectures and we are discussing how move to monitored production state for potential long-term business relations.

With regards to formal agreements (MoUs) with private partners, we completed and signed the agreement with CloudSME, and started initial technical testing to enable their users to access EGI resources. We held a first kick-off meeting (call) to define priorities and timelines and published news item announcing collaboration⁴⁴. We have also finalised the agreement with Terradue that is awaiting for signature. The document articulates previous/ongoing work and defines future

³⁸ http://i4ms.eu/events/amsterdam_event_detail.php?post_id=196

³⁹ <https://www.surveymonkey.com/r/business-engagement>

⁴⁰ <http://www.digitalinfrastructures.eu/content/sme-dynamics-workshop-%E2%80%93-towards-integrated-approach>

⁴¹ <http://www.terrasigna.com/>

⁴² <http://sinergise.com/>

⁴³ <http://www.eoproc.de/>

⁴⁴ http://www.egi.eu/news-and-media/newsfeed/news_2016_023.html

activities⁴⁵. We are also drafting the business case document, the first of a list of brochure to disseminate the success stories of private organisations working with EGI.

With regards to the big data value chain activity, we continued our collaboration with the BDVA by participating in the BDVA Activity Group meetings and we attended the BDVA summit (2-3 March 2016, Den Haag). We also established collaboration with the Big Data Europe project⁴⁶ and are discussing how EGI can offer infrastructure services to serve their big data platform supporting societal challenges. In this area, we should also link the activities with DG Growth (see NA2.2 report) and the pilot with Sinergise/EOProc mentioned above. We also promoted the EGI services and possible role to support the exploitation of Sentinel data at the EARSC workshop on the Marketplace for EO services (22 June 2016, Brussels).

CNRS

CNRS took part in the EGI-Engage SME / Industry Engagement activities, namely the regular meetings held, leaflet contributions/feedback, and actively participated in the Workshop: Design your e-Infrastructure. 4 new contacts were added to the SME contact database from attendance at a French conference, nevertheless there are not yet concrete results.

CSC

CSC has signed the Pay-for-use Letter of Intent (LoI)⁴⁷ stating its ability to offer cloud resources⁴⁷ for SMEs and other paying customers. A number of national customers, especially in the biomedical sector, has previously taken benefit of the CSC cloud offering. With the LoI the opportunity for EGI.eu to broker the resources is now opened. CSC has participated in meeting and discussions organised within the activity.

GRNET

During the first half of PY2, the main effort has been focused on authoring D2.10 “Market analysis report of selected sectors” mainly on agri-food sector that is due by PM18 (Aug 2016). However, GRNET has participated in all meetings and requests circulated via the mailing list.

IICT-BAS

Established new contacts with companies dealing with real-estate and discussed possibilities for them to employ big data technologies. These companies are one of small and one of medium size with different use cases. Discussions about implementation of blockchain technology, feasibility, performance, and applicability with industrial partners. Dissemination during the recently organized training on how to use our Xeon Phi-based cluster, where developers from several SMEs were involved. During the development of the Business Plan for the Bulgarian HPC and distributed infrastructure in the next few years, we made provisions for the access from SMEs to our infrastructure and in this way to EGI. These organisations have been recorded in the EGI database as well as in RT.

⁴⁵ <https://documents.egi.eu/document/2887>

⁴⁶ <https://www.big-data-europe.eu/>

⁴⁷ <https://research.csc.fi/cloud-computing>

Engineering

According to the plan, the main effort was used during PY1. The remaining hours were used as follows: 1) post D2.7 analysis and report to translate findings to EGI requirements; 2) revision of D2.7 Market Report on the Fishery and Marine Sciences Data Analysis Sector for Minor typographical errors as per request from EC reviewers.

SWING/FMI

There has been work with Institut Curie in Paris to develop a resource sharing system that increases awareness and facilitates the sharing of resources between academic and commercial organizations. From discussion, there is interest to expose EGI resources. The system can be found at <http://iris.curie.fr>, which shares publicly a list of resources at Institut Curie. In addition, the system is used to engage SMEs to take advantage of research resources at Institut Curie or to facilitate collaborations between academic and commercial entities. The platform at Institut Curie now has 739 active internal users with 172 registered resources. The test platform is already used to facilitate resource sharing and services to users at commercial entities such as Nikon, L'Oréal, ogd2pharma, and Roper Scientific. In addition, users from local academic institutions such as Institut Pasteur, Université Paris-Sud, Institut Jacques Monod, Institute of Plant Science of Paris Saclay, Institut Gustave Roussy, and Université Paris Diderot are using the platform to receive services and use resources from Institut Curie. This represents a potential large opportunity for EGI, which will continue to be pushed into the second half of PY2.

BIFI

The institute is launching a public presentation of the Centre of Supercomputing of Aragon (CESAR) with research groups, government agencies and companies invited. We are participating in the working group RESxPYME associated to the Supercomputing National Network, in the task of identifying other national and European initiatives (with EGI-Engage among others). There has been work in preparing a proposal for funding a thematic national network on Big Data with a total of 10 Spanish partners in order to bring this technology closer to companies. The cooperation with European initiatives such as EGI-Engage is one of the tasks to carry out. We held direct talks with business environments: ITA, Industry 4.0 Forum in Bilbao and are discussing how to participate in <http://cf2016.holacloud.eu>

IFCA-CSIC

IFCA concluded the work in common with several SMEs and companies in Spain to support different Cloud solutions for the LifeWatch initiative using FedCloud architecture as reference. This provides an important experience about this joint work and transfer process. Contacts have been established with these companies, and in particular with GMV (engineering) on the possibility to collaborate in new projects like the PCP launched by Helix Nebula. IFCA is also trying to launch new initiatives with SMEs related to the exploitation of ESA Earth Observation program, COPERNICUS, using FedCloud tools and resources.

INFN

INFN contributed to Torus Workshop. TORUS is a project funded by the European Union in the framework of Erasmus+ Capacity Building involving Italian SME on Distributed Computing Architectures And Environmental Science Applications 6-10 June 2016 University of Ferrara.

We also established good contacts and relationship with several SMEs in Emilia Romagna. The Emilia Romagna region is investing remarkably on Research and technology transfer in all major domains of innovation, and in creating value from Big Data. Attention to innovation in the Big Data field comes mainly from a widespread interdisciplinary and dynamic research system (4 public universities, prominent research infrastructures and facilities and the territorial branches of national research centres) well integrated into the local industrial landscape mainly composed by SMEs. The contact with the SME is ongoing, has not reached any concrete results waiting for approval of high level training initiatives foreseen for Sept 2016.

SURFsara

SURFsara has restructured teams internally that have led to identifying new contacts to push forward this activity. A call was held with project management to define objectives and potential input from SURFsara for PY2. We reviewed EGI Business Leaflet Flyer content and design and provided our feedback based on real world experience in working with the private sector.

LIP

The Portuguese NGI was restructured, which is now a new independent legal entity dedicated to deliver scientific computing and data services to the scientific and academic sectors. The infrastructure access policy foresees the delivery of services to the private and public sectors with focus on SMEs and government related bodies. A project to fund the new Portuguese distributed computing infrastructure (INCD) approved under the Portuguese Scientific Infrastructures Roadmap, has been submitted in July. The project will fund the infrastructure over the next years, and foresees economic activity up to 20%, having special focus on SMEs. The project includes the development of business and sustainability models that will include these economic activities. INCD will pursue the coordination of these activities with EGI. INCD encompasses LIP together with the Portuguese NREN and the Portuguese Civil Engineering Laboratory (LNEC). The close relations between these entities are being exploited to facilitate the contact with the private sector, and the use of existing facilities (eg. NREN network and datacenters) encompassing changes to AUP policies to allow commercial activities.

Contacts with potential private customers in the civil engineering and close related domains have been established (via LNEC) namely (structural, coastal and estuary simulations). There is interest in cloud and HPC capacity for simulations from small companies. Other contacts with the Portuguese earth observation community have been established that includes private entities have been recently established although in a very early stage. Currently, there is preparation of a small scale provisioning of services for simulation of radiation effects on electronic components under an ESA contract with LIP. Services to be used by ESA via a portal that can submit to the Portuguese infrastructure. This activity will help assess costs and impact of providing such paid services. This means evaluating the possibility of delivering pay-per-use services to public entities

under the Portuguese government cloud framework, an activity that may enable a better understanding of the delivering of this type of services to public or private entities.

2.3 Issues and Treatment

No issues have been identified during reporting period.

2.4 Plans for next period

2.4.1 Communication and Dissemination

Early activities will focus on launching the new website, publishing a new edition of the Inspired Newsletter and creating a new brochure on the EGI service catalogue. Regarding events, we will support the organisation and evaluation of the conference DI4R, open the bid for a conference in autumn 2017 and start the organisation of the EGI Conference 2017 in spring. Other activities will be identified and defined in the updated communication plan.

2.4.2 Strategy, Business Development and Exploitation

In the area of strategic planning and evaluation, we plan to complete evaluation of EGI strategy implementation and update the plan for 2017. We will also revise the EGI strategy in collaboration with the stakeholder. Concerning the governance, we will focus on completing the setup of SIB and organise first meeting. In the area of impact assessment, we will complete the first version of the EGI impact report and publish it.

In the area of service and solution strategy, the work will focus on completing the implementation of the IMS (Integrated Management System) and supporting the work towards achieving the combined certification for the ISO 20000/9000 standards. With regards to the service portfolio, we will complete the first version with the details for each service portfolio entry and also with the various service options.

Concerning the contribution to the DI4R conference, we will organise and run the session “Joint service catalogue for research session”, the co-located training “FitSM Foundation Training” (and related poster), and we will contribute to the Workshop on “Design your e-Infrastructure”.

In the area of cross-border procurement of e-Infrastructure services, we will organise a session at the DI4R conference to finalise the analysis with the stakeholders and publish the final report.

Concerning the pay-for-use implementation, we plan to move the e-GRANT pay-for-use instance in production and continue to improve the system based on user experience and arising new business scenarios/questions, to investigate and implement a billing function, to define a full broker business model for EGI Foundation (the latter was previously put on hold due to VAT questions being analysed by Dutch authorities; now that this has been resolved, this activity can move forward with more in-depth analysis and implementation). We will also support the analysis of the work for the EGI marketplace.

2.4.3 SME/Industry Engagement and Big Data Value Chain

EGI Foundation

Regarding business engagement activities, the second half of PY2 will be completing the SME onboarding steps as part of the EGI IMS based on experience from discussions with SMEs from the first year. In order to harmonize statistics and the information obtained from business contacts the aim is to complete at least 20 questionnaires, which will be supported by all activity partners. Final organisation and delivery of an EGI for SME presentation will be carried out in a dedicated session called the “SME Dynamics Workshop” at the DI4R Conference in Krakow. Efforts will focus on identifying 3 SMEs to work on success stories to be achieved by the end of the project; ideally the three SMEs should cover the following sectors: Earth observation, agri-food, fishery & marine. In terms of the big data value chain, participation will continue in the BDVA Activity Group meeting and breakout sessions, push forward collaboration with Terrasigna and coordinate collaboration with Big Data Europe.

As the majority of the market analysis concludes in M18 with the last related deliverable (D2.10), work moving forward will be to analyse the recommendations outlined in order to define specific offers, create any material required, and conduct suggested outreach to both the fisheries/marine and agri-food sectors.

Activities outlined in the current 3 formal agreements (MoU) with UberCloud, CloudSME and Terradue will continue of the next year. A webinar to present EGI for SME is currently being planned through UberCloud, tentatively scheduled on 18 Oct as a call for expressions of interest with the aim to identify at least 1 company to run pilot. Technical integration of the EGI FedCloud with the CloudSME platform will continue, start to match technical requirements from customers with EGI service offers and run at least 1 pilot. A business case study with Terradue is currently being drafted with the plans to publish in the coming months.

CNRS

France Grilles Success Days⁴⁸ will be held in November in Paris. CNRS plans to invite local industry to attend and continue to push forward contacts.

CSC

Part from contributing to the WP2 activities priority is to actively seek opportunities through the pay-for-use activity. A collaboration with the VTT Technical Research Centre of Finland Ltd is about to be formalised. VTT has a large network of SMEs and together we have the intention to better meet the demands of SMEs in terms of simulation knowledge and actual computing capacity.

GRNET

As main authors of D2.10, support will be provided around the technical evaluation of requirements defined within the report. This also includes input to communication activities such as solutions to promote and how to reach the selected sectors. We will continue to add identified companies in the contact database for evaluating business cases for collaboration and any activities defined through the group.

⁴⁸ <https://succes2016.sciencesconf.org/>

IICT-BAS

Concentration will be put on clarifying the modus of cooperation with the newly established SME contacts and start to determine the technical persons from them that are to be involved. In addition, we will continue to broaden the contacts with startups, especially during an upcoming training event in October or November to be held at IICT-BAS, as it has been found that newly established startups have more needs of infrastructure and expertise.

SWING/FMI

Continue the discussions and work with the Institut Curie in Paris to understand how EGI resources can be included. This will be done in collaboration with CNRS leveraging the national connection, experience as an EGI provider and also partner within this project task. Once completed, any opportunities with commercial providers will be fed through the group mailing list as well.

Engineering

All effort consumed as per project planning. No further actions are scheduled.

IFCA-CSIC

IFCA will continue the collaboration with different SMEs and companies on e-infrastructure projects. The effort on producing a first integrated modelling of a water reservoir includes now the use of GPUs for predictive NN, a work in collaboration with the SME Ecohydros that now aims to be extended to the Sanabria Lake. The offer will also extend more directly to the HPC area, with the integration of supercomputing resources using dockers for genetic tools (commercial contact yet pending). IFCA will also keep contact with Open Science Cloud initiatives at national level, and with other ones related to structural funds where companies can and should get involved.

INFN

Contact the SME involved in the Emilia Romagna high level training initiative and investigate their interest in using the EGI Services.

SURFsara

We will continue to contribute to the documents and the web site information available for industry, disseminate EGI possibilities and services to attract SMEs through SURFsara Private Section Program, Fortissimo and other related SURFsara activities. This includes completing the EGI industry questionnaire for statistical analysis and to try to establish formal collaborations/pilots.

LIP

Now the major restructuring of the Portuguese NGI is complete, we will focus on pursuing the contacts within civil engineering, earth observation, and also biology where interest on potential delivery of services to SMEs has been observed. Further development of the INCD administrative structure will aim to enable the delivery of services on a pay-per-use basis via this new legal entity. This will require the preparation of training and dissemination activities to take place over the next months through which contacts with potential SMEs can be established and/or deepened in order to identify and provide the opportunities more suitable to EGI.

3 E-Infrastructure Commons

3.1 Summary

This work package coordinates the development of the e-Infrastructure Commons - an ecosystem of services that constitute the foundation layer of any distributed e-Infrastructures. The technical development of the e-Infrastructure Commons services is planned to be user-driven to ensure full interoperability with other e-Infrastructures and RIs. The existing set of services is going to be expanded to include a Service Registry and Marketplace. The main targets of this task are:

- Provide viable methods for authentication and authorisation in the EGI ecosystem.
- Simplify the access to the infrastructure services through technological innovation and new services in the area of Service Registry and Marketplace and resource allocation.
- Evolve the EGI accounting system to manage the data deluge expected over the next years, including new types of accounting metric (e.g. data accounting) and redesigning of the presentation layer to improve the user experience.
- Adapt the operations tools to new technologies and to satisfy new requirements emerging from service providers and user communities.
- Define interfaces to create a network of analogue tools that provides users with integrated view of all the infrastructures involved.

The WP3 manager organises monthly meeting to check the progress of the developments and discuss about issues and tool dependencies⁴⁹. The WP3 roadmap is published in the activity wiki page⁵⁰ and it has been updated as described in the M3.3 Operational tools development roadmap agreed⁵¹.

During this reporting period, the e-Infrastructure Commons work package focussed its effort on finalising the design of the solutions depicted in the first year of the project and on starting the development of such features:

- Task JRA1.1 Authentication and Authorisation Infrastructure enhanced the pilot of the new AAI infrastructure with the introduction of new modules and features in the IdP/SP proxy and the support of early adopters;
- Task JRA1.2 Service Registry and Marketplace continued the assessment of the technologies to implement the service registry and marketplace following the criteria defined during the first year of the project.
- Task JRA1.3 Accounting enhanced the accounting repository and released the first prototype of the new accounting portal.

⁴⁹ <https://indico.egi.eu/indico/category/152/>

⁵⁰ <https://wiki.egi.eu/wiki/EGI-Engage:WP3>

⁵¹ <https://documents.egi.eu/document/2836>

- Task JRA1.4 Operational Tools released new versions of Operations Portal, ARGO and GOCDDB and made several improvements on the SECANT tool.
- Task JRA1.5 Resource Allocation – e-GRANT worked on a first prototype of a pay-for-use platform that integrates two EGI processes, pay-for-use and resource allocation, in a unique system.

3.2 Main Achievements

3.2.1 Authentication and Authorisation Infrastructure

The core component of the new AAI infrastructure, designed during the first year of the project, is an IdP/SP Proxy based on SimpleSAMLphp technology, which acts as a Service Provider (SP) for the supported identity federations. At the same time it acts as an Identity Provider (IdP) for the services deployed in the EGI Infrastructure.

After the conclusion of the design phase of the new AAI infrastructure and the beginning of the deployment of the pilot, the AAI team focussed its activities on this period on two main areas: (1) gradual introduction of new modules and features in the IdP/SP proxy, (2) identification and support of early adopters to run test campaigns, refine the design and, consequently, improve the pilot infrastructure.

As result of the work achieved in the first area, the IdP/SP proxy is now offering several new features such as the support for

- different Levels of Assurance (LoA);
- user authentication via OAuth2 and OpenID Connect, including social credentials (Google, Facebook, LinkedIn and ORCID);
- querying Attribute Authorities via SAML2 Attribute Queries and REST APIs;
- generating long-lived, non-reassignable, non-targeted, opaque and globally unique user identifiers based on the attributes received from the identity provider;
- user enrolment (flows to be revised based on feedback) and account linking via the integration with the CManage Registry⁵²;

In addition, the IdP/SP proxy has been interconnected with a CILogon-like Token Translation Service able to generate both End-Entity Certificates and Per-User Sub-Proxies (PUSPs), to the GOCDDB as Attribute Provider (via REST API), to RCauth.eu Where Are You From (WAYF) Service and to three EGI tools as service providers (AppDB, GGUS and GOCDDB). The team is also currently testing the integration with the EGI FedCloud Service Provider using SAML.

A specification of eduPersonEntitlement attributes for encapsulating authorisation information about VO/group membership has been also defined and is currently under review.

Regarding the activities in the second area, an extensive collaboration has been established with ELIXIR that has brought to the full integration of the ELIXIR SAML Identity Provider Proxy with the

⁵² <https://spaces.internet2.edu/display/CManage/Home>

EGI IdP/SP proxy and, then, with the new AAI infrastructure. As result of this, ELIXIR users can now access the EGI GOCDDB and AppDB services with their ELIXIR credential⁵³.

3.2.2 Service Registry and Marketplace

The first half of the second project year has been devoted to advance in the assessment of the technologies to implement the service registry and marketplace following the criteria described in the deliverable D3.2⁵⁴ such as the completeness of solutions against the requirements, compliance with the AAI architecture of EGI.eu, costs of development incurred to support a set of priority requirements and maintenance, etc.

To evaluate the completeness of the solutions and adequacy of the current features, it has been decided to evaluate live service instances for two possible solutions: OpenIRIS and Prestashop.

OpenIRIS is the solution that has been identified has the most promising during the first evaluation phase. Instead, Prestashop is a well-known solution to develop marketplace, largely diffuse in the commercial world. In both cases, the first assessment was done with just the documentation.

A draft data model has been defined to allow the implementation of both prototypes that are currently under testing and evaluation. This data model is now taking into account only the services published in the EGI service catalogue. In addition, it has been started the definition of the user workflow within the marketplace that will be an additional criteria that will be used to perform the assessment of the live instances (easiness/complexity to implement the wished workflow in the two prototypes).

In the meantime, an online service catalogue, listing all the EGI services, has been included into the new EGI website that will be published in the coming weeks.

3.2.3 Accounting

During this reporting period, both the components of the EGI accounting system, the repository and the portal, were enhanced with several improvements.

The repository has been enhanced with new features for Cloud, CPU and Storage accounting. For Cloud accounting, changes required to support version 0.4 of the Cloud Accounting Usage Record were implemented and will be rolled out into production imminently at the same time as changes to support the accounting of long running VMs. For CPU accounting, support for a new log format used by Torque has been added. For Storage accounting, third party storage scripts have been improved and the central repository has been updated to handle these changes. In addition, work has been done to introduce a new metric in the repository, the Dataset accounting. After having collected input from stakeholders, a proposed usage record to use for data set accounting has been produced and has been disseminated for feedback and revision.

⁵³ https://wiki.egi.eu/wiki/EGI_AAI_integration_with_ELIXIR_AAI

⁵⁴ <https://documents.egi.eu/document/2658>

After the conclusion of the design phase in project year one, activities related on the accounting portal has been focused on the development of the new portal, using new web tools allowing improving look and feel and creating a modern interface that offer a better experience to the end-users. A first release of the portal has been released in May⁵⁵. It will replace the current production instance⁵⁶ in the next months.

3.2.4 Operations Tools

This task includes the work related to the Operations Portal, ARGO, GOCDB and the Security Monitoring tools. The technological evolution of these tools has been driven by the need to support new technologies (e.g. cloud) and to satisfy new requirements emerging from service providers and user communities, in particular from the Research Infrastructures (RIs) contributing to EGI-Engage via the EGI Competence Centers (CCs) and the Resource Providers (RPs) who contribute infrastructure services to the federation.

During this period, the version 4.0 of the Operations Portal has been released with the evolutions on different thematises such as

1. “frameworks & JS libraries” to increase the maintainability and the efficiency of the applications with the adoption of standard libraries (e.g. Google Chart)
2. “module and project modifications” to increase the performances through a review of the structure of the code
3. “ergonomics” with a standardization of the presentation layer
4. “new Downtime module” with a new subscription page for the downtime announcements and new visualisation facilities (timelines charts, tables and a search tool)
5. enhancements on the “Continuous Integration process” with the definition of a new procedure of developments⁵⁷.

In addition, the VO Administration and operations PORTal (VAPOR), which now supports cloud technology, has been extended with a GLUE2 based resource browser and is fully integrated with the Operations Portal, has been further enhanced with a new ergonomic and a better look and feel.

About ARGO activities, a new version (v1.6.7-1) of the ARGO Compute Engine was released and deployed in production. This version introduces the support for multiple monitoring engines, it adds the framework for the real-time ingestion service (not used in production yet), it introduces authorization capabilities in the Web API and includes documentation, stability and performance improvements. Apart from the production release, there have been 10 new development releases focusing on the ingestion API.

The ARGO Web UI added a new report for the ELIXIR community and updated the SAML engine in order to integrate eventually with the EGI AAI. Several graphics improvements have been added to the Web UI.

⁵⁵ <https://accounting-next.egi.eu/>

⁵⁶ <https://accounting.egi.eu>

⁵⁷ https://forge.in2p3.fr/projects/opsportaluser/wiki/Development_Procedure

The POEM service has been actively developed in order to introduce initial support for the management of the probes that are used in the POEM profiles. A new development version has been released and is tested in the development testbed.

Finally, the beta version of ARGO Messaging Service was introduced at the EGI conference in Amsterdam. During this period we have completed a full implementation of the core Google PubSub API specification and currently we are testing it on the development testbed.

The work of the GOCDDB product team focussed on the following core functionalities, released fewer than three main production versions (v5.5, 5.6 and 5.6.1):

- Role Abstractions for Multi-Tenant: in order to satisfy EGI business rules, GOCDDB requires a comprehensive role based permissions model. The old role/permissions logic was abstracted and improved to allow all roles and the actions they permit to be defined in a single role-action-mapping XML file. This approach allows roles to be applied to different projects hosted in the same GOCDDB instance.
- Reserved Scope Tags and Rules: new 'Reserved-tags' were developed to restrict the uncontrolled tagging of resources. This includes new inheritance rules to allow reserved tags to be cascaded to child resources. This mechanism prevents non-authorised sites and services from using tags not intended for them and were added in response to discussions with the WLCG Information-Systems working group. A large-scale GGUS ticketing campaign was completed in order to add the relevant reserved tags to all WLCG resources for the core WLCG VOs (atlas, alice, lhcb, cms). This mechanism was also applied for the ELIXIR resource centre.
- Downtime Calendar View: a new calendar that allows fine-grained resource filtering according to scope tags and many other optional filter parameters.
- Bulk addition/upload of multiple custom properties: a number of new interfaces were added to allow the batch upload/definition of multiple custom properties for different resource types.
- Integration of GOCDDB with the EGI Proxy-IdP for federated access: users can login to GOCDDB via the EGI proxy IdP after they authenticate to the IdP with a scheme that provides an adequate level of assurance (LoA). GOCDDB serves as both an attribute provider and client to the IdP. A number of investigations were completed regarding personal-data protection rules and opacity/re-publishing of federated attributes. Findings were published on the GOCDDB wiki and fed back to EGI for the IdP work.
- Forced paging on API queries: selected API queries perform paging by default even if the 'page' URL parameter is not specified. This was needed to manage server-loading and concurrency.

Finally, EGI security experts made several improvements on SECANT, an automated framework to detect common flows in VM images. It is now able to create a VM in a contained environment and run a series of probes such as “nmap”, “nessus”, “weak passwords”, “Pakiti”, “configuration flaws (Lynis)”, etc.

3.2.5 Resource Allocation – e-GRANT

The period was devoted to preparations and final development of the second release of e-GRANT. In cooperation with EGI pay-for-use working group and EGI Operations team e-GRANT's development team focused on several areas:

- integration of two EGI processes: Pay-for-Use and Resource Allocation in one system
- testing and development of improvements in the Pay-for-Use process
- modifications in the pay-for-use process due to changes in the negotiation process
- test deployment of joint system (PfU and RA) to production
- initial plan for integration with EGI Market Place
- integration of two EGI processes

In order to support two processes which rely on different procedures, the business logic implemented in e-GRANT needed to be extended. The documents' machine state upon which the system relies was upgraded and customized to support numerous business scenarios and act as a unanimous system for EGI Customers.

Testing and development of improvements in the Pay-for-Use process

After the development of the system in accordance to the technical documentation the Release and Deployment process was started and groups of interest were asked to perform the user tests in the system. Remarks were collected and major points involved in the system's final development

Modifications in the Pay-for-Use process due to changes in the negotiation process

Further changes to the e-GRANT's business logic needed to be implemented due to simplifications of Pay-for-Use negotiation scenario. The involvement of the EGI Broker in the process was removed and, as a result, the negotiation is taking place between the Customer and Resource Administrator.

The data schema was changed to improve users' experience in resource discovery. Resource pools become visible in the same way as other documents in the system.

The function to search and summarize resources available in the system for a given resource request was improved to increase performance and show more information to the user.

The SLA signing action is triggered only after user's action of consent (checkbox).

Test deployment of joint system (PfU and RA) to production

The system was released in beta version to allow first business customers to engage in the Pay-for-Use process and test its execution. At the same time testing of the joint system from the operational side is taking place.

Initial plan for integration with EGI Market Place

Integration with new EGI Market Place was discussed and few options were considered. Decision on the final scenario will be taken by EGI.eu and plans for e-GRANT development will be adapted accordingly.

3.3 Issues and Treatment

Issue: In June, 2 part time developers working for the ARGO team at GRNET, left the project.

Treatment: Developers have been partially replaced and the time is expected to hire also another person within 2016 Q3.

3.4 Plans for next period

3.4.1 Authentication and Authorisation Infrastructure

In the next period, the team will work (1) on refining the current deployment of the pilot according to the user feedback, such as the refinement of user enrolment flows and the optimisation of Perun Attribute Authority connector in multi-VO use cases, (2) on completing the deployment of the IdP/SP proxy with the last lacking features, support for X509 authentication and OIDC services and (3) on integrating further service providers, such as the EGI FedCloud sites and the OpenData platform.

Furthermore, collaboration with EUDAT will be launched with the aim to deploy an EGI/EUDAT AAI integration pilot that would allow EGI and EUDAT users to exploit both EGI and EUDAT services with a unique credential.

3.4.2 Service Registry and Marketplace

In the next period, the assessment of the OpenIRIS and Prestashop live instances will be completed and the data model will be further detailed and finalised. Furthermore, the team will start to work on the integration with the other EGI tools.

A study to publish thematic platforms within the marketplace, in addition to the EGI service catalogue, will be done to understand how providers of these platforms can benefit from the marketplace. Thematic platforms will be an example of category of services not owned by EGI that could be published in EGI marketplace.

Finally an analysis of specific cloud marketplace enablement tools, such as AppCara, App Marketplace, Juju, Alien4Cloud & Cloudify, will be done to understand their integrability with the EGI marketplace.

3.4.3 Accounting

In the next period, the accounting team will roll out storage accounting to production. This will require the introduction of storage summaries to keep the load on the portal manageable. The team will also investigate other storage publishers with the relevant developers.

In addition, once the metrics to use for data set accounting have been revised and agreed upon, the accounting team will develop a prototype that will initially extract accounting information from the OpenData platform.

Improvements will be also made to the way we summarise cloud records so that more information is available in the Accounting Portal. This period will also see the second release of the Accounting Repository which will include many small improvements as well the larger changes covered earlier.

Finally, the work required to move the new Accounting Portal in production and to phase out the old one will be achieved. It includes the validation of the new portal from its main stakeholders.

3.4.4 Operations Tools

Operations Portal

The portal will be integrated with the new EGI AAI and will be accessible through the EGI IdP/SP proxy. The integration with Perun will be enhanced with the automatization of several interactions. A new version of the dashboards will be released that will include a refactorization of the code, the simplification of the interfaces and user customizations. In VAPOR, the job monitoring scripts will be improved and two studies will be done to verify the possibility (1) to collect information about Cloud Storage and (2) to monitor virtual machines.

ARGO

The work of the ARGO team will be focused on four main areas:

- Service for managing probes: the POEM service will be extended and authorized users will be able to upload and manage monitoring probes from a web based services, guaranteeing a faster management/deployment of new probes. In addition, a built-in testing environment will be available to validate a new probe before it goes to production. Pre-release: 2016Q3 / First release: 2016Q4
- Real-time status results: introduction of a Streaming Layer in the ARGO Compute Engine that will allow to process and publish status results as they arrive. Ability to create composable computation pipelines. Pre-release: 2016Q3 / First release: 2016Q4
- Overhaul of the notification system: the new streaming layer will be used to move notifications from the Monitoring Engines to the Compute Engine. Pre-release: 2016Q4 / First-release: 2017Q1
- New messaging service: the specification will be finalised and the “Per topic/subscription authorization” (2016Q3) and the “Messaging metrics” (2017Q1) will be developed. In addition the service will be connected to the new EGI AAI (2017Q1). Pre-release: 2016Q3 / First-release: 2016Q4 / Second-release: 2017Q1

GOCDB

Future work will follow the GOCDB roadmap. This includes:

- Write API: provide a REST API intended for client scripts for performing CRUD operations on selected resources (sites/services/endpoints/custom-properties). This will require site administrators to manage their own access control lists (ACLs) via the GOCDB portal. The ACLs will declare which (other) certificates are allowed to post updates to a particular resource (e.g. robot certificates). Other mechanisms for defining ACLs such as

defining API access Keys will also be considered. The first release of this work is planned for testing end August/Sept.

- Work will also be informed by the requirements raised by WLCG in response to their plans to deprecate their usage of the BDII and to move more information from the BDII into GOCDB. Similar requests have been raised by other collaborators such as using GOCDB to define Grid-service endpoints and attributes rather than publishing this data in BDII.
- Other work includes previous targets that were previously de-prioritised. This includes capturing more detail in automatic object-diff auditing, addition of new business rules, and MVC/GUI refactoring.

Security Monitoring

The team will proceed with the pilot of Secant and will examine different possibilities to integrate it with AppDB and/or used to start assessing VA's available for fedcloud. A new version of the SECANT tool will be released at the end of the year 2 of the project.

3.4.5 Resource Allocation – e-GRANT

Due to intensified work around new joint platform and its production deployment, some of the tasks planned for the current period have been moved to the last quarter of 2016:

- plan the integration with EGI monitoring framework
- plan the integration with EGI Accounting System and Accounting Portal

Rest of the tasks planned:

- integration with new EGI AAI
- project and initial development of a new front-end according to EGI's recommendations
- changes to the Resource Allocation process in accordance with the new requirements
- improvements in technology and architecture of the platform

4 Platforms for the Data Commons

4.1 Summary

The main purpose of this work package is to advance the current technical infrastructure of EGI by expanding the capabilities of the current platforms, and by integrating new ones. The result of the activity will be an integrated solution of data and compute services that will contribute to the Open Commons solution. It will do so by further evolving the EGI Federated Cloud infrastructure platform to provide the integrating services and users with greater flexibility and elasticity in the overall use of the platform, as well as ensuring continuity in the support for Cloud Middleware Frameworks.

This work package will also introduce an Open Data Access platform that will provide capabilities to publish, use and reuse openly accessible data (including, but not limited to, scientific data sets released into the public domain, publicly funded research papers and project deliverables, and software artefacts and demonstrators coming out of public research projects). Ensuring support for a broad number of use cases and data commons needs, activities in this work package will also include integration activities of a number of partner e-Infrastructures both located in Europe and worldwide. This will include integrating existing cloud infrastructures with the EGI Federated Cloud platform (e.g. the Canadian CANFAR infrastructure) and accelerated computing facilities (e.g. GPGPUs – general-purpose computation on graphics processing unit).

During the reporting period of this report, the following summarizes the work of each activity: Within JRA2.1 (Federated Open Data), essential functionality for the EGI Open Data Platform (ODP) has been developed, in addition to starting preparations for the launch of the ODP prototype in November 2016 and an initial demonstration of the ODP at the DI4R conference in September 2016. Within JRA2.2 (Federated Cloud), development has progressed on key cloud standards (OCCI) and components (OOI, Keystone-VOMS) in addition to work done on the cloud accounting infrastructure. Finally, work has been done developing new VM management functionality within the FedCloud AppDB component. Within JRA2.3 (e-Infrastructure Management) work has continued in integrating EUDAT and CANFAR. The technical integration of gCube/D4Science is now complete and an SLA is now being negotiated between EGI and D4Science. Within JRA2.4 (Accelerated Computing) work has progressed on both the High Throughput Computing (HTC) and Cloud computing aspects of GPGPU computing. For HTC side, prototypes have been designed and implemented for a number of batch systems (HTCondor, Slurm, SGE and LSF). The GLUE information system has also been extended to allow for GPGPU computing as part of this work. For the Cloud side, work has progressed to bring GPGPU computing to both OpenStack and OpenNebula by configuring and optimising Federated Cloud resources at different EGI research centres.

4.2 Main Achievements

4.2.1 Federated Open Data

The goal of Open Data Platform is to provide a solution to integrate various data repositories managed by different user communities within EGI. The Open Data Platform will focus on providing a federated, open access, unified platform for users and application developers in order to foster research results reuse in Europe.

During the course of the reporting period, the main activities within JRA2.1 have been oriented towards development of first prototype version of the Open Data Platform, based on the requirements analysis and architecture design performed in the previous period. Over the first months of this period, the main activities included ensuring that necessary low level functionality necessary for open data access were introduced into the ODP underlying solution, i.e. Onedata. This included creation of snapshots from data spaces, advanced metadata editing capabilities, publishing of snapshots as open data sets, previewing open data sets in an anonymously available web interface and comprehensive REST API allowing for integration with existing platforms and infrastructures. All these functionalities have required development on both the data management system level as well as at the level of user interface.

Furthermore, integration with selected communities have commenced, most importantly with LifeWatch, in order to demonstrate a complete open data cycle from inception and ingestion into the data management system, through curation and enrichment with metadata to publishing in open access registries and finally reuse of this data by end users.

Onedata has also been used as a testbed for evaluating Human Brain Project data visualization service, which has demonstrated the capability of Onedata to meet the requirements of a data-intensive neuroinformatics use case, both in terms of reliability and speed.

The demonstration of the prototype of Open Data Platform on LifeWatch use case will be demonstrated during the Digital Infrastructures 2016 conference Krakow.

4.2.2 Federated Cloud

The main objective for this task is the evolution of the existing IaaS Cloud with new capabilities, focusing on satisfying the incoming requirements from the Competence Centers (CCs). In order to track and prioritize the needed developments, the "Requirements" queue in the EGI RT has been used. Whenever possible and appropriate, these improvements are being pushed into the Open Cloud Computing Interface (OCCI) standard, so as to have a way to expose the newly developed capabilities. Moreover, this task also performs the maintenance of the existing integration tools, required for a resource center to be part of the EGI cloud, such as the OCCI interfaces, authentication and authorization modules, VM image management, accounting, and information system.

During this period JRA2.2 has discussed and elaborated an implementable draft (D4.4) describing the interactions between resource providers and users in order to reallocate a paused VM from one resource provider to another.

Regarding OCCI support for the various Cloud Management Frameworks:

- rOCCI have been updated so as to support OCCI 1.2, implementing the VM snapshot and VM resize extensions developed in the previous period (D4.2 and D4.3).
- OpenStack OCCI Interface (ooi) has been enhanced so as to support the OpenStack Networking component (Neutron) as well as nova-compute.

The OpenStack Keystone-VOMS integration module has been updated to support the latest OpenStack production versions (Liberty, Mitaka), released during this period. Moreover, this component has been completely refactored, migrating from M2Crypto into PyOpenSSL, making possible the support of Python 3.

In addition, Keystone support for the V3 has been successfully tested as a proof of concept. Modifications to other software components are needed (namely rOCCI-cli to support the new authentication flow, and the Apache mod_gridsite to be able to pass down the VOSM FQANs using the HTTPD environment variables. The usage of Keystone V3 does not require any additional module in the server side but only a specialized configuration.

Regarding the accounting collection tools:

- A new version of cASO, the accounting collection tools for OpenStack has been released: <https://appdb.egi.eu/store/software/caso>, supporting the V0.4 of the accounting record.
- A new version of oneacct-export, the accounting collection tools for Opennebula has been released: <https://appdb.egi.eu/store/software/oneacct.export>, supporting the V0.4 of the accounting record.

A prototype implementation of the GLUE 2.1 Cloud profile in the information system provider⁵⁸ has started. This task is being carried out in coordination with the Open Grid Forum GLUE working group. The information provider has been refactored so as to ease the implementation of new cloud profiles and information renderings (like JSON) thus making possible to reuse the same tool to create different representations of the collected information.

The current VMI synchronization component, vmcatcher, is being phased out and is being replaced with a new component developed by the new tool maintained within EGI by CESNET. This new component is still in the design phase, but it will remain interface-compatible with the existing toolchain.

The plans for extending the EGI Applications DataBase (AppDB) have resulted in a new VM Operations dashboard⁵⁹, integrated with the EGI AAI. Through this new portal, EGI users are able to launch and manage and launch standalone VMs as well as more complex virtual infrastructures

⁵⁸ <https://github.com/EGI-FCTF/cloud-bdii-provider>

⁵⁹ <https://dashboard.appdb.egi.eu/>

(topologies), based on the VM Appliances registered in the AppDB (thus available at the individual sites). The VM Operations dashboard leverages the Infrastructure Manager (IM)⁶⁰ and its OCCl connectors to perform the complex VM topology deployment and management.

4.2.3 e-Infrastructures Integration

This task objective is to further expand the EGI capacities and capabilities with the integration of other e-Infrastructures. The proposal included three target e-Infrastructures and technical solutions for integration: EUDAT, the Canadian Advanced Network for Astronomical Research (CANFAR), and the gCube/D4Science infrastructure (D4Science), but the task is not limited to those and has started to contact external partners for the integration of their solutions and e-Infrastructures into EGI.

JRA2.3 has continued work with the integration of identified platforms in the proposal:

- EGI-EUDAT integration task continues to develop the selected use cases to support the use of services of both infrastructures. ICOS is the most relevant use case for this collaboration, and current work is focused in accessing selected EUDAT services from EGI Federated Cloud resources. A call for participation for cross e-Infrastructure case studies was produced by EGI (M4.2) and is waiting for EUDAT comments to be finally made public and collect new use cases.
- The CANFAR integration progresses as defined in D4.1 where the roadmap of the integration activities is detailed. There is a working implementation of the GMS (Group Management System) ready and an enhanced version is undergoing fixing the identified bugs and issues during this period. Work on the implementation of the VOspace storage system has also started.
- The technical integration of gCube/D4Science is completed and pilot use cases can make use of the EGI FedCloud to scale up and execute their models when needed. Currently an SLA is being negotiated to formalise the use of EGI resources by the D4Science platform.

Apart from these platforms, the EGI Federated Cloud is testing the integration of the new EGI AAI based on SAML and OpenID Connect to allow certificate-less access to the cloud resources to users. The work has started with OpenStack, which is the predominant Cloud Management Framework deployed at the EGI infrastructure, and using upstream developments to facilitate the integration for providers not familiar with EGI. Contacts with SURFSara (NL), CSC (FI) and SNIC (SE) have been established to provide them with the necessary information to evaluate the current federation and proceed with the integration.

4.2.4 Accelerated Computing

Driven by the user communities represented in the EGI-Engage Competence Centres (MoBrain and LifeWatch), and other ones making use of the EGI infrastructure (Virgo, LHCb, and Molecular

⁶⁰ <http://www.grycap.upv.es/im/index.php>

Dynamics), this task is aimed at enabling GPGPU support both on High Throughput Computing (HTC) and Cloud platforms.

For the HTC platform, the plan is to provide full GPGPU support for CREAM-CE with the most popular batch systems, also by exploiting the new attributes defined in the GLUE2.1 schema. For the cloud side, the plan is to investigate the current Passthrough and vGPU technologies to prepare specialised VM images for enabling GPGPU support in the EGI Federated Cloud.

Over the reporting period of this report, the following has been achieved on the HTC side:

- A GPGPU-enabled CREAM-CE prototype designed to work with the HTCondor Local Resource Management System (LRMS) has been implemented and tested at the GRIF HTCondor-based GPU/MIC cluster. This paves the way for adding support to two new JDL attributes - MICNumber and GPUModel - expected to be supported by latest version of HTCondor, Slurm and LSF batch systems. With these two new attributes, the user is able to select the computing nodes which host one or more MIC cards. It also addresses the case where a computing node hosts different models of GPGPU cards.
- A GPGPU-enabled CREAM-CE prototype designed to work with the Slurm and SGE LRMSes has been implemented and tested at ARNES Slurm based GPU cluster and QMUL SGE based GPU cluster
- The most recent activities of this task were presented at the EGI Conference 2016 in “Accelerated Computing: latest developments and use cases” session⁶¹
- The GLUE2.1 draft has been updated by adding new accelerator card specific attributes, for both grid and cloud infrastructures. Representatives of the task participated to the OGF GLUE-WG meeting of 17th of May and an updated version of the GLUE2.1 draft was produced, together with plans to implement a prototype for the infosys based on it. Future official approval of GLUE 2.1 would occur after the specification is revised based on lessons learned from the prototype

The work for enabling support of accelerated computing in the grid environment officially stopped at the end of May 2016 within EGI-ENGAGE. However, the CREAM developers’ team are committed to producing a major CREAM-CE release for the SL6 and CentOS7 operating systems with the new accelerated computing capabilities by the end of 2016⁶². This work is being funded separately from EGI-ENGAGE.

Over the reporting period of this report, the following has been achieved on the Cloud side:

- Tests were performed to verify the possibility to migrate GPU-enabled VMs within the OpenStack based cloud infrastructure. VM migration is needed for the maintenance and applying of security updates on worker nodes. The VM migration was successful with cold restart of GPGPU-enabled VMs. Live migration is not supported due to current limitations within Linux kernels and drivers.

⁶¹ <https://indico.egi.eu/indico/event/2875/session/40/?slotId=0#20160408>

⁶² <https://wiki.italiangrid.it/CREAM>

- A status report of the cloud related activities was presented at the EGI Conference 2016 in “Accelerated Computing: latest developments and use cases” session⁶³
- The GLUE2.1 draft was discussed at the OGF GLUE-WG meeting and updated with the relevant accelerator card specific attributes.
- GPGPU experimental support was enabled at the CESNET-Metacloud site, based on the OpenNebula cloud management framework. VMs with Tesla M2090 GPU cards were instantiated and tested with a various molecular dynamics programs.
- An experimental cloud site to enable GPGPU support with LXC/LXD “hypervisor” with OpenStack has been created. LXC/LXD is full container supported by Linux, and could provide better performance and stability than KVM (must faster startup time, better integration with OS), especially in terms of GPGPU support (simpler site setup, more stable than KVM PCI passthrough) however, that requires a new VM format with the same kernel at the hosts.
- A new OpenNebula site with GPGPUs has been created at IISAS and is being integrated into the EGI federated cloud. Users are thus able to benefit from equal support for GPGPUs on both CMFs (OpenStack and OpenNebula).

4.3 Issues and Treatment

No issues have been identified during reporting period.

4.4 Plans for next period

4.4.1 Federated Open Data

In the next period, the mains goals of JRA2.1 will be concerned with extending the open data management functionality and community outreach.

In particular the further improvements will include integration with open access sites such as OpenAIRE⁶⁴, automation of generating handlers for open data sets, either DOI or PID and support for Data Management Plans related to open data sets.

The community outreach will be at first targeted towards EGI-Engage communities and then extended to other research communities which could potentially benefit from open data access management solution. The outreach will be organized by a series of webinars with various levels of complexity focusing on various features of the Open Data Platform (ODP) from the point of view of the typical end user, through the perspective of Principal Investigators and data curators and to the system administrators responsible for storage resource management.

Furthermore, Open Data Platform will be integrated into the EGI DataHub service, a single point of entry for data management functionality provided by EGI, where it will be responsible for allowing users to easily discover and manage open data.

⁶³ <https://indico.egi.eu/indico/event/2875/session/40/?slotId=0#20160408>

⁶⁴ <https://www.openaire.eu/>

During the next reporting period, first implementation of the Open Data Platform will be made available to selected communities for preliminary testing and evaluation. In the next periods further integration with the user communities is planned as well as extending the Open Data Platform functionality in order to meet their functional as well as non-functional requirements as well as possible.

4.4.2 Federated Cloud

The expected steps for the upcoming period are focused on the evolution of the existing tools and integration modules to support the latest standard and interfaces versions.

In this context, the information system modules will implement the GLUE 2.1 Cloud profile, as defined by the OGF, once it becomes final. During this period it is expected close collaboration with the OGF in order to finalize the drafting phase, solving any possible problem that might arise. Besides, we expect to complete the OCCI 1.2 support in Synnefo, as long as fix any possible issue present in the OCCI 1.2 implementations of ooi and rOCCI.

As we already stated, the VMI synchronization tool is being rewritten, maintaining the existing interfaces thus preserving the compatibility with the current tool chain. We expect to have a released in the upcoming months, easing its deployment and improving its reliability.

Regarding authentication and authorization, we expect to elaborate a configuration guide for enabling OpenStack Identity (Keystone) V3 to leverage VOMS authentication without using external modules like Keystone-VOMS. Once this is done, we will add support for V3 in the client libraries (both the OpenStack native libraries and the OCCI clients) and design a deprecation plan of the Keystone VOMS V2 authentication. Moreover, it is expected that the EGI-AAI solution becomes available in this new period, so, in parallel, we will integrate its support in the V3 developments. This way, whenever the VOMS V2 authentication is deprecated we will be able to migrate users to the new identity APIs.

4.4.3 e-Infrastructures Integration

EGI-EUDAT work will continue supporting the selected use cases and further develop the integration focusing on AAI integration and the use of the new EUDAT APIs. As soon as EUDAT provides comments to the joint call for case studies, this will be released and work on any new cases will be started.

The collaboration with CANFAR will release a second version of the Group Management System (GMS) that enables both Canadian and European researchers to login and interact with the system. Monitoring on this service will be developed and enabled so to obtain availability and reliability metrics.

The SLA with gCube/D4Science should be agreed and VREs will be able to start using the EGI Federated Cloud in production.

Integration with SAML and OpenID Connect for OpenStack resources will be completed and for OpenNebula resources there will be an evaluation of the possible ways to perform a similar integration.

Follow-up meetings with several resource providers will be held to understand the feasibility of their integration with the EGI Federated Cloud in the short-term.

4.4.4 Accelerated Computing

In the next period, a new site with GPGPU and OpenNebula will be created and integrated into the EGI FedCloud. That will provide support for accelerated computing on both major Cloud Middleware Frameworks (OpenStack and OpenNebula) in EGI Federated Cloud.

Work will continue on enabling GPGPU support on containers (LXC, Docker). Discussions with user communities are expected to ensure that user requirements are met and that required support is provided.

A modified accounting scheme with included information about usage of GPGPU will be created. Site BDII with GPGPU information included according to GLUE2.1 will be created

5 Operations

5.1 Summary

WP5 has continued in the past six months in supporting the EGI production infrastructure by:

- Coordinating the operational activities and technology provisioning
- Developing the security activities
- Work towards the deployment in production of new services in EGI

Task SA1.1 has successfully provided coordination of the operational activities of the operations centres through regular OMB and Operations meetings. Operational procedures and OLAs have been updated to reflect the new requirements of the production infrastructure.

Technology provisioning has continued with regular UMD updates, three updates of UMD-4 and one update of UMD-3 during the reporting period. SA1.1 has liaised with the developers to have a smooth transition to EL7 based products. Decommission of SL5 from the production infrastructure has been completed well before the end of life of the operating system (2017). The architecture of the middleware distribution for cloud components (CMD) have been agreed with the relevant stakeholders, the technical infrastructure for software provision for the federated cloud is ready, and first release is expected at the end of summer.

Task SA1.2 progressed in the evolution of the security operations of EGI. The security risk assessment has been used to identify the main security threats and the first mitigations have been put in place.

A new security policy (Virtual Machines Endorsement) was approved, and the security incident handling procedure has been updated to reflect specific for the cloud use case. The processes to handle software security vulnerabilities have been further improved.

SA1.2 has continued to liaise with external projects, OSG and XSEDE, and security bodies, such as the WISE community, euGridPMA and IGTF.

The task is closely collaborating with JRA1.1 in the definition of the security framework to support the new EGI AAI services.

Task SA1.3 focused on preparing the support for three main platforms: gCube, ESA e-challenge and long tail of science.

gCube will use the federated cloud resources to run generic computations and data-mining applications. Two virtual appliances have been built specifically for the EGI Fedcloud use, and tested using the incubator federated cloud catch-all VO. Production is bound to start in October 2016, and there are already a number of use cases that are interested to use fedcloud resources through gCube: Scalable mining, BiodiversityLab, Stock assessment, Aquaculture, Protected Area Impact Maps.

Through the integration of gCube other use cases currently using D4Science can benefit from EGI resources, in particular expressions of interest have been collected from ENVRIPPlus and SoBigData.

Terradue, responsible for running the ESA exploitation platforms for the hydrology and geohazards use cases, has finalized the integration with fedcloud and successfully testing.

Long tail of science platform continued the work of integration with new tools, like science gateways, to enable LTOS users to access EGI resources. New cloud resources have been added to the pool to reach the minimum critical mass to allow the first users to access cloud services.

5.2 Main Achievements

5.2.1 Operations Coordination

Operations coordination has been provided by EGI Foundation and INFN.

The task has provided coordination for the operational activities of the EGI production infrastructure. The activities can be grouped in the following categories.

Coordination of the NGI and Resource centres operations

The Operations Management Board⁶⁵ continued to be held monthly with representatives of the operations centres and main core activities. The meeting's agenda included the discussion for approval of new policies and procedures (described in Documentation subsection), as well as more forward looking topics about the evolution of the operational services of EGI. Another monthly meeting, the Operations meeting⁶⁶, involves the operations centre staff and the agenda focuses on more technical topics. Both meeting are chaired by the operations coordination team, and are usually held, respectively, in the last Thursday and in the second Monday of the month.

Operations coordination supported resource centres and user communities participating to a VO SLA, helping in configuring the resource centre services following the VO requirements, making sure the users are able to productively access the resource centres listed in the VO SLA. During the reporting period the following VOs have been supported in the after-sale process: DRIHM, Mobrain, BILS, ESA geohazards and hydrology, EXTRAS.

Operations coordination liaised with the operational activities in the WLCG user community to ensure a strong collaboration and an alignment of the operational activities. This has been mostly achieved through an EGI representation in the WLCG technical working groups.

Technology management and software provisioning

The task coordinates the release of new software through the EGI software provisioning process, and also the decommissioning of outdated versions from the production infrastructure. Operations coordination also liaise with the product teams who develop the middleware released in UMD. Fortnightly UMD release team meetings are organised with the software provisioning team and the developers to discuss release timelines, technical plans, and issues.

⁶⁵ <https://indico.egi.eu/indico/category/19/>

⁶⁶ <https://indico.egi.eu/indico/category/32/>

Starting from September 2015 and until March 2016, UMD3 supported only security fixes for SL5 operating system. A decommissioning campaign started in January 2016 and concluded in the months after April 2016, when UMD suspended SL5 support for updates and bug fixes.

Almost all the resource centres with few exceptions handled centrally upgraded to SL6 or CentOS7. Two resource centres were suspended for not responding to operational tickets.

The software released in UMD is produced by a number of product teams, who work independently. To provide an aggregated view of the latest updates for the UMD software EGI Operations provide a “Preview” repository created with the aim of making available in a single point the updates of the middleware products as soon as they are released by the product teams. The same updates undergo to the usual UMD verification and staged rollout processes, which require several weeks to be completed, and through the Preview repository the resource centres administrators can deploy software as soon as it is made available. Preview Repository is provided without any additional quality assurance process because the products are released as they are provided by the product teams, therefore EGI recommends using it only for very urgent operational reasons.

The Preview Repository is updated once every month, collecting contributions from the product teams⁶⁷.

The preparation of the technical infrastructure for the Cloud Middleware Distribution is completed, now the software release team is preparing a dry run to test it before the first release. The task is also collecting information about the product teams who will contribute with software for the federated cloud. The product teams are from the EGI community, for example from the JRA2 work package, and also external, for example the INDIGO DataCloud project.

EGI has also agreed an underpinning agreement with the INDIGO project⁶⁸ for the support of the software that will be released in CMD.

INFN, the developers of the gLite-WMS software, a workload management and broker service for HTC computing (largely used by VOs accessing HTC through opportunistic usage), announced to EGI Operations that their long-term plans do not include the support for the WMS. There will be no glite-WMS porting for EL7, and bug fixing will continue for a limited time. INFN is ensuring support for the short term to avoid disruptions. Operations coordination is investigating possible alternatives to replace the service, at the moment the most promising is DIRAC, a workload management tool already used by some EGI virtual organizations.

Coordination of the core activities

The second phase⁶⁹ of provisioning of the core activities has started in May 2016. A set of OLAs have been agreed with the providers who were assigned to the service provisioning through the second bidding process. Where necessary, operations coordination ensured that the proper handover was carried out between leaving providers and new providers, to minimise any effect on

⁶⁷ <http://repository.egi.eu/category/preview-repository>

⁶⁸ <https://www.indigo-datacloud.eu/>

⁶⁹ https://wiki.egi.eu/wiki/EGI_Core_activities:PHASE_II

service provisioning. A new core service, the registry of virtual appliances AppDB, has been added to the list of supported services, another service “Operations support” in phase I has been re-scoped to focus only on the support for resource allocation and long tail of science.

Availability and Reliability computation for Cloud services has been integrated with the processes already in place to monitor the A/R of the other high throughput computing services. This has been possible as the quality of the service provided by the Federated Cloud resource centres has grown in stability and efficiency, making feasible the merging of the monitoring profile of cloud services with the profile used for the overall A/R calculations in EGI.

Currently the set of probes monitoring the status of cloud services include both tests for the standard virtual machine management interface OCCl, and for the OpenStack APIs. In the process of integration probes that require improvements have been modified and tested for production. In the June OMB⁷⁰ the new critical profile was approved and it became effective since July 1st 2016.

During the reporting period the ARGO monitoring system have been fully centralised. A fully central deployment allows having faster development of the monitoring infrastructure, a more reliable service and to reduce the effort required for the NGIs to monitor their sites, maintaining unaltered the effort cost of the core activity. Previously the monitoring engine was distributed and the result database and availability computation engine was centralised. In July 2016 all the monitoring data is produced by a set of monitoring services deployed in high availability by the partners providing the monitoring core activity. This change required a careful coordination with the NGIs and the ARGO team to avoid any loss of monitoring data and a validation of the results produced by the new monitoring service. The migration has been completed without major disruptions, minor issues have been tracked and - mostly- solved, and any pending issue is going to be fixed within the end of August.

Documentation

A new procedure⁷¹ has been produced, and approved by the OMB, to regulate the deployment in production of new releases of the central operations services. The goal of the new procedure is to track the changes in the operational framework of EGI, and to keep the stakeholders informed of any variations in the tools.

Operations coordination is also participating to the federated cloud documentation maintenance, with a focus on the documentation for the resource providers⁷². The wiki pages have been reorganised to be more effective and to present information more effectively without duplications, and a new FAQ for resource providers has been produced by aggregating the important facts and questions about joining federated cloud.

New policies and procedures, created by other tasks or activities, have been managed pushing the documents through the process of OMB approval. In the reporting period OMB approved also the following policies and procedures:

⁷⁰ <https://indico.egi.eu/indico/event/2807/>

⁷¹ <https://wiki.egi.eu/wiki/PROC23>

⁷² https://wiki.egi.eu/wiki/Federated_Cloud_resource_providers_support

- LTOS AUP⁷³
- LTOS Security policy⁷⁴
- VMI Endorsement and operations policy⁷⁵
- Procedure SEC01⁷⁶

Evolution of the information system

The cloud services information are published using the definitions of GLUE 2 Schema, in the process of strengthen the operational integration of the cloud services SA1.1 has continued in the work of releasing the GLUE2.1 schema, with the introduction of new classes and attributes that better describe the cloud services. In May 2016 there was a meeting with the glue-wg for reviewing the cloud extensions and the outcome was to implement a prototype of the new schema for testing it, before getting the final approval by the glue-wg. This prototype is planned for September 2016.

UMD Releases

- UMD 4.1.0, released on May 27, 2015, including:
 - CentOS (Frontier-Squid 2.7.STABLE9-24.2, dCache 2.13.28)
 - SL6: APEL SSM 2.1.5, APEL 1.4.1, site BDII 1.2.1, top BDII 1.1.4, MyProxy 6.1.15, GridSite, canl 2.2.6, emi-wn 3.1.0, emi-ui 3.1.0, dCache SRM Client 2.2.27, srm-ifce 1.23.1, CGSI-gSOAP 1.3.8, gfal2-plugin-xrootd 0.4.0, gfal2-utils 1.2.1, dpm-argus 1.8.10, xroot 4.2.3, DPM 1.8.10, StoRM 1.11.11, Globus Gram5 6.1.0, cream 1.16.5, cream-ge 2.3.2, ARC 15.03.4, WMS utils 3.6.6, globus-default-security 6.2.0, Globus GridFTP 9.1.0, FTS 3.3.1, CernVM FS 2.1.20, glexec 1.2.3, Frontier-Squid 2.7.STABLE9-24.2, VOMS client 3.0.6, davix 0.5.0, argus-pap 1.6.4
- UMD 4.1.1, release on June 26, including:
 - SL6: edg-mkgridmap 4.0.3, globus-net-manager 0.12.1
- UMD 3.14.2: released on June 2016 including
 - SL6: StoRM 1.11.11, dCache 2.10.56, edg-mkgrid 4.0.3, globus-net-manager
- UMD 4.2.0 released on August 4, 2015, including:
 - CentOS7 (edg-mkgridmap 4.0.3, umd-release 4.1.1, FTS3 3.4.3, QCG Computing 4.0.0, QCG Notification 4.0.0, dCache 2.13.35)
 - SL6 (FTS3 3.4.3, umd-release 4.1.1, CREAM GE 2.0.2, CREAM Slurm 1.0.2, CREAM LSF 2.0.4, WN Slurm 1.0.0, dCache 2.13.25)

5.2.2 Development of Security Operations

The aim of this task is to evolve the security activities in EGI to support new technologies and resource provisioning paradigms, while maintaining a secure trustworthy infrastructure and supporting new use cases and ways to access the resources.

⁷³ <https://documents.egi.eu/document/2635>

⁷⁴ <https://documents.egi.eu/document/2769>

⁷⁵ <https://documents.egi.eu/document/2729>

⁷⁶ <https://wiki.egi.eu/wiki/SEC01>

Ideas and plans for PY2 were discussed and agreed at the January 2016 EGI CSIRT meeting in Prague. Work on policy development for EGI AAI takes place in WP3 (JRA1) and work on the development of security training takes place in WP6 (SA2).

During this period, several meetings were attended by one or more of the partners. Participation in these provides essential input to our activities, for discussion with others and for gaining feedback from EGI, from other e-Infrastructures and from Research Infrastructures and Communities. The meetings and conferences attended included:

- Cloudscape2016, Brussels, 8-9 March 2016;
- The International Symposium on Grids and Clouds (ISGC2016), Taipei, 13-18 March 2016;
- Face to face meeting of the EGI CSIRT, Amsterdam 4-5 April 2016;

Security requirements and risk assessments (STFC)

Work here has been building on the security risk assessment completed in PY1 working on the required controls to mitigate the highest risk issues, and to the participation to the technical working groups in EGI, such as the federated cloud meetings. A document is being prepared on the "Federated Cloud Software, Security and AAI". This will include guidelines to address lengthy discussions, for example on how best to control the contextualisation of running Virtual Machines, and the VM Operator role.

Evolution of operational security procedures (CERN)

Good progress has been made on the development of security procedures and operations. This included work on an updated procedure⁷⁷ for Security Incident Response⁷⁸. Other operational improvements were made to the Resource Provider certification procedure, in the context of re-certification of sites suspended for security reasons. Security operations staff are also working on adopting improvements being made elsewhere, including on the automated security checks of Virtual Machine images and in the whole area of Virtual Machine and User management, e.g. how best to find the list of sites running instances of a particular vulnerable virtual machine image. The IRTF has responded to improvements made in the JRA1 work package on monitoring, such as the security dashboard and the central emergency suspension system.

Development of new trust frameworks and security policies (STFC and Nikhef)

Work in trust frameworks and policy issues for federated AAI has been done jointly with EGI-Engage JRA1. Work on this sub-task has included pushing several trust and policy issues through to approval by IGTF at EUGridPMA and TAGPMA meetings. This included the accreditation by IGTF of a new "CILogon-like" CA for Europe (RCauth.eu) for use in the new EGI-Engage AAI service, an update to the IGTF IOTA profile, work on private key protection models for centralized credential stores and new video-supported identity vetting models.

⁷⁷ <https://wiki.egi.eu/wiki/PROC09>

⁷⁸ <https://wiki.egi.eu/wiki/SEC01>

An updated version of the Security Policy for the Endorsement and Operation of Virtual Machine Images⁷⁹ was produced and approved. Other policy development work, still ongoing, included starting on revisions to the top-level Security Policy document to make it more general and more obviously applicable to all current EGI services, and work on a complete revision of the Data Protection policy framework to eventually replace the old policy on user-level job-accounting data.

Collaboration with other e-Infrastructures (GEANT, EUDAT, PRACE, XSEDE and others) on security trust and development of best practice continues to take place under the auspices of the recently formed WISE community⁸⁰. The partners of SA1.2 have leading roles in this important community, e.g. membership of the steering committee and leading work groups, such as SCIV2-WG, which is defining best practices, trust and policy standards for security collaboration. EGI-Engage partners actively participated in and gave talks at the WISE BoF session at TNC16 and the WISE workshop at the XSEDE2016 and played a leading role in planning for the next WISE workshop in September 2016 just before the DI4R conference in Krakow.

Development of new security challenge frameworks (Nikhef)

Work has continued on the preparation of a security challenge to the EGI Federated Cloud service. The framework is ready and running but more work is still needed on the injection of malware/malicious content, on the integration with SSC monitoring, and the usage of contextualization.

Development of software vulnerability handling processes (STFC)

Investigations have taken place on the Cloud Middleware Distribution (CMD) and how it ties in with the EGI Software Vulnerability Group (SVG) processes. A new advisory template is now in use, while other templates still need further updating. Procedures are being worked on, together with the EGI Federated Cloud group and the EGI CSIRT as to how best to maintain contact with Virtual Machine endorsers & operators so that advisories about vulnerabilities in AppDB VM images can be communicated.

5.2.3 Integration, Deployment of Grid and Cloud Platforms

iMarine and gCube integration

CNR designed, developed, and tested two Virtual appliances that can be exploited by the community to perform data analysis on biological, social, satellite, and environmental data. To this aim, two types of services have been properly packaged and made available as Virtual Appliances.

The first one is composed by a distributed set of web services embedding computational processes for discovering patterns in large data sets involving methods at the intersection of artificial intelligence, machine learning, and statistics (DataMiner Worker).

The second one exploits a set of "lightweight" workers (Generic Worker) that are used for computations, in which one computation is divided into several parts that are low-resources

⁷⁹ <https://documents.egi.eu/document/2729>

⁸⁰ <https://wise-community.org>

demanding. Generic Workers employ one core for service operations and another one for processes executions. Applications include for example the combination of geographical and biological information to produce projections of environmental scenarios, assessing fisheries activities in the global oceans or predicting the impact of climate changes on biodiversity.

In the context of the Fishery and Marine Sciences, CNR promoted the exploitation of those appliances on FedCloud resources. The following expressions of interests were collected by the following operational Virtual Research Environments:

- **Scalable Data Mining:** The Scalable Data Mining is a VRE designed to apply Data Mining techniques to biological data. The algorithms are executed in a distributed fashion on the e-Infrastructure nodes or on local multi-core machines. Statistical data processing can be applied to perform Niche Modelling or Ecological Modelling experiments. Other applications can use general purpose techniques like Bayesian models. Time series of observations can be managed as well, in order to classify trends, catch anomaly patterns and perform simulations. The idea under the distributed computation for data mining techniques is to overcome common limitations that can happen when using statistical algorithms: the training and projection procedure timing, the linear or non-linear time increase when the number of data to process increases, the multiple runs needed for reducing overfitting or local minima problems, the multiple models topologies to be evaluated for assessing the optimal model's configuration.

All the above issues strongly limit the amount of time a scientist can dedicate to the evaluation of the results and to the combination and comparison of the outcomes of different experiments. On the other side the Statistical Data Mining VRE adds advantages in using a distributed e-Infrastructure endowed with many data sources. Some of these are: efficiency and time saving in computations; availability of a set of data sources containing environmental or species features; reliability of the quality of the features; certification of compliance between e-Infrastructure data sources and algorithms inputs\outputs; sharing of results and users' files.

- **BiodiversityLab:** The BiodiversityLab is a VRE designed to provide a collection of applications that allow scholars to perform complete experiments about single individuals or groups of marine species. The VRE allows to: inspect species maps; produce a species distribution map by means of either an expert system (AquaMaps) or a machine learning model (e.g. Neural Networks); analyse species observation trends; inspect species occurrence data; inspect species descriptions and characteristics; perform analysis of climatic changes and of their effects on species distribution; produce GIS maps for geo-spatial datasets; discover Taxa names; cluster occurrence data; estimate similarities among habitats.
- **Stock Assessment:** the Stock Assessment VRE is designed to facilitate collaborative stock assessment where a role based workflow produces stock assessments in several clear steps: from data collation, and harmonization, through model driven analysis and dissemination. Planned solutions are oriented to serve both resource managers in need of harmonized stock-related information and scientists seeking a powerful infrastructure for new analytical approaches.

- Performance Evaluation In Aquaculture: the Performance Evaluation In Aquaculture VRE is designed to support capturing and confronting the dual challenge of understanding (a) the performance of an aqua-farming operation, allowing on one hand investors and entrepreneurs to conform to environmental rules and optimize the use of resources (monetary and non-monetary ones) and (b) the pressure on investment and on environment produced by such operations, so that scientists and policy makers can craft guidelines or even regulations, taking into account the economic interest of those operations.
- FAO Tuna Atlas: the FAO Tuna Atlas VRE offers tools to centralize, harmonize and standardize multiple data source on tuna and bill fishes catches to feed 2 FAO tools: the Atlas of Tuna and Billfish Catches; the global tuna catches by stock
- Protected Area Impact Maps: the Protected Area Impact Maps VRE supports spatial planning projects by providing existing rich analytical data platforms with additional EO based information products related to human impacts on natural systems, with an initial focus on Natura2000 sites and Marine Protected Areas.

Moreover, since the Fishery and Marine Sciences community is powered by the D4Science infrastructure, the same approach was promoted towards the other communities powered by D4Science. Positive feedback has been so far collected by:

- SoBigData⁸¹: a set of VREs to support the SoBigData community. SoBigData proposes to create the Social Mining & Big Data Ecosystem: a research infrastructure (RI) providing an integrated ecosystem for ethic-sensitive scientific discoveries and advanced applications of social data mining on the various dimensions of social life, as recorded by 'big data'. Building on several established national infrastructures, SoBigData opens up new research avenues in multiple research fields, including mathematics, ICT, and human, social and economic sciences, by enabling easy comparison, reuse and integration of state-of-the-art big social data, methods, and services, into new research.
- ENVRIPlus⁸²: the ENVRIPlus VRE aims to support Environmental and Earth System Research Infrastructures, projects and networks together with technical specialist partners to create a more coherent, interdisciplinary and interoperable cluster of resources exploitable across Europe.

To support those VREs and empower the communities that are exploiting them, the SLA process between D4Science and EGI is started. Details about this SLA are still subject to verification and approval. However it is worth to report the typology of resources that can be exploited in production by D4Science for the cited VREs:

- VMs with 8 virtual CPU cores with 16GB of RAM and 100GB of storage (to host a DataMiner Worker)
- VMs with 16 virtual CPU cores with 32GB of RAM and 100GB of storage (to host a DataMiner Worker)

⁸¹ <http://www.sobigdata.eu/>

⁸² <http://www.envriplus.eu/>

- VMs with 2 virtual CPU cores with 8GB of RAM and 50GB of storage (to host a Generic Worker)

Services for the long tail of science

The long tail of science services platform (LTOS) is composed by a user registration service (mostly developed in the first project year), a set of science gateways and a pool of resources supporting the catch-all VO created for the LTOS. The subtask worked in the progressive integration of these components.

New science gateways have been integrated in the platform during the reporting period: WS-PGRADE/gUSE⁸³, QCG Portal⁸⁴, EC3 portal⁸⁵.

A REST API has been implemented in the user registration service to expose information to the gateways.

The notification system, that reports both to users and administrators about activities and actions to be taken, have been improved to provide more clear information and to tune the amount of notifications. Also the front-end have been improved to be more user-friendly. The mechanism to handle customers' resource requests for URP Broker/administrator has been enhanced, and also the process for resources allocation has been simplified.

From a non-technical point of view, EGI Foundation and Terradue have agreed to sign memorandum of understanding and an SLA to support this collaboration, also beyond the project boundaries. The document is under finalization at the moment of writing.

At the moment of writing the sites successfully tested are providing enough resources to run the first production use cases that is planned for September 2016. There are no blocking issues.

5.3 Issues and Treatment

Issue: Not enough resources for gCube SLA

Mitigation: A new call for interest has been circulated through the EGI partners to identify the cloud providers who could participate to the SLA for the iMarine/gCube use case and support the beginning of production in October 2016. Operations coordination will focus on making sure that new providers who agree to provide resources will be ready for production on time.

Issue: gLite WMS being decommissioned

Mitigation: DIRAC WMS will be tested as a replacement. In case it is selected as the EGI endorsed solution for workload management in HTC, Operations coordination will work on an agreement with the DIRAC team for long-term support of the software and the documentation for new communities approaching the tool.

⁸³ <http://guse.hu/about/architecture/ws-pgrade>

⁸⁴ <http://www.qoscosgrid.org/trac/qcg>

⁸⁵ <http://servproject.i3m.upv.es/ec3/>

5.4 Plans for next period

5.4.1 Operations Coordination

Operations coordination will continue the support for the daily operations of the infrastructure, through the OMB meetings and Operations meetings, the coordination of the core services and through the technology deployment coordination.

Coordinate the roll-out in production of the storage accounting, which is being pre-tested at the moment of writing.

Release the first CMD major release. The first releases aim at making available in the EGI software provisioning framework, through the same quality assurance process, all the components that are deployed by cloud providers to federate in the EGI infrastructure.

Collaborate with one or more VO to test DIRAC as a replacement of gLite-WMS. In particular tests will focus on the migration to the DIRAC API, from the WMS client, and on the use of the private pilots, where computational tasks run on behalf of the actual user as opposite to the robot credentials commonly used with pilot systems like DIRAC.

Prepare a detailed cost analysis of the current set of core activities, including cost elements not usually considered like hardware and housing costs. Begin to assess the composition of the future core activities, for the next bid.

Continue the support of rolling in production of new services, for example the software provided by the INDIGO project and the JRA2 activity, and new VOs use cases.

5.4.2 Development of Security Operations

Task SA1.2 will continue on mitigation of risks identified in the PY1 security risk assessment. New EGI services may also require risk assessment and these will be performed.

Review and revise procedures based on experience. Risk levels for vulnerabilities concerning Virtual Machines will be reviewed, and as required appropriate contact lists for those who need to be involved in security operations will be established, e.g. Virtual Appliance owners and VM Endorsers.

The task team members will complete the work on risk mitigation for the EGI Federated Cloud, and new operational security procedures for Federated Cloud will be developed, approved and adopted. A new WISE sub-group on Cloud Security has been proposed by staff in SA1.2 to investigate current best practices on the tools and methods used for monitoring and control of Virtual Machine instances, particularly those where the end-user has privileged access. When approved work on this will start soon afterwards.

SA1.2 will investigate the possibility of deploying a stable and monitored central security emergency suspension system for Federated Clouds.

We will continue to work with EGI-Engage JRA1 on the AAI policy and trust issues.

New or revised security policy documents will be produced, consulted on and passed all the way through to formal adoption as required. This will include a revised Top-level security policy, revision of the three existing Virtual Organisation policies, and finalising the new policy framework on the protection of personal data.

A security service challenge of the EGI Federated Cloud Service will be carried out. Analysis of this will be carried out and any required improvements will be made.

All partners will work on the preparation of the deliverable D5.3, “Evolution of security policies, procedures and best practices”, which is due just after the end of PY2.

5.4.3 Integration, Deployment of Grid and Cloud Platforms

Integration of iMarine and gCube platforms with fedcloud

The plan for the next period is organized according to priority actions as follow:

- To support the 6 VREs of the Fisheries and Marine community reported above (1077 active users);
- To continue the promotion of Cloud resources through EGI towards the other VREs of the same community. So far the Fisheries and Marine Sciences community operates 28 VREs, out of those 15 have increasing computational needs. Therefore there is an outlook to at least double the number of VREs that could exploit the FedCloud resources. Moreover, the community plans to start operating 5 more VREs before the end of 2016 and all of them should have considerable computational needs.
- To support the SoBigData community and its VREs computational needs (150 active users). This community is growing fast and one of their first VRE, named TagMe, is heavily used to support a huge amount of short-lived processes. TagMe processes reached 1.3 million in a single day in July on its current instances hosted on D4Science and in average it performs 3.5 million processes per month in its first 3 months of exploitations.
- To finalize the training and support to the ENVRIPplus community to start operating VREs for them in the production environment.

The plan is clearly heavily dependent on the SLA that will be signed between D4Science and EGI.

Integration of the ESA e-challenge platform in fedcloud

In September Federated cloud will support the first production run of the geohazard e-challenge. Tests will continue to validate all the sites participating to the SLA, to add the resources needed for the other use cases planned later in the year, and next year.

Testing scenarios will be extended including the network configuration that will be available through the new OCCi implementation that supports openstack and opennebula network configuration services.

From a non-technical point of view EGI Operations and user support team will continue in negotiating new resources being added to the SLA to support the increase in resource demand that the e-challenge use case plans for 2017.

6 Knowledge Commons

6.1 Summary

During the reporting period WP6 Competence Centres continued with the development and integration of community-specific services into the EGI federation, while the other part of the WP (task 1 and 2) offered consultancy and training services for new communities. Highlights of the progress and achievements:

- 4 Webinar trainings on specialised topics, and 1 face-to-face training & consultancy event were delivered to new communities (primarily representatives of Research Infrastructures). Two technical tutorials were delivered by the MoBrain Competence Centre, and two by NGIs.
- Two new user guides/tutorials have been prepared as a response to emerging needs and user requests: Docker Swarm on EGI FedCloud; Hadoop on EGI FedCloud
- Technical collaboration advanced with the ICOS, EuroBioImaging, Euro-Argo, ENES, EMSO research infrastructures, as well as with 8-10 mid-size project and institutional communities. In this consultancy work the topics were dominated by various ways of exploiting the EGI Federated Cloud.
- 3 Competence Centres produced four deliverables:
 - The MoBrain CC integrated the DISVIS, POWERFIT and AMBER applications with GPGPUs in EGI, and started offering these ‘as services’ through their online portals (D6.4 and D6.7).
 - The Disaster Mitigation CC released two web application portals (WRF and iCOMCOT) that are connected to HTC resources in Asia-Pacific, and are offered ‘as services’ for simulation wave propagation and weather simulation (D6.9).
 - The ELIXIR CC setup a federated cloud infrastructure, using cloud resources from CC members, and federation services (AAI, GOCDB, AppDB) from EGI (D6.10).
- Key achievements in other CCs:
 - The BBMRI CC reached an integrated BiobankCloud-EGI FedCloud release, and initiated deployment at 3 affiliated biobanks.
 - The DARIAH CC continued finalising its service portfolio, and restructured its Wiki page <https://wiki.egi.eu/wiki/CC-DARIAH> into a ‘promotional page’ that will be used in the next period to promote EGI-based services for digital humanities researchers and groups.
 - The LifeWatch CC deployed applications on its infrastructure, setup monitoring systems to collect usage statistics. The details of application usage will be documented in the next period, in the D6.18 deliverable.
 - The EISCAT_3D CC continued the customisation of the DIRAC system, and had various rounds of reviews/assessment with EISCAT members. The feedback helps both the fine tuning of the EISCAT data model, and the DIRAC implementation

towards an EISCAT-specific data catalogue and application/workload management system.

- The EPOS CC continued working on the AAI, Computing and Data federation use cases. (releases will be in Month 24)

6.2 Main Achievements

6.2.1 Training

During the reporting period the activities in this task were continuing operation of the cloud-based training infrastructure and development of user tutorial about how to run Docker container-based applications in the EGI Federated Cloud⁸⁶.

We've started preparation for a DI4R tutorial titled the 'EGI Federated Cloud for developers'. The tutorial will target application and platform developers who want to create software on top of the EGI Federated Cloud system. The tutorial will be the first of this type, and is planned to result in reusable training package for similar online and f2f events of the future.

The first release of the OneData software is in preparation, which is the core element of the new 'EGI Open Data Platform' service. The testing will continue in the next period and aims to result in online guide/tutorial for users about how to use this system in various data sharing and federation use cases.

We've been preparing the DI4R tutorial titled 'Federated AAI meets reality, Security Incident Handling Role Play in DI4R'. The tutorial will go through different defensive and offensive techniques to protect grid and cloud resources, and to respond to various incident types.

The testing and evaluation of the first release of the INDIGO-Datacloud software (released in July) has been started. Testers are now allocated to assess the different elements of this software in the EGI Federated Cloud, focusing on components that can underpin scientific use cases that are in the engagement pipeline (See Task SA2.2 below for further details). Based on the outcome of the assessment user guides and tutorials will be prepared in collaboration with INDIGO-Datacloud project, specifically for EGI users.

In collaboration with the ENVRIplus project, the task delivered an 'EGI tutorial and use case consultancy workshop' the 2nd ENVRIplus Week, in May 10th. The event was attended by scientific communities from the Environmental sciences RIs of ENVRIplus.

We've supported the setup and delivery of four EGI Webinars:

- EGI Open Data Platform by Lukasz Dutka (CYFRONET), Matthew Viljoen (EGI Foundation)
- WS-PGRADE by Peter Kacsuk and Zoltan Farkas (MTA SZTAKI)
- DIRAC webinar by Andrei Tsaregorodtsev (CNRS)
- Hadoop on the EGI Federated Cloud by Tamas Kiss (Uni. of Westminster) and Carlos Blanco (Uni. of Cantabria)

⁸⁶ https://wiki.egi.eu/wiki/Federated_Cloud_user_support#Docker_containers

The following training events have been also delivered in the broader EGI-Engage and EGI context (Without SA2.1, but all registered in the Training Marketplace at <http://training.egi.eu>):

- Two courses by the MoBrain Competence Centre:
 - Cryo EM course - Solving the Structure of Macromolecular Complexes: A Hands on Workshop in Taiwan on 2016-03-14
 - INSTRUCT practical course: Advanced methods for the integration of diverse structural data with NMR data: A tutorial in the Netherlands between 2016-04-11 - 2016-04-15.
- SURF Research Boot Camp by SURFsara in the Netherlands on 2016-04-21
- The 6th DIRAC User Workshop in France between 2016-05-23 - 2016-05-25

6.2.2 Technical User Support

In the second project year - in response to the recommendations of the 1st EC review - the EGI outreach and user support activity will focus on structured communities with whom technical collaboration already agreed/started during PY1. Progress with these support activities was the following during the reporting period:

1. Euro-Argo RI: Prototype of the data cloud with data subscription and delivery capability for users. The prototype is under implementation by members of University of Cantabria (Spain) and members of the Euro-Argo RI. The setup is envisaged to use the following technologies: Hadoop, SWIFT, WS-PGRADE, and HIVE.
2. ICOS RI: Two use cases are under implementation, using similar technical setup:
 - a. Integration of a community portal in the EGI cloud, with scalable back-end system based on Docker, and with data stage-in/out capabilities towards EUDAT.
 - b. Integration of station data into data products.
3. EuroBioImaging: UPV in the INDIGO-Datacloud project has designed and now developing a 'VRE as a service' architecture that can support this use case. EGI works on the integration of the INDIGO components with EGI Federated Cloud (in JRA2.2 and SA2.1). SLA will be signed with the community about the EGI and INDIGO services that would underpin the setup.
4. ENES: Similar to EuroBioImaging, but with using the Ophidia data analytics framework instead of the UPV system. (Both are from INDIGO-Datacloud project)
5. EMSO RI: There was no progress with this collaboration during this period, because EMSO was focussing on its ERIC application. We aim to sign MoU with EMSO-Dev project and deploy the EMSODEV Data Management Platform prototype in the EGI Federated Cloud. A 'Competence Centre' with countries/NGIs common to EMSO and EGI was already established. (IT,FR,UK,GR,PT,IE. DE and NL are initially observers; the inclusion of Turkey is under discussion.) SLA agreed for the EGI services that will underpin the EMSO platform.

Besides the above reported communities, progress was made during the reporting period with supporting the uptake of cloud technologies within the following use cases / communities:

- Molecular modelling - University of Catania: the community is interested on deploy some bioinformatics applications on the EGI FedCloud. First tests were performed on the FedCloud to evaluate the performance on running such applications into a distributed environment.
- BioISI (Biosystems and Integrative Sciences Institute): initial assessment with the community was performed and SLA negotiation is ongoing. Resources on LIP (PT) already being provided.
- Big Data Europe: BDE uses docker swarm as underlying technology for their applications. An evaluation of Docker Swarm deployment scenarios for FedCloud was done and documented in the EGI wiki. A test Docker Swarm cluster is available for Big Data Europe to test their pilot applications.
- ENVRIplus: environmental researchers want to process data using own algorithms, but most of current ENVRIplus research infrastructures only provide data products using standardized processing algorithms. The use cases are testing D4Science VRE to support such a need.
- HBP imaging service: The Imaging Service is a software component developed by the Customer to run on multiple Docker containers, load balanced for resiliency and scalability purposes. For this use case a SLA negotiation is already in place but the status is on hold due to the change of the HBP leadership.
- HBP iPython/Jupyter Notebook: Initial tests have been conducted to use Docker containers in the EGI e-Infrastructure and evaluate performance. They are currently evaluating other solutions such as: Docker Swarm (or Kubernetes).
- EXTRAS project⁸⁷: the community deployed in production a Science Gateway connected with the EGI FedCloud. This SG spawns a VM each time a user requires a new computation. EXTRAS agreed an SLA with EGI valid from 01/05/2016 to 01/01/2017 with 50 cores, 200 GB of RAM and around 1 TB of storage offered by two resource providers.
- DIRAC use cases:
 - The main user of the DIRAC4EGI service is WeNMR, with running 6 million jobs in the last 12 months.
 - Within the EISCAT_3D Competence Centre DIRAC is extended into a community-specific environment that offers file cataloguing and computing services.
 - The complexsystems.eu VO is a new DIRAC user, with mostly users from France, sharing the common interest of simulating the behaviour of complex, networked systems.
 - Other main DIRAC installations are for/by ILC, CTA, and Belle-2 communities. (particle physics, astrophysics)
- CMS: CMS is interested to extend its set of cloud resources and approached EGI to evaluate the FedCloud. A first meeting was organised to present the FedCloud features to the community and to agree on starting collaboration. CMS is currently using EC2 interface to access cloud resources, including OpenStack sites.
- CHIPSTER: a custom portlet within the LToS has been developed to let users to generate a chispter account and access the server. The CESGA cloud provider has been configured to

⁸⁷ <http://www.extras-fp7.eu/index.php>

support the vo.access.egi.eu VO. Configuration of the Chipster server at CESGA is in progress.

6.2.3 ELIXIR

Much of the work has been done during the reporting period (1.3 -31.7. 2016) and is described in the document "Infrastructure tests and best usage practices for life science service providers." that was published in June 2016⁸⁸.

This document summarizes the status of EGI Federated Cloud integration in the ELIXIR CC member institutes. It also includes experiences of the sites that have integrated their cloud environments to the EGI Federated Cloud as well as best practices instructions for integrating Open Stack, OpenNebula and Synnefo based cloud services to the EGI Federated Cloud. The goal of this Integration work is to make EGI Federated Cloud compatible service for the ELIXIR Compute Platform⁸⁹.

During the period discussed here significant progress has been achieved with AAI integration. Key achievement so far is a credential translation service that converts ELIXIR AAI Identities into X.509 certificates. The recently released EGI AAI proxy service has been integrated to the ELIXIR AAI. This integration enables ELIXIR user community to use GOCDB, AppDB and OpenStack resources with their ELIXIR identities.

In the same time, the work with the scientific use cases has started. Test systems for the computing tools used by the use cases have been under preparation in local cloud environments and in some cases first tests to use federated cloud environment has been carried out. Progress so far with the use cases is as follow:

- cBioPortal
- Marine metagenomics: The first scripts for setting up a virtual cluster running the META-pipe analysis pipeline are now ready. At the moment these scripts have been constructed for the OpenStack cloud environment of CSC. The next step will be modifying these scripts so that they can be applied to EGI Federated Cloud.
- Insyght Comparative Genomics: the Bacterial Genomics preconfigured image, compatible with OpenStack clouds, is now available⁹⁰ and includes CloudInit for contextualization. It comes with a script that launches the application in a clustered mode. Next steps are to integrate the image in the EGI cloud marketplace (AppDB) and to automatically deploy a configured cluster without the need of the user intervention.
- PhenoMeNal: a version of Gophercloud (Go library to access OpenStack used in Terraform) modified to support EGI AAI is now available. This library will allow extending Terraform to be used directly on OpenStack resources federated in EGI as required by the PhenoMeNal project.
- JetStream interoperability

⁸⁸ <https://documents.egi.eu/document/2802>

⁸⁹ <http://bit.ly/elixirtech>

⁹⁰ <https://github.com/enolfc/gophercloud>

6.2.4 BBMRI

The work focused on two major aspects related to integration of BiobankCloud platform with EGI FedCloud infrastructure as a part of the 'D6.11 Security toolset release for BBMRI-ERIC' due in November 2016:

1. implementation of OCCI into BiobankCloud and complementary adaptations needed on the side of FedCloud and
2. integration of BiobankCloud with Shibboleth federated authentication infrastructure.

BiobankCloud integration with the OCCI is a fundamental prerequisite for deployment on any cloud resource of the FedCloud infrastructures, both for private clouds of the biobanks (the default for sensitive data processing) and in public clouds (for anonymous data processing with low risk of reidentification). The integration has been finished by April 2016.

Integration with the Shibboleth turned out to be more complicated than anticipated because of complex internal authentication infrastructure in BiobankCloud related to its advanced authentication and authorization infrastructure in order to solve the multitenancy problem with respect to users and their projects (the user may be entitled to access two different data sets for two different project, but not allowed to merge or correlated the datasets mutually). This requires prepending the Shibboleth service provider components in front of the internal authentication systems and integrating the identities coming from the federated IdP into the local person-project identities. This work is still ongoing as of end of the reporting period.

At the end of the reporting period the pilot biobanks have been contacted via their national nodes (BBMRI.cz, BBMRI.nl, BBMRI.se) in order to start preparation of the pilot deployments.

6.2.5 MoBrain

The MoBrain CC aims to lower barriers for scientists to access modern e-Science solutions to simulate life science processes from micro to macro-scales. The CC builds on grid- and cloud-based infrastructures and on the existing expertise available within WeNMR/INSTRUCT and NeuGrid4You. During the reporting period MoBrain:

- Deployed Scipion at SURFsara for a real CryoEM testing and Instruct training. Visited the NeCEN center in Leiden (cryo-EM experimental facility) to help with Scipion installation both locally and at SURFsara cloud.
- Worked on a Scipion image to be published in the EGI AppDB marketplace.
- Started installing Scipion on the STFC Scarf cluster.
- Built Relion on the STFC Scarf cluster, and will make it available to users
- Benchmarked the AMBER and GROMACS software packages for molecular dynamics (MD) simulations to compare their performance on single and multi-core CPUs, which are widely available within the EGI infrastructure, with respect to their performance on GPGPUs, which are a scarcer resource (described in D6.7).
- Implemented GPGPU-enabled web interfaces for the AMBER package for molecular dynamics simulations and for the DisVis tool for quantifying and visualizing the accessible

interaction space of distance-restrained binary complexes (described in D6.12)) and PowerFit software for rigid-body fitting in cryo-electron microscopy map.

- The CIRMMP GPGPU resources were updated with the latest NVIDIA driver 352.93 supporting now CUDA 7.5 version. DisVis and PowerFit docker images have been updated accordingly and are now the ones maintained by the INDIGO-DataCloud project and distributed via dockerhub at indigodatacloudapps repository⁹¹.
- Together with DIRAC developers successfully tested a DIRAC VM cloud execution of a Gromacs job on the EGI FedCloud
- Together with the INDIGO-Datacloud project, demonstrated the use of the FutureGateway for job submission from the web interfaces for the AMBER package
- Maintained the mobrain web portal hosted on the EGI FedCloud
- Completed the migration of the N4U software to CVMFS
- Organized/contributed to several training and dissemination events:
 - Cryo-EM tutorial in Taipei, Taiwan ROC, at ISGC 2016⁹²
 - EGI conference in April in Amsterdam
 - INSTRUCT practical course on “Advanced methods for the integration of diverse structural data with NMR data – 2nd Edition” in Utrecht, the Netherlands, in April⁹³
 - PhD School on “Molecular Modelling for Life Sciences”, Pula, Italy, in June⁹⁴
 - 13th Annual Meeting of the Bioinformatics Italian Society in Salerno, Italy in June⁹⁵
 - EMBO Global Exchange lecture course on “Structural and biophysical methods for biological macromolecules in solution” in Suwon, South Korea in June⁹⁶
 - Chianti Workshop "Magnetic resonance for cellular structural biology", Principina di Terra, Italy, in June⁹⁷
 - EMBO Practical course on “Integrative Modelling of Biomolecular Interactions” in Barcelona, Spain, in July⁹⁸
 - Pasteur workshop on “Integrative Structural Biology in Paris, France in July⁹⁹
 - XXVII International Council on Magnetic Resonance in Biological Systems (ICMRBS), Kyoto, Japan, in August¹⁰⁰

6.2.6 DARIAH

The main activities of the DARIAH Competence Centre during the period between March and August 2016 were the finalization of the 2 mini-projects including releasing of the CC's technical

⁹¹ <https://github.com/indigo-dc/ansible-role-disvis-powerfit>

⁹² <http://event.twgrid.org/isgc2016/cryo.html>

⁹³ <http://www.bonvinlab.org/education/INSTRUCT-practical-course>

⁹⁴ <http://molmodel2016.dsf.unica.it/wp>

⁹⁵ <http://bits2016.bioinformatics.it>

⁹⁶ <http://events.embo.org/16-macromolecule>

⁹⁷ <http://www.cerm.unifi.it/chianti/>

⁹⁸ <http://events.embo.org/16-biomol-interact/>

⁹⁹ <http://www.pasteur.fr/en/education/integrative-structural-biology>

¹⁰⁰ <http://www.icmrbs2016.org/>

solutions for DARIAH, work on the sustainability, and user involvement plan and dissemination activities. In July 2016, the final version of the Bavarian Dialects repository, provided by the Austrian Academy of Science was released. The repository is developed by INFN and is based on the gLibrary repository framework, which allows contents to be stored across numerous cloud and grid sites. Currently, the repository is freely accessible via Catania Science Gateway for the Long Tail of Science¹⁰¹. The Liferay-based portlet for accessing the repository can download from the github¹⁰². The Parallel Semantic Search Engine is enriched with new databases such as DBpedia. The ongoing actions are installation of the Bavarian Dialects repository portlet and embedding Simple and Parallel Semantic Search Engines (SSE/PSSE) into the DARIAH Science Portal.

The final version of the SIR application, developed by GWDG, for real-time search of multi-sourced structured and unstructured data, will be released during August 2016. The SIR application is based the CDSTAR technology, previously developed by GWDG. Furthermore, the OCR use case is under development and is expected to be available late August 2016.

The DARIAH Science Portal, the central service developed and provided to digital humanists, is setting up and is under registration as a service provider in EduGain. The portal is up and running, but, currently, no applications or services being developed by DARIAH CC are integrated or embedded into portal. The work plan is to integrate the services: PSSE, Bavarian dialects portlet and SIR/OCR portlet, by the end of September 2016. The portal is accessible on the following link: <https://dariah-gateway.lpds.sztaki.hu/>

In June 2016 the first version of the “User involvement and sustainability plan” is published. However, with the forthcoming final releases of the DARIAH CC applications and services, the plan will be updated and modified accordingly.

During the report period, 4 teleconference meetings were organized, following the scheme agreed at the beginning of the project, one meeting per month. In addition, the DARIAH CC outcomes and services were presented at 3 events: EGI Conference in Amsterdam in April 2016 (presentation), e-Infrastructure Day in Zagreb in June (presentation), and on the Digital Humanities (DH2016) conference in Krakow in July (poster). Furthermore, a short paper was submitted and is currently under revision for the EURALEX conference (6-10 Sep 2016, Tbilisi, Georgia).

6.2.7 LifeWatch

The EGI LifeWatch CC has progressed in the last months on the implementation of solutions on top of FedCloud resources for the research teams involved, as reported in the presentations made at the dedicated meeting in Amsterdam in the framework of the EGI Conference. In particular significant advances on the integration of data from observatories, comparison with models, and also the implementation of workflows using python tools were discussed, along more technical points like the solution for AAI or the use of noSQL databases.

LifeWatch also contributed in the EGI session on GPU usage, given the interest for automated image recognition/processing. Several technical working groups have been created with the aim of

¹⁰¹ <http://csgf.egi.eu/sade>

¹⁰² <https://github.com/csgf/sade-portlet>

long-term fora for collaboration, discussion and exchange of experiences, using the OpenProject framework. The EGI LW CC was also proposed as an example to a new initiative, EMBRC, on marine biology, at the e-infrastructure working group meeting in Paris on 5th July.

6.2.8 EISCAT_3D

The establishment of the DIRAC portal has been continued in the period. A first version was released at the end of March, which were evaluated within the CC and after minor modifications presented at the EISCAT_3D User Meeting in Uppsala, Sweden 18-19 May. The presentation included a step-by-step live show of the portal, from selection to download of data and also a first attempt of data processing. The interest from the users was high, and the comments encouraging continuing the development.

After this first version, the system has been developed further to be more robust, and to allow authorization and access rights according to the EISCAT statutes. This has proven to be not so straightforward in the present version of DIRAC. A temporary workaround have been written, that will be integrated in the next DIRAC version. The setup for data processing have also been improved, and will give the users a new and wanted way of displaying the data to study interesting events. Our main analysis tool is written in Matlab, and we have had a couple of discussions with Mathworks how to solve the license problem for the cloud.

The data model for the EISCAT_3D has also been improved after inputs from users and EISCAT engineers. Now it is ready to be applied to the present data and tested in the portal environment.

In the weekly teleconferences the progress is followed up and discussed to solve upcoming issues and planning of the continuing work. There have also been face-to-face meetings of different constellations whenever the opportunity arise. Presentations and discussions took place at RDA 1-3 March in Tokyo, DIRAC future meeting 9 March in Kiruna, the EGI conference 4-8 April, ENVRI week in Zandvoort 9-13 May, EISCAT_3D User meeting 18-19 May and at the DIRAC conference 23-25 May in Montpellier.

6.2.9 EPOS

The EPOS Competence Centre is collaboration between EGI and partners of the European Plate Observing System (EPOS). It aims at establishing a knowledge transfer between the EGI and EPOS communities, concerning the possibilities offered by the EGI common services and their adoption in real scenarios. Activities include collecting, analysing and comparing Earth Science community needs with EGI services and evaluating their integration and interoperability with the EPOS infrastructure.

During the reporting period the EPOS competence Centre was working on the implementation of the three pilot use cases that have been identified in PY1 and documented in M6.4¹⁰³:

- AAI use case: to date the most important task was to identify and validate AAI services which might possibly be useful to be integrated within the EPOS infrastructure. To test possible

¹⁰³ <https://documents.egi.eu/document/2676>

solutions a simple use case scenario has been designed. It covers the basic interaction workflow between the EPOS central hub (ICS-C) and two thematic core services (TCSs). As a result of its implementation we expect basic data transfer to be achieved with help of the full AAI implementation. The AAI prototype will be integrated in the computational earth sciences use case where, in addition to the basic data transfer, further data analysis will be performed. The goal of this work is to test the interactions between ICS-C and ICS-D (Distributed Core Services).

The EPOS CC work on AAI has been presented to the EPOS IT groups in Nottingham. Based on further consultation with the EPOS working groups 6 and 7, a working prototype has been developed. The prototype has been presented to EPOS IT group in Prague and has been chosen to be implemented within EPOS central hub. Furthermore, a study has been started to understand how to make interoperable the EPOS AAI designed in this task with the new EGI AAI infrastructure. A meeting was organised where an integration roadmap was defined¹⁰⁴.

- Computational Earth Science and Seismology: Besides addressing interdisciplinary data discovery and access in solid earth science, EPOS aims at establishing a model to integrate and sustain community-driven computational earth science services (CES). This will allow the development of new computational tools, as well as the continuative support of the existing ones. This requires the development of new procurement policies for processing resources (ICS-D) and the investigation of homogeneous technological solutions facilitating their exploitation. More specifically, the computational use case aims at evaluating ease of access, integration and management of the FedCloud resources within the VERCE platform based on the WS-PGRADE/gUSE framework.

For these purposes, the latest release of gUSE had to be deployed on the VERCE infrastructure and evaluated for proper operation, especially in those features supporting the access to the EGI FedCloud infrastructure. This activity highlighted a basic challenge that needs to be addressed: Virtual Machines started in some cloud resources (e.g. the CNRS FedCloud resource) are not reachable from the public internet by default, requiring an additional request to associate a public IP with the VM (a very common way of offering this service). After consultation with the gUSE development team, it was decided to pursue the deployment of an additional middleware provided for gUSE at the resource provider's site. Due to additional challenges met during the deployment and configuration (most of which can be found in the issue tracking of gUSE), this is a work in progress requiring some bug fixes in gUSE itself. In the meantime, a contextualization and deployment script has been created, to prepare a virtual machine based on the EGI CentOS base image provided on the AppDB with the prerequisites for the VERCE processing tasks.

Another activity for the use case is exploiting the services provided by EUDAT for data-staging and persistent storage of the research results. The work just recently started on this topic. This will demonstrate the integration of the EGI and EUDAT computational and data-management e-infrastructures, into a specific community-based platform. The work is

¹⁰⁴ <https://indico.egi.eu/indico/event/3029/>

coordinated by EGI across several communities who need similar integrated setup between EGI and EUDAT.

- Satellite Data use case: it is related to the services that will be offered by the EPOS satellite data Thematic Core Service (TCS) to the wide range of EPOS users. This TCS deals with the processing of the Earth Observation datasets collected by various satellites, including the Sentinels of the Copernicus programme to address several societal challenges. Aim of this use case is putting together expertise of ICT experts and platforms operators with advanced knowledge on Earth Observation systems with EGI infrastructure expertise to create an environment where new added value services could be easily developed and integrated in the Satellite Data TCS offer. This work is linked to the integration of the Geohazard ESA Thematic Exploitation Platform (TEP) with the EGI FedCloud is currently running under the EGI-Engage task SA1.3. Implementation is almost completed and the first test campaign will start in September 2016. In addition, Terradue, leader of the consortium developing the Geohazard TEP for ESA, agreed an SLA with EGI to exploit its cloud resources and is going to sign an MoU with the EGI foundation that defines the plan for the this collaboration for the next years.

6.2.10 Disaster Mitigation

For the typhoon Haiyan case study (by PH and TW), DMCC finished the complete storm surge propagation process in Open Ocean by coupled WRF model and COMCOT storm surge model. Based on this case study, a storm surge simulation web portal will be developed in the future. Preliminary simulation and sensitivity test on Malaysia flood case had been carried out. Finer resolution and model tuning will be implemented with further observation data from Malaysia. Advanced visualization with portable headset was demonstrated by DE at ISGC 2016. DMCC hosted Disaster Mitigation Workshop at APAN42 meeting on August 4 in Hong Kong to promote the DMCC collaborations to all Asia countries.

6.3 Issues and Treatment

No issues have been identified during reporting period.

6.4 Plans for next period

The next period will be a critical one for WP6. 6 Competence Centres will finish technical development and integration activities (services and applications on top of EGI services), resulting 8 deliverables (EPOS, LifeWatch, DARIAH, ELIXIR, BBMRI, Disaster Mitigation). The success of these services will be key influencers of long-term relationship between EGI and the respective Research Infrastructure communities. Right after the development each of the CCs will have to enter into a new phase of life with service promotion, training, usage monitoring and feedback evaluation from their connected research infrastructure communities. The SA2.1 activity will support this with baseline training materials (11 so far from the first 18 month of EGI-Engage), and with offering Webinar, cloud and event management infrastructure for CC events. Outreach to new communities is planned at key events, such as the upcoming DI4R conference and the co-

located Design your e-infrastructure workshop, and various events organised by RI communities (e.g. ENVRIplus week, DARIAH Annual Forum).

Technical developments with/for new communities continue, expecting services to be deployed in EGI for Euro-Argo, ICOS, EuroBioImaging, ENES and EMSO.

6.4.1 Training

In upcoming period the task is planning to implement the remaining items from the PY2 training plan (The plan was documented in M6.5¹⁰⁵):

- Delivery of two tutorials at DI4R:
 - EGI Federated Cloud for developers
 - Federated AAI meets reality, Security Incident Handling Role Play in DI4R
- Content development:
 - Create online tutorial on how to use the EGI AAI proxy service (CheckIn).
 - Finish testing of the OneData and EGI Open Data Platform, create user guides and if needed online user tutorial about them.
 - Create guides about the platform for the long-tail. (One guide for users; One for NGI support teams). Deliver online/f2f course based on these if needed.
 - Complete the guides about high level platforms and tools that exist on top of the EGI Federated Cloud. (Do this in collaboration with the Marketplace activity)
 - Assess the status of GPGPU resources in the EGI cloud and grid infrastructures; complete user guides and tutorials to support the uptake.
 - Finish testing of the first INDIGO-Datacloud release in the EGI Federated Cloud, create user guides and if needed online user tutorial about the use of this software in the EGI cloud.
- Support the training activities of Competence Centres. (All CCs will deliver training in PY2. See details in M6.5)

6.4.2 Technical user support

EGI will focus technical engagement activities in the remaining part of PY2 (and also in PY3) on the following structured communities, aiming at achieving the listed tangible outcomes: Euro-Argo; ICOS (Two use cases); EuroBioImaging and ENES (in collaboration with Indigo-Datacloud); EMSO.

Other community engagement activities that will be carried out, targeting specific groups:

- Deliver the next 'Design your e-infrastructure workshop' in collaboration with EUDAT, OpenAire and GEANT. (Co-located with DI4R). The workshop will provide opportunity to establish technical collaboration with 4-5 new, structured community or project.
- Long-tail of science: Help SA1 finalise and release the EGI platform by integrating generic and thematic services into it. Support NGI user support teams in uptake of the platform; Support the communication team promote the platform directly to the long-tail.

¹⁰⁵ <https://documents.egi.eu/document/2810>

- Initiate collaboration with VRE H2020 projects, using the 'Design your e-infrastructure workshop and DI4R as meeting opportunities.
 - Multiscale Genomics (MuG) H2020 VRE project
 - Recognition and Enrichment of Archival Documents (READ) H2020 VRE project
 - OpenDreamKit H2020 VRE project
 - EVER-EST H2020 VRE project
- Anticipated changes in the DIRAC service (besides continue supporting the communities who are already in the pipeline):
 - Identify VOs and start testing with them DIRAC as a WMS alternative on DIRAC4EGI.
 - Discuss with UK NGI, France NGI, Cyfronet to assess interest for a common deployment (instead of the separate installations) to lower operational cost.

6.4.3 ELIXIR

During the next period of the project, ELIXIR CC will continue implementing the virtual machines, developed and used by the 5 scientific use cases, to the EGI Federated Cloud. In the same time the work with the AAI integration and applying the results will be continued so that the ELIXIR user community can access the EGI resources through the ELIXIR VO.

There are technical and administrative issues requiring more work in the future that are blocking some of ELIXIR CC members (CSC, CNRS, SURFsara) to provide resources to ELIXIR VO through the EGI Federated Cloud. These issues will be followed up within the CC. In addition to the technical development work, EGI ELIXIR CC needs to define policies and protocols for using the resources of ELIXIR VO in collaboration with the ELIXIR-EXCELERATE project.

6.4.4 BBMRI

BBMRI competence centre will proceed toward the pilot deployment of the enhanced BiobankCloud platform into the pilot biobanks. This will be accompanied by the report analysing the workflows in the biobanks and analysis of their expected porting to the BiobankCloud. Finalization of Shibboleth integration with BiobankCloud is also expected.

6.4.5 MoBrain

The MoBrain activities are following the original plan. Note that most activities and deliverables were concentrated in the first 18 months of the project. MoBrain will remain very active in dissemination and training.

6.4.6 DARIAH

DARIAH CC will proceed with the integration of the PSSE, Bavarian Dialects, SIR and OCR applications into the DARIAH Science Portal and the consolidation of the services. The process of registering the DARIAH Science Portal as a service provider in EduGain will be finished by September and will allow the access to the EGI Cloud Federation resources to the members of the DARIAH IdP. The full integration of PSSE and Bavarian Dialects is planned to be finished by early September

2016. The final release of the DARIAH Science portal is planned for February 2016 with the EGI-Engage deliverable D6.17.

For the next project period, DARIAH CC will concentrate on the dissemination actions towards the digital humanities communities. The first actions are planned for September/October by organizing the workshop during the EURALEX conference (6-10 Sep 2016) on which the DARIAH Science Portal, Bavarian Dialects and PSSE will be demonstrated. Also, DARIAH CC will demonstrate service being developed on central DARIAH event, the DARIAH-EU Annual meeting (10-12 Oct Ghent). The participation on other DH-related conferences and events is planned as well as writing a paper for a journal. Furthermore, CC plan to extend the network of customers of our services and to establish more contacts with various DARIAH and Digital Humanities-related projects and research groups and offer them the provided services and applications. As a part of dissemination actions, we plan to publish short user-stories and best practices in DARIAH newsletters, broadcasting news on new services via DARIAH web pages as well as other digital humanities research projects. Also, once the applications and services are finalized, the demonstrators will be used to disseminate the benefits and usefulness of the cloud-based services to attract new research communities to use and to increase the number of DH-enabled applications benefiting from using the cloud infrastructure and accessible via the DARIAH science portal.

6.4.7 LifeWatch

LifeWatch activities will concentrate in the consolidation of the technical solutions being developed and implemented on FedCloud resources to support the application targets: data flow from observatories, data processing workflows and citizen science. We aim to put into production the AAI solution to be used in the future by the LifeWatch ESFRI, along the lines being proposed in the AARC project, supporting the access of users using ORCID. We also aim to complete the deployment of the R and python tools as a service for data processing and the access to GPU resources for citizen science applications (image/sound recognition). The CC will continue monitoring its deployed applications and report usage in the D6.18 deliverable by the end of PY2.

6.4.8 EISCAT_3D

For the coming period, we will continue to develop the portal to implement further the new EISCAT data model into the environment, and improve the visualisation process interface and to introduce the Matlab analysis. We will discuss the DIRAC EISCAT portal in a session at DI4R in Krakow 28-30 September, and expect more of the ad-hoc meetings at other conferences. After a holiday break in August, we will continue the weekly teleconferences.

6.4.9 EPOS

Plans for progress with the three pilot use cases:

- AAI: The demonstrated prototype setup will be developed into a production system within the EPOS central hub.

- Earthquake simulation use case: The IP allocation subsystem of the VERCE portal will be improved/extended; The integration of EUDAT storage system will continue.
- Satellite Data use case: The integration of the Geohazard ESA Thematic Exploitation Platform (TEP) with the EGI FedCloud will continue.

6.4.10 Disaster Mitigation

DMCC will focus on the Malaysian flood and Indonesian forest fire two case studies. Based on the experiences of advanced visualisation on local sourcing, the 2nd 3D visualization case on typhoon Morakot has been kicked off by TW and DE. DMCC will collaborate with DRIHM project starting from the weather simulation portal by integrating WRF portal of DRIHM with the DMCC regional e-Infrastructure. With the EGI-APAN (Asia Pacific Advanced Network) collaboration framework on disaster mitigation from January 2016, some Asia countries such as Nepal, Vietnam, and Thailand will bring more case studies and work with Disaster Mitigation Competence Centre in the future.

7 Consortium Management

7.1 Summary

Consortium management includes project finance and administration, technical coordination and WP management and quality management. All these activities are delivered through WP1.

7.2 Main Achievements

7.2.1 Administrative and Financial Management

During this period a collaboration board has been held in Amsterdam on April 8 and a second has been organized. Discussions regarding the project reporting activities for the period PM13-PM30 and the planned review agenda took place with the EC.

The Project office has achieved the following tasks:

- Preparation of PY1 effort consumptions table and analysis across WPs and partners
- Preparation of PY1 financial report:
 - Submission of Periodic Report including PY1 financial statement within deadline 30-04-2016;
 - Clarification to PY1 provided to EC: July 2016
 - Periodic Report and Costs acceptance: 13-07-2016
 - EC payment: 20/07/2016
- Management of project funding:
 - Transfer of Pre-financing advance 2: 14-03-2016
 - PY1 interim payment: 16-09-2017 approved by PMB on 8/09/2016
- Preparation of project Amendment N2:
 - Preparation of overview budget and efforts tables in liaison with the AMB, PMB and the Technical Director.
 - Drafting of the change log and the update of the DoA
 - Discussion of proposed changes with the PMB and CB
- Preparation of an intermediate project report for the period M13-18 (this document) including the overview of efforts and costs consumed.

Documentation is maintained to provide access to updated contractual information and procedures, financial and reporting activities of the consortium and projects minutes.

7.2.2 Technical Management

The main activities of project technical management in PM13-PM18 are:

- The preparation of PY1 review, which took place in April 2016 in Brussels, and the preparation of the first periodic report.

- The preparation of the PY2 yearly work plan: the work plan, organized by Work Package, defines the detailed list of internal milestones and the respective timing, and is used by the Technical Director and the WP Managers to properly time work during period PM13-PM24, to define dependencies across outputs of different work packages and to monitor the execution of the work plan on a monthly basis. The PY2 work plan was prepared taking into account recommendations from the project review.
- The analysis of the review report, the definition of necessary measures to address the reviewers' comments (see below) and the implementation of various changes feeding into Amendment 2.
- The review and update of the technical work plan concerning the end-user facing tools: the marketplace (service discovery and access), e-GRANT (service access management, including SLAs/OLAs management) and Long-Tail of Science (discovery of services for the long-tail of science and user authentication and authorization management). The new approved technical work plan, to be finalized in November, requires a single GUI entry point for service discovery and access and the definition of the business logic that will enable service access across all services of the external portfolio.
- Two project letters were prepared in March and February to inform all members of the consortium about the outcomes of the project after the first 12 project months, and to inform about the outcome of PY1 review and the proposed actions to address the review report.

PY2 Activity Plan in response to the review

The changes introduced to answer the request of increasing activities that maximize impact of the project, concern the project exploitation plan, the engagement towards Research Infrastructures and commercial entities (like IaaS and PaaS providers), the communications and quality plan, and the EGI strategy.

To address the reviewers' comments, **the engagement towards commercial entities has been focused on the big data value by ensuring success stories in three sectors by the end of the project: 1) fishery and marine resource conservation, 2) earth observation, and 3) agri-food.** As suggested by the reviewers, we focused on use cases where EGI can deliver a unique value. Communications activities to underpin technical and business development plans, were defined. These include the creation of promotion material for the updated EGI service catalogue, the delivery of tailored messages and materials to specific target groups, and the presence at key community events to promote the project outcomes.

The engagement with Research Infrastructures has been reinforced by establishing **new support activities for ENES, EuroBioImaging, EMSO, EuroArgo, and ICOS**, leveraging the work of collaborating projects like ENVRIplus, CORBEL, and INDIGO DataCloud. This action is expected to further increase the number of Research Infrastructures adopting distributed computing and data management in their service architecture or working on proof of concepts enabled by EGI services. During PM13-PM18 the number of integrated Research Infrastructures increased from 11 to 15.

Concerning the evolution of the strategic thinking, EGI has been updating the vision for the next years in relationship to the new EC policy developments (e.g. European Cloud Initiative and European Open Science Cloud). This activity has involved the EGI Council, has required an update of the strategy implementation plan and will be a necessary step for the update of the EGI strategy planned on a 2-year basis.

7.2.3 Quality management

Task NA1.3 in reporting period was responsible for updating and execution of quality plan described in Report of quality status and quality plan for Period 2 taking into account Project review recommendations, as well as executing risk management described in D1.2 Risk analysis and risk response for Period 1. In addition Review of project outputs procedure has been defined to gather, review and execute dissemination plan and exploitation plan of project outputs in structure and repetitive way.

During reporting period Task 1.3 has been highly involved in ISO 20k and ISO 9k certification preparation. A series of meeting have been held with process managers to define and prepare missing procedures.

7.3 Issues and Treatment

Issue: During the preparation of the periodic report and financial statement the project office has identified several partners which were deviating to the activity plan.

Mitigation: A detailed analysis has been made and reported to the AMB and PMB. Finally a document was circulated to the PMB with an overview of the DoA changes required to take the necessary corrective actions, which have led to the preparation of Amendment N2 recently accepted by the PMB.

7.3.1 Milestones and Deliverables

Table 1 – Preferred colour scheme

Id	Activity No	Deliverable / Milestone title	Type (***)	Lead partner	Original Delivery date(*)¹⁰⁶	Revised delivery date	Status (**)
D1.3	WP1	Report of quality status and quality plan for Period 2	R	EGI Foundation	PM15	21/06/2016	<i>PMB approved</i>
D2.10	WP2	Market Analysis Report of selected sectors	R	GRNT	PM18	17/09/2016	<i>PMB approved</i>
M2.3	WP2	The EGI website is reviewed	DEC	EGI Foundation	PM13	11/03/2016	<i>PMB approved</i>
M2.4	WP2	EGI Service management practices are audited	OTHER	EGI Foundation	PM15	31/05/2016	<i>PMB approved</i>
M3.3	WP3	Operational tools development roadmap revised	R	EGI Foundation	PM16	15/07/2016	<i>PMB approved</i>
M3.4	WP3	Pilot services and best practices to enable federated AAI solutions released	DEM	EGI Foundation	PM16	30/06/2016	<i>PMB approved</i>
M3.5	WP3	The first version of the EGI Marketplace is demonstrated	DEM	EGI Foundation	PM18	-----	<i>In preparation</i>
D3.5	WP3	First release of the new Accounting Portal deployed in production	OTHER	CSIC	PM14	25/05/2016	<i>PMB approved</i>
D3.7	WP3	First release of the EGI Service Registry and Marketplace prototype	DEM	SWING	PM18	-----	<i>Pending internal review</i>
D6.10	WP6	Infrastructure tests and best usage practices for life science service providers	R	CSC	PM15	29/06/2016	<i>PMB approved</i>
D6.12	WP6	GPGPU-enabled web portal(s) for MoBrain	OTHER	CIRMMP	PM16	11/07/2016	<i>PMB approved</i>
D6.7	WP6	Implementation and evaluation of AMBER and/or GROMACS	R	CIRMMP	PM13	29/06/2016	<i>PMB approved</i>
D6.9	WP6	Web portals for tsunami wave propagation simulations and for WRF-based weather simulation	OTHER	AS	PM14	24/05/2016	<i>PMB approved</i>
M6.5	WP6	Joint training program for the second period is agreed	R	EGI Foundation	PM15	31/05/2016	<i>PMB approved</i>

¹⁰⁶ (*) Dates are expressed in project month (1-30).

(**) Status = Not started – In preparation – Pending internal review – PMB approved

(***) Type = **R** = Document, report **DEC** = Website, press & media actions, events **DEM** = Demonstrators, pilots, prototypes, plan design **OTHER** = Software, technical diagram etc.

Id	Activity No	Deliverable / Milestone title	Type (***)	Lead partner	Original Delivery date(*) ¹⁰⁶	Revised delivery date	Status (**)
M4.2	WP4	Launch of call for cross e-Infrastructure case studies	OTHER	EGI Foundation	PM15	16/06/2016	<i>PMB approved</i>
M5.2	WP5	e-CEO challenges run on EGI resources	OTHER	EGI Foundation	PM15	21/06/2016	<i>PMB approved</i>
M5.3	WP5	VREs for the fishery and marine sciences community are deployed in production	OTHER	EGI Foundation	PM18	31/08/2016	<i>PMB approved</i>

7.3.2 Consumption of Effort

This section reports on the effort consumed in the period PM13 to PM18. The report was extracted on October 14 and provides effort views per WP and per partner in each WP.

Overall the efforts used are in line with the plan; some partners deviate but the activity performed is satisfactory and no corrective action is envisaged at this stage.

During the coming quarter and for the preparation of M21 report upon which a subsequent payment will be released, the WP and task leaders will watch these deviations, request clarification and take the appropriate measures. Payment will be retained until clarification is completed.

Table 1. Effort consumed during M13 to M18 per WP.

Work Package	P2 M13 - M24		
	Planned PMs months period 12	Actual PM Months period 6	%
WP01	31.8	16.5	52%
WP02	85.9	49.7	58%
WP03	61.0	39.5	65%
WP04	66.2	48.4	73%
WP05	46.1	17.3	38%
WP06	194.0	93.3	48%
Grand Total	484.9	264.8	55%

Table 2. Effort consumed during M13 to M18 per WP and per task.

		P2 M13 - M24		
Work Package	Task	Planned PMs 12 months period	Actual PM Months period 6	%
WP01	T01.01	7.83	3.81	49%
	T01.02	16.06	8.59	53%
	T01.03	7.94	4.13	52%
WP01 Total		31.83	16.53	52%
WP02	T02.01	20.57	10.91	53%
	T02.02	36.58	19.71	54%
	T02.03	28.72	19.07	66%
WP02 Total		85.87	49.69	58%

WP03	T03.01	11.1	4.23	38%
	T03.02	9.13	2.42	27%
	T03.03	11.31	8.66	77%
	T03.04	22.38	14.33	64%
	T03.05	7.05	9.91	141%
WP03 Total		60.97	39.55	65%
WP04	T04.01	19.13	19.58	102%
	T04.02	27.95	19.23	69%
	T04.03	12.91	4.90	38%
	T04.04	6.18	4.71	76%
WP04 Total		66.17	48.42	73%
WP05	T05.01	24.5	11.47	47%
	T05.02	12.48	3.37	27%
	T05.03	9.1	2.50	27%
WP05 Total		46.08	17.34	38%
WP06	T06.01	23.52	10.60	45%
	T06.02	16.24	9.57	59%
	T06.03	19.72	7.11	36%
	T06.04	13.33	5.08	38%
	T06.05	24.1	15.06	62%
	T06.06	19.24	15.41	80%
	T06.07	25.64	15.19	59%
	T06.08	13.56	3.39	25%
	T06.09	15.88	6.53	41%
	T06.10	22.77	5.31	23%
WP06 Total		194	93.26	48%
Grand Total		484.92	264.78	55%

7.3.3 Overall Financial Status

This section provides effort views aggregated by partner.

Participant	P2 M13 - M24			
	Planned PMs 12 months period	Actual PM 6 Months period	Used PMs %	Total estimated costs
P01: EGI.eu	123.5	57.7	47%	512,837.7
P02: OEAW	1.2	0.1	4%	653.5
P03: VLIZ	2.7	0.2	7%	2,666.9
P04: IICT-BAS	1.8	0.9	50%	5,860.2
P05: SwiNG	6.6	-	0%	-
P06: CESNET	26.4	10.9	41%	4,931.3
P07: Fraunhofer	2.7	1.0	35%	6,776.3
P08: GWDG	1.0	2.4	240%	16,653.6
P09: CSIC	21.9	15.8	72%	87,731.9
P09-1: UAB	5.1	1.9	38%	8,851.3
P09-2: UNIZAR	7.7	3.0	39%	12,000.0
P09-3: UPVLC	6.8	10.8	160%	36,468.9
P09-4: CESGA	10.6	7.4	70%	29,897.0
P10: CSC	8.7	2.9	34%	25,156.4
P11: CNRS	14.8	13.3	90%	69,772.3
P12: INRA	3.5	1.1	31%	-
P13: GRNET	22.9	9.3	40%	57,505.6
P13-1: IASA	6.9	3.4	50%	21,008.8
P13-2: AUTH	5.4	3.1	58%	19,247.7
P14: SRCE	3.8	2.2	58%	8,667.4
P14-1: RBI	8.1	4.9	60%	19,376.0
P15: MTA SZTAKI	4.4	3.9	88%	21,325.2
P16: INFN	35.8	24.0	67%	131,771.2
P16-1: INAF	5.6	-	0%	-

P16-2: INGV	2.4	-	0%	-
P17: CIRMMP	7.8	5.4	69%	15,461.1
P18: IRCCS FBF	-	-	-	-
P19: CNR	5.5	2.1	39%	17,963.7
P20: Engineering	1.7	3.1	189%	-
P21: SURFsara BV	3.0	1.4	45%	12,106.2
P21-1: KNAW	1.0	0.8	82%	7,961.8
P21-2: KNMI	1.6	-	0%	-
P21-3: NIKHEF	7.5	3.8	50%	47,877.1
P21-4: BCBR	3.2	-	0%	1,893.5
P22: CYFRONET	30.1	29.4	98%	132,992.7
P23: LIP	8.1	4.0	50%	18,580.8
P24: ICETA	1.2	-	0%	-

	P2 M13 - M24			
Participant	Planned PMs 12 months period	Actual PM 6 Months period	Used PMs %	Total estimated costs
P25-1: KTH	3.0	-	0%	-
P25-2: UMEA	2.5	0.1	6%	1,258.1
P26: IISAS	3.4	2.1	63%	14,962.9
P27: TUBITAK	2.0	1.0	52%	6,327.7
P28: STFC	17.5	9.8	56%	73,534.0
P29: BBMRI-ERIC	3.1	-	0%	-
P29-1: BBMRI-CZ	2.0	1.7	85%	4,527.2
P29-2: BBMRI-SE	0.5	-	0%	-
P29-3: BBMRI-NL	2.5	2.0	81%	15,135.7
P30: EMBL	2.2	0.4	17%	3,628.6
P31: CERN	3.1	2.8	92%	35,465.4
P32: EISCAT	5.6	2.3	41%	30,355.4
P33: FAO	4.5	6.0	133%	17,912.8
P34: Agro-Know	-	-	-	510.6
P35: GNUBILA	1.6	1.0	63%	7,210.0
P36: IU (OSG)	2.3	-	0%	-
P37: AS	19.2	4.3	22%	26,875.0
P38: ASTI	0.8	0.5	63%	3,125.0
P39: ITB	1.6	0.2	13%	1,325.0
P40: UPM	1.2	0.3	25%	1,875.0
Grand Total	484.9	264.8	55%	1,598,024.3

The estimate of the personnel and travel costs is given for information purpose only. Some of these personnel costs have been calculated using the average PMs costs of the budget. Not all travel costs have been reported at the time of writing. The actual costs will be provided at M24 for the preparation of the 2nd periodic report.

Partners' report and clarifications:

- P08: GWDG: PMs reported includes unfunded efforts that are performed in addition to the planned efforts.
- P09-3: UPVLC: exceeded plan in WP4-JRA2
- P12: INRA: PMs costs not provided
- P20: Engineering: PMs costs not provided
- P20: Engineering: activity peak in Y2 and reduced in Y3
- P22: CYFRONET: PMs reported includes unfunded efforts that are performed in addition to the planned efforts.
- P25-1: KTH & P29-2: BBMRI-SE: no activity to report due to lack of staff after departure of the person in charge of the activity. The recruitment is ongoing.
- P33: FAO: activity ending in Y2
- P05 & P36 – P40: Unfunded Partners

7.4 Issues and Treatment

There is no significant deviation that requires a corrective action at this stage. The WP and Task leader will keep watching the activity performed. A payment related to PY1 costs was released in September 2016 and for those partners who have not yet reached the 85% threshold an additional payment will be calculated at PM21. Persistent deviations will lead to a corrective action to ensure that PY2 plan can be realised as defined in the DoA.

7.5 Plans for period PM19-PM24

Amendment 2 of the project is under discussion with the Project Management Board. It will be developed and submitted to the Collaboration Board mid November. Then it will be finalized and submitted to the EU.

8 Project Metrics

8.1 Overall metrics

Objective	Metric ID	Impact and Metric	Value PM18	Target PM24
O5	KPI.2.SA1.Integration	Number of RIs and e-Infrastructures integrated with EGI	15	15
O3	KPI.3.SA1.Software	Number of new registered software items and VM appliances	7/18	30/90
O3	KPI.4.SA1.Cloud	Number of providers offering compute and storage capacity accessible through open standard interfaces	23	25
O1	KPI.5.SA2.Users	Estimated total number of researchers served by EGI	48169	48 000
O2	KPI.7.SA2.Users	Number of new research communities served	8	20
O2	KPI.8.SA1.Users	Number of VO SLAs established and number of long tail of science SLA (num1/num2) <i>Policies, tools and procedures for the long tail of science are complete. The service catalogue for the long tail of science user group is under finalization and will be disseminated during PM19-PM24</i>	5/0	10/100
O5	KPI.16.SA2.Support	Number of international support cases (for/with RIs, projects, industry)	56	30
O4	KPI.18.NA2.Industry (New)	Number of SME/Industry that successfully implemented a use case involving EGI services <i>Two SLAs under completion: Terradue (Earth Observation data and data products exploitation, SLA started on Jan 2016) and Peachnote (musicology, SLA started on March 2014)</i>	2	2
O2	KPI.19.NA2.Partnerships (New)	Number of PaaS providers that are EGI partners: EGI Marketplace partners/Technology partners (num1/num2)	5	10

8.2 Activity metrics

8.2.1 NA1 – Project Management

Metric ID	Metric	Value PM06
M.NA1.Quality.1	Number of days of Deliverable, milestone delay per WP	<ul style="list-style-type: none"> • Wp1: 21 • Wp2: 17 • Wp3: 98 • Wp4: 15 • Wp5: 20 • Wp6: 154
M.NA1.Quality.2 (new)	Percentage of delayed deliverables and milestones per WP	<ul style="list-style-type: none"> • Wp1 100% • Wp2 33% • Wp3 80% • Wp4 100% • Wp5 50% • Wp6 80%

8.2.2 NA2 – Strategy, Policy and Communication

Metric ID	Metric	Value PM06
M.NA2.Communication.1	Percentage of articles, news, blog posts about or contributed by user communities and NGIs/EIROs with respect to the total of items published in EGI's channels	39%
M.NA2.Communication.2	Number of unique visitors to the website	17244
M.NA2.Communication.3	Number of page views on the website	61879
M.NA2.Communication.4	Number of news items published	23
M.NA2.Communication.6	Number of case studies published	1
M.NA2.Communication.7	Attendee-days per event	666
M.NA2.Communication.8	Number of scientific publications supported by EGI	/gathered yearly/
M.NA2.Strategy.1	Number of EGI impact assessment reports circulated to the stakeholders <i>Interim impact report prepared for EC and consortium members, final version circulated in October</i>	1
M.NA2.Strategy.2	Number of MoUs involving EGI.eu or EGI-Engage as a project	4
M.NA2.Strategy.3	Number of contracts established with paying customers <i>Financial advice from Dutch authorities</i>	0

	<i>on applicable tax regime for the EGI Foundation obtained during the current reporting period</i>	
M.NA2.Strategy.4 (exKPI)	Number of relevant authorities informed of the policy paper on procurement	0
M.NA2.Industry.1	Number of engaged SMEs/Industry contacts	81
M.NA2.Industry.2	Number of establish collaborations with SMEs/Industry (e.g. MoU)	11
M.NA2.Industry.3	Number of requirements gathered from market analysis activities	30
M.NA2.Industry.4 (exKPI)	Number of services, demonstrators and project ideas running on EGI for SMEs and industry	43

8.2.3 JRA1 – E-Infrastructure Commons

Metric ID	Metric	Value PM06
M.JRA1.AAI.1	Number of communities adopting federated IdP	1
M.JRA1.Marketplace.1	Number of entries in the EGI Marketplace (i.e. services, applications etc.) <i>First prototype of Marketplace due in the next reporting period</i>	0
M.JRA1.Accounting.1	Number of kinds of data repository systems being integrated with the EGI accounting software <i>Dataset usage record standard defined during the current reporting period</i>	0
M.JRA1.Accounting.2	Number of kinds of storage systems being integrated with the EGI accounting software	2
M.JRA1.OpsTools.1	Number of new requirements introduced in the roadmap	9
M.JRA1.OpsTools.2	Number of probes developed to monitor cloud resources	2

8.2.4 JRA2 – Platforms for the Data Commons

Metric ID	Metric	Value M06
M.JRA2.Cloud.1	Number of VM instances managed through AppDB GUI	70

M.JRA2.Cloud.2	Percentage of cloud providers providing snapshot support <i>The interface implementing OCCl 1.2 was released during this reporting period</i>	0
M.JRA2.Cloud.3	Percentage of cloud providers providing VM resizing support <i>The interface implementing OCCl 1.2 was released during this reporting period</i>	0
M.JRA2.Cloud.4	Number of OCCl implementation supporting OCCl 1.2	1
M.JRA2.Integration.1	Number of European cloud providers in the federated Astronomy community cloud <i>CANFAR integration activities are in progress</i>	0
M.JRA2.Integration.2	Number of virtual appliances shared	93
M.JRA2.Integration.4	Number of EUDAT services integrated with the HTC and Cloud platforms of EGI	2
M.JRA2.Integration.6	Number of research clouds that interoperate with EGI federated cloud: community clouds, integrated, peer	2
M.JRA2.Integration.7 (New)	Number of models executed on Federated Cloud resources	0
M.JRA2.Integration.8 (New)	Number of CPUs consumed on Federated Cloud resources	0
M.JRA2.AcceleratedComputing.1	Number of batch systems for which GPGPU integration is possible to be supported through CREAM	5
M.JRA2.AcceleratedComputing.2	Number of Cloud Middleware Frameworks for which GPGPU integration is supported and implemented	1
M.JRA2.AcceleratedComputing.3	Number of level 3 disciplines with user applications that can use federated accelerated computing	3
M.JAR2.OpenData1 (exKPI)	Number of open research datasets that can be published, discovered, used and reused by EGI applications/tools <i>Open Data Platform to released as prototype in the next reporting</i>	0

	<i>period. Testing conducted with brain atlas information (Human Brain Project), biodiversity datasets (LifeWatch) and Copernicus Earth Observation data.</i>	
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------	--

8.2.5 SA1 – Operations

Metric ID	Metric	Value M06
M.SA1.Operations.3	Amount of allocated resources (storage) allocated through a EGI centrally managed pool of resources to Long tail of science	1.2TB
M.SA1.Operations.4	Amount of allocated resources (logical cores) allocated through a EGI centrally managed pool of resources to Long tail of science	3360
M.SA1.Operations.5	Number of new products distributed with UMD <i>Preparation of new release for cloud middleware (CMD) completed in current reporting period</i>	0
M.SA1.Operations.6 (exKPI)	Number of CPU cores available to international research communities and long tail of science (all user communities included, HTC and Cloud) [HTC/Cloud]	HTC: 826497. Cloud: 6646
M.SA1.Operations.7 (exKPI)	Number of storage available to international research communities and long tail of science (disk and tape, HTC and Cloud) [PB] [HTC/Tape/Cloud]	HTC Disk: 285.3 PB . HTC Tape: 281 PB. Cloud: 425.5 TB
M.SA1.Operations.8 (exJRA1)	Number of user requests handled in e-GRANT	3
M.SA1.Operations.9 (exKPI)	Number of compute resources available to the long tail of science	18249
M.SA1.Operations.10 (New)	Number of CPU time consumed by robot certificates	519,037,353 hours
M.SA1.SecurityOperations.1	Number of security policies and procedures updated reviewed and adapted to support new services	9
M.SA1.Platforms.1	Number of gCUBE VREs instantiated on the Federated Cloud for the iMARINE community <i>Technical integration completed in the current reporting period</i>	0

M.SA1.Platforms.2	Number of CPU time consumed by e-CEO challenges (hours * cores)	2593
M.SA1.Platforms.3 (New)	Amount of computing resources used by long tail of science, both Cloud and HTC	3800 CPU * h
M.SA1.Users1 (exKPI)	User satisfaction	/gathered yearly/

8.2.6 SA2 – Knowledge Commons

Metric ID	Metric	Value M06
M.SA2.UserSupport.1	Number of training modules produced and kept up-to-date	11
M.SA2.UserSupport.2	HTC Absolute normalized CPU time to a reference value of HEPSEC06 (excluding OPS and dteam) per 1 level disciplines IN HOURS	<ul style="list-style-type: none"> Engineering and Technology: 313,491,026 Humanities: 7,631,637 Medical and Health Sciences: 18,364,563 Natural Sciences: 2,042,070,248 Social Sciences: 7,631,637 Support Activities: 473,055,286
M.SA2.UserSupport.3	HTC Relative increase normalized CPU time to a reference value of HEPSEC06 (excluding OPS and dteam) per 1 level disciplines	<ul style="list-style-type: none"> Agricultural Sciences -23% Engineering and Technology -16% Humanities 38% Medical and Health 152% Natural Sciences -25% Social Sciences 38% Support Activities -74%
M.SA2.UserSupport.4	Relative increase of users per 1 level disciplines	<ul style="list-style-type: none"> Engineering and Technology : 11% Medical and Health Sciences: 9% Natural Sciences: 4% Agricultural Sciences: 17% Social Sciences: 13% Humanities: 12% Support Activities: -2% Others: -8%
M.SA2.UserSupport.5	HTC Number of Low/Medium/High Activity VOs and total	Low: 110 Medium: 100 High: 80 Total: 290
M.SA2.UserSupport.6	Number of VM instantiated in Federated Cloud per 1 level discipline	117929 in total

M.SA2.UserSupport.7 (exKPI)	Number of delivered knowledge transfer events	5
M.SA2.UserSupport.8 (New)	Number of robot certificates used in EGI Infrastructure	180

Dissemination report

The report on the dissemination plan is provided as a separate document at <https://documents.egi.eu/document/2832>.



This material by Parties of the EGI-Engage Consortium is licensed under a [Creative Commons Attribution 4.0 International License](#).

The EGI-Engage project is co-funded by the European Union (EU) Horizon 2020 program under Grant number 654142 <http://go.egi.eu/eng>