D-NA1.3.1: Management and progress report at PM 36

10/10/2014

Project acronym:	VERCE
Project n°:	283543
Funding Scheme:	Combination of CP & CSA
Call Identifier:	FP7-INFRASTRUCTURES-2011-2
WP:	WP1/NA1, Management
Filename:	D-NA1.3.1.pdf
Author(s):	J.P. Vilotte
Location:	http://www.verce.eu/Repository/Deliverables/RP4
Type of document:	EU intermediate report
Dissemination level:	Public
Status:	Final
Due date of delivery:	30/09/2014 – revised 20/11/2015
Keywords:	NA1, Management and progress report.

Version	*Author*		*Date*	*Comments*
1	A. Mulle	CNRS	23/09/2014	Initial draft for comments.
2	J.P. Vilotte	CNRS-IPGP	25/09/2014	Project's objectives, work packages progress
3	A. Weexsteen	CNRS	26/09/2014	Review and additional editing.
4	J.P. Vilotte	CNRS-IPGP	28/09/2014	Review and additional editing.
5	J.P. Vilotte	CNRS-IPGP	1/10/2014	Additional editing.
6	A. Mulle	CNRS	3/10/2014	Review.
7	A. Weexsteen	CNRS	3/10/2014	Additional review.
8	J.P. Vilotte	CNRS-IPGP	5/10/2014	Final editing and review.
9	J.P. Vilotte	CNRS-IPGP	15/11/2015	Improvements and additional editing.

Copyright notice

Declaration by the scientific representative of the project coordinator

Grant Agreement number: 283543

Project acronym: VERCE

Project title: Virtual Earthquake and seismology Research Community e-science environment in Europe

Funding Scheme: Combination of CP & CSA

Date of latest version of Annex I against which the assessment will be made: 16/08/2011

Intermediate Periodic Report: 4th period

Period covered: final overview

Name, title and organisation of the scientific representative[1]**:** Professor Jean-Pierre Vilotte, Institut de Physique du Globe de Paris

Tel: +33 1 83 95 75 83

Fax: +33 1 83 95 77 17

E-mail: vilotte@ipgp.fr

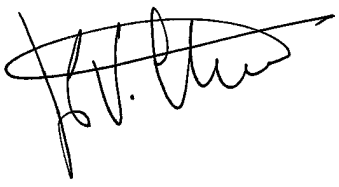
Project website: http://www.verce.eu/

I, as scientific representative of the coordinator of this project and in line with the obligations as stated in Article II.2.3 of the Grant Agreement declare that:

- The periodic report represents an accurate description of the work carried out in this project for this reporting period.
- The project (tick as appropriate)[a]:
  - ☐ has fully achieved its objectives and technical for the period;
  - × has achieved most of its objectives and technical goals for the period with relatively minor deviations;
  - ☐ has failed to achieve critical objectives and/or is not all in schedule.
- The public website, if applicable
  - × is up to date
  - ☐ is not up to date
- To my best knowledge, the financial statements which are being submitted as part of this report are in line with the actual work carried out and are consistent with the report on the resources used for the project (section 3.4) and if applicable with the certificate on financial statement.
- All beneficiaries, in particular non-profit public bodies, secondary and higher education establishments, research organisations and SMEs, have declared to have verified their legal status. Any changes have been reported under section 3.2.3 (Project Management) in accordance with Article II.3.f of the Grant Agreement.

Name of the scientific representative of the coordinator: **Jean-Pierre Vilotte**

Date: 05/10/2014



For most of the projects, the signature of this declaration could be done directly via the IT reporting tool through an adapted IT mechanism.

[a] If either these boxes below is ticked, the report should reflect these and any remedial actions taken.

Contents

List of Figures

List of Tables

Executive summary

Context and objectives

Seismology addresses both fundamental problems in understanding Earth's internal wave sources and structures and augmented societal applications; and put a great premium on open-access data archives integrated globally. Today our ability to acquire data outpaces our ability to explore and analyse them. New discoveries will emerge from statistical analysis and modelling of the wealth of data generated by observation and monitoring systems.

The objective of VERCE is to provide a comprehensive and integrated *virtual research environment* (VRE) delivering to a broad base of seismology researchers in Europe easy to use high-performance full waveform simulations together with a data-intensive framework for the collaborative development of innovative statistical data analysis methods.

The present VERCE E-Infrastructure involves the following major elements moving from the researchers' point of contact to the contextual digital resources:

- The *scientific gateway* that is web accessible from anywhere, provides an integrated view of all available resources, handles continuity between sessions and supports collaboration with shared data, methods, and codes, and with pervasive data-access control.
- The *framework for computational science*, a user-oriented full waveform simulation (FWS) service within the science gateway, accommodating simulation codes, arranges for their use including the supply of input data and the acquisition of results with associated metadata.
- The *framework for data-intensive analysis*, a developer-oriented service linked with the science gateway, allowing advanced researchers to develop and compose new data analytic methods, to deliver them as fine-grained workflows, handling data streams with pervasive provenance management system, to integrate diverse data management systems, and to automate their mapping to different Distributed Computing Infrastructures (DCIs) architectures.
- The *research platform*, an operational platform that federates and combines diverse computing and data resources provided by the VERCE partners and the European DCI providers within the VERCE virtual organisation (VO), together with a number of monitoring software components, balancing flexibility with performance and resilience with scalability.
- The *mapping to multiple distributed computing infrastructures* handling identity management, authority controls, transformations between representations and protocols and access to resources.
- An *extensive collection of data handling services and components* that are used at different stages of the above contributions, as well as extensible coupling with the seismology community data services.
- A *set of provenance tools* exploiting pervasive provenance records to support validation, quality control of data and methods.

The design of the science gateway and the VERCE platform incorporates a number of '*intellectual ramps*", that ensure new users encounter minimum hurdles at the start, and can progress to sophisticated use of the full facilities without encountering steep learning curves. A set of training sessions complements those ramps, together with best-practice guides and exemplars, such as reference workflows and data processing pipelines.

Following the last review meeting, a number of recommendations were made for the last-mile implementation, and led to a roadmap of planned services, tools and E-Infrastructure. The roadmap and the VERCE strategy for their implementation are detailed in a separate document.

Work carried in the last six months

Pilot applications and use cases

Priority during this period was given to: (1) finalise the integration of the FWS service within the science gateway and extend the user-oriented functionalities for induction of new seismology researchers; (2) further develop and evaluate the data-intensive framework for the analysis of both observations and simulated results, i.e., including event and continuous waveforms analysis; provide training material, exemplars and documentation for the set of training sessions open to external seismology users.

The main achievements during this period are:

- Forward wave simulation (FWS): The first demonstrated integration within the science gateway was internally evaluated together with external users' feedback collected after the online training events. This led to refine the FWS use case and services to accommodate new workflow components including: user-oriented interfaces for the configuration and the parameterisation of the SPECFEM3D[2] input files; library of shared meshes and earth models; new visualisation and presentation tools for the FWS results integrating external geospatial standards and Google Earth KML file generation; new automated scripts and functionalities for automated check and validation of user submitted meshes.

- Data-intensive framework: a number of physical and remote meetings accelerated the integration of the `dispel4py` framework and `ObsPy`[3], together with the design of shared processing elements (PEs) and pre-configured pipelines for the analysis of event and continuous waveforms, *i.e.* misfit analysis of observed and simulated event waveforms, continuous wave form analysis for seismic noise correlation and seismic source detection applications. This led to a suite of test cases and exemplars for the evaluation of the scalability and of the performance of the `dispel4py` framework onto different DCIs architectures.

- Training material and user documentation: new training material, exemplars and user documentation were developed for the online training sessions introducing the VERCE platform and its computational and data-intensive framework functionalities.

Training and user documentation

Priority was given during this period to: (1) "intellectual ramps" supported by the science gateway for the FWS induction of external seismology users; (2) a set of online training sessions, *e.g.* webinars, open to external users with hands-on practicals and exemplars. "intellectual ramps" and online training were developed hand-in-hand to ensure that students and researchers, with different levels of expertise, can progress incrementally to sophisticated use of the computational and data-intensive frameworks and develop new data analysis skills and computational thinking in their research practice.

The main achievements are:

- Two online training sessions were organised in July 2014, open to external users. They provided an exhaustive introduction to the VRE and the E-Infrastructure, together with hands-on sessions and exemplars of the integrated FWS services within the science gateway. User documentation were designed including a science gateway and FWS portlet installation guide, A feedback questionnaire was distributed to the ten participants and analysed. Overall the feedback was extremely positive, and the style of training well received. Further flexibility needs were identified and will be integrated. Training videos were made available in the VERCE website.

[2] geodynamics.org/cig/software/specfem3d/
[3] github.com/obspy/obspy/wiki

- A second online training session has been prepared and will be held in October 2014. It will focus on the data-intensive framework, providing an introduction to `dispel4py`, together with hands-on sessions and workflow exemplars of pre-processing and seismic noise correlation pipelines. It is expected that at the end of this training session attendees will take home their own developed workflow. User documentation and training videos will be published in the website.
- During the next period, two training workshops are already planed in March 2015 at LRZ and in July 2015 at the university of Liverpool. Those workshops will as well be advertised and open to external users from different institutions in Europe.

Table 1 – Training plan

HPC on-line workshop	**July 10 & 17, 2014** 10 registered external users	10 MScs, PhDs and Post-Docs interested in the wave-simulations use case for their research	Open to external beta-testers	September 2014 – feed back session and reports on how to improve the system and new services
Data-intensive analysis on-line workshop	**October 15 & 16, 2014** 9 registered external users end of September	Up to 10 MScs, PhDs and Post-Docs interested in the `dispel4py` framework for seismic waveform analysis	Open to external beta-testers	December 2014 – feed back session on how to improve the system, performance, and scalability
Training workshop (LRZ)	**March 2015** 3 days session	Up to 25 participants interested in the VERCE VRE and E-Infrastructure for their research	Open to a wide community including the related European EPOS and ERC projects and E-Infrastructure providers	April 2015 – feedback on how to improve the system for their research practice
Training workshop (ULIV)	**July 2015** 3 days session	Up to 25 participants interested in the VERCE VRE and E-Infrastructure for their research	Open to a wide community including the European EPOS and ERC projects and E-Infrastructure providers	August 2015 – feedback on the VERCE platform and architecture

Dissemination and public outreach

During this period, the priority was given to improve and refine the dissemination material and the communication tools to promote VERCE including: digital tools (website, general social media, video social media), electronic dissemination materials, metric tools.

The main achievements are:

- The VERCE website has been redesigned and restructured. The website traffic is continuously monitored and increasing.
- Videos of the training sessions were added to the VERCE website.
- The science gateway was externally promoted through tutorial videos.

Management and operation of the research platform

Priority during this period was to: (1) improve the monitoring and the continuity of services provided by the research platform federating distributed of computing and data resources provided by the VERCE partners and European DCIs, *e.g.* such as LRZ and CINECA; (2) operate and extend the Virtual Organisation (VO) that provides for the increasing number of registered users access to the federated resources, tools and services, handling and minimising the burden of the diversity of existing authentication, authorisation and identification (AAI) policies and protocols across the autonomous providers that VERCE federates; (3) explore Cloud resources provided by SCAI and the EGI Cloud federation.

The main achievements during this period are:

- The administration and the definition of users' roles in the VO are now managed by tools in the science gateway to insure semantic consistency across contexts, and `VOMS` made possible to support single sign in on and automated delegation to establish access to the federated resources. An `OTRS` ticketing support has been integrated within the science gateway with different levels of support.
- The operation and the management of the number of integrated instantiation, management and monitoring software components, both internal and external ones, that are distributed across the platform sites, *i.e.* not all the components reside at any one site, has been improved. Resilience and sustainability rely on appropriate insulation from changes in software services, data representations, network protocols and hardware architecture. In particular, operation, resilience and scalability of the science gateway, together with auxiliary services, have been improved via virtual machines running in an `OpenNebula`[4] Cloud environment, together with a backup strategy via an `OpenNebula API`, which hosts the portal front-end and the `gUSE`[5] backend.
- The research platform integrates diverse data-management systems, including file systems, on the many computing resources in use, database management systems and shared data services. The underlying mapping system needs to choose the right transfer mechanism, *e.g.* `GridFTP` for large volumes as it handles recovery and retry, but `ftp` and `http` for small transfers.
- To provide caches and persistent storage of users' workspaces, and of the FDSN[6] downloaded data, an `iRODS`[7] federation of storage sites is integrated into the research platform and managed. This system supports association of metadata and persistent identifiers with files. To provide redundancy for data access and mitigate availability issues, *e.g.* data locality, replication strategy has been investigated between the different `iRODS` sites. Incremental transfer to the services provided by EUDAT[8] is anticipated, provided they can also be set up in autonomous partners' contexts. However they were not yet mature enough in this period.
- The distributed system for results and their metadata is closely linked with provenance, which uses `MongoDB`[9] as it distributed `NoSQL` database management system. The broad user bases of `iRODS` and `MongoDB` motivated their selection, support from EUDAT will help sustainability.
- A proof-of-concept Cloud application has been developed allowing users to discover and download seismic traces from data archives through web service compliant with the FDSN standards, and to process them using `ObsPy` on Cloud resources. It is now in evaluation by seismologists.
- The next objectives are: (1) improve the maintenance and the backup strategy of the science gateway virtual machines and services; (2) simplify and improve the data replication performance on

[4] opennebula.org
[5] guse.hu/about/home
[6] www.fdsn.org
[7] irods.org
[8] eudat.eu
[9] www.mongodb.com

the data platform; (3) improve the operation of the X509-based authentication infrastructures, *e.g.* for the PRACE[10] resources.

Integration and evaluation of the platform services

Priority was given during this period to: (1) the fifth release of the platform services and documentation; (2) the preparation of the next release cycle.

The main achievements are:

- The fifth PDCA cycle was completed by the end of July, 2014 and released, *i.e.* four components of the data-intensive framework were evaluated and approved.
- The next six months roadmap and the schedule for the next release cycle, *e.g.* October, 2014 - March, 2015, have been defined. Requirements of tools and services to be evaluated are being collected through the Request Form. Significant efforts will be deployed for enabling the integration of the data-intensive misfit analysis and its integration within the science gateway.

Scientific gateway and user interfaces

During this period, efforts have focused on: (1) providing a number of "intellectual ramps" incorporated into the science gateway for induction of new seismology researchers, and for facilitating the external evaluation of the functionalities by beta users; (2) extending the FWS services and functionalities of the science gateway in response to the users' evaluation feedback; (3) first integration of advanced tools for data processing and analysis of the dispel4py framework; (4) improving the performances and the robustness of the science gateway services and tools. This has been helped by active collaborations between VERCE and the SCI-BUS[11] and ER-Flow[12] projects.

The main achievements are:

- The science gateway has been improved to support the required modes of interaction and to provide convenient and understandable access to tools and functions that enable all stages of the FWS research method. A first set concerns the front-end with interactive and context dependent user-oriented interfaces that automate marshalling inputs and collecting results, and present it as a simple operation where seismologists can select a region, choose an existing mesh and Earth model from a shared library, configure and manage their runs for a selected number of events and stations. It integrates FDSN services with external geospatial standards, *e.g.* typically those supporting OGC[13] standards mandated by the INSPIRE[14] directives, as well as visualisation services. A second set concerns the refactoring of the FWS workflow, which now integrates advanced tools of the dispel4py framework, offers runtime feedback and context dependent exploration of the provenance information collected at different stages of the execution of the workflow, together with data and metadata sharing functionalities.
- The science gateway is now linked with the SHIWA workflow repository[15], and two FWS workflows have already been exported and listed among the seismology collection hosted by the repository and downloadable by other portals.
- Components of the dispel4py framework are integrated for some critical data processing components of the FWS workflow, *e.g.* post-processing pipelines of the simulated waveforms includ-

[10] www.prace-ri.eu
[11] www.sci-bus.eu
[12] www.erflow.eu
[13] www.opengeospatial.org
[14] inspire.ec.europa.eu
[15] shiwa-repo.cpc.wmin.ac.uk/shiwa-repo/

ing the production of images, plots and other products. This brought performance improvement, thanks to the dispel4py automated mapping of fine-grained data streaming workflows onto diverse computing architectures.

- The number of registered researchers on the science gateway is rapidly growing, *e.g.* more than fourty today, and others are waiting to be approved.
- In the next months, priority will be given to further integration of the dispel4py framework in support of the misfit analysis, *e.g.* analysis of both observed and simulated waveforms, and to extend the Provenance system, its scalability and representation using the W3C-PROV[16] recommendation, with flexibility and efficiency based on JSON[17].

Harnessing data-intensive application-software

Priority was given during this period to: (1) FWS application-software tuned for multiple DCIs architectures; (2) data-intensive processing elements and pipelines; (3) "intellectual ramps" and FWS demonstrators for the set of training sessions and the induction of new seismology researchers.

The main achievements are:

- The FWS workflows have been mapped onto new HPC and DCIs resources of the research platform, together with a number of improvements in SPECFEM3D that is now accommodated in multiple HPC and institutional resources.
- The shared library of meshes and Earth models has been extended in phase with the users' research practice, together with new tools allowing advanced users to submit their own meshes and Earth models with a quality check procedure.
- Pre- and post-processing elements and pipelines have been developed and productised for the misfit analysis of observed and simulated event waveforms within the ObsPy and dispel4py frameworks extending the FWS workflow toward the orchestrated FWS-Misfit workflow that will be integrated within the science gateway. This includes new synthetic waveforms file format and provenance information supporting the W3C-PROV model, and improved provenance-based data movements between the VERCE caches and persistent storage layer, and the distributed data storage resources provided by the federated DCIs, handling the diversity of remote data-management systems and of their access and security protocols.
- The next steps are: (1) finalise the integration of the pre- and post-processing components within the science gateway for the FWS-Misfit workflows; (2) further developments within the ObsPy Python library; (3) pre-configured pipelines and exemplars of data-intensive analysis of continuous waveforms that will be provided by the data-intensive framework, including a pervasive provenance system; preparation of the planed training workshops in 2015.

Architecture and platform tools for data-intensive analysis and modelling

Through UEDIN and ULIV, a new collaborative framework has been set between VERCE and Terra-Correlator, a NERC-funded project at Edinburgh, for data-intensive seismic analysis. Collaboration with ER-Flow was finalised by a MoU.

The main achievements during this period are:

- dispel4py: a Python framework for distributed data-intensive fine-grained abstract streaming workflows, demonstrated in real use-cases for seismic noise cross correlation analysis, post processing of forward wave simulation misfit analysis. dispel4py supports mappings on several

[16] www.w3.org/TR/prov-overview/
[17] www.json.org/json-fr.html

computing architectures, including shared memory and multicore MPI architectures and Cloud environment. `dispel4py` is being linked and integrated with the science gateway in support of the data analysis stages of the FWS and misfit workflows.

- Provenance system: this system collects metadata associated to data transformation during the end-to-end workflows and their mapping onto different DCIs. The current implementation, apart provenance generation mechanisms, provides a prototype, browser-based interface and a web API built on top of a `NoSQL` storage technology. This offers combined operations for accessing and downloading data, which may be selectively stored at runtime, into dedicated archives.
- Science gateway, a production-ready `SCI-BUS`-powered science gateway is now in operation and linked to the SHIWA workflow repository.
- Improved integration and operation of certificates, in collaboration with SA1 and SA3, making authentication on the VERCE gateway easier.
- The next steps will be: (1) integration of the `dispel4py` framework for data-analysis components of the FWS-Misfit workflows; (2) `dispel4py` registry prototype to promote method and component description and sharing.

Impact of the VERCE project

A key asset is the agile strategy of VERCE deployed in a multi-organisational context. It engages seismologists, data scientists, ICT researchers, HPC and data resource providers, system administrators into short-lived task forces each with a goal, *e.g.* that is a seismology priority, and intimately coupling research thinking with technical innovation. This changes the focus from HPC production environments and community data services, where middleware standards, security procedures and connectivity dominated, to user-focussed scenarii. This avoids wasteful bouts of technology centricity where technologists collect requirements and develop a system that is not used because the ideas of the planned users have moved on.

This mutual understanding and active networking is an invaluable asset for the future of the larger community, with a new generation of solid-Earth scientists with data-use skills. It is also a key long-term goal, and the EPOS implementation phase, starting in 2015, will build on this momentum to deliver an e-science environment where a growing number of solid-Earth scientists will continue to advance their science and increase the power of their platform and of their research practice tackling data-intensive challenges.

The VERCE VRE enables seismologists from anywhere in Europe, wether or not they have local support, to harvest the full information content from large dense seismic networks inducing new research practices and improving the full path of data use.

By delivering this cutting edge technology to a broader seismological base, VERCE improves research productivity of the European seismological community. This improved research productivity is a world-leading advance, and provides a leading edge relative to many of the research competitors world wide.

Website: http://www.verce.eu

1 Project objectives and work progress for the period

This document, *e.g.* D-NA1.3.1, is the intermediate management and progress report covering the first six months of the fourth reporting period, from month 31 to 36.

VERCE is structured into interacting work packages (WPs), as shown in Fig. 1, which are beside the Management activities (NA1/WP1): Network activities (NA2/WP2 to NA4/WP4), Service activities (SA1/WP5 to SA3/WP7), and Joint Research & Development activities (JRA1/WP8 and JRA2/WP9).

Beside this structure, establishing short-lived, agile transversal task forces, which engage seismologists, data scientists, ICT researchers, HPC and Data resource providers, and system engineers, to discover solutions and ways forward that closely match seismology researchers' requirements has been an invaluable asset. This strategy is imported from modern commercial software engineering practice, but here it is deployed in a multi-organisational contexts.

These task forces worked intensively to co-develop an integrated view and to tease out a path to a solution that will be used. They cross barriers which would prove insuperable without this intense co-operation. As well as delivering prototype and demonstration to the productisation team, they also develop a very effective communication network– a common language and a commitment to respond positively to each other needs. This mutual understanding and active networking is an invaluable asset for the future of the project and of the larger community.

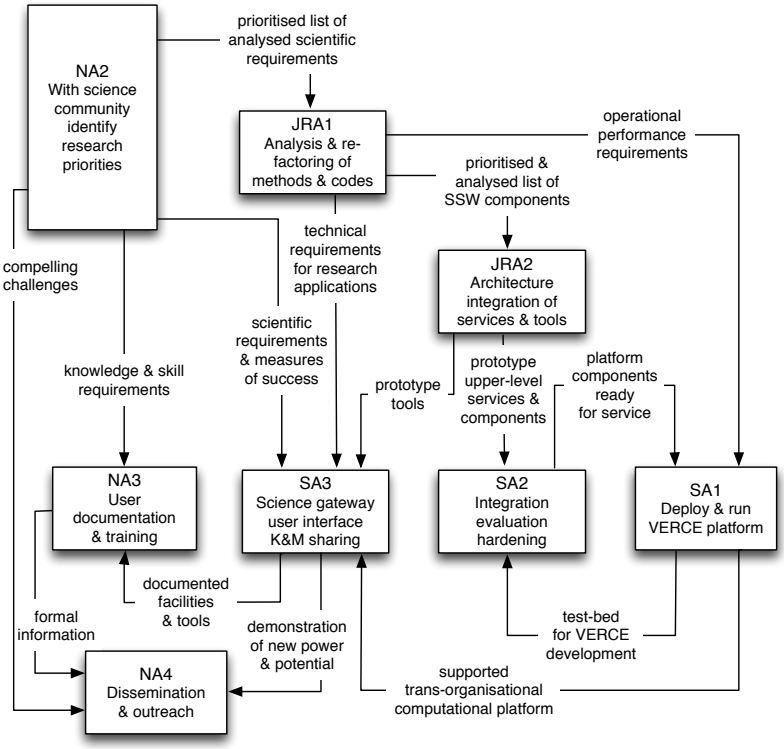


Figure 1 – Work packages in VERCE

1.1 Network activities

The network activities are user-and-research driven horizontal orchestration activities. The main strategy is to:

- Smooth the path from theoretical research through proof-of-concept to a dependable research e-science environment in phase with the research practice in seismology.

- Create and share data-intensive analysis and modelling methods, tools and practices for exploiting the wealth of seismological observations.
- Provide "intellectual ramps" and training, developed hand-in-hand, for induction of new seismology researchers.

The main achievements in the last six months are:

- Evaluation and validation of the enabled application use case and of their integration within the Scientific Gateway (D-NA2.4). This includes: a number of improvements and extension of the FWS use case and of its application-software, together with accommodating in the data-intensive framework, *e.g.* `ObsPy` and `dispel4py`, the diversity of the methods and of their implementation for the analysis of seismic waveforms, *e.g.* like in seismic noise correlation and seismic source detection applications.
- Demonstrators for dissemination and outreach (DNA2.4.1). This includes: FWS demonstrators using the science gateway services and tools, new user documentation and training material for the training sessions.
- Training sessions and material, together with the analysis of users' feedback (DNA3.2.2, NA3). This includes two online training sessions in July and October 2014, together with two planed training workshops during the year 2015, as well as users' guides and training videos available on the VERCE website.
- Dissemination and public outreach (DNA4.3.2). This includes: continuous update and monitoring of a number of social media (`Facebook`, `Twitter`, `Google+`) and professional media (`Linkedin`) tools, and improvement of the website.
- The activity is on time and the next six months priorities are: (1) new components and functions of the FWS workflows toward orchestrated FWS-Misfit workflows, including pre- and post- processing of seismic waveforms for the misfit evaluation integrating `dispel4py` processing elements and pipelines; (2) further evaluation of the `dispel4py` capabilities and performance for continuous waveform data-intensive analysis in the context of seismic noise correlation and seismic source detection; (3) documentation and demonstrators for the forthcoming set of training workshops.

1.2 Service activities

The service activities are at the interface between the user-oriented coordination activities and the research and development activities. The main strategy is to:

- Evaluate and integrate a platform of tools, services and application software components, relevant to the selected applications and to the platform operation including internal and external software components already adopted by the existing European E-Infrastructure providers, *e.g.* EGI[18], PRACE, IGE[19], or supported by other EU projects, *e.g.* EUDAT, SCI-BUS, ER-flow.
- Deploy and operate the successive versions of the platform providing a framework that eases the access across the federated data and computing resources and the continuity of service.
- Manage the release process through a Plan-Do-Check-Act (PDCA) cycle.
- Manage and operate the Virtual Organisation (VO) providing for the registered users a consistent and flexible access to the federated resources, tools and services, handling and minimising the burden of the diversity of existing AAI policies and protocols across multiple autonomous data and computing providers.

[18] www.egi.eu
[19] www.ige-project.eu

The main achievements in the last six months were:

- Operation and management of the VERCE platform (DSA1.4). This includes: extension and improvement of the iRODS-based data layer providing caches and persistent storage of users' workspaces, with redundancy and replication functionalities; improvement of the resilience and scalability of the science gateway through virtual machines in a Cloud environment, hosting the gUSE backend and portal front-end services together with the upgrade of the gUSE services to support the SHA2 encryption algorithm for EGI certificates; improvement and extension of the monitoring system integrating CINECA and INGV; integration of the OTRS ticketing system and new shared end-user support organisation within VERCE; a Cloud data analysis application prototype integrating ObsPy and FDSN data web services.
- Research platform integration and release (DSA2.4). This includes: the fifth release of the integrated tools and services with three new data-intensive components.
- Scientific gateway and use case integration (D-SA3.4). This includes: new components with improved usability and workflow management; refactoring of the FWS workflows offering runtime feedback and provenance-based action triggering; integration of dispel4py components for pre- and post-processing of simulated waveforms and provenance API; improved performance and mapping to multiple distributed computing infrastructures handling AAI controls and transformations between representation and protocols; incorporation of a number of "intellectual ramps" that ensure users encounter minimum hurdles at the start and can progress to sophisticated use of the full facilities.
- The activity is on time and the next six months priorities are: (1) improve and extend the iRODS-based data layer in support of the orchestrated FWS-Misfit workflow integration, (2) improve the operation of the X509-based computing resources of the research platform, (3) improve and extend the pervasive provenance system supporting W3C−PROV recommendations for extended interoperability, (4) extension of the data-intensive components in support of the FWS-Misfit workflow integration within the science gateway.

1.3 Research and development activities

The Research and Development activities draw on the selected data-intensive pilot applications to enable the transition from proof-of-concept demonstration to dependable research e-science environment in the seismology community. The main strategy is to:

- Analyse and adapt the software implementation of the pilot applications software to facilitate their adoption and sharing by a broader users community through reusable data and work flow components on the VERCE platform;
- Define and provide an application driven e-science environment providing a flexible hub between the seismology research-oriented world and the infrastructure-oriented world of data archives, HPC, Grid and Cloud computing.
- Maintain a balance between long-term sustainability considerations and fast science case implementation for external users' evaluation.

The main achievements in the last six months are:

- Enabling pilot applications and validation of the VERCE architecture (D-JRA1.2.3). This includes: new developments in the ObsPy library extending the support of seismology data and file formats, as well as a new scalable simulated data format for improving parallel I/Os and wave simulation codes performance; a prototype implementation of the complete full waveform inver-

sion, *i.e.* LASIF[20], with a SPECFEM3D interface, and provenance system, which is intended to steer the development on the E-Infrastructure; improved mapping of SPECFEM3D onto different DCIs architectures together with an improved input file generator; productisation of new processing elements for visualisation of wave propagation results; provision of new wave simulation application-software, *i.e.* SES3D-NT and SeisSol[21], to be supported by the platform.

- VERCE architecture and tools for data-intensive applications (D-JRA2.1.3). This includes: delivery of the dispel4py data-intensive framework; new provenance generation mechanisms and context dependent browser-based user-oriented interface supporting user visualisation and exploration of the provenance information that can trigger runtime data movement.
- The activity is on time and the next six months priorities are: (1) productisation and evaluation of the dispel4py components for the analysis of observed and simulated event waveforms to be integrated in the FWS-Misfit workflows; (2) evaluation of the performance of dispel4py onto different DCI architectures; (3) further integration of dispel4py within the science gateway and the data layer.

[20] www.lasif.net
[21] www.seissol.org

2 Milestones

Table 2 – VERCE milestones.

Number	Name	WP	Lead beneficiary	Delivery date	Achieved	Achievement date	Comments
MS1	M-NA1.1	1	CNRS	Month 6	Yes	25/05/2012	Approved
MS2	M-NA2.1	2	INGV	Month 6	Yes	25/05/2012	Approved
MS3	M-NA3.1	3	ULIV	Month 6	Yes	25/05/2012	Approved
MS4	M-NA1.1.1	4	CNRS	Month 6	Yes	31/03/2012	Approved
MS5	M-NA4.1.1	4	EMSC	Month 6	Yes	25/05/2012	Approved
MS6	M-SA1.1	5	CNRS	Month 6	Yes	25/05/2012	Approved
MS7	MSA3.1	7	KNMI	Month 6	Yes	25/05/2012	Approved
MS13	M-JRA2.1	9	UEDIN	Month 6	Yes	25/05/2012	Approved
MS8	M-NA1.1.2	1	CNRS	Month 12	Yes	01/10/2012	Approved
MS9	M-NA2.4	2, 4	INGV	Month 12	Yes	01/10/2012	Approved
MS10	M-SA2.1	5, 6	LRZ	Month 12	Yes	01/10/2012	Approved
MS11	M-SA3.2	7	KNMI	Month 12	Yes	01/10/2012	Approved
MS12	M-JRA1.1	2, 8	LMU	Month 12	Yes	01/10/2012	Approved
MS24	M-JRA2.2	9	UEDIN	Month 12	Yes	01/10/2012	Approved
MS14	M-NA1.2	1	CNRS	Month 18	Yes	01/04/2013	Approved
MS15	M-NA2.2	2, 3	INGV	Month 18	Yes	01/04/2013	Approved
MS16	M-NA3.4.1	3	ULIV	Month 18	Yes	01/04/2013	Approved
MS17	M-SA2.2	5, 6	LRZ	Month 18	Yes	01/04/2013	Approved
MS18	M-SA3.3	7	KNMI	Month 18	Yes	01/04/2013	Approved
MS19	M-JRA1.2	2, 6, 8	LMU	Month 18	Yes	01/04/2013	Approved
MS20	M-NA1.2.1	1	CNRS	Month 24	Yes	01/10/2013	Approved
MS21	M-SA2.3	5,6	LRZ	Month 24	Yes	01/10/2013	Approved
MS22	MS-SA3.4	7	KNMI	Month 24	Yes	01/10/2013	Approved
MS25	M-NA1.3	1	CNRS	Month 30	Yes	01/04/2014	Approved
MS26	M-NA3.4.2	3	ULIV	Month 30	Yes	01/04/2013	Approved
MS27	M-SA2.4	5, 6	LRZ	Month 30	Yes	01/04/2013	Approved
MS28	M-SA3.5	7	KNMI	Month 30	Yes	01/04/2014	Approved
MS29	M-JRA1.4	2,6,8	LMU	Month 30	Yes	01/04/2014	Approved
MS30	M-NA1.3.1	1	CNRS	Month 36	Yes	08/10/2014	On time
MS31	M-SA2.5	5,6	LRZ	Month 36	Yes	30/09/2014	On time
MS32	M-SA3.6	7	KNMI	Month 36	Yes	30/09/2014	On time
MS33	M-JRA1.5	8	LMU	Month 36	Yes	30/09/2014	On time

3 Deliverables

Table 3: VERCE deliverables.

Number	Name	WP	Lead	Nature	Dissemination	Delivery date	Achievement date	Status	Contractual
D2.1	D-NA2.1	2	INGV	Report	Public	Month 6	25/05/2012	Approved	Yes
D3.1	D-NA3.1	3	ULIV	Report	Public	Month 6	25/05/2012	Approved	Yes
D4.1	D-NA4.1	4	EMSC	Report	Public	Month 6	25/05/2012	Approved	Yes
D5.1	D-SA1.1	5	CNRS	Report	Public	Month 6	25/05/2012	Approved	Yes
D6.1	D-SA2.1	6	LRZ	Report	Public	Month 6	25/05/2012	Approved	Yes
D7.1	D-SA3.1	7	KNMI	Report	Public	Month 6	25/05/2012	Approved	Yes
D8.1	D-JRA1.1	8	LMU	Report	Public	Month 6	25/05/2012	Approved	Yes
D9.1	D-JRA2.1	9	UEDIN	Report	Public	Month 6	25/05/2012	Approved	Yes
D1.1.1	D-NA1.1.1	1	CNRS	Report	Public	Month 12	01/10/2012	Approved	Yes
D2.2	D-NA2.2	2	INGV	Report	Public	Month 12	01/10/2012	Approved	Yes
D2.2.1	D-NA2.2.1	2	INGV	Report	Public	Month 12	01/10/2012	Approved	Yes
D3.2	D-NA3.2	3	ULIV	Report	Public	Month 12	01/10/2012	Approved	Yes
D4.3	D-NA4.3	4	EMSC	Report	Public	Month 12	01/10/2012	Approved	Yes
D5.2	D-SA1.2	5	CNRS	Report	Public	Month 12	01/10/2012	Approved	Yes
D6.2	D-SA2.2	6	LRZ	Report	Public	Month 12	01/10/2012	Approved	Yes
D7.2	D-SA3.2	7	KNMI	Report	Public	Month 12	01/10/2012	Approved	Yes
D8.2.1	D-JRA1.2.1	8	LMU	Report	Public	Month 12	01/10/2012	Approved	Yes
D9.1.1	D-JRA2.1.1	9	UEDIN	Report	Public	Month 12	01/10/2012	Approved	Yes
D2.2.2	D-NA2.2.2	2	INGV	Report	Public	Month 18	01/04/2013	Approved	Yes
D5.2.1	D-SA1.2.1	5	CNRS	Report	Public	Month 18	01/04/2013	Approved	Yes
D6.2.1	D-SA2.2.1	6	LRZ	Report	Public	Month 18	01/04/2013	Approved	Yes
D7.2.1	D-SA3.2.1	7	KNMI	Report	Public	Month 18	01/04/2013	Approved	Yes
D1.2.1	D-NA1.2.1	5	CNRS	Report	Public	Month 24	01/10/2013	Approved	Yes
D2.3	D-NA2.3	2	INGV	Report	Public	Month 24	01/10/2013	Approved	Yes
D2.3.1	D-NA2.2.1	2	INGV	Report	Public	Month 24	01/10/2013	Approved	Yes
D3.2.1	D-NA3.2.1	3	ULIV	Report	Public	Month 24	01/10/2013	Approved	Yes
D4.3.1	D-NA4.3.1	4	EMSC	Report	Public	Month 24	01/10/2013	Approved	Yes
D5.3	D-SA1.3	1	CNRS	Report	Public	Month 24	01/10/2013	Approved	Yes
D6.3	D-SA2.3	6	LRZ	Report	Public	Month 24	01/10/2013	Approved	Yes
D6.3.0	D-SA2.3	6	LRZ	Report	Public	Month 24	01/10/2013	Approved	Yes
D7.3	D-SA3.3	7	KNMI	Report	Public	Month 24	01/10/2013	Approved	Yes
D8.2.2	D-JRA1.2.2	8	LMU	Report	Public	Month 24	01/10/2013	Approved	Yes
D9.1.2	D-JRA2.1.2	2	UEDIN	Report	Public	Month 24	01/10/2013	Approved	Yes
D2.3.2	D-NA2.3.2	2	INGV	Report	Public	Month 30	01/04/2014	Approved	Yes
D5.3.1	D-SA1.3.1	5	CNRS	Report	Public	Month 30	01/04/2014	Approved	Yes
D6.3.1	D-SA2.3.1	6	LRZ	Report	Public	Month 30	01/04/2014	Approved	Yes
D7.3.1	D-SA3.3.1	7	KNMI	Report	Public	Month 30	01/04/2014	Approved	Yes
D1.3.1	D-NA1.3.1	1	CNRS	Report	Public	Month 36	05/10/2014	Submitted	Yes
D2.4	D-NA2.4	2	INGV	Report	Public	Month 36	30/09/2014	Submitted	Yes

Continued on next page

Concluded from previous page

Number	Name	WP	Lead	Nature	Dissemination	Delivery date	Achievement date	Status	Contractual
D2.4.1	D-NA2.4.1	2	INGV	Report	Public	Month 36	30/09/2014	Submitted	Yes
D3.2.2	D-NA3.4	3	ULIV	Report	Public	Month 36	30/09/2014	Submitted	Yes
D4.3.2	D-NA4.3.2	4	EMSC	Report	Public	Month 36	30/09/2014	Submitted	Yes
D5.4	D-SA1.4	5	CNRS	Report	Public	Month 36	30/09/2014	Submitted	Yes
D6.4	D-SA2.4	6	LRZ	Report	Public	Month 36	30/09/2014	Submitted	Yes
D7.4	D-SA3.4	7	KNMI	Report	Public	Month 36	30/09/2014	Submitted	Yes
D8.2.3	D-JRA1.2.3	8	LMU	Report	Public	Month 36	30/09/2014	Submitted	Yes
D9.1.3	D-JRA2.1.3	9	UEDIN	Report	Public	Month 36	30/09/2014	Submitted	Yes

4 Project management during the period

4.1 Consortium, management structure and governance

During this period, there has been no change in the consortium. The project management office (PMO) has not changed, and its composition is: Jean-Pierre Vilotte (IPGP-CNRS) – Project Coordinator; Arthur Mulle (CNRS-INSU) – Project Manager; Antoine Weexsteen (CNRS-INSU) – Legal officer; Rosa Bernal-Carrera (IPGP); Geneviève Moguilny (CNRS-INSU).

Since the last review meeting, the Steering Committee (SC) has met on a three-week regular basis through remote meetings, while the Project Executive Board (PEB) is meeting on a two-week regular basis also through remote meetings. Beside these regular PEB meetings, smaller size meetings have been organised on particular topics, *e.g.* such as the iRODS data platform, the science gateway management and operation, data-intensive waveforms analysis and integration of `dispel4py`.

A short-lived task force has been initiated around the Misfit analysis implementation. It involves two computational seismologists, an expert developer of the `ObsPy` Python library, two `dispel4py` experts, a data expert, an expert in provenance and the developer of the interaction templates and portlets. After intensive preparation, it had a three-day meeting at which current systems were tried and ideas for improvements were hatched and critical steps evaluated. This initiated an intense development and collaboration for a period of two months, with daily communication between pairs attacking subgoals. A further face-to-face is planed to accelerate the integration of the components, towards exhaustive tests and refinement before production deployment. This task force has adopted agile development methods but has greater focus on teasing out what Misfit analysis components are becoming feasible as the VERCE digital context improves.

Monitoring of the project activities

Some partners faces man power issues. This situation has been closely monitored by the PMO to find a solution.

- WP4/NA4 work package: The EMSC correspondant has been away since August 2014 but should return January 2015.
- WP5/SA1 work package: Two persons recruited at IPGP and involved in JRA1, SA1 and SA2 left the project end of September 2014 and November 2014, respectively, for new job positions. At this time of the project, *i.e.* a year before the end of the project, the Steering Committee has agreed the best solution was to transfer part of their responsibility to the SCAI partner. This implies transferring 8 PMs from CNRS-IPGP to SCAI. This minor modification is being discussed with the project officer.

4.2 Internal communication

Overall the use of the VERCE Redmine online collaborative environment is now well adopted, and partners are quite familiar and at ease with it.

The Redmine wiki section is at the heart of the internal communication system, and is growing fast with minor time-to-time reorganisation to reduce the number of pages regrouping some of its content.

The NA1 (WP1) wiki section contains the minutes of the SC and PEB meetings as well as many of the transversal issues, *e.g.* collaboration with other projects – each project has a wiki page that the partners and the coordinator keep updated, reporting deadlines, procedures, templates, state of the deliverables, state of the art, risk management, sustainability strategy, *etc.*

Participation of partners at related events and relevant seminars, and the representation of VERCE in national and international contexts are tracked in an *ad hoc* wiki page in the NA4 section.

The Meeting section is where all information on on-going and past meetings, *i.e.* both within and across the WPs and at the project level, can be found. The Doodle tool for online polls has proved to be very useful, especially for the organisation of the various VERCE meetings.

The Document section is mainly updated by the PMO to share official documents or project internal templates. The Repository (SVN) is used for uploading/downloading the different versions of the deliverables, codes and other drafts that need a version tracking. Forums are not used frequently but allow keeping track of important on-going discussions. An *ad hoc* Issues Tracking System has been set up for NA3 (WP3) as part of the training strategy allowing to deal with direct user requests on technical questions regarding the VERCE tools and technologies.

A significant part of the internal communication still relies on emails, especially for communications from the PMO to the project's boards and for communication within WPs. The mailing lists system works well and helps targeting groups. Mailing lists are regularly updated by the PMO on the basis of a manpower file that lists all personnel working on VERCE and their roles and involvement in the different WPs.

External meetings and communication

In the last six months, the VERCE science gateway was presented and demonstrated in a number of public events, aiming at improving the visibility of the project within the international communities of e-science experts, seismologists, and therefore potential users. These events contribute to the visibility of the project and its services.

A full and continuously updated list of these events is kept on the Redmine by WP4. A selected list includes:

1. *VERCE and ENVRI meeting*, 23-25 April 2014, at KNMI (De Bilt) and UvA (Amsterdam), Netherland. Participants: M. Atkinson, I. Klampanos (UEDIN)
2. *SCI-BUS annual meeting*, 4-8 May 2014, CESME, Turkey. Participant: M. Atkinson
3. *EGU meeting*, 28 April -3 May 2014, Vienna, Austria. Participants: H. Scwichtenberg, A. Germünd (SCAI), G. Moguilny, J.-P. Vilotte (CNRS-IPGP), A. Michellini, E. Casarotti (INGV), A. Rietbrock, Th. Garth (ULIV), T. van Eck, A. Spinuso (KNMI), R. Bossu, C. Etivant-Dernoncour (EMSC), A. Frank (LRZ), 10 presentations, and three organised sessions.
4. *EGI Community Forum*, 19-23 May 2014, Helsinki University, Finland. Participants: Horst Schwichtenberg (SCAI), André Germünd (SCAI), Geneviève Moguilny, Visakh Muraleedharan (CNRS-IPGP), four presentations.
5. *IWSG14 International workshop on science gateways*, 3-4 June 2014,Trinity College, Dublin, Ireland. Participant: A. Spinuso (KNMI). one presentation.
6. *ISC'14 meeting*, 22-26 June 2014, Leipzig, Germany. Participants: G. Erbacci (CINECA). One presentation and a poster at the CINECA booth.
7. *General Assembly of the European Seismological Commission*, 24-29 August 2014, Istanbul, Turkey. Participants: Alessandro Spinuso and Torild van Eck (KNMI), one talk and poster.
8. *EGI conference on Challenges and Solutions for Big Data processing on Clouds*, 16-26 September 2014, Amsterdam, Netherland. Participants: Horst Schwichtenberg, André Germünd (SCAI), one talk.

9. *ISWG 14*, 19-23 October 2014, Trentino, Italy. Participant: A. Spinuso (UEDIN). One communication.

10. *SC'14 conference*, 16-21 November 2014, New Orleans, USA. Participants: R. Filgueira, M. Atkinson, A. Krause (UEDIN), M. David (CNRS-IPGP). one communication.

4.3 Cooperation with other projects

A first aspect of the VRE strategy is to create synergies and coordination with other projects in the seismology and the solid-Earth sciences.

- Collaboration between VERCE and EPOS[22] is of primary importance.
- VERCE builds on collaborations with international NPOs in seismology like IRIS[23] and Earthscope[24] in the USA, and JAMSTEC[25] and NIED[26] in Japan.
- VERCE developed synergies and coordination with a number of strategic FP7 projects in seismology, i.e., NERA[27], SHARE[28], REAKT[29], and TIDES.
- VERCE developed research-driven synergy and coordination with a number of European Research Council projects, i.e., WaveTomo[30], WHISPER, and more recently with the NERC-funded Terra-Correlator project[31].
- VERCE is contributing to the E-infrastructure Coordinated Research Action[32] in environmental sciences, launched by the Belmont Forum[33] within the framework of the Future Earth[34] initiative.

Another aspect of the strategy is to foster active collaborations with the European infrastructures, i.e. EGI, PRACE and EUDAT.

- VERCE contributes to two pilot case studies within the EGI-PRACE-EUDAT task force.
- VERCE collaborates with EUDAT with regard to iRODS based data management federation and data movement using `GridFTP`.

The last aspect of this strategy was the collaboration with other related projects that develop methodologies and software components, that were ultimately integrated within the VERCE platform.

- SCI-BUS: providing science gateway/portal technology to integrate access to computing, storage and other facilities and E-infrastructures. The VERCE Science Gateway is based on widely used production quality frameworks and solutions (`Liferay`, `WS-PGRADE/gUSE`). A Memorandum of Understanding (MoU) and a collaboration roadmap has been agreed between the two projects.
- IGE: providing tools to share computing resources, databases and other on-line tools. IGE, has now become the European Globus Community Forum[35] (EGCF) and is a member of the Globus

[22] epos-eu.org
[23] iris.edu
[24] earthscope.org
[25] jamstec.go.jp/e/
[26] bosai.go.jp/e/
[27] nera-eu.org
[28] share-eu.org
[29] reaktproject.eu
[30] ipgp.fr/en/sismo/erc-wavetomo
[31] gtr.rcuk.ac.uk/project/F8C52878-0385-42E1-820D-D0463968B3C0
[32] bfe-inf.org
[33] belmontforum.org
[34] icsu.org/future-earth
[35] egcf.eu

Alliance. Collaborations between VERCE and IGE have been formalised by a MoU. This collaboration has been important and has led to the adoption by VERCE of a number of components provided by IGE.

- ER-Flow[36]: a follow-on project to SHIWA, which developed technologies to allow interoperability between workflow systems. Collaboration between ER-flow and VERCE has been formalised by a MoU. The VERCE Science Gateway is now connected to the SHIWA workflow repository.
- Finally active collaboration between VERCE and the developer team of `ObsPy` has lead to a new version of the VERCE supported `ObsPy` library with the new Adaptable Seismic Data Format (ASDF) encapsulating HDF5 container formats, together with extended metadata for end-to-end data processing and analysis workflows, as well as the Large-scale Seismic Inversion Framework (LASIF) prototype.

This strategy is continuously reviewed and updated as the project evolves keeping track in order to collect experience and acquired know-how not only for the VERCE project but also for further projects of the solid Earth sciences.

4.4 Risk management plan

The Risk Management Plan is continuously actualised and refined. The Risk Management Plan and the monitoring strategy and tools were detailed in the Management and Progress Report D-NA1.1.1.

Each risk is assessed in terms of likelihood and impact, and is recorded in the Risks Register. This metric allows defining a Risk Exposure factor, which help to prioritise the different risks for management control. This dynamic document is continuously reviewed and updated throughout the project in interaction with the different work package leaders.

4.5 Sustainability elements

A sustainability strategy plan is under preparation through regular SC/PEB meetings. The VERCE sustainability strategy plan is linked to: (a) the sustainability of the European computing and data infrastructures, e.g. PRACE, EGI; (b) the implementation plan of EPOS that recognises the VERCE platform has a component of the seismology core services and a major contribution to the integrated core services; (c) the international framework of the seismology community, in particular with the FDSN, the COOPEUS[37] project, and the recent E-Infrastructure Coordinated Research Action of the Belmont Forum in environmental sciences.

4.6 Last review recommendations

The last review report made a number of explicit recommendations:

1. *The consortium should provide a revised work plan defining the goals to be achieved by the end of the project.*
   - This recommendation has been taken into account. A roadmap of the VERCE planned services, tools and E-Infrastructure has been set up and documented in a separate document.
2. *The VERCE architecture should be freeze and its final version thoroughly documented.*
   - This recommendation has been taken into account. The detailed roadmap of the VERCE E-Infrastructure up to the end of the project is provided in the separate document mentioned above.

[36] erflow.eu
[37] coopeus.eu

3. *External users should be invited to test both use cases and provide feedback. This feedback will ultimately decide whether the VERCE is successful or not. The list of the beta-testers and their feedback should be made available regularly.*

    - This recommendation has been taken into account. A first two-day online training session has been open to external users and held July 10th and 17th, providing an overview of the science gateway and of the enabled FWS science case, with hands-on practical sessions. Ten external users attended and provided feedback, which is being analysed by NA3, NA2 and SA3. A second two-day online training session open to external users will be held October 15th and 16th, providing an overview of the `dispel4py` framework for seismic waveform data-intensive analysis, with hands-on practicals. Nine external users are now registered and will provide feedback that will be analysed by NA3, NA2 and JRA2. The number of registered users in the science gateway grew to forty, while others are waiting to be approved. It is worth mentioning that more than half of the accounts belong to external seismologists interested in the use of the platform, others split between users coming from the VERCE consortium and technologists interested in exploring the platform's capabilities. Therefore the evaluation of the VRE and E-Infrastructure is well engaged now. Finally, it is worth to mention that prior review recommendations emphasised the priority to be given to the FWS science case for which the external research community is organised around well-established high performance simulation codes, *e.g.* such as SPECFEM3D. We followed this recommendation, keeping in our priority the data-intensive analysis science case, but reshaping it toward the research practice of the external user community.

4. *The two fully functional use cases (compute intensive –* SPECFEM3D*) and data-intensive (seismic noise correlation) should be finalised.*

    - As mentioned above, priority has been given in the last months to the forward waveform simulation science case, as recommended by the reviewers. This science case is well advanced now, in production and open to external users. Finalisation of the science case will involve data-intensive analysis of observed and simulated event waveforms for the evaluation of the misfit. The latter is being supported by the `dispel4py` framework and the `iRODS`-based federated platform. Data-intensive analysis science case is being refined to take into account the research practice of the external user community and to engage them in the evaluation of the VRE and E-Infrastructure. The VRE flexibility intends to enable researchers active researchers to invent and refine scalable methods for innovative statistical analysis of seismic waveforms in a wide range of applications contexts, *e.g.* detection and restoration of a variety of seismic sources, time dependent imaging of the Earth's interior properties. The VRE will support the development and the deployment of this diversity of methods and implementations, with the `dispel4py` data-intensive framework integrated with its platform and data-handling services. This enables seismologists to identify new targets and to develop new methods to achieve them, remaining in full control and using their familiar Python tools, *e.g.* such as the open-source community toolbox `ObsPy`, in their usual working and visualisation environment, *e.g.* their laptop. Once seismologists have refined their methods, they can exploit `dispel4py` to encode them as abstract streaming data workflows, that can be automatically mapped onto HPC, data-intensive clusters or Clouds, to maximise reliability and scalability. The `dispel4py` supports fine-grained, user-controlled provenance tracking to help seismologists analyse, understand, improve and steer their data-intensive work, based on the same standards, tools and interaction mechanisms used for FWS, so that they fit well together. `dispel4py` is smoothing the path from theoretical research through proof of concept to sustainable use and sharing. The roadmap toward the computational and data-intensive frameworks is explained in more details in the roadmap document provided as a separate document.

5. *Following the completion of beta testing by external users, the consortium should prepare videos documentation for advertising the system to the wider scientific community.*

   - This recommendation has been taken into account. At this stage, a first step has been the production and the dissemination of a number of training videos, related to the two online training sessions. It is worth to mention here, that in the last months, the science gateway achievements have been presented and demonstrated in a number of public events. It has been also presented and demonstrated at a number of meetings of the EPOS project. This led EPOS to positively evaluate the VERCE achievements and to build on VERCE for its implementation phase starting 2015.

6. *The consortium should clarify the users eligible for using the VERCE platform, type of credential needed and how to be obtained. Possibly for users to try VERCE should be foreseen.*

   - This recommendation has been taken into account and is being addressed. Improved integration and operation of certificates, making authorisation in VERCE easier has been addressed. Management of the VERCE VO has been clarified and improved among the different work packages. The next steps are evaluation, by the same team on larger problems or using different DCI, by other VERCE researchers, or by external researchers. This we identify as the evaluation phase and it can take some time as new issues are detected, *e.g.* such as the facilities offered to seismologists, the way they are presented, interactions with other parts of the VERCE E-Infrastructure or scalability. This evaluation phase also engages researchers external to VERCE in synergy with EPOS. The availability of HPC resources is limited. Gaining access to HPC resources on PRACE is complex. However, the new call system of PRACE starts to recognise scientific communities. This may improve on these restrictions. The situation is slightly easier with EGI Cloud and Grid resources. Part of these issues is being addressed in the SA1 and SA3 reports.

7. *A continuity of services plan including appropriate procedures and deployment topologies should be drafted.*

   - This recommendation has been taken into account and is being addressed by SA1, SA3 and NA3. Progress can be found in the SA1 and SA3 reports, especially with regards to the science gateway management and continuity of services. A number of documentation, *e.g.* platform deployment, user and site customisation, have been finalised and made available in the Redmine platform. A continuity of services plan is being drafted by SA1 and SA3 and will be available for the next review meeting. Another component here is the on going work within EPOS related to the Integrated Core Services management and operation, during its implementation phase starting in 2015, and in which the VERCE VRE and E-Infrastructure will is included.

4.7 Next steps

During the next reporting period, the PMO will continue monitoring and update the Risk Register. It will also continue to focus on the drafting of a Sustainability Plan in a moving landscape. A draft of the sustainability plan will be presented in the next review meeting. VERCE is closely following and contributing to the RDA[38] activities for Open-Data and was also invited to contribute to the Belmont initiative on E-Infrastructure in environmental sciences, that includes sustainability issues and funding by the national research agencies.

The VERCE PMO will keep coordinating and monitoring the partners' contributions and project activities making sure they are coherent with the DoW while at the same time guaranteeing the needed flexibility to adapt to research developments, changing needs of the community, and collaboration with

[38] rd-alliance.org

other projects. The PMO will make sure all deliverables and milestones are ready for the next deadline on 1 April 2015.