Management and Progress Report At PM 42

27/04/2015

Project acronym:	VERCE
Project n°:	283543
Funding Scheme:	Combination of CP & CSA
Call Identifier:	FP7-INFRASTRUCTURES-2011-2
WP:	WP1/NA1, Management
Filename:	D-NA1.3.2.pdf
Author(s):	A. Mulle and A. Weexsteen (CNRS), J.P. Vilotte (CNRS-IPGP)
Location:	http://www.verce.eu/Repository/Deliverables/RP4/
Type of document:	EU intermediate Report
Dissemination level:	Public
Status:	Final
Due date of delivery:	30/03/2015

Version	Author	Date	Comments
1	A. Mulle	08/03/2015	Initial draft for comments
2	J.P. Vilotte	10/03/2015	Project's objectives, Work Progress
3	A. Weexsteen	12/03/2015	Review and additional editing
3	J.P. Vilotte	15/03/2015	Review and additional editing
4	J.P. Vilotte	20/03/2015	Additional editing
5	H. Schwichtenberg & M. Galea	27/03/2015	Review and additional editing
6	M. Atkinson	28/03/2015	Review and additional editing
5	J.P. Vilotte	02/04/2015	Final review

Copyright notice

VERCE ("Virtual Earthquake and seismology Research Community e-science environment in Europe") is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 7th Framework Programme. VERCE began in October 2011 and will run for 4 years.

Table of Contents

Declaration by the scientific representative of the project coordinator 4
Publishable summary 6
1. Project objectives and work progress for the period 13
1.1. Network activities 13
1.2. Service activities 15
1.3. Research and development activities 16
1.4. VERCE roadmap update 18
1.5. Milestones and Deliverables 20
2. Project management during the period 27
2.1. VERCE consortium: management, structure and governance 27
2.2. Internal communication 27
2.3. Internal meetings 27
2.4. Training sessions and webinars 27
2.5. External meetings 28
2.6. Cooperation with other projects 28
2.7. Project activities monitoring 29
2.8. Risk management plan 29
2.9. Last review recommendations 30
3. Elements for a VERCE sustainability plane and strategy 32
3.1. The role of VERCE 32
3.2. Quantitative needs for a sustainable VERCE 34
3.3. Elements of the VERCE sustainability strategy 35
3.4. Training and dissemination outreach 36
4. Next steps 36

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: **1st** **2nd** **3rd** **4th** X

Period covered: **from 1st October 2014 to 31st March 2015**

Name, title and organisation of the scientific representative of the project's coordinator[1]:

Prof. 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 address: http://www.verce.eu/

[1] Usually the contact person of the coordinator as specified in Art. 8.1. of the Grant Agreement.

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 attached periodic report represents an accurate description of the work carried out in this project for this reporting period;
- The project (tick as appropriate) [2]:

  □ has fully achieved its objectives and technical goals for the period;

  X 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 at all on schedule.

- The public website, if applicable

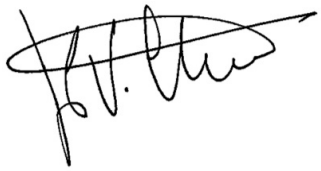
  X is up to date

  ☐ 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 scientific representative of the Coordinator: **Jean-Pierre Vilotte**

Date: 5/04/2015



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

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

Publishable summary

Context and objectives

Seismology addresses both fundamental problems in understanding the Earth's internal wave sources and structures and augmented societal applications; and puts 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 VERCE strategy is to provide an e-science environment and services driven by the needs of the earthquake and seismology research community, delivering:

- Computational seismology services for simulating seismic wave-field generated by earthquake events and for predicting associated seismic waveforms and ground motion;
- A toolkit for seismologists analysing data generated by observational systems and simulations, including both event waveforms analysis, i.e., comparison between observed and simulated event waveforms, and continuous waveform analysis, i.e., such as seismic noise correlation and seismic source detection;
- An extensive collection of data handling services, components and algorithms that are used at different stages in the above contributions and which are suitable for use in Earth sciences and other domains in many cases;
- An e-Infrastructure framework that provides seismologists with an integrated means of using all these services, tools and algorithms, and that will facilitate further advances in data-intensive computational seismology, i.e., simulation and data analysis; and accommodates users with different levels of expertise.

Important efforts are devoted to training sessions and "intellectual ramps" providing safe and supported means for seismology researchers to engage incrementally with the methods and the tools of data-intensive research; and the use of the VERCE e-science environment. This includes a number of training sessions oriented toward external users of the seismology research community, and promotion of the VERCE Science Gateway through communication at IT and Earth sciences conferences. Feedback from a selected number of external users are analysed and have guided the evolution and the improvements of the Science Gateway tools and services.

Following the third review meeting of the VERCE project, held on the 6th June 2014, a number of recommendations were made for implementation and finalisation. A roadmap of planned services, tools and e-infrastructures has been approved by the steering committee. This was provided at the previous reporting period and referred hereby as the VERCE roadmap.

Work carried during the last reporting (1st October 2014 – 31st March 2015)

Pilot applications and use cases

Taking into account the review recommendations, priority has been given to:

- Finalise the integration, and extend the functionalities of the Forward Wave Modelling (FWM) tools and services within the VERCE Science Gateway;
- Finalise a toolkit for Data-intensive analysis (DIA) of data retrieved from observations and simulations, i.e., including event wave form analysis for HPC wave simulation, and continuous wave form analysis for seismic noise and event detection analysis;
- Training sessions and webinars open to external users of the seismology research community, together with training material and documentation.

The main achievements are:

- Forward Wave Modelling: new features and extended flexibility of the FWM tools have been integrated based on users' feedback recorded mainly in particular during training events. This includes: multi-

purpose workflows and extended provenance tools accommodating users with different level of expertise; integration of new data FDSN compliant web services exposing earthquake parameters; improved management and customisation tools of models and meshes allowing the use of user-customised ones in addition to those stored in the shared library of earth models and meshes; GUI UX interfaces translating the file-centric user interface of SPECFEM3D in the structure of the gateway; new visualisation scripts of wave-field propagation and seismograms including Google Earth KML file generation. New processing elements (PEs) have been defined including: creation of input files based upon user selection of seismic stations, seismic source events and simulation parameters; ObsPy and *dispel4py* library tools for misfit analysis between observed and simulated waveforms.

- Data-intensive analysis: the strategy has been refined based on research practice analysis and users' feedback. This includes: an extensive data-intensive analysis toolkit, *dispel4py*, tailored toward event waveform and continuous waveforms stream analysis; a suite of test cases and examples for the evaluation of the scalability and the performance of *dispel4py* functionalities including Cloud DCIs and MPI mappings; a dedicated taskforce engaged, through regular meetings and code sprints, into the implementation of the misfit calculation use case integrating key DIA components in line with the VERCE roadmap.
- Training material and user documentation: new training material and user documentation have been provided for the intense training session activities providing an overview of the VERCE platform and of the Science Gateway services and tools, i.e. specifically FWM - based on SPECFEM3D - and DIA – based on *dispel4py* – tools and their practical user functionalities through the Science Gateway. All the videos on the training webinars are available on the VERCE website[3].
- Analysis of external users' feedback recorded at the training sessions: this led to recommendations for improving and extending flexibility of the Science Gateway services and tools; and delivering a smooth path between different levels of expertise and research practice especially for FWM.

Training and user documentation

In line with previous review recommendations, the priority was given to providing user-oriented "intellectual ramps", through training sessions, videos and user documentation allowing seismology researchers to engage incrementally with the VERCE platform of tools and services in phase with their research practice and expertise.

The main achievements are:

- A number of webinars and training sessions providing means for seismology researchers to engage with the VERCE tools and services. This includes: two-day online training session July 10th & 17th, 2014, introducing the VERCE platform and the Scientific Gateway FWM tools with practical and demos; two-day online training session October 15th & 16th, 2014, introducing the VERCE tools, specifically *dispel4py* library and framework, for data-intensive research with practical allowing attendees to develop their own data processing and analysis workflows and evaluate mappings onto different computing architectures; three-day face-to-face training session March 9th-11th, 2015, in Munich, hosted by LRZ, covering the VERCE platform and Scientific Gateway functionalities, the FWM and *dispel4py* tools, novel paradigms including validation of results using metadata and provenance exploration, access to HPC and Grid computing resources, VERCE data management resources through a number of practical exercises.
- A questionnaire was designed for these training events, allowing recording and analysis of feedback and impressions on how the VERCE services and tools will impact on their research practices and common needs. The overall feedback was extremely positive, and the style of training well received. Most participants recognised the potential offered by the VERCE FWM tools and services in support of their own research or the one they intend to undertake. Participants showed definite eagerness for more models and meshes promoting the importance of the shared library of earth models and meshes. In addition, they positively evaluated the extended tools allowing validation, re-parameterisation and use of their private customised models and meshes, with the support of the VERCE provenance system and data management layer with authorisation-based access control. The data-intensive analysis library, *dispel4py*, was well received by the participants that showed interest in adopting this development framework for developing their own research data analysis workflows, metadata and provenance

[3] http://www.verce.eu/Training/UseVERCE.php

management. Obtaining and using certificates was recognised as the main hurdle for users approaching for the first time production distributed computing infrastructures with different security and access policies. Even though the support provided by VERCE in minimising such hurdles and smoothing the path of scientists in the adoption and use of their own certificates was positively evaluated, these security hurdles are the main impediment to rapid uptake and need to be addressed beyond VERCE.

- New upcoming events include: (i) training sessions June 1st-5th, 2015, solicited by the TIDES-COST project in Bertinoro (Italy), aiming at educating young seismologists about the latest generation of tools for data analysis and modelling; (ii) data-intensive and computational modelling training material solicited by the international summer school of the NSF-funded Open-Science Data Cloud PIRE project, June 8th-12th, 2015, in the University of Amsterdam; and (iii) the final VERCE face-to-face training event, July 1st-3rd, 2015, in the University of Liverpool, in coordination with the H2020 EPOS-IP ESFRI project.
- Training videos and material are available on the VERCE website[4], and are regularly extended and updated. A range of user documentation has been produced and made also available. Professional videos are in preparation at the University of Liverpool and shall be finalised for the final training event in July 2015.

Dissemination and Public Outreach

During the reporting period the communication tools and channels were continuously updated and monitored. The main achievements are:

- The VERCE Science Gateway has been demonstrated and its awareness disseminated to multi-disciplinary audience during a number of international meetings, i.e., the AGU fall meeting (December 15th-19th, 2014, San Francisco), The e-Infrastructures for Earth Sciences workshop with GEANT, PRACE, EGI, EUDAT (January 22nd -23rd, 2015, Amsterdam); the EGU meeting (April 12th-17th, 2015, Vienna), the EGI Forum (May 18th-22th, 2015, Lisbon)

Management and operation of the research platform

During this period, integration of platform components, such as data management, DCI middleware and security mechanisms have been improved and consolidated, together with the integration of new resources and the support of new developments in line with the VERCE roadmap. Efforts were devoted to continuous upgrade, in response to feedback from daily operation and training sessions, of the VERCE tools and services, and to the continuity of services for: accessing and managing the VERCE platform data and computing resources, integrating and evaluating the new releases of the VERCE platform.

The main achievements during this period are:

- Integration of new resources within the VERCE Scientific Gateway and its abstract DCI Bridge layer handling X.509 Eurogrid PMA certificates: (I) HPC-CINECA-04 providing external access - based on UNICORE 7 - to the new GALILEO tier-1 HPC system at CINECA, which has replaced the decommissioned PLX system (HPC-CINECA-02); (ii) DEP-SCAI-01 providing external access – via the Globus middleware - to a new institutional local large data-intensive parallel system at SCAI Fraunhofer, where *displel4py* and ObsPy are available; (iii) DEP-INGV-01, providing external access to INGV data and compute resources, is under validation tests after some delay due to local policies; European Grid Infrastructure resources, via the VERCE Virtual Organization and the gLite middleware – is under progress.
- Higher performance and reliability of the VERCE data platform and services, based on federated iRODS servers, a scientific metadata catalogue (an instance of MongoDB) and a provenance web service to support workflow tools and application data provisioning through: restructuration of the Cloud storage infrastructure at UEDIN now supporting the GridFTP iRODS DSI through GlusterFS; revision and performance improvement of data replication policies, waiting for other solutions from the

[4] http://www.verce.eu/Training/UseVERCE

EUDAT project, based on "resource groups" - of local and remote iRODS resources – and leveraging iRODs' micro services to achieve transparent local files replication to remote resources within the same group.

- Improvement of the operation, scalability and reliability of the VERCE Science Gateway, and of its hosted auxiliary services, via two Linux Kernel-based virtual machines (KVM) running in an OpenNebula Cloud environment at SCAI: hosting the portal frontend (based on WPS-PGRADE and Liferay) and the backend (gUSE) both running as Java processes within an Apache Tomcat container; particular emphasis has been placed on compatibility with certificates using SHA2 mechanisms for the gUSE backend services; backup strategy automatically generates snapshots of the full virtual machines via the OpenNebula API.
- Portal support is currently hosted at IPGP and uses OTRS for first level support (account, security issue, service availability) and related issues with job submission (software request).
- The Inca grid-resource, hosted at LRZ on a virtual machine, monitors the integrated resources of the VERCE platform and provides http access to the outside for resources status.
- New GCMT FDSN web services for querying Global Centroid Moment Tensor catalogue data, hosted now at KNMI via a FDSN server implementation, are now integrated and supported within the VERCE Science Gateway.
- Training sessions and end-user fully functional support through: IGTF compliant X.509 training certificates, VERCE portal and iRODS accounting and GSI authentication, configuration iRODS folder layout and access rights.

Integration and evaluation of the platform services

This period corresponds to the sixth completed Plan-Do-Check-Act (PDCA) cycle. Release schedule and recommended work practices have been documented.

The main achievements are:

- The sixth PDCA cycle was completed January 31st, 2015 and five components were evaluated and approved: SPECFEM3D 2.0.2, requiring a change in the format of the configuration file and of the job batch script; *dispel4py*, a VERCE development from UEDIN for data-intensive applications; *h5Py*, that both ObsPy and *dispel4py* can take advantage of; ParaView, for some visualisation components of the VERCE Science Gateway, and which support Python and mpi4Py; FFmpeg, for parallel (*mpi4Py*) videos generation via the VERCE Science Gateway.
- The previous fifth PDCA cycle, completed July 2014, already released: *ObsPy* 0.9.2; *MPI* for Python 1.3.1, i.e., an important *dispel4py* component for mappings on both HPC and Grid resources; Network X 1.8.1, both on HPC and Grid resources with additional *dispel4py* scripts; and iRODS 3.3.1, i.e., an important component of the VERCE data layer.
- The next and last six months schedule, April 1st – September 30th, 2015 is now defined according to the VERCE roadmap. Requirements of tools and services to be evaluated will be collected through the Request Form. Significant efforts will be deployed for enabling the misfit data-intensive analysis component of the FWM use case.

Scientific gateway and user interfaces

During this period, efforts have focused on: training and external evaluation of the Science Gateway functionalities; rolling out new functionalities of the Science Gateway in line with the VERCE roadmap and in interaction with users' evaluation feedback; addition of new data services and implementation of new multi-purpose workflows; integration of advanced tools for data-intensive analysis (*dispel4py*); improving performance and robustness of the Science Gateway services and tools.

The main achievements during this period are:

- The Science Gateway has been upgraded to the latest release of the gUSE framework, and is operated now in the OpenNebula Cloud environment hosted in SCAI. This upgrade was crucial to improve the security, management and robustness of the gateway, which now support the SHA2 encryption

algorithm adopted by the newly issued grid-certificates, fostering the integration within the available DCIs.

- Extended functionalities of the VERCE Scientific Gateway, in line with the VERCE roadmap and in response to users' feedback, supporting a more flexible usage and a variety of FWM scenarios in complete control of the researchers and allowing: custom models which can be developed, re-parameterised, modified and used in the gateway in addition to the ones offered by the VERCE shared library of Earth models and meshes, taking advantage of the VERCE provenance system and data management layer offering a private space with authorisation-based control; integration of new data sources with in particular the Global Centroid Moment Tensor (GCMT) FDSN compliant web service exposing the GCMT parameters in QuakeML; refactoring of the FWM workflows, which now offer runtime data and metadata updates, and adoption of advanced tools for data processing (*dispel4py*); multi-purpose workflows supporting users with different levels of expertise, in particular by enabling the interactive production and gathering within the gateway of input files, models, meshes and event data required by FWM scenarios, which can then be filtered and shipped to targeted computing resources that are already or not yet integrated in the VERCE platform; extended functionalities of the GUI FWM setup interface supporting FDSN and map services; improved solver configuration allowing solver, mesh-and-earth models selection from a shared library for less advanced users; extended provenance information and tools supporting users in the analysis of their experiments with metadata auto-composition and improved portability and interoperability (W3C-PROV-JSON representation) with other provenance tools enabling hand-over of a final research product to institutional or multidisciplinary archives.
- The *dispel4py* technology has been adopted and integrated for critical data processing parts, i.e., post-processing pipelines of raw simulated data to produce images, plots and other products, bringing performance improvement thanks to the *displel4py* capability of mapping pipelines and workflow graphs to parallel computing resources.
- The VERCE Science Gateway already has nearly 100 registered users - lecturers, researchers, MSc and PhD students in Earth sciences and also, in a few cases, oil exploration – coming from nearly everywhere in Europe while others are waiting to be accepted as new features will be rolled out: the most active users are those who attended one of the training or webinar sessions.
- Users' feedback and interactions have been very positive recognising: the potential offered by the VERCE Science Gateway for their current research practice and for tackling new challenging problems; the importance of a shared library of earth models and meshes providing safe and supported means for seismology researchers to engage incrementally in the learning process to produce new ones and possibly contribute to expand that library; the interest in adopting the VERCE tools, specifically *dispel4py*, for data intensive research; the support provided by VERCE in smoothing the path of scientists over the many different policies, certification, and validation of the e-infrastructure providers, which is the main hurdle and impediment to rapid uptake and for many of these issues outside the project's control and that will be receiving attention from new Horizon 2020 projects and others.
- In the next months, priority will be given to the full implementation of the misfit analysis component in FWM workflows, with further integration of the *dispel4py* data analysis functionalities: a dedicated task force has been established and participants are engaged in regular meetings and code sprints – Rome, November 2014, Edinburgh, April 2015 - aiming at finalising its implementation as indicated in the VERCE roadmap.

Harnessing intensive applications

Taking into account the reviewers' recommendations, important efforts were devoted to the dense schedule of training and webinar activities, together with further flexibility improvements and extensions of the FWM and DIA tools in line with the VERCE roadmap and in interaction with users' feedback.

The main achievements during this period are:

- Extension of the mapping of the FWM simulations to run on new resources of the VERCE platform, HPC (CINECA) and DCIs (SCAI and INGV);
- Extension of the Earth models and meshes library tools and services in support of the FWM use cases, now allowing users to re-parameterise and change meshes when testing them, taking advantage of the

VERCE provenance system and data layer management tools and services their own models together with a control quality checking procedure, and possibly contribute to the extension of the shared library.

- Extension of the data analysis tools for evaluating simulated event waveforms against real observations, making use of the ObsPy and *dispel4py* libraries; definition of new synthetic waveforms file format and seismic data provenance complying with the W3C PROC model.
- Improved data staging and data movements using the iRODS technology, across the distributed data storage resources of the VERCE data layer, enabling users to produce interactively the input files required by the simulation solver, gathering and organising event data from FDSN services, and then filtering and shipping those to target resources hosting specific solver installation, thanks to the gateway tools that exploit the provenance recordings.
- Extension of the library of supported wave simulation solvers beside SPECFEM3D, i.e., SES3D-NT and SeisSol in support of the forward wave simulation use case.
- Support and training material for the dense set of training and webinar activities that took place July and October, 2014, March, 2015, and will take place in June and July, 2015.
- The next steps are to finalise the integration of the post processing data analysis, a challenge addressed by a new task force involving meetings and code sprints, i.e., misfit calculation within the Science Gateway and multi-purpose workflows according to the VERCE roadmap, making use the *dispel4py* functionalities; further development and provision of the ObsPy Python library; seismological examples of data-intensive analysis of continuous waveforms using *dispel4py;* support of solvers and tools on additional computing resources recently integrated or being integrated to the VERCE Science Gateway.

Architecture and platform tools for data analysis and modelling

A new collaboration has been setup with the NERC-funded TerraCorrelator project through the UEDIN and ULIV partners of VERCE around data-intensive seismic analysis. During the last period, collaboration with SCI-BUS and ER-Flow has been very fruitful and led to the connection of the VERCE Science Gateway with the SHIWA workflow repository. The main activities were in line with the VERCE roadmap, and significant effort was devoted to the training and webinar sessions together with the analysis of the users' feedback leading to new extensions and flexibility of the VERCE tools and services.

The main achievements during this period are:

- Production-ready SCI-BUS-powered VERCE Science Gateway, which is now connected to the SHIWA workflow repository.
- Integration of the VERCE architecture on the OpenNebula Cloud system hosted at SCAI, as part of the EGI FedCloud system, including upgrading to SCI-BUS 3.6.6 and EGI SHA2 user certification.
- Improved integration and operation of certificates minimising the identification, authentication and authorisation hurdles in order to smooth the path for scientists when they first approach e-infrastructure such as VERCE that composes many different authorisation policies. This authorization hurdle has been identified from users' feedback as the main impediment to rapid uptake: many of the issues are outside the project's control and are receiving attention with new Horizon 2020 projects and others such as EGI-Engage and EU-DATA2020 to which VERCE is already providing feedback through the participation in their workshops.
- Extension of the *dispel4py* system enabling distributed data-intensive fine-grained abstract data stream workflows, allowing seismologists to construct graphs of connected processes – for example ObsPy functions or Python scripts – called "*Processing Elements*" (PE) to transform, analyse and handle data streams and collection of data from files on storage services, and supporting their mapping on several computing architectures, including shared memory and multicore MPI architectures, and Cloud environments: *dispel4py* is being integrated with the VERCE Science Gateway to support the data analysis stages of the FWM workflows.
- Extended metadata collection and provenance metadata recording the relationships between data according to the W3C-PROV2 standards during the different data transformation stages within end-to-end workflows combining FWM and data analysis using internal identifier (PID) and longer-term persistent identification under seismologists' request calling European PID services or mechanisms established by EUDAT. The current implementation includes: (i) the provenance generation mechanisms; (ii) browser-based interfaces and tools built on top of a NoSQL storage technology

enabling combined operations for accessing and downloading data, which may be selectively stored at runtime, into dedicated archives. Efforts also involve improving the gUSE/WPS-PGRADE workflow with better integration with *dispel4py* system;

- Improvement of the iRODS and Provenance services that are currently hosted on the Data Intensive Research Cluster (EDIM1) at EPCC containing: the iRODS server; a webserver running the iRODS web frontend; a GridFTP server with the GRIdFTP iRODS DSI; an instance of MongoDB storing the scientific metadata associated with data stored in iRODS, and the provenance web service.
- Development of *dispel4py* tutorial library of examples and practical for training and webinar sessions (Edinburgh, October 2014; Munich, March 2015; Amsterdam, July 2015, Liverpool, July 2015), together with running dedicated webinar sessions.
- The next steps will be to fully support the *dispel4py* integration for data-analysis pipeline components of FWM workflows, i.e., the misfit calculation stages, for which a dedicated task force has been established within the project with regular meetings and code sprints (Rome, November 2014; Edinburgh, April 2015), case and the data-intensive continuous waveform analysis, i.e., seismic noise correlation analysis and seismic source detection; extending families of *dispel4py*-enabled data analysis PEs and tutorial examples used during training and webinar sessions.

Impact of the VERCE project

The overall users' feedback was extremely positive, recognising the potential offered by the VERCE services and tools in support of their own research or the one they intend to undertake. VERCE lay the basis for transformative development in data use and modelling of the seismology research community in Europe, integrating a broad set of innovations which are now endorsed by the new EPOS-IP project funded this year under the H2020-INFRADEV-1-2015-1 call, for thematic and integrated core services, in particular under the WP6 "ICS-TCS integration and interoperability", the WP7 "ICS design and development", and the WP8 "TCS-Seismology" work packages. It will contribute key elements concerning computation and data to the EPOS-IP computation work package.

Website: http://www.verce.eu

1. Project objectives and work progress for the period

This document, i.e., D-NA1.3.2, is the second intermediate management and progress report covering the 1st October 2014 – 31st March 2015 period of the fourth reporting period.

VERCE is structured into nine work packages (WPs): Management activities (NA1/WP1), Network activities (NA2/WP2 to NA4/WP4), Service activities (SA1/WP5 to SA3/WP7), and Research & Development activities (JRA1/WP8 and JRA2/WP9).

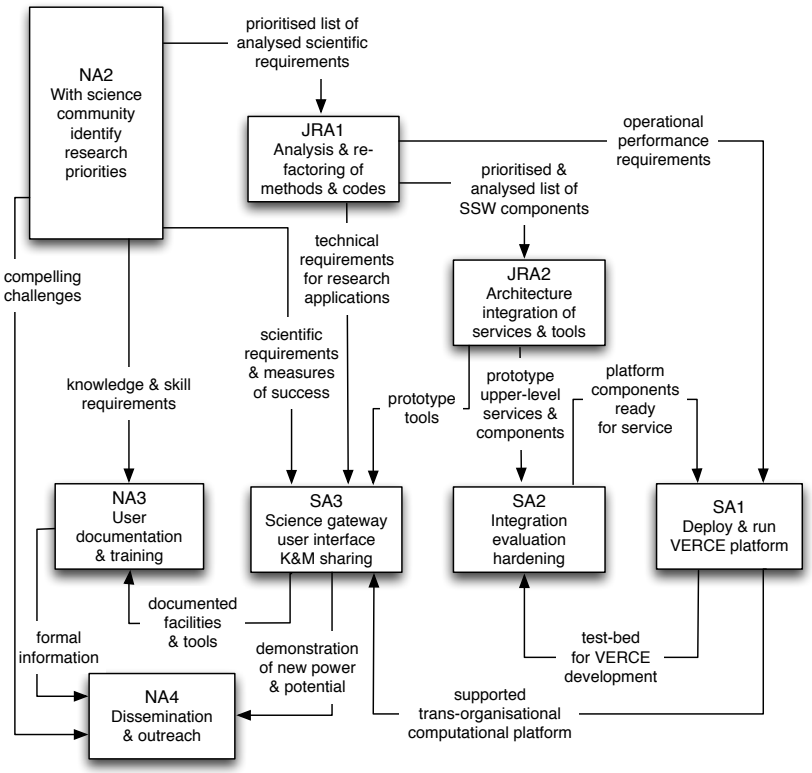


Figure 1 - Work packages in VERCE

1.1. Network activities

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

- Smooth the path from theoretical research 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 training and "intellectual ramps" providing safe and supported means for seismology researchers to engage incrementally with VERCE tools and services for their research practice.

The main deliverables and achievements in the last six months are in line with the VERCE roadmap:

- Application use cases on the VERCE platform, i.e., examples and demonstrators (M-NA2.3), in coordination with SA3, JRA1, SA1 and SA2: this includes a number of improvements and increased flexibility of the FWM use case, i.e., custom model upload, refinement and validation, FWM derivatives and multi-purpose workflows, together with the evaluation of the VERCE toolkit for data-intensive processing and analysis of continuous and event waveforms based on the ObsPy and *dispel4py* libraries.
- Updated demonstrators for dissemination and training, in coordination with NA3, SA3, this includes: FWM demonstrators using the VERCE Science Gateway, Data-intensive analysis and processing training use cases for *dispel4py*; user documentation, training material and videos for the webinar and training sessions (July, 2014; October, 2014; March, 2015) open to external users of the seismology research community.
- The NA2 activity is on time according to the VERCE roadmap and in phase with the users' feedback received in particular during the training sessions program. The next six months roadmap and priorities are: (1) the next July training session in Liverpool including FWM and Data-intensive analysis (*dispel4py*) for continuous evaluation and feedback from external seismology researchers and students; (2) extension of the FWM workflows with the implementation of the misfit calculation components, involving joint data analysis of the simulated and observed event waveforms, and making use of the *dispel4py* library and functionalities (data streaming and DCI mappings). For the later, a dedicated task force has been established, together with SA3 and JRA1, with regular meetings and code sprints (Rome, November 2014; Edinburgh, April 2015).
- Training sessions and evaluation of users' feedback (e.g., D-SA3.4.1) in coordination between NA3, NA2, SA3, JRA1. This includes both FWM and *dispel4py* use cases. User documentation has been made available, including installation guide of the gUSE framework within the VERCE Science Gateway; Customisation of the VERCE portal, Configuration of GT5 resources in the portal, installation, deployment and use guide for the FWM portlet.
- The NA3 training activity has been quite intense and is on time according to the VERCE roadmap. This webinar and training sessions agenda is summarised in the following table:

Training sessions and workshops				
VERCE FWM training webinars	**July 10th & 17th, 2014** 10 registered users	PhDs and Post-Docs Science Gateway and FWM tools	Open to external users	September 2014 -Feed back evaluation
VERCE DI analysis webinars	**October 15th & 16th, 2014** 9 registered users	PhDs and Post-Docs *dispel4py* library for seismic waveform processing and analysis	Open to external users	December 2014 – feed back evaluation
VERCE training workshop (Munich, Germany)	**March 9th-11th, 2015** 20 registered users	MSc students, PhDs and Post-Docs VERCE e-science environment, FWM tools, and dispel4py library	Open to external users	April 2015 – feedback evaluation
TIDES-VERCE Training workshop (Bertinoro, Italy) *Solicited by the TIDES-COST project (EU-funded)*	**June 1st-5th, 2015** ~20-25 participants	Researchers, Post-Docs, PhDs and MSc students VERCE platform and Science Gateway, FWM and Data-intensive analysis tools	Open to the European computational seismology community	September 2015 – joint feedback evaluation between VERCE and TIDES projects
PIRE-VERCE Training and outreach sessions (Amsterdam & Edinburgh)	**June 8th-12th, 2015**	Researchers, PhDs and Post-Docs Data-intensive and computational modelling e-environment and methodologies	Open to users and external scientific communities Will be followed by six-weeks practical in the university of Edinburgh for several students	September 2015 – joint feedback evaluation between VERCE and PIRE projects.

Solicited by the Open-Science Data Cloud PIRE project (NSF-funded)				
VERCE training workshop (Liverpool, UK)	**July 1st-3rd, 2015** ~25 participants	Researchers, PhDs and Post-Docs VERCE platform and Science Gateway; FWM tools including first misfit calculation implementation, Data-intensive analysis, *dispel4py,* toolkit	Open to a wide seismology community from European ERC projects, EPOS-IP ESFRI project and infrastructure projects/providers	September 2015 – feedback evaluation

1.2. Service activities

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

- Evaluate-and-integrate a platform of tools, services and application software components for enabling science applications and platform operation, including software components already adopted by the existing European e-Infrastructure providers (e.g., EGI, PRACE, IGE) or supported by EU related projects (e.g., EUDAT, SCI-BUS, ER-flow).
- Manage the release process of the research platform through a Plan-Do-Check-Act (PDCA) cycle.
- Deploy-and-operate the successive versions of the platform, providing access across a set of public and private data-and-compute resources, and a continuity of services.
- Manage and operate the VERCE Virtual Organization (VO) providing a flexible and integrated framework to users and resource providers, including AAI mechanisms.
- Define and integrate a user-oriented Science Gateway enabling: data-intensive seismology applications; access to private and public resources while hiding some of the complexity and heterogeneity of task submission and AAI policies; execution monitoring and provenance tools and services.

The main deliverables and achievements in the last six months are in line with the VERCE DOW and roadmap:

- Operation and management of the VERCE platform (D-SA1.4.1, SA1), in coordination with SA2, SA3 and JRA1. This includes: access integration of new HPC, Grid and institutional resources to the VERCE platform and the Science Gateway, i.e., HPC-CINECA-04, DEP-SCAI-01, DEP-INGV-01, extending support of UNICORE in addition to GLOBUS by the Science Gateway and the available computing resources for users and training events; extension of the number of sites integrated within the VERCE's INCA monitoring system; improved robustness and performance of the VERCE data layer backend for the Science Gateway, i.e., involving storage restructuration of the VERCE's iRODS platform and new replication policies by leveraging iRODS' micro services; virtualisation of all VERCE services, with improved continuity of services and backup mechanisms, that may now be hosted in commercial or public Cloud environments; integration of new FDSN data services, i.e., GCMT FDSN services for earthquake events; improved certificates support, based on X.509 and IGTF certificates, minimizing the AAI hurdles in order to smooth the path of new users approaching an e-infrastructure like VERCE that composes many different IAA regimes and policies.
- VERCE platform integration and release (DSA2.4.1, SA2), in coordination with SA1, SA3 and JRA1. The sixth release of the integrated tools and services was successfully completed on time, end of March 2015, reaching the M-SA2.6 milestone. Five components were evaluated and all approved, this includes: a new version 2.0.2, of SPECFEM 3D, h5py a python component in support of ObsPy and *dispel4py*; *dispel4py* 1.0 system to construct graphs of connected processes to transform, analyse and handle data streams and collection of data from files or storage services, with automatic mappings to different computing and storage infrastructures. The later supporting the misfit calculation implementation within FWM multi-purpose workflows.

- Scientific Gateway update and user feedback evaluation (D-SA3.4.1, SA3), in coordination with JRA2, JRA1, SA2, SA1, NA3 and NA2. This includes in addition to the new FWM tools described in the previous report (D-SA3.4) extended functionalities and flexibility in response to users' feedback: i.e., improved management of meshes and models in addition to the VERCE shared library, allowing the use and the management of custom models; integration of new data sources, i.e., specifically the new FDSN compliant service exposing GCMT parameters in QuakeML; multi-purpose workflows allowing advanced users to make use of the gateway tools and services, i.e., in particular the Provenance tools, to gather, produced and distilled interactively FWM input data which can be retrieved, filtered and shipped to targeted resources supporting specific solver installation; extended Provenance and exploration tools, supporting the W3C-PROV JSON representation and improving interoperability with other provenance exploration tools, to explore workflows' outcome, i.e., simulated waveforms, metadata and dependencies, data sources and diagnostics. Those make substantial use of the VERCE data management layer with authorisation-based access control.
- The service activities, together with NA3 and JRA2, have been quite active in preparing, supporting and running an intense training and outreach agenda, together with interacting with users, collecting and analysing feedback for refining the implementation strategy of the VERCE roadmap. The robustness of the VERCE platform, tools and services has been successfully demonstrated during the different training sessions and workshops where multiple users with different level of expertise were working simultaneously. During the last six months, iterations over the users' feedback, thanks to the training events, face-to-face internal meetings and code sprints involving seismologists of VERCE have been driving the overall service activities, and the close technical interactions with the ER-FLOW and SCI-BUS projects.
- The VERCE Science Gateway already has nearly 100 registered users, and others waiting to be accepted as new features are being rolled out: typically they are lecturers, researchers, Post-Docs and PhD students in research institutions, and in few cases exploration geophysics professionals, and they are coming from all over Europe or from other continents; the most active ones are those who have attended one of the training sessions in addition to researchers within the VERCE consortium; they provide continuous feedback and suggestions which are analysed to guide and prioritise tools and services development that are on time with the VERCE roadmap and meet the M-SA3.7 milestone.
- The next period priorities, in line with the VERCE roadmap, are devoted to: the finalisation of the misfit data analysis implementation and integration to the FWM multi-purpose workflows, making use of *dispel4py* processes, an activity which has already started with a dedicated taskforce engaged in regular meetings and code sprints, i.e., Rome, November 2014 and Edinburgh, April, 2015; continued integration and validation of additional computational resources; deployment of the VERCE operated services, i.e., Science Gateway and associated web services, iRODS and provenance services, portal support, on other institutional resources for continued operation and sustainability of the VERCE services as thematic core services and integrated core services contribution in the framework of the H2020 funded EPOS-IP project; new tools and services improvements suggested by users' feedback and in line with the VERCE roadmap.

1.3. Research and development activities

The Research and Development activities draw on the data-intensive 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 wider users community through reusable workflow 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, i.e., data and HPC-Grid-Cloud computing.
- Maintain a balance between long-term sustainability considerations and fast implementation of new scientific methods for external users to evaluate.

The main objectives and deliverables during the last 6 months were:

- Validation of a new version of the FWM solver, i.e., SPECFEM3D, and its support, in collaboration with SA1, SA2 and SA3, on new resources with the VERCE Science gateway; extension the earth models and meshes library tools and services and of the synthetic waveform data exploration and visualisation tools.
- In coordination with JRA2, SA3, SA2, SA1 and NA2, extended flexibility has been added to FWM multi-purpose workflows with in particular, an improved interactive FWM input file generator integrated with the VERCE data layer; new extension of the ObsPy library for the support of seismology data and file formats, in particular the NDK file format of the GCMT earthquake catalogue and the version 1.1 of the FDSN web services; a prototype implementation of the complete full waveform inversion, i.e., LASIF, with SPECFEM3D interface, and provenance tracking, that is intended to be gradually mapped onto the VERCE platform.
- Contribution to integrated international activities of the computational seismology community for developing and adopting new high-level data and meta-data structure representation and container formats such as the *Adaptable Seismic Data Format* (ASDF) that support HDF5 and the *Adaptable IO System* (ADIOS) - in collaboration with the NSF-funded CGI project and Oakridge National Lab (DOE) – and better exploitation of parallel I/O from multiple streams and storage architectures.
- Together with JRA2 and NA2, extension of preconfigured data analysis pipelines using the *dispel4py* Python system that can be used as exemplars for researchers wishing to build their own or use some simple pipeline reconfiguration options. They include: those used to post-process FWM results and generate visualisation; those used in the FWM data gathering and distilling phase; a number of noise-correlation analysis pipelines examples.
- Support and training material and exemplars for the dense VERCE training and outreach activity program including the last 9th-11th March, 2015, VERCE training workshop, the next May 18th-22nd EGI Forum meeting, June 1st-5th training session with the TIDES-COST project, and the July 1st-3rd VERCE training workshops.
- The JRA1 activity is on time according the VERCE roadmap and meet the M-JRA1.6 milestone.
- The next steps are: to finalise the post-processing data analysis FWM workflow components, i.e. the misfit components, and its integration within the Science Gateway making use of the *dispel4py* system; further extension of the ObsPy Python library; new seismological examples of data-intensive analysis of continuous waveforms using *dispel4py;* support of solvers and tools on additional computing resources recently integrated or being integrated to the VERCE Science Gateway.
- Close collaboration between JRA2, SA3 and SA1 led to: improvement of the performance and the robustness of the iROD-enabled data management layer, through storage restructuring of the Cloud infrastructure at UEDIN and leveraging iROD's micro services with new replication; extension of iRODS and Provenance services; extension of the gUSE/WS-PGRADE Science Gateway framework to support higher user-oriented flexibility of FWM workflows and better integration of the dispel4py technology for data analysis components.
- Extension of the *dispel4py* system and of its mapping functionalities on a number of parallel computing architectures. The *dispel4py* Python system is now released to seismologists who can easily installed it and use their usual Python programming tools and set their own data analysis pipelines up to graph of connected processes to transform, analyse and handle data streams and data from files or storage services, that can be mapped and run different distributed computing infrastructure (DCI) environments. Currently VERCE supports *dispel4py* mappings for sequential, multiprocessing and MPI/OpenMP parallel environments. In collaboration with JRA1 and NA2, the data-intensive processing-element library for *dispel4py* has been extended, many of them exploiting ObsPy functions, together with a number of preconfigured pipelines that can be used as exemplars for seismologists wishing to build their own, or be the basis for some simple pipeline user-oriented reconfiguration.
- JRA2 has been actively involved in the production of training material and user documentation, i.e., especially *dispel4py* for data intensive research, and in running *dispel4py* sessions and practical during the VERCE training sessions, especially in October, 2014, March 2015, and the next June and July sessions. The overall feedback on the *dispel4py* system has been very positive.
- The JRA2 activity is on time and in line with the VERCE roadmap. The next steps will involve extended integration of *dispel4py* within the Science Gateway in support of FWM data analysis

components both for pre- and post- processing, and for continuous waveforms analysis in support of seismic noise correlation studies.

1.4. VERCE roadmap update

The VERCE roadmap of planned services, tools and e-infrastructure framework was issued in the last reporting period, and articulated around four elements:

1. A computational seismology service for forward waveform modelling simulation;
2. A toolkit for seismologists analysing data;
3. An extensive collection of data handling services;
4. An e-Infrastructure framework, providing an integrated means of using all these services, tools and algorithms.

	Status		
	Evaluation	Production	Comment
Computational seismology: FWM			
SPECFEM3D package simulation service		31/03/2015	On time
Metadata collection from model runs	31/12/2014	30/06/2015	On time
Integration with FDSN and GCMT web services	31/03/2015	30/09/2015	In advance
Misfit analysis library tools and services	31/12/2014	31/03/2015	In progress with a slight delay: expected 31/07/2015
Pre-processing and post-processing pipeline configuration	31/03/2015	30/09/2015	On time
Mesh and model upload services and shared library	31/03/2015	30/09/2015	In advance: support users' custom models and re-parameterisation and provenance tools
Seismic data analysis			
ObsPy Python library		Since 2012	Continuous extended functionalities
Dispel4py python system	30/09/2014	31/12/2014	On time and already released, continuous extension
Data-processing element libraries for dispel4py	31/12/2014	30/06/2015	On time

dispel4py mapping to DCI architectures	30/09/2014	31/03/2015	On time: include sequential, multiprocessing and MPI architectures
Preconfigured library of seismological example pipelines	31/12/2014	30/09/2015	On time
Data handling services			
Federated iRODS for observation and synthetic data	31/03/2014	31/03/2015	On time: performance and robustness improvements, new replication policies
GridFTP bulk data transport	31/03/2014	31/03/2015	On time
Linking secondary seismic data with descriptive metadata	31/03/2014	30/06/2015	On time
Provenance metadata	31/12/2014	31/03/2015	On time: extended to support W3C-PROV JSON reprsentation
Tools to exploit data provenance	31/12/2014	30/06/2015	On time
Use of iRODS and MongoDB services	31/12/2014	31/03/2015	On time
Ingest tools and services of data collections	30/06/2015	30/09/2015	In progress
e-Infrastructure framework			
Science Gateway with gUSE/WS-PGRADE framework		30/09/2014	On time with already 100 registered users providing continuous feedback Include extended HPC and institutional resources
SCI-BUS and ER-FLOW		30/09/2014	Continuous collaboration within signed MOUs
FWM Job submission		30/09/2014	On time: adding new institutional and Grid resources
Gateway running on private or public Clouds using VMs		31/12/2014	On time
Forward modelling on HPC systems	30/09/2014	31/12/2014	On time: include a number of existing community services and exploration and visualisation tools
FWM workflows	31/09/2014	30/03/2015	On time: new multi-purpose functionalities with increased user-oriented flexibility

Provenance Management	30/09/2014	31/12/2014	On time: provenance API extended to support W3C-PROV JSON representation
Document store	30/09/2014	31/12/2014	On time
iRODS frontend	30/09/2014	31/12/2014	On time: continuous improvement
Security credential handling	30/09/2014	30/12/2014	On time: continuous extension and improvement in response to users' feedback

1.5. Milestones and deliverables

Table 1- Milestones

MILESTONES							
Milestone no.	**Milestone name**	**Work package**	**Lead beneficiary**	**Delivery date from Annex I**	**Achieved**	**Actual / Forecast achievement date**	**Comments**
MS1	M-NA1.1	1	CNRS	Month 6	Yes	25/05/2012	
MS2	M-NA2.1	2	INGV	Month 6	Yes	25/05/2012	
MS3	M-NA3.1	3	ULIV	Month 6	Yes	25/05/2012	
MS4	M-NA1.1.1	4	CNRS	Month 6	Yes	31/03/2012	
MS5	M-NA4.1.1	4	EMSC	Month 6	Yes	25/05/2012	
MS6	M-SA1.1	5	CNRS	Month 6	Yes	25/05/2012	
MS7	MSA3.1	7	KNMI	Month 6	Yes	25/05/2012	
MS13	M-JRA2.1	9	UEDIN	Month 6	Yes	25/05/2012	Slight deviation; see paragraph 2.8 in periodic report 1. "WP9 - JRA2: Tools for data analysis on modeling"
MS8	M-NA1.1.2	1	CNRS	Month 12	Yes	01/10/2012	
MS9	M-NA2.4	2, 4	INGV, EMSC	Month 12	Yes	01/10/2012	

MS10	M-SA2.1	5, 6	CNRS, BADW-LRZ	Month 12	Yes	01/10/2012	
MS11	M-SA3.2	7	KNMI	Month 12	Yes	01/10/2012	
MS12	M-JRA1.1	2, 8	INGV, LMU	Month 12	Yes	01/10/2012	
MS24	M-JRA2.2	9	UEDIN	Month 12	Yes	01/10/2012	
MS14	M-NA1.2	1	CNRS	Month 18	Yes	01/04/2013	
MS15	M-NA2.2	2, 3	INGV	Month 18	Yes	01/04/2013	
MS16	M-NA3.4.1	3	ULIV	Month 18	Yes	01/04/2013	
MS17	M-SA2.2	5, 6	BADW-LRZ	Month 18	Yes	01/04/2013	
MS18	M-SA3.3	7	KNMI	Month 18	Yes	01/04/2013	
MS19	M-JRA1.2	2, 6, 8	LMU	Month 18	Yes	01/04/2013	
MS20	M-NA1.2.1	1	CNRS	Month 24	Yes	01/10/2013	
MS21	M-SA2.3	5,6	CNRS, BADW-LRZ	Month 24	yes	01/10/2013	
MS22	MS-SA3.4	7	KNMI	Month 24	Yes	01/10/2013	
MS25	M-NA1.3	1	CNRS	Month 30	Yes	01/04/2014	
MS26	M-NA3.4.2	3	ULIV	Month 30	Yes	01/04/2013	
MS27	M-SA2.4	5, 6	LRZ	Month 30	Yes	01/04/2013	Slightly revised platform components
MS28	M-SA3.5	7	KNMI	Month 30	Yes	01/04/2014	
MS29	M-JRA1-4	2,6,8	LMU	Month 30	Yes/no	01/04/2014	Slightly delayed and revised following reviewer's suggestions
MS30	M-NA1.3.1	1	CNRS	Month 36	Yes	08/10/2014	
MS31	M-SA2.5	5,6	LRZ	Month 36	Yes	30/09/2014	

MS32	M-SA3.6	7	KNMI	Month 36	Yes	30/09/2014	
MS33	M-JRA1.5	8	LMU	Month 36	Yes	30/09/2014	
MS35	M-NA1.3.2	1	CNRS	Month 42	Yes	01/04/2015	
MS36	M-SA2.6	5,6	LRZ	Month 42	Yes	01/04/2015	
MS37	M-NA2.3	2	INGV	Month 42	Yes	01/04/2015	
MS38	M-SA3.7	7	KNMI	Month 42	Yes	01/04/2015	
MS39	M-JRA1.6	8	LMU	Month 42	Yes	01/04/2015	

Table 2- Deliverables

Del	Name	WP	Lead	Nature	Diss. level[5]	Delivery date Annex I	Delivery date	Status	Contractual Yes/No
D2.1	D-NA2.1	2	INGV	Report	Public	Month 6	25/05/2012	Submitted	Yes
D3.1	D-NA3.1	3	ULIV	Report	Public	Month 6	25/05/2012	Submitted	Yes
D4.1	D-NA4.1	4	EMSC	Report	Public	Month 6	25/05/2012	Submitted	Yes
D5.1	D-SA1.1	5	CNRS	Report	Public	Month 6	25/05/2012	Submitted	Yes
D6.1	D-SA2.1	6	LRZ	Report	Public	Month 6	25/05/2012	Submitted	Yes
D7.1	D-SA3.1	7	KNMI	Report	Public	Month 6	25/05/2012	Submitted	Yes
D8.1	D-JRA1.1	8	LMU	Report	Public	Month 6	25/05/2012	Submitted	Yes
D9.1	D-JRA2.1	9	UEDIN	Report	Public	Month 6	25/05/2012	Submitted	Yes
D1.1.1	D-NA1.1.1	1	CNRS	Report	Public	Month 12	01/10/2012	Submitted	Yes
D2.2	D-NA2.2	2	INGV	Report	Public	Month 12	01/10/2012	Submitted	Yes
D2.2.1	D-NA2.2.1	2	INGV	Report	Public	Month 12	01/10/2012	Submitted	Yes
D3.2	D-NA3.2	3	ULIV	Report	Public	Month 12	01/10/2012	Submitted	Yes
D4.3	D-NA4.3	4	EMSC	Report	Public	Month 12	01/10/2012	Submitted	Yes
D5.2	D-SA1.2	5	CNRS	Report	Public	Month 12	01/10/2012	Submitted	Yes
D6.2	D-SA2.2	6	LRZ	Report	Public	Month 12	01/10/2012	Submitted	Yes

D7.2	D-SA3.2	7	KNMI	Report	Public	Month 12	01/10/2012	Submitted	Yes
D8.2.1	D-JRA1.2.1	8	LMU	Report	Public	Month 12	01/10/2012	Submitted	Yes
D9.1.1	D-JRA2.1.1	9	UEDIN	Report	Public	Month 12	01/10/2012	Submitted	Yes
D2.2.2	D-NA2.2.2	2	INGV	Report	Public	Month 18	01/04/2013	Submitted	Yes
D5.2.1	D-SA1.2.1	5	CNRS	Report	Public	Month 18	01/04/2013	Submitted	Yes
D6.2.1	D-SA2.2.1	6	LRZ	Report	Public	Month 18	01/04/2013	Submitted	Yes
D7.2.1	D-SA3.2.1	7	KNMI	Report	Public	Month 18	01/04/2013	Submitted	Yes
D1.2.1	D-NA1.2.1	5	CNRS	Report	Public	Month 24	01/10/2013	Submitted	Yes
D2.3	D-NA2.3	2	INGV	Report	Public	Month 24	01/10/2013	Submitted	Yes
D2.3.1	D-NA2.2.1	2	INGV	Report	Public	Month 24	01/10/2013	Submitted	Yes
D3.2.1	D-NA3.2.1	3	ULIV	Report	Public	Month 24	01/10/2013	Submitted	Yes
D4.3.1	D-NA4.3.1	4	EMSC	Report	Public	Month 24	01/10/2013	Submitted	Yes
D5.3	D-SA1.3	1	CNRS	Report	Public	Month 24	01/10/2013	Submitted	Yes
D6.3	D-SA2.3	6	LRZ	Report	Public	Month 24	01/10/2013	Submitted	Yes
D6.3.0	D-SA2.3.0	6	LRZ	Report	Public	Month 24	01/10/2013	Submitted	Yes
D7.3	D-SA3.3	7	KNMI	Report	Public	Month 24	01/10/2013	Submitted	Yes
D8.2.2	D-JRA1.2.2	8	LMU	Report	Public	Month 24	01/10/2013	Submitted	Yes

D9.1.2	D-JRA2.1.2	2	UEDIN	Report	Public	Month 24	01/10/2013	Submitted	Yes
D2.3.2	D-NA2.3.2	2	INGV	Report	Public	Month 30	01/04/2014	Submitted	Yes
D5.3.1	D-SA1.3.1	5	CNRS	Report	Public	Month 30	01/04/2014	Submitted	Yes
D6.3.1	D-SA2.3.1	6	LRZ	Report	Public	Month 30	01/04/2014	Submitted	Yes
D7.3.1	D-SA3.3.1	7	KNMI	Report	Public	Month 30	01/04/2014	Submitted	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
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
D1.3.2	D-NA1.3.2	1	CNRS	Report	Public	Month 42	30/03/2015	Submitted	Yes
D5.4.1	D-SA1.4.1	5	CNRS	Report	Public	Month 42	30/03/2015	Submitted	Yes

D6.4.1	D-SA2.4.1	6	LRZ	Report	Public	Month 42	30/03/2015	Submitted	Yes
D7.4.1	D-SA3.4.1	7	KNMI	Report	Public	Month 42	30/03/2015	Submitted	Yes

2. Project management during the period

2.1. VERCE consortium: management, structure and governance

No change in the Consortium composition and beneficiaries'.

Project Management Office (PMO): no change, composition of the PMO 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)

2.2. Internal communication

No change has occurred in the organisation of internal communication.

2.3. Internal meetings

Since the last reporting period, the Steering Committee has regularly met on a three-weeks regular basis through remote meetings, while the Project Executive Board is meeting on a two-week regular basis. Minutes of those meetings were made as usual available on the Redmine platform.

A joint SC-PEB two-day face-to-face meeting took place at IPG Paris March 26th-27th, 2015, with a primary focus on the VERCE roadmap and the new VERCE training sessions.

A dedicated task force has been established across a number of related work packages (NA2, SA3, JRA1, JRA2), for the implementation of the misfit data analysis components in the FWM workflows in line with the VERCE roadmap. This transverse activity has gained a good momentum through regular meetings and code sprints, i.e., Rome, November 2014, Edinburgh, April 2015, and share a wiki section in the Redmine platform.

Other internal meetings were organised during international events in relation with the VERCE outreach and dissemination activity, i.e., EGU, EGI Forum for example. Those meetings also provide opportunities to organize coordination meetings with other related projects, and make sure VERCE develops coherently with the Community's needs and wishes.

2.4. Training sessions and webinars

VERCE has engaged users in training sessions and webinars, recording feedback and impressions on how the Science Gateway services will impact their research practices and common needs.

After the two-day training webinar, July 10th and 17th 2014, providing an overview of the VERCE Science Gateway and FWM workflows practical, a two-day training webinar took place October 15th and 16th 2014, dedicated to the VERCE tools, i.e., specifically *dispel4py*, for data-intensive analysis with a number of practical exercises illustrating the dispel4py functionalities for extending user's processing elements – often Python scripts, and build up workflows to transform, analyse and handle data streams and collections of data from files or storage services. A three-day training workshop was organised, in coordination between JRA1, NA1, SA3 and NA3, in Munich from the 9th to the 11th of March 2015. This extensive training provided an overview of the Science Gateway tools and services, of the FWM workflow and data management tools and services, and of the *dispel4py* system tools, with a number of practical sessions.

The overall feedback was quite positive recognising the potential offered by the VERCE Science Gateway in support of their own research practices. The extended flexibility for managing Earth models and meshes, together with gathering earthquake data and distilling all those as input files for FWM simulation scenarios was very much appreciated. The *dispel4py* system was well received and quite in line with the present day research practices and common needs which are quite different from the FWM research community.

A number of training videos and user documentation have been produced for these two training sessions and are openly available on the VERCE website.

The VERCE Science Gateway already has nearly 100 registered users, i.e., lecturers, researchers, Post-Docs, PhD students from everywhere in Europe and in a few cases exploration geophysics professional.

2.5. External meetings

In the last six months, the VERCE members participated in a number of scientific and outreach events where the VERCE Science Gateway, together with the FWM workflow tools and the *dispel4py* system were demonstrated to a multi-disciplinary audience.

A selected list includes:

- **EGI conference on Challenges and Solutions for Big Data processing on Clouds**, 16th-26th September 2014, Amsterdam, Netherland. Participants: H. Schwichtenberg, A. Germünd (SCAI), one talk.
- **The 13th International Sematic Web Conference (IWSC 2014)**, 19th–23rd October 2014, Trentino, Italy. Participant: A. Spinuso (KNMI). One communication.
- **Super Computing 2014 conference (SC'14),** 16th-20th November 2014, New Orleans, USA. Participants: R. Filgeira, I. Klampanos, M. Atkinson, A. Krause (UEDIN), M. David (CNRS-IPGP). One communication at the DISCS'14 workshop, and a video on the FWM VERCE activity, with SPECFEM3D, at the CINECA boot.
- **American Geophysical Union (AGU) fall meeting,** 15th-19th December 2014, San Francisco, US. Participants: A. Michelini, E. Casarotti (INGV); H. Igel, L. Krischer (LMU), A. Rietbrock, T. Garth (ULIV); J.-P. Vilotte (IPGP-CNRS); A. Spinuso (KNMI) … four posters and two oral communications.
- **E-Infrastructure for Earth Sciences – workshop (EGI, PRANCE, GEANT, EUDAT),** 22nd–23rd January 2015, Amsterdam, Netherland. Participants: A. Spinuso (KNMI), H. Schwichtenberg, A. Germünd (SCAI). One oral communication
- **5th Munich Earth Science School,** 22nd-27th February 2015, Sudelfeld, Germany. Participants: H. Igel, L. Krischer (LMU), A. Michelini, E. Casarotti (INGV). One oral communication and demonstrator.
- **European Geophysical Union (EGU) General Assembly 2015,** 12th-17th April 2015, Vienna, Austria. Participants: R. Filgueira, M. Atkinson (UEDIN); A. Michelini, E. Casarotti (INGV); A. Rietbrock, T. Garth (ULIV); H. Scwichtenberg, A. Germund (SCAI), R. Bossu (EMC); J.P. Vilotte, G. Moguilny (CNRS-IPGP); A. Spinuso (KNMI) … 3 posters and 6 oral communications, dissemination material at the EPOS boot.
- **11th IEE International conference on e-science,** 31st August – 4th September 2015, Munich, Germany. Participants: M. Atkinson, A. Krause, I. Klampanos, R. Filgueira (UEDIN); A. Frank, S.H. Leong (LRZ); M. Carpené (CINECA); J.-P. Vilotte (CNRS-IPGP); F. Magnoni, A. Michelini (INGV); H. Igel, L. Krischer (LMU); A. Spinuso (KNMI); H. Schwichtenberg, A. Germünd (SCAI). Two accepted papers. A co-located workshop on IT for Research Infrastructures (IT4RIs) is being organised by ENVRI, with M. Atkinson (UEDIN) as one of the three organisers and several VERCE members on its PC. It will provide communication from VERCE to 22 research infrastructures.

2.6. Cooperation with other projects

Part of the effort to build a sustainable and interoperable infrastructure is the investigation of possible collaborations with other sharing relevant projects.

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

- Of particular importance is the contributions of VERCE to EPOS, the solid Earth science ESFRI project, i.e. especially with regard to the thematic (seismology, WP8) and integrated (WP6 & WP7) core services of the EPOS e-science environment for the EPOS implementation phase which has been funded under the H2020-INFRADEV-1-2015-1.
- VERCE has collaborations with international NPOs in seismology like IRIS and Earthscope in the US, and JAMSTEC and NIED in Japan.

- Another important aspect is the synergy and the collaborations between VERCE and a number of strategic FP7 projects in seismology, i.e., NERA (2010-2014), SHARE, and REAKT.
- Synergy and contribution to the Belmont Forum e-infrastructure Coordinated Research Action in environmental sciences within the G8 framework and the Future Earth global challenges.
- Finally, with regard to the research applications and the dissemination of VERCE in the seismology research community, synergy and coordination with a number of European Research Council projects, i.e., WaveTomo (ERC), WHISPER (ERC), and more recently with the NERC-funded TerraCorrelator projects.

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

- VERCE is participating to two pilot case studies within the EGI-PRACE-EUDAT synergy.
- VERCE has been collaborating with EUDAT with regard to iRODS based data management federation and data movement using GridFTP. VERCE is now collaborating with and contributing to the H2020 follow-on project EUDAT2020.
- VERCE is also contributing to the H2020 EGI-Engage project, funded under the H2020-EINFRA-2014-2 call, i.e., specifically within the EPOS Competence Centre (CC-EPOS), with a special focus on AAI harmonization across multiple e-infrastructures in coordination with EPOS.

The last aspect of this strategy is the synergies and coordination with other related projects that develop methodologies and software components 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 common understanding and a collaboration roadmap have been finalized and formalized as a MOU between the two projects.
- IGE: providing tools to share computing resources, databases and other on-line tools. IGE is a member of the Globus Alliance. Collaborations between VERCE and IGE have been formalized by a signed MOU. This collaboration has been important and has led to the adoption by VERCE of a number of components provided by IGE.
- ER-Flow: a follow-on project to SHIWA, which developed technologies to allow interoperability between workflow systems. Collaboration between ER-flow and VERCE has been formalized 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 supporting the new Adaptable Seismic Data Format (ASDF) encapsulating HDF5 and ADIOS container formats, together with extended metadata associated with 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 acquire know-how not only for the VERCE project but also for further projects of the solid Earth sciences.

2.7. Project activities monitoring

The situation is continuously monitored by the PMO.

2.8. Risk management plan

The Risk Management Plan is continuously reviewed and refined by the SC and the PEB. It has received particular revision within the VERCE roadmap document delivered in the last reporting period.

Each risk is assessed in terms of likelihood and impact, and is recorded in the Risks Register. This metric allows a Risk Exposure factor to be defined, which helps to prioritize the different risks for management control.

2.9. Last review recommendations

- ***The consortium should provide a revised work plan defining the goals to be achieved by the end of the project.***
  - ❖ A VERCE roadmap of planned services, tools and e-infrastructure has been produced and documented in a separate document delivered in the last reporting period. In this document, progress status is provided and will be continuously updated.

- ***The VERCE architecture should be "freeze" and its final version thoroughly documented.***
  - ❖ A comprehensive description and analysis of the VERCE e-infrastructure components is provided in the VERCE roadmap document, together with an analysis of the main issues and risks. The critical components are all in production since December 2014. The D-SA1.4.1 report provides in addition a technical analysis of the operated VERCE services, taking into account the different VERCE stakeholders and defining activity areas: Science Gateway and associated web services; iRODS and provenance services; portal support; Inca monitoring service; FDSN services. Each individual activity and the support it currently gets are described to provide an element of the sustainability assessment. Nevertheless, continuous improvements and flexibility will be rolled out until the end of the project. This activity is critical to keep VERCE in phase with the research practices and common needs, and engaged more of them while facilitating the transition from proof-of-concept demonstration to dependable research facility within the EPOS implementation phase.

- ***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.***
  - ❖ In the last 12 months, a dense programme of training webinars and sessions has been successfully implemented, i.e., webinars in July and October 2014, training workshops March, June and July 2015, providing an overview of the VERCE Science Gateway tools and services, of the FWM workflows and data management tools, of the data-intensive analysis tools with the *dispel4py* system through hands-on practical, demonstrators and documentation. Users' feedback was recorded and analysed leading us to roll out improvements, extend flexibility and introduce new features. The VERCE Science Gateway already has nearly 100 registered users and others are waiting to be accepted as we roll out new features. Typically they are from academic research, i.e., lecturers, researchers, Post-Docs and PhD students, and in few cases from exploration geophysics industry. Recorded feedback shows a real interest in adopting the VERCE Gateway services and tools for their own research, as well as in adopting the VERCE data-intensive research tools, i.e., specifically *dispel4py*. These events will be followed by two additional training sessions. The first has been solicited by the TIDES Cost-project, and will be held 1st-5th of June in Bertinoro (Italy), aiming at exposing young seismologists to the latest generation of tools for data analysis and modelling. The second one is the last VERCE training workshop in Liverpool 1st-3rd of July 2015, including both the Science Gateway FWM workflows and tools, and the *dispel4py* system for data-intensive continuous and event waveforms analysis. In addition, VERCE will present data-intensive analysis and FWM training material to the international school, organised by the NSF-funded Open-Science Data Cloud PIRE project, which will be held at the University of Amsterdam 8th-12th of June 2015, and for some of the participants, followed by a six-week practical at the University of Edinburgh. Training will continue beyond VERCE, in the VERCE partners, as well as in the newly funded EPOS-IP project.

- ***The two fully functional use cases (compute intensive – SPECFEM) and data-intensive (seismic noise correlation) should be finalized.***
  - ❖ In the strategy of VERCE, attention is focused on the co-evolution of the e-science environment with research practice in seismology, through continuous feedback from the seismology researchers. FWM research practice in the international computational seismology community is today dependent upon the use of complex and high-performance "community" wave propagation solvers, such as SPECFEM3D, as a service. The VERCE Science Gateway provides today extended and flexible tools for gathering and distilling earthquake earthquake's event data, earth models and meshes, parameterising simulation scenarios and submitting them to external computing resources together with retrieving and analysing the simulated waveforms. The feedback recorded from the training sessions attendees and the Science Gateway users show that the potential offered by the VERCE Science Gateway is really well evaluated and in phase with the research

practice of advanced and less advance users, facilitating the use of the latest generation of simulation tools in their research.

- ❖ With regard to data-intensive analysis research, i.e., specifically the seismic noise correlation applications, the international seismology community is far less organised around standardised processing and analysis pipelines, and organised around the development of their own analysis pipelines around specific scientific approaches. The ObsPy Python library, providing Python functions for processing and analysing many common forms of seismic data in different formats, and Python programming tools are widely used by this community. Therefore the strategy of VERCE has evolved and adapt, integrating the ObsPy library with the tools and services supported by VERCE, and contributing to its extension with new functionalities, i.e., in particular with the support of the new Adaptable Seismic Data Format (ASDF) encapsulating HDF5 and ADIOS container formats, together with extended metadata associated to the end-to-end data processing and analysis pipelines. ObsPy has been extended to support new FDSN services. The VERCE strategy has been to develop the *dispel4py* system allowing seismologists to construct graphs of connected processes to transform, analyse and handle data streams and collections of data from files or storage services. Seismologists can now write their own algorithms and methods with their usual Python programming tools, making use of ObsPy functions, and set them up to be used as stages, called "Processing Elements" in these graphs to compose their own *dispel4py* abstract workflow graphs that can then be mapped onto different DCIs architectures. These work together with metadata and provenance management tools already developed for the FWM applications. A basic library of data-intensive PEs and of example pipelines has been designed and is available for download to help seismology researchers wishing to build their own. The flexibility of these tools have been well received by users who found there a strategy that is well aligned with their current research practice and needs.
- ❖ Finally, the last goal of VERCE is to provide also misfit data analysis components, making use of the *dispel4py* functionalities, integrated within the FWM workflows functionalities as part of the VERCE Science Gateway. This is again in line with the research practice of the international computational seismology community who are seeking for standardised and automatized misfit analysis components to evaluate FWM scenarios against recorded event waveforms. This is well advanced and organised with a dedicated VERCE taskforce to meet the VERCE roadmap.

- ***Following the completion of beta testing by external users, the consortium should prepare videos documentation for advertising the system to the wider scientific community.***
  - ❖ At this stage, a first step toward this recommendation has been the production and the dissemination of a number of training videos, related to VERCE methods, tools and services. It is worth mentioning here that the VERCE Science Gateway achievements have been presented and demonstrated to a multi-disciplinary audience at a number of outreach events, such as the EGI Forum (the next one being the Lisbon meeting 18th-22th of May, 2015), and international conferences, such as the AGU Fall meeting (December 2014), the EGU meeting (April 2015), or the next eleventh IEEE international eScience Conference in September 2015). It has also been presented and demonstrated at a number of internal meetings of the EPOS ESFRI project, where it has been positively evaluated and integrated with the thematic core services (seismology, WP8) of the newly funded EPOS IP project. Today a professional video is being prepared by ULIV as part of the NA3 activity, and should be delivered after the last training meeting in July 2015. This would be used, together with the documentation already available on the VERCE website, to advertise VERCE system even after the end of this project.

- ***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.***
  - ❖ Improved integration and operation of certificates, making authentication on the VERCE gateway easier is described in the D-SA1.4.1 and D-SA3.4.1 reports and in the last reporting period, seeking to minimise these security hurdles and smoothing the path for scientists when they first approach an e-infrastructure such as VERCE that composes many different authorisation regimes. Many of these issues are outside the project's control and are receiving attention from new H2020 projects and others. These residual issues with access to the VERCE platform are due to the current authorisation and accounting hurdles, which, despite the homogeneous technical solutions adopted by VERCE across institutional and international DCIs, still have to comply with different policies, certification and validation steps. This can indeed delay the uptake from interested users. VERCE is actually contributing to newly funded H2020 projects like EUDAT2020, and EGI-Engage, where VERCE contribute today to the EPOS Competence Centre (EPOS-CC), an activity that will

continue as part of the EPOS-IP project. More information can be found in the D-SA3.4.1 report. Today the VERCE Science Gateway already has nearly 100 registered users, while others are waiting to be accepted as we roll out new features. Typically, they are lecturers, researchers, Post-Docs and PhD students in Earth Sciences in universities and national agencies from all other Europe. In few cases, those are also exploration geophysics professionals from well know industries. One has however to bear in mind that the availability of HPC resources are limited. Gaining access and providing HPC resources on PRACE architectures is complex and must applied for individually. For those granted access VERCE is providing, via the VERCE Science Gateway, tools to make efficient use of those granted resources, as it has been illustrated for the researchers and students on each training courses. The new PRACE system call recognises communities and reduce these restrictions. This will be explored during the EPOS-IP phase. The situation is slightly easier with EGI Cloud and Grid resources.

- ***A continuity of services plan including appropriate procedures and deployment topologies should be drafted.***
  - ❖ This is addressed in the D-SA1.4.1 report, which provides insight into the effort and resources required to operate the VERCE platform, particularly with regard to the planning of continued operation and sustainability of the services that VERCE provides to users. Continuity of services was also addressed in the previous D-SA1.4 report. This has also been integrated in the management and sustainability plan of the newly funded H2020 EPOS-IP project where those services are part of the seismology thematic core services within the WP8 work package.

3. Elements for a VERCE sustainability plane and strategy

3.1. The role of VERCE

A primary goal of VERCE is to provide to the earthquake and seismology research community in Europe an e-infrastructure platform, tools and services enabling data-intensive computing and analysis methods exploiting the increasing wealth of open data generated by observational and monitoring systems in Europe and globally for addressing fundamental problems in understanding the Earth's internal wave sources and structures, and augmented societal applications. Another goal is to deliver "intellectual ramps" providing safe and supported means for seismology researchers and students to incrementally engage with new data-intensive methods and tools, and better exploit European data and computing infrastructures.

As described in the VERCE roadmap document, the main achievements of VERCE are the following:

- *Computational seismology services* for forward wave-field simulation in response to earthquake events;
- *Data-intensive toolkit for transforming and analysing data* from observations and simulations, including event and continuous waveforms;
- *Extensive collection of services to handle data streams and collection of data* from files or storage devices, that are used at different stages in the above contributions and which are suitable for use in Earth sciences and other domains in many cases;
- *E-Infrastructure framework and Science Gateway* that provides seismologists with an integrated means of using all these services, tools and algorithms, and which will facilitate further advances in data-intensive computational seismology and accommodate users with different levels of expertise.

These contributions adopt and build on standards that encourage wider international and cross-disciplinary collaboration. They are designed to work effectively on the e-infrastructures platforms seismologists currently use and to continue to work on new institutional, national, and multi-national distributed computing infrastructures (DCIs). They depend also on resources, services and technologies provided by many organisations and contemporary European projects.

The ***international seismological level*** is of course of vital importance, as VERCE will only work with uniform agreement on standards and protocols within the seismology community. Co-operation and agreement on interoperability standards is managed through the Federation of Digital Seismological Networks (FDSN) in coordination with the Observatories and Research Facilities for European Seismology (ORFEUS), in particular

the European Integrated Data Archives (EIDA) initiatives, under the umbrella of the Group on Earth Obervations (GEO). VERCE paid particular attention to liaison with other international initiatives, such as the Research Data Alliance (RDA) for Open data and the e-infrastructure and Data Management for Global Change Research Collaborative Research Action of the Belmont Forum.

At the ***national level***, the situation is different for each of the VERCE partners. Each of their countries has a seismological community, and potential users of VERCE for its scientific research. Moreover observatories, institutions or research teams that operate observational and monitoring networks, are interested in sharing data analytics and computational methods, tools and services for exploiting and valorising these data, and supporting up-take by their scientists and data providers. National initiatives are where the bulk of work takes place. These initiatives are all different, and their aims and organisation depend strongly on their national/agency priorities, and on the way their national research and funding schemes are organised. National funding must provide the core support for the data and computing infrastructures, which populate VERCE. In addition, the national initiatives are close to their national science communities.

At the ***European seismological level***, the European Plate Observatory System (EPOS) is the ESFRI initiative in solid Earth sciences binding together the national efforts in solid Earth sciences and increasing their effectiveness. VERCE represents the European virtual earthquake and seismology research community, an important element of the EPOS landscape, and has proven to be indispensable as an effective multi-level collaborative and coordination framework between European partners of the seismology research community, the data-science and the IT community, and the data-and-compute infrastructure providers.

VERCE paid particular attention to liaisons with the strategic EPOS-ESFRI initiative; "nearby" European projects of the seismology community, such as the NERA, SHARE and REAKT projects. Each of those projects pursued their own aims but learned from each other, building a global strategy and minimizing duplication of efforts. This helped national and European activities to be better focused and to have much greater impact.

At the ***wider European level***, effective liaison with projects dealing with the generic aspects of data and computing infrastructures, i.e., PRACE and EGI, and of interoperability, i.e., specifically EUDAT, was sought to make sure that the VERCE framework is able to interoperate with the European generic multidisciplinary framework, and that specific requirements of the data-intensive research community in seismology are taken into account.

The VERCE framework has never been an isolated island, and generic building blocks were used whenever possible. Following the strong recommendation from the reviewers and from the European commission, VERCE established liaison and collaboration with the SCI-BUS and ER-Flow initiatives. This collaboration was formalised through a Memorandum of Understanding (MoU) and a collaboration roadmap. In particular, the VERCE Science Gateway is compliant with SCI-BUS and ER-Flow technologies. The VERCE data management layer is based upon the iRODS technology supported by EUDAT. Provenance services and metadata are compliant with the W3C-PROV2 standards recommendation, while data services are compliant with the FDSN standards and protocols, supported by ORDEUS-EIDA. The use of current, open standards ensures that VERCE applications will be able to tap into generic and EPOS related resources. So far as those standards keep being supported at the European level.

Each VERCE partner brings different kind of resources with different management methods, and different expertise. Each VERCE node has its own organisation depending on its national priorities, which in turn depend of its funding sources and on the way seismology, data and computing infrastructures are organized in each country. In some countries, for instance, most resources are provided through dedicated projects; in others, there is a national coordination of the resources provided at the institute and national levels; in others, coordination is provided directly by a national funding agency. Many combinations of the above schemes are possible. Even though the balance between those contributions to the e-science VERCE framework is different for the different partners, the common characteristic is that all partners are strongly committed to VERCE development.

VERCE governance has been agile and pragmatically performed throughout the years of the project, and shaped by the research priorities and practices of the seismologists in VERCE, in collaboration with a number of "nearby" European Research Council (ERC) projects, for instance WHISPER and WAVETOMO.

3.2. Quantitative needs for a sustainable VERCE

During this reporting period, the current level of VERCE effort has been compared to an estimation of the resources required for a sustainable VERCE-VO, assuming that most services are frozen and no dramatic change occur, e.g. no very serious security issues and software components continue to be maintained (e.g., still upgrade support by the gUse community).

Activity	Current (for development)	Required (for sustainability)
Core service support and operation	2.0	1.5
Infrastructure support		
Science software and data-intensive tools support	1.5	1.0
Users community support	4.2	1.0
Total	4.5	3.5

In the **Core service and infrastructure support** and operation, D-SA1.4.1 provides an overview of the technical sustainability of the VERCE platform and services, identifying and describing the Core Services that need to be operated for a continuity of the services VERCE provides to scientific users.

The Core services include: the Science Gateway and associated web services (operated currently at SCAI), the iRODS platform and Provenance services (operated currently at UEDIN), the support portal (operated currently at CNRS-IPGP), the Inca monitoring services (operated currently at LRZ), and the FDSN-compliant services (operated currently at KNMI/ORFEUS). The VERCE core services are running in virtual machines and can be duplicated and deployed on new seismology resources providers, e.g. contributing to EPOS-IP such as INGV, CNRS, LMU and ULIV, who have already offered to host and support the operation some of those services within the EPOS implementation phase.

The Infrastructure support estimation assumes that the resource providers stay fixed, i.e., no new sites are joined to the VERCE platform. Additional effort is required for the integration of additional sites to the VERCE platform and to the VERCE-VO, honouring access policy requirements, which depends on the expertise and environment of the resource provider. When middleware (unicore, globus, glite) exists at the new site, effort could be estimated as ~ 1 PM per site. When the site has no or little experience in providing external access via globus or Unicore, it is difficult to estimate, as it would include basic middleware support. It is worth to note that seismological VERCE institutions have all experienced IT-persons and support external access to their resources. Additional effort is associated with the VERCE-VO coordination and management to maintain and to adapt the VERCE framework of technical standards, and the interaction with generic infrastructure initiatives, such as PRACE, EGI and EUDAT. Based on VERCE's experience of the required effort needed to develop and adopt standards and services, and to gather and transform community requirements into real infrastructure components, this is estimated to ~1 PM per year.

In the **support of the science software and tools**, i.e., FWM software (SPECFEM3D) and data analysis tools (ObsPy, *dispel4py*), the analysis highlights the range of approaches employed by the community in the development and support of forward waveform solvers and data-intensive analysis tools that are distributed across international community-driven organisation, such as the Computational Infrastructure for Geodynamics (CIG) or ObsPy, or research initiatives, such as *dispel4py*. The required resources to support the high impact software, tools and services are well identified, and there is a strong case for coordination of this European wide

efforts between EPOS-IP, and the European PRACE centres in order to promote community codes and analysis tools in the seismology domains, rather than unsustainable monolithic models. VERCE-VO support to those codes and data analysis tools, i.e., for example SPECFEM3D, ObsPy and dispel4py, and coordination is estimated at ~ 1 FTE, assuming that installation at the resource providers is done with provider's support.

The analysis of the VERCE activities for seismology users support shows that the community engagement, training and education events have a high impact, but are in need of European wide coordination via the VERCE VO during the EPOS implementation phase. While training will continue beyond VERCE, in the VERCE institutions for their staff and students, sustained coordination of the training activities, including the development of the necessary training materials, lead to an estimation at ~ 1 FTE

3.3. Elements of the VERCE sustainability strategy

EPOS continues to evaluate VERCE as one of the important infrastructures for seismology, and the recent history of the VERCE-VO shows that European coordination is essential, as well as participation of strong knowledgeable teams mostly funded at the national level, i.e., the current VERCE-VO and EPOS partners. The VERCE project has succeeded in building a European-wide community of data providers, scientific users and data scientists; in making data-intensive computational seismology visible amongst the European disciplinary and generic data and computing infrastructures, and in having the VERCE platform recognized as a research infrastructure for seismology in the European roadmap established by the EPOS implementation phase that has just been funded under the Call H2020-INFRADEV-1-2015-1, Topics INFRADEV-3-2015. VERCE achievements and extension have been integrated in the thematic core services of the WP8 work package: Seismology, i.e., task-8.7 led by ULIV; and will contribute to the WP7 work package: ICS design and development, while integration within the EPOS cross-disciplinary e-infrastructure framework will be one focus for the WP6 work package: ICS-TCS integration and interoperability, in coordination with WP7.

It is worth mentioning here that all the seismological partners of VERCE are strongly committed to EPOS-IP and in particular to the WP8 and WP6 work packages. VERCE sustainability is therefore included in the agenda of the ESFRI-EPOS-IP project - and of its European stakeholders - beyond the end of the project in October 2015. The initial VERCE aim at providing data-intensive computational research methods, tools, services and an e-infrastructure framework for the seismology research community in Europe can be considered to be fulfilled.

One important new feature in EPOS-IP is that the VERCE teams and the large solid earth data providers will work *together*, even though EPOS-IP will only bring a limited additional man power for the thematic core services. Recent history of VERCE shows that European coordination is essential, as well as the contribution of strong knowledgeable teams and institutions mostly funded at the national level. Therefore the required funding scheme, as in most of the ESFRIs, is a combination of EC and national funds from the EPOS stakeholders. EPOS-IP also comes with additional tasks to ensure supported coordination with the ESFRIs and their pathfinders. This will nevertheless contribute to close the gap between the required and existing support levels.

On the VERCE-VO side, an MOU will be prepared by the end of the project between VERCE/EPOS partners, and SCAI, CINECA and LRZ partners, establishing their common interest in the support of the VERCE framework within EPOS-IP, and their commitment maintaining the existing support levels during the 2016 transition period.

Another important new element in the VERCE sustainability landscape is the successful submission of the EGI-Engage and EUDAT2020 proposals to the Call H2020-EINFRA-2014-2, TOPICS EINFRA-1-2014.

- EUDAT2020, involving the VERCE partners UEDIN, INGV, CINECA, is of importance to bring additional manpower and improving the integration and the development of the VERCE data layer management and resources, based on the iRODS platform and micro services.
- The EPOS Competence Centre (EPOS-CC) - WP6/SA2.9 – in EGI-Engage involves VERCE partners CNRS, KNMI, INGV and SCAI, and will bring an additional manpower estimated to an addition of 30 PM. The VERCE FWM use-case, including the misfit data analysis component, has been selected as "Computational Use Case". This will improve the integration of EGI and EUDAT resources within the VERCE platform, taking advantage of VERCE developments integrating UNICORE resources (from CINECA) in the platform.

EPOS-CC will also bring new synergy between EGI, EUDAT and PRACE to tackle the current authorisation and accounting hurdles, which despite the homogeneous technical solutions adopted by VERCE, still have many issues outside the VERCE project's control.

Finally, the availability of HPC computing resources is limited and gaining access to such resources at the national and European levels is mostly on an individual research project basis, and may vary from one country to another. However, the new PRACE system call recognises communities, such as EPOS, which will ameliorate these restrictions. It is worth mentioning here that the VERCE user community is also quite active and visible in the PRACE science challenges, for example newly funded PRACE project "Imagine-IT" with for PI E. Casarotti, who is one of the VERCE leaders in the development of the FWM use case.

The worst impact of the lack of EC-funding for the SCI-BUS and ER-Flow follow up projects in the last H2020 Calls is to introduce a high level of risk for the sustainability and quality of the gUSE/WPS-PGRADE VERCE Science Gateway framework, services and workflow tools, which are at the very core of the VERCE framework. At least this technology stack and standards will continue to be supported by national funding level for the next few years. The adoption of these generic building blocks was strongly encouraged in the VERCE project reviewer's recommendations and the European commission and stated as "Intensify collaboration with members of the SCI-BUS project. Aim at making full use of their technology and frameworks and possibly adopt their solutions for the VERCE platform". This was indeed followed in the strategic implementation of the VERCE framework, with MoUs signed between VERCE, SCI-BUS and ER-Flow, together with a detailed collaboration roadmap.

3.4. Training and dissemination outreach

One major point for the consolidation of activities towards training and dissemination in the VERCE context has been the dense program of training webinars and workshops, together with the production of a comprehensive set of training material, and contribution, through talks and demonstrations, within a number of disciplinary and multi-disciplinary international meetings.

These activities at the European level have to be redefined within the EPOS implementation phase. The development of a European network for VERCE-related education activities, in liaison with "nearby" projects, such as the TIDES COST-project, and with the new H2020 initiatives, such as EGI-Engage and EUDAT2020, through for example the EPOS-CC. Training will also continue beyond VERCE, in the VERCE institutions for their staff and students, in particular through task 8.7 of the EPOS-IP WP8.

Finally, two papers summarizing the framework, the science tools and the achievements of VERCE were accepted for the eleventh IEEE Conference on e-Science. Those papers are of interest for other disciplines on their way to build their own interoperability framework in the international and European contexts, and especially within the EPOS implementation phase.

4. Next steps

Detailed next steps have been presented in the sections above, we only present here the over-arching project-management activities.

During the next reporting period, the PMO will continue monitoring the updates of the Risk Register. It will also continue to focus on refining the Sustainability Plan in a rapidly moving landscape, in coordination with EPOS.

It 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 other projects.

The PMO will make sure all deliverables (and milestones) are ready for the next deadline on 1st October 2015 and the final review meeting that should take place November 2015 in Paris at the CNRS or IPGP.