D-NA3.1: first defined documentation and training

21/05/2012

Project acronyme:	VERCE
Project n°:	283543
Funding Scheme:	Combination of CP & CSA
Call Identifier:	FP7-INFRASTRUCTURES-2011-2
WP:	WP3/NA3, User documentation and training
Filename:	D-NA3.1.pdf
Author(s):	X. Wang (ULIV), A. Rietbrock (ULIV), M. Galea (UEDIN)
Location:	http://www.verce.eu/Repository/Deliverables/RP1/
Type of document:	Deliverable
Dissemination level:	Public
Status:	Final
Due date of delivery:	15/05/2012
Reviewer:	A. Frank (LRZ)
Keywords:	NA3, Documentation, Training

Version	Author	Date	Comments
0.11	X. Wang (ULIV)	09/04/2012	Initial draft for comments
0.12	M. Galea (UEDIN)	12/04/2012	Updating and comments
0.13	A. Rietbrock (ULIV)	15/04/2012	Updating
0.2	X. Wang (ULIV)	15/04/2012	Internal review draft
0.21	A. Frank (LRZ)	24/04/2012	Internal review
0.22	X. Wang (ULIV)	28/04/2012	Accept the review and extend chapter 4
0.23	X. Wang (ULIV)	07/05/2012	Finalize after F2F meeting
0.24	X. Wang (ULIV)	08/05/2012	Add Glossary and extend Executive Summary
0.25	A. Rietbrock (ULIV)	09/05/2012	Review executive summary
0.26	X. Wang (ULIV)	09/05/2012	Re-edit
0.27	X. Wang (ULIV)	14/05/2012	Harmonize Glossary, Keywords and author list
0.3	X. Wang (ULIV)	14/05/2012	Change the dissemination level and status, release final version
0.31	X. Wang (ULIV)	15/05/2012	Update cover page, centre images and captions
0.32	X. Wang (ULIV)	16/05/2012	Update Glossary and tables

Copyright notice

Table of Contents

List of figures

List of tables

Executive summary

To encourage and support the use of the VERCE infrastructure, the NA3 (Training and user documentation) work package is defined to provide users and developers of the earthquake and seismology community with appropriate documentation, tutorial examples, web accessible self-tuition material and training in the use and exploitation of the VERCE platform. The main aim of this reporting period is to define the development strategy of NA3 and outline procedures of the initial work plan.

Chapter 1 outlines the NA3 development strategy. NA3 will collect necessary information for developing the training and documentation program from surveys and direct input from NA2 and SA3. After analysing this information, the documentation and training program will be developed and updated on a regular basis. Additional members from other WPs will be invited to review the documentation and training program. With the review result, NA3 will collect further information or update documentation and training. Formal information will be delivered to NA4 for a timely dissemination. The documentation and training will be constantly updated during whole project period. Two types of training needs are identified: internal and external project training. The internal project training is aimed at VERCE project members to enable them to develop and share the skills necessary to build an integrated e-Infrastructure for seismology. The external project training is aimed at the earthquake seismology community. It targets both existing and potential users of the evolving e-Infrastructure, to encourage and enable them to understand how to pursue their research and advance science using the added benefits of the VERCE e-Infrastructure. We also make use of external training available by other relevant EU projects such as PRACE, EGI training marketplace, QUEST and Munich Winter School programme, and EUDAT wherever possible.

The first NA3 survey focused on the internal requirements of the project. The questionnaire is designed to understand the requirements of internal project members to gain additional knowledge and skills. Additionally, the questionnaire covers the possibility of documents and training events that could be provided from project partners as well. The result of survey is analyzed in Chapter 2.

Following the analysis of the survey, the development plan is illustrated in Chapter 3. Four mediums are adopted to support user documentation and training: 1) a range of documents will be provided on the project website. 2) a series of workshop will be organized. 3) a set of videos will be put on the project website that includes usage of the computing facilities and webinars collected and recorded from EU workshops and VERCE project workshops. 4) questions and training requests can also be addressed via a helpdesk.

The work in the next six months is described in Chapter 4. NA3 Wiki and repository will be updated to reflect the carried out work. The first internal training workshop will be organized in September 2012. We will cooperate with NA2 and SA3 to identify a document list of necessary knowledge and skills, VERCE scientific gateway usage, and publish the documentation on the VERCE webpage. We will look in detail into training events offered by other EU projects and notify VERCE project member. We will also make the workshop and documentation information available to NA4 for dissemination to the wider community.

1. Introduction

1.1. Role of NA3

The mission of the NA3 work package is to provide users and developers of the earthquake and seismology community with appropriate documentation, tutorial examples, web accessible self-tuition material and training in the use and exploitation of the VERCE platform. NA3 collects the necessary knowledge and skills from NA2 and the user manual of the VERCE platform and its user interface from SA3. NA3 also provides all training/documentation-related information to NA4 for dissemination purposes.

1.2. NA3 Strategy

NA3 is responsible for VERCE training and user documentation. The necessary information for developing the training and documentation program is collected from surveys and direct input from NA2 and SA3. After analysing this information, the documentation and training program will be developed and updated on a regular basis. Additional members from other WPs will be invited to review the documentation and training program. With the review result, NA3 will collect further information or update documentation and training. Formal information will be delivered to NA4 for a timely dissemination. The documentation and training will be constantly updated during whole project period. The NA3 strategy is illustrated in Figure 1.

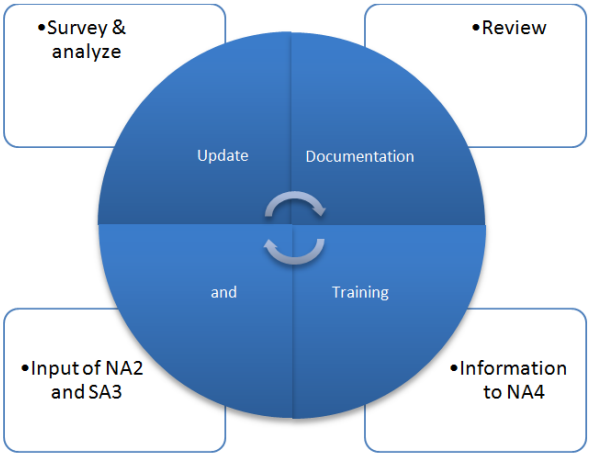

Figure 1. NA3 strategy workflow

Two types of training and documentation are identified: a) internal project training and b) external user and community training.

The internal project training is aimed at VERCE project members to enable them to develop and share the skills necessary to build an integrated e-Infrastructure for seismology. The external user and community training is aimed at the earthquake seismology community. It targets both existing and potential users of the evolving e-Infrastructure, to encourage and enable them to understand how to pursue their research and advance science using the added benefits of the VERCE e-Infrastructure.

User documentation is provided in two phases. In the first phase, the user documents will support the VERCE project members of seismologists and IT developers to understand the computing facilities and the requirement of processes and practices used in seismology. With the release of the platform and user interface from SA3, the user manual of how to use the platform and interface will be added in the user documentation, in addition to the previous material. Project members and external users from the seismological community will benefit by using the provided documentation.

As part of NA3 strategy, we make use of external training available by other relevant EU projects such as PRACE, EGI training marketplace, QUEST and Munich Winter School programme, and EUDAT wherever possible.

In cooperation with NA4, NA3 could also provide tutorials about the project and platform at external workshops or conferences.

Normally the workflow in Figure 1 will be cycled each half year, but documentation will be updated quickly when a new platform/module is released. A notification email will be circulated to project mailing list before the work plan and documentation of NA3 is updated. Project members will therefore be aware of upcoming surveys and contribute or receive respective training material (e.g., documents, manuals, tutorials).

2. Documentation and training requirement

The initial documentation and training provision has to meet the requirements of the project members. The NA3 surveys are designed to understand the requirements needed by both the IT and seismology specialists present in all work packages.

2.1. First NA3 survey result

The first NA3 survey focused on the internal requirements of the project. The questionnaire is designed to understand the requirements of internal project members to gain additional knowledge and skills. Additionally, the questionnaire covers the possibility of documents and training events that could be provided from project partners as well.

The survey started from Friday, 2nd March 2012 and ended on Friday, 16th March 2012. We received inputs of 14 project members from INGV, LRZ, LMU, UEDIN, EMSC and CINECA. The input partners are illustrated in Figure 2. The full survey input data can be accessed from link of http://www.verce-project.eu/projects/verce1/repository/changes/verce/All/NA/NA3/Survey/First-NA3-Survey.xls.

We also contacted project partners who did not respond to this survey to check whether they do have training requirements and investigated how to meet their requirements. The process to ensure that we consider all requirements from all project partners will be an ongoing effort.

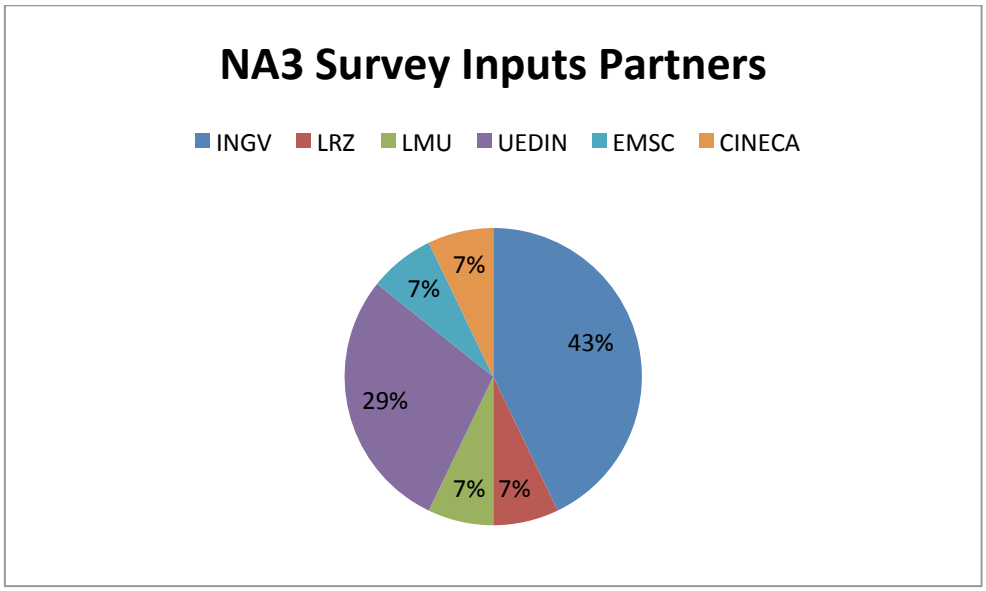

Figure 2. NA3 survey input partners

The survey result shows that more IT knowledge and skills are required in the initial project stage as outlined in Figure 3. The detailed list is described in Table 1. The majority of requirements are not urgent, only the introductory course on seismology in question 8) is required as soon as possible.

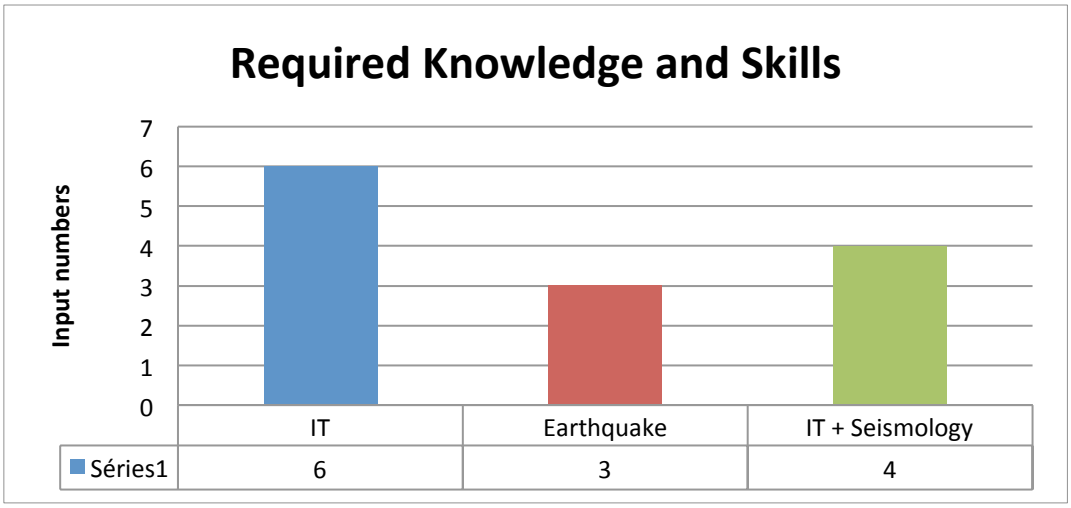

Figure 3. Required knowledge and skills

1). To use HPC
- Compilers and compiler options, MPI flavours
- Memory restrictions of various HPC machines
- Batch schedulers
- Available optimised libraries on respective HPCs
- Grid tools

2). To use Grid resources (e.g. EGI)
- Setup and maintenance of a Grid infrastructure
- Grid tools
- Compiler and compiler options
- Batch schedulers

3). Integration of codes, tools and services
- Information about the code
- How to integrate
- Restrictions and requirements of the code (compiler, memory, optimised?)
- Requirements
- Dependencies
- Security risks
- ADMIRE

4). To use VERCE infrastructure
- Basic usage of seismology software relevant to VERCE

5). Using distributed data and computing. Efficient parallelisation techniques for high volume computing. Interface protocols. Web services

6). For the HPC use cases and associated pilot applications
- Numerical methods for wave propagation (SEM, FD, etc)
- Basic knowledge about parallel programming
- Geophysical models and model generation (i.e. computational grid generation)
- EQ data access, analysis and visualization (e.g., with ObsPy)

7). IT (profiling, optimization, multicore-GPU programming)

8). Knowledge:
- Introductory course on seismology research so as to better understand requirements discussion.
- The HPC processes and practice used in seismology applications.
- Available and planned low-level e-Infrastructures in Europe on which VERCE can build.
- An overview of the current data and metadata standards used by seismologists.

Table 1. List of required knowledge and skills

The preferred training and documentation mediums are outlined in Figure 4. The feedback of preferred training and documentation mediums expresses that attending workshop and watching/downloading documents from the website have top priority, while watching training material via video, solving problem via helpdesk and mailing list are followed.

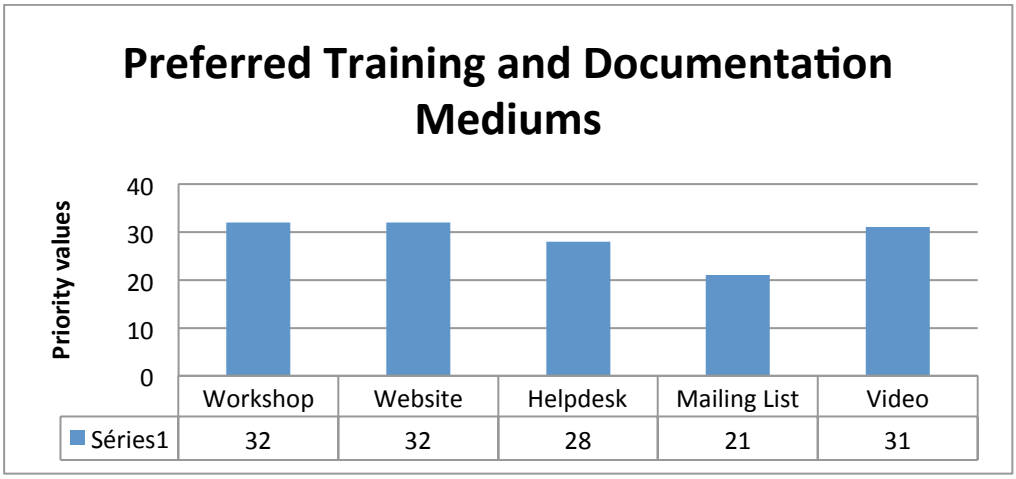

Figure 4. Preferred training and documentation mediums

While new skills are required for some project members, other project members can contribute their knowledge in types of documents and training. They also can cooperate to organize workshops to disseminate the necessary know-how. Table 2 lists the contributed knowledge that will actually help the requirements of other partners.

1). How to use HPC, SuperMIG/SuperMUC
2). Compilers, MPI and OpenMP
3). Batch schedulers: Loadleveler, PBS and SLURM
4). Optimised libraries, e.g. HDF5 and MPI IO
5). GLOBUS tools (via IGE if possible)
6). Procedure to compute crustal velocity variations by means of cross correlation of seismic noise
7). SPECFEM3D meshing (CUBIT)
8). Knowledge: data-intensive architectures, technologies and methods
9). Computer science: API to access Web services with the collaboration of ORFEUS
10). Profilers, e.g. SCALASCA
11). UNICORE, iRODS

Table 2. List of contributed knowledge and skill

3. Work plan

With the analysis of the survey result, four mediums are adopted to support user documentation and training:

1) A range of documents will be provided on the project website.

2) A series of workshop will be organized.

3) A set of videos will be put on the project website that includes usage of the computing facilities and webinars collected and recorded from EU workshops and VERCE project workshops.

4) Questions and training requests can also be solved via helpdesk.

However, the above four mediums are not exclusive as they complement each other. More than one option allows users to achieve the skills needed in their best way.

3.1. Documentation

3.1.1. Proposed document list

In this initial stage, we aim at providing required documentation to the IT developers and seismologists to provide background knowledge for developing the VERCE e-infrastructure platform. The proposed document list is displayed in the following enumeration:

1) VERCE platform

a) Use case and pilot application: numerical methods for wave propagation, geophysical models and model generation, earthquake data access, analysis and visualization.

b) Integration of application code: code information, dependencies and integration methods.

c) Integration of tools and services: requirements, dependencies and security risks.

2) General information

a) Grid infrastructure: setup and maintenance, workflow, data-intensive architecture and introductory examples.

b) HPC usage: compiler, parallel computing, batch scheduler, grid tools, interface protocols and web services.

c) Seismology: seismology research introduction, overview of current data and metadata standards.

3.1.2. Document access

The document list is set up on the VERCE website which can be accessed using the link http://www.verce.eu/Training/KnowledgeBase.php.

As these documents are hosted by project partners and external third parties, we only provide a document name with a link. We leave all maintenance and updating work to the project partners and third parties where appropriate.

At the time of the project's own platform releases, the documents of applications running on the platform will be uploaded to the project website and will be hosted and updated by NA3. In addition, NA3 will take the maintenance and updating work that will last for whole project period.

3.2. Workshops

3.2.1. Workshop plan

VERCE training workshops are organized annually. Two types of workshop are identified: internal project workshops and combined internal and external project workshops. The internal project training workshops are planned to start this year. For the combined internal and external project training workshops the appropriate time for introducing the VERCE platform will be discussed among the project partners.

External sources for training contributions (e.g., the PRACE Advance Training Centres (PATC), EGI Training Marketplace, QUEST and Munich Winter School programme) will be taken into account.

3.2.2. Proposed first workshop

The first internal project workshop will be at Liverpool this summer or autumn. We will combine IT-related and seismology-related training to enable both communities participating in VERCE to develop one common language in order to ease the developments on the VERCE platform and the deployment of initial use cases. The discussion about the workshop programme will start very soon now.

3.3. Helpdesk

Project members are allowed to request training and document related questions. We aim to adopt existing running tool within VERCE project website to implement this function. Under the New Issue section of Redmine, a new tracker Training and a new Watcher of NA3 Trainer are added. When members ask question through New Issue section, the ticket will be allocated to NA3 Trainer that will be redirected to some responding answers to process this ticket. The possibility of coordinating of the helpdesk system with the helpdesks of the European e-Infrastructures (EGI, PRACE) will be investigated.

3.4. Videos and webinars

For providing training videos the following steps are planned:

1) Project partners are encouraged to contribute related training material videos.
2) Training videos from other EU projects will be collected.
3) The VERCE project will record its own videos.
4) Webinars will be recorded and made available after the webinar dates.

4. Future work

In the next six months, the following tasks are proposed:

1) Keep NA3 Wiki and repository updated

   NA3 Wiki and repository will be updated to reflect the carried out work.

2) Organize first internal project training workshop

   We hope as many as possible project members can attend this workshop. The doodle tool is adopted for them to select date they can attend. So the workshop will be organised on the date of majority project members can attend. The workshop programme will be discussed soon. Once the preparation is ready, we will notify the participants by using project mailing list and project webpage.

3) Cooperate with NA2 and SA3 to identify a document list

   Once NA2 figures out the research priorities and SA3 releases template of gateway user interface, communication will be built with these two WPs to identify document list and document providers.

4) Regularly update document list on project webpage

   Every modification of document list will be reflected under Knowledge Base tab of VERCE project webpage.

5) Identify appropriate training opportunities offered by other EU projects and e-Infrastructure programs

We will look in detail into training events offered by other EU projects and notify VERCE project member.

6) Notify NA4 about workshop and documentation information

The workshop and documentation information will be available to NA4 for dissemination to the wider community.

Glossary and Links

CINECA	Consorzio Interuniversitario Cineca
EGI	European Grid Infrastructure – http://www.egi.eu
EMSC	Euro-Mediterranean Seismological Centre
EQ	EarthQuake
EUDAT	EUropean DATa is a project currently funded by the EC for the development of the Common Data Interface - http://www.eudat.eu
FD	Finite-Difference wave propagation
GPU	Graphics Processing Unit
High-performance computing (HPC)	Use of powerful processors, high-speed networks and parallel supercomputers for running computationally intensive applications
IGE	Initiative for Globus in Europe
INGV	Istituto Nazionale di Geofisica e Vulcanologia
iRODS	Integrated Rule-Oriented Data-management System - https://www.irods.org/
LMU	Ludwig-Maximilians-Universitaet Muenchen
LRZ	Leibniz-Rechenzentrum
MPI	Message Passing Interface
NA2	Equivalent to Work Package 2 (WP2)
NA3	Equivalent to Work Package 3 (WP3)
NA4	Equivalent to Work Package 4 (WP4)
OpenMP	Open Multi-Processing
ORFEUS	Observatories and Research Facilities for European Seismology
PBS	Portable Batch System
PRACE	Partnership for Advanced Computing in Europe - http://www.prace-project.eu/
QUEST	QUantitative Estimation of Earth's Seismic Sources and Structure
SA3	Equivalent to Work Package 7 (WP7)
SCALASCA	A software tool that supports the performance optimization of parallel programs by measuring and analyzing their runtime behaviour
SEM	Spectral Element Method wave propagation
SLURM	A high scalable resource manager

SPECFEM3D	A software to simulate seismic wave propagation in sedimentary basins or any other regional geological model
SuperMIG/SuperMUC	The name of a new supercomputer of the LRZ
UEDIN	The University of Edinburgh
UNICORE	Interface to Computing Resources - http://www.unicore.eu/