D-JRA1.2.1: Enabling pilot applications: first report and validation of the VERCE architecture

Thursday 18th October, 2012

Project acronym:	VERCE
Project n°:	283543
Funding Scheme:	Combination of CP & CSA
Call Identifier:	FP7-INFRASTRUCTURES-2011-2
WP:	WP8/JRA1, Harnessing data-intensive applications
Filename:	D-JRA1.2.1.pdf
Author(s):	M.Simon, H. Igel
Location:	http://www.verce.eu/Repository/Deliverables/RP2/
Type of document:	Deliverable
Dissemination level:	Restricted
Status:	Final
Due date of delivery:	01/10/2012
Reviewer:	Alberto Michellini
Keywords :	HPC, seismology, e-science environment, service-oriented architecture, Grid, earthquake, data infrastructure, seismology, report

---

Version	*Author*	*Date*	*Comments*
1	M. Simon , H.Igel LMU	14/09/2012	Initial draft for comments
2	M. Simon LMU	27/09/2012	Corrections after review
3	M.Simon LMU	28/09/2012	Corrections after second 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.

Contents

List of Figures

List of Tables

1 Executive Summary

The role of JRA1 is to take the NA2 pilot cpu- and data-intensive applications and their software implementation through to 'productisation', enabling them to be used with the model and architecture developed by JRA2, leading to an improved use of the applications by the community. It will establish a prioritised list of technical goals to be met by the architecture and high-level services, which will be remitted to JRA2. It will deliver initial analyses and lists of components and tools that should be ingested by SA2. It will identify required presentational views, standard assemblies of services and their descriptions that should be developed by SA3. It will establish operational criteria that will need to be addressed by SA1 before scientific goals can be met. It will provide NA4 with a roadmap outlining how the e-Infrastructure can be improved and therefore empower scientists. This roadmap will also identify and inform NA3 of internal and external training requirements.

For this reporting period, the objective, as described in the DoW, is to analyze, improve and check the methods and their implementation. During the project, the JRA1 experts will ensure continuous number of adaptation and optimization cycle to productise the pilot applications on the platform. For the first release cycle, the suggested applications have been evaluated, in close collaboration with the responsible work packages and deployed where the applications were accepted. The JRA1 domain experts were assisting in installing and setting up tests to verify the installation. During the first release period two applications have been accepted by SA2, namely the numerical wave propagation application Seissol as well as the data handling and analysis code Obspy. Extensive feedback was given by SA2 on the other prioritized applications 1) Specfem3D, 2) Ses(c)3D 3) Axisem and 4) Comcot which need to be improved to match the requirements for the VERCE platform during this release cycle and the efforts are currently in progress.

A performance analysis scheme for the cpu-intensive use case applications has been developed by JRA1 (CINECA) and it has been started to be applied to pilot applications, starting with Specfem3D. It will continue to be applied to Seissol, Ses3d, Sec3d and Axisem and Comcot as well. Additionally, the applications will be supplied with a regressive testing scheme to ensure performance in accuracy as well, where not already in place. Furthermore a benchmark for the 'Ambient Noise Cross-Correlation' use case applications will be set up, in order to validate performance as well as a reference for the implementation in the VERCE platform.

Furthermore the work flows of the proposed use cases have been further analyzed and broken down into detail. A partial prototype implementation of the 'Forward Modelling and Inversion', focussing on the forward modelling part has been set up to run on the Verce test bed, involving certificate based authentification, remote job submission/file transfer to and from HPC site SuperMUC using the Globus toolkit as well as remote visualisation. Additionally the ingestion of Python/Obspy routines into the VERCE workflow enactment engine DISPEL was demonstrated and analyzed.

In the upcoming release period JRA1 will continue to work on modifying the application to fulfill the requirements based on the feedback received from the other work packages as well as continue to develop and analyze the workflow and respective requirements.

2 Introduction

In the course of the first phase of this reporting period the implementation of the selected use case has started and JRA1 is now in the process of integrating it into the envisioned VERCE architecture. A first foundation was laid out and for the upcoming months JRA1 will be responsible for mapping the applications to the architecture in an iterative feedback cycle, formulating requirements to the architecture and modifying the applications to suite the possible solutions. A first validation of the current release was successfully performed and JRA1 will continue in assistance and development of the applications. In this mid-term report we will summarize the past, current and future efforts during this stage.

3 Content of the Report

3.1 Refactoring/Reengineering

3.1.1 Providing applications to SA2, response to feedback on applications

In the first phase of this release cycle JRA1 supported SA2 in deploying and testing the pilot applications and tools, resulting from the first deliverable (i.e. Obspy, Seissol, Specfem, Sec3d, Ses3d, Axisem and Comcot) on the VERCE resources. For the initial release, only Obspy, Seissol were accepted immediately while the other applications were delayed due to failure of meeting the requirements for immediate deployment, which was to be expected. For a detailed review of the performed tests, please refer to the SA2 deliverable. So far, the issues have been identified and JRA1 has started to re-engineer and modularize the failing applications to meet the requirements for deployment, starting with Ses3D, Specfem and Comcot.

An overview over the activities:

- Support for SA2 in installing and testing the following applications: Obspy, Seissol, Specfem3d, Sec3d, Ses3d, Axisem (Sec3d, Ses3d, Axisem, were delayed due to failure of meeting the requirements). JRA1 has started to re-engineer Ses3d to modularize the code (S. Mauerberger in cooperation with G. Brietzke), adpoting parallelization strategies suitable for the VERCE platform.
- Test setups for Seissol have been developed and provided to other partners. In particular a mesh ($\sim$1.5 million tetrahedral elements) and material model of the European region for a test simulation of the 2009 L'Aquila earthquake have been tested on SuperMIG and PLX. The tests were successful and only generic requirements (Documentation and Version number) had to be delivered subsequently.
- A test package for Obspy has been provided and tested.
- Test packages for forward simulations using Specfem3d have been provided to SA2. In particular, a relatively small test case (mesh of about 0.2 million of hexahedral elements and about 13.5 million of grid points) has been tested on PLX (Graziella Ferini - CINECA) and a larger test case (mesh of about 1.6 MHex and about 104 M of grid points) has been tested on PLX, SuperMIG and SuperMUC (Siew Hoon Leong - LRZ, and David Weissenbach - IPGP). Tests performed with Specfem3d on SuperMIG and SuperMUC were successful.
- Planned possible changes in the code Specfem3d to facilitate setting of some simulation parameters that are presently hard coded.
- Linear module parallelisation of COMCOT code by menas of MPI in shared memory configuration and analysis of performances by means of SCALASCA (Graziella Ferini - CINECA, Fabrizio Romano - INGV)

- COMCOT code has been debugged, its portability extended and uploaded to the VERCE repository (Graziella Ferini - CINECA, Fabrizio Romano - INGV)
- Planned linear and nonlinear modules parallelisation of COMCOT code by means of MPI in distributed memory configuration by using the domain decomposition technique (Graziella Ferini - CINECA, Fabrizio Romano - INGV)

3.2 Workflow development

3.2.1 The work flow for CI and DI intensive use case have been further implemented using either DISPEL (DI) or Python (CI)

The prototype of the workflow of the 'e-tomo' use case has been further elaborated and successfully implemented in Python using the approved applications and tools deployed on the VERCE test bed. The demonstration of the functionality of the current implementation was successfully performed during the VERCE Developers Training in Liverpool. The current workflow prototype implementation is relying on Web-Services (NERIES, ArcLink) and other online resources (GlobalGMT) to gather most information neccessary to initiate a forward simulation. However, the mesh is still to be supplied manually to the server, but partitioning using Metis is already being done on the server. Job submission as well as file transfer have been implemented using Globus Tools (GridFTP and GRAM) and certificate based authentication to the VERCE-test server. Although only job submission and file transfer in between the VERCE-test server and SuperMUC was demonstrated, due to the certificate based nature of authentication, this functionality is extensible to other compatible resources in the PRACE and EGI initiatives as well as other supporting resource providers. In a next step, the post processing unit of the workflow will be transfered to run on a cluster-like ressource. Since job submission protocols might differ, the workflow will be extended to be compatible with those. A first design to a graphical user interface for initialisation and Visualisation of the output has been set-up and the remote visualisation possibilites using paraview via TightVNS have been demonstrated.

Additionally, the steps of the workflow have been worked out in detail (forward modelling and inversion workflow) together with SA2 and the requirements were given to JRA2 within the established task force as can be seen from the Fig.1 and Fig.2.

3.3 Validation of the VERCE architecture

3.3.1 Validation of VERCE test bed

When considering the VERCE architecture as the current state of the test bed and the tools suggested by SA2, an initial limited validation has been successful using and demonstrating gridftp for file transfer and GRAM for HPC access through the VERCE test server. The capabilities of the Globus toolkit seem very promising and well suited for the current state of the test bed (Verce-test server and SuperMUC) although other certificate based grid-middleware tools (Unicore, gLite etc) can also be supported, if necessary. The suggested component JSAGA seems to deliver the support for all of these. Also first promising test on the visualisation resources were successfull and the integration of those into the current prototype will be further pursued. The first validation of the integration of the DISPEL workflow enactment engine into the VERCE test bed has been demonstrated on EDIM1. Another deployment on the ULIV Linux cluster is currently pending in test status. As a next step it is suggested to integrate the pilot applications proposed by NA2 and to run a benchmark of to test the performance and deployment.

3.4 Validation of VERCE Components

- The suggested applications support the requirements for the 'Forward and Inverse Simulation' use case.

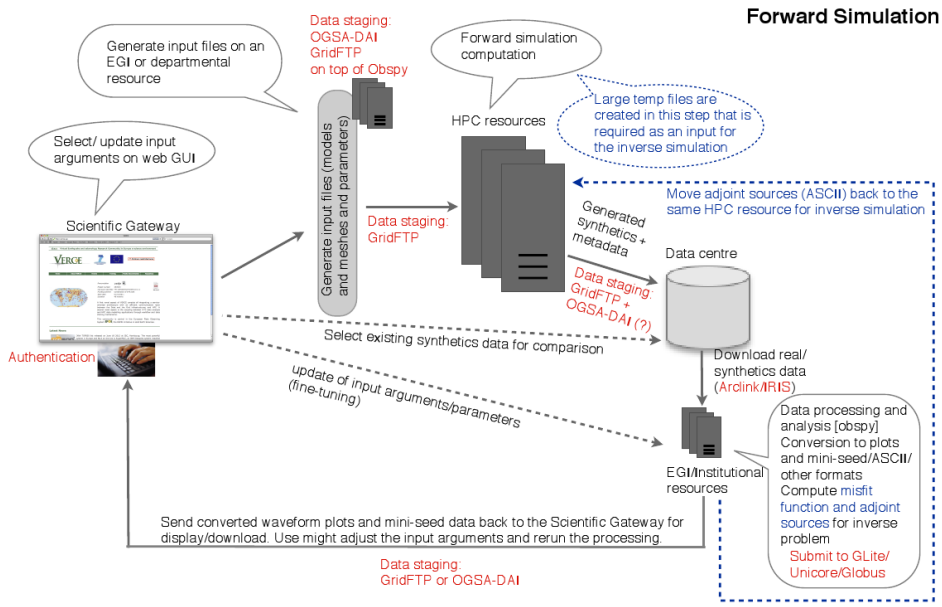

Figure 1 – Forward Modelling Workflow

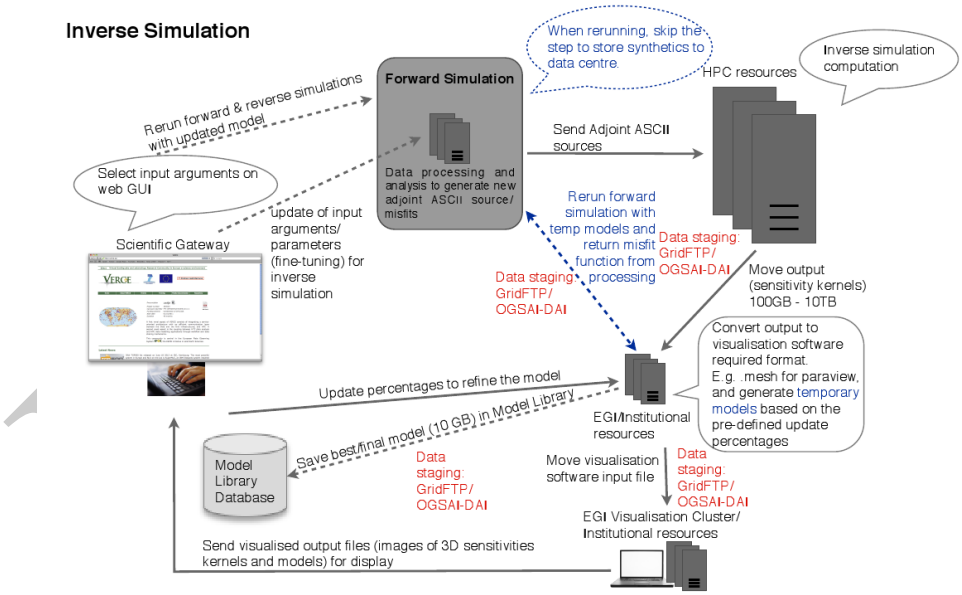

Figure 2 – Inverse Modelling Workflow

- An issue with the Obspy instrument correction, diverging from SAC instrument correction has been identified and elaborated (Kasra Hosseini, LMU).
- The applications for the 'Ambient noise cross correlation' use case have been provided to JRA2 for implementation in the DISPEL framework (INGV).
- JRA1 has started with assistence of JRA2 in integrating the Obspy-functionalities into the DISPEL framework.

3.5 Performance measurements of cpu-intensive (CI) and data-intensive (DI) applications

Please also refer to the Deliverable of SA2.

3.5.1 Benchmarking for CI case in place - regressive testing

Tests in progress with Specfem3d on PLX as well as on SuperMUC. Tests for performance scaling of the code still need to be completed.

3.5.2 Benchmarking for DI required

The necessity of setting up a benchmark suite for the DI use case has been identified at the Liverpool meeting. Such a benchmark should contain the data transfer protocols and a representative data volume with which the performance of the underlying applications (e.g., cross-correlations) can be quantified. The benchmark is work in progress (INGV, LMU, IPG, UEDIN).

3.6 Productisation: Regressive testing

In order to ensure proper development during the re-engineering of the applications, regressive testing schemes to evaluate accuracy and proper implementation of the changes, will be integrated into the selected pilot applications.

3.6.1 Ses3d/Sec3d

The effort to implement a regressive testing scheme for Ses3d has started and has already produced an inital working version.

3.6.2 Axisem

Regressive testing for Axisem has been been implemented as well as additional features. Initial bugs in the post processing routines have been fixed. The code is now able to run on up to 256 cores.

3.6.3 Obspy and Seissol

Seissol and Obspy both rely on a regressive testing scheme using over-night builds and comparison with benchmark results to ensure accurate results during the development process. These regressive tests return data on performance, quality control on a daily base.

3.6.4 Specfem3d

Setting up regressive testing scheme for Specfem3d was also planned. So far each installation is verified by comparison with a benchmark example manually. However, according to the developers, they will not continue to support this version since there will be a new GPGPU-version available soon and supported.

3.7 Other Activities

3.7.1 QUEST Discussion

During the 3rd workshop of the training network QUEST VERCE was introduced to an audience of approx. 100 international seismologists. A discussion followed on how the VERCE platform could best help the community.

3.7.2 JRA1-steered Meeting

A two-day meeting with a specific focus on the implementation of the use cases prioritized by NA2 was held in Munich, June 24th and 25th, 2012. The meeting led to the definition of two task forces designed to optimize the implementation of cpu-intensive and data-intensive use cases.

3.7.3 Meeting with representatives of KNMI

The requirements for the generated synthetic waveforms produced in the 'Forward Simulation and Inversion ' use case were discussed with Luca Trani and Reinoud Sleeman from KNMI aiming at an integration in an extendend Data Quality Evaluation Platform and hosting of synthetic waveforms alongside real data at the data centers.

4 Glossary and Links

2D-surface Wavefields	A collection of "synthetic seismogramms" on every-surface point of the simulated geographical region
ArcLink	A protocol for data transfer from geographically distribute data archives based on time windows - http://www.seiscomp3.org/wiki/doc/applications/arclink
DI	data-intensive
CI	cpu-intensive
Dispel	Verce Workflow Enactment Engine
Earth Model	Assumed one to three dimensional parameter sets of the earth's interior on which a simulation is based.
EIDA	European Integrated Data Archives infrastructure - http://www.verce.eu//ITCoordinationMeetingFebruary2012/EIDA-Overview.pdf
EMSC	Euro-Mediterranean Seismological Centre
Event-data	Catalog entry, from an event-catalog (usually < 10lines ascii or QuakeML), including location, magnitude and/or moment values and type of an earthquake.
Forward Simulation	Simulation of seismic wave-propagation, results in synthetic seismograms.
Full-Waveform Inversion	Tomographic inversion of the real seismograms (or differences between real and synthetic seismograms) to determine the underlying earth model.
globalCMT	Global Centroid-Moment-Tensor Project - http://www.globalcmt.org/
Globus Toolkit	An open source software toolkit used for building grids - http://www.globus.org/toolkit/
GRAM	Job submission and control function of the Globus Toolkit - http://www.globus.org/toolkit/docs/2.4/gram/
GridFTP	GridFTP is an extension of the standard File Transfer Protocol (FTP) for use with Grid computing - http://www.globus.org/toolkit/docs/latest-stable/gridftp/
high-performance computing (HPC)	Use of powerful processors, high-speed networks and parallel supercomputers for running computationally intensive applications.
IRIS	Incorporated Research Institutions for Seismology (Data-Center)
JSAGA	A Java implementation of the Simple API for Grid Applications (SAGA) specification from the Open Grid Forum (OGF) - http://grid.in2p3.fr/jsaga/
NERIES, NERA	Seismological I3 projects supported by the EC
ObsPy	A Python framework for processing seismological data - http://obspy.org

PRACE	Partnership for Advanced Computing in Europe - http://www.prace-project.eu/
QUEST	an Initial Training Network in computational seismology funded within the 7th Framework People Programme by the European Commission - www.quest-itn.org
RapidSeis	Portal for interactively running C++ scripts on seismological waveform data Not yet ready for Python.
Real (or observed) Seismograms	Data recorded at one or more seismic stations and made available by data-centers (size of data depends on duration and sampling rate, also meta-data)
Regressive Testing	A regression test aims to uncover regressions (bugs) in existing area of a system (computational code) after changes has been made.
SEED, mSEED, SAC	Standard seismic data formats
STF	Source Time Function. Shape of the original "signal-wave" of the earthquake, calculated by adjoint source inversion (or other methods). The "signature" of the STF is present in the recorded seismograms.
metadata	Data that describes data. Metadata may include references to schemas, provenance, and information quality. In Seismology, metadata may also refer to data required in order to sanitise a seismograph's response.
Synthetic Seismograms	Waveform(time series) calculated in a computer simulation (size of data depends on duration and sampling rate, also meta-data). It is dependent on the solver, the computational grid(mesh), the earth model, the event parameters, and the location of "observation".
TightVNC	A free remote control software package. With TightVNC, you can see the desktop of a remote machine and control it with your local mouse and keyboard, just like you would do it sitting in the front of that computer - http://www.tightvnc.com/
WP leader	The institution that has the responsibility for a certain WP, not the single person. e.g. JRA1 leader is LMU.
NA2	WP2
NA3	WP3
NA4	WP4
SA1	WP5
SA2	WP6
SA3	WP7
JRA1	WP8
JRA2	WP9