SBGf 30 anos

EIAGRID: In-field optimization of seismic data acquisition by real-time subsurface imaging using a remote GRID computing environment.

Björn Z. Heilmann*, Ana M. Vallenilla Ferrara*, Guido Satta*, and Ernesto Bonomi*
*CRS4:SURL, Imaging and Numerical Geophysics

This paper was prepared for presentation at the 11th International Congress of the Brazilian Geophysical Society, held in Salvador, Brazil, August 24-28, 2009.

Summary

The constant growth of contaminated sites, the unsustainable use of natural resources, and, last but not least, the hydrological risk related to extreme meteorological events and increased climate variability are major environmental issues today. Finding solutions for these complex problems requires an integrated cross-disciplinary approach, providing a unified basis for environmental science and engineering. In computer science, grid computing is emerging worldwide as a formidable tool allowing distributed computation and data management with administratively-distant resources. Utilizing these modern high performance computing technologies, the GRIDA3 project bundles several applications from different fields of geoscience aiming to support decision making for reasonable and responsible land use and resource management. In this abstract we present the geophysical application of the GRIDA3 project named EIAGRID. As illustarted in Figure 1, it aims at supporting environmental research by enabling the use of grid computing facilities to provide real-time subsurface imaging directly during data acquisition. Furthermore, shared data storage using georeferencing tools and data grid technology facilitates acquisition design, workflow construction and the interpretation of results.

Introduction

Seismic reflection profiling has a broad application range spanning from shallow targets at a few meters depth to targets at a depth of several kilometers. Nevertheless, it is primarily used by the hydrocarbon industry and hardly for environmental purposes. The complexity of data acquisition and processing poses severe problems for environmental and geotechnical engineering: Professional seismic processing software is expensive to buy and demands large experience from the user. In-field processing equipment needed for real-time data Quality Control (QC) and immediate optimization of the acquisition parameters is often not available for this kind of studies. As a result, the data quality can be suboptimal. In the worst case, a crucial parameter such as receiver spacing, maximum offset, or recording time turns out later to be inappropriate and the complete acquisition campaign has to be repeated.

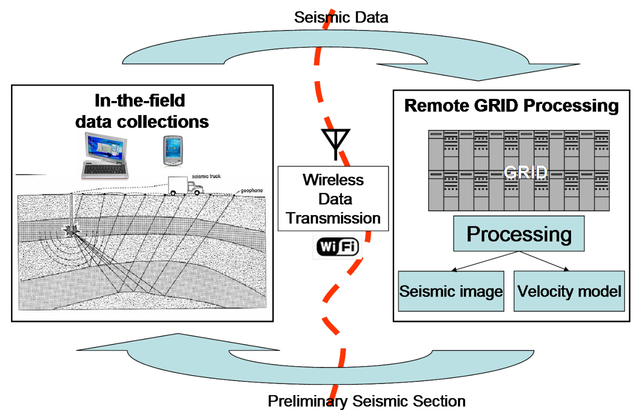

Figure 1: Main scheme of the the EIAGRID service: In field-processing equipment is substituted by remote access to high performance grid computing facilities which can be controlled by a user-friendly web-browser interface, accessed from the field using wireless data transmission technology.

The EIAGRID portal provides an innovative solution to this problem combining state-of-the-art data processing methods and modern remote grid computing technology. In-field processing equipment is substituted by remote access to high performance grid computing facilities. The latter can be ubiquitously controlled by a user-friendly web-browser interface accessed from the field by any mobile computer using wireless data transmission technology. The complexity of data-manipulation and processing and thus also the time demanding user interaction is minimized by a data-driven, and highly automated velocity analysis and imaging approach based on the Common-Reflection-Surface (CRS) stack. Furthermore, the huge computing power provided by the grid deployment allows parallel testing of alternative processing sequences and parameter settings, a feature which considerably reduces the turn-around times.

A shared data storage using georeferencing tools and the data grid technology iRODS (i Rule Oriented Data System) is under current development. It will allow to publish already accomplished projects, making results, processing workflows and parameter settings available in a transparent and reproducible way. Creating a unified georeferenced database shared by all users will facilitate complex studies and enable the use of data-crossing techniques to incorporate results of other environmental applications hosted on the GRIDA3 portal.

In the following we discuss the three principal tasks of the EIAGRID service: data upload and conversion, data visualization and preprocessing, and seismic reflection imaging.

Data upload and conversion

A typical seismic 2D survey is carried through by applying a multitude of so-called common-shot experiments. The EIAGRID service allows to upload the recorded data to the remote computing facilities immediately after each shot. This is done via wireless data transmission and a web-browser-based Graphical User Interface (GUI). In this way a single shot-gather or the complete range of shots can be preprocessed and visualized while the acquisition still takes place. For the wireless data transmission (see, e.g., Halonen et al., 2003) high speed protocols such as EDGE (Enhanced Data Rates for GSM Evolution), UMTS (Universal Mobile Telecommunications System), or HSUPA/HSDPA (High-Speed Uplink/Downlink Packet Access) are required, particularly for the raw data upload. Supported data formats for the upload to the EIAGRID portal are the SEGY and the SEG2 format. At any stage of the data collection, the uploaded shot-gathers can be concatenated to a single file, converted to the internal data format and subjected to further processing.

Data visualization and preprocessing

Looking at a raw shot seismogram as depicted on the left side of Figure 2, mainly different kinds of noise that superimpose the searched for reflections are visible. Raw data contains a multitude of different wave types and only primary bodywaves of a specified wave mode, usually compressional waves, are considered as signal for seismic reflection imaging. All other wave types are considered as noise and have to be removed during preprocessing and processing. Accurate preprocessing has much influence on the reliability of the final subsurface image. However, during data acquisition only a preliminary result suitable for in-field QC of the acquired data is needed and thus a basic collection of preprocessing tools with a limited choice of parameter settings is sufficient. For data visualization and preprocessing we chose the free software package Seismic Un*x (SU) provided by the Colorado School of Mines (Cohen and Stockwell, 2000). This package features a multitude of data visualization, manipulation and processing tools that are applied via command line in a Unix-like manner. We included a small subset of this tools with a customized choice of options into the EIAGRID GUI (see Figure 3). In this way a technician without any experience in SU can perform the following tasks from any hardware configuration that supports a web-browser:

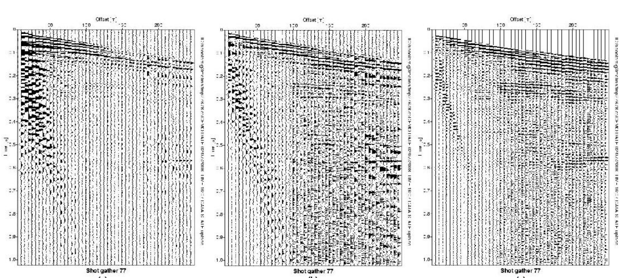

Figure 2: Shot-gather of seismic data recorded in the Muravera delta (South-East Sardinia). Before preprocessing (a); after the application of gain and trace-balancing (b), after bandpass filter and muting (c).

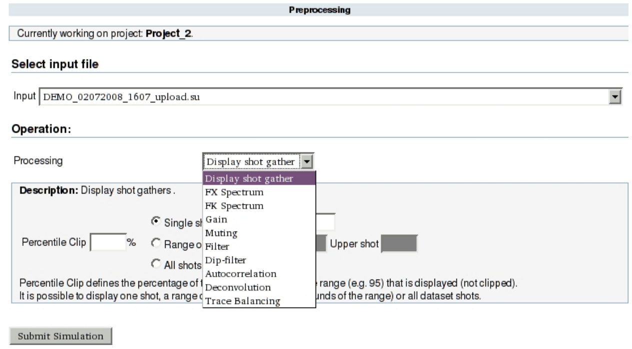

Figure 3: EIAGRID tools for data preprocessing and visualization. Since the output of previous steps can be used as input of the current process, a complete preprocessing workflow can be carried through.

- visualize traces in shot-gathers, view frequency, frequency-wavenumber and autocorrelation spectra.
- mute early arrivals not related to reflection events.
- apply a time dependent gain function to remove the effect of spherical divergence.
- balance the amplitude content of traces to correct for variations along the line caused, e. g., by laterally changing near-surface conditions and varying geophone-ground coupling.
- band-pass filter the data to suppress noise that lies outside the expected signal bandwidth.
- dip filter the data in the f-k domain to discriminate coherent noise by its traveltime-versus-offset dip.
- apply deconvolution to increase the temporal resolution and/or to remove reverberations and short periode multiples.

CRS stack and velocity analysis

The Common-Reflection-Surface (CRS) stack method (see, e. g., Jäger, 1999; Mann, 2002; Heilmann, 2007) can be seen as a generalized multi-dimensional and multi-parameter stacking and velocity analysis tool based on a second-order approximation of reflection events and high density coherence analysis. It provides a simulated zero-offset (ZO) section of high signal-to-noise ratio and valuable kinematic wavefield attributes that include and complement the conventional stacking velocity.

The CRS theory provides a framework to derive the basic stacking formula which approximates the time of flight of signals traveling in a 2D non-homogeneous medium, using properties that are related to a fictitious ZO ray. Expressed in midpoint and half-offset coordinates x_m and h this formula reads

$$t_{\mathrm{hyp}}^2(x_m,h) = \left[t_0 + \frac{2\sin\alpha\,(x_m-x_0)}{v_0}\right]^2 + \frac{2t_0\cos^2\alpha}{v_0}\left[\frac{(x_m-x_0)^2}{R_N} + \frac{h^2}{R_{NIP}}\right],$$

where v_0 represents the near-surface velocity, t_0 and x_0 specify traveltime and coincident source/receiver position of the ZO reference ray, and α, R_{NIP}, and R_N denote the three kinematic wavefield attributes.

Stacking is applied in a data-driven manner: For each sample of the ZO section the three wavefield attributes are determined from the seismic signals by means of a data-driven optimization loop as depicted in Figure 4. As a result, a high degree of automatization can be achieved. Neither manual picking in velocity spectra nor a priori information concerning the unknown macro-velocity model are required. This makes the CRS method very suited for real-time processing where time-consuming user-interaction has to be minimized as far as possible.
For the EIAGRID portal, a grid computing version of the CRS stack was developed that divides the work into a large number of subprocesses that are submitted to the grid using a job scheduler. The CRS stack GUI with parameter values used for a test data set acquired close to Muravera, in the Flumendosa River Delta, Sardinia (Italy), is depicted in Figure 5. The job control window, which shows the status of all subprocesses running on the computing cluster, is displayed in Figure 7. Finally, the resulting stack section obtained after a couple of minutes is shown in Figure 8. The three physically interpretable wavefield-attribute sections (pictures omitted) are used: (1) for residual static corrections resulting in a corrected dataset which can be subject to a second processing cycle (see, e. g., Heilmann, 2007). Source and receiver statics for the Muravera data are depicted in Figure 6. (2) to derive time migration velocities v_{tm} that read, according to Mann (2002),

$$v_{tm}^2 = \frac{2v_0^2 R_{NIP}}{2R_{NIP}\,sin^2\alpha + v_0\,t_0\,cos^2\alpha}.$$

The latter are used for creating a smooth migration velocity model by means of an iterative 2D smoothing algorithm which fills the gaps between reflections where no reliable wavefield-attributes can be obtained.

Time migration

Migration reverses the effects of wave propagation in order to transform the preprocessed data into an image which resembles the distribution of the geological interfaces in the subsurface. Different to depth migration in which the migrated image displays the actual geological structure in depth, the time-migrated image is still defined in the time domain and requires a time-to-depth conversion for a final interpretation. Here we use a parallelized Kirchhoff time migration scheme, based on an integral solution of the wave equation. For each ZO location an optimum migration aperture defined by the projected Fresnel zone and centered around the stationary point, where migration operator and reflection event are tangent, is estimated from the previously obtained CRS results (Spinner, 2007). The prestack time migration results for the Muravera data overlayed on the migration velocity model is depicted in Figure 9.

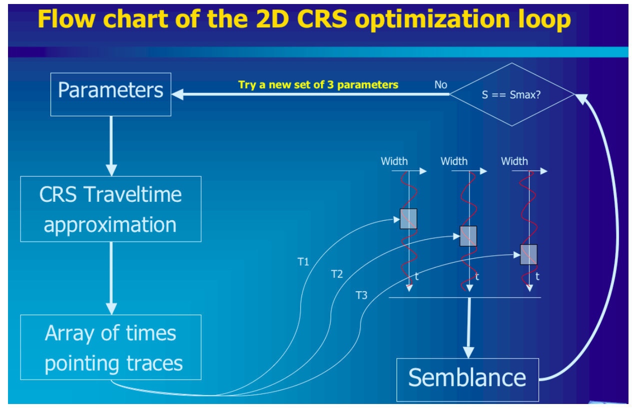

Figure 4: Data-driven stacking parameter determination. Each stacking parameter triple within a given search range defines a hypothetical second-order reflection response. The optimum parameter triple maximizes the coherence between this prediction and the actually measured data.

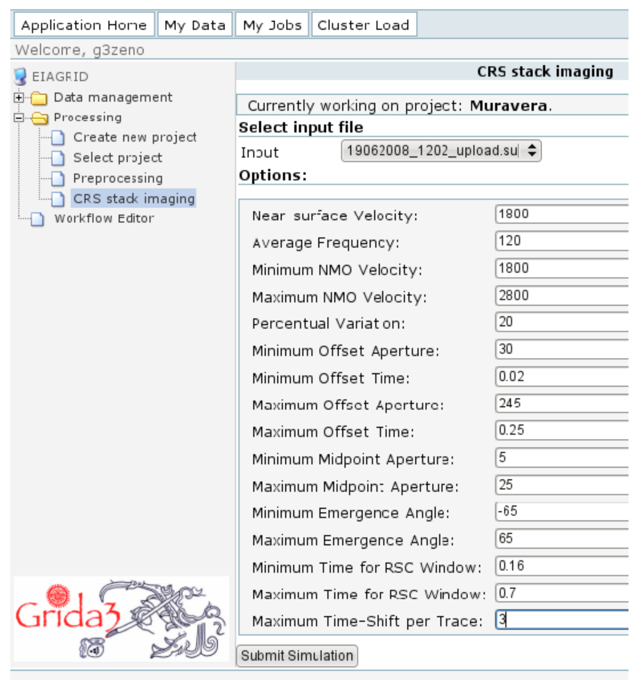

Figure 5: CRS stack interface with processing parameters used for the Muravera data.

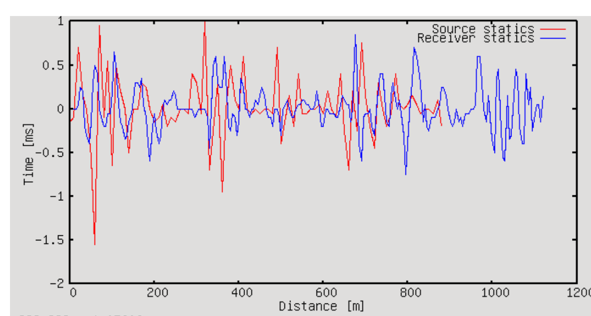

Figure 6: Residual static correction: Source statics (red) and receiver statics (blue) for the Muravera dataset.

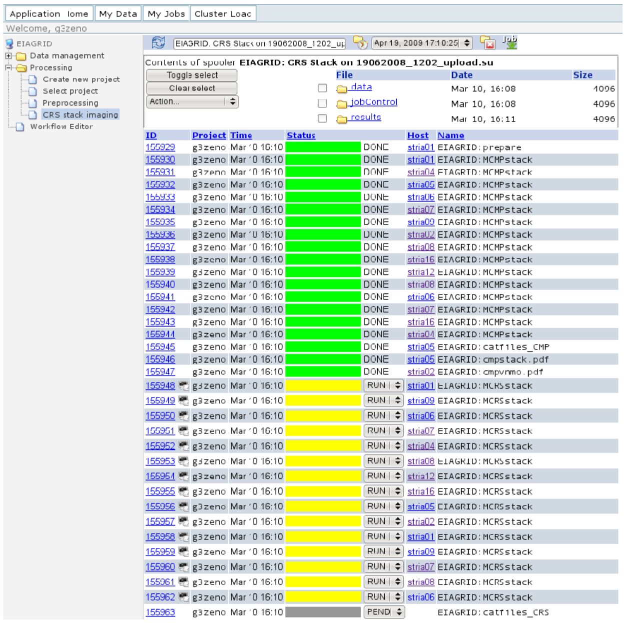

Figure 7: Job control window of the EIAGRID portal, showing fifteen CRS stack subprocesses running on the cluster.

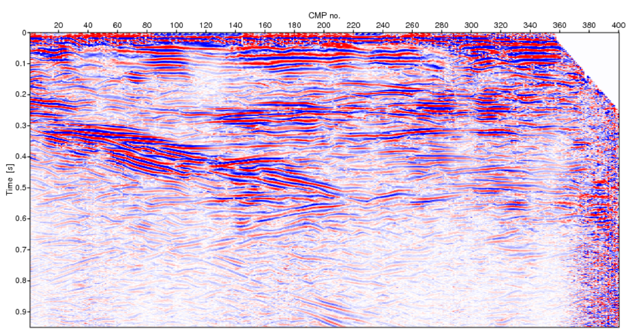

Figure 8: CRS stacked ZO section for an environmental data set recorded close to Muravera in the South-East of Sardinia (see also Deidda et al., 2006).

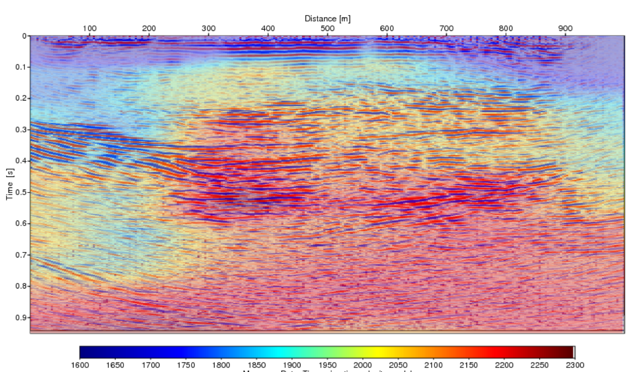

Figure 9: Prestack time migration result obtained for the Muravera data overlayed on the velocity model. The maximum traveltime is related to a depth of approximately 1 km.

Conclusions and Outlook

The extent and the complexity of environmental problems and their consequences on the ecosystem, human health and economic development make an integrated, cross-disciplinary approach necessary. The Grid approach is emerging worldwide as a formidable tool to tackle such problems. It reflects a conceptual framework rather than a physical resource employing central management of both data grid (shared storage clusters) and grid computing facilities (shared computing clusters). This abstract focused on data processing and visualization using distant grid computing facilities. Nevertheless, the large potential comprised in combining and sharing, e. g., hydrological, meteorological and geophysical data within a multidisciplinary data-grid environment using Geo Information System (GIS) technologies, mesh generators and data-crossing techniques should be mentioned as well. The functionality of the EIAGRID service was demonstrated by means of an acoustic reflection seismic data set comprizing compressional waves from explosive sources. The EIAGRID service was also successfully tested for a shear-wave data set from Cardedu, Sardinia and for a multi-offset Ground Penetrating Radar (GPR) data set from Larreule, France. The implementation of a graphical workflow manager is under current development. It will allow the user to construct entire workflows that can be executed in parallel using the large computing power provided by the grid deployment.

Acknowledgments

This work is funded by the government of the Regione Autonoma della Sardegna and by the Italian Ministry of Research and Education. Furthermore, we would like to thank Miriam Spinner for providing the time migration code, Martin Tygel, Hervè Perroud and Gian Piero Deidda for providing data, and all colleagues working in the GRIDA3 project.

References

Cohen, J. K. and Stockwell, J. J. W. (2000). Seismic Un*x Release 34: a free package for seismic research and processing. *Geophysical Journal International*, 125:431–442.

Deidda, G. P., Ranieri, G., Uras, G., Cosentino, P., and Martorana, R. (2006). Geophysical investigations in the Flumendosa River Delta, Sardinia, Italy—Seismic reflection imaging. *Geophysics*, 71(4):121–128.

Halonen, T., J., R., and J., M. (2003). *GSM, GPRS, and EDGE performance: Evolution towards 3G/UMTS*. Wiley.

Heilmann, Z. (2007). *CRS-stack-based seismic reflection imaging for land data in time and depth domains*. [online] http://digbib.ubka.uni-karlsruhe.de/volltexte/1000005850.

Jäger, R. (1999). The Common-Reflection-Surface stack - theory and application. Master's thesis, Karlsruhe University.

Mann, J. (2002). *Extensions and applications of the Common-Reflection-Surface stack method*. Logos Verlag, Berlin.

Spinner, M. (2007). *CRS-based minimum-aperture Kirchhoff migration in the time domain*. [online] http://digbib.ubka.uni-karlsruhe.de/volltexte/1000007099.