

Kaleidoscope Project

Francisco Ortigosa
REPSOL-YPF

RATIONALE

The sediments below the deep and ultra-deep waters of the US Federal waters of the Gulf of Mexico shelter rich oil reserves, sometimes as much as 40,000 ft from the surface. Minerals Management Service (MMS), the federal agency in the U.S. Department of the Interior that manages the nation's oil, natural gas and other mineral resources on the outer continental shelf in federal offshore waters, estimates that the Gulf of Mexico holds 37 billion barrels of "undiscovered, conventionally recoverable" oil.

These reserves are very difficult to find and reach due to thick layers of salt that preclude the imaging and visualization of the oil-bearing sands underneath. The oil industry uses sophisticated technologies to locate and visualize these exploratory objectives. These technologies are computing intensive and the success to properly "see underneath" depends largely on the power of the supercomputers used. It is remarkable that public benchmarks show that the Cell/BE Processors perform the computation of algorithms central to seismic imaging, 40 times faster than leading brand processors used in today's supercomputers. That increase in computer power makes feasible the application of imaging technologies that until today have been considered as a utopia in the oil industry, allowing more reliable exploration.

The Kaleidoscope Project seeks exploitation of Cell/BE unparalleled properties for the creation of the next generation seismic imaging technologies specifically tailored to the Processor for the visualization of the earth interior and the adaptation of existing imaging technologies used in oil exploration by exploiting the Cell/BE Processor unparalleled properties. The output from the Kaleidoscope Project will be faster tools, by several orders of magnitude, more reliable software to visualize below the thick layers of salt present in the Gulf of Mexico and therefore reducing significantly the exploration risks and making accessible oil reserves that otherwise would be invisible to the industry.

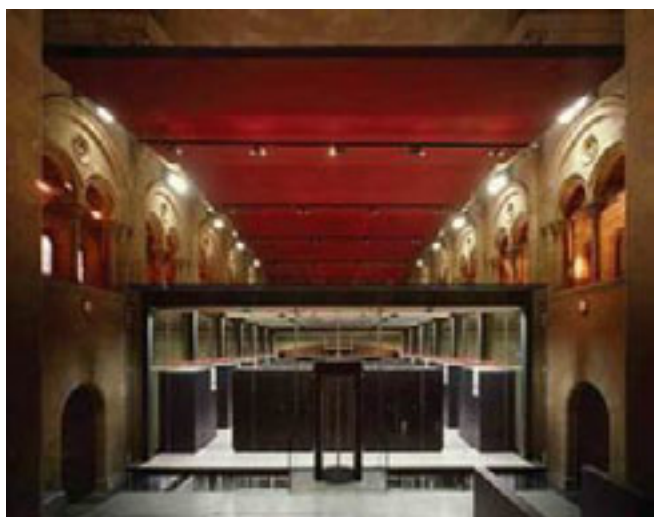


Figure 1. The MareNostrum supercomputer was installed in a former cathedral. Courtesy of Barcelona Supercomputing Center.

The Kaleidoscope Project is a "dream team" partnership of top geophysicists, computer scientists and organizations from around the world has been created by Repsol YPF, a Spanish integrated oil company with important assets in the US Gulf of Mexico, 3DGeo, a leader Houston-based imaging

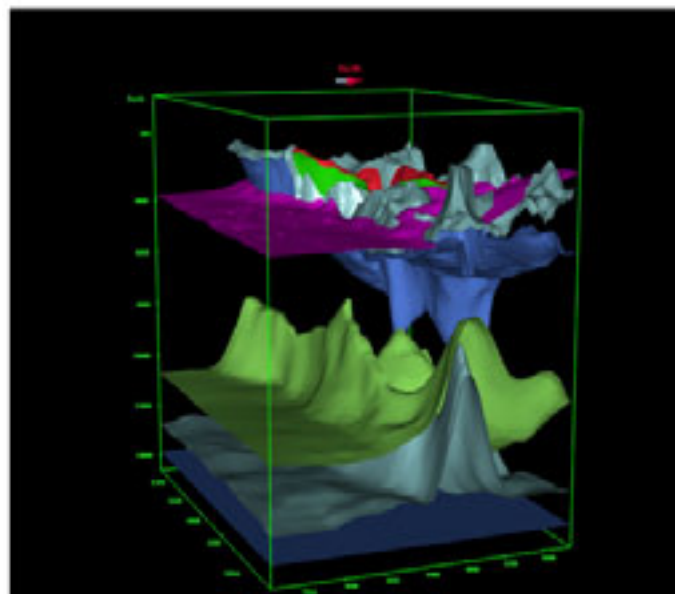


Figure 2. Geological Repsol YPF model with SALT bodies (blue) and other structures

company formed by Stanford University professor and seismic imaging pioneer, Biondo Biondi, and the Barcelona Supercomputer Center (BSC). The BSC hosts the MareNostrum, powered by IBM, the third largest computer of the in Europe. The Kaleidoscope Project has privileged access through the BSC to Cell/BE based systems and technology because the BSC is one of the few research centers in the world developing libraries and codes for such processors.

SOURCES OF VALUE IN SEISMIC IMAGING

The Earth's Interior can be decomposed everywhere in two elements. First, the geometry of all rock layers underneath the surface, called *Structure*, and, second the velocity distribution of these rocks. Such velocities refer to the speed of sound when transmitted through a particular rock across the structure. That velocity is an intrinsic property to the nature of the rocks, and altogether constitutes the *Velocity Field*. *Structure* and *Velocity Field* are independent variables of the Earth's Interior that when combined give place to the *Velocity Model* of a given place on the Earth.

Subsurface Seismic Imaging consists on determining simultaneously the *Structure* and the *Velocity Field* somewhere on the Earth from a single seismic experiment.

Since there are two unknown variables, *Structure* and *Velocity Field* and a single experiment, the solution to the problem is indeterminate. The only way to solve it is by iterating from an initial *Velocity Model* that somehow has to be guessed before the process start.

There are three Sources of Value in Seismic Imaging: Velocity Model, Algorithms used in the iterations, and the Capacity of the computers used to iterate.

VELOCITY MODEL

Focusing seismic energy in Seismic Imaging is exactly the same process as focusing light in Optics. The Velocity Model in Seismic Imaging plays the

EUROPEAN REGION NEWSLETTER

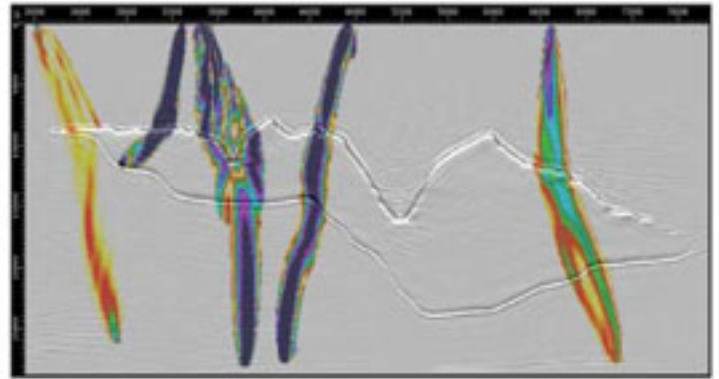
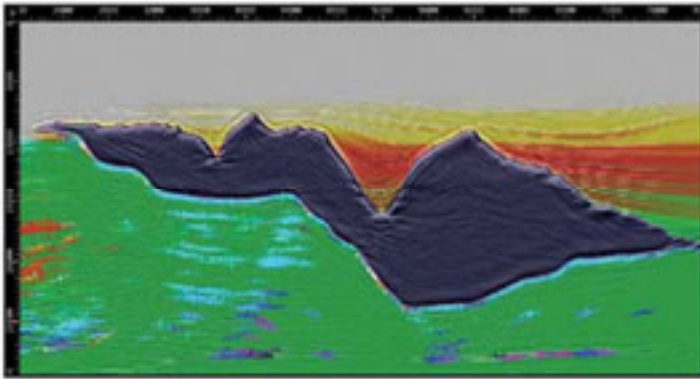


Figure 3. Sigsbee synthetic model

same role of a lens. Repsol-YPF strongly believes that the greatest value added during the Seismic Imaging process comes from the time spent "crafting the lens", i.e. building the Velocity Model. Moreover, as clearer subsalt images are provided by new algorithms, there is still the problem of spatially accurate images, resolvable only with a detailed velocity model.

To maximize the time for the data to be in the hands of the interpreters to build the Velocity Model, the computation time has to be minimized, therefore requiring a very fast computational solution.

Within Kaleidoscope, the Velocity Model Building process is improved with the development of tools that reduce the need for subjective and sometimes inconsistent human interaction, and at the same time, increasing turnaround for large, dense 3D seismic datasets.

ALGORITHM

Algorithms are crucial for the quality of the final image. Seismic Imaging algorithms have been known even from the sixties, but they had not been coded due to the lack of adequate or affordable hardware. The affordability of the hardware is growing faster than the capacity to progress in the algorithm implementation and coding.

But because of lack of coordination between algorithm builders and hardware manufacturers, there is a real race, to provide with newer and faster algorithms and codes through taking shortcuts and compromises from the original algorithms. This competition is leading to a situation that somebody quoted as "algorithm pollution" referring to the amount of different flavors for the implementation of Seismic Imaging technologies.

The Kaleidoscope Project tackles this problem from a different perspective: Collaboration between algorithm coders (3DGeo, BSC) and hardware manufacturers (IBM). In doing so, there is a parallel simultaneous research in hardware and software. Therefore, there is no compromise in the imaging quality by making trade-offs between algorithm accuracy and algorithm speed. The speed will come from the Cell/B.E. (*) processor, the I/O improvements and the tailoring of codes to the Cell/B.E. (*) processor.

The imaging algorithms being developed are grouped in three categories:

- Anisotropic One-pass and Two-pass Shot Profile Migration
- Plane-Wave Migration & Migration in Tilted Coordinates (PWSPM)
- Anisotropic Reverse Time Migration (RTM)

CAPACITY

Every iteration of an average Seismic Imaging production processing project requires 10^{20} floating point operations. That means that every iteration in an average 10 Tflops machine requires four months. On the other hand, such iteration would only require one and a half day in a peta-scale machine.

The need for Capacity is obvious in Seismic Imaging. The amazing price-performance ratio of Linux PC-Clusters made Seismic Imaging Technology a reality. Wave Equation algorithms have been only widely available to the industry from around year 2003. The evolution of the algorithms, and the application to exploration in increasingly geologically complex areas is a consequence of the ever increasing performance of PC-Clusters.

During the last five years, the need in computing power needed for seismic imaging in the oil industry has increased two orders of magnitude and the storage, and I/O needs, three orders of magnitude. At present time, and considering the new algorithms to come (Waveform Inversion, Plane Wave Reverse Time Migration,...) there is no indication that this rate will decrease.

Capacity evolution is predictable. By looking at the Top500 computers by year on a log-log scale, it can be predicted that in three years peta-scale capacity will be widely available.

The problem is not when to achieve the peta-scale capacity widely available, but how to do it. It is clear that widely available peta-scale capacity will require:

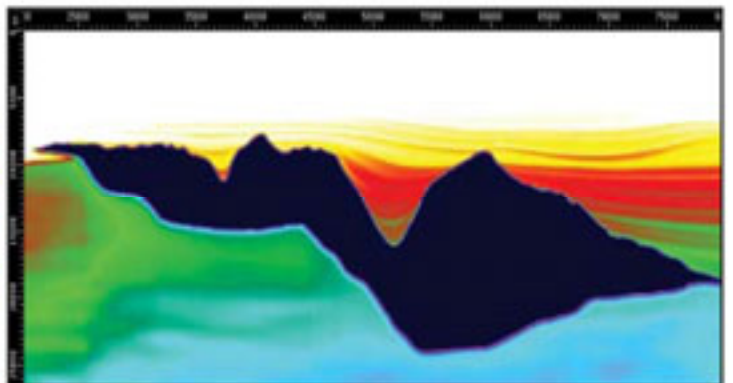
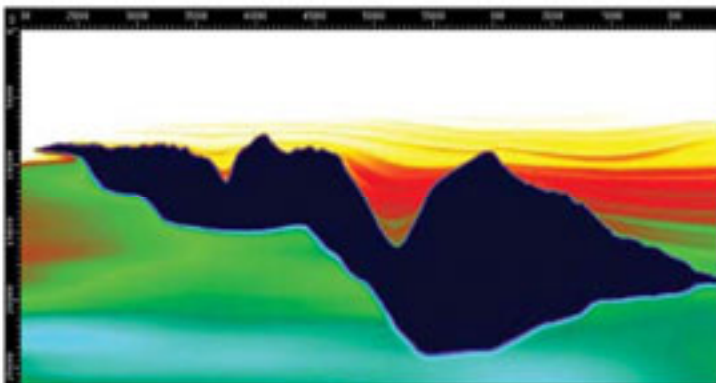


Figure 4. Two iterations of Wave Path Tomography (Sigsbee starting velocity model)

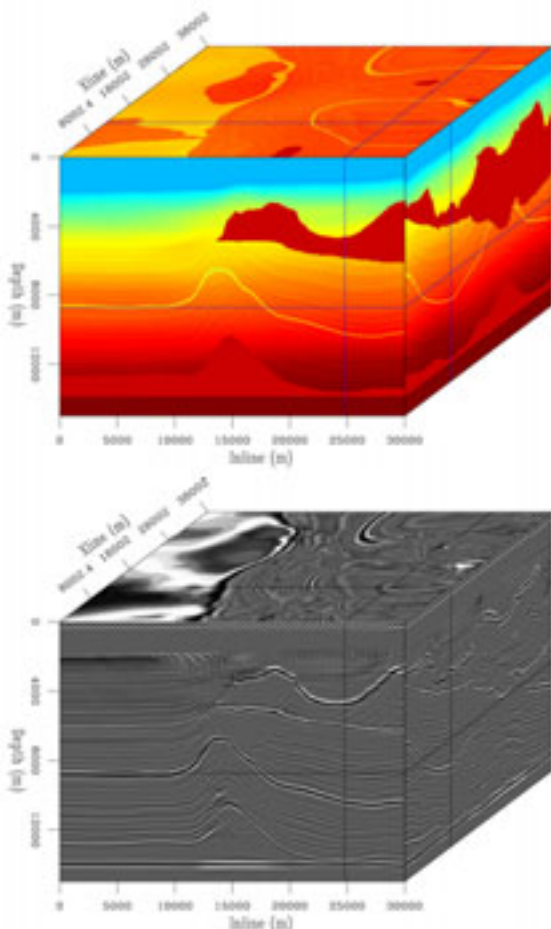


Figure 5. (a) 3D velocity model, (b) Shot-Profile Migration (SPM) of the modeled data with 4,047 shot gathers and limited cross-line aperture. AGC has been applied to the image

- A multielement(multicore) processor, given that CPU frequency has reached a limit. Any multielement processor will be difficult to program.
- This multielement processor must have very low power consumption to make peta-scale capacity economically viable and technically feasible.
- Widely available peta-scale capacity requires a very cheap processor. That implies mass production in figures larger than the current production for processors in the PC industry.

The Cell/BE processor met these three characteristics and in addition, testing on the processor evidences that the performance of present Cell/BE vs. present Superscalar Processors is superior in FFTs (up to 40 times), Stencil Computations (more than 15 times), with a power consumption reduction of 10 times.

Migration software must give simultaneously the maximum performance and the maximum flexibility to simulate different scenarios. It is critical to exploit the different levels of parallelism with maximum efficiency. In Kaleidoscope we manage the following levels of parallelism:

- **Grid level:** All the shots in a migration algorithm could be process in parallel. The different computational nodes of a cluster could be used to run simultaneously different shots. This is the called Grid or Application level parallelism. We use the Grid-SuperScalar programming model to exploit it. This level requires manage workflows with some data dependencies defined by input/output files. Moreover is critical to have a fault tolerance mechanism at this level due to the duration of a complete migration execution. The parallel efficiency of this level is 100%. The key issue at this level is the simplicity to express the workflows.

- **Process level:** Each individual shot requires some hardware resources. If these resources are larger than the available resources in a single computational node of the cluster, the shot execution must be splitted between some computational nodes using domain decomposition techniques. We use MPI programming model to exploit this level. The scalability of this level is limited by the use of Finite Differences as discretization technique. However, this is not a problem because the number of domains needed to have enough hardware resources is quite small. We always work with a parallel efficiency at this level greater than 90%. Moreover, at this level we must manage the IO needed by RTM. Using asynchronous IO and checkpointing capability we are able to minimize the IO time in a RTM execution.
- **Thread level:** In the present supercomputers, a single computational node use to be a shared memory multiprocessor. In order to use efficiently all the processors in a computational node we use the OpenMP programming model. This allows us to use all the memory in a single node for a single shot or for a domain from a single shot. The key issue is to manage properly the memory access in order to have a good thread load balancing and the minimum thread memory interferences. We have obtained a 94% parallel efficiency at this level using IBM JS21 blades as computational nodes.
- **Processor level:** Because migration algorithms use to be limited by memory bandwidth it is critical to minimize the cache miss ratio of the computational kernel. This is accomplished using blocking algorithms. Other important point at this level is to exploit vector capabilities of the processor.

PUBLICATIONS

- *Technologies Improve Subsalt Imaging*, Dimitri Bevc, Francisco Ortigosa, Antoine Guitton, Bruno Kaelin and Moritz Fliedner. The American Oil & Gas Reporter, February 2008
- *Modeling of wide-azimuth towed-streamer surveys with high performance computing*, B. Kaelin, J. Higginbotham, C. A. Fernandez, F. Ortigosa, B. Fontecha & J. M. Cela, SEG International Exposition and 77th Annual Meeting, San Antonio, Sept 2007
- *3D Migration of a Simulated Wide-Azimuth Towed Streamer Survey*, A. Guitton, B. Kaelin & F. Ortigosa, SEG International Exposition and 77th Annual Meeting, San Antonio, Sept 2007
- *Next Generation Seismic Imaging: High Fidelity Algorithms and High-End Computing*, Dimitri Bevc, Francisco Ortigosa, Antoine Guitton & Bruno Kaelin, AGU General Assembly, Mexico, May 2007
- *Least-Square Attenuation of Reverse-Time-Migration Artifacts*, Antoine Guitton, Bruno Kaelin, Biondo Biondi, Geophysics, Vol. 72, N. 1, February 2007
- *Imaging Condition For Reverse Time Migration*, Bruno Kaelin, Antoine Guitton, SEG 2006 International Exposition and 76th Annual Meeting, New Orleans, Oct 2006
- *Least-Square Attenuation of Reverse-Time Migration Artifacts*, Antoine Guitton, Bruno Kaelin, Biondo Biondi, SEG 2006 International Exposition and 76th Annual Meeting, New Orleans, Oct 2006
- *Robust Imaging Condition for Shot-Profile Migration*, Antoine Guitton, Alejandro Valenciano, Dimitri Bevc, SEG 2006 International Exposition and 76th Annual Meeting, New Orleans, Oct 2006
- *Robust Illumination Compensation for Shot-Profile Migration*, Antoine Guitton, Alejandro Valenciano, Dimitri Bevc, EAGE 68th Conference & Exhibition, Vienna, Austria, June 2006
- *Imaging methods in complex overburden*, Antoine Guitton, Francisco Ortigosa, Bruno Kaelin, EAGE 69th Conference & Exhibition, London, UK, June 2007
- *Illumination effects in reverse time migration*, Bruno Kaelin, Antoine Guitton, Francisco Ortigosa, EAGE 69th Conference & Exhibition, London, UK, June 2007