Using large GPU clusters and the race towards exaflops to improve high-resolution acoustic imaging

Dimitri Komatitsch
Laboratory of Mechanics and Acoustics
CNRS, Marseille, France

GPU’2016, Roma, Italia
September 28, 2016

with some slides from Emanuele Casarotti et al. (INGV Roma, Italy)
Application domains

Earthquakes

Ocean acoustics

Non destructive testing
Earthquake hazard assessment

Use parallel computing to simulate earthquakes

Learn about structure of the Earth based upon seismic waves (tomography)

Produce seismic hazard maps (local/regional scale) e.g. Los Angeles, Tokyo, Mexico City, Seattle

2001 Gujarati (M 7.7) Earthquake, India

20,000 people killed
167,000 injured
≈ 339,000 buildings destroyed
783,000 buildings damaged
Equations of motion (solid)

Differential or *strong* form (e.g., finite differences):

\[ \rho \partial^2_t \mathbf{u} = \nabla \cdot \sigma + \mathbf{f} \]

We solve the integral or *weak* form in the time domain:

\[
\int \rho \mathbf{w} \cdot \partial^2_t \mathbf{u} \, d^3 \mathbf{r} = -\int \nabla \mathbf{w} : \sigma \, d^3 \mathbf{r} \\
\quad + \mathbf{M} : \nabla \mathbf{w} (\mathbf{r}_s) S(t) - \int_{F-S} \mathbf{w} \cdot \sigma \cdot \hat{n} \, d^2 \mathbf{r} \\
\quad + \text{attenuation (memory variables) and ocean load}
\]
Spectral-Element Method

- Developed in Computational Fluid Dynamics (Patera 1984)
- Accuracy of a pseudospectral method, flexibility of a finite-element method
- Extended by Komatitsch and Tromp, Chaljub et al.
- Large curved “spectral” finite-elements with high-degree polynomial interpolation
- Mesh honors the main discontinuities (velocity, density) and topography
- Very efficient on parallel computers, no linear system to invert (diagonal mass matrix)
Our SPECFEM3D software package

Goal: model acoustic / elastic / viscoelastic / poroelastic / seismic wave propagation in the Earth (earthquakes, oil industry), in ocean acoustics, in non destructive testing, in medical acoustic tomography...

The SPECFEM3D source code is open (GNU GPL v2)

Mostly developed by Dimitri Komatitsch and Jeroen Tromp at Harvard University, Caltech and Princeton (USA) and later University of Pau (France) since 1996.

Improved with INRIA (Pau, France), CNRS (Marseille, France), the Barcelona Supercomputing Center (Spain) and University of Basel (Switzerland).
Earthquakes

6 April 2009
Mw 6.2 L’Aquila (Italy)

310 casualties
~ 1000 injured
~ 26000 homeless

Collaboration with
Emanuele Casarotti and Federica Magnoni (INGV Roma, Italy)
Mw 6.2 L’Aquila
L’Aquila, Italy, April 6, 2009 (Mw = 6.2)

Location of the epicenter (© Google Maps)

Mesh defined on the JADE supercomputer on April 7, 2009
Scenario

1D flat - max PGV 45 cm/s

1D w topo - max PGV 48 cm/s

3D - max PGV 74 cm/s

Max PGV data in the epicentral area ~ 65 cm/s

INGV ShakeMap: CENTRAL ITALY – AQUILANO

<table>
<thead>
<tr>
<th>PERCEIVED SHAKING</th>
<th>Not felt</th>
<th>Weak</th>
<th>Light</th>
<th>Moderate</th>
<th>Strong</th>
<th>Very strong</th>
<th>Severe</th>
<th>Violent</th>
<th>Extreme</th>
</tr>
</thead>
<tbody>
<tr>
<td>POTENTIAL DAMAGE</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>Very light</td>
<td>Light</td>
<td>Moderate</td>
<td>Heavy</td>
<td>Very Heavy</td>
<td></td>
</tr>
<tr>
<td>PEAK ACC.(%g)</td>
<td>&lt;0.17</td>
<td>0.17–1.4</td>
<td>1.4–4.6</td>
<td>4.0–9</td>
<td>9–17</td>
<td>17–32</td>
<td>32–61</td>
<td>61–114</td>
<td>&gt;114</td>
</tr>
<tr>
<td>PEAK VEL.(cm/s)</td>
<td>&lt;0.12</td>
<td>0.12–1.1</td>
<td>1.1–3.4</td>
<td>3.4–8</td>
<td>8–16</td>
<td>16–31</td>
<td>31–59</td>
<td>59–115</td>
<td>&gt;115</td>
</tr>
<tr>
<td>INSTRUMENTAL INTENSITY</td>
<td>I</td>
<td>II–III</td>
<td>IV</td>
<td>V</td>
<td>VI</td>
<td>VII</td>
<td>VIII</td>
<td>IX</td>
<td>X+</td>
</tr>
</tbody>
</table>

Scale based upon Wald et al., 1989

(Faenza et al., 2011)
Oil industry applications

- Elastic wave propagation in complex 3D structures,
- Often fluid / solid problems: many oil fields are located offshore (deep offshore, or shallower).
- Anisotropic rocks, geological faults, cracks, bathymetry / topography…
- Thin weathered zone / layer at the surface ⇒ model dispersive surface waves.
Building a cluster

Year 2000, Caltech (USA).

Parallel calculations with message passing (MPI).
320 processors, 160 Gb of memory, Linux.
Huge progress in 10 years

**SC2003 Gordon Bell Award**

**Dimitri Komatitsch**

California Institute of Technology

A 14.6 Billion Degrees of Freedom, 5 Teraflops, 2.5 Terabyte Earthquake Simulation on the Earth Simulator

Earth Simulator: Peak 40 Teraflops; we won the Gordon Bell supercomputing award with SPECFEM3D for a run at 5 teraflops sustained (!!) (OK, with 15 billion degrees of freedom…)
Results for load balancing: cache misses (J. Labarta, BSC)

⇒ it is crucial to reuse common points by keeping them in the cache
GPU graphics cards

Why are they so powerful for scientific computing?

Compute all pixels simultaneously, massive multithreading.
Porting SPECFEM3D on GPUs

At each iteration of the serial time loop, three main types of operations are performed:

- update (with no dependency) of some global arrays composed of the unique points of the mesh

- purely local calculations of the product of predefined derivative matrices with a local copy of the displacement vector along cut planes in the three directions (i, j and k) of a 3D spectral element

- update (with no dependency) of other global arrays composed of the unique points of the mesh
**BLAS 3 (Basic Linear Algebra Subroutines)**

Can we use highly optimized BLAS matrix/matrix products (90% of computations)?

- **For one element**: matrices (5x25, 25x5, 5 x matrices of (5x5)), BLAS is not efficient. Overhead is too expensive for matrices smaller than 20 to 30 square.

- **If we build big matrices** by appending several elements, we have to build 3 matrices, each having a main direction (x,y,z), which causes a lot of cache misses due to the global access because the elements are taken in different orders, thus destroying spatial locality.

- Since all arrays are static, the compiler already produces a very well optimized code.

⇒ No need to, and cannot easily use BLAS

⇒ Compiler already does an excellent job for small static loops
Porting to GPUs: mesh coloring

Key challenge: ensure that contributions from two local nodes never update the same global value from different warps

Use of mesh coloring: suppress dependencies between mesh points inside a given kernel

Use of “atomic” is OK and sufficient these days
High-frequency ocean acoustics, inverse problems in seismology, acoustic tomography, reverse-time migration in seismics: high resolution needed, and/or large iterative problems to solve ⇒ Large calculations to perform.

⇒ GPU computing: code needs to be rewritten, but large speedup can be obtained (around 20x-30x for our finite-element codes, but it is difficult to define speedup).
Adjoint methods for tomography and imaging

Problem is self-adjoint, thus no need for automatic differentiation (AD, autodiff)

\[
\chi_1(m) = \frac{1}{2} \sum_{r=1}^{N_r} \int_0^T w_r(t) \| s(x_r, t; m) - d(x_r, t) \|^2 \, dt,
\]

\[
\delta \chi_1 = \int_V \left[ K_\rho(x) \delta \ln \rho(x) + K_\mu(x) \delta \ln \mu(x) + K_\kappa(x) \delta \ln \kappa(x) \right] \, d^3 x,
\]

\[
K_\kappa(x) = -\int_0^T \kappa(x) [\nabla \cdot s^\dagger(x, T - t)] \nabla \cdot s(x, t) \, dt,
\]

Theory: A. Tarantola, Talagrand and Courtier.

‘Banana-Donut’ kernels (Tony Dahlen et al., Princeton)

Close to time reversal (Mathias Fink et al.) but not identical, thus interesting developments to do.

Idea: apply this to tomography of the full Earth (current ANR / NSF contract with Princeton University, USA), and in acoustic tomography: ocean acoustics, non destructive testing.
L-BFGS method

- Iterative Gauss-Newton algorithm

\[ m_{k+1} = m_k + \left( G_k^t G_k + C_m^{-1} \right)^{-1} \nabla J(m_k) \]

L-BFGS (Low-memory Broyden–Fletcher–Goldfarb–Shanno):

Approximate \( \delta m_k \) from:

\[ m_{k-1}, m_{k-2}, m_{k-3}, \ldots, m_0 \]

\[ \nabla J(m_{k-1}), \nabla J(m_{k-2}), \ldots, \nabla J(m_0) \]

\( \rightarrow \) no need to invert or even build a big matrix
The PYROPE experiment

- French/Spanish initiative, supported by the French ANR
- ~150 temporary + 50 permanent BB stations
- Interstation spacing ~ 60 km
- Dense transects across the Pyrénées
A hybrid approach: Coupling global and regional propagations

A hybrid technique for 3-D waveform modeling and inversion of high frequency teleseismic body waves

Global propagation
Spherically symmetric Earth model

Regional propagation
3-D spherical shell

S. Chevrot, V. Monteiller, D. Komatitsch, N. Fuji & R. Martin
San Francisco, USA, American Geophysical Union Fall Meeting, December 2011
**Synthetic full waveform inversion example**

Full waveform modeling:
- Direct P wave
- Converted waves

With hierarchical frequency content
Imaging the Pyrénées Mountains

- Drastically-improved quality of the images thanks to the high frequencies involved
- This results in a much more precise and therefore much more interesting geological interpretation (how the Earth formed and keeps evolving)

Undoing attenuation on GPUs

- Constitutive relationship:
  \[
  T(t) = \int_{-\infty}^{t} \partial_{t} c(t - t') : \nabla s(t') \, dt'
  \]

  Difficult in time domain methods because of convolution

- Use $L$ Zener body standard linear solids to make an absorption-band model:
  \[
  \mu(t) = \mu_{R} \left[ 1 - \sum_{\ell = 1}^{L} \left( 1 - \frac{\tau_{\ell}^{\varepsilon}}{\tau_{\ell}^{\sigma}} \right) e^{-t/\tau_{\ell}^{\sigma}} \right] H(t)
  \]
Undoing attenuation

Conclusions and future work

- On modern computers, large 3D full-waveform forward modeling problems can be solved at high resolution in the time domain for acoustic / elastic / viscoelastic / poroelastic / seismic waves.

- Inverse (adjoint) tomography / imaging problems can also be studied, although the cost is still high.

- Useful in different industries in addition to academia: oil and gas, medical imaging, ocean acoustics / sonars, non destructive testing (concrete, composite media, fractures, cracks).

- Hybrid (GPU) computing is useful to solve inverse problems in seismic wave propagation and imaging.

- **PRACE project with INGV Roma** to image the Italian lithosphere: 40 million core hours on a petaflop machine.

- Some future trends: high-frequency ocean acoustics, tomography of buried objects, wavelet compression.
About the path to exascale

- We are highly interested and involved in the effort, but we are not 100% experts (we are in acoustics or geophysics labs, not computer science)

- In most cases we will run hundreds of semi-independent runs on different parts of the machine rather than a single big run; big data is thus becoming an issue

- We are in the process of adding OpenMP support in addition to MPI; not too challenging in our application, only a few critical routines impacted

- We tried higher-level directive models (OpenAcc, StarSs and OmpSs from Barcelona BSC). So far the code we get is always significantly slower than our pure MPI code, but the programming models are flexible and interesting

- We successfully used GPUs, including for realistic inverse problems

- INRIA (Franck Cappello) and we added fault-tolerant MPI (SC’11 paper)

- We also recently used ARM boards (MONTBLANC European project) to target lower energy-to-solution models.

*The SPECFEM3D code is freely available open source at www.geodynamics.org*
Acknowledgements and funding

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 641943.