Geophysical Research Abstracts
Vol. 16, EGU2014-16606-1, 2014
EGU General Assembly 2014


# Big Data and High-Performance Computing in Global Seismology

Ebru Bozdag (1), Matthieu Lefebvre (2), Wenjie Lei (2), Daniel Peter (3), James Smith (2), Dimitri Komatitsch (4), Jeroen Tromp (2,5)

(1) Geoazur, University of Nice Sophia Antipolis, Valbonne, France (bozdag@geoazur.unice.fr), (2) Department of Geosciences, Princeton University, Princeton, NJ, USA, (3) Institute of Geophysics, ETH Zurich, Switzerland, (4) Laboratory of Mechanics and Acoustics, CNRS/University of Aix-Marseille, France, (5) Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ, USA

Much of our knowledge of Earth's interior is based on seismic observations and measurements. Adjoint methods provide an efficient way of incorporating 3D full wave propagation in iterative seismic inversions to enhance tomographic images and thus our understanding of processes taking place inside the Earth. Our aim is to take adjoint tomography, which has been successfully applied to regional and continental scale problems, further to image the entire planet. This is one of the extreme imaging challenges in seismology, mainly due to the intense computational requirements and vast amount of high-quality seismic data that can potentially be assimilated. We have started low-resolution inversions (T > 30 s and T > 60 s for body and surface waves, respectively) with a limited data set (253 carefully selected earthquakes and seismic data from permanent and temporary networks) on Oak Ridge National Laboratory's Cray XK7 "Titan" system. Recent improvements in our 3D global wave propagation solvers, such as a GPU version of the SPECFEM3D_GLOBE package, will enable us perform higher-resolution (T > 9 s) and longer duration (∼180 m) simulations to take the advantage of high-frequency body waves and major-arc surface waves, thereby improving imbalanced ray coverage as a result of the uneven global distribution of sources and receivers. Our ultimate goal is to use all earthquakes in the global CMT catalogue within the magnitude range of our interest and data from all available seismic networks. To take the full advantage of computational resources, we need a solid framework to manage big data sets during numerical simulations, pre-processing (i.e. data requests and quality checks, processing data, window selection, etc.) and post-processing (i.e. pre-conditioning and smoothing kernels, etc.). We address the bottlenecks in our global seismic workflow, which are mainly coming from heavy I/O traffic during simulations and the pre- and post-processing stages, by defining new data formats for seismograms and outputs of our 3D solvers (i.e. meshes, kernels, seismic models, etc.) based on ORNL's ADIOS libraries. We will discuss our global adjoint tomography workflow on HPC systems as well as the current status of our global inversions.