US-China Collaborative Development of Scalable and Accurate Earthquake Simulations on Tianhe-1A Supercomputer (SEAS)

1. Introduction

Large-scale efficient and accurate numerical simulation of seismic wave propagation is an important tool for seismic exploration and seismology, which are critical to earthquake prevention, disaster reduction and petroleum exploration. AWP-ODC, an anelastic 4th-order staggered-grid finite-difference (FD) wave propagation code (originally developed by Kim Bak Olsen at University of Utah), is a highly scalable earthquake application for seismic Petascale simulations. It is capable of simulating accurate dynamic earthquake rupture propagation on planar faults as well as producing detailed physics-based anelastic ground motions at frequencies pertinent to safe building design. This application has enabled ground-breaking science results in recent years including San Andreas fault scenarios TeraShake and ShakeOut-D, revealing order-of-magnitude Los Angeles wave-guide amplification (Olsen et al. 2006, 2008, 2009), and a Pacific Northwest megathrust scenario (Olsen et al., 2008). The application achieved "M8" in 2010, the largest-ever fully dynamic simulation of a magnitude-8 earthquake on the southern San Andreas fault for frequencies up to 2 Hz, sustained 220 Tflop/s for 24 hours on NCCS Jaguar (Figure 1, and became an ACM Gordon Bell Finalist in 2010.

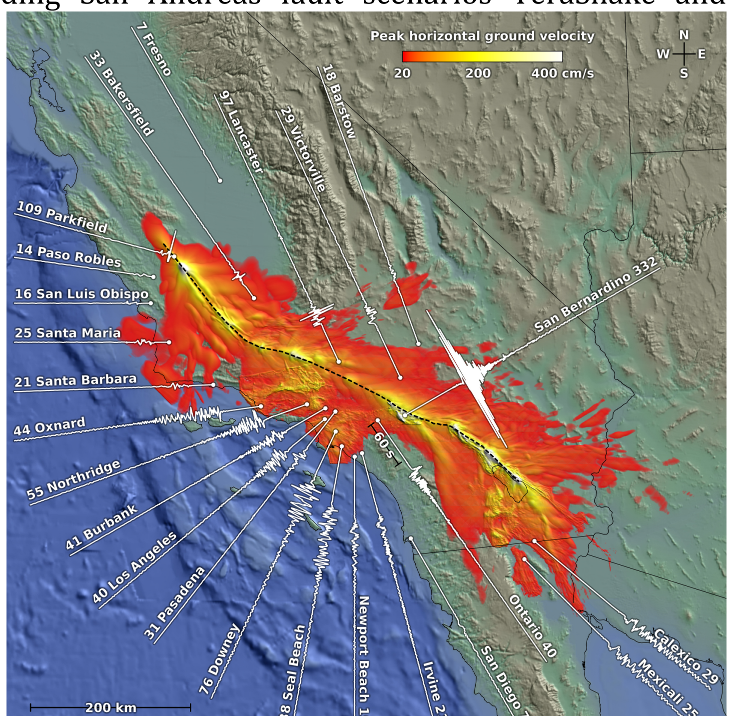

Figure 1. Peak ground velocities for M8, a magnitude-8 earthquake simulation using 223K Jaguar cores, sustaining 220 teraflops (Cui et al., 2010)

The current CPU version of AWP-ODC application has achieved excellent scalability performance through many optimizations. However, to reach the target resolution and accuracy, we have to use very fine-scale mesh to calculate the dynamic simulations, which leads to a huge amount of data and computation. Significant CPU-core hours are required for fixed mesh ground motion simulations, which cannot meet the growing computational need. For example, the M8 simulation required 5.5 millions CPU-hours on Jaguar.

Due to the large computational power and wide memory bandwidth with relatively low cost, exploiting GPU for scientific computation can accelerate computation several dozens time faster than a conventional CPU, which plays a significant role in the improvement of computational efficiency. The emergence of GPU as an alternative to conventional general-purpose processors has led to significant interest in these architectures by the HPC community in recent years. China currently houses the Tianhe-1A, the largest and fastest CPU-GPU cluster in the world. This collaborative project is aimed to extend the application capabilities to take advantage of the availability of the GPU system to accelerate, in particular, seismic hazard calculations.

Moreover, the proposed work will advance Earth system science by delivering a highly efficient and scalable code for Petascale computing. We will accomplish this goal with two steps. The first step is to accelerate the code in the waveform modeling mode on CPU-GPU clusters to largely reduce the allocation hours needed for production calculations. The second step is to combine GPU acceleration and adaptive mesh refinement (AMR) technology to reduce the time-to-solutions to 1/100 of the current value, and expand software capabilities including more advanced features such as in elastic off-fault response and nonlinear soil response.

The outcome of this project is to establish a new simulation capability for use on the largest CPU-GPU clusters, with improved scalability, robustness, and high parallel efficiency. The enhancement of software capabilities will eventually further Earth system research in both countries and provide improved seismic hazard analysis and a road map to engineer safe structures.

2. Description of the development

Our code development targets the Tianhe-1A platform capabilities. The Tianhe-1A Supercomputer, located at the National Supercomputer Center, Tianjin, has set a performance record of 2.57 PFlops, as measured by the LINPACK benchmark (Yang, et al., 2011), ranked at No. 2 in the TOP500 list (the 38rd edition, November 2011). Tianhe-1A epitomizes modern heterogeneous computing by coupling massively parallel GPUs with multi-core CPUs, enabling significant achievements in performance, size and power. The system uses 7,168 NVIDIA® Tesla™ M2050 GPUs and 14,336 CPUs, has 262TB memory totally, and 2PB Capacity of I/O storage. The applications running on the system include petroleum exploration, biological medicine research, simulation of large aircraft design, remote sensing data processing, data analysis of financial engineering and simulation of environment research. This project will extend the scalability and usability of Tianhe-1A system to support earthquake applications at the Petascale level.

The AWP-ODC code has been developed into an initial protocol version of a CUDA based GPU code, which was well tested on both GPU NVIDIA C2050 and M2090 systems. The single-GPU based benchmarking on NVIDIA M2090 demonstrated outstanding performance that is comparable to the original optimized MPI code running on 64 CPU cores of a 8 dual-socket quad-core Intel Xeon E5530 (SDSC Triton, Intel Nehalem 2.40GHz), or 128 CPU core of a cluster of 11 dual-socket six-core AMD Opteron processors (NICS Kraken, AMD Istanbul 2.6GHz) (Zhou, et al., 2012). The MPI-CUDA version has moved to a fine-tuning stage. Advanced CUDA features and overlapping algorithms have been implemented to reduce the communication latency caused by data exchange between CPU and GPU or MPI communication between CPUs. The multi-GPU version has been tested and verified on four NVIDIA M2090 GPUs and migrated to Georgia Tech Keeneland for further tuning.

For GPU acceleration, we will initially analyze the current algorithm and draw up a suitable scheme for GPU computing. The CPU computing resources in the nodes will be also utilized completely to realize the effective cooperated computing on both GPU and CPU. The splitting computation ratio between CPU and GPU will be adjusted automatically based on the number of CPUs and GPUs, which ensures the load balance between CPU and GPU through hybrid solution to reduce runtime overhead. Aggressive optimization techniques will be considered, such as cache optimization, SSE acceleration, loop unrolling and vectorization, to catch up the powerful GPU acceleration capability. Our final goal is to implement a highly scalable AWP-ODC GPU code running at full machine scale on Tianhe 1A to achieve sustained high flops by the end of 2012.

We propose to incorporate AMR and discontinuous meshes (DM) into AWP-ODC in order to enhance the accuracy of the dynamic rupture models and wave propagation and reduce computing demands by orders of magnitude. Stable DM schemes have been developed for staggered grids (e.g., Kristek et al., 2010) which allow for finer grid steps in the low-velocity near surface material. We will use AMR to refine and coarsen the grid steps in space and time in the vicinity of the fault to increase the accuracy of the dynamic rupture. Thus, a version of AWP-ODC with AMR and DM incorporated (AWP-ODC-AMR-DM) will provide a much more powerful research tool to the seismology community, in China, USA, and other parts of the world. For a MPI version of AWP-ODC-AMR-DM, we plan to use PARAMESH (MacNeice et al., 2000) as the mesh management core. We choose this framework for its simplicity and promising scalability. PARAMESH provides a set of F90 subroutines that facilitate parallel Octree-structured adaptive mesh management based on MPI, including mesh refinement and coarsening, load balancing via block migration, neighboring info gathering, data communication between blocks and some common prolongation, restriction, and interpolation schemes. We propose to develop a pre-processing module to convert the mesh, media, source etc. to a block-structured form which simplifies remaining efforts to realizing the FD schemes on one block.

We expect to complete the implementation of the AMR/DM code in 1 year. However, considering the testing/verification and possible algorithm level issues such as the mesh refine/coarsen criterion, interpolation issues on coarse-fine grid interface etc., a project duration of 2-3 years is likely more realistic. Finally, we are interested in developing parallel AMR software/infrastructure on modern hardware architectures. Here, we will consider AMR schemes for dynamic rupture and wave propagation simulation using a hybrid-programming model for a longer term project (5 years).

3. Importance of the Topic to Particular Scientific Applications and/or Computer Science Research

Since 2000, earthquakes have claimed the lives of more than 730,000 people worldwide and have affected the global economy (e.g., the 2011 M9.0 Tohoku, earthquake caused disruption in supplies for the electronic and automotive industry). Both USA and China include regions of extraordinarily high seismic risk. For example, the United States suffered $20 billion+ losses and 70+ fatalities during the 1994 Northridge earthquake in California, but an overdue M7.5+ event on the southern San Andreas fault is expected to be significantly more devastating. China has suffered from major destruction of life and properties during the Wenchuan Earthquake in 2008, the largest impact of any earthquake to strike China since the catastrophic 1976 Tangshan Earthquake. The heightened awareness of the value of risk management is the opportunity to leverage the strength of modeling to encourage mitigation and reduce the losses for future earthquakes. Improved seismic modeling through this collaborative project will help inform urban planning, seismic design provisions in building codes, performance-based earthquake engineering, and disaster planning.

The GPU and AMR/DM developments proposed here will promote highly scalable and efficient earthquake simulations in high performance computing environments. Large-scale simulations using AMR/DM have been increasingly used in the dynamic rupture modeling and wave propagation community. AMR provides an advantage over fixed mesh simulations by refining the mesh only where necessary, allowing a greatly reduced computational cost. However, PDE solvers using AMR are among the most challenging applications to adapt to massively parallel computing environments. This project will collaborate the effort and develop a highly efficient algorithm using off-the-shelf AMR framework, supported by AMR experts at the Chinese Academy of Sciences that will allow the local timesteps to vary with spatial resolution.

In addition to the computational enhancements to the ADP-ODC code discussed above, we propose several critical Earth-science-related additions to the code. Inelastic off-fault response has been shown to significantly affect both rupture propagation (in particular, slip velocities) and the resulting ground motion estimates (Ma and Andrews 2010). We propose to implement a pressure-dependent Drucker-Prager yield criterion in a depth-dependent stress model (Drucker and Prager, 1952), which provides a convenient tool to include inelastic off-fault response. Furthermore, we propose to implement fully-nonlinear 3D soil response in AWP-ODC, an effort dictated by the need to push deterministic simulations of wave propagation to frequencies beyond about 1 Hz. We plan to use the dilatancy model of Iai et al. (1990), which is implemented in the fully nonlinear 3D finite difference code NOAH2D (Bonilla et al., 2006). Such implementation will allow us to study the nonlinear effects due to surface waves, which caused the strongest shaking in the Los Angeles basin for linear simulations of large earthquake on the San Andreas fault (e.g., Cui et al., 2010; Olsen et al., 2006, 2008, 2009).

4. Objectives of a Particular Collaboration

Major objectives of this collaboration project are to develop and test the proposed computational rupture dynamic and wave propagation platform and make the platform and its products available across domains. The enhanced platform will establish well-defined features, allowing geoscientists, earthquake engineers, and risk managers to create a full-scale, time-dependent hazard model. Four objectives will enable research on Petascale computing facilities:

- Improve parallel performance of the existing AWP-ODC-GPU code and enable it to run efficiently on China's Tianhe-1A system with a goal of scaling up to the full system scale. Effective algorithms will be developed to speed up the C code on Intel x5670 CPUs and Nvidia Tesla M2050 GPUs on Tianhe-1A (Chi). The benchmarks will then be compared to the performance on DOE ORNL's Titan system by end of 2012 when Titan becomes available (Cui).
- Develop AMR/DM capabilities added to AWP-ODC, capable of resolving the rupture process and wave propagation over orders of magnitude in length and time scales (Chi).
- Incorporate inelastic off-fault response into AWP-ODC (Olsen, Day).
- Implement fully nonlinear 3D soil response into AWP-ODC (Olsen, Day).

During the proposed AWP-ODC software development process, we will verify correctness of our software by running well-established reference earthquake wave propagation reference problems. Validation testing of AWP-ODC will include simulation of well-observed historical earthquakes and comparison of the simulation results against observed ground motions.

5.Benefits to the Bilateral and Global Communities

China will gain from this project to access a state-of-the-art seismic application for regional scale earthquake simulations and seismology expertise via experts from the Southern California Earthquake Center (SCEC). Tuning the application to run on Tianhe-1A system will advance large-scale seismic simulations in China. US seismologists will assist in this transition process, gaining a scalable code to run on start-of-the-art heterogeneous architecture for seismic hazard analysis. The final product of this GPU-based code will be considered to produce a California state-wide CyberShake 3.0 map, a project coordinated by SCEC and USGS. The planned 5000 site Strain Green Tensor calculations for the CyberShake originally required 700+ millions of allocation hours to complete the calculations based on current US-based system. The added DM capability will push the efficiency further ahead, preparing for next generation Petascale systems for energy-aware calculations. Finally, faculty and graduate students will gain valuable international collaboration experience from this joint effort.

6. Identification of Research Activities in China and US that Form Basis of New Collaborations

Dr. Yifeng Cui, director of High Performance GeoComputing Lab at San Diego Supercomputer Center and adjunct Professor at San Diego State University, and the AWP-ODC co-developer. He has extensive experience in High Performance

Computing and has optimized and maintained the code to current status. Dr. Cui will take the leadership in this software development. HPGeoC team members, including Prof. Dong Ju Choi, visualization expert Amit Chourasia, and graduate student Jun Zhou, will participate in the project and provide the development solutions.

Dr. Xuebin Chi, full Professor and the director of Supercomputing Center of Chinese Academy of Sciences. He has broad experience in Parallel Computing and Computational Mathematics. His team will develop GPU code and add AMR/DM to AWP-ODC, and evaluate and tune the implementation for the Tianhe-1A system. SC team members, Dr. Jian Zhang and Dr. Ningming Nie will develop GPU code and add AMR/DM to AWP-ODC, and evaluate and tune the implementation for the Tianhe-1A system.

Prof. Xiang-chu Yin, full Professor of the Institute of Earthquake Science, Chinese Earthquake Administration，LNM (State Key Laboratory of Nonlinear Mechanics), Institute of Mechanics, Chinese Academy of Sciences, former director of the department of Earthquake physics at the Institute of Geophysics and ISB member of ACES (APEC Cooperation for Earthquake Simulation). He has experience in numerical simulation of the seismogenic process. Professor Yongxian Zhang, Ziping Song and graduate student Yue Liu will participate in the project.

Dr. Kim Olsen, a Professor of Seismology at San Diego State University and developer of the original AWP software. He has extensive experience in dynamic rupture and ground motion calculations, high performance computing and web-based dissemination of simulations. He will be responsible for seismology-related development including the inelastic off-fault response and nonlinear 3D soil modeling.

Dr. Steven Day, a Professor of Seismology at San Diego State University and co-developer of the dynamic rupture components and efficient coarse-grained anelastic attenuation scheme in the AWP-ODC code. He has broad experience in computational seismology, including nonlinear mechanics of earthquake rupture and large-scale, linear and nonlinear stress wave propagation in heterogeneous earth models. He will be supporting the algorithm-level development by contributing to the concept, design and verification of AMR/DM-related development. He will also support the simulations of the dynamic rupture and wave propagation.

Mr. Philip Maechling, Associate Director and IT architect of SCEC. He will serve as the representative of SCEC, and be responsible for coordinating the US efforts to ensure the new model capabilities will be integrated to SCEC computational platforms. He is also responsible for overall data management of the project.

7. Mechanisms to Enable Sustained Collaborative Research

The project results will be managed, analyzed and disseminated via different mechanisms ranging from community-specific information portals, to bilateral center websites. The results of the software will be released to both nation’s communities for production simulations. An important mode of dissemination of results of this project to the scientific community will be through publications in peer-reviewed journals and presentations at scientific conferences. In addition, one workshop per year will be organized that will specifically involve scientists external to the project to facilitate use of the software, and solicit their input on project direction. The software resulting from the project will be made available to the wider scientific community both during and after the conclusion of the project. We will initiate bi-weekly teleconference calls via Skype, and monthly webinars, and in-person project workshops. Student exchange and/or PI on-site meetings will be planned on an annual basis. In addition to the internal project website, we will develop more generally accessible websites hosted at CAS and SDSC, that would be of interest to both scientific communities and the general public.

8. References

Bonilla, L., P.C. Luis, and S. Nielsen (2006). 1D and 2D linear and nonlinear site response in the Grenoble area, in *Third International Symposium on the Effects of Surface Geology on Seismic Motion*, Grenoble, France, 30 August – 1 September 2006, 82.

Cui, Y., K.B. Olsen, T. H. Jordan, K. Lee, J. Zhou, P. Small, D. Roten, G. Ely, D.K. Panda, A. Chourasia, J. Levesque, S. M. Day, P. Maechling (2010). Scalable earthquake simulation on petascale supercomputers, in Procs. *Supercomputing Conference* New Orleans, November 2010, ACM Gordon Bell Prize Finalist

Jiang, C.,X.S. Feng, J. Zhang and D. Zhong (2010). AMR Simulations of Magneto-hydrodynamic Problmes by the CESE Method in Curvilinear Coordinates, *Solar Physics*, 267, 463-491.

Drucker, D.C., and W. Prager (1952). Soil mechanics and plastic analysis or limit design, Q. *Appl. Math* 10, 157-165.

Iai, S., Y. Matsunaga, and T. Kameoka (1990). Strain space plasticity model for cyclic mobility, *Report of the Port and Harbour Institute* 29, 27-56.

Kristek, J., P. Moczo, and M. Galis (2010). Stable discontinuous staggered grid in the finite-difference modling of seismic motion, *Geophysical Journal* 183, 1401-1407

Ma, S., and D. J. Andrews (2010), Inelastic off-fault response and three-dimensional earthquake rupture dynamics on a strike-slip fault, *J. Geophys. Res.*, 115, B04304, doi:10.1029/2009JB006382.

MacNeice, P., Olson, K.M., Mobarry, C., de Fainchtein, R., Packer, C.: 2000, *Comput. Phys. Commun.* 126, 330.

Olsen, K. B., S. M. Day, L. Dalguer, J. Mayhew, Y. Cui, J. Zhu, V.M. Cruz-Atienza, D. Roten, P. Maechling, T.H. Jordan, and A. Chourasia (2009). ShakeOut-D: Ground Motion Estimates Using an Ensemble of Large Earthquakes on the Southern San Andreas Fault With Spontaneous Rupture Propagation, *Geophysical Research Letters* **36** L04303, doi:10.1029/2008GL036832.

Olsen, K.B., S. M. Day, J. B. Minster, Y. Cui, A Chouasia, R. Moore, P. Maechling, T. Jordan (2008), TeraShake2: simulation of Mw7.7 earthquakes on the southern San Andreas fault with spontaneous rupture description, *Bull. Seismol. Soc. Am.* **98**, 1162-1185, doi: 10.1785/0120070148.

Olsen, K. B., S. M. Day, J. B. Minster, Y. Cui, A. Chourasia, M. Faerman, R. Moore, P. Maechling, and T. Jordan (2006). Strong shaking in Los Angeles expected from southern San Andreas earthquake, *Geophysical Research Letters*, **33**, L07305, doi:10.1029/2005GL025472.

Olsen, K.B., W.J. Stephenson, and Andreas Geisselmeyer (2008). 3D Crustal Structure and Long-period Ground Motions From a M9.0 Megathrust Earthquake in the Pacific Northwest Region, *Jour. Seismol.* DOI 10.1007/s10950-007-9082-y.

Yang, X.J., X.K. Liao and J.Q. Song, The Tianhe-1A Supercomputer (2011): Its Hardware and Software, *J. of computer Science and Technology,* 26(3), May 2011, 344-351, DOI 10.1007/s11390-011-1137-4.

Zhou, J., Unat, D., Choi, D.J., Guest, C.C. and Cui, Y. (2012). Hands-on Performance Tuning of a 3D Finite Difference Earthquake Simulation on GPU Fermi Chipset, *International Conference on Computational Science*, Omaha, June 4-6, 2012 (accepted).