Progress of porting AWP-ODC to next generation HPC architectures and a 4-Hz Iwan-type nonlinear dynamic simulation of the ShakeOut scenario on TACC Frontera

Yifeng Cui (yfcui@sdsc.edu)[1], Daniel Roten [1], Akash Palla[1], Anish Govind[1], Scott Callaghan[2], Mattew Norman [3], Lars Koesterke[4], Wenyang Zhang[4], Philip Maechling[2]
[1]San Diego Supercomputer Center, [2]University of Southern California, [3]Oak Ridge National Laboratory, [4]Texas Advanced Computing Center

Summary

AWP-ODC is a 4th-order finite difference code used by the SCEC community for linear wave propagation[1,3,8], Iwan-type nonlinear dynamic rupture and wave propagation[10,11], and Strain Green Tensor simulation[2]. We have ported and verified the CUDA-version of AWP-ODC, including a reciprocal version used in the SCEC CyberShake project, to HIP so that it can also run on AMD GPUs [5,6]. This code achieved sustained 32.6 Petaflop/s performance and 95.6% parallel efficiency at full scale on Frontier, a Leadership Computing Facility at Oak Ridge National Laboratory. The readiness of this community software on AMD Radeon Instinct GPUs and EPYC CPUs allows SCEC to take advantage of exascale systems to produce more realistic ground motions and accurate seismic hazard products.

We present progress made for an early science project for the *Leadership-Class Computing Facility* (LCCF) at Texas Advanced Computing Center (TACC)[13], funded by NSF's *Characteristic Science Applications* (CSA) program. We also report some initial porting effort to Microsoft Azure[9], a cloud based system supported by Microsoft's Azure Accelerator for Research Program. In addition, we have ported Iwan nonlinear AWP-ODC code onto Nvidia A100, in the process of porting to HIP. The CPU-based dynamic modeling AWP-ODC code is ported to TACC Frontera and used to carry out a 4-Hz Iwan-type nonlinear dynamic simulation of the ShakeOut scenario at full system scale[11].

AWP-ODC Software Progress on Heterogeneous Machines

Newly added feature in AWP-ODC[4] includes GPU-Aware, which supports for passing GPU buffers directly to MPI calls, with 15% performance gain observed compared to the original configuration setup. This feature has been implemented on both HIP and CUDA AWP-ODC versions. OSU NOWLAB's on-the-fly message compression is enabled in AWP-ODC through the enhancement of lossless and lossy compression algorithms, MPC and ZFP, respectively[14]. The redesign in MVAPICH2 MPI library results in 19% and 37% improvement in the GPU computing flops in AWP on V100s, with enhanced MPC-OPT and ZFP-OPT schemes respectively[14]. This is the first work that leverages the GPU-based compression techniques to significantly improve the GPU communication performance in a real application[14]. On TACC Lonestar-6 A100s, we observed 48%-64% benefits using on-the-fly MPC compression using MPC over GDR[6]. Combined MVAPICH2-GDR enhancement over IMPI, including both CUDA-aware support and on-the-fly compression, improves application performance by 154% on 16 nodes[6]. We have also ported AWP-ODC to NVIDIA's latest H100 on Lonestar-6 at TACC, the initial benchmarking results are presented in the MLUPS Table below.

AWP-ODC benchmarks on TACC Lonestar-6 A100 nodes with 160x160x2048 per GPU configuration*

Lonestar6	mvapich2-2.3.7			mvapich2-2.3.7-gdr			mvapich2-2.3.7-gdr-compresson		
a100	gcc11.2.0			gcc11.2.0			gcc11.2.0		
nodes	Tflop/s	sec/step	parall eff.	Tflop/s	sec/step	parall eff.	Tflop/s	sec/step	parall eff.
2	2.0250	0.0488	100.0%	2.2960	0.0399	100.0%	3.7710	0.0261	100.0%
4	4.0270	0.0494	99.4%	4.5260	0.0436	98.6%	6.8510	0.0288	90.8%
8	7.8250	0.0510	96.6%	9.3250	0.0425	101.5%	13.7560	0.0288	91.2%
16	14.4130	0.1543	89.0%	17.1360	0.0460	93.3%	27.5580	0.0288	91.3%
	impi19.0.9			mvapich2-plus-3.0a2			mvapich2-plus-latest		
	gcc11.2.0			gcc11.2.0			gcc11.2.0		
	Tflop/s	sec/step	parall eff.	Tflop/s	sec/step	parall eff.	Tflop/s	sec/step	parall eff.
2	1.6800	0.0585	100.0%	2.391	0.0411	100.0%	3.151	0.0311	100.0%
4	3.4800	0.0572	103.6%	4.579	0.0431	95.8%	5.399	0.0366	85.7%
8	5.8170	0.0686	86.6%	7.796	0.0509	81.5%	10.136	0.0391	80.4%
16	10.8380	0.0737	80.6%	15.214	0.0523	79.5%	20.097	0.0395	79.7%

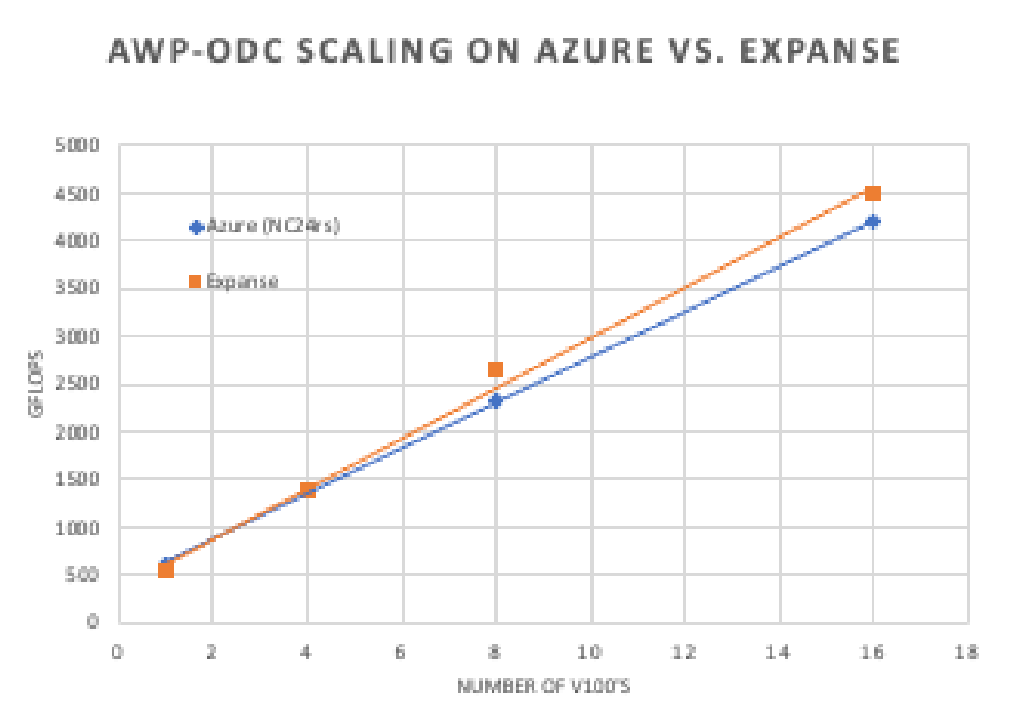
***mvapich2-2.3.7-gdr: runs with CUDA-aware; mvapich2-2.3.7-gdr-compresson: runs with CUDA-aware+on-the-fly compression (MPC)**

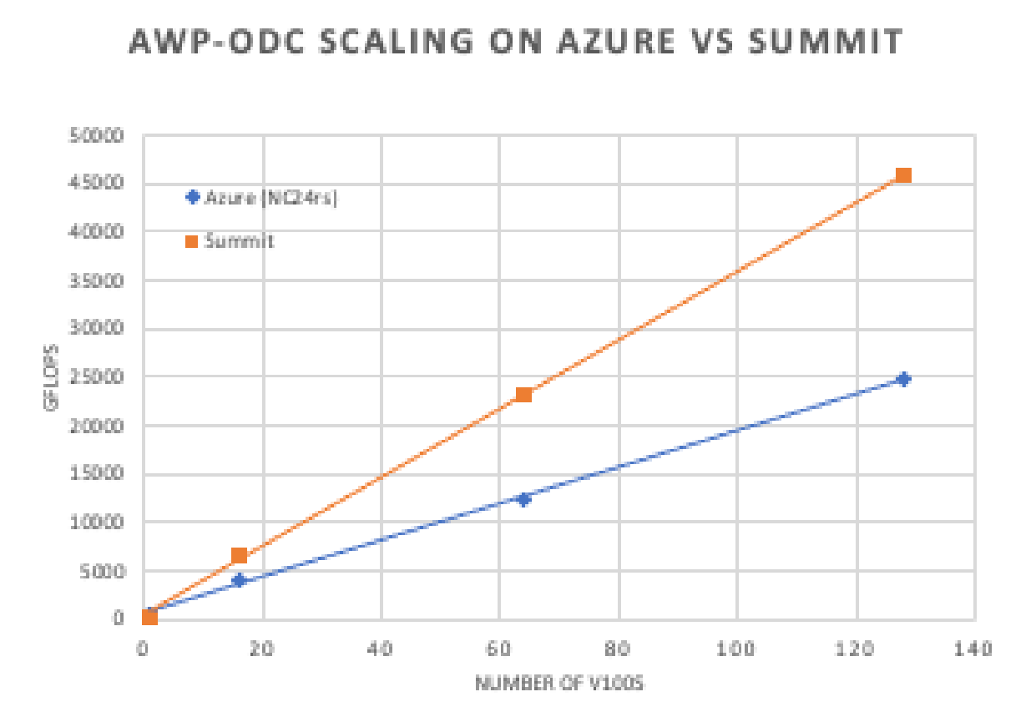
AWP-ODC	K20X	KNL7250	V100 (NVLink)	V100 (PCIe)	V100 (PCIe+Opt)	A100 (NVLink)	H100 (PCIe)	H100 (PCIe+Opt)	MI250X (Slingshot)
MLUPS**	552	1092	1598	1074	2009	1937	3713	5145	1711
Speedup	1x	1.98x	2.89x	1.95x	3.64x	3.51x	6.72x	9.32x	3.10x

****Millions of lattice point update completed per second**

Accelerating AWP-ODC on Microsoft Azure Cloud

We have deployed AWP-ODC to Azure to leverage the array of tools and services that Azure provides for tightly coupled HPC simulation on commercial cloud. This project is collaborated with Internet 2/Azure Accelerator supporting team, as part of Microsoft Internet2/Azure Accelerator for Research Fall 2022 Program. Azure credits were awarded through Cloudbank, an NSF-funded initiative.

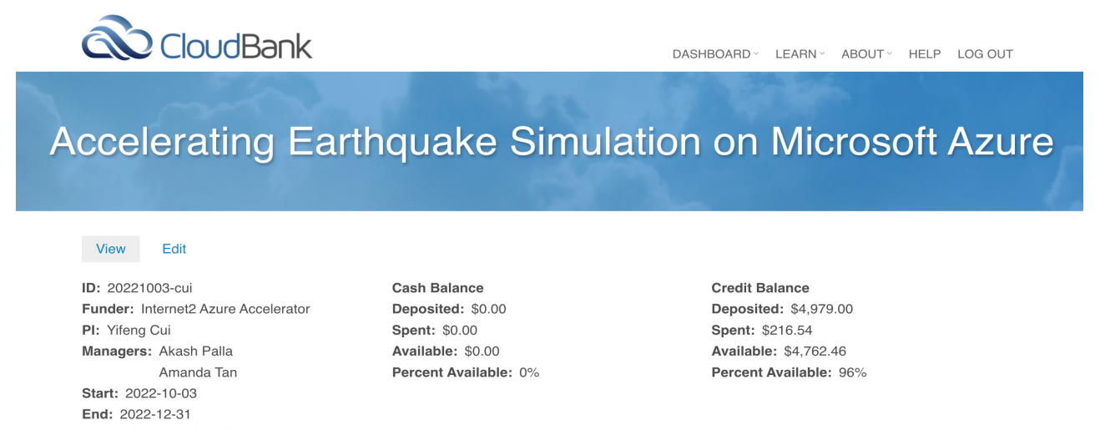




We demonstrated the AWP performance with a benchmark of ground motion simulation on various GPU based cloud instances, and a comparison of the cloud solution to on-premises bare-metal systems[9].

Challenges

- Digesting the wide breadth of options and configurations
- Higher threshold of initial setup needed
- Lack of comprehensive forums for debugging errors



Accelerating Earthquake Simulation on Microsoft Azure

ID: 20221003-cui
Funder: Internet2 Azure Accelerator
PI: Yifeng Cui
Managers: Akash Palla
Amanda Tan
Start: 2022-10-03
End: 2022-12-31
Requested Clouds: Microsoft Azure

Cash Balance
Deposited: $0.00
Spent: $0.00
Available: $0.00
Percent Available: 0%

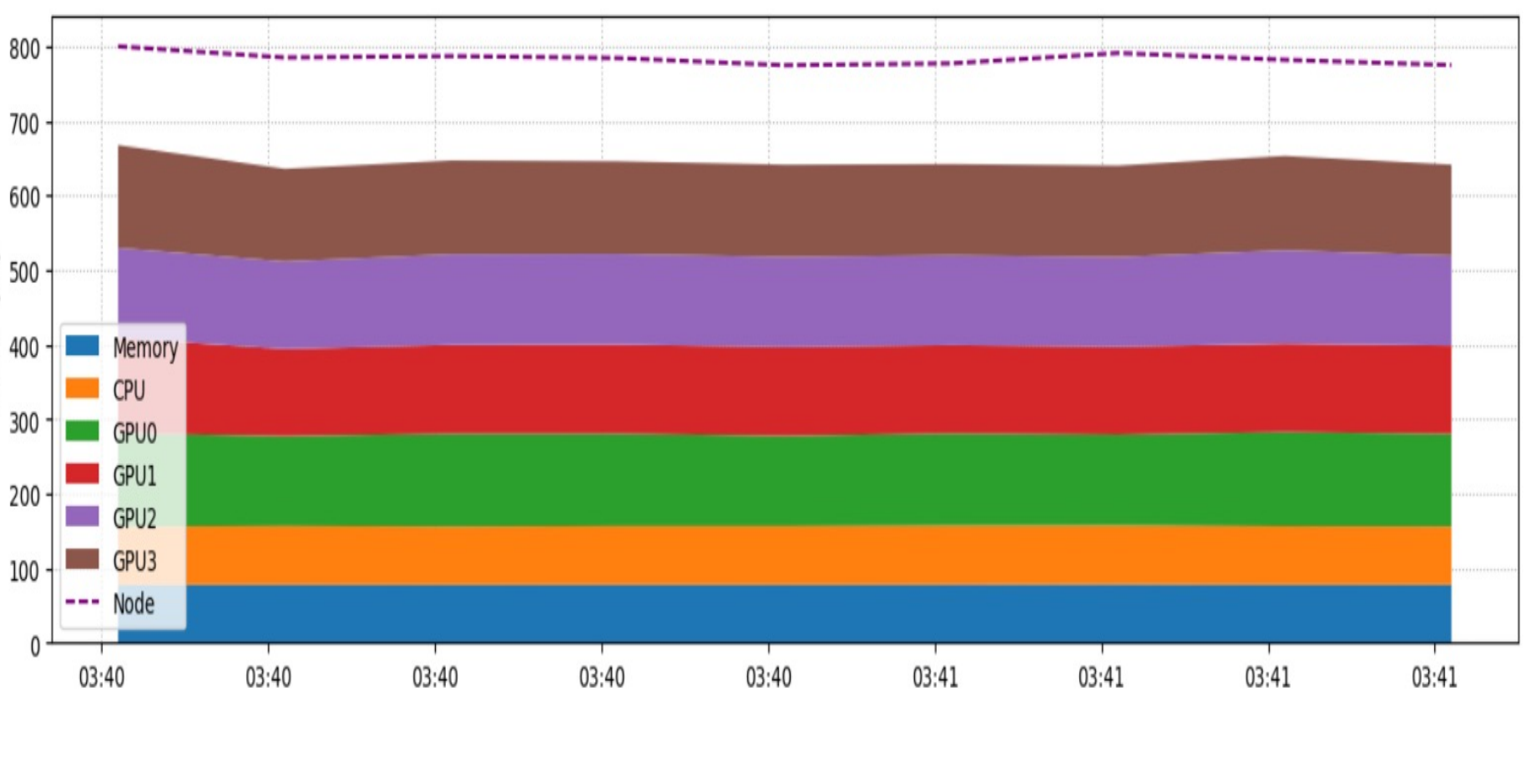
Credit Balance
Deposited: $4,979.00
Spent: $216.54
Available: $4,762.46
Percent Available: 96%

Microsoft Internet2/Azure Accelerator for Research Fall 2022 program, credits awarded through NSF Cloudbank program[9]

Benefits

- Wide flexibility and options of hardware and software allows infrastructure to be tailored to specific workload
- Spin up large VM instances instantly without waiting in a queue/system quotas

AWP-ODC power consumption study on NERSC Perlmutter (Govind, 2023)

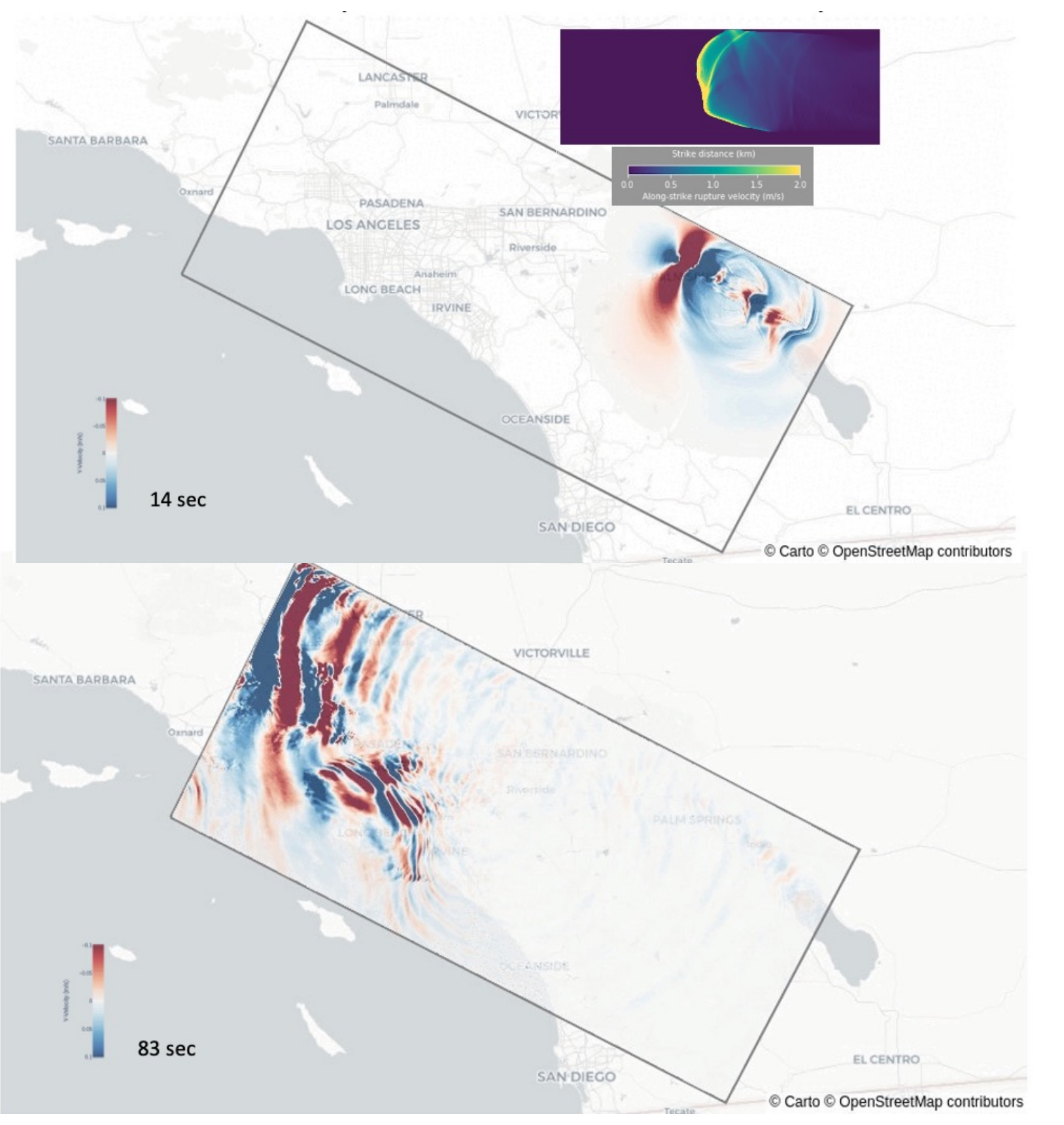


4-Hz Iwan-type nonlinear dynamic simulation of the ShakeOut scenario

Large-scale simulation assumed linear rheology, but it is well known that waves with sufficiently large amplitudes propagating into soft near-surface materials are affected by nonlinear behavior, which can modify the seismic ground motion significantly. Until recently, calculations which simulate the nonlinear response of soft soils neglects the interplay between multi-dimensional and potentially nonlinear effects, caused by phenomena such as long period-surface waves, finite fault effects, and fault damage zones. Using TACC LSCP allocation, we carried out an advanced Iwan-type nonlinear earthquake dynamic modeling for a Mw7.8 earthquake on the southern San Andreas fault during Texascale Days April 13-14, 2023[12].

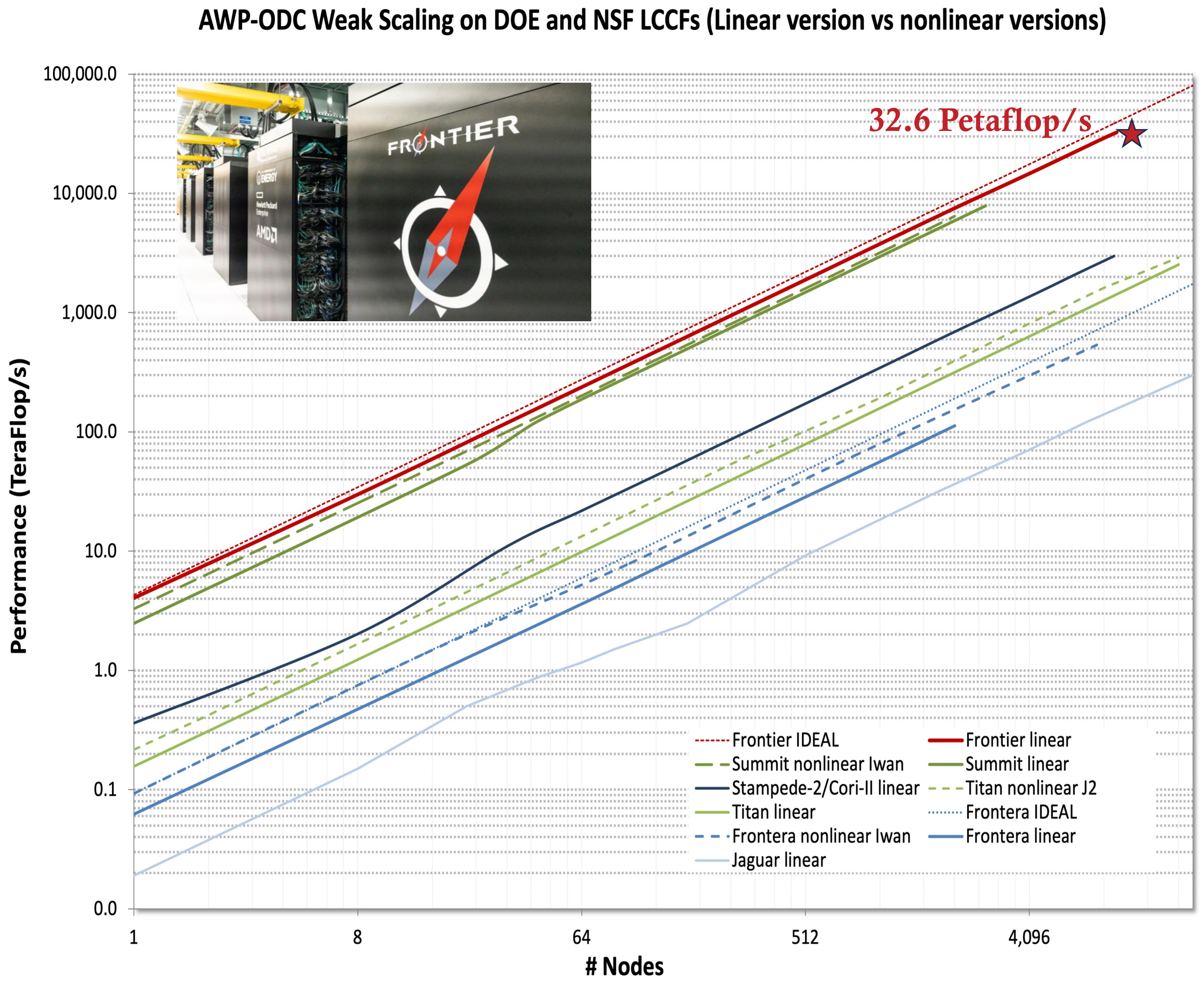
The dynamic simulation was the M7.8 ShakeOut scenario with 10 yield surfaces using Iwan nonlinear model using a resolution of 25m and minimum shear-wave velocity of 500 m/s. Rupture propagation is simulated from SE to NW to obtain maximum excitation of the waveguide due interconnected basin in the northern Los Angeles area[12].

An Iwan nonlinearity calculation is quite expensive compared to a linear calculation or even some commonly-used nonlinear calculations. It increases memory and computation requirements significantly, since 6 stress components and 6 memory variables associated with each individual yield surface need to be tracked and updated at run-time[11]. With respect to nonlinear simulations using a single yield surface (J2 model), the memory consumption of the Iwan model 5x-13x larger, and the computational time is 10x-30x longer. In this simulation, Masing unloading and reloading behaviors was simulated by tracking an overlay of 10 concentric von Mises yield surfaces, and the stress-strain relationship was approximated by taking the sum (overlay) of the stress associated with each surface[11]. The simulation was the first large-scale dynamic simulation using AWP-ODC, with combined spontaneous rupture and wave propagation calculations in a single step to simulate a SE-NW rupturing M7.8 earthquake with realistic nonlinear rheology in the sedimentary infill of the San Gabriel and Los Angeles basins[11]. The result is compared with earlier simulations of the same scenario that used either a 100m resolution or a simple nonlinear J2 model.



Y-direction velocity snapshot of 0-4 Hz dynamic simulation of a M7.8 earthquake scenario on the southern San Andreas fault, fault-parallel velocity at 83 seconds using the Iwan nonlinear model at a resolution of 25m and minimum shear-wave velocity of 500 m/s. The simulation took 22.5 hours on TACC Frontera's 7,680 nodes (Simulation by Yifeng Cui and Daniel Roten of SDSC, image by Akash Palla of UCSD).

CUDA AWP-ODC Ported to HIP on AMD MI250X and Verified



Frontier[7] at ORNL is the current No. 1 system in the June 2023 TOP500 list. This HPE Cray EX system based on the AMD Radeon Instinct GPUs and EPYC CPUs is the first US system with a peak performance exceeding one ExaFlop/s[5]. Both CUDA based AWP-ODC-GPU and AWP-ODC-SGT have been ported and verified on this AMD MI250X based system. *Frontier* weak scaling efficiency is achieved in 95.6% on 9,248 nodes, with 32.6 Petaflop/s benchmarked in sustained performance[6]. This implementation also includes GPU-aware MPI. The Iwan and discontinuous-mesh based AWP code is being ported to HIP.

Acknowledgements

This work is supported by the NSF CSA program under Award #2139536, and SCEC RSA #162830509. SCEC is funded by National Science Foundation (NSF) Cooperative Agreement EAR-1600087 and U.S. Geological Survey (USGS) Cooperative Agreement G17AC00047. The authors thank Dr. Ossian O'Reilly of AMD for providing HIP porting advice, and Prof. DK Panda NOWLAB team of OSU, especially Qinghua Zhou, for CUDA-Aware and on-the-fly message compression support. The authors acknowledge the Office of Science of the U.S. Department of Energy (DOE) for providing computing resource supports resources through the ASCR Leadership Computing Challenge (ALCC) allocation award and OLCF for DD allocation; TACC for Frontera LSCP allocation and CSA allocation, ACCESS for Delta allocation, SDSC for Director's Discretionary allocation on Expanse, and AMD for providing resource access on AMD Accelerator Cluster (AAC). This work is also supported by SDSC core program.

References

[1] Cui, Y., K.B. Olsen, T.H. Jordan, K. Lee, J. Zhou,, P. Small, G. Ely, D. Roten, DK Panda, A. Chourasia, J. Levesque, S.M. Day and P. Maechling, Scalable Earthquake Simulation on Petascale Supercomputers, SC10, New Orleans, Nov, 13-19, 2010

[2] Cui, Y., Poyraz, E., Zhou, J., Callaghan, S., Maechling, P., Jordan, T., Shih, L. and Chen, P., Accelerating CyberShake Calculations on XE6/XK7 Platforms of Blue Waters, Proceeding of 2013 Extreme Scaling Workshop, 8-17, I EEE Xplore Digital Library, ISSN 2381-1986, doi: 10.1109/XSW.2013.6, 20, August 15-16, Boulder, 2013.

[3] Cui, Y., E. Poyraz, K.B. Olsen, J. Zhou, K. Withers, S. Callaghan, J. Larkin, C. Guest, D. Choi, A. Chourasia, Z. Shi, S.M. Day, P.J. Maechling, T.H. Jordan (2013), Physics-based seismic hazard analysis on petascale heterogeneous supercomputers, SC13, Denver, CO, November 18-21, 2013.

[4] Cui, Y., Zhou, J., Poyraz, E., Choi, D. J. (2016). AWP-ODC-OS (v1.0), Open source releases under BSD-2 clause license, available from https://github.com/HPGeoC/awp-odc-os

[5] Cui, Y., S. Callaghan, O. O'Reilly, Y. Ren, L. Koesterke, W. Zhang, M. Norman, D. Roten, C. Goulet and P. Maechling. Porting AWP-ODC to next -generation architectures: the Characteristic Scientific Application Project. SCEC Annual Meeting, Sept 11-13, Palm Springs, 2022. https://www.scec.org/publication/12322.

[6] Cui, Y. , Extreme-scale Earthquake simulation with MVAPICH, MUG'23, Columbus, Aug 21-23, 2023.

[7] Frontier: https://www.olcf.ornl.gov/frontier/

[8] Olsen, K.B., Simulation of Three-Dimensional Wave Propagation in the Salt Lake Basin, doctoral dissertation, Univ. of Utah, 1994, p. 157.

[9] Palla, A. and Y. Cui. Internet2/Azure Accelerator 2022 Research Project Final Presentation, Dec 7, 2022. https://internet2.edu/internet2-microsoft-azure-accelerator-2022-researchers-in-action/.

[10] Roten, D., Y. Cui, K. Olsen, S. Day, K. Withers, W. Savran, P. Wang and D. Mu, High-frequency nonlinear earthquake simulations on petascale heterogeneous supercomputers, SC'16, 1-10, Nov 13-18, Salt Lake City, 2016.

[11] Roten, D., T. Yeh, K. Olsen, S. Day and Y. Cui. Implementation of Iwan-type nonlinear rheology in a 3D high-order staggered-grid finite-difference method. accepted to BSSA, 2023.

[12] Roten, D. and Cui Y., Nonlinear dynamic modeling for a M7.8 earthquake on the southern San Andreas fault, Frontera User Meeting, Austin, Aug 3-4 2023.

[13] TACC CSA Award In the News. Available at: https://www.tacc.utexas.edu/-/21-scientific-codes-selected-for-new-high-performance-software-improvement-program.

[14] Zhou, Q. N. Kumar, P. Kousha, S. Ghazimirsaeed, H. Subramoni and DK Panda, Designing High-Performance MPI Libraries with On-the-fly Compression for Modern GPU Clusters, *2021 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 2021, pp. 444-453, doi: 10.1109/IPDPS49936.2021.00053.