AWP-ODC with Topography and Discontinuous Mesh:
Extreme-Scale Earthquake Simulation using MVAPICH

Yifeng Cui, San Diego Supercomputer Center
Te-Yang Yeh, San Diego State University
MUG’24, Aug 19-21, 2024

Acknowledgments

HPGeoC Team

Akash Palla

Arnav Talreja

Daniel Roten

Collaborators

Philip Maechling

Ossian O'Reilly

Kim Olsen

Lars Koesterke

NOWLAB Team

DK Panda

Hari Subramoni

Qinghua Zhou

Lang Xu

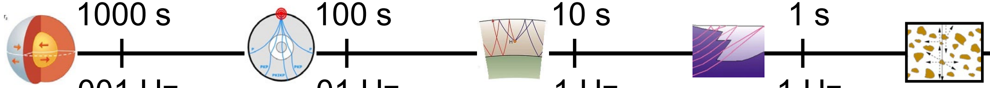
Computing Allocation OLCF DD, TACC LSCP and CSA, ACCESS Delta, SDSC Expanse, AMD AAC, DOE INCITE & ALCC

Funding NSF CSSI, LCCF/CSA, NSF/USGS SCEC Core, SDSC

Advanced Earthquake Modeling with Nonlinearity and Topography in AWP-ODC

Seismic band

low-order free oscillations *mantle waves* *crustal waves* *basin waves* *strongly scattered waves*



1000 s	100 s	10 s	1 s	0.1 s	period
.001 Hz	.01 Hz	.1 Hz	1 Hz	10 Hz	frequency

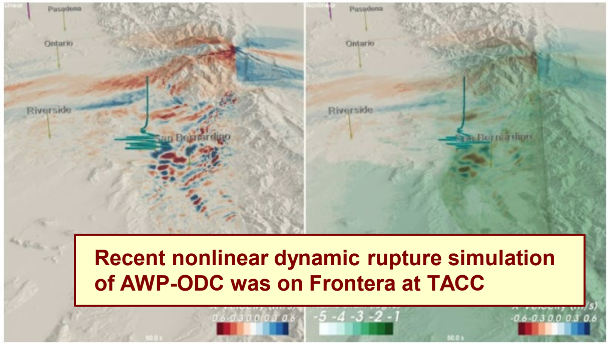
tall buildings *houses* *stiff structures*

Earthquake engineering band

SAF simulation 2023

physics-based deterministic — **4 Hz** — *empirical stochastic*

CyberShake 1 Hz



Recent nonlinear dynamic rupture simulation of AWP-ODC was on Frontera at TACC

(Snapshots from linear (left) and nonlinear (right) simulations using AWP-ODC showing wave propagation during a magnitude 7.7 SAF earthquake, Roten, SC'16)

Must validate new physics

- fault roughness
- near-fault plasticity
- frequency-dependent attenuation
- topography — **1.3x more expensive vs linear**
- small-scale near-surface heterogeneity
- multi-surface nonlinearity — **20x more expensive vs linear**

AWP-ODC

- Started as personal research code (Olsen 1994)
- 3D velocity-stress wave equations

$$\partial_t v = \frac{1}{\rho}\nabla\cdot\sigma \qquad \partial_t\sigma = \lambda(\nabla\cdot v)\mathrm{I} + \mu(\nabla v + \nabla v^{\mathrm{T}})$$

solved by explicit staggered-grid 4th-order FD

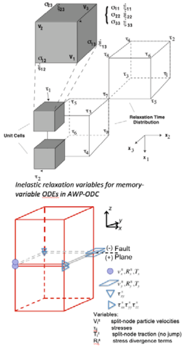
- Memory variable formulation of inelastic relaxation

$$\sigma(t) = M_u\left[\varepsilon(t) - \sum_{i=1}^{N}\varsigma_i(t)\right] \qquad \tau_i\frac{d\varsigma_i(t)}{dt} + \varsigma_i(t) = \lambda_i\frac{\delta M}{M_u}\varepsilon(t)$$

$$Q^{-1}(\omega) \approx \frac{\delta M}{M_u}\sum_{i=1}^{N}\frac{\lambda_i\omega\tau_i}{\omega^2\tau_i^2+1}$$

using coarse-grained representation (Day 1998)

- Dynamic rupture by the staggered-grid split-node (SGSN) method (Dalguer and Day 2007)
  - Displacement nodes split at fault surface: explicitly discontinuous displacement & velocity
  - All interactions between sides occur through traction vector at displacement node
- Absorbing boundary conditions by perfectly matched layers (PML) (Marcinkovich and Olsen 2003) and Cerjan et al. (1985)



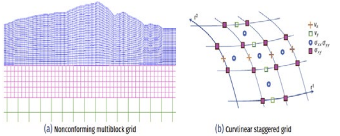
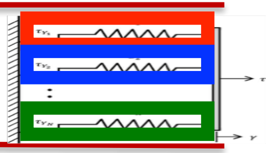


Figure 1: (a) Curvilinear grid, used for discretizing topography, overlaying cartesian grids with decreasing grid resolution with depth. (b) Arrangement of velocity and stresses in a curvilinear staggered grid

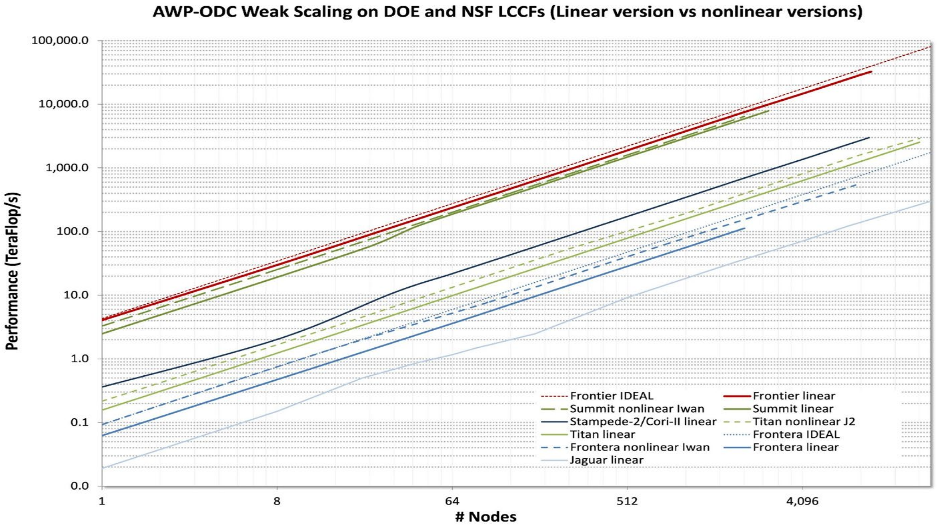


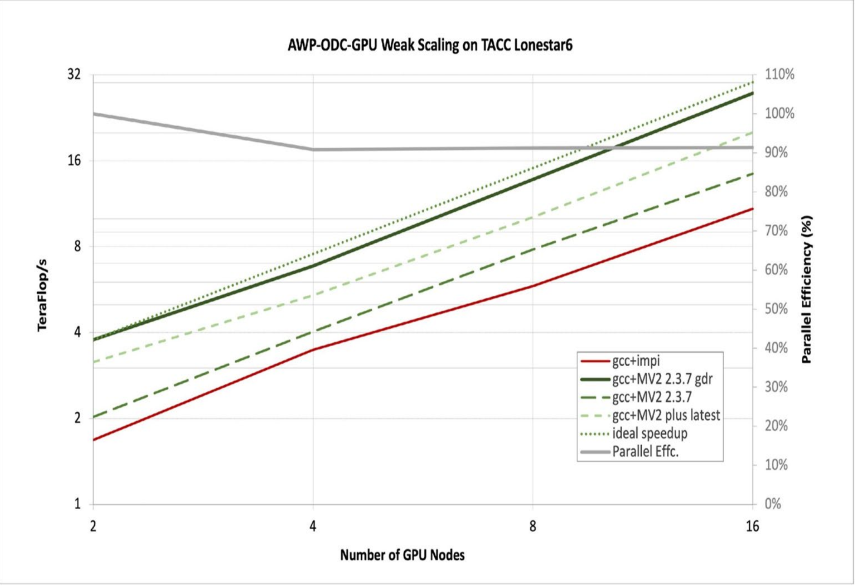
8 Hz

SAF Simulation 2026

physics-based deterministic — *physics-based stochastic* — *empirical stochastic*

AWP-ODC Performance and MV2 Evaluation



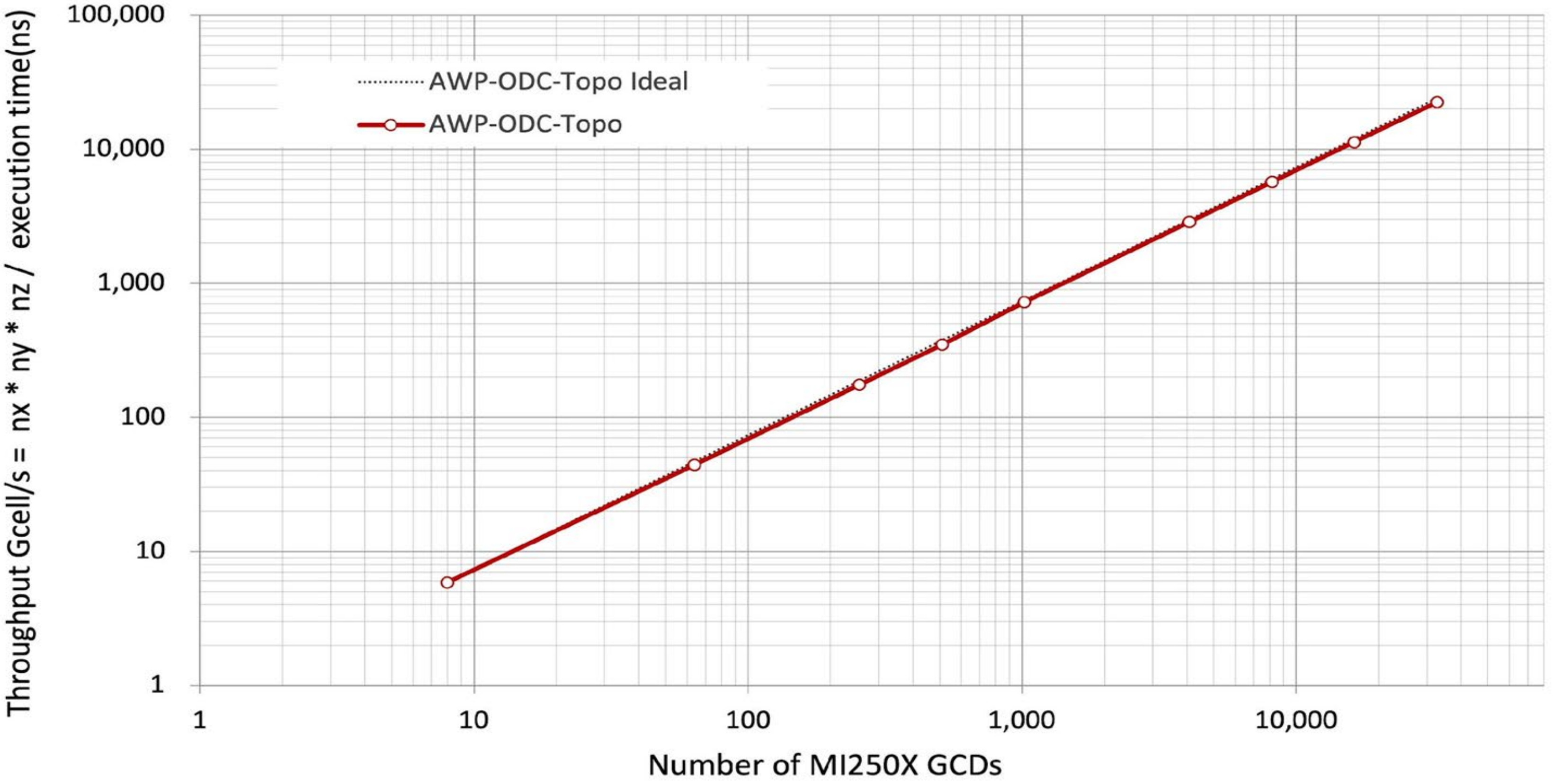


AWP-ODC simulation allocation annually ca. 200-300M core-hours in recent years, supported by DOE INCITE/ALCC and NSF LSCR (TACC) computing programs

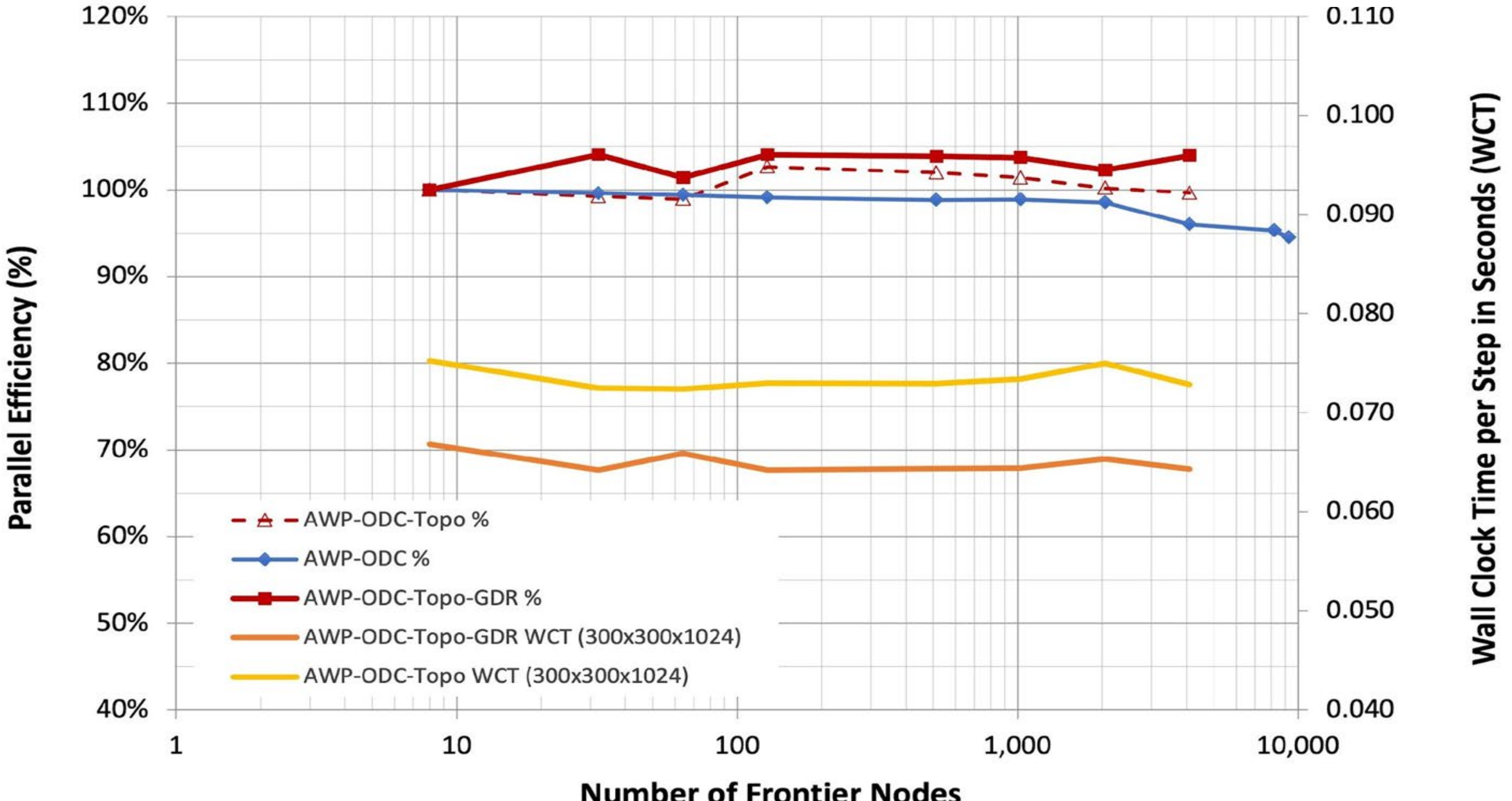
AWP-ODC	K20X	KNL7250	V100 (NVLink)	A100 (NVLink) Openmpi	A100 (PCIe) impi	A100 (PCIe+Opt) MV2-gdr-MPC	H100 (PCIe)	H100 (PCIe+Opt)	MI250X (Slingshot)	GH200
MLUPS**	552	1092	1598	1937	896	2009	3713	5145	1711	8480
Speedup	1x*	1.98x	2.89x	3.51x	1.62x	3.64x	6.72x	9.32x	3.10x	15.36x

* 160x160x2048 per GPU configuration ** Millions of lattice point update completed per second

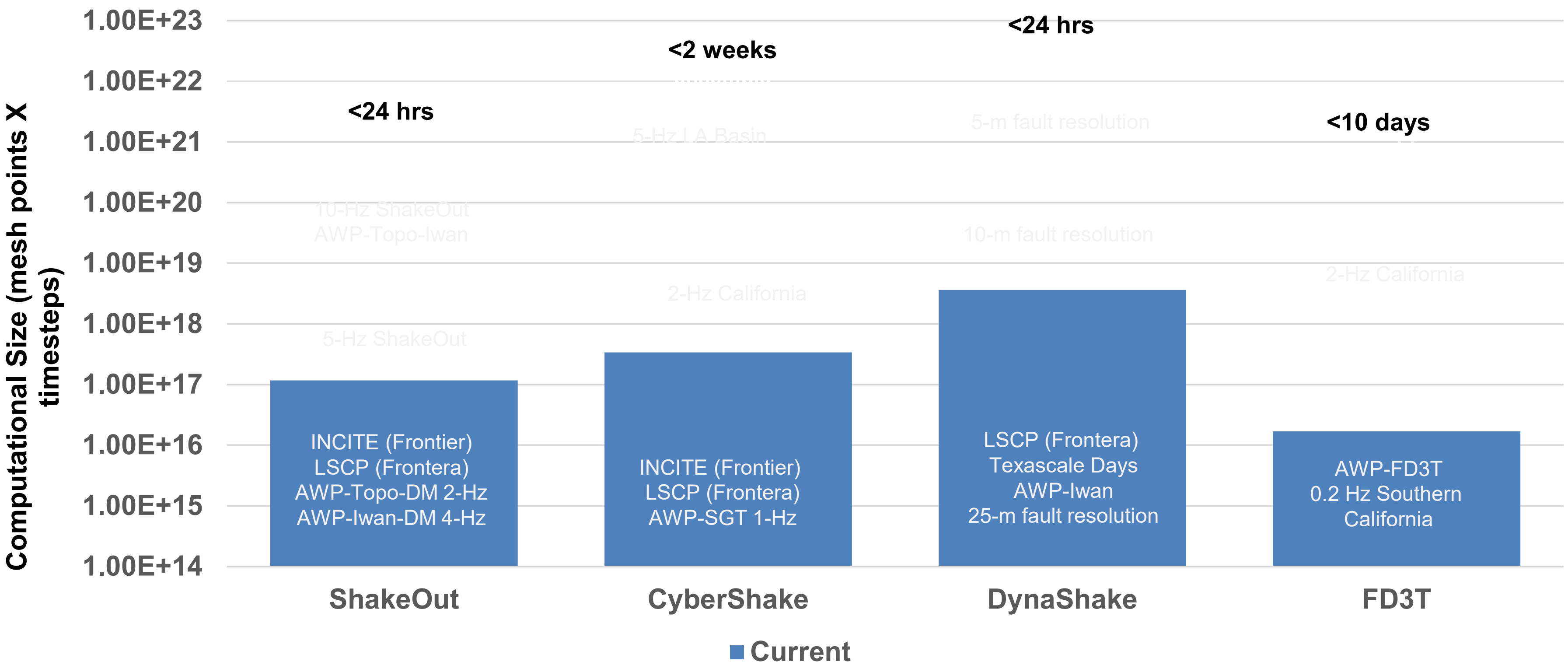
AWP-ODC-Topo Weak Scaling on Frontier



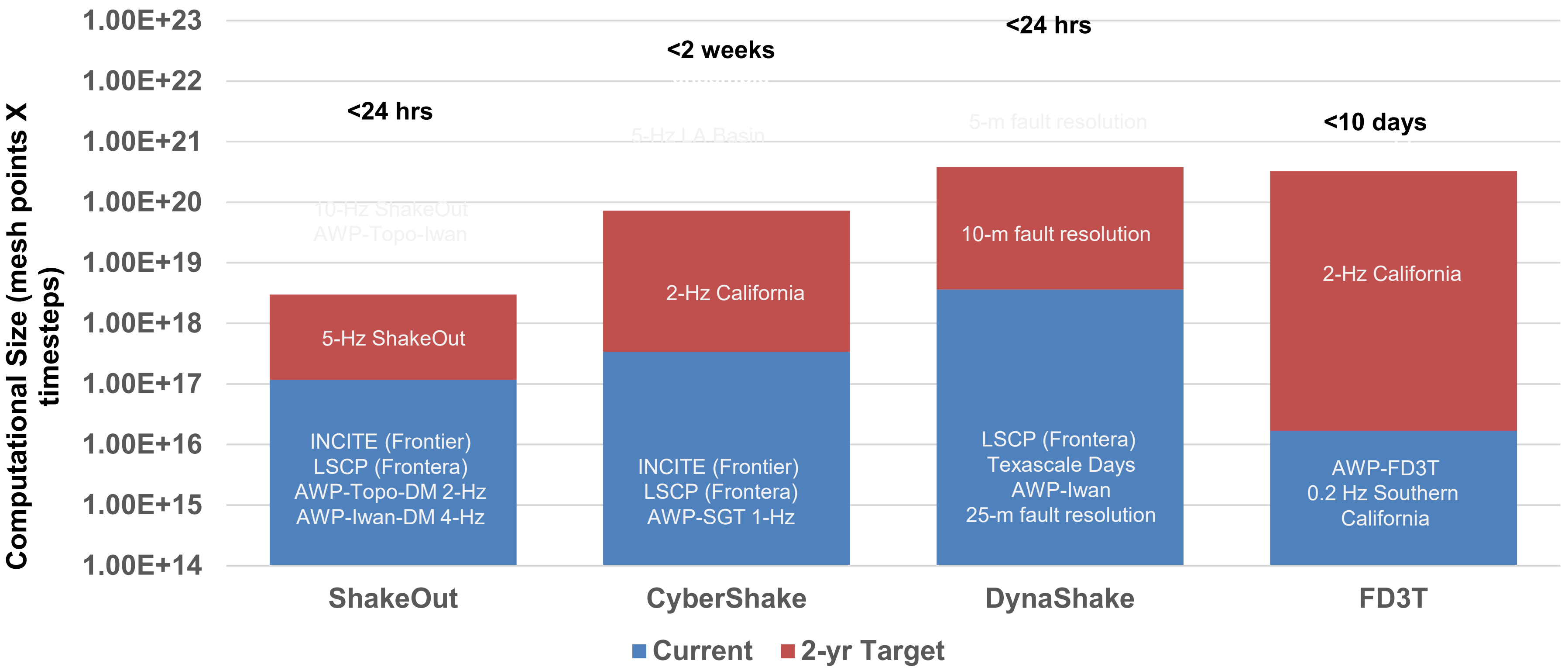
AWP-ODC-Topo w/ and w/o ROCm-Aware on Frontier



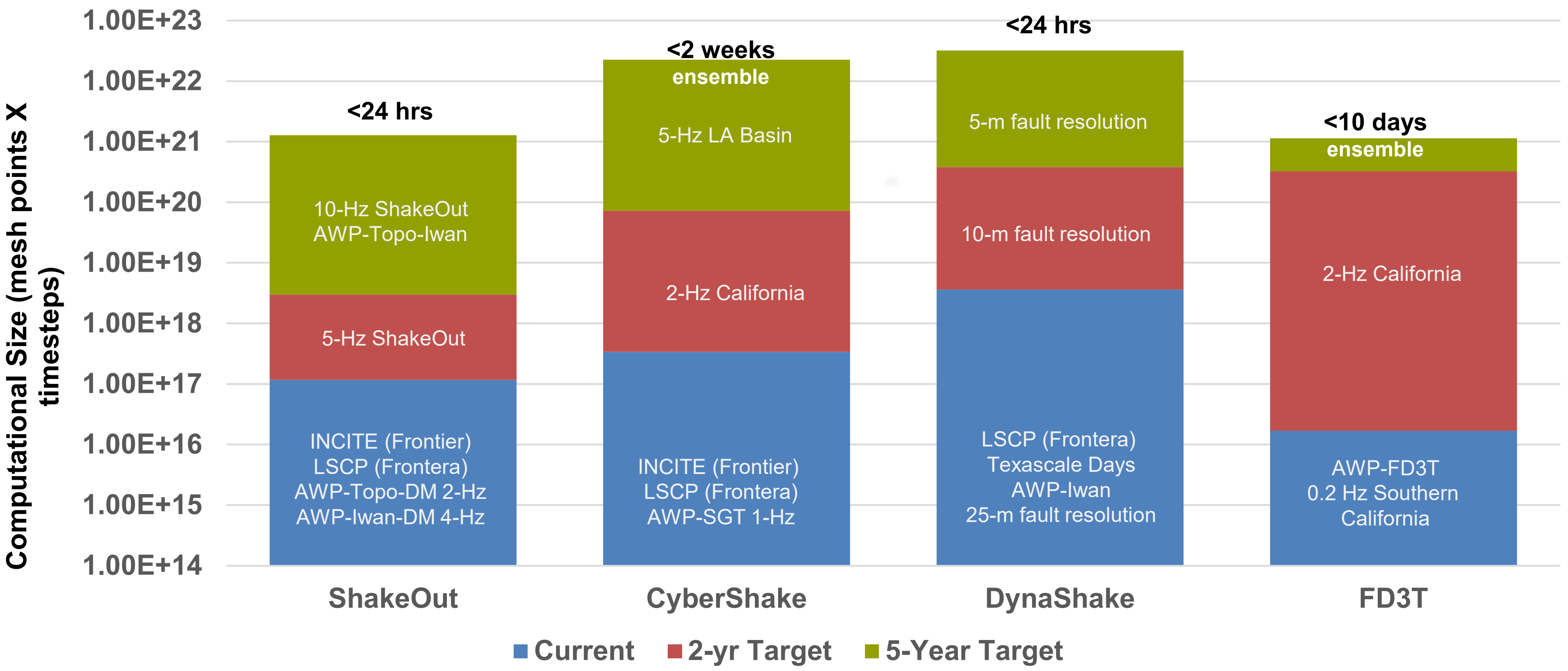
Computational Requirements of AWP-ODC



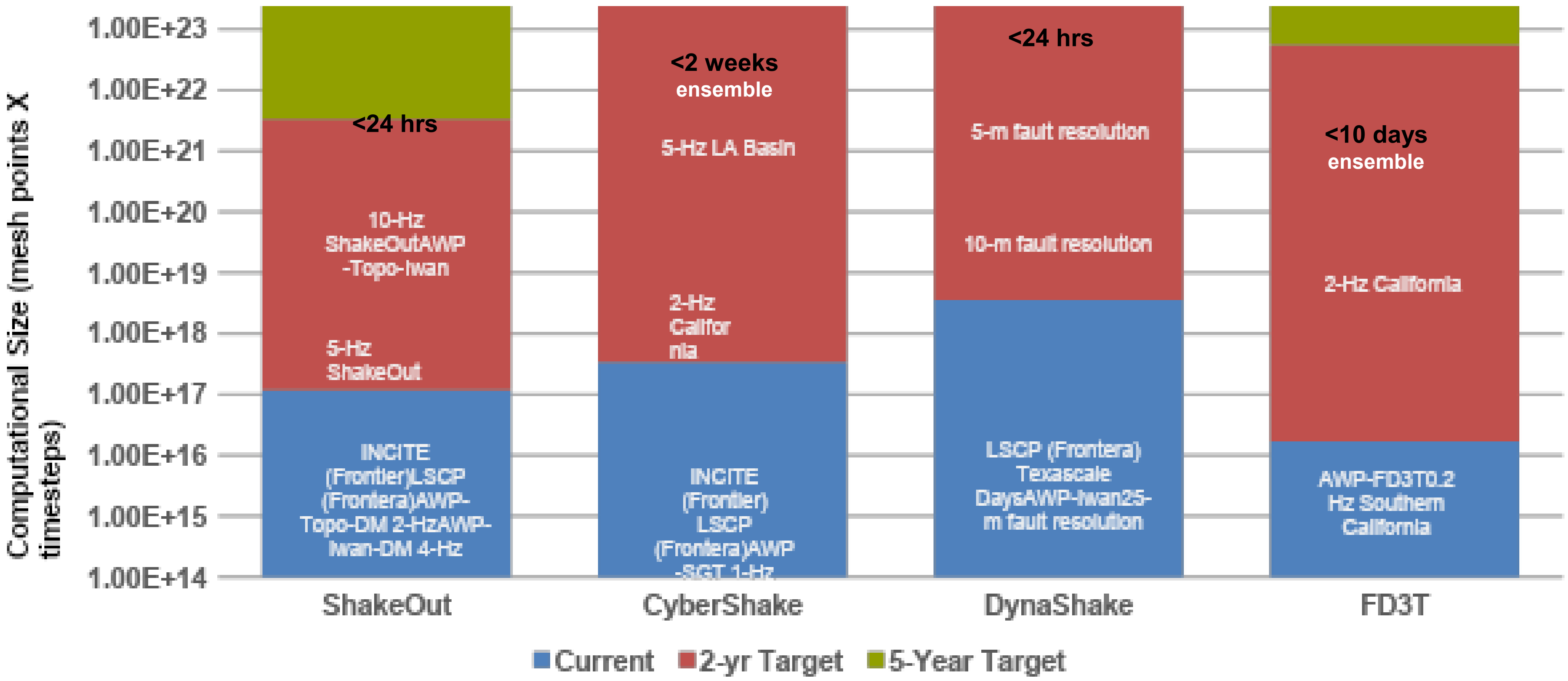
Computational Requirements of AWP-ODC



Computational Requirements of AWP-ODC



Computational Requirements of AWP-ODC



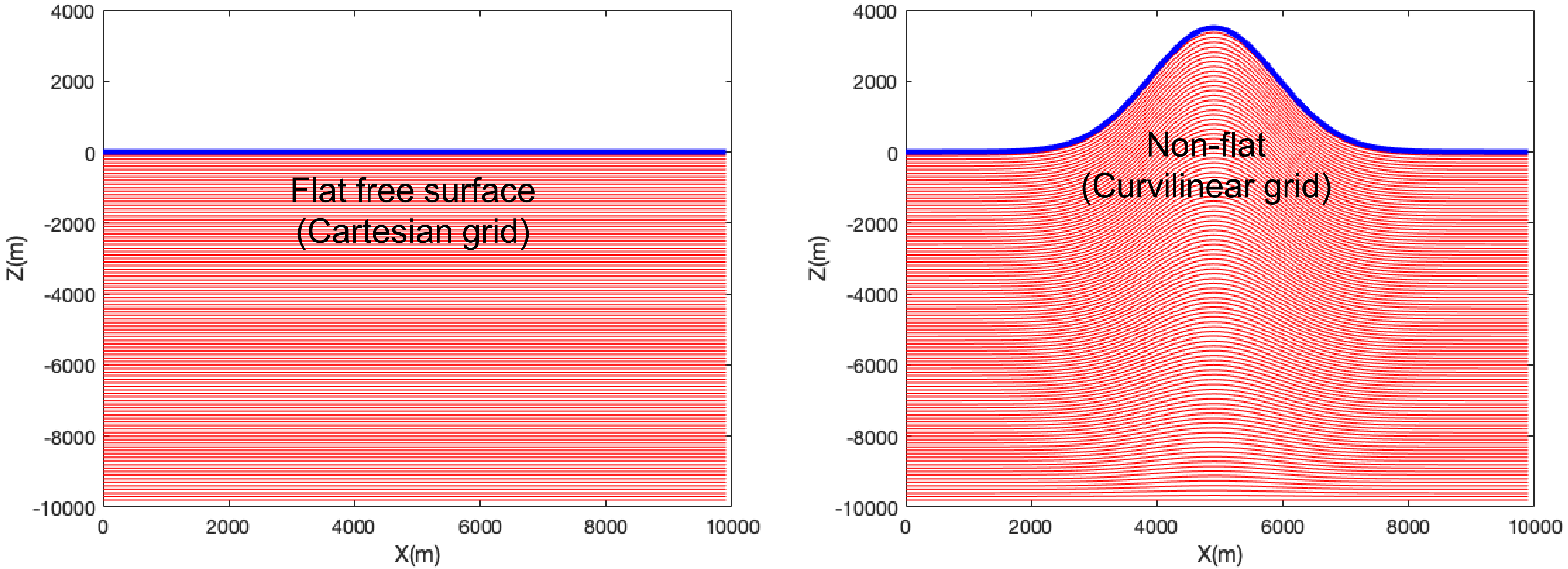
AWP-ODC - Recent Advancements

- 4th-order accurate staggered-grid finite-difference code
- GPU-enabled and highly scalable (Cui et al., 2013), with **curvilinear grid** and **discontinuous mesh** features
- Developed and verified on Summit at Oak Ridge Leadership Computing Facility (OLCF), with Nvidia GPUs
- Verified ported HIP version, full production runs on Frontier at OLCF (AMD GPUs)
- Phenomenal scalability with new features

Implementation of Curvilinear Grid

- Implementing traction free boundary using curvilinear grid (O'Reilly et al., 2021)
- Horizontal locations of grids remain the same
- Vertical grid stretching
- Curvilinear mapping

$$x_1 = F_1(r^1, r^2, r^3), \quad x_2 = F_2(r^1, r^2, r^3), \quad x_3 = F_3(r^1, r^2, r^3)$$



Governing Equations with Curvilinear Grid

Covariant basis vector (tangential):

$$\mathbf{a}_i = \begin{bmatrix} \frac{\partial x_1}{\partial r^i} \\ \frac{\partial x_2}{\partial r^i} \\ \frac{\partial x_3}{\partial r^i} \end{bmatrix}, \quad a_{ij} = \frac{\partial x_j}{\partial r^i}$$

Contravariant basis vector (orthogonal):

$$\mathbf{a}^j = \begin{bmatrix} \frac{\partial r^j}{\partial x_1} \\ \frac{\partial r^j}{\partial x_2} \\ \frac{\partial r^j}{\partial x_3} \end{bmatrix}, \quad a^{ji} = \frac{\partial r^j}{\partial x_i}$$

Establishing orthogonality:

$$\mathbf{a}_i \cdot \mathbf{a}^j = \delta_{ij}$$

Governing equation in Cartesian coordinate:

$$\rho \frac{\partial v_i}{\partial t} = \sum_{j=1}^{3} \frac{\partial \sigma_{ij}}{\partial x_j},$$

$$\frac{\partial \sigma_{ij}}{\partial t} = \lambda \sum_{k=1}^{3} \frac{\partial v_k}{\partial x_k} \delta_{ij} + \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right)$$

$$\downarrow$$

Governing equation in curvilinear coordinate:

$$\rho \frac{\partial v_i}{\partial t} = \frac{1}{J} \sum_{k,j} \frac{\partial}{\partial r^k} (J a^{kj} \sigma_{ij}),$$

$$\frac{\partial \sigma_{ij}}{\partial t} = \lambda \sum_{k,l} \frac{\partial v_k}{\partial r_l} a^{lk} \delta_{ij} + \mu \sum_{l} \left(\frac{\partial v_i}{\partial r^l} a^{lj} + \frac{\partial v_j}{\partial r^l} a^{li} \right)$$

$$\rho \frac{\partial v_i}{\partial t} = \frac{1}{J}\left(\frac{\partial}{r^1}\left(Ja^{11}\sigma_{i1}\right) + \frac{\partial}{r^2}\left(Ja^{21}\sigma_{i1}\right) + \frac{\partial}{r^3}\left(Ja^{31}\sigma_{i1}\right)\right)$$
$$+ \frac{1}{J}\left(\frac{\partial}{r^1}\left(Ja^{12}\sigma_{i2}\right) + \frac{\partial}{r^2}\left(Ja^{22}\sigma_{i2}\right) + \frac{\partial}{r^3}\left(Ja^{32}\sigma_{i2}\right)\right)$$
$$+ \frac{1}{J}\left(\frac{\partial}{r^1}\left(Ja^{13}\sigma_{i3}\right) + \frac{\partial}{r^2}\left(Ja^{23}\sigma_{i3}\right) + \frac{\partial}{r^3}\left(Ja^{33}\sigma_{i3}\right)\right),$$

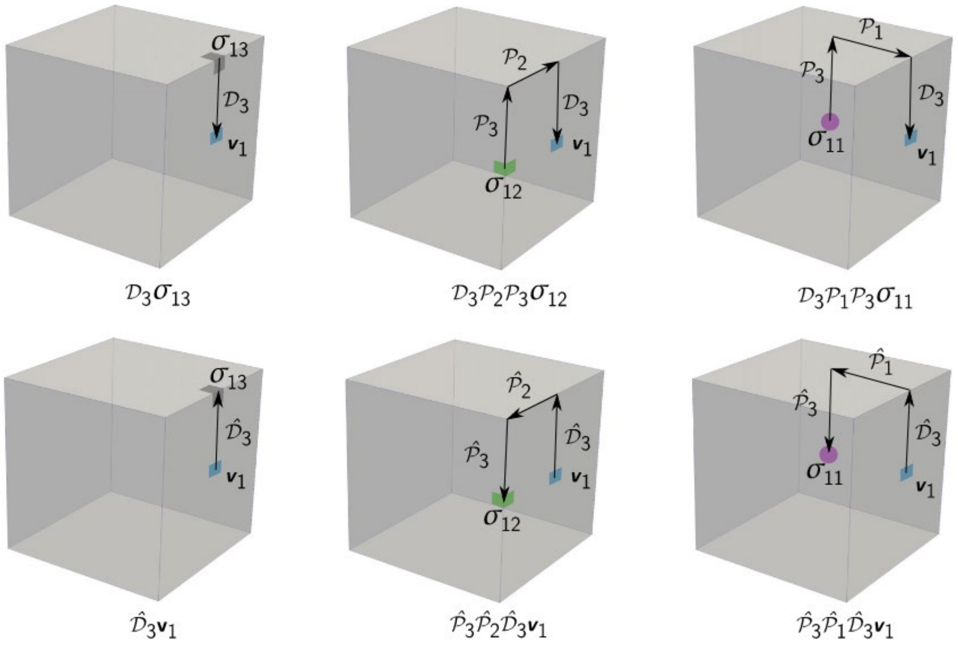
$$\frac{\partial \sigma_{ij}}{\partial t} = \lambda\left(\left(\frac{\partial v_1}{\partial r_1}a^{11} + \frac{\partial v_1}{\partial r_2}a^{21} + \frac{\partial v_1}{\partial r_3}a^{31}\right)\right.$$
$$+ \left(\frac{\partial v_2}{\partial r_1}a^{12} + \frac{\partial v_2}{\partial r_2}a^{22} + \frac{\partial v_2}{\partial r_3}a^{32}\right)$$
$$\left. + \left(\frac{\partial v_3}{\partial r_1}a^{13} + \frac{\partial v_3}{\partial r_2}a^{23} + \frac{\partial v_3}{\partial r_3}a^{33}\right)\right)\delta_{ij}$$
$$+ \mu\left(\left(\frac{\partial v_i}{\partial r^1}a^{1j} + \frac{\partial v_i}{\partial r^2}a^{2j} + \frac{\partial v_i}{\partial r^3}a^{3j}\right)\right.$$
$$\left. + \left(\frac{\partial v_j}{\partial r^1}a^{1i} + \frac{\partial v_j}{\partial r^2}a^{2i} + \frac{\partial v_j}{\partial r^3}a^{3i}\right)\right).$$

Recovering the Cartesian form when $a^{kj} = \delta_{kj}$ and $J = 1$

a^{kk} Diagonal terms, easy to implement

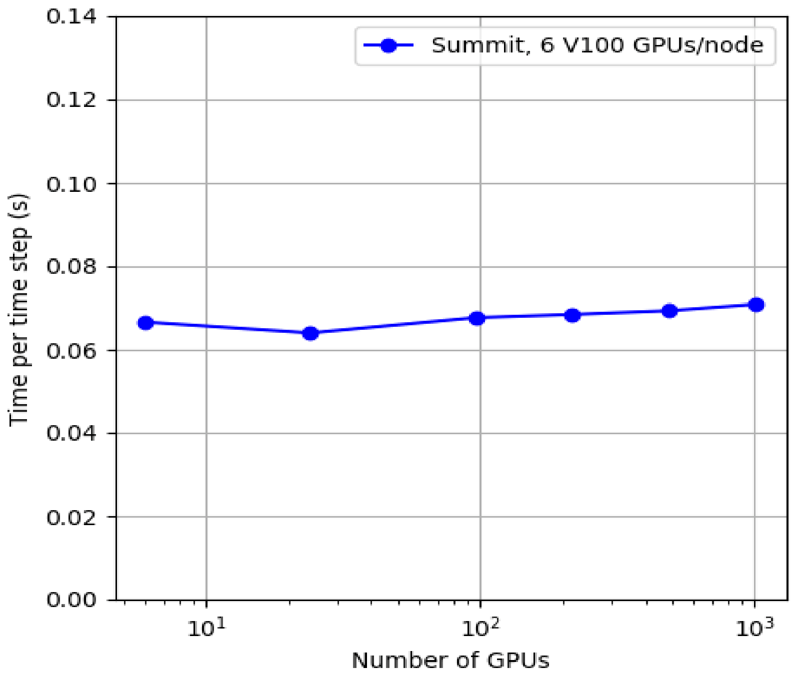
$a^{kj},\ k \neq j$ Off-diagonal terms, not so easy, need interpolation

Estimation of velocity and stress components using numerical differencing operator

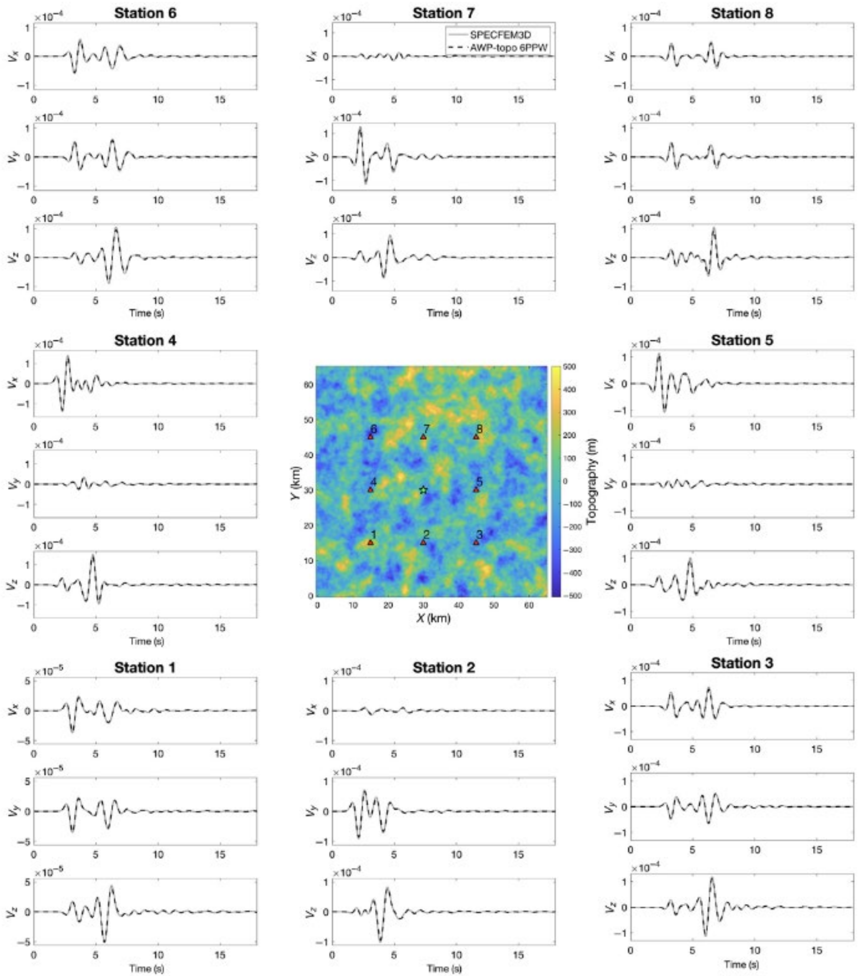


Weak Scaling and Verification

Weak scaling result with Nvidia V100 GPUs

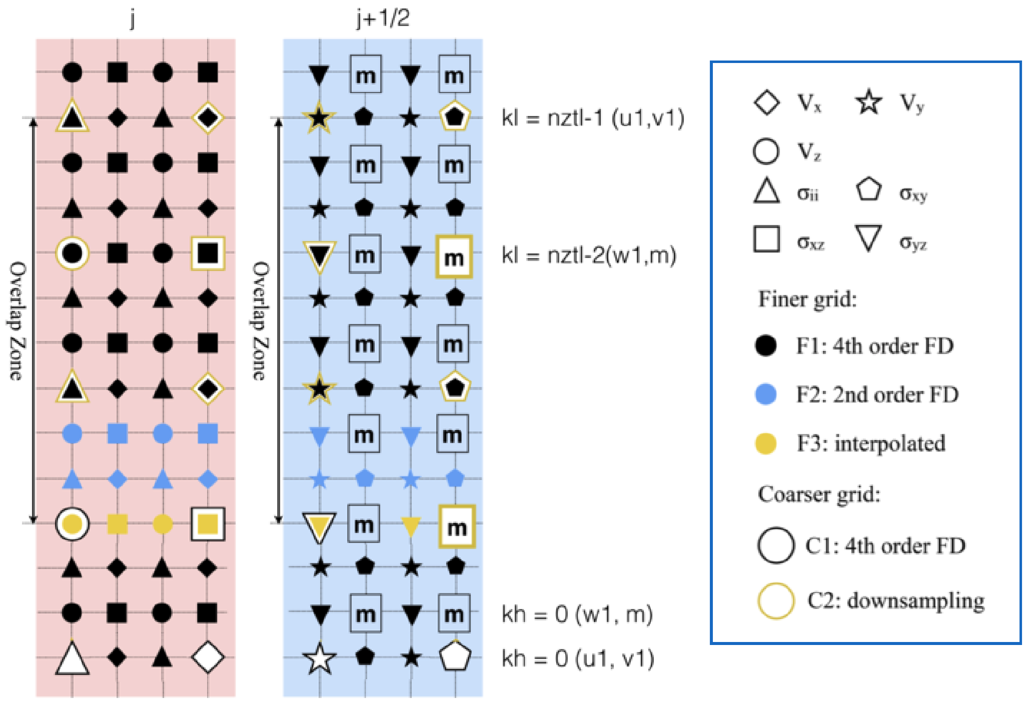


Verification with SPECFEM3D



Implementation of Discontinuous Mesh (DM)

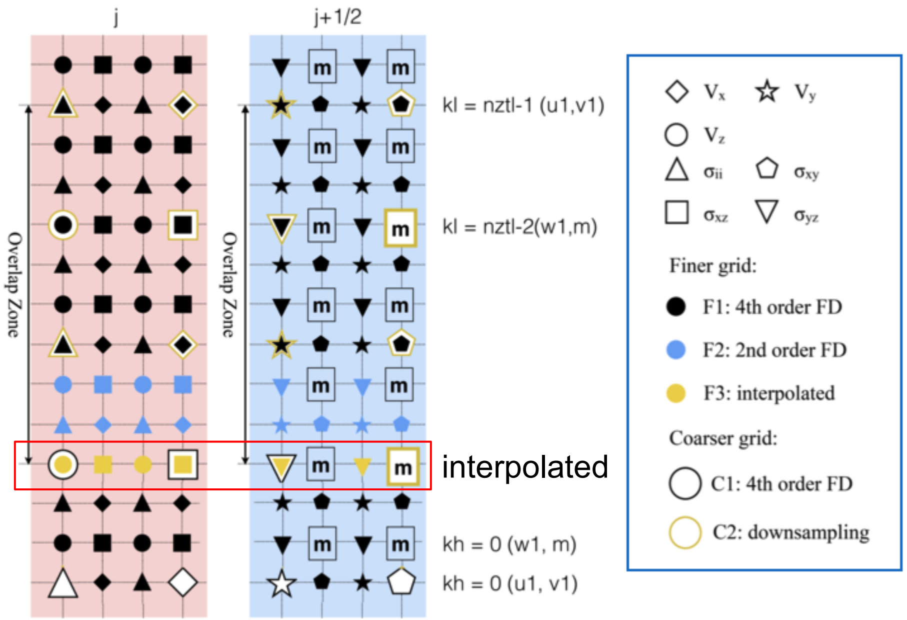
- Constant grid spacing is too fine for greater depths
- Factor-of-three ratio coarsening (Nie et al., 2017)
- Wavefield Estimation Using a Discontinuous Mesh Interface (WEDMI)
- Overlap zone for data exchange



Implementation of Discontinuous Mesh (DM)-cont'ed

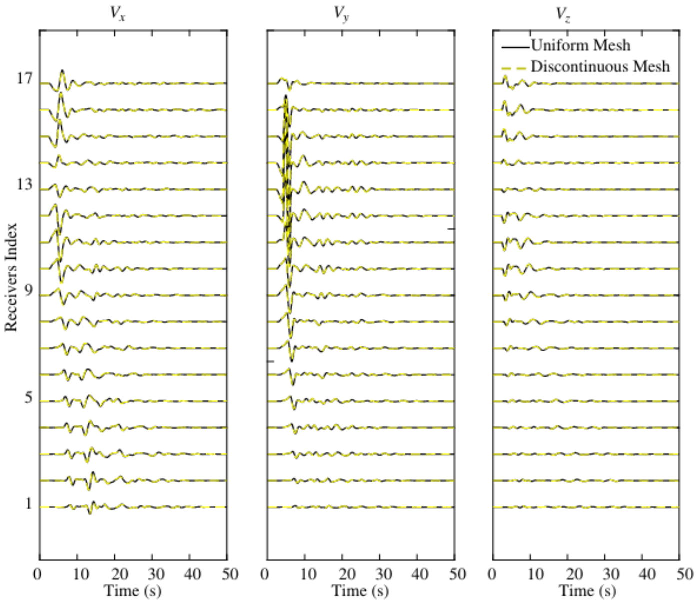
- Downsampling process needed

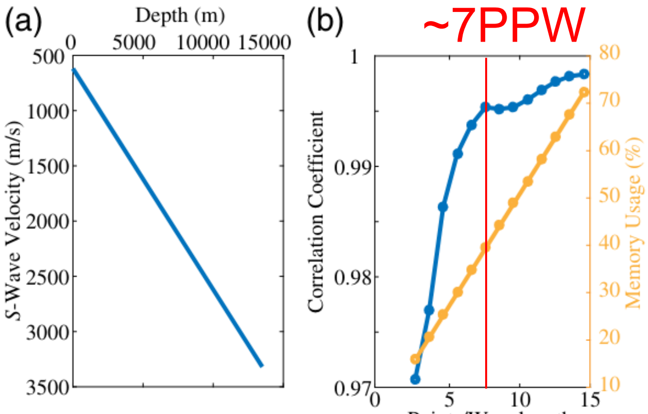
- Directly passing values to collocated coarser grids
  -> proven unstable
- Filter first and interpolate (Nie et al., 2017)



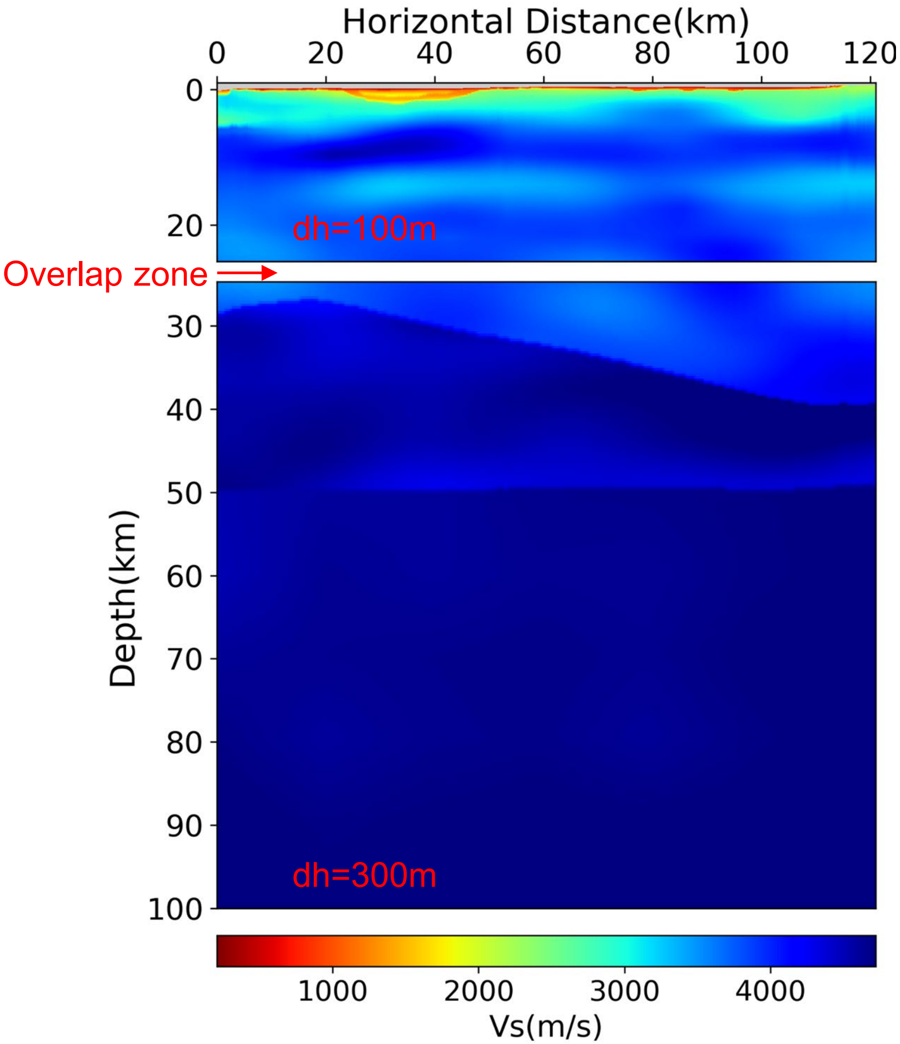
Velocity field update:
1. fourth-order velocities update in the coarser region (C1)
2. fourth-order velocities update in the finer region (F1)
3. second-order velocities update in the finer region (F2)
4. free surface calculation
5. interpolation of velocities in the finer region (F3)
6. downsampling of velocities in the coarser region (C2)

Stress field update:
1. fourth-order stresses update in the coarser region (C1)
2. fourth-order stresses update in the finer region (F1)
3. second-order stresses update in the finer region (F2)
4. interpolation of stresses in the finer region (F3)
5. downsampling of stresses in the coarser region (C2)
6. apply source





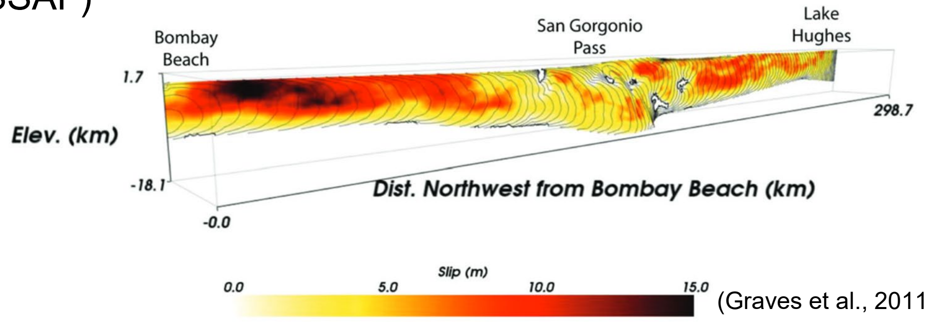
Nie et al. (2017)



The ShakeOut Scenario

- M7.8 scenario along southern San Andreas Fault (SSAF)
- Best possible science for fault geometry and source rupture characteristics in 2008

ShakeOut source model



(Graves et al., 2011)

USGS
science for a changing world

CGS
CALIFORNIA GEOLOGICAL SURVEY

The ShakeOut Scenario

By Lucile M. Jones, Richard Bernknopf, Dale Cox, James Goltz, Kenneth Hudnut, Dennis Mileti, Suzanne Perry, Daniel Ponti, Keith Porter, Michael Reichle, Hope Seligson, Kimberley Shoaf, Jerry Treiman, and Anne Wein

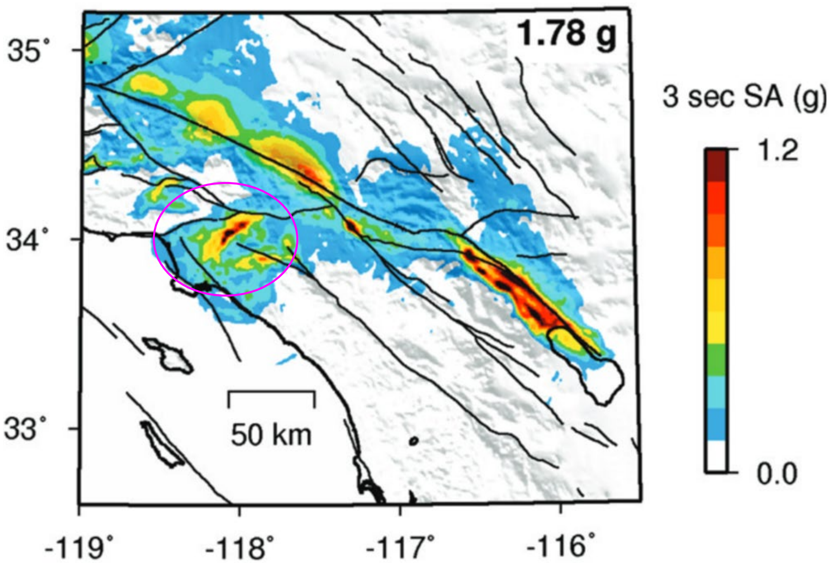
USGS Open File Report 2008-1150
CGS Preliminary Report 25
Version 1.0

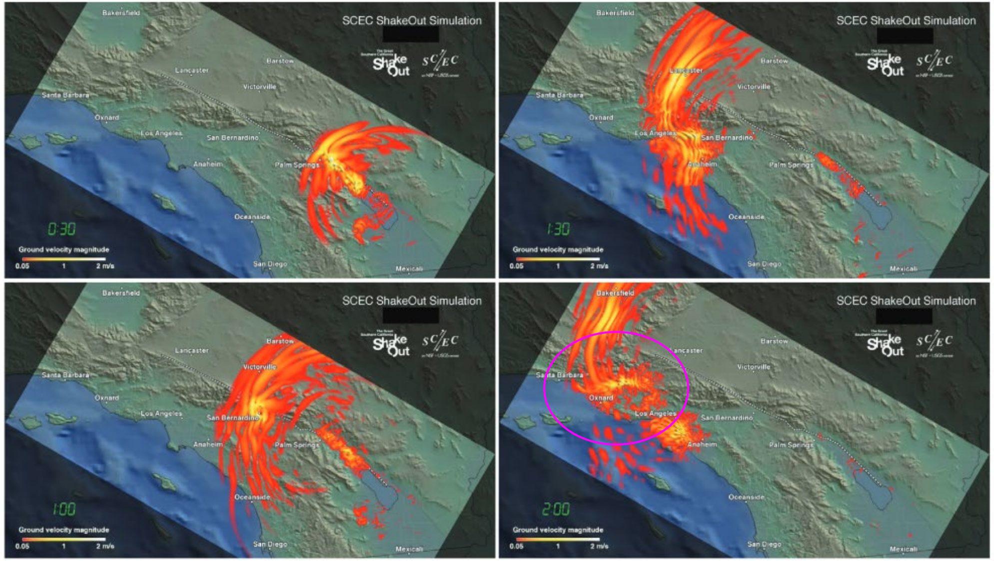
2008

Expected Strong Ground Motions

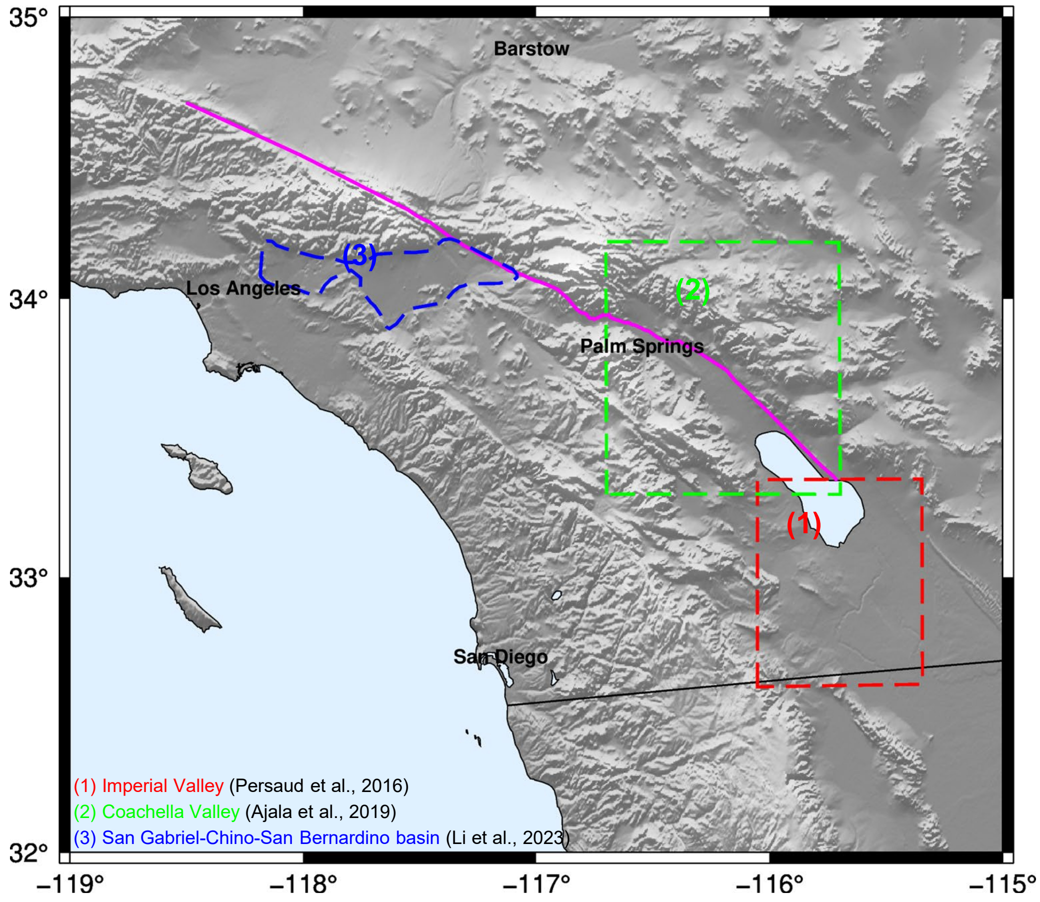
- Physics-based 3D wave propagation simulation: **AWP-ODC**
- Coherent long-period (3s-period and longer) waveguide channeling into Los Angeles basin
- Is best available science in 2008 still the best?

ShakeOut (Graves et al. 2011)

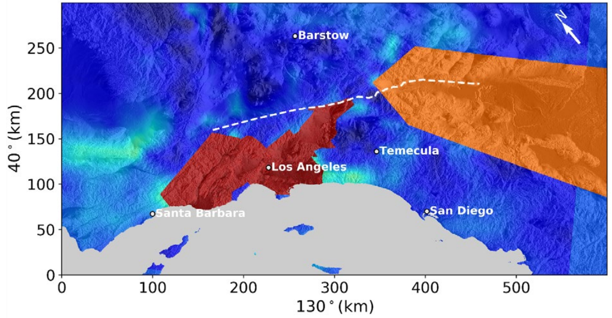




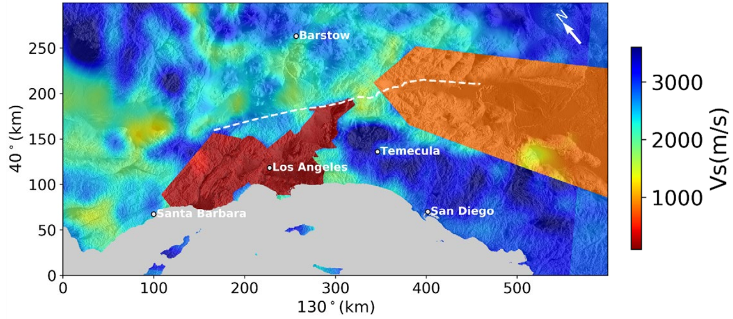
Improved/New Model Features



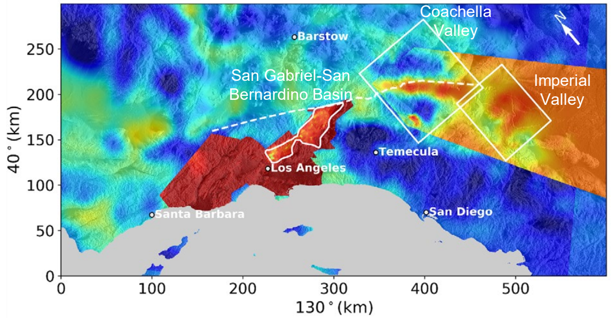
CVM-S4 (Used in the 2008 study)



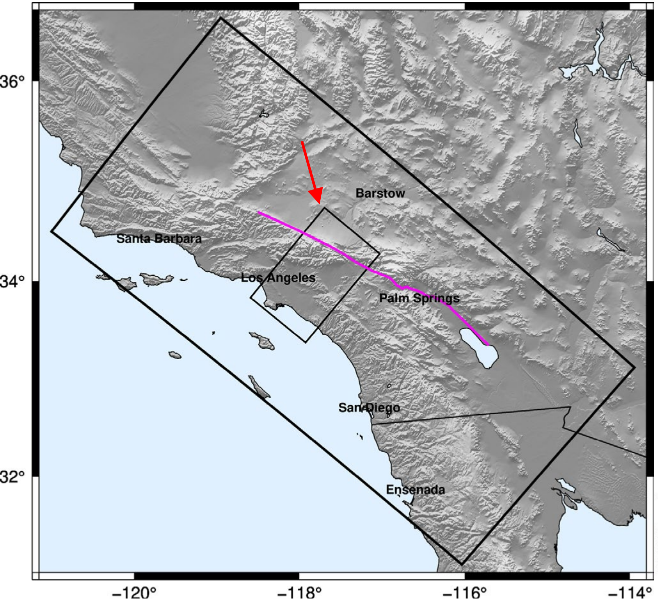
CVM-S4.26.M01 (Updates from Lee et al., 2014)

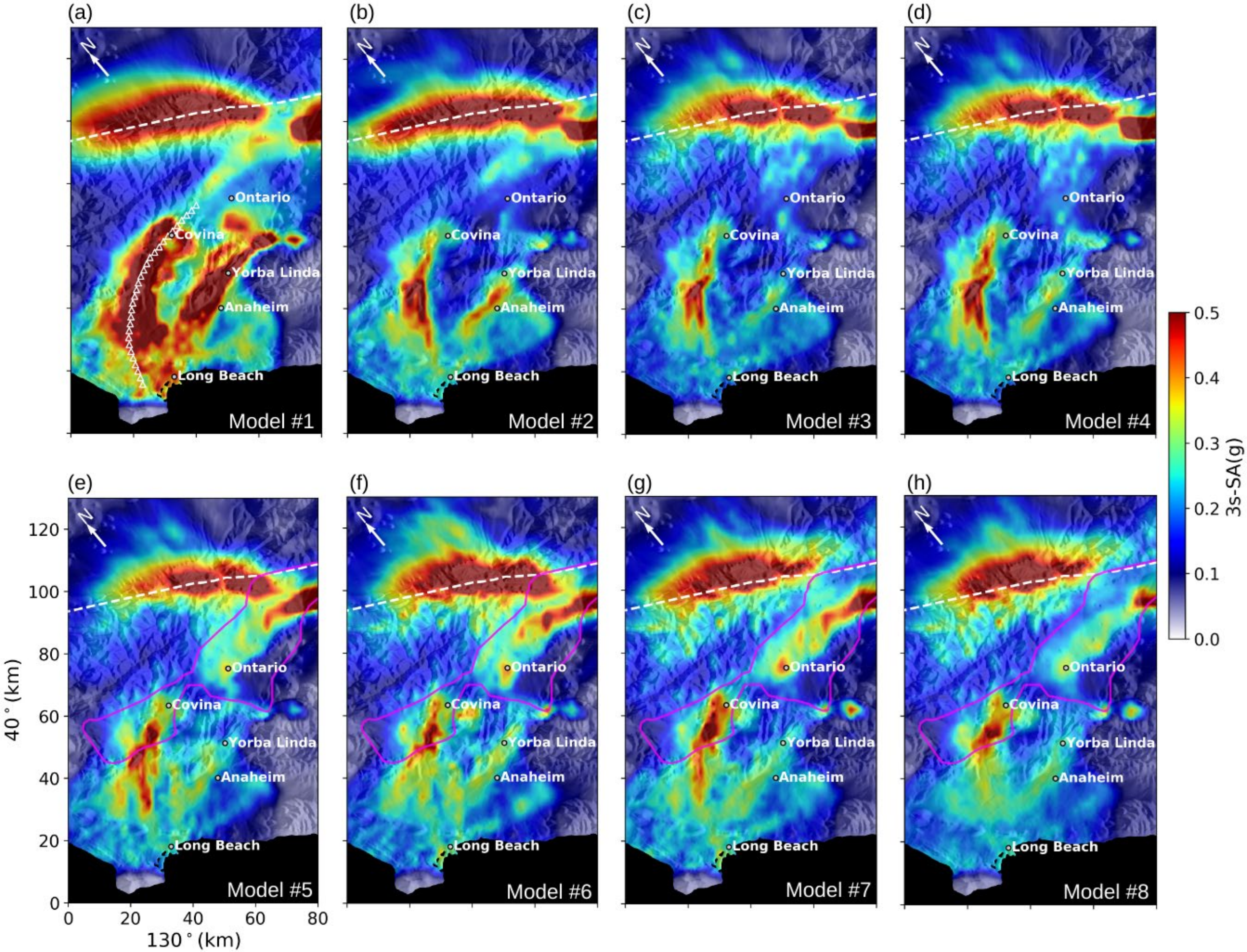


CVM-S4.26.M01+Local Models



Ground Motions Fade with Model Updates



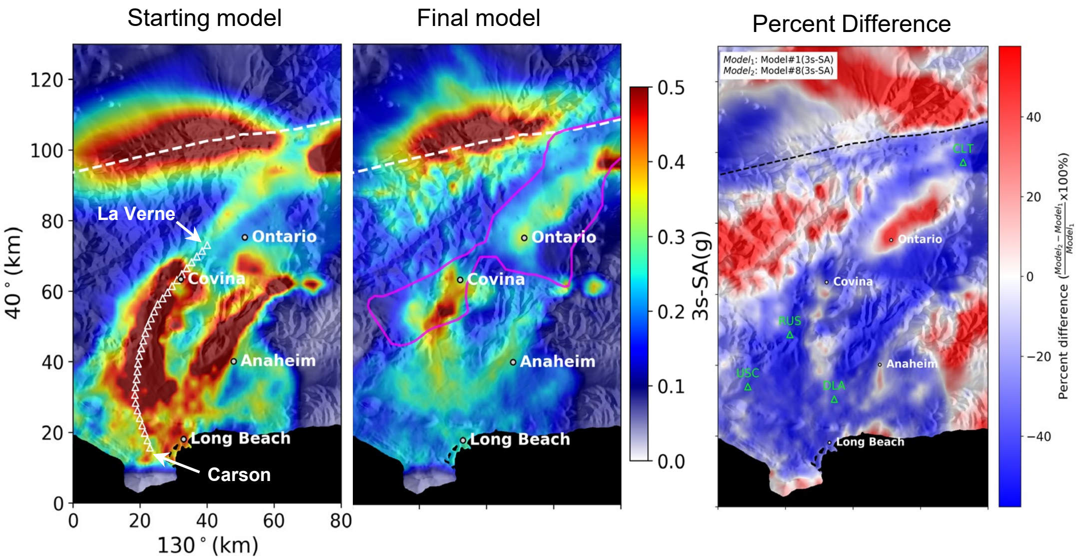


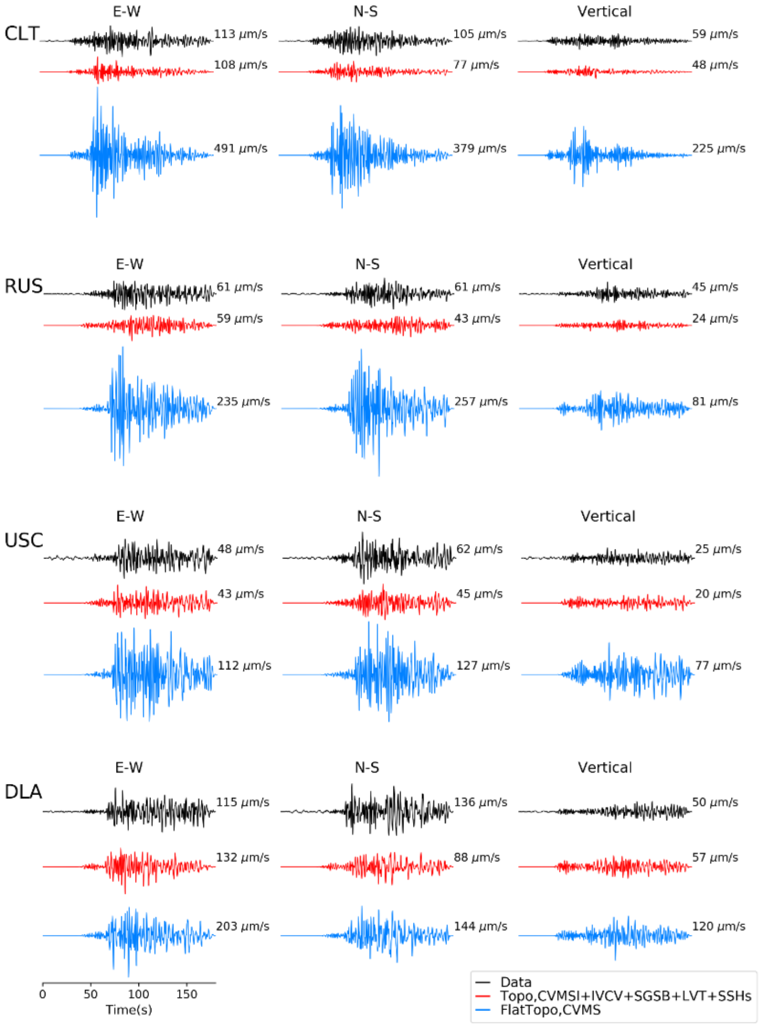
The Final Model...

- The exceptional scalability and performance of the AWP-ODC due to the most recent advancements allows for examination of plenty earth models including the actual surface topography at low computational cost

- Combining all model features, the predicted ground motions along both waveguide branches are reduced by 50-70% relative to the starting model

- The validation of the final model confirms the robustness of the final model up to 1 Hz.





Starting model　　　　Final model