

The EQSIM Earthquake Simulation Framework and Plan for Open Access Data Sharing

2025 PEER-LBNL Workshop on the Regional Scale Simulated Ground Motion Database (SGMD) for the San Francisco Bay Area

D. McCallen
Energy Geosciences Division
Lawrence Berkeley National Lab



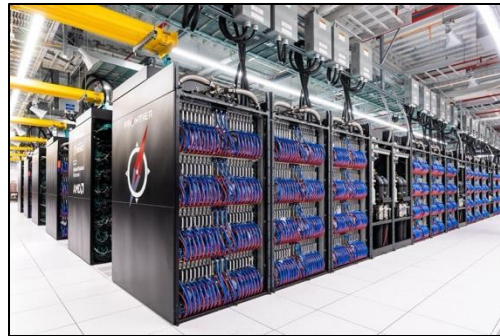
U.S. DEPARTMENT OF
ENERGY

Office of
 Cybersecurity, Energy Security,
 and Emergency Response

Two major Department of Energy projects have supported computational advancements

Computational R&D (2017-2023)

Provided the simulation tools



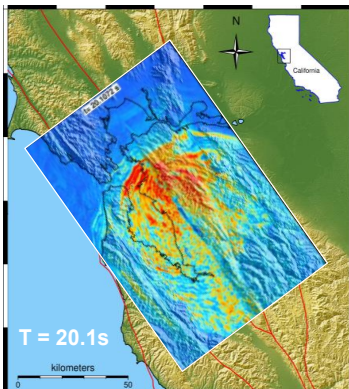
DOE developed the first *exascale* computational ecosystem based on GPU accelerator technologies

Energy System Applications (2023-2025)

Providing the mission "pull"



Office of Cybersecurity, Energy Security, and Emergency Response



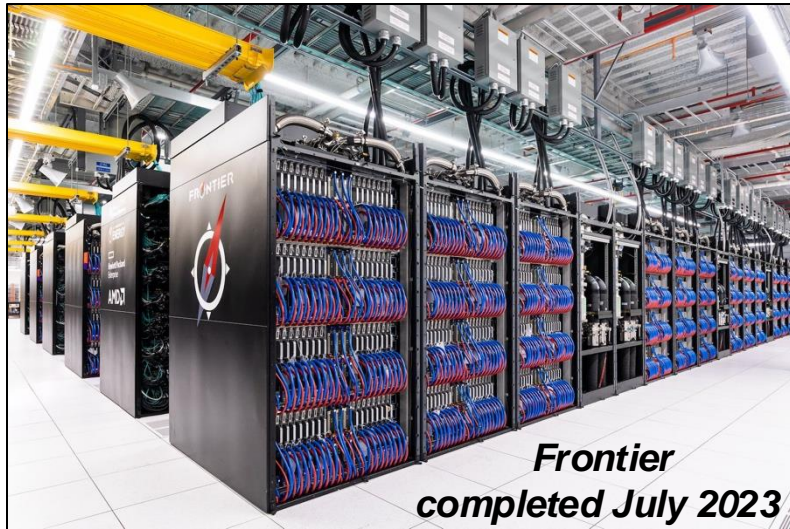
Name	Longitude	Latitude	Elevation (m)	X (m)	Y (m)	Vx0 (m/sec)	Vx1 (m/sec)	Vx2 (m/sec)	Vx3 (m/sec)	Vx4 (m/sec)	Vx5 (m/sec)	Realization 1					
												R1 (km)	R2 (km)	R3 (km)	R4 (km)	PGA (g)	PGV (m/sec)
S_01_01	-122.1816	38.5322	257.7	2000	2000	400.0	0.34	3.51	605.0	414.2	416.7	44.26	23.82	46.20	0.07	0.10	0.08
S_01_02	-122.2022	38.3417	430.8	2000	4000	780.9	0.23	0.86	780.9	784.4	785.4	44.70	33.80	44.70	0.04	0.09	0.09
S_01_03	-122.2207	38.3312	342.2	2000	6000	780.9	0.18	0.87	780.9	784.4	785.4	43.19	31.80	43.19	0.03	0.10	0.09
S_01_04	-122.2392	38.3206	257.7	2000	8000	780.9	0.47	0.87	780.9	784.4	785.3	43.72	29.79	41.72	0.07	0.14	0.10
S_01_05	-122.2578	38.3101	52.2	2000	10000	780.9	0.22	0.87	780.9	784.4	785.3	40.29	27.77	40.29	0.06	0.10	0.10
S_01_06	-122.2763	38.2995	6.2	2000	12000	780.9	0.29	0.87	780.9	784.4	785.3	38.91	25.76	38.92	0.06	0.13	0.13
S_01_07	-122.2948	38.2889	7.6	2000	14000	780.9	0.47	2.10	780.9	784.4	785.3	37.59	23.75	37.59	0.08	0.19	0.13
S_01_08	-122.3133	38.2784	780.9	2000	16000	780.9	0.13	0.87	780.9	784.4	785.3	36.34	21.74	36.34	0.05	0.23	0.23
S_01_09	-122.3318	38.2678	33.9	2000	18000	780.9	0.04	0.87	780.9	784.4	788.5	35.15	19.74	35.15	0.06	0.09	0.13
S_01_10	-122.3503	38.2572	41.3	2000	20000	780.9	0.25	0.87	780.9	784.4	785.3	34.04	17.73	34.04	0.07	0.13	0.13
S_01_11	-122.3688	38.2466	82.1	2000	22000	780.9	0.18	0.87	780.9	784.4	785.3	33.01	15.72	33.01	0.07	0.17	0.14
S_01_12	-122.3873	38.2360	37.7	2000	24000	780.9	0.47	1.13	780.9	784.4	785.3	32.08	13.71	32.08	0.09	0.21	0.15
S_01_13	-122.4058	38.2255	2.0	2000	26000	780.9	0.47	1.13	780.9	784.4	785.3	31.26	11.70	31.26	0.06	0.20	0.16
S_01_14	-122.4243	38.2149	1.6	2000	28000	780.9	0.47	1.66	780.9	784.4	785.3	30.54	9.69	30.54	0.08	0.24	0.18
S_01_15	-122.4427	38.2043	18.4	2000	30000	780.9	0.47	1.57	780.9	784.4	785.3	29.93	7.69	29.93	0.09	0.19	0.17
S_01_16	-122.4612	38.1937	201.9	2000	32000	780.9	0.46	1.63	780.9	784.4	785.4	29.45	5.68	29.45	0.09	0.19	0.16
S_01_17	-122.4797	38.1831	45.3	2000	34000	780.9	0.09	0.87	780.9	784.4	785.3	29.10	3.68	29.10	0.08	0.12	0.12
S_01_18	-122.4981	38.1725	88.8	2000	36000	780.9	0.14	0.87	780.9	784.4	785.3	28.88	1.66	28.88	0.10	0.17	0.12
S_01_19	-122.5166	38.1619	2.3	2000	38000	264.3	0.27	0.87	273.2	249.5	249.5	28.81	-0.55	28.81	0.20	0.34	0.15
S_01_20	-122.5351	38.1513	0.0	2000	40000	432.2	0.16	0.87	270.0	249.5	249.5	29.07	-0.41	29.07	0.23	0.39	0.14
S_01_21	-122.5536	38.1406	0.1	2000	42000	501.7	0.14	0.87	501.7	504.8	505.7	29.06	-4.36	29.06	0.17	0.26	0.12
S_01_22	-122.5719	38.13	264.0	2000	44000	1015.2	0.00	0.87	1015.2	1019.1	1019.1	29.40	-6.73	29.40	0.11	0.17	0.10
S_01_23	-122.5903	38.1184	29.4	2000	46000	574.2	0.02	0.87	574.2	574.0	577.9	29.87	-8.38	29.87	0.11	0.16	0.10
S_01_24	-122.6087	38.1088	41.0	2000	48000	501.7	0.11	0.87	501.7	504.8	505.7	30.46	-10.81	30.46	0.18	0.22	0.12
S_01_25	-122.6271	38.0982	107.8	2000	50000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	31.17	-12.38	31.17	0.08	0.12	0.10
S_01_26	-122.6456	38.0886	141.0	2000	52000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	31.99	-14.40	31.99	0.08	0.12	0.10
S_01_27	-122.6641	38.0789	222.2	2000	54000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	32.91	-16.41	32.91	0.12	0.13	0.08
S_01_28	-122.6824	38.0683	303.7	2000	56000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	33.93	-18.41	33.93	0.08	0.15	0.09
S_01_29	-122.7008	38.0577	384.0	2000	58000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	35.03	-20.41	35.03	0.09	0.11	0.08
S_01_30	-122.7192	38.0471	237.6	2000	60000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	36.20	-22.42	36.20	0.09	0.12	0.08
S_01_31	-122.7375	38.0364	139.7	2000	62000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	37.44	-24.43	37.44	0.08	0.12	0.08
S_01_32	-122.7559	38.0257	175.4	2000	64000	1015.2	0.00	0.87	1015.2	1028.1	1031.7	38.75	-26.43	38.75	0.06	0.10	0.08
S_01_33	-122.7743	38.0151	139.1	2000	66000	501.7	0.20	0.71	501.7	504.8	505.7	40.12	-28.43	40.12	0.11	0.12	0.08
S_01_34	-122.7927	38.0044	135.2	2000	68000	501.7	0.41	0.71	501.7	504.8	505.7	41.54	-30.44	41.54	0.07	0.13	0.09
S_01_35	-122.8111	37.9938	101.1	2000	70000	501.7	0.41	0.71	501.7	504.8	505.7	43.03	-32.44	43.03	0.06	0.10	0.08
S_01_36	-122.8294	37.9831	0.0	2000	72000	501.7	0.63	0.79	501.7	507.3	507.3	44.51	-34.45	44.51	0.07	0.12	0.10
S_01_37	-122.8477	37.9725	0.0	2000	74000	501.7	0.64	0.99	505.8	505.8	505.8	46.05	-36.45	46.05	0.06	0.12	0.11
S_01_38	-122.8661	37.9619	0.0	2000	76000	501.7	0.64	1.23	505.8	505.8	505.8	47.63	-38.45	47.63	0.06	0.14	0.11
S_01_39	-122.8844	37.9511	0.0	2000	78000	501.5	0.65	1.49	501.5	501.6	501.6	49.23	-40.46	49.23	0.04	0.14	0.12
S_01_40	-122.9028	37.9403	301.7	4000	2000	400.0	0.34	2.90	605.0	614.2	614.2	41.05	33.96	41.05	0.06	0.12	0.08

Simulated high-fidelity earthquake ground motions across key regions for engineering risk applications

The DOE Exascale Computing Project created the opportunity to advance earthquake simulations

The DOE Exascale Computing Project (ECP) developments (2017-2023)

Advanced computer platforms – Frontier was the world's first exaflop (1×10^{18} Flops) GPU-accelerated supercomputer



Frontier completed July 2023

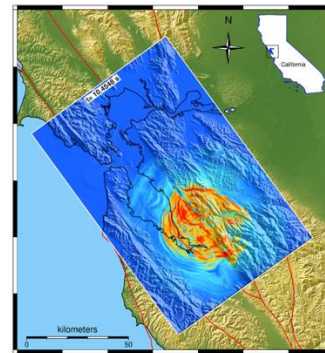
9,400 nodes, 1.194 exaflops



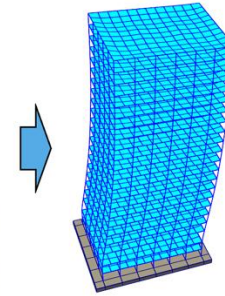
1,000,000,000,000,000 Flops

24 competitively selected software apps

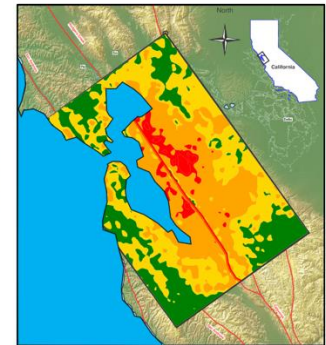
Advanced software – the *EQSIM* framework for regional-scale fault-to-structure simulations purpose built for GPU-accelerated platforms



Ground motion simulations



Infrastructure response simulations

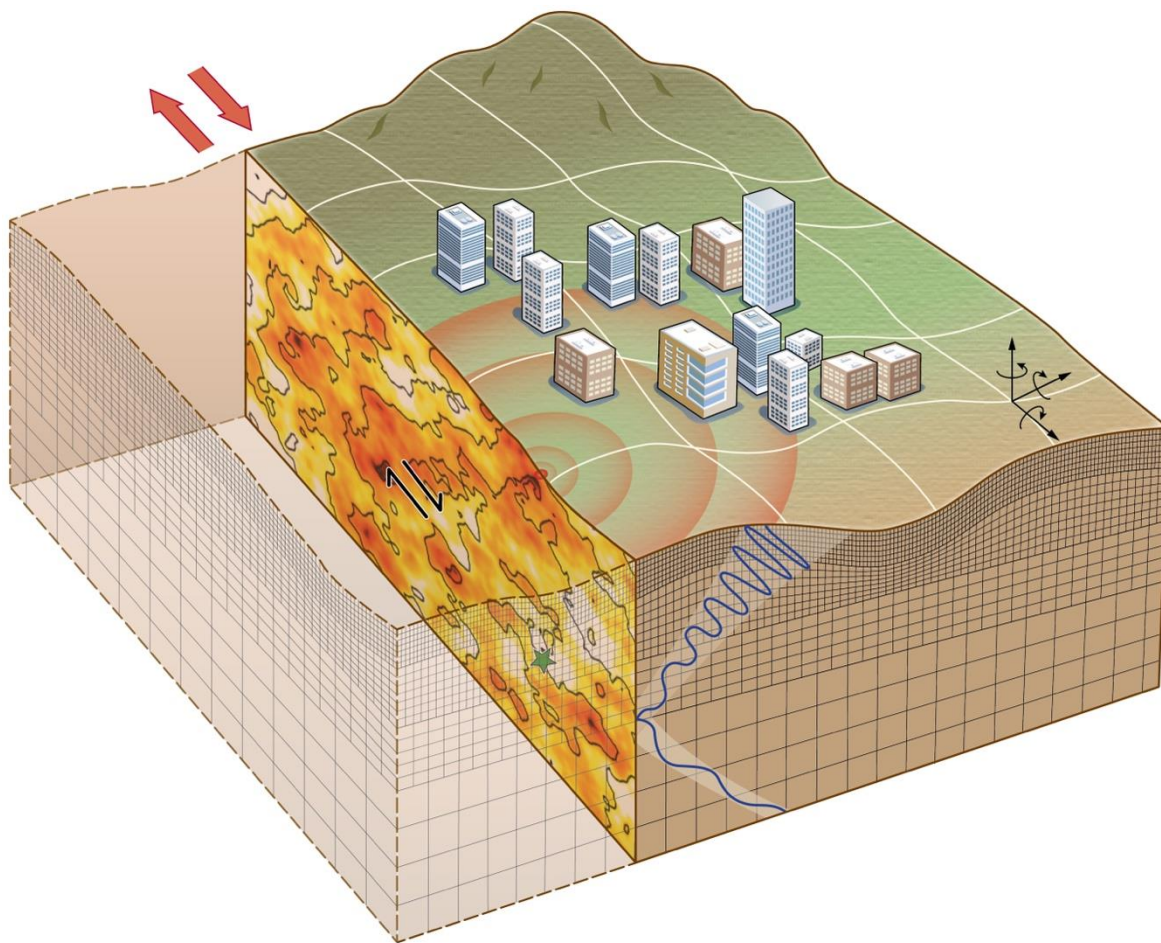


Deeper insight into regional infrastructure risk



ECP provided the support to complete many key advancements to the SW4 geophysics code

SW4 – Fourth order in space and time
(Pettersson, Sjogreen, Pankajakshan)



Advanced algorithms

- Mesh refinement in curvilinear and Cartesian grids

Readiness for GPU-based platforms

- Decompose computation loops into subtasks / tuning code for GPUs

Massive data management

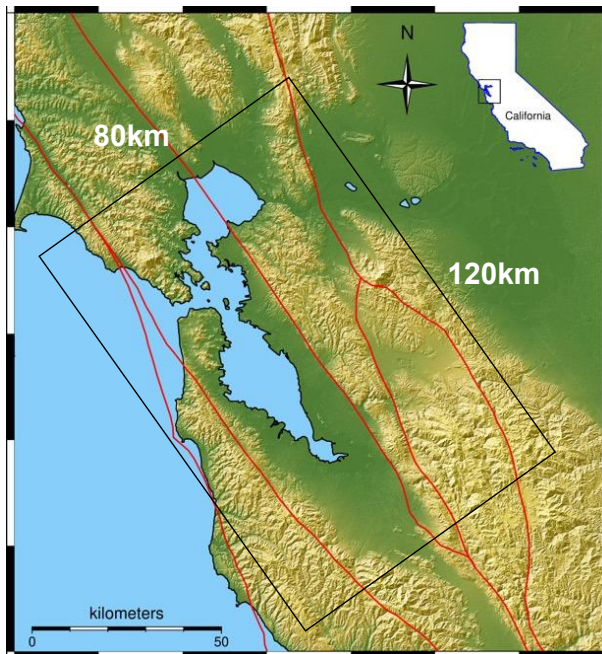
- Transition to HDF5-based IO (from SW4 homebrew)
- Utilization of ZFP for data compression

Enhanced physics models

- Enhancements to the Graves - Pitarka model

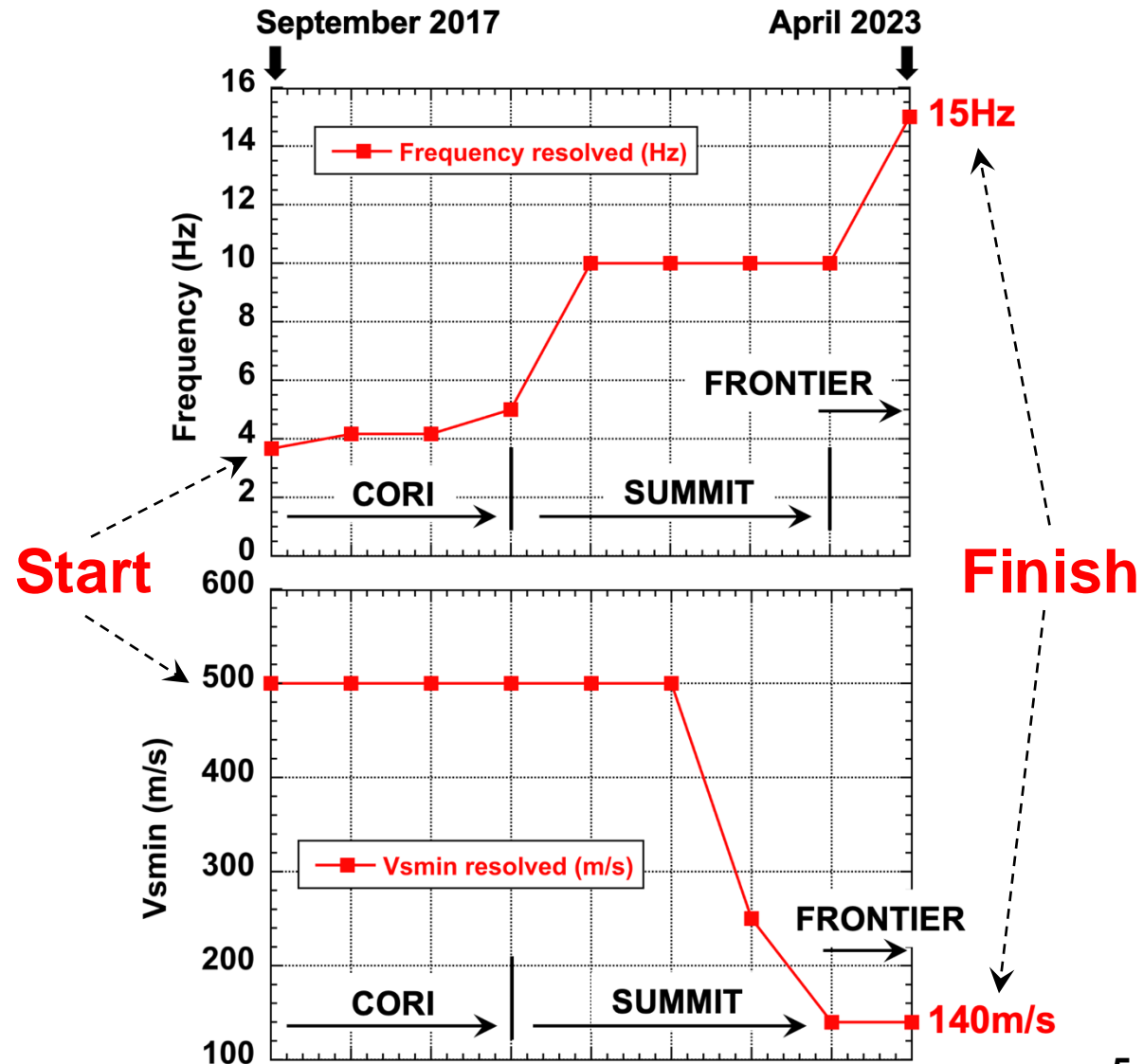
On GPU systems EQSIM has pushed the computational edge of simulation *resolution*

San Francisco Bay regional model

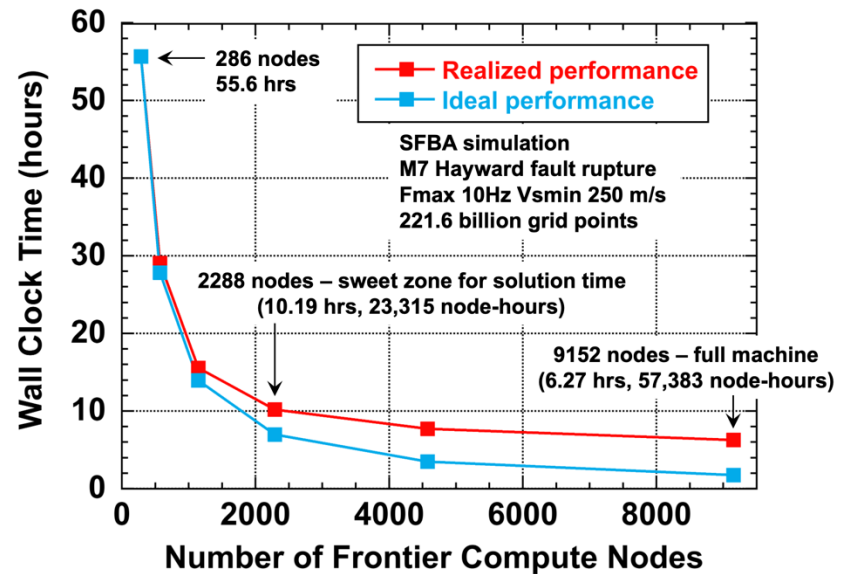
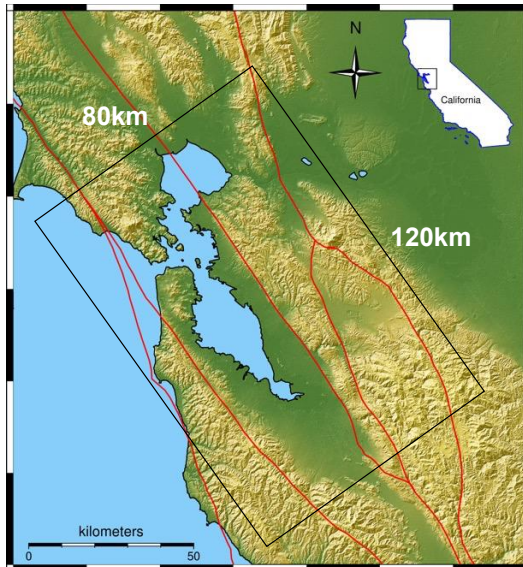


Computational effort...

- f (domain size)
- f (max frequency resolved)
- f (min Vs resolved)

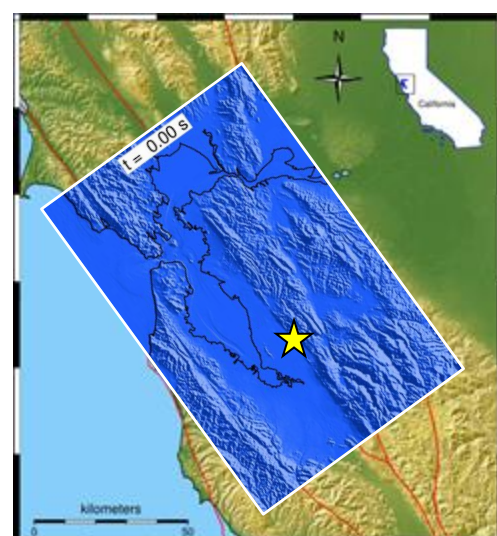
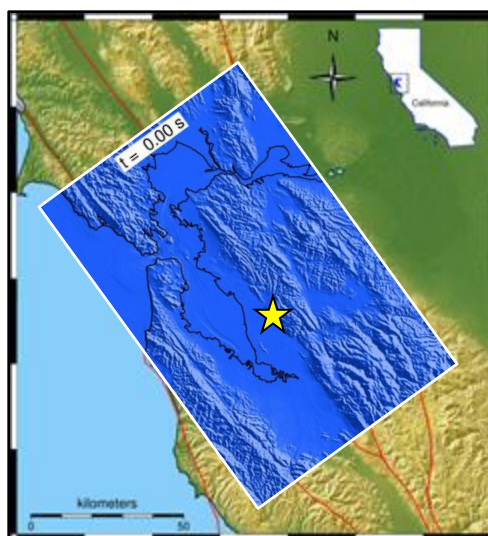
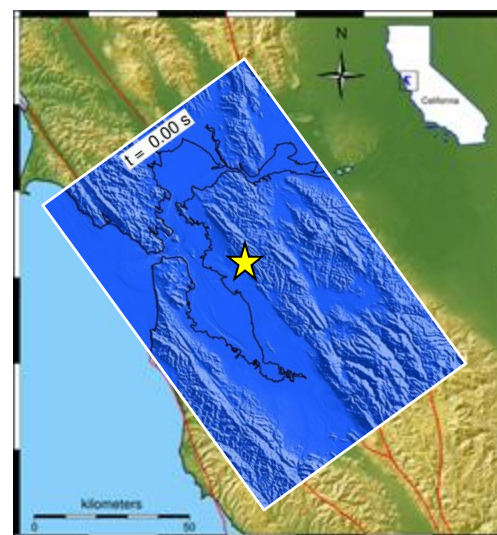
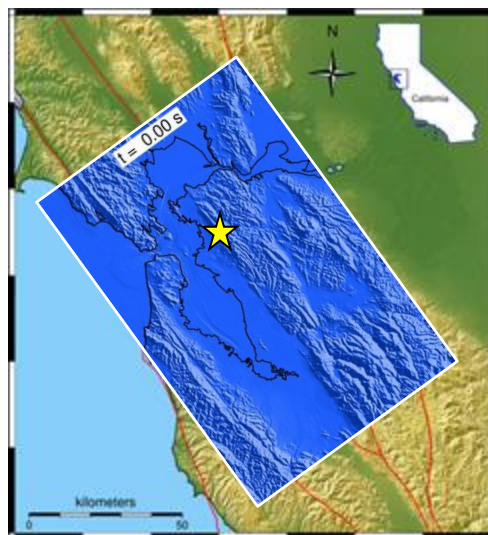
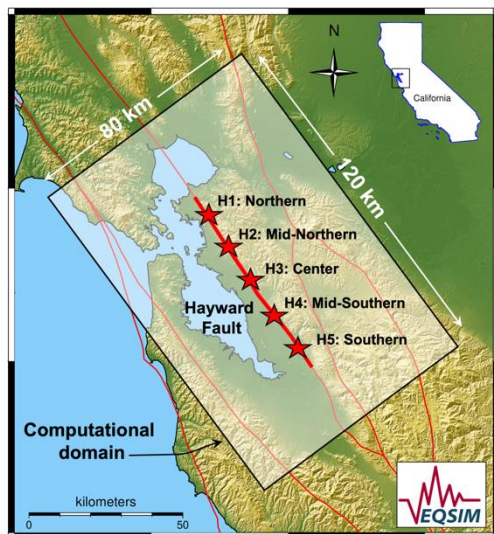


On GPU systems EQSIM has pushed the computational edge of simulation *speed*




Model	# Frontier nodes	Wall clock time (hrs)
Fmax 5 Hz Vsmin 250 m/s	3072	1.8
Fmax 10 Hz Vsmin 250 m/s	9152	6.3
Fmax 15 Hz Vsmin 140 m/s	9152	29.2

Take away – virtual exploration of the parameter space impacting ground motions is within grasp



With the transition toward applications, we have teamed with PEER to frame the path forward

PEER Pacific Rim Forum
June 2021



PACIFIC EARTHQUAKE ENGINEERING
RESEARCH CENTER

The PEER International Pacific Rim
Forum 2021: Regional-Scale Simulations
of Earthquake Ground Motions and
Infrastructure Response for
Performance-Based Earthquake Engineering

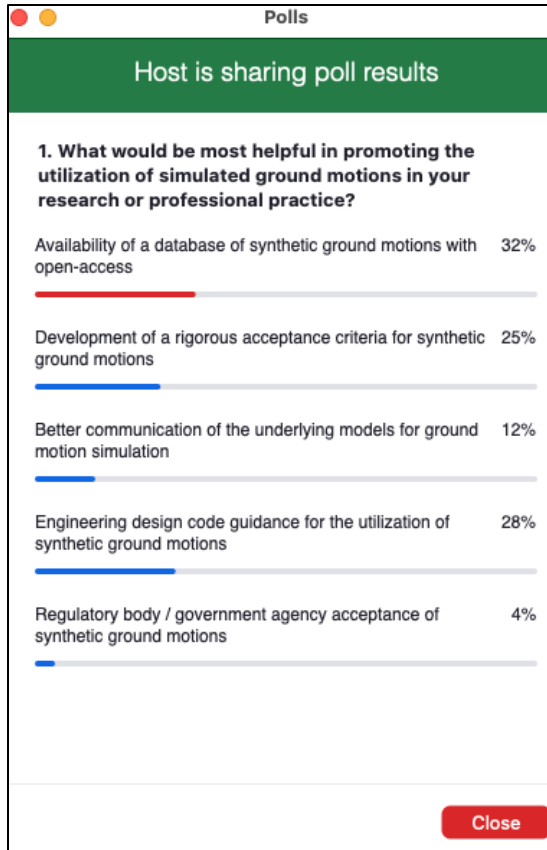
David McCallen
Floriana Petrone
Elnaz Esmailzadeh Seylabi
Arben Pitarka
Norman Abrahamson
Sherif Elfass

PEER Report No. 2022/04

Pacific Earthquake Engineering Research Center
Headquarters at the University of California, Berkeley
July 2022

PEER 2022/04
July 2022

261 International Participants
41 International Speakers



Attendees
recommended
priorities

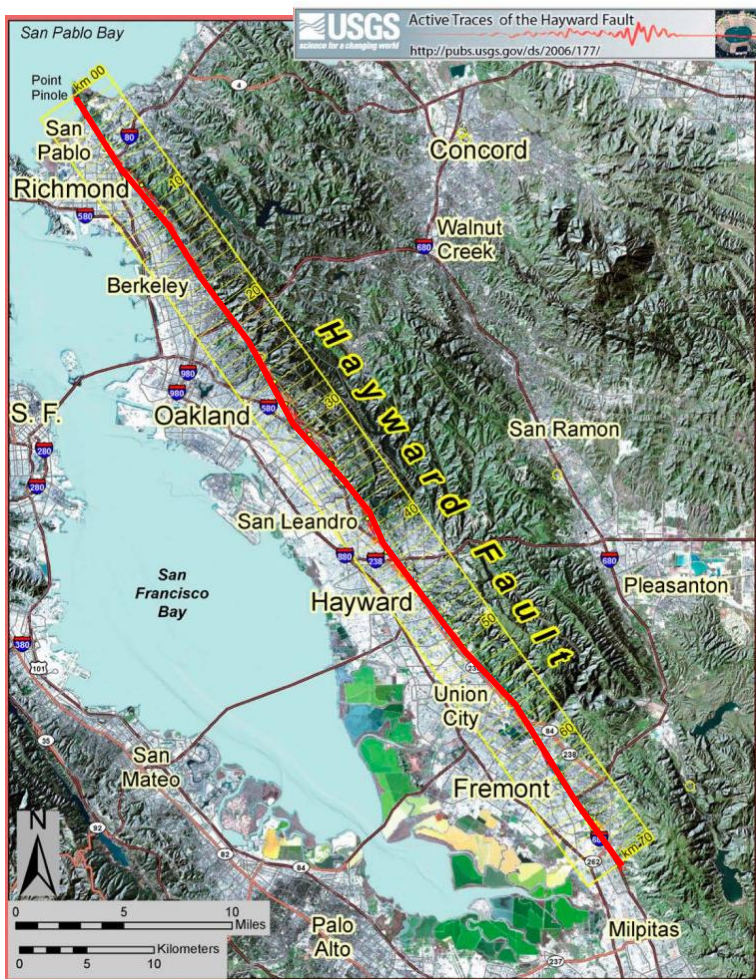
Availability of a database
of synthetic motions
with open-access

Assessment and
acceptance criteria for
synthetic motions

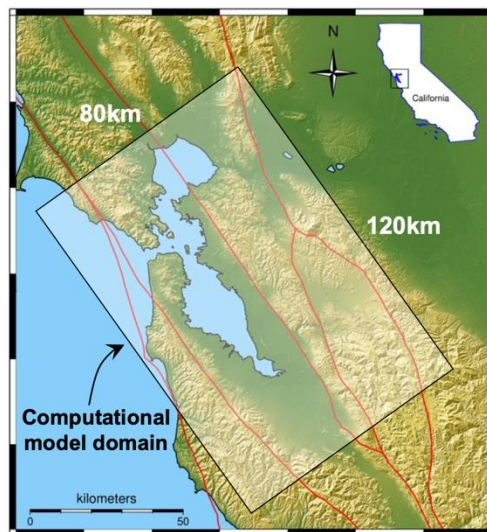
An effective and
operationally efficient
user interface

Our initial baseline simulated ground motion database has focused on the SFBA

The inaugural Simulated Ground Motion Database (SGMD)



San Francisco Bay Area simulation model



Model domain 120 x 80 x 30km

29 billion grid points

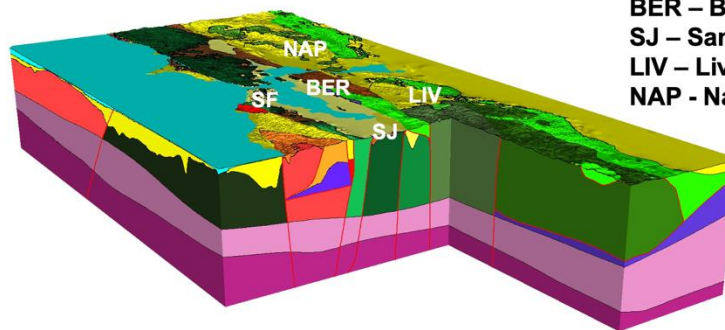
Fmax 5Hz

Vsmin 250m/s

Linear viscoelastic material model

Frequency independent attenuation factors Qp and Qs for P and S waves

USGS SFBR 3D geologic and seismic velocity model V21.1



SF – San Francisco
BER – Berkeley
SJ – San Jose
LIV – Livermore
NAP - Napa

Activity 1 - completion of many SFBA regional-scale simulated earthquake ground motions

Engineering

D. McCallen



Berkeley Lab

F. Petrone



Univ. Nevada

M. Miah



Berkeley Lab

Seismology

A. Pitarka



Livermore Lab

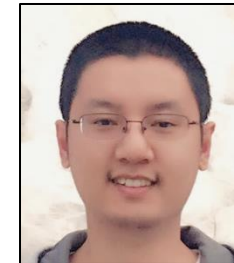
R. Nikata



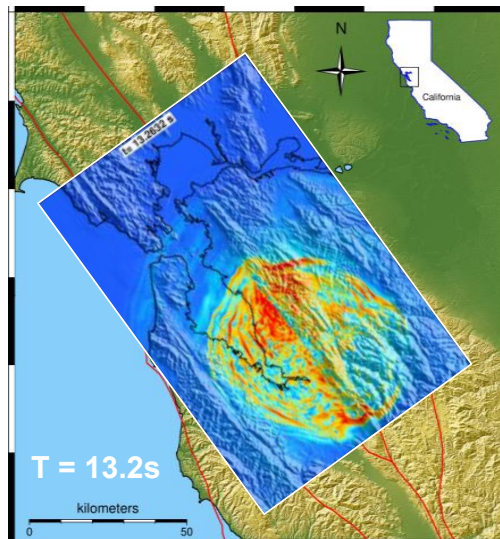
Berkeley Lab

Computer Science

H. Tang



Berkeley Lab



E. Taciroglu



UCLA Civil Engineering

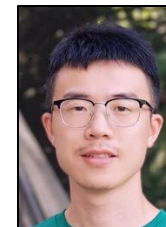
Current postdoctoral scholars and PhD students

Kostantinos Tsalouchidis



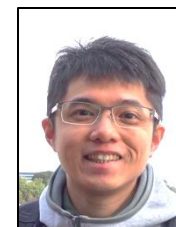
LBLN

Junfei Huang



LBLN

Clifford Yen



UCLA

Flora Xia



CALTECH



UCLA Samueli School of Engineering

Activity 2 - development of the interactive open access simulated ground motion database

K. Mosalam



PEER Director &
UC Berkeley

S. Gunay



G. Vargas



A. Kasalanati



C. Perez
PhD student



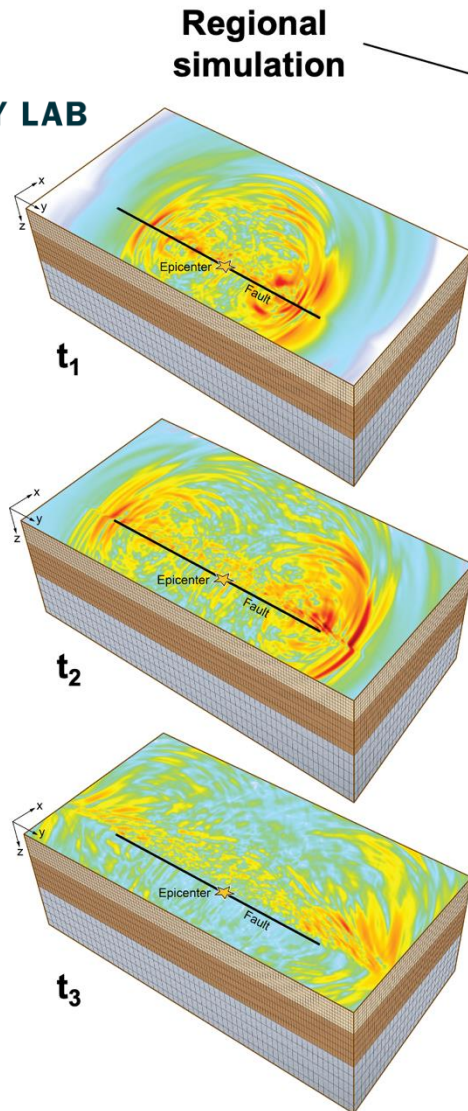
Existing – Measured ground
motion database

The screenshot shows the PEER Ground Motion Database website. The header includes the logo and text: "PEER Ground Motion Database Pacific Earthquake Engineering Research Center NGA-West2". Navigation links include HOME, DOCUMENTATION, HELP, SUBSCRIBE, and PEER. A user is signed in, with the email DMCCALLEN@UNR.EDU and a SIGN_OUT link. A "Signed in successfully." message is displayed. The "Target Spectrum" section has a "Select Spectrum Model" dropdown set to "No Scaling" and a "Select models to generate target spectrum" input field. There are also links for "Show/Hide GMM Notation", "Show/Hide GMM Regions", and "Show/Hide GMM Figures".

New – Simulated ground
motion database

The screenshot shows the PEER-LBNL Simulated Ground Motion Database website. The header includes the logo and text: "PEER-LBNL Simulated Ground Motion Database". Navigation links include Home, Signup, and Login. A 3D diagram of a simulated ground motion database is shown on the left. The "Log In" section has a "Username*" field with "mccallen" entered and a "Password*" field with "....." entered. There is a "Log In" button, a "Forgot Password?" link, and a "Need An Account? Sign Up Now Account Activation" link.

Workflow from simulation to SGMD server - developing a flexible data schema that can scale



Down sampling

Processing

Storage for open access
SGMD server



EQSIM
HDF5
Ground velocities
data options
- Compressed
- Uncompressed

Spatial
and
temporal
down
sampling
options

Numerical
differentiation
with 2nd order
central
difference

Ground
accelerations

Ground
velocities

Ground
velocities

- Fault normal
- Fault parallel
- Vertical

Ground
displacements

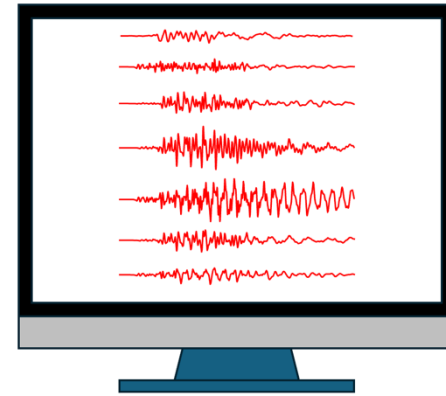
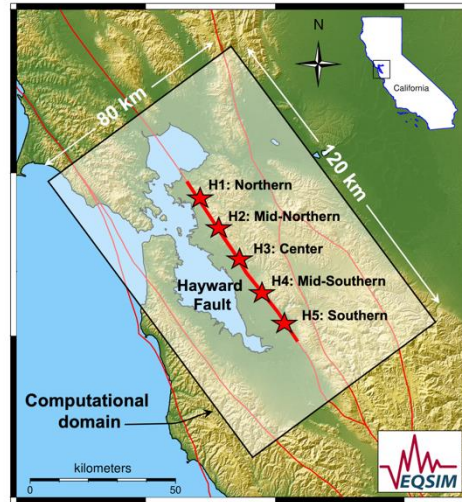
Numerical
integration
with trapezoidal
rule



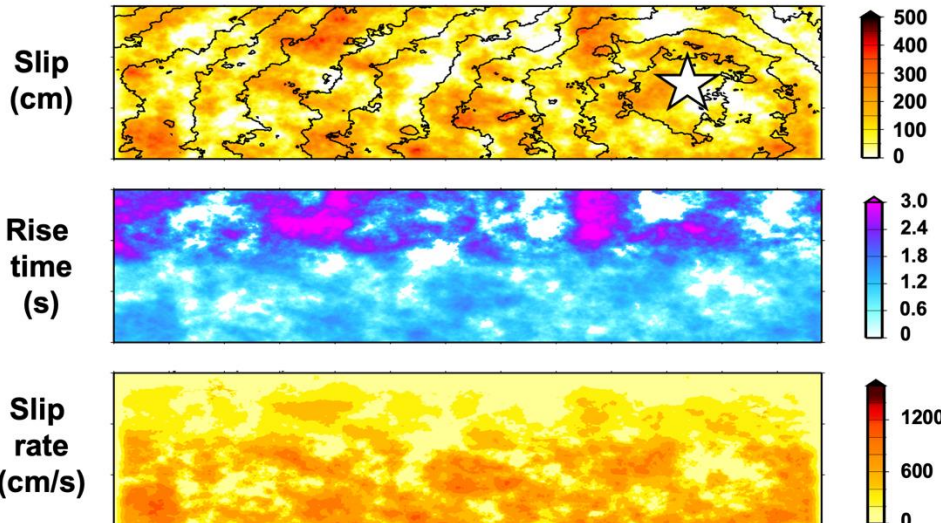
Linux server
64 core AMD
processor
128 GB RAM
68TB of SSD storage
RAID 6 storage
configuration

Focus on representing source rupture variability with the Graves - Pitarka rupture model

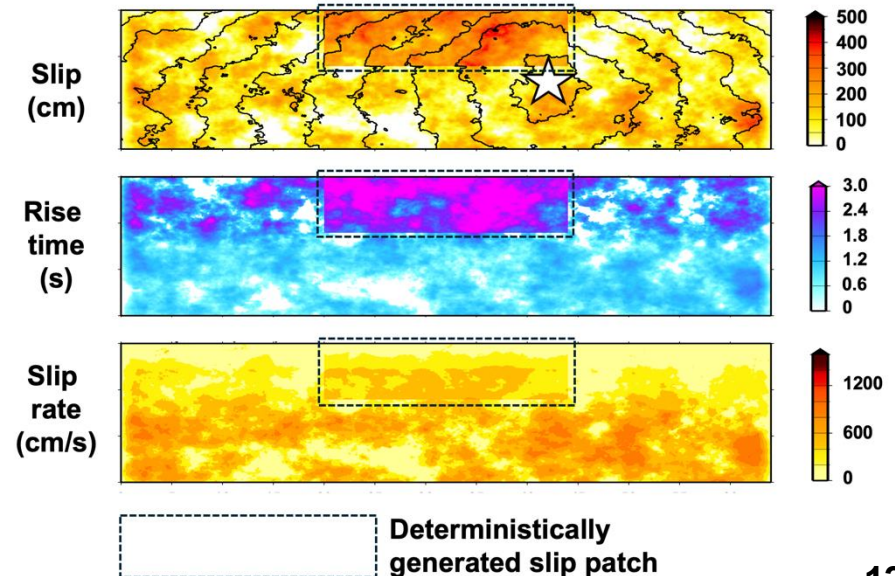
50 Hayward fault rupture realizations



Pure stochastic rupture models

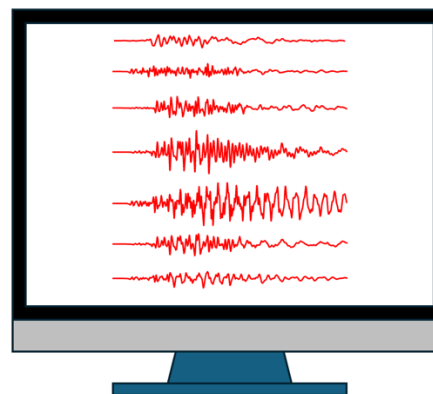
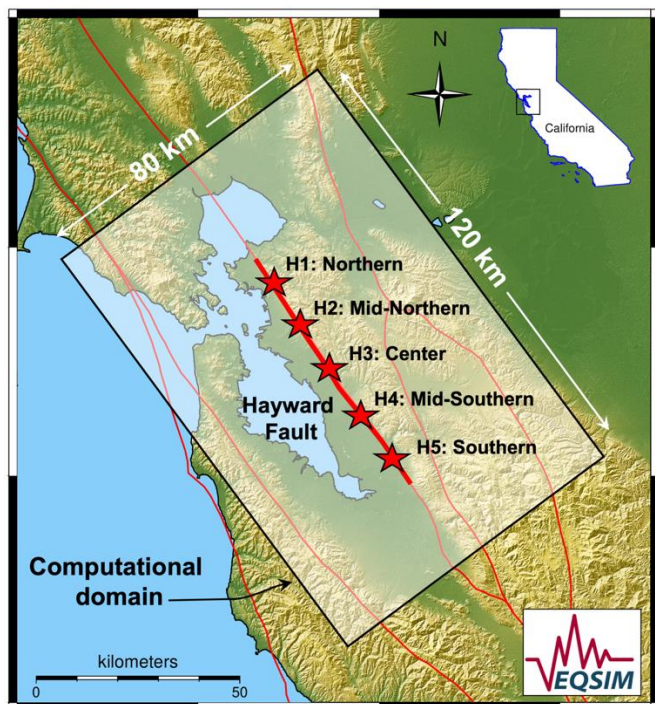


Stochastic plus deterministic patch rupture models

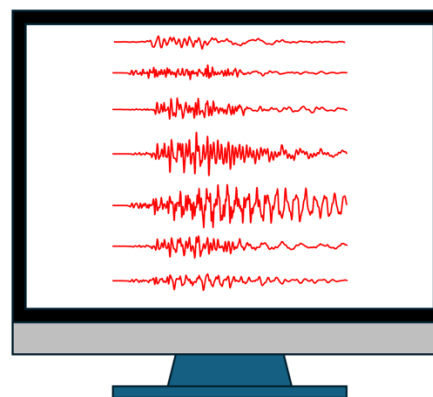


Simulated ground motions for 50 M7 Hayward ruptures are complete and on the SGMD server

50 Hayward fault rupture realizations



1,021,500 time series @ 2 km spacing
(uncompressed HDF5 files - 330GB
uncompressed ASCII files - **700GB**)

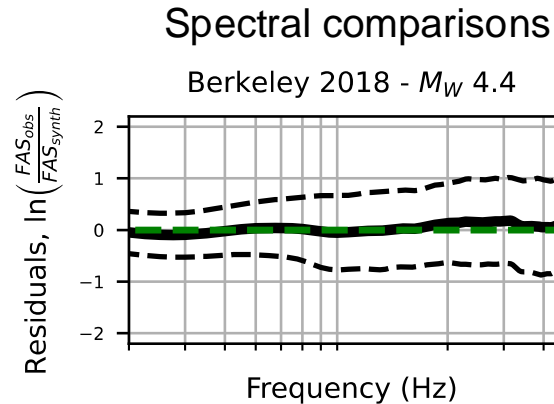
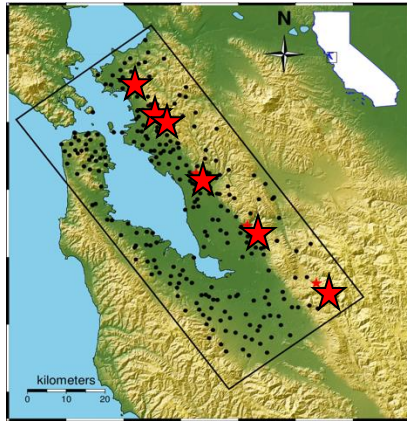


110,606,400,000 time series @ 6.25 m spacing
(compressed and time down sampled HDF5 files - **42TB**)

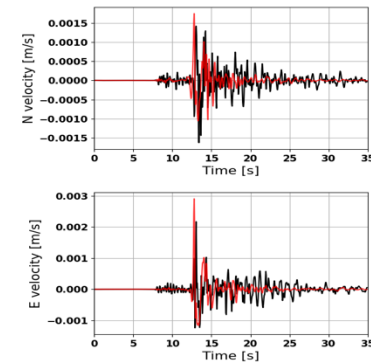
We have tested our simulations in multiple ways

1) Testing the EQSIM Bay Area model – 7 small Hayward event simulations

C. Pinilla Ramos



Comparison to measured ground motion waveforms

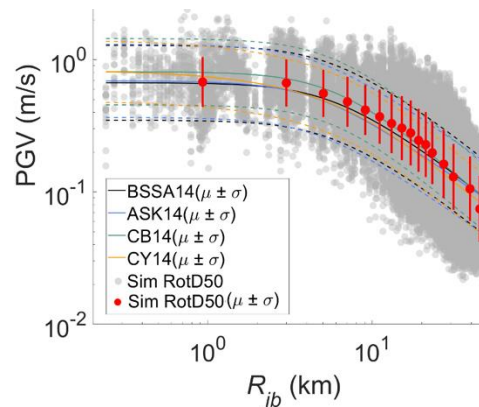


2) Evaluating the simulated large events - 50 M7 Hayward fault realizations

F. Petrone

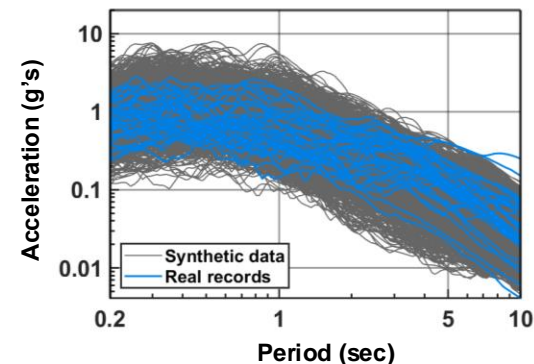


Comparison to existing empirical GMP equations



Comparison to existing commensurate ground motion data (near-fault sites < 10 km)

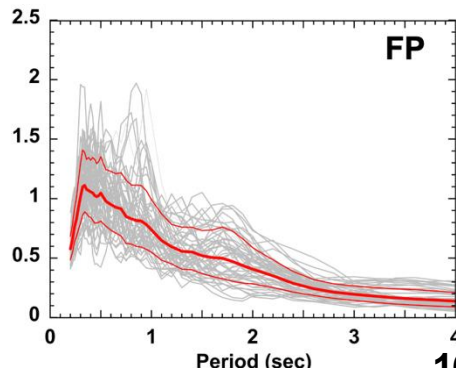
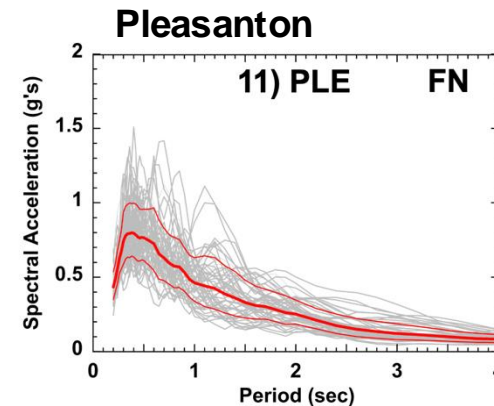
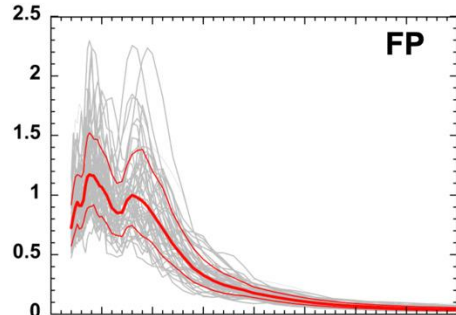
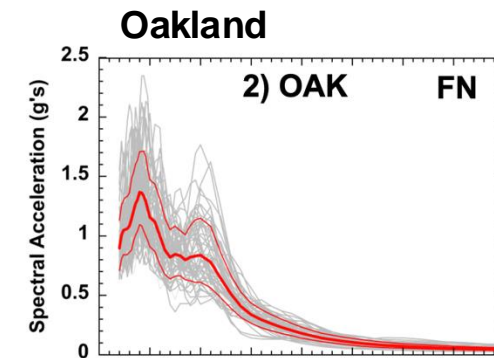
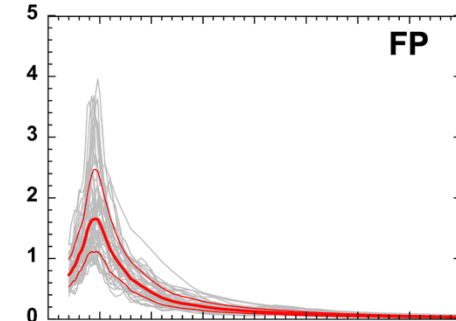
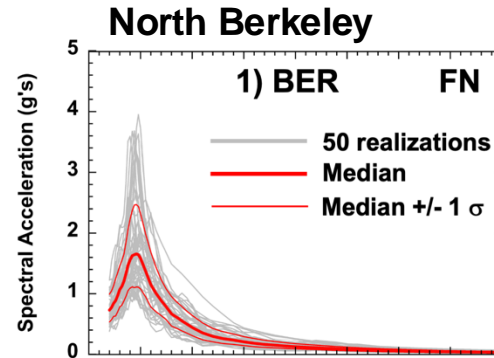
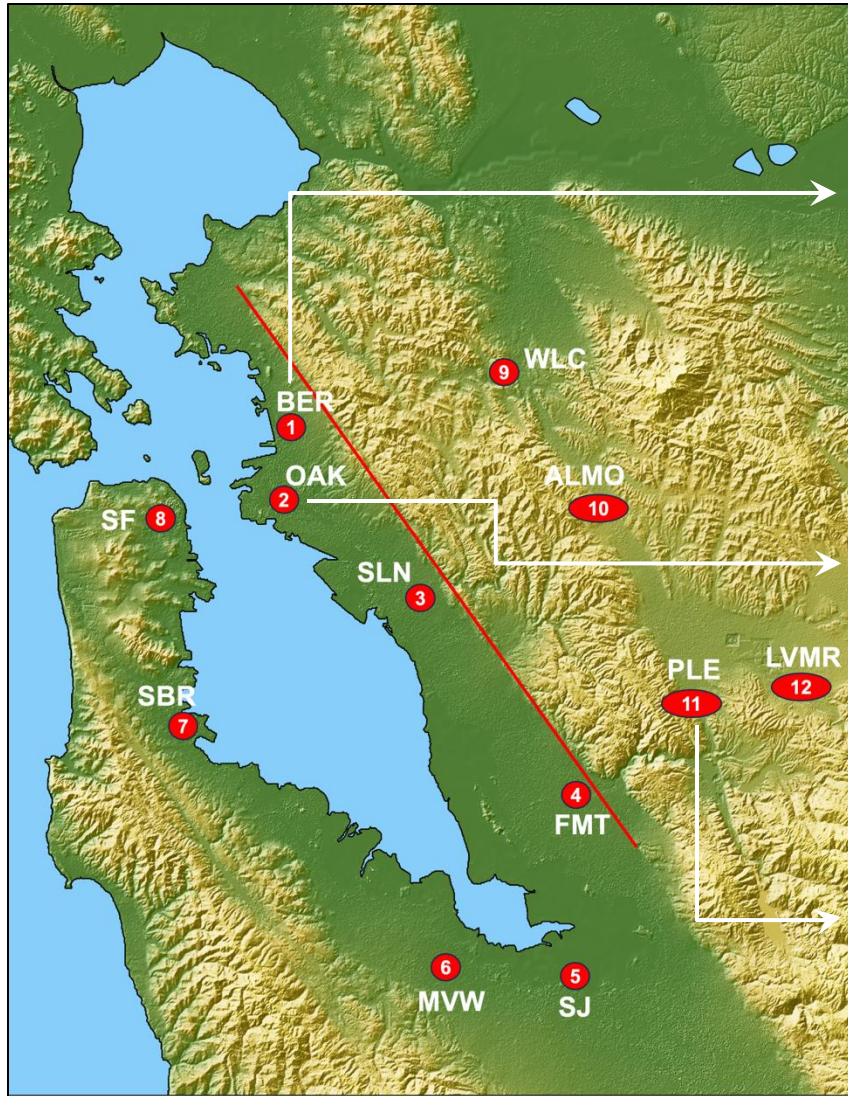
Spectral scatter - FaultNormal



R. Nikata



There is a *LOT* to analyze, e.g. complex within- and between-event distributions of motions



Since January 2025, 15 SGMD *Beta Users* have downloaded 5,633,712 time series

P. Arduino



UW/SimCenter

F. McKenna



UCB/SimCenter

S. Muin



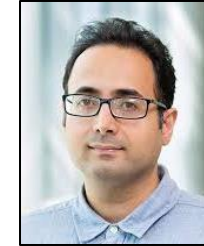
USC

T. Opabola



UC Berkeley

M. Esteghamati



Utah State

J. Dai



SoCal Edison

K. Hudnut



SoCal Edison

B. Low



PG&E

A. Kottke



PG&E

A. Shepard



FEMA

I. Kim



Degenkolb

M. Kenawy



OK State

B. Rowshandel



CA EQ Authority

H. Burton



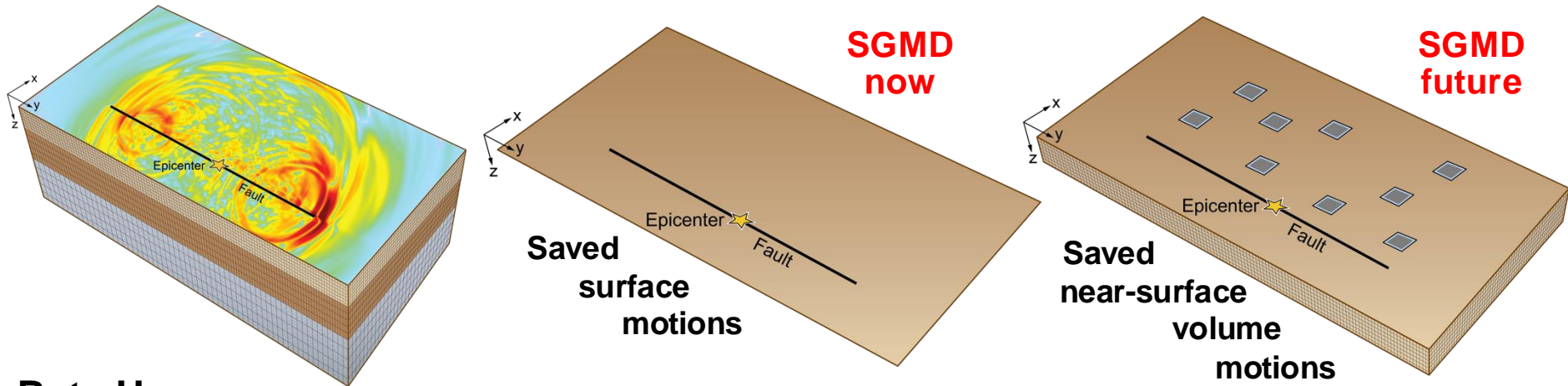
UCLA

E. Miranda



Stanford

The SGMD can continue to evolve to support advanced engineering risk applications

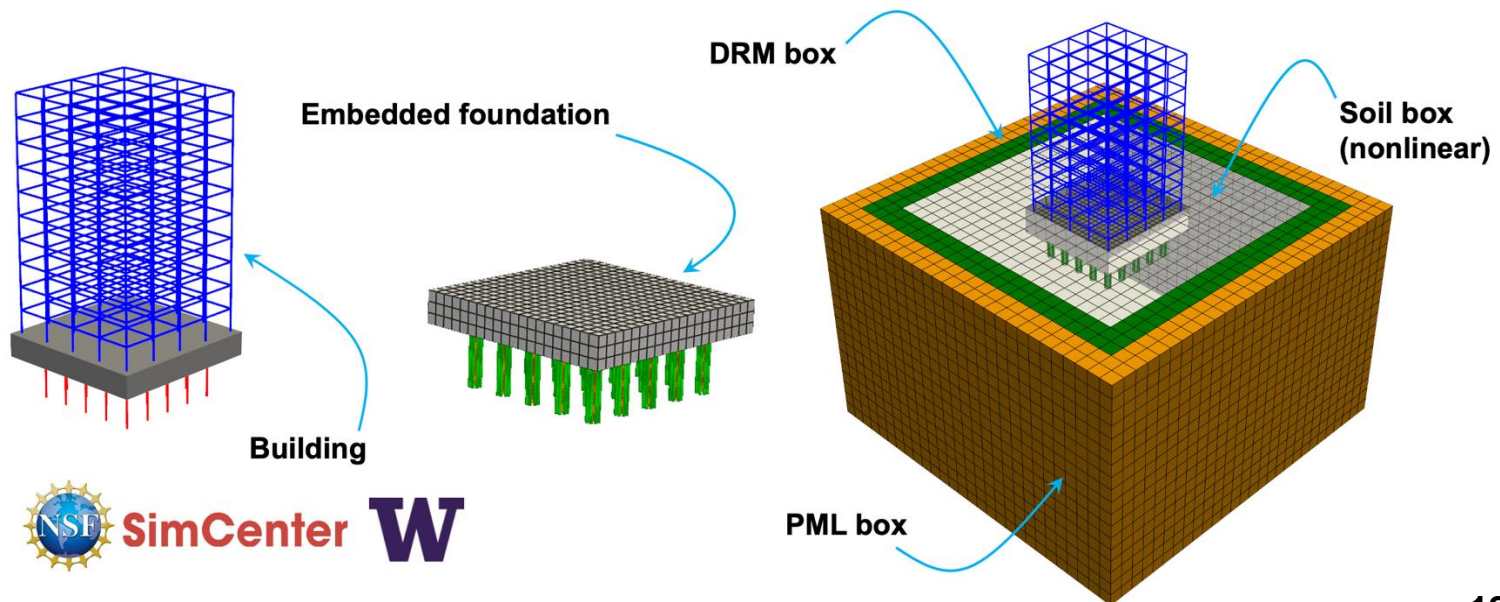


Beta Users

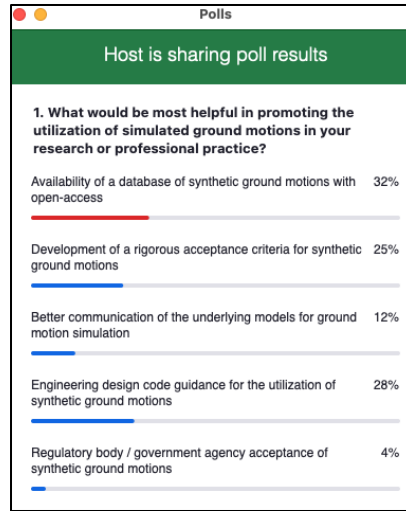
P. Arduino



F. McKenna



Community input for defining additional capabilities and enhancements is crucial



Links to more detailed info are on the SGMD website

Research Paper

Performance evaluation of the USGS velocity model for the San Francisco Bay Area

Camilo Pinilla-Ramos¹, Arben Pitarka², David McCallen, M. EERI³, and Rie Nakata⁴

Abstract
In this study, we evaluated the performance of the United States Geological Survey velocity model developed for the San Francisco Bay Area (SFBA), version 21.1. The evaluation was performed through high-resolution three-dimensional physics-based ground motion simulations of seven small-magnitude earthquakes (ranging from magnitude 1.8 to 4.4) that occurred on the eastern side of the San Francisco Bay. The simulations were performed in the frequency range from 0 to 5 Hz with a minimum shear-wave velocity of 250 m/s, which allowed the capture of wave propagation effects of the near-source soil materials that characterizes local basins. Based on the direct comparison of Fourier amplitude spectra between recorded and simulated ground motions for more than 250 stations, we found that the velocity model generally performs well in the frequency range of 0.2–5 Hz. The median value of the Fourier amplitude residuals was found to be near zero for all seven earthquakes. The slight over-prediction of 0.2 log-amplitude units at frequencies above 3 Hz in our simulations was attributed to the potentially inaccurate representation of the source radiation pattern by a double-couple point source model, and simple representation of shallow small-scale underground structural complexity in the velocity model. Maps of spectral amplitude differences between the simulated and recorded data were used to identify areas responsible for systematic ground motion over-predictions or under-predictions. For example, while some sub-domains over soft sediments show over-prediction patterns, the block east of the Hayward fault is prone to exhibit patterns of under-prediction. These maps can be used to guide future refinements of the SFBA velocity model. Since our simulation methodology allows for the decoupling of the source and wave propagation effects, the ground motion data generated by our simulations can also be used to quantify the epistemic uncertainty due to the velocity model, in empirically based ground motion estimates for the SFBA.

¹Lawrence Berkeley National Laboratory, Berkeley, CA, USA
²Lawrence Berkeley National Laboratory, Livermore, CA, USA
³University of Nevada, Reno, NV, USA
⁴University of Nevada, Reno, NV, USA

Corresponding author:
Camilo Pinilla-Ramos, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
Email: cpinilla@berkeley.lbl.gov

Research Paper

Ground-motions site and event specificity: Insights from assessing a suite of simulated ground motions in the San Francisco Bay Area

Floriانا Petrone, EERI¹, Arsam Taslimi, EERI¹, Majid Mohammadi Nia¹, David McCallen, EERI³, and Arben Pitarka²

Abstract
This article presents the results of a research that is part of a larger collaborative effort between the Lawrence Berkeley National Laboratory and the Pacific Earthquake Engineering Research Center, funded by the US Department of Energy Office of Cybersecurity, Energy Security and Emergency Response. The main objective of this study is to assess a suite of near and far-field simulated ground motions obtained from 20 realizations of an M7 Hayward Fault earthquake in the San Francisco Bay Area, California, USA, and inform the selection of rupture simulation parameters leading to strong motions. To this aim, comparisons are conducted with NGA-W2 and directivity ground-motion models and a selected population of records. An archetypal steel moment-resisting frame is utilized to assess infrastructure response distributions. The analyses carried out for each simulated event and subdomain with consistent properties in terms of shallow shear-wave velocity proved to be instrumental for better interpreting the differences between simulated motions and empirical models. The main reasons identified for variances between simulations and empirical relationships included (1) directivity effects fully captured by the simulations across the full breadth of rupture models; (2) site vicinity to regions that incorporate large-scale patches, particularly those in the forward-directivity direction; and (3) presence of geologic structures that can “trap” seismic waves and produce ground motions with large amplitude and long signal duration. The analyses carried out in this work provide a path for interpreting ground-motion site and event specificity obtained from a suite of physics-based simulations, differing only in the rupture model characterization, to inform the selection of simulation scenarios for site-specific engineering analyses.

¹University of Nevada, Reno, Reno, NV, USA
²Lawrence Berkeley National Laboratory, Berkeley, CA, USA
³Lawrence Livermore National Laboratory, Livermore, CA, USA

Corresponding author:
Florianna Petrone, University of Nevada, Reno, 1664 N. Virginia St., Reno, NV 95577, USA.
Email: florianna@unr.edu

Research Paper

Regional-scale fault-to-structure earthquake simulations with the EQSIM framework: Workflow maturation and computational performance on GPU-accelerated exascale platforms

David McCallen, M. EERI¹, Arben Pitarka², Houjun Tang³, Ramesh Pankajakshan³, N Anders Peterson³, Mamun Miah³, and Junfei Huang³

Abstract
Continuous advancements in scientific and engineering understanding of earthquake phenomena, combined with the associated development of representative physics-based models, is providing a foundation for high-performance, fault-to-structure earthquake simulations. However, regional-scale applications of high-performance models have been challenged by the computational requirements at the resolutions required for engineering risk assessments. The Earthquake Simulation (EQSIM) framework's software application development under the US Department of Energy (DOE) Exascale Computing Project, is focused on overcoming the existing computational barriers and enabling routine regional-scale simulations at resolutions relevant to a breadth of engineered systems. This multidisciplinary software development—drawing upon expertise in geophysics, engineering, applied math and computer science—is preparing the advanced computational workflow necessary to fully exploit the DOE's exascale computer platforms coming online in the 2023 to 2024 timeframe. Achievement of the computational performance required for high-resolution regional models containing upward of hundreds of billions to trillions of model grid points requires numerical efficiency in every phase of a regional simulation. This includes run time start-up and regional model generation, effective distribution of the computational workload across thousands of computer nodes.

¹Lawrence Berkeley National Laboratory, Berkeley, CA, USA
²Lawrence Livermore National Laboratory, Livermore, CA, USA
³University of Nevada, Reno, Reno, NV, USA

Corresponding author:
David McCallen, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA.
Email: dmccallen@lbl.gov