2023 INCITE Proposal Submission
Proposal

Title: EQSIM Regional-Scale Simulations for Earthquake Hazard and Risk Assessments

Principal Investigator: David McCallen

Organization: Lawrence Berkeley National Lab

Date/Time Generated: 6/15/2022 8:04:20 PM

Section 1: PI and Co-PI Information

Question #1

Principal Investigator: *The PI is responsible for the project and managing any resources awarded to the project. If your project has multiple investigators, list the PI in this section and add any Co-PIs in the following section.*

Principal Investigator

First Name

David

Last Name

McCallen

Organization

Lawrence Berkeley National Laboratory

Email

dbmccallen@lbl.gov

Work Phone

5102893286

Address Line 1

1 Cyclotron Road

Address Line 2

(No answer given.)

City

Berkeley

State

California

Zip Code

94720

Question #2

Co-PI (s)

First Name

Arben

Last Name

Pitarka

Organization

Lawrence Livermore National Laboratory

Email

pitarka1@llnl.gov

First Name

Houjun

Last Name

Tang

Organization

Lawrence Berkeley Natioal Laboratory

Email

htang@lbl.gov

First Name

Rie

Last Name

Nakata

Organization

Lawrence Berkeley National Laboratory

Email

rnakata@lbl.gov

First Name

Ramesh

Last Name

Pankajakshan

Organization

Lawrence Livermore National Laboratory

Email

pankajakshan1@llnl.gov

Question #3

Institutional Contact: *For the PI's institution on the proposal, identify the agent who has the authority to review, negotiate, and sign the user agreement on behalf of that institution. The person who can commit an organization may be someone in the contracts or procurement department, legal, or if a university, the department head or Sponsored Research Office or Grants Department.*

Institutional Contact

Institutional Contact Name

Jens Birkholzer

Institutional Contact Phone

510 486 7134

Institutional Contact Email

jtbirkholzer@lbl.gov

Section 2: Project Information

Question #1

Select the category that best describes your project.

Research Category

Earth Science: Geological Sciences

Question #2

Please provide a project summary in two sentences that can be used to describe the impact of your project to the public (50 words maximum)

Project Summary

Advanced high performance computing is becoming a key tool for the prediction of earthquake hazard and risk on a regional scale. This project will push the limits of simulation based prediction of regional scale ground motions and infrastructure response

Section 3: Early Career Track

Question #1

Early Career

Starting in the INCITE 2022 year, INCITE is committing 10% of allocatable time to an [Early Career Track](#) in INCITE. The goal of the early career track is to encourage the next generation of high-performance computing researchers. Researchers within 10 years from earning their PhD (after December 31st 2012) may choose to apply. Projects will go through the regular INCITE Computational Readiness and Peer Review process, but the INCITE Management Committee will consider meritorious projects in the Early Career Track separately.

***Who Can Apply**: Researchers less than 10 years out from their PhD that need LCF-level capabilities to advance their overall research plan and who have not been a previous INCITE PI.*

How to Apply:

In the regular application process, there will be a check-box to self-identify as early career.

- *The required CV should make eligibility clear.*
- *If awarded, how will this allocation fit into your overall research plan for the next 5 years?*

Projects will go through the regular INCITE review process. The INCITE Program is targeting at least 10% of allocatable time. When selecting the INCITE Career Track, PIs are not restricted to just competing in that track.

- *What is the Early Career Track?*
  - *The INCITE Program created the Early Career Track to encourage researchers establishing their research careers. INCITE will award at least 10% of allocatable time to meritorious projects.*
- *Will this increase my chances of receiving an award?*
  - *Potentially, this could increase chances of an award. Projects must still be deemed scientifically meritorious through the review process INCITE uses each year.*
- *What do I need to do to be considered on the Early Career Track?*
  - *In the application process, select 'Yes' at 'If you are within 10 years of your PhD, would you like to be considered in the Early Career Track?' You will need to write a paragraph about how the INCITE proposal fits into your 5-year research and career goals.*
- *What review criteria will be used for the Early Career Track?*
  - *The same criteria for computational readiness and scientific merit will be applied to projects in the Early Career Track as will be applied to projects in the traditional track. The different will be manifest in awards decisions by the INCITE management committee.*

Early Career Track

If you are within 10 years of your PhD, would you like to be considered in the Early

Career Track? Choosing this does not reduce your chances of receiving an award.

No

If ‘yes’, what year was your PhD? If 'no' enter N/A

N/A

If ‘yes’, how will this allocation fit into your overall research plan for the next 5 years? If 'no' enter N/A.

N/A

Section 4: INCITE Allocation Request & Other Project Funding/Computing Resources

Question #1

OLCF Summit (IBM / AC922) Resource Request - 2023

Question #2

OLCF Frontier (Cray Shasta) Resource Request – 2023

Question #3

OLCF Frontier (Cray Shasta) Resource Request – 2024

Node Hours

1,500,000

Storage (TB)

810TB

Off-Line Storage (TB)

2,700TB

Question #4

OLCF Frontier (Cray Shasta) Resource Request – 2025

Question #5

ALCF Theta (Cray XC40) Resource Request - 2023

Question #6

ALCF Polaris Resource Request - 2023

Question #7

ALCF Polaris Resource Request - 2024

Question #8

ALCF Polaris Resource Request - 2025

Question #9

ALCF Aurora (Intel X^e) Resource Request – 2023

Question #10

ALCF Aurora (Intel X^e) Resource Request – 2024

Question #11

ALCF Aurora (Intel X^e) Resource Request – 2025

Question #12

List any funding this project receives from other funding agencies.

Funding Sources

Question #13

List any other high-performance computing allocations being received in support of this project.

Other High Performance Computing Resource Allocations

Section 5: Project Narrative and Supplemental Materials

Question #1

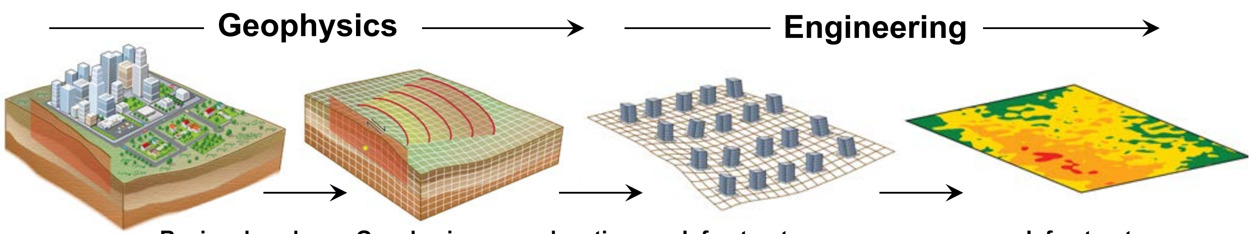
Using the templates provided here, please follow the [INCITE Proposal Preparation Instructions](#) to prepare your proposal. Elements needed include (1) Project Executive Summary, (2) Project Narrative, (3) Personnel Justification and Management Plan, (4) Milestone Table, (5) Publications Resulting from prior INCITE Awards (if appropriate), and (6) Biographical Sketches for the PI and all co-PI's. Concatenate all materials into a single PDF file. Prior to submission, it is strongly recommended that proposers review their proposals to ensure they comply with the proposal preparation instructions.

Concatenate all materials below into a single PDF file.

1. **Project Executive Summary (One Page Max)**
2. **Project Narrative (15 Pages Max)**
3. **Personnel Justification and Management Plan (1 Page Max)**
4. **Milestone Table**
5. **Publications resulting from prior INCITE Awards (if appropriate)**
6. **Biographical Sketches for the PI and all co-PI's.**

INCITEMcCallenetal.pdf
The attachment is on the following page.

EQSIM Regional-Scale Simulations for Earthquake Hazard and Risk Assessments



A New Proposal to FY2023 DOE INCITE Leadership Computing

David McCallen[1], PI
Lawrence Berkeley National Laboratory
& University of Nevada, Reno

Arben Pitarka, Co-I
Lawrence Livermore National Laboratory

Houjun Tang, Co-I
Lawrence Berkeley National Laboratory

Rie Nakata, Co-I
Lawrence Berkeley National Laboratory

Ramesh Pankajakshan, Co-I
Lawrence Livermore National Laboratory

Submitted by the Lawrence Berkeley National Laboratory

June 17, 2022

[1] Corresponding Investigator: dbmccallen@lbl.gov, (510) 289-3286

Table of Contents

Project Summary

The large-scale research simulations proposed herein occur at the nexus of earth science and engineering and are focused on advancing the utilization of large-scale high-performance simulations in assessing site-specific regional earthquake hazard and risk. The proposed simulation activities consist of multidisciplinary, coupled geophysics and engineering simulations for regional-scale fault-to-structure modeling that starts from an earthquake fault rupture, propagates seismic waves through a heterogenous earth and finally represents the complex interaction between incident seismic waves and critical infrastructure systems at the surface/subsurface boundary. The proposed simulations have only recently become feasible as a result of major software and workflow advancements in the DOE Exascale Computing Project EQSIM application development, coupled with the advancements of the FRONTIER exaflop system. The EQSIM project has been focused on developing capabilities for the emerging exaflop systems with demonstrated performance on SUMMIT and early successful testing on CRUSHER. The simulation tasks proposed will yield new information on ground motion severity, its spatial variability, and building response in the near field of major earthquakes, the manner in which incident seismic waves actually interact with critical infrastructure (as opposed to legacy simplified engineering idealizations) and the number fault rupture realizations that are required to fully characterize the risk to critical infrastructure. Finally, the unique suite of simulated ground motion datasets generated will be shared via open source for utilization by the broad earthquake research community of earth scientists and engineers to help promote additional research and practical utilization of synthetic ground motions.

INCITE Allocation Request

The INCITE allocation request along with other funding and computing resources and support are summarized in Table 1. The core of the INCITE request includes FRONTIER resources for a number of large earthquake simulations of the San Francisco Bay Area.

Table 1. INCITE request for FRONTIER resources and complementary support, these numbers are based on EQSIM performance experience to-date on CRUSHER.

INCITE Request	
Item	**Resource Request**
~10 San Francisco Bay Area regional-scale ground motion simulations for a M7 Hayward fault earthquake (max frequency = 10 Hz, Vsmin = 150 m/s) and 1 exploratory simulation of a M7.5 San Andreas fault earthquake	1,500,000 FRONTIER Node-Hours

ECP and LBNL Complementary Support	
Item	**Resource Request (EQSIM ECP Project)**
Project team member support	Provided by the EQSIM project
Selected nonlinear building response simulations for the 10 SFBA simulations	70,000 PERLMUTTER Node-Hours (Director's allocation)

(A) Executive Summary

Large earthquakes present a significant risk around the world and are a major societal challenge for virtually every element of the built environment. Traditional earthquake hazard and risk assessments for critical infrastructure have relied on empirically based approaches that use historical earthquake ground motions from many different locations to estimate future earthquake ground motions at a specific site of interest. Given the fact that ground motions for a particular location are strongly influenced by the physics of the specific earthquake processes, including the earthquake fault rupture mechanism, seismic wave propagation through the heterogeneous earth, and site response at the location of a particular facility, earthquake ground motions are very complex with significant spatial variation in both frequency content and amplitude. The homogenization of historical records from many disparate sites in traditional empirically based ground motion estimates cannot fully capture the complex site-to-site variability of ground motion that is necessary characterize in order to ensure the most resilient built environment. Over the past decade, interest in utilizing high performance simulations to characterize earthquake ground motions (earthquake hazard) and infrastructure response (earthquake risk) has appreciably accelerated. However, the extreme computational demands required to execute hazard and risk simulations at regional scale have presented a significant barrier.

Over the past four years, the DOE Exascale Computing Project (ECP), has supported the development of a unique application workflow for fault-to-structure simulations of earthquake processes at regional scale. The **E**arth**Q**uake **SIM**ulation (EQSIM) framework development has made significant advancements towards high-fidelity regional earthquake simulations. The focus of EQSIM development has been to provide the capabilities required to remove existing computational barriers and realize the full potential of simulation-based contributions to regional-scale hazard and risk evaluations. Based on the ECP project developments, EQSIM now allows computationally efficient regional ground motion simulations at frequencies relevant to a large breadth of engineered systems with resolution of frequencies up to 10 Hz.

The INCITE proposal described herein will significantly leverage the most recent major software and workflow developments of EQSIM to provide an unprecedented exploration of deterministic, physics-based simulations of both earthquake ground motions and corresponding infrastructure response. Based on EQSIM advanced capabilities and the preparation and demonstrated of readiness to fully exploit the FRONTIER platform, this work will utilize existing large regional-scale models of the San Francisco Bay Area (SFBA) generated for the ECP project to perform a full suite of scenario earthquake simulations for large SFBA earthquakes. This will include a number of fault rupture realizations and parametric variations (including fine-scale, stochastic geologic heterogeneities) for each earthquake scenario. This work will represent the next important and necessary step towards the full development of simulation-based earthquake hazard and risk assessments, and will yield unprecedented new knowledge on the complex characteristics of the regional distribution of earthquake risk. The simulated ground motion datasets will be broadly shared through open access with the earthquake science and engineering communities to promote the broad dissemination and exploitation of synthetic earthquake ground motion data.

(B) Project Narrative

Impact of Proposed Work and Time Relevance

The ability to understand the regionally distributed, site-specific ground motions that will occur in future earthquakes is essential to reliably ensuring the seismic adequacy and resilience of major infrastructure systems. Currently, empirical models that are based on historically observed and measured earthquake motions form the basis for future ground motion estimates. These empirical models are typically hampered by the lack of historical data from large earthquakes in the region of interest, combined with the fact that even for earthquakes that have been observed and measured, the data available is typically quite sparse in terms of spatial coverage especially at short distances from the fault. For example, Figure 1 shows the historical earthquake data available in the existing Pacific Earthquake Engineering Research Center strong motion database. Each circle in this plot represents one measured ground motion dataset from an historical earthquake. Inspection of this plot immediately shows that in the region of highest interest for infrastructure damage, the near-field region of large earthquakes (0-20 Km and magnitudes 5 to 8), the historical data is quite sparse. Consequently, while these empirical models can provide some statistical insight on potential future earthquake motions, they cannot fully inform and predict the site-to-site variability of motions.

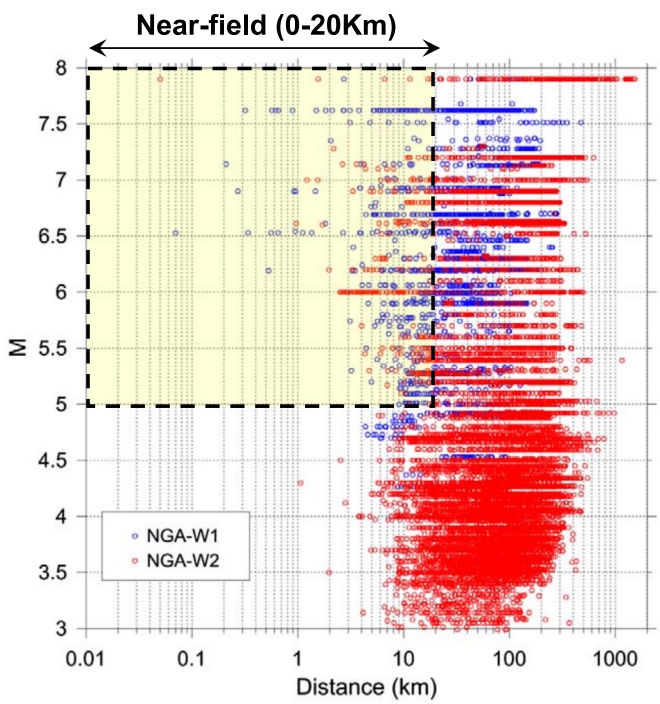


Figure 1. Existing datasets of historical measured earthquake ground motion time histories in the PEER strong-motion database.

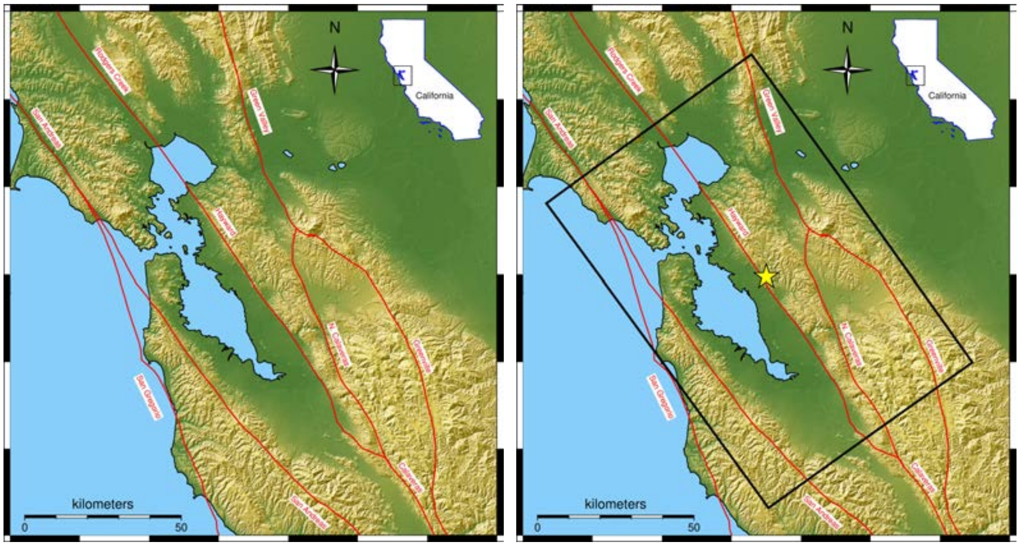
With the limitations of empirically based ground motion models, and the need to accurately understand the site-specific motions impacting specific structures, there has been a significant and growing interest in utilizing physics-based, high-performance simulations to predict future ground motions and infrastructure earthquake response. However, the computational requirements of regional-scale simulations are extreme. It is particularly challenging to execute ground motion simulations at the frequency resolutions that are relevant to engineered infrastructure. As an example, consider regional-scale modeling of the SFBA in California as shown in Figure 2. For a geophysics wave propagation model encompassing the entire urban area and major earthquake faults such as the Hayward fault (a 120Km x 80Km x 30Km domain), the computational effort for modeling an earthquake fault rupture and resulting regional propagation of seismic waves is proportional to the model physical volume, the duration of the earthquake simulation, the fourth power of the frequency being resolved in the simulation and the inverse of the forth power of the minimum shear wave speed being resolved in the geologic model as indicated in Figure 2. These fourth order dependencies provide a steep computational demand to overcome in order to get to earthquake simulations of high value and utility.

Until recently, the ability to compute ground motions to higher frequencies and to resolve soft near-surface soil deposits with low shear wave speed (Vsmin) has been extremely limited. Historical simulations of the SFBA for example have been computationally limited to 2-3 Hz, with a very limited number of fault rupture realizations. In light of the fact that many engineered structures can have natural vibration frequencies on the order of 5 to 10 Hz, and that near surface soft sediments can amplify earthquake ground motions, these computational demands have been limiting. In terms of realizing the ultimate utility of simulated motions for infrastructure risk assessments, extending simulations into these very computationally demanding regimes is essential.

Some research groups have addressed this challenge through a hybrid method whereby physics-based wave propagation models are only used to simulate motions to 1 or 2 Hz and these resulting deterministic waveforms are then merged with fully stochastic representations of higher frequency motion waveforms. The EQSIM Exascale application development project was specifically aimed at achieving the computational advancements necessary to:

- Allow regional-scale, physics-based, fully deterministic ground motion simulations at the higher frequencies relevant to engineered systems (i.e. up to 10 Hz), which is challenged by the fact that computational effort varies as frequency to the fourth power;
- Computationally resolve near-surface sediments with shear wave speeds (Vsmin) down to 150 to 200 m/s, which is challenged by the fact that computational effort varies inversely with the fourth power of Vsmin;
- Achieve earthquake scenario simulation wall clock times that are of sufficiently short duration to allow the necessary multiple realizations of a given scenario earthquake to account for multiple realizations of the fault rupture process, geologic variations etc.

Over the past four years in the EQSIM ECP application development project, there have been major enhancements in the ability to execute high-fidelity regional-scale ground motion simulations. Starting with the existing SW4 fourth order seismic wave propagation code, advancements have included a number of algorithmic improvements including advanced mesh refinement to optimize the computational mesh in alignment with the natural depth variation of geologic properties, major enhancements to parallel I/O based on efficient data compression and utilization of efficient HDF5 data containers, and optimization of SW4 for execution on GPU-based platforms. In addition, the full workflow for fault-to-structure simulations has been completed with rigorous coupling between global geophysics and local engineering system models. The adopted coupling workflow, based on the Domain Reduction Method (DRM), allows full representation of a complex wavefield, consisting of inclined body waves and surface waves, on local soil/structure systems.



San Francisco Bay Area and Major Faults

San Francisco Bay Area EQSIM Computational Domain (120 x 80 x 30 Km)

Computational Effort $\propto$ **(Model Volume) x (Earthquake Duration) x (Freqmax / Vsmin)^4**

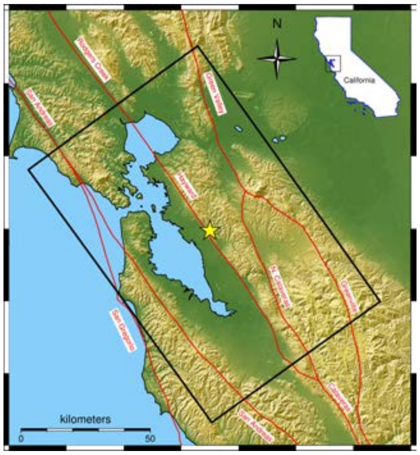
	↓	↓	↓
Examples	**120 x 80 x 30** **288,000 Km^3**	**M = 7 Hayward Fault** **Event ~ 90 seconds**	**Freqmax = 10 Hz** **Vsmin = 150 m/s**

Figure 2. Regional scale model of the SFBA California and computational effort associated with a regional-scale simulation.

As part of the EQSIM development, annual regional simulation performance assessments are tracked with a DOE ECP Project approved application Figure of Merit (FOM). For the EQSIM project the FOM is expressed as shown in Figure 3. The FOM includes a measure of computational effort in terms of maximum frequency and minimum shear waves speed resolved, and the wall clock time for an earthquake simulation. The EQSIM exascale simulation challenge problem was defined in terms of a fixed SFBA regional model and simulation of a M7 Hayward fault earthquake. The 7.6 factor in the FOM formula was included to yield a FOM of 1.0 in the initial regional-scale simulation performed with the EQSIM framework at the start of the project.

The advancements in EQSIM performance have been documented with an annual fiscal year performance evaluation and update of the FOM. The progress achieved in frequency and Vsmin resolution since the inception of EQSIM development are shown in Figure 4. The simulation capabilities have increased from a frequency resolution of 3.5 Hz and Vsmin of 500 m/s at the start of EQSIM development to a frequency resolution of 10 Hz and Vsmin of 140 m/s. In addition, as shown in Figure 5, the peak FOM has increased from an initial value of 1 to a value of 5013, reflecting a major leap in regional simulation capability. The major advancements in FOM reflect the integrated effects of advanced algorithm development, enhanced workflow and I/O, and transition from CPU based platforms (CORI) to GPU accelerator platforms (SUMMIT and CRUSHER).

Challenge problem of high frequency regional simulations (SFBA model)



Frequency resolved

$$\text{FOM} = \frac{(\text{Freq})^4}{\text{Wall Clock Time x 7.6}}\left[\frac{500}{\text{Vsmin}}\right]^4$$

Run time **Softest sediment resolved**

- Reflects the challenge of resolving high frequencies
- Reflects the challenge of resolving low velocity near-surface sediments
- Reflects the time of solution

Figure 3. The EQSIM Figure of Merit for tracking computational progress; regional-scale model of the SFBA provides the domain for regional-scale simulations and determination of the EQSIM FOM.

To put this newly achieved simulation capability into proper perspective, the computational throughput achieved with EQSIM relative to all previous regional earthquake ground motion simulations of the SFBA is shown in Figure 6. The steep sequential advancements in the EQSIM regional simulations are evident starting in 2017 and progressing to 2021.

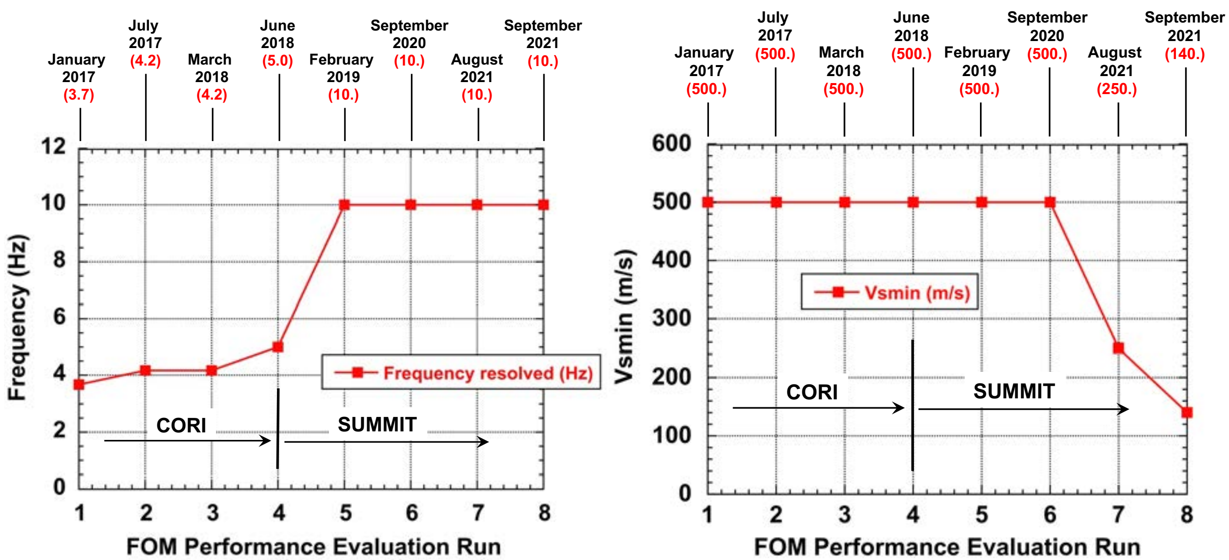
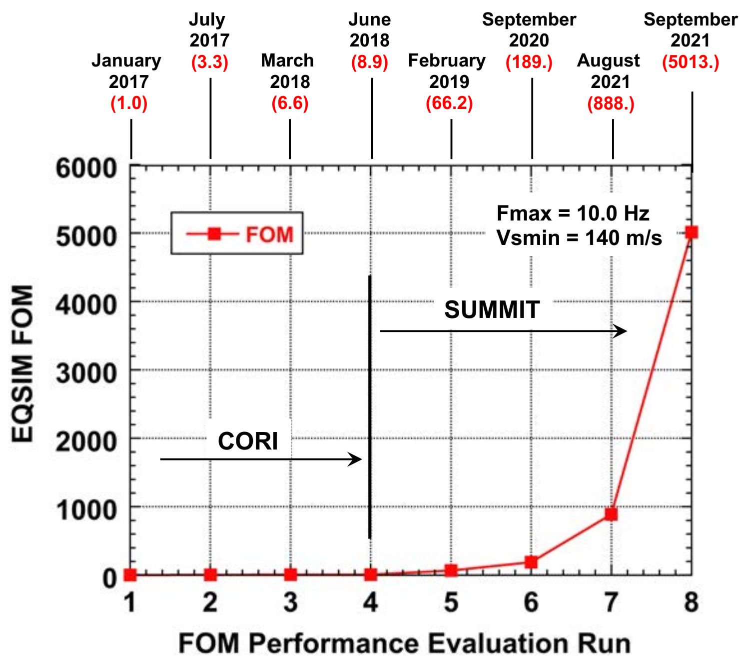


Figure 4. Advancements in frequency content and Vsmin resolution of SFBA earthquake simulations since the inception of the EQSIM application development.

With the EQSIM computational advancements, combined with emerging exascale capabilities embodied in the FRONTIER platform, we are positioned to perform a set of unprecedented deterministic simulations of future SFBA earthquakes. The overarching objectives of the proposed simulation work are to provide new knowledge and insight into the complex regional-scale distribution of earthquake ground motions and infrastructure risk, and to provide a unique large-scale set of simulated regional ground motions available to the earth science and engineering communities in an open-source format. It is noted that historically the Hayward fault experiences a major earthquake on average every 150 years, and it has been 153 years since the last Hayward fault event, therefore any and all new information on Hayward fault earthquake ground motions and the resulting risk to infrastructure will be of high value and interest.



$$\text{FOM} = \frac{(\text{Freq})^4}{\text{Wall Clock Time} \times 7.6}\left[\frac{500}{\text{Vsmin}}\right]^4$$

Figure 5. Advancements in EQSIM Figure of Merit since the inception of the EQSIM ECP application development project.

Objectives

The EQSIM application development capabilities have provided the integrated algorithms, data management and scalability to overcome the computational barriers to physics-based, high-fidelity regional-scale simulations of earthquake processes. The uniqueness and impact potential of the work described in this proposal for the earth science and earthquake engineering communities is fully dependent on these timely advancements. The simulations defined in this proposal will exploit these capabilities in confluence with arriving exaflop platforms to provide new computationally-based scientific and engineering discovery and knowledge including:

- Deeper insight and understanding of the complex regional spatial distribution of earthquake ground motions and infrastructure risk for major earthquakes, with a particular focus on the near-field of large earthquakes where historical measured ground motion data is extremely sparse (Figure 1);

- Exploit the detailed 3D simulations for new understanding of the interaction between complex incident seismic waveforms, including inclined body waves and surface waves, and infrastructure systems. This will provide a critical evaluation of the realism of decades-old engineering simplifying idealizations based on the assumption that all incident seismic waves arriving at a site consist of one-dimensional, vertically propagating shear and compressional waves;
- Provide a unique high-fidelity dataset of spatially dense simulated earthquake ground motions for the SFBA that will be placed on the Lawrence Berkeley National Laboratory SPIN web server system (described below) for open community access and broad utilization by the research and earthquake engineering practitioner communities;
- Provide new insight into the number of fault rupture realizations that should be included in a simulation-based risk assessment to account for uncertainties in the fault rupture process for a specific scenario earthquake event. Risk to a particular structure will be defined in terms of median building response with an associated statistical distribution across all of the fault rupture realizations;
- Explore the ability to extend high frequency fault-to-structure simulations to even larger earthquake scenarios in the range of a M7.5 event.

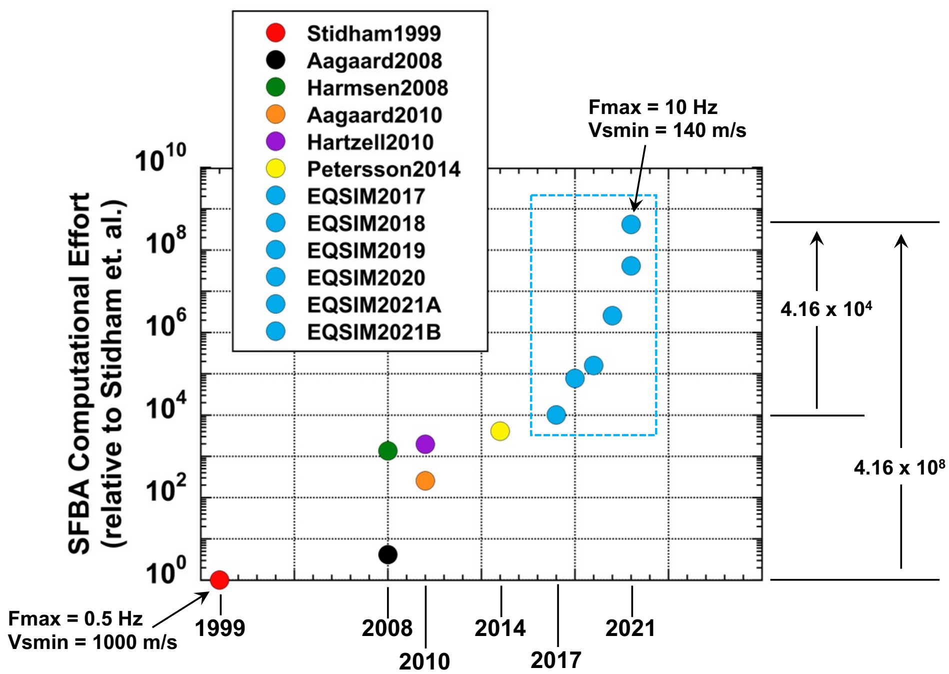


Figure 6. All regional-scale simulations performed for the SFBA (EQSIM advancements shown in blue).

To achieve these objectives, the existing regional geophysics model of the SFBA, along with a set of representative steel frame building models that have been created, will be utilized in a significant set of Hayward fault earthquake realizations. In general, the precise location of the hypocenter and rupture evolution for a specific earthquake is not well constrained so a number of different fault rupture realizations must be executed to span the potential rupture possibilities.

Based on investigations recently performed by the EQSIM team, it is anticipated that on the order of 10 fault rupture realizations will be necessary to capture the statistics of risk for building structures within the parametric variability of the kinematic fault rupture model incorporated in EQSIM. Thus, this work proposes executing approximately 10 high-fidelity Hayward fault rupture simulations (Fmax = 10Hz, Vsmin ~150 m/s) for a magnitude 7 Hayward fault earthquake. The work will also explore the limits of high-fidelity regional simulations and the ability to further scale-up the simulations to a much larger M7.5 earthquake on the San Andreas fault as shown in Figure 7, including identification of the highest frequency resolution that can be resolved with current software and hardware for this very large event.

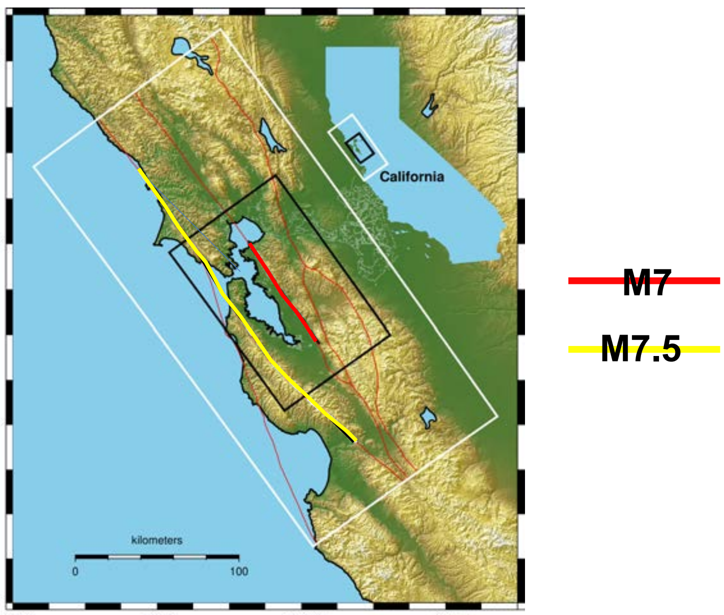


Figure 7. Comparative computational domains for a M7 Hayward fault earthquake and a M7.5 San Andreas fault earthquake.

Benchmarking and Execution Readiness

In addition to individual code advancements, major work has focused on completing the development and testing of an effective and efficient fault-to-structure workflow in EQSIM. This workflow includes two major components as outlined in Figure 8 and Figure 9. The first component of workflow is associated with the ground motion simulations. This includes efficient auto-generation of regional models of 100's of billions of grid points at runtime, effective time stepping to solve the three-dimensional wave equation with the efficient creation of large checkpoint files at user specified internals, and the efficient writing of compressed ground motion waveform data into an HDF5 container. The second component of workflow consists of utilization of generated ground motion data to perform infrastructure response simulations for specific infrastructure systems at a large number of ground surface sites. This workflow allows traditional fixed-based simulations whereby the synthetic ground motions are applied directly to a fixed-base structural model, and a second option for generating a soil island - infrastructure model to allow the representation of soil-structure-interaction and complex incident seismic waves.

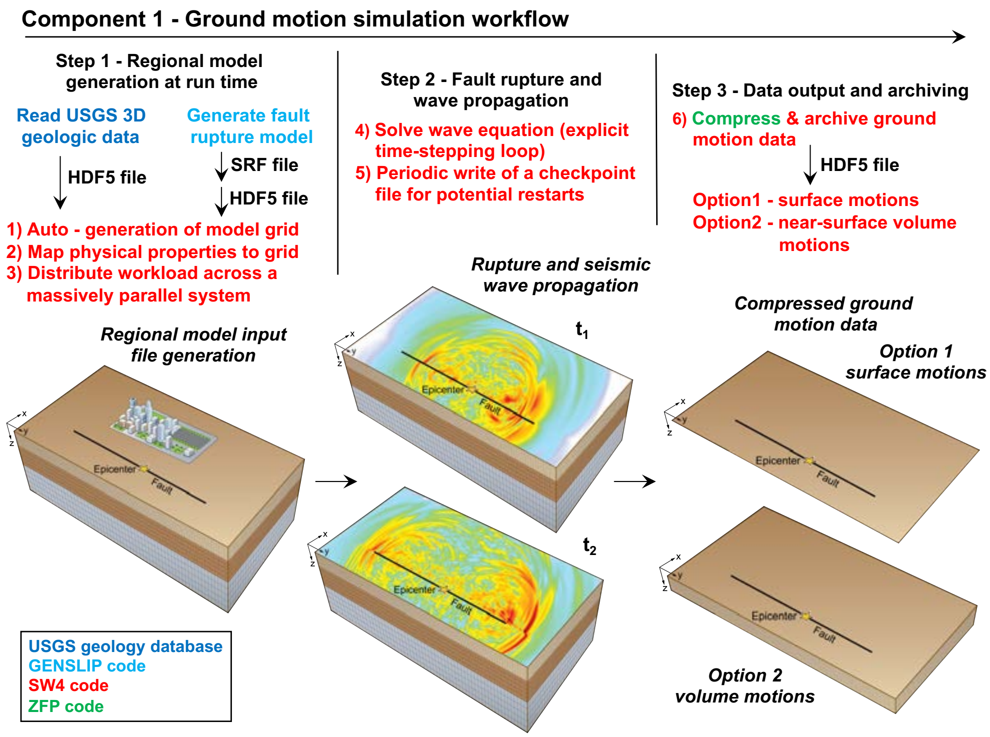


Figure 8. Component 1 workflow elements developed for the EQSIM fault-to-structure workflow including software code elements, this component on regional ground motion simulation will be executed on the FRONTIER platform.

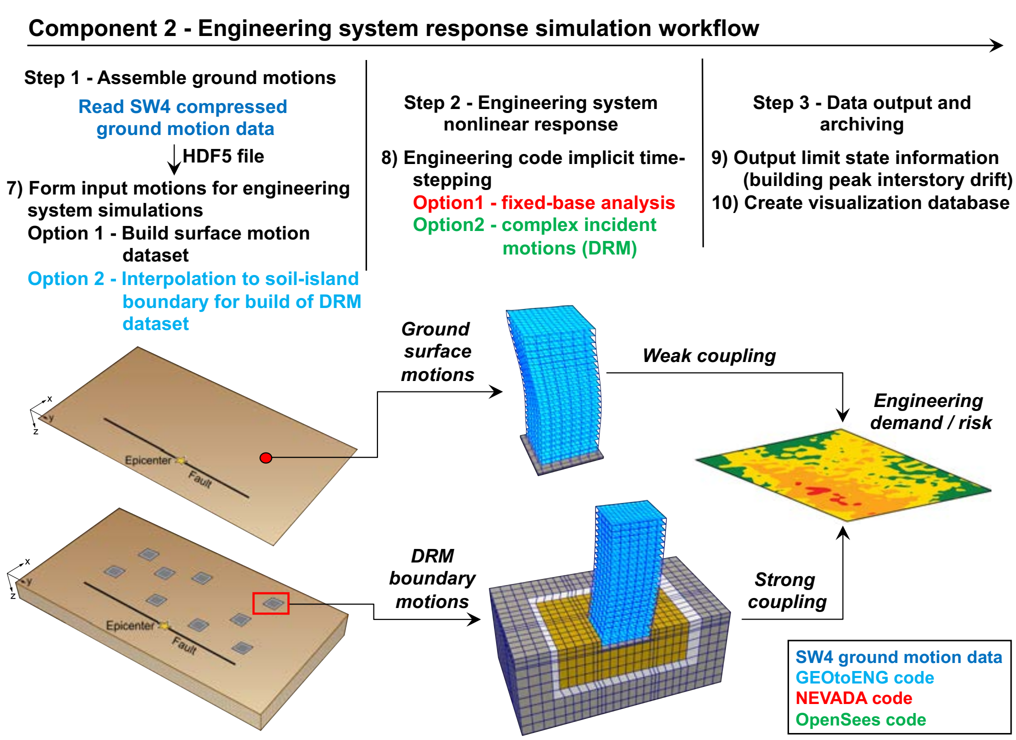
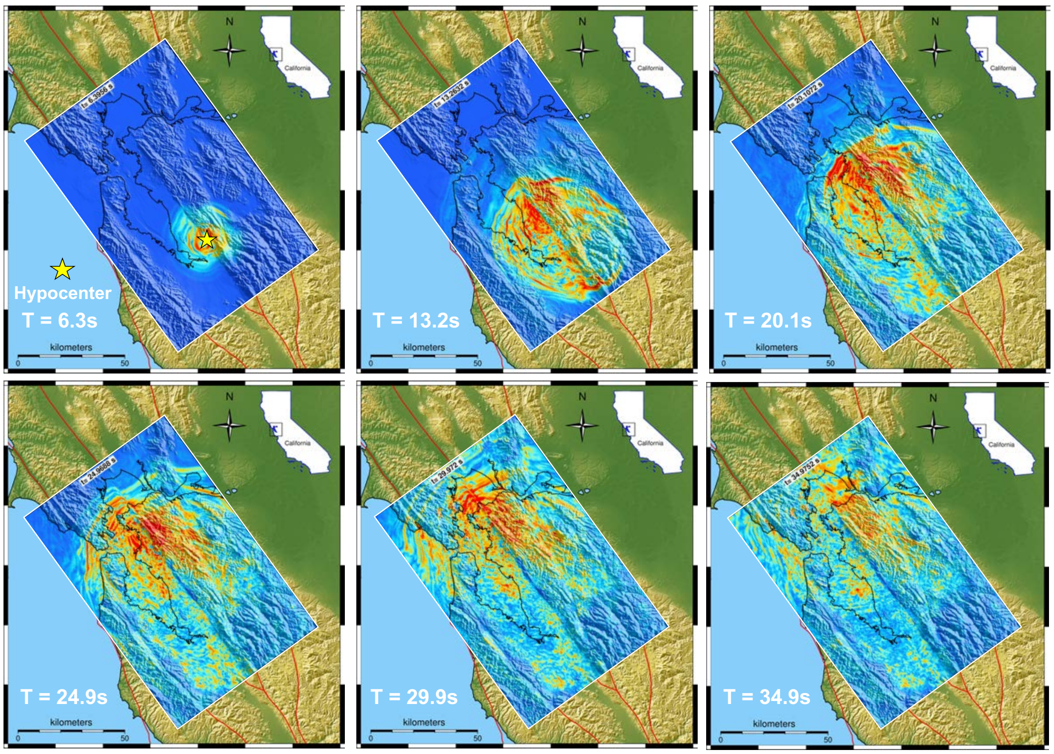


Figure 9. Component 2 workflow elements developed for the EQSIM fault-to-structure workflow including software code elements, this component on infrastructure response to simulated ground motions would be executed on the PERLMUTTER platform.

The EQSIM capabilities and the fault-to-structure workflow have been extensively tested and exercised on the SUMMIT platform in high-fidelity ground motion simulations and full fault-to-structure simulations. For example, Figure 10 illustrates a high-fidelity simulation of a M7 Hayward fault earthquake which exercised a 391 billion grid point model for a 90 second duration earthquake simulation using 3,600 nodes of SUMMIT in a 43-hour run (155,000 node-hours) generating 267 TB of ground motion data. The generated ground motion datasets can subsequently be transferred to CORI or PERLMUTTER to complete infrastructure system simulations to yield information on the complex regional distribution and degree of risk to specific infrastructure as illustrated in Figure 11.



Fmax = 10 Hz
Vsmin = 140 m/s
Computational domain 120 x 80 x 30 Km
Grid points = 391 billion
Grid size at ground surface = 1.75 m
3,600 Summit nodes
42.7 hour run time
Ground motion data = 267 TB (2.6 TB comp)

Figure 10. Workflow component 1 - a large-scale SFBA M7 Hayward fault earthquake ground motion simulation and performance metrics on SUMMIT.

In preparation for the earliest demonstration of achievement of the EQSIM FOM on FRONTIER, EQSIM has been undergoing extensive testing on CRUSHER. The current evaluations of EQSIM on CRUSHER have included successful compilation, successful completion and correctness confirmation for a full suite of SW4 wave propagation test problems and testing of the CRUSHER network performance through execution of a suite of increasingly large problems. Figure 12 illustrates the EQSIM FOM obtained as a function of number of nodes employed on CRUSHER, SUMMIT and PERLMUTTER for a maximum modeled frequency of 5Hz and minimum Vs of 250 m/s SFBA simulation of a M7 Hayward fault event. Early scaling tests on CRUSHER, utilizing the limited number of nodes available, are shown in Figure 13. CRUSHER is still undergoing significant performance optimization work that will ultimately transfer to FRONTIER, but this early performance testing provides confidence that EQSIM will be able to exploit the FRONTIER network and achieve large-scale performance.

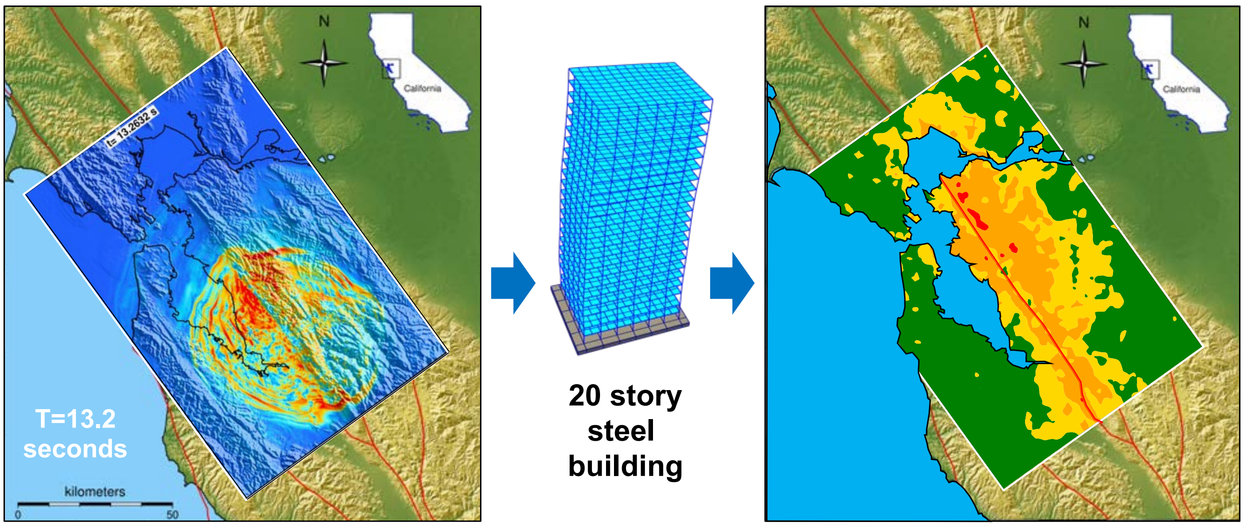


Figure 11. Workflow component 2 - completion of fault-to-structure simulation workflow resulting in determination of the complex SFBA regional distribution of building risk (in terms of peak interstory drift) for typical 20 story steel frame buildings with building response simulations performed on CORI.

A single EQSIM regional earthquake simulation generates a large volume of earthquake ground motion synthetic data (~260TB uncompressed). To optimally understand the information contained in these datasets and extract the most insight, a special web-based interrogator has been created using the SPIN system at LBNL. The SPIN system, specifically created to allow web-based display of large datasets, stores videos and images of data on a community file system at user selected spatial density, and interactively displays pre-generated plots of ground motion data as shown in Figure 14. The EQSIM-vis software has been developed to efficiently manipulate and display ground motion data in real-time on a Google maps background on the SPIN system. Operationally, EQSIM-vis allows the user to use mouse functions (drag and click) to zoom-in, zoom-out and selectively displace ground motion plots in png format in a highly interactive and low-latency interface. This provides the user with a very powerful tool for interrogation of large datasets. The EQSIM-vis software will be utilized for the proposed INCITE simulations to extract the maximum insight from the large suite of earthquake simulations and will provide the open access to the generated ground motion data.

In summary, the proposed INCITE simulations will be performed by an existing, dedicated and high performing team, fully exploiting four years of major simulation capability advancements, to define the path forward and leading edge of high-fidelity regional-scale simulations.

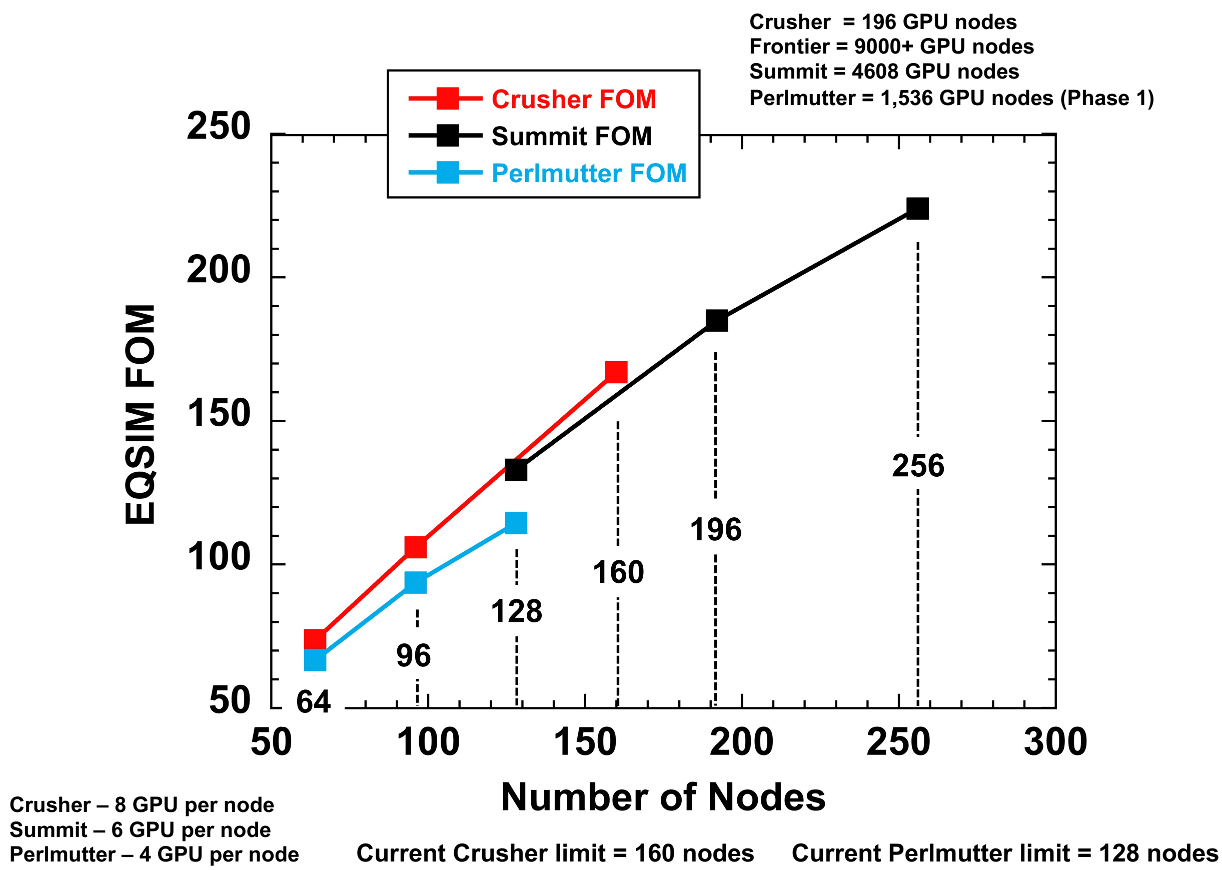


Figure 12. EQSIM performance for a regional-scale SFBA simulation, early runs on CRUSHER nodes (Fmax=5Hz, Vsmin=250m/s, 16.47 billion grid points).

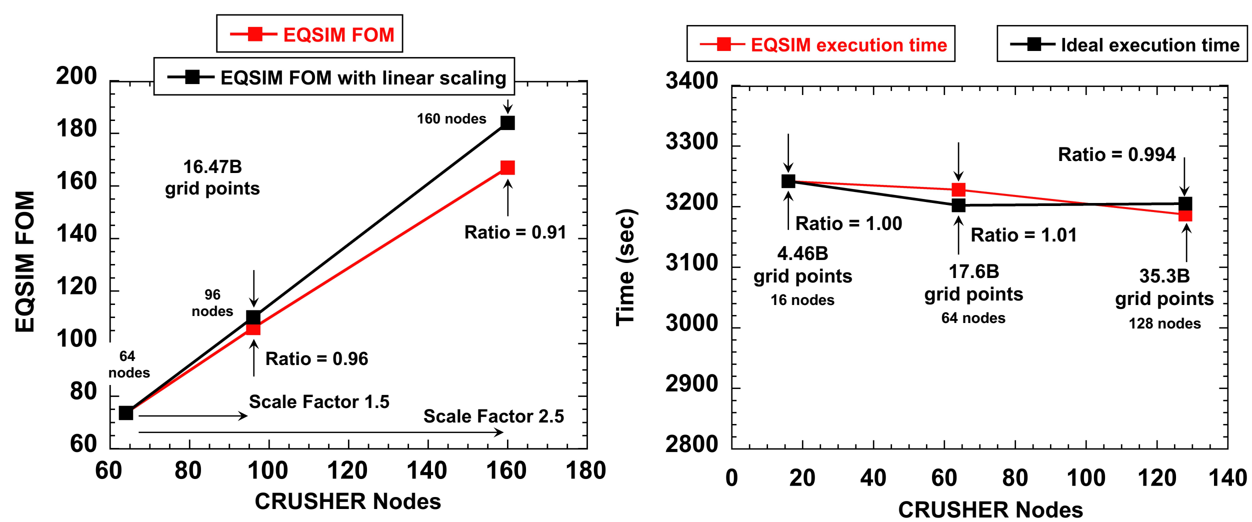


Figure 13. Early scaling tests on CRUSHER within the current limited access to 160 CRUSHER nodes; strong scaling (left), weak scaling (right).

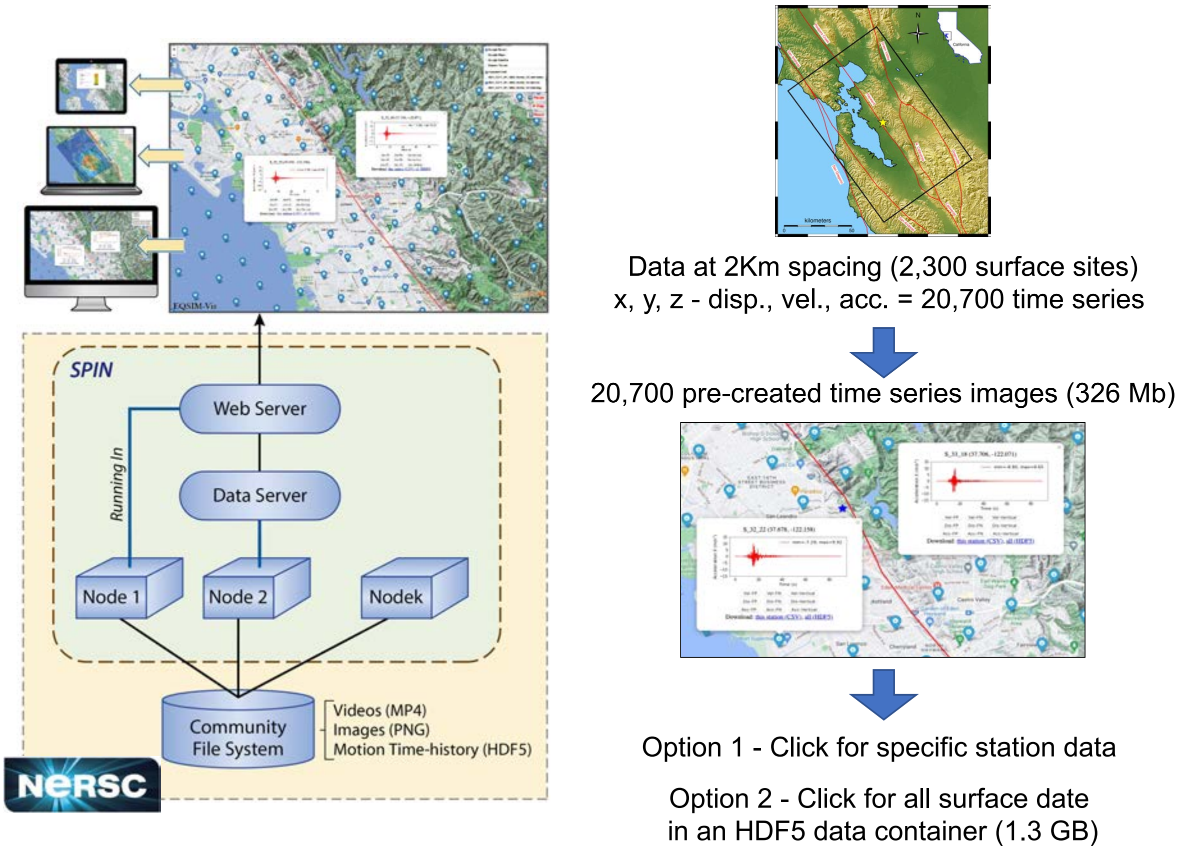


Figure 14. The EQSIM-vis software running on the LBNL SPIN server system allows efficient rapid display, interrogation and selective download of simulated ground motion data.

(C) Personnel Justification and Management Plan

The simulation execution team will consist of a subset of experts from the existing EQSIM ECP application development team, the full EQSIM team is shown in Figure 14, those outlined in red will be responsible for executing large-scale INCITE simulations. These experienced team members have been core contributors to the EQSIM framework development and have worked closely together for the past four years and will be responsible for the successful execution of the proposed earthquake simulations. The effort will draw upon other members of the EQSIM development team if needed. The INCITE simulation team members include:

Dr. David McCallen (PI) - ECP EQSIM application development PI and expert in earthquake engineering and simulation of the earthquake response of infrastructure systems. Dr. McCallen will serve as project PI and oversee overall execution of the large-scale simulations.

Dr. Arben Pitarka (CO-I) - Seismologist and expert in the numerical simulation of seismic wave propagation and co-developer of the widely recognized and utilized Graves-Pitarka fault rupture model. Dr. Pitarka will assist in running simulations and will be the creator of a representative suite of M7 Hayward fault rupture models.

Dr. Rie Nakata (CO-I) - Seismologist and expert in the numerical modeling of seismic wave propagation. Dr. Nakata has expertise and experience executing the SW4 seismic wave propagation code and the characterization of geologic structure for use in regional-scale simulations.

Dr. Houjun Tang (CO-I) - Computer Scientist and expert in large-scale parallel data management. Dr. Tang has been the designer and implementor of the SW4 parallel data management structures and developer of the EQSIM-vis software. Dr. Tang also works in the LBNL HDF5 group and has deep knowledge of HDF5 libraries as well as data compression with the U.S. DOE developed ZFP compression software.

Dr. Ramesh Pankajakshan (CO-I) - Computer Scientist with extensive experience with large-scale simulations on leadership class HPC systems. Dr. Pankajakshan has been instrumental in porting and performance optimization of the SW4 application to GPU based systems including the implementation of RAJA library functions and profiling and debugging tool applications for GPUs.

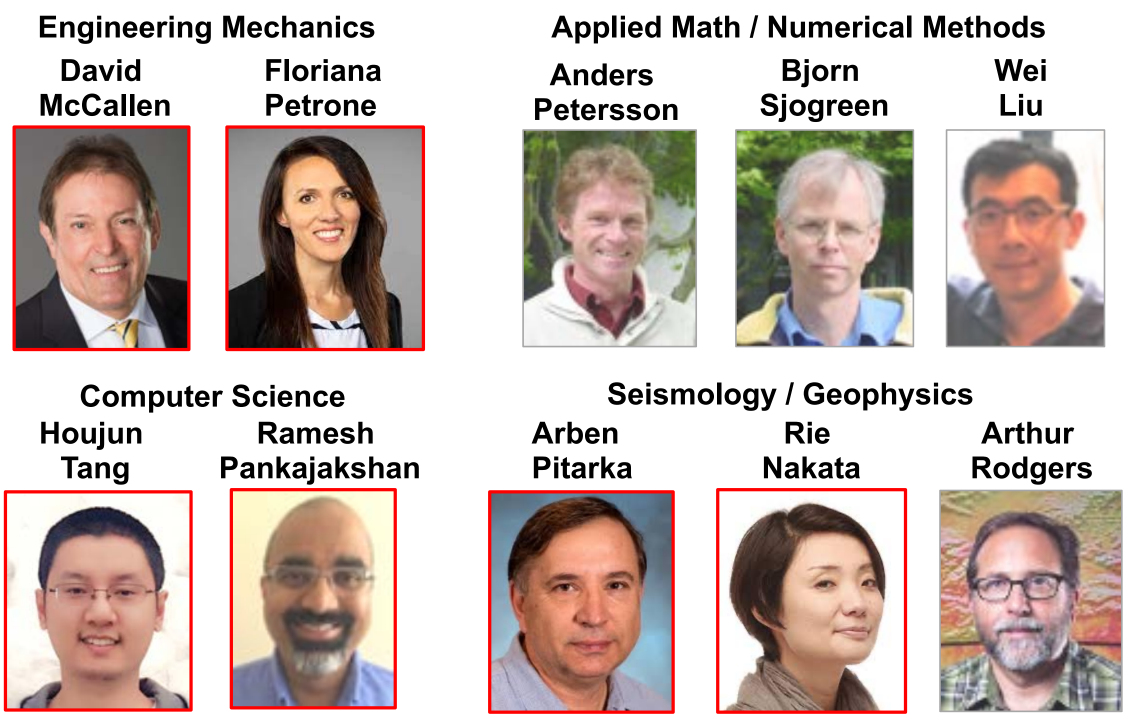


Figure 14. The EQSIM ECP application development team.

Team member collaborators

1) David McCallen
Floriana Petrone UNR
Anders Petersson LLNL
Bjorn Sjogreen LLNL
Wei Liu LLNL
Anders Petersson LLNL
Arthur Rodgers LLNL
Norman Abrahamson UCB
Jenna Wong SFSU
Frank McKenna UCB
Boris Jeremic UCD

2) Arben Pitarka
Keith Koper U of Utah
Christine Goulet SCEC
Benchun Duan T A&M
Christian Poppeliers SNL
Floriana Petrone UNR
Anders Petersson LLNL
Arthur Rodgers LLNL

3) Houjun Tang
Stephen Bailey LBNL
Norman Abrahamson UCB
Jay Lofstead SNL
Zarija Lukic LBNL
Sarp Oral ORNL
Anders Petersson LLNL
Florina Petrone UNR
Arthur Rodgers LLNL
Robert Ross ANL
Kesheng Wu ORNL

4) Rie Nakata
David Lumley UT Dallas
Kimi Mochizuki U of Tokyo
Masa Kinoshita U of Tokyo
Nori Nakata LBNL
Gary Hampson Downunder Geo
Kurt Nihei Chevron
Arthur Rodgers
Anders Petersson

5) Ramesh Pankajakshan
Ian Karlin NVIDIA
John Gyllenhaa LLNL
David Richards LLNL
Judith C Hill LLNL
Edgar Borja LLNL
Trent D' Hooge LLNL
Olga Pearce T A&M
Benjamin Casses LLNL
Adam Bertsch LLNL
Mathew Leininger LLNL
Adam Moody LLNL

(D) Milestone Table

Proposal Title **EQSIM Regional-Scale Simulations for Earthquake Hazard and Risk**		**Lead PI** **David McCallen**
Year 1	Total number of FRONTIER node hours: 1,500,000	
Milestone	Details	Dates
High fidelity SFBA regional-scale simulations of M7 Hayward fault earthquake realizations	Resource: Frontier Total node-hours: 1,200,000 Production runs (# of nodes): ~10 (~3,600 Frontier nodes) Filesystem storage (TB and dates): 810TB (12/22 – 12/23) Archival storage (TB and dates): 2,700TB (12/22 – 12/23) Software Application: EQSIM (SW4) Tasks: Execute large-scale FRONTIER runs, verify ground motions validity, migrate data to LBNL for building simulations and enter compressed datasets into LBNL SPIN system for community access Dependencies: RAJA, HDF5, ZFP, MPI, HIP, ROCm, Umpire	December 22 - December 23
Exploratory simulation of a M7.5 Hayward fault earthquake	Resource: Frontier Total node-hours: 300,000 Production runs (# of nodes): 1 (~6,000 Frontier nodes) Filesystem storage (TB and dates): 1,300TB (3/23 – 12/23) Archival storage (TB and dates): 1,300TB (3/23 – 12/23) Software Application: EQSIM (SW4) Tasks: Perform exploratory large run for a M7.5 earthquake, migrate data to LBNL for building simulations and enter compressed dataset into SPIN system for community access Dependencies: RAJA, HDF5, ZFP, MPI, HIP, ROCm, Umpire	March 23 - December 23
Building response simulations for regional scale risk and understanding of seismic wave - structure interaction	To be completed by the EQSIM project utilizing the PERLMUTTER platform at LBNL	

(E) Publications Resulting from the EQSIM ECP Development Project

McCallen, D., Tang, H., Wu, S., Eckert, E., Huang, J., & Petersson, N. A. (2022), Coupling of Regional Geophysics and Local Soil-Structure Models in the EQSIM Fault-to-Structure Earthquake Simulation Framework, The International Journal of High-Performance Computing Applications, 36(1), 78-92.

McCallen, D., Petersson, N.A., Rodgers, A., Pitarka, A., Miah, M., Petrone, F., Sjogreen, B., Abrahamson, N., Tang (2021), H., EQSIM - A Multidisciplinary Framework for Fault-to-Structure

Simulations on Exascale Computers Part I: Computational Models and Workflow, Earthquake Spectra, 37(2), 707-735.

Kenawy, M., McCallen, D., Pitarka, A., Rodgers, A. (2022), Seismic Risk to Buildings in the San Francisco Bay Area Predicted by Broadband Physics-based Hayward Fault Rupture Simulations, Proceedings of the12th National Conference on Earthquake Engineering, Salt Lake City, Utah,

McCallen, D., Petrone, F., Miah, M., Pitarka, A., Rodgers, A., Abrahamson, N. (2021), EQSIM - A Multidisciplinary Framework for Fault-to-Structure Earthquake Simulations on Exascale Computers Part II: Regional Simulations of Building Response, Earthquake Spectra, 37(2), 736-761.

Tang, H., Byna, N., Petersson, N.A., McCallen, D. (2021), Tuning Parallel Data Compression and I/O for Large-scale Earthquake Simulation, Proceedings of the 2021 IEEE International Conference on Big Data, 2992-2997.

Petrone, F., Abrahamson, N., McCallen, D., Pitarka, A., & Rodgers, A. (2021). Engineering evaluation of the EQSIM simulated ground-motion database: The San Francisco Bay Area Region, Earthquake Engineering & Structural Dynamics, 50(15), 3939–3961.

Kenawy, M., McCallen, D., & Pitarka, A. (2021). Variability of Near-Fault Seismic Risk to Reinforced Concrete Buildings Based on High-Resolution Physics-Based Ground Motion Simulations, Earthquake Engineering & Structural Dynamics, 50(6), 1713–1733.

Wu, S., Eckert, E., Huang, J., McCallen, D. (2021), The Influence of Rotational Components of Earthquake Ground Motions on Building Response, Proceedings of the 17th World Conference on Earthquake Engineering, Sendai Japan.

Petrone, F., Abrahamson, N., McCallen, D., & Miah, M. (2020). Validation of (Not-Historical) Large-Event Near-Fault Ground-Motion Simulations for Use in Civil Engineering Applications, Earthquake Engineering & Structural Dynamics, 50(1), 116–134.

McCallen, D., Petersson, N.A., Rodgers, A., Miah, M., Pitarka, A., Petrone, F., Tang, H. (2020), The Earthquake Simulation (EQSIM) Framework for Physics-Based Fault-to-Structure Simulations, Proceedings of the 17th World Conference on Earthquake Engineering, Sendai, Japan.

Rodgers, A., Petersson, N.A., Pitarka, A., McCallen, D., Sjogreen, B., Abrahamson, N. (2019), Broadband (0–5 Hz) Fully Deterministic 3D Ground-Motion Simulations of a Magnitude 7.0 Hayward Fault Earthquake: Comparison with Empirical Ground-Motion Models and 3D Path and Site Effects from Source Normalized Intensities, Seismological Research Letters, 90(3), 1268–1284.

Rodgers, A., Pitarka, A., McCallen D. (2019), The Effect of Fault Geometry and Minimum Shear Wavespeed on 3D Ground-Motion Simulations for an Mw 6.5 Hayward Fault Scenario Earthquake, San Francisco Bay Area, Northern California, Bulletin of the Seismological Society of America, 109(4), 1265-1281.

Wang, S., Petersson, N.A. (2019), Fourth Order Finite Difference Methods for the Wave Equation with Mesh Refinement Interfaces, SIAM Journal on Scientific Computing, 41(5), A3246-A3275.

Miah, M., Petrone, F., Wong, J., McCallen, D. (2018), Regional Scale Earthquake Risk Estimation based on Broadband Ground Motion Simulations, Proceedings of the 11th National Conference on Earthquake Engineering Integrating Science Engineering and Policy, Los Angeles, CA, 3354-3363.

Rodgers, A., Petersson, N.A., Pitarka, A., Miah, M., McCallen, D., Jeremic, B. (2018), HPC Simulations of Broadband Near-fault Ground Motions for Engineering Applications, Proceedings of the 11th National Conference on Earthquake Engineering Integrating Science Engineering and Policy, Los Angeles, CA, 3343-3353.

Rodgers, A., Petersson, N.A., Pitarka, A., McCallen, D. (2018), Simulation of Large Scenario and Recorded Moderate Earthquakes in the San Francisco Bay Area, Proceedings of the 11th National Conference on Earthquake Engineering Integrating Science Engineering and Policy, Los Angeles, CA, 4094-4103.

Rodgers, A., Pitarka, A., Petersson, N. A., Sjögreen, B., McCallen, D. (2018), Broadband (0–4 Hz) Ground Motions for a Magnitude 7.0 Hayward Fault Earthquake with Three-Dimensional Structure and Topography, Geophysical Research Letters, 45, 739-746.

Johansen, H., A. Rodgers, N. A. Petersson, D. McCallen, B. Sjogreen, M. Miah (2017), Toward Exascale Earthquake Ground Motion Simulations for Near-Fault Engineering Analysis, Computing in Science & Engineering, 19(5), 27–37.

(F) PI / CO-I Biographical Sketches

DAVID B. McCALLEN, PH.D.

Senior Scientist
Energy Geosciences Division
Lawrence Berkeley National Laboratory
Professor
Department of Civil and Environmental Engineering
University of Nevada, Reno

A. Professional Preparation

Ph.D. – Civil Engineering	University of California, Davis
Master of Science – Civil Engineering	University of California, Davis
Bachelor of Science – Civil Engineering	California State University, Chico

B. Appointments

2014 – Present	Senior Scientist; Energy Geosciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, California;
2018 – Present	Professor, Department of Civil and Environmental Engineering; University of Nevada, Reno;
2013 – 2018	Associate Vice President; Office of the National Laboratories, University of California Office of the President, Oakland, California;
2009 – 2013	Deputy Principle Associate Director, Global Security Directorate, Lawrence Livermore National Laboratory, Livermore California;
2004 – 2009	Program Leader for Nuclear Science and Engineering, Lawrence Livermore National Laboratory, University of California, Livermore, California;
2002 – 2004	Division Leader, Laser Science Engineering Division, Lawrence Livermore National Laboratory, University of California, Livermore, California;
2000 – 2002	Director, Engineering Research Center for Complex Distributed Systems, Lawrence Livermore National Laboratory, University of California, Livermore, California;
1982 – 2000	Structural Mechanics Engineer, Structural and Applied Mechanics Group, Lawrence Livermore National Laboratory, University of California, Livermore, California.

C. Selected Publications

D. McCallen, H. Tang, S. Wu, E. Eckert, J. Huang, N.A. Petersson (2021). Coupling of Regional Geophysics and Local Soil-Structure Models in the EQSIM Fault-to-Structure Earthquake Simulation Framework, The International Journal of High-Performance Computing Applications, https://doi.org/10.1177/10943420211019118.

D. McCallen, A. Petersson, A. Rodgers, A. Pitarka, M. Miah, F. Petrone, B. Sjogreen, N. Abrahamson, H. Tang (2021). EQSIM—A Multidisciplinary Framework for Fault-to-Structure Earthquake Simulations on Exascale Computers Part I: Computational Models and Workflow, Earthquake Spectra, 37(2), 707–735.

D. McCallen, F. Petrone, M. Miah, A. Pitarka, A. Rodgers, N. Abrahamson (2021). EQSIM—A Multidisciplinary Framework for Fault-to-Structure Earthquake Simulations on Exascale Computers, Part II: Regional Simulations of Building Response, Earthquake Spectra, 37(2), 736–761.

M. Kenawy, **D. McCallen,** A. Pitarka (2021). Variability of Near-Fault Seismic Risk to Reinforced Concrete Buildings Based on High-Resolution Physics-Based Ground Motion Simulations, Earthquake Engineering & Structural Dynamics, 50(6), 1713–1733.

F. Petrone, N. Abrahamson, **D. McCallen**, M. Miah (2021). Validation of (not-historical) Large Event Near-Fault Ground-Motion Simulations for use in Civil Engineering Applications, Earthquake Engineering and Structural Dynamics (50), 116-134.

F. Petrone, N. Abrahamson, **D. McCallen**, A. Pitarka, A. Rodgers (2021), Engineering Evaluation of the EQSIM Simulated Ground-Motion Database: The San Francisco Bay Area Region, Earthquake Engineering and Structural Dynamics, https://doi.org/10.1002/eqe.3540.

D. McCallen, F. Petrone (2019), An Optical Technique for Measuring Transient and Residual Interstory Drift as Seismic Structural Health Monitoring (S2HM) Observables, Chapter 11 for Seismic

Structural Health Monitoring; From Theory to Successful Practice, Limongelli and Celebi Editors, Springer Tracts in Civil Engineering.

A. Rodgers, A. Pitarka, **D. McCallen** (2019), The Effect of Fault Geometry and Minimum Shear Wavespeed on 3D Ground Motion Simulations for an Mw 6.5 Hayward Fault Scenario Earthquake, San Francisco Bay Area, Northern California, Bulletin of the Seismological Society of America, Vol. 109, No.4, pp 1265-1281.

A. Rodgers, A. Petersson, A. Pitarka, **D. McCallen**, B. Sjorgreen and N. Abrahamson (2018), Broadband (0-5 Hz) Fully Deterministic 3D Ground-Motion Simulations of a Magnitude 7.0 Hayward Fault Earthquake: Comparison with Empirical Ground-Motion Models and 3D Path and Site Effects from Source Normalized Intensities, Seismological Research Letters, 45, pages 739-747.

D. McCallen, F. Petrone, J. Coates, N. Repanich (2017), A Laser-Based Optical Sensor for Broad-Band Measurements of Building Earthquake Drift, Earthquake Engineering Research Institute, Earthquake Spectra, Vol. 22, No.4, pages 1573-1598

H. Johansen, A. Rodgers, A. Petersson, **D. McCallen**, B. Sjogreen, M. Miah, (2017), Toward Exascale Earthquake Ground Motion Simulations for Near-Fault Engineering Analysis, Computing in Science and Engineering, IEEE, Vol. 19, No. 5.

J. Abell, N. .Orbovic, **D. McCallen**, B. Jeremic (2017), Earthquake Soil-Structure Interaction of Nuclear Power Plants, Differences in Response to 3-D, 3x1-D, and 1-D Excitations, Earthquake Engineering and Structural Dynamics, Vol. 47, pages 1478-1495.

D. McCallen, A. Astaneh-Asl, S. Larsen, L. Hutchings (2009), The Response of Long-Span Bridges to Near- Fault Earthquake Ground Motions, Proceedings of the ASCE Technical Conference on Life Line Earthquake Engineering, Oakland, California, June 2009.

D. Synergistic Activities

Research Computational simulation of the nonlinear response of building, bridge and energy systems, regional-scale simulation of ground motions and infrastructure response, development of advanced sensors for monitoring the seismic response of infrastructure

Teaching Numerical Methods in Structural and Earthquake Engineering
Advanced Structural Dynamics

Prof. Activity

- Member, Science and Technology Committee, Los Alamos National Laboratory
- Chair, External Advisory Board, NSF SimCenter, University of California, Berkeley
- Member, Nuclear Engineering External Advisory Board, Department of Nuclear Engineering, University of California, Berkeley
- Peer reviewer – Journal of Earthquake Engineering, ASCE Journal of Engineering Mechanics, ASCE Journal of Structural Engineering, International Journal of Solids and Structures, Bulletin of the Seismological Society of America, International Journal for Structural Engineering and Mechanics, Earthquake Engineering and Structural Dynamics, Nuclear Engineering and Design

E. Major Research Projects

(i) Project Leader, A Modern Computational Framework for the Nonlinear Seismic Analysis of Nuclear Systems, U.S. DOE, 2015-2022

(ii) Principal Investigator, EQSIM - High Performance Multidisciplinary Simulations for Regional Scale Earthquake Hazard and Risk Assessments, U.S. DOE Office of Science Exascale Computing Project, 2017-2023

(iii) Principal Investigator, Regional-Scale Simulation of Earthquake Hazard and Risk to Critical Energy Systems, U.S. DOE Office of Cybersecurity, Energy Security and Emergency Response, 2021-2022

Arben Pitarka

Deputy Seismology Group Leader
Atmospheric, Earth and Energy Division
Physical and Life Sciences Directorate
Lawrence Livermore National Laboratory (LLNL),
Livermore, CA 94550
(email) pitarka1@llnl.gov
(phone) 925-424-3010

Professional Preparation

Ph.D. Seismology	1997	Kyoto University, Kyoto, Japan
Ph.D. Engineering Seismology	1993	Tirana University, Tirana, Albania,
M.S. Engineering Seismology	1991	Tirana University, Tirana, Albania,
B.S. Geophysics	1984	IPG, University of Strasbourg, France

Appointments

2011-present	Seismologist, LLNL, Livermore, CA
2010-2011	Principal Scientist, BAE Systems/AFTAC, Patrick AFB, FL
2009-2010	Senior Scientist, QTSI, Cocoa Beach, FL
1996-2009	Principal Seismologist, URS Corporation, Pasadena, CA
1984-1993	Project Seismologist, Seismological Institute, Albania

Research Interests and Experience

Extensive experience in computational seismology. Research interests include development of numerical techniques for modeling rupture dynamics and elastic wave propagation using 3D Finite Difference methods. Has made extensive contribution to development of physics-based earthquake rupture models and development of hybrid methods for simulating earthquake broadband ground motion with engineering applications. In the field of nuclear explosion monitoring has made technical contributions to development and testing numerical techniques for coupling hydrodynamic modeling of the explosion source and elastic wave propagation.

List of Selected Publications

Pitarka, A., R. Graves, K. Irikura, K. Miyakoshi, C. Wu, H. Kawase, A. Rodgers, and D. Mccallen (2021), Refinements to the Graves-Pitarka Kinematic Rupture Generator, Including a Dynamically Consistent Slip-Rate Function, Applied to the 2019 Mw 7.1 Ridgecrest Earthquake, *Bull. Seismo. Soc. Am.,* 1-20, doi: 10.1785/0120210138.

Pitarka, A., A. Akinci, P. DeGori, and M. Buttinelli (2021). Deterministic 3D Ground-Motion Simulations (0-5Hz) and Surface Topography Effects of the 30 October 2016 Mw 6.5 Norcia, Italy Earthquake, *Bull. Seismo. Soc. Am.,1-25,* doi: 10.1785/0120210133

Pitarka, A., and R. Mellors (2021), Using Dense Array Waveform Correlations to Build a Velocity Model With Stochastic Variability, *Bull. Seismo. Soc. Am.,*1–21, doi: 10.1785/0120200206

Dr. Rie Nakata

Also known as Dr. Rie Kamei
Lawrence Berkeley National Laboratory
1 Cyclotron Road, Berkley, CA, 94720
T: +1 405 406 4172; E: rnakata@lbl.gov

EDUCATION

2008-2012	Ph.D. in Geophysics Department of Earth Sciences, The University of Western Ontario
2003-2005	M. Eng. Department of Civil and Earth Resources Engineering, Kyoto University
1999-2003	B. Eng. Undergraduate School of Global Engineering, Kyoto University,

WORK EXPERIENCE

2021-	Research Scientist, Lawrence Berkeley National Laboratory
2019-	Assistant Professor, Earthquake Research Institute, University of Tokyo
2019	Visiting Scientist, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology
2017-2018	Visiting Researcher, ConocoPhillips School of Geology and Geophysics The University of Oklahoma
2013-2018	Research Fellow, Centre for Energy Geoscience, School of Earth and Environment, The University of Western Australia
2013	Post-doctoral fellow, Department of Earth Sciences, The University of Western Ontario
2008-2013	Research Assistant Department of Earth Sciences, The University of Western Ontario
2002-2005	Research Assistant Kyoto University

PUBLICATIONS

1. **Nakata, R**., Jang, U.G., Lumley, D., Mouri, T., Nakatsukasa, M., Takanashi, M., Kato, A., 2022, Seismic time-lapse monitoring of near-surface microbubble injection by full waveform inversion, Submitted to Geophysical Research Letters.
2. Nakata, N., **Nakata, R**., Kato, A., Xue, Z., 2022, Cross-well seismic modeling with inclusion of tube waves and tube-wave-related arrivals, Submitted to Journal of Geophysical International
3. Glubokovskikh, S., Saygin, E., Shapiro, S., Gurevich, B., Isaenkov, R., Lumley, D., **Nakata, R.**, Drew, J., Pevzner, R., 2022, A small CO2 leakage may induce seismicity on a sub-seismic fault in a good-porosity clastic saline aquifer, Geophysical Research Letters (accepted)
4. **Nakata, R.**, Lumley, D., Hampson, G., Nihei, K., Nakata, N., 2020, Waveform-based estimation of Q and scattering properties for zero-offset VSP data, Geophysics, 85, R365-R379.
5. Tsuji, T., Minato, S., **Kamei, R.,** Tsuru, T., Kimura, G. 2017 Detachment fault from fully coupled plate interface; Insights from seismic profiles at the hypocenter of the 2016 Off-Mie earthquake in the Nankai Trough, submitted to Earth and Planetary Science Letters
6. **Kamei, R.,** Lumley, D. 2017 Full waveform inversion of repeating seismic events to estimate time-lapse velocity changes, Geophysical Journal International, 209(2), 1239-1264.
7. Lumley, D., King. A., Pevzner, R., Bona, A., Dautriat, J., Esteban, L., Hauser, T., Hoskin, T., Issa, N., **Kamei, R.**, Langhi, L., Miyoshi, T., Mueller, T., Potter, T., Sarout, J., Siggins, T., Shragge, J., Tertyshnikov, K., Urosevic, M., 2016, Feasibility and Design for Passive Seismic Monitoring at the SW Hub CO2 Geosequestration Site, Report for Australian National Low Emissions Coal Research & Development Project 7-0212-0203 (available from http://anlecrd.com.au/wp-content/uploads/2016/08/ANLEC-Project-7-0212-0203_Revised-Final-Report-24.06.2016.pdf).

Houjun Tang

EDUCATION

North Carolina State University – Raleigh, NC *2012 – 2016*
Ph.D. in Computer Science

Shenzhen University – Shenzhen, China *2008 - 2012*
B.Eng. in Computer Science and Technology

RESEARCH EXPERIENCE

Computational Research Division, Berkeley Lab
Computer Research Scientist *2019 - Current*

Computational Research Division, Berkeley Lab
Postdoctoral Researcher *2016 - 2019*

Computer Science Department, NC State University
Research Assistant *2012 - 2016*

PUBLICATIONS

1. Xiaoxia Zhang, Degang Chen, Hong Yu, Guoyin Wang, **Houjun Tang**, and Kesheng Wu. "Improving nonnegative matrix factorization with advanced graph regularization." Information Sciences 597 (2022): 125-143.
2. **Houjun Tang,** Bing Xie, Suren Byna, Phillip Carns, Quincey Koziol, Sudarsun Kannan, Jay Lofstead, and Sarp Oral, "SCTuner: An Auto-tuner Addressing Dynamic I/O Needs on Supercomputer I/O Sub-systems ", PDSW, in conjunction with SC21, 2021
3. **Houjun Tang**, Suren Byna, N. A. Petersson and David McCallen, "Tuning Parallel Data Compression and I/O for Large-scale Earthquake Simulation," IWBDR, in conjunction with BigData 21, 2021
4. Jean Luca Bez, **Houjun Tang**, Bing Xie, David Williams-Young, Rob Latham, Rob Ross, Sarp Oral, and Suren Byna, "I/O Bottleneck Detection and Tuning: Connecting the Dots using Interactive Log Analysis", PDSW, in conjunction with SC21, 2021
5. **Houjun Tang**, Quincey Koziol, John Ravi, and Suren Byna, "Transparent Asynchronous Parallel I/O using Background Threads", IEEE TPDS - Special Section on Innovative R&D toward the Exascale Era, 2021
6. David McCallen, **Houjun Tang,** Suiwen Wu, Eric Eckert, Junfei Huang, N Anders Petersson, "Coupling of regional geophysics and local soil-structure models in the EQSIM fault-to-structure earthquake simulation framework", The International Journal of High Performance Computing Applications, 2021
7. Bing Xie, **Houjun Tang**, Suren Byna, Jesse Hanley, Quincey Koziol, Tonglin Li, Sarp Oral, "Battle of the Defaults: Extracting Performance Characteristics of HDF5 under Production Load", CCGrid 2021
8. **Houjun Tang**, Suren Byna, Bin Dong, and Quincey Koziol "Parallel Query Service for Object-centric Data Management Systems", The 6th IEEE International Workshop on High-Performance Big Data and Cloud Computing (HPBDC) 2020, in conjunction with IPDPS 2020.
9. David McCallen, Anders Petersson, Arthur Rodgers, Arben Pitarka, Mamun Miah, Floriana Petrone, Bjorn Sjogreen, Norman Abrahamson, **Houjun Tang,** "EQSIM-A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers part I:

Computational models and workflow", Earthquake Spectra, December 1, 2020, 37:707-735

10. Suren Byna, M. Scot Breitenfeld, Bin Dong, Quincey Koziol, Elena Pourmal, Dana Robinson, Jerome Soumagne, **Houjun Tang**, Venkatram Vishwanath, and Richard Warren, "ExaHDF5: Delivering Efficient Parallel I/O on Exascale Computing Systems", Journal of Computer Science and Technology 2020, 35(1): 145-160
11. **Houjun Tang**, Suren Byna, Stephen Bailey, Zarija Lukic, Jialin Liu,Quincey Koziol, Bin Dong, and Teng Wang, “Tuning Object-centric Data Management Systems for Large Scale Scientific Applications”, HiPC 2019
12. Bin Dong, Kesheng Wu, Suren Byna, **Houjun Tang**, “SLOPE: Structural Locality-aware Programming Model for Composing Array Data Analysis”, ISC 2019
13. Jingqing Mu, Jerome Soumagne, Suren Byna, Quincey Koziol, **Houjun Tang**, Richard Warren, “Interfacing HDF5 with A Scalable Object-centric Storage System on Hierarchical Storage”, CUG 2019
14. Tonglin Li, Quincey Koziol, **Houjun Tang**, Jialin Liu, Suren Byna, “I/O Performance Analysis of Science Applications Using HDF5 File-level Provenance”, CUG 2019
15. **Houjun Tang**, Suren Byna, François Tessier, Teng Wang, et al, “Toward Scalable and Asynchronous Object-centric Data Management for HPC”, CCGRID 2018
16. Bin Dong, Teng Wang, **Houjun Tang**, Quincey Koziol, Kesheng Wu, Suren Byna, ARCHIE: Data Analysis Acceleration with Array Caching in Hierarchical Storage, Big Data 2018
17. Jialin Liu, Quincey Koziol, Gregory Butler, Neil Fortner, Mohamad Chaarawi, **Houjun Tang**, et al. “Evaluation of HPC Application I/O on Object Storage Systems”, PDSW-DISCS 2018
18. Wei Zhang, **Houjun Tang**, Suren Byna, Yong Che, “DART: Distributed Adaptive Radix Tree for Efficient Affix-based Keyword Search on HPC Systems”, PACT 2018
19. Jingqing Mu, Jerome Soumagne, **Houjun Tang**, Suren Byna, Quincey Koziol, Richard Warren, “A Transparent Server-Managed Object Storage System for HPC”, CLUSTER 2018
20. Teng Wang, Suren Byna, Bin Dong, **Houjun Tang**, “UniviStor: Integrated Hierarchical and Distributed Storage for HPC”, CLUSTER 2018
21. **Houjun Tang**, Suren Byna, Bin Dong, Jialin Liu, Quincey Koziol, “SoMeta: Scalable Object-centric Metadata Management for Hgh Performance Computing”, CLUSTER 2017
22. Jialin Liu, Quincey Koziol, **Houjun Tang**, Francois Tessier, et al, “Understanding the I/O Performance Gap Between Cori KNL and Haswell”, CUG 2017
23. Wenzhao Zhang, **Houjun Tang**, Stephen Ranshous, Suren Byna, Daniel Martín, Kesheng Wu, et al, “Exploring Memory Hierarchy and Network Topology for Runtime AMR Data Sharing Across Scientific Applications”, BigData 2016
24. **Houjun Tang**, Suren Byna, Steven Harenberg, Xiaocheng Zou, Wenzhao Zhang, et al, "Usage Pattern-Driven Dynamic Data Layout Reorganization", CCGrid 2016
25. Wenzhao Zhang, **Houjun Tang**, Steven Harenberg, Suren Byna, et al, "AMRZone: A Runtime AMR Data Sharing Framework for Scientific Applications", CCGrid 2016
26. Wenzhao Zhang, **Houjun Tang**, Xiaocheng Zou, Steven Harenberg, Qing Liu, Scott Klasky, Nagiza Samatova, “Exploring Memory Hierarchy to Improve Scientific Data Read Performance”, CLUSTER 2015
27. David Boyuka, **Houjun Tang**, Kushal Bansal, Xiaocheng Zou, Scott Klasky, Nagiza Samatova, “The Hyperdyadic Index and Generalized Ondexing and Query with PIQUE”, SSDBM 2015
28. Xiaocheng Zou, David Boyuka, Dhara Desai, Daniel Martin, Suren Byna, Kesheng Wu, Kushal Bansal, Bin Dong, Wenzhao Zhang, **Houjun Tang**, et al, "AMR-aware In Situ Indexing and Scalable Querying", The 24th High Performance Computing Symposium (HPC 2016), **HPC 2016 Best Paper Award**

COMPUTATIONAL SCIENTIST

Ramesh Pankajakshan

PROFILE

Experienced research scientist with over two decades of experience with large scale simulations on leadership class HPC systems.

EXPERIENCE

COMP. SCIENTIST, LAWRENCE LIVERMORE NATL. LABS; LIVERMORE, CA – 2016-PRESENT

Porting and performance optimization of seismic code to GPUs(Nvidia V100, AMD MI250X) and other current architectures(Fujitsu A64X)

Ran large-scale seismic simulations on quarter of Summit(#2 on Top 500 list) with a 100X improvement in Figure of Merit w.r.t the best CPU platform

Initial port of BoomerAMG(Algebraic MultiGrid) in the Hypre math library to GPUs

Consultant for internal and external code porting and optimization

Lab lead for Math libraries and Machine Learning Working Group for the CORAL2 Exascale procurement. Member of Tools and Message Passing Working Groups

Member of benchmarking and acceptance teams for Livermore Computing HPC procurements

RES. PROFESSOR, UNIVERSITY OF TENNESSEE; CHATTANOOGA, TN – 2002-2015

Lead development efforts for faster than realtime plume predictions on laptops, coupled agent-based modeling and fluid dynamics simulations for urban evacuation studies, design optimization for particulate flows and drag reduction devices for Class 8 trucks

PDF/ASST. RES. PROF/ASSOC. RES. PROF; MISSISSIPPI STATE UNIVERSITY, STARKVILLE, MS – 1997-2002

Lead developer for the Office of Naval Research funded UNCLE incompressible flow solver used for maneuvering submarine and ship simulations by the Naval Surface Warfare Center Carderock Division

EDUCATION

Mississippi State University, Starkville, MS – Ph.D. Computational Engg. 1997

Mississippi State University, Starkville, MS – M.S. Mechanical Engineering 1993

Indian Institute of Technology, Madras,India –- B.Tech Mechanical Engineering 1991

SKILLS

- Extensive experience with C/C++ and Python
- Extensive knowledge of MPI and OpenMP

- Regular user of profiling and debugging tools for GPUs(Nsight Compute/Systems, rocprof, Totalview, valgrind4hpc), build systems and revision control
- Visualization tools: FieldView, Tecplot, Paraview, POV-Ray
- Experienced in writing proposals and making presentations to prospective clients

SELECTED PUBLICATIONS

Pankajakshan, R., Lin, P-H., and Sjögreen, B. “Porting a 3D Seismic Modeling Code(SW4) to CORAL machines,” IBM Journal of Research and Development, Dec. 2019

Epstein, J. M., Pankajakshan, R., and Hammond, R. A., “Combining Computational Fluid Dynamics and Agent-Based Modeling: A New Approach to Evacuation Planning,” PLoS ONE, May 2011.

Pankajakshan, R., Mitchell, B.J., and Taylor, L.K., "Simulation of Unsteady Two-Phase Flows Using a Parallel Eulerian-Lagrangian Approach," Computers & Fluids, Volume 41, Issue 1, February 2011, Pages 20-26

Pankajakshan, R., Remotigue, M.G., Taylor, L.K., Jiang, M. Briley, W.R., and Whitfield, D.L., “Validation of Control-Surface Induced Submarine Maneuvering Simulations using UNCLE,” 24th Symposium on Naval Hydrodynamics, Fukuoka, Japan, July 8-13, 2002.

Newman III, J.C., Pankajakshan, R., Whitfield, D.L., and Taylor, L.K., “Computational Design Using RANS,” 24th Symposium on Naval Hydrodynamics, Fukuoka, Japan, July 8-13, 2002.

INVITED PRESENTATION

Pankajakshan R., “CFD For Tractor-Trailer Aerodynamics: Current Trends and Future Outlook,” presented to the Committee on Assessment of Technologies and Approaches for Reducing Fuel Consumption of Medium-and Heavy-Duty Vehicles, Phase 2, Transportation Research Board, The National Academies, Washington DC. Dec. 2014

925 577 5947 RAMESHES@ICLOUD.COM

Software Applications and Dependent Packages

In the EQSIM framework a number of software applications are connected to assemble the overall workflow as indicated in Figure 8 and Figure 9. These include:

- The SW4 (*Seismic Waves 4th Order*) explicit finite-difference application for seismic wave propagation[2,3,4,5];
- The NEVADA implicit nonlinear finite element application for the inelastic analysis of building systems;
- The OPENSEEs implicit nonlinear finite element application for the inelastic analysis of soil and structure systems.

For the ground motion simulations that will be executed on FRONTIER within the proposed INCITE resource request, SW4 will be utilized with subsequent transfer of ground motion data to PERLMUTTER for performing infrastructure response simulations. SW4 is built upon a strong theoretical foundation provided by the summation by parts principle which guides derivation of high order-of-accuracy finite-difference methods such that resulting time integration scheme is provably stable without requiring artificial dissipation or ad-hoc filtering. SW4 has been a major focus of the ECP application development and both performance enhancing features and new and improved physics models have been incorporated. The critical features residing within SW4 are shown in Figure 15. In addition, SW4 has been thoroughly exercised on SUMMIT and tested on CRUSHER and is ready to transition to the full FRONTIER platform.

Additional applications which SW4 utilizes on SUMMIT and CRUSHER include:

- RAJA C++ routines to enable implementation across multiple GPU platforms. RAJA systematically encapsulates platform-specific code to enable applications to be highly portable across diverse platforms with minimal code disruption. RAJA is supported by the ECP software technology program, is available on SUMMIT and CRUSHER and will be available on FRONTIER;
- Hierarchical Data Format (HDF5) data containers for efficient parallel I/O, the HDF5 library is supported by the ECP software technology program and are available on SUMMIT and CRUSHER;
- The ZFP compression library which works seamlessly and transparently with the HDF5 I/O library. ZFP is also supported by the ECP software technology program, is available on SUMMIT and CRUSHER and will be available on FRONTIER;

---

[2] Petersson, N.A., and Sjogreen, B., 2012. Stable and Efficient Modeling of Anelastic Attenuation in Seismic Wave Propagation, *Comm. Comput. Phys.*, **12**, 193–225.
[3] Petersson, N.A, and Sjogreen, B., 2015. Wave Propagation in Anisotropic Elastic Materials and Curvilinear Coordinates Using a Summation-by-Parts Finite-Difference Method, *J. Comput. Phys.*, **299**, 820–841.
[4] Wang, S., Petersson, N.A. (2019) Fourth order finite difference methods for the wave equation with mesh refinement interfaces, SIAM J. Sci. Comput. 41(5), A3246-A3275.
[5] Zhang, L., Wang, S., Petersson, N.A. (2020) Elastic Wave Propagation in Curvilinear Coordinates with Mesh Refinement Interfaces by a Fourth Order Finite Difference Method, SIAM J. Sci. Comp. 43(2), A1472-A1496.

- Additional dependencies include MPI, HIP, ROCm, and UMPIRE.

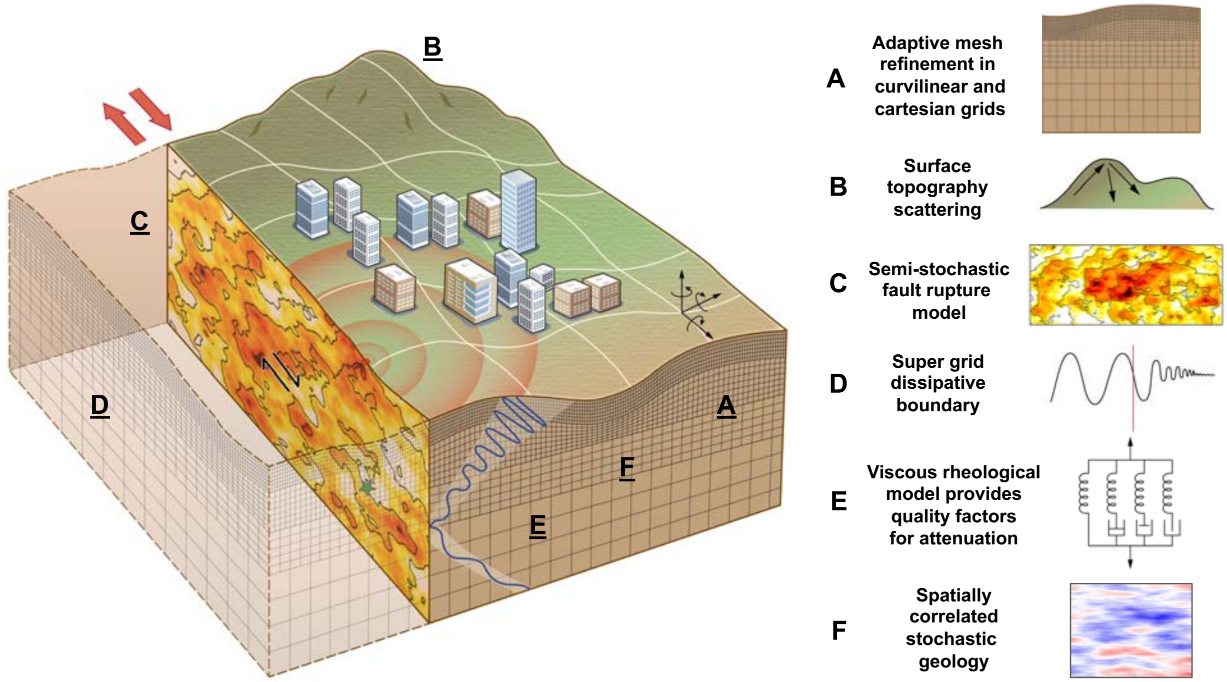


Figure 15. Capabilities implemented in the SW4 application for seismic wave modeling.

Proprietary and Sensitive Information

There is no proprietary, sensitive or export-controlled information associated with this proposed work.

Potential Subject Matter Expert Reviewers

Dr. Ertugrul Taciroglu
Professor and Chair
Department of Civil and Environmental Engineering
University of California Los Angeles
etacir@g.ucla.edu

Dr. Brendon Bradley
Professor of Earthquake Engineering
Director of QuakeCoRE
University of Canterbury
Christchurch, New Zealand
brendon.bradley@canterbury.ac.nz

Dr. Ruth Harris
Research Scientist
United States Geologic Survey
Mountain View, California
harris@usgs.gov

Section 6: Software Applications and Packages

Question #1

Please list any software packages used by the project, and indicate if they are on open source or export controlled.

Application Packages

Package Name

SW4

Indicate whether Open Source or Export Controlled.

Open Source

Package Name

RAJA

Indicate whether Open Source or Export Controlled.

Open Source

Package Name

ZFP

Indicate whether Open Source or Export Controlled.

Open Source

Section 7: Wrap-Up Questions

Question #1

National Security Decision Directive (NSDD) 189 defines Fundamental Research as "basic and applied research in science and engineering, the results of which ordinarily are published and shared broadly within the scientific community, as distinguished from proprietary research and from industrial development, design, production, and product utilization, the results of which ordinarily are restricted for proprietary or national security reasons." Publicly Available Information is defined as information *obtainable free of charge (other than minor shipping or copying fees) and without restriction, which is available via the internet, journal publications, textbooks, articles, newspapers, magazines, etc.*

The INCITE program distinguishes between the generation of proprietary information (deemed a proprietary project) and the use of proprietary information as input. In the latter, the project may be considered as Fundamental Research or nonproprietary under the terms of the nonproprietary user agreement. Proprietary information, including computer codes and data, brought into the LCF for use by the project - but not for generation of new intellectual property, etc., using the facility resources - may be protected under a nonproprietary user agreement.

Proprietary Information

Are the proposed project and its intended outcome considered Fundamental Research or Publicly Available Information?

Yes

Will the proposed project use proprietary information, intellectual property, or licensing?

No

Will the proposed project generate proprietary information, intellectual property, or licensing as the result of the work being proposed?

If the response is Yes, please contact the INCITE manager, INCITE@doeleadershipcomputing.org, prior to submittal to discuss the INCITE policy on proprietary work.

No

Question #2

The following questions are provided to determine whether research associated with an INCITE proposal may be export controlled. Responding to these questions can facilitate - but not substitute for - any export control review required for this proposal.

PIs are responsible for knowing whether their project uses or generates sensitive or restricted information. Department of Energy systems contain only data related to scientific research and do not contain personally identifiable information. Therefore, you should answer "Yes" if your project uses or generates data that fall under the Privacy Act of 1974 U.S.C. 552a. Use of high-performance computing resources to store, manipulate, or remotely access any national security information is prohibited. This includes, but is not limited to, classified information, unclassified controlled nuclear information (UCNI); naval nuclear propulsion information (NNPI); and the design or development of nuclear, biological, or chemical weapons or of any weapons of mass destruction. For more information contact the Office of Domestic and International Energy Policy, Department of Energy, Washington DC 20585, 202-586-9211.

Export Control

Does this project use or generate sensitive or restricted information?

No

Does the proposed project involve any of the following areas?

i. Military, space craft, satellites, missiles, and associated hardware, software or technical data

ii. Nuclear reactors and components, nuclear material enrichment equipment, components (Trigger List) and associated hardware, software or technical data

iii. Encryption above 128 bit software (source and object code)

iv. Weapons of mass destruction or their precursors (nuclear, chemical and biological)

No

Does the proposed project involve International Traffic in Arms Regulations (ITAR)?

No

Question #3

The following questions deal with health data. PIs are responsible for knowing if their project uses any health data and if that data is protected. Note that certain health data may fall both within these questions as well as be considered sensitive as per question #2. Questions regarding these answers to these questions should be directed to the centers or program manager prior to submission.

Health Data

Will this project use health data?

No

Will this project use human health data?

No

Will this project use Protected Health Information (PHI)?

No

Question #4

The PI and designated Project Manager agree to the following:

Monitor Agreement

I certify that the information provided herein contains no proprietary or export control material and is correct to the best of my knowledge.

Yes

I agree to provide periodic updates of research accomplishments and to acknowledge INCITE and the LCF in publications resulting from an INCITE award.

Yes

I agree to monitor the usage associated with an INCITE award to ensure that usage is only for the project being described herein and that all U. S. Export Controls are complied with.

Yes

I understand that the INCITE program reserves the right to periodically redistribute allocations from underutilized projects.

Yes

Section 8: Outreach and Suggested Reviewers

Question #1

By what sources (colleagues, web sites, email notices, other) have you heard about the INCITE program? This information will help refine our outreach efforts.

Outreach

Question #2

Suggested Reviewers

Suggest names of individuals who would be particularly suited to assess the proposed research.

Dr. Ertugrul Taciroglu

Professor and Chair

Department of Civil and Environmental Engineering

University of California Los Angeles

etacir@g.ucla.edu

Dr. Brendon Bradley

Professor of Earthquake Engineering

Director of QuakeCoRE

University of Canterbury

Christchurch, New Zealand

brendon.bradley@canterbury.ac.nz

Dr. Ruth Harris

Research Scientist

United States Geologic Survey

Mountain View, California

harris@usgs.gov

Section 9: Testbed Resources

Question #1

The ALCF and OLCF have test bed resources for new technologies, details below. If you would like access to these resources to support the work in this proposal, please provide the information below. (1 Page Limit)

The OLCF Quantum Computing User Program is designed to enable research by providing a broad spectrum of user access to the best available quantum computing systems, evaluate technology by *monitoring the breadth and performance of early quantum computing applications, and Engage the quantum computing community and support the growth of the quantum information science ecosystems. More information can be found here: https://www.olcf.ornl.gov/olcf-resources/compute-systems/quantum-computing-user-program/quantum-computing-user-support-documentation.*

The ALCF AI Testbed provides access to next-generation of AI-accelerator machines to enable evaluation of both hardware and workflows. Current hardware available includes Cerebras C-2, Graphcore MK1, Groq, Habana Gaudi, and SambaNova Dataflow. New hardware is regularly acquired as it becomes available. Up to date information can be found here: https://www.alcf.anl.gov/alcf-ai-testbed.

Describe the experiments you would be interested in performing, resources required, and their relationship to the current proposal. Please note, these are smaller experimental resources and a large amount of resources are not available. Instead, these resources are to explore the possibilities for these technologies might innovate future work. This request does not contribute to the 15-page proposal limit.

We do not have a need for test bed resources.pdf
The attachment is on the following page.

We do not have a need for test bed resources