

# Sustainable Computing for Earthquake Forecast Testing Centers

Philip Maechling

Statewide California Earthquake Center

**September 8-11, 2024**

Hilton Palm Springs, California

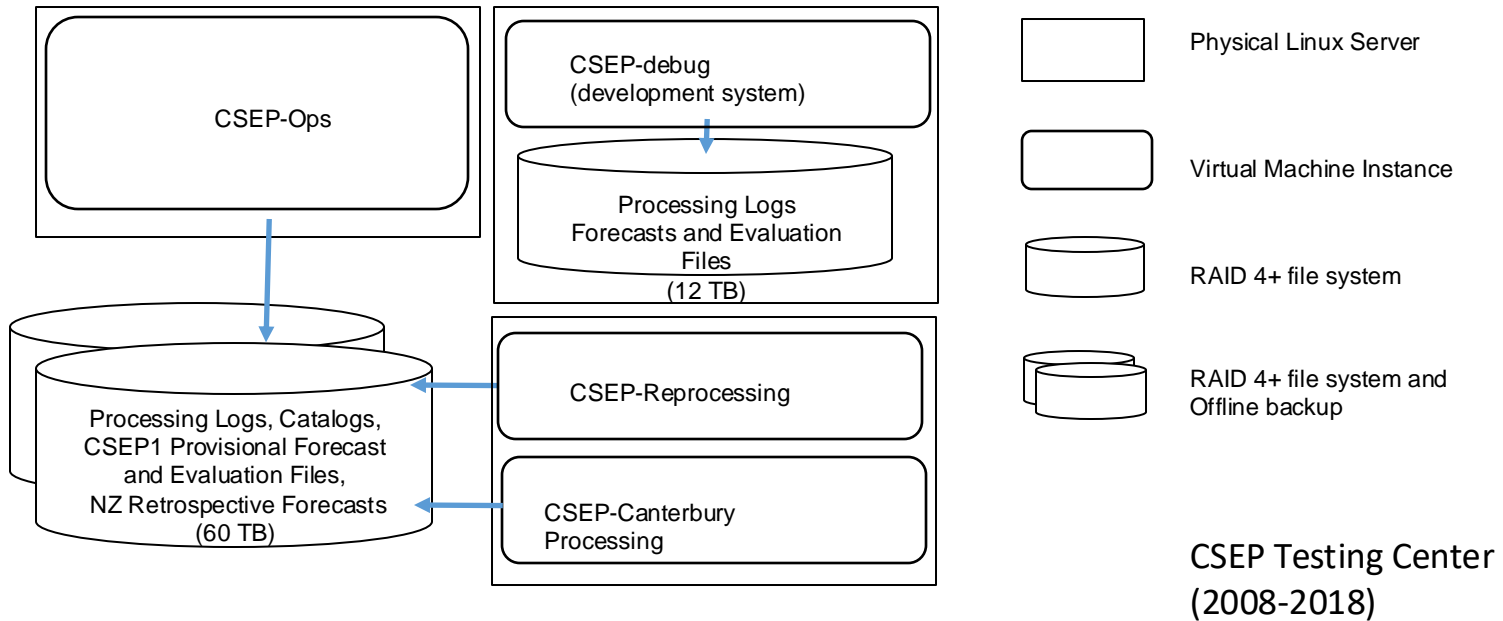


# Testing Essential In Scientific Process

The SCEC research collaboration is working to develop a predictive understanding of earthquake processes.

Community-based testing programs have advanced several areas:

- 3D Motion in Basins
- Dynamic Rupture Modelers
- Transient Detection
- ShakeAlert algorithm performance
- Collaboratory for the Study of Earthquake Predictability
- Broadband Platform Ground Motion Model Comparison



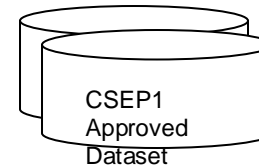
CSEP-Ops is current operational testing center with 60TB file system, with backup.

CSEP-debug is the current test system with CSEP data through 2017.

CSEP-Reprocessing is new image used for any needed CSEP1 reprocessing.

CSEP-Canterbury Processing is new image used for any needed NZ reprocessing.

With this configuration, we then stop using other CSEP VMs and computer systems. VMs shut down. Data moved onto offline backups. Computers repurposed: CSEP-debug2, CSEP Testing, CSEP Certification, Northridge, CSEP-generic, CSEP-X. Once CSEP 1 reprocessing has been completed, CSEP-Ops becomes CSEP-Ops on demand processing system until all supported models are moved to CSEP 2. CSEPTesting.org web site, CSEP trac site, CSEP SVN server, and CSEP Publishing are operated until CSEP1 data is online.



Release the CSEP1 dataset through a data hosting service that requires us to format the data and metadata but does not require us to operate any computers.



# Testing Essential In Scientific Process

One of the lessons learned is the challenge of operating and maintaining computer servers, and data storage.

Most new research projects will avoid purchasing physical servers and storage and will outsource computing and storage to academic service centers (e.g. USC Center for Advance Research Computing (CARC)), open-science vendors (e.g. SDSC SGX3), or commercial cloud vendors (e.g. Amazon, Google, Microsoft).

Sustainable future Earthquake Forecast Testing Centers will almost definitely use open-source software repos (e.g. GitHub), persistent scientific data storage and identifies (e.g. Zenodo DOIs), and and cloud-based computing technologies such as Docker, Kubernetes, Kafka, and others.



# Testing Essential In Scientific Process

Original Goals of CSEP System were defined as:

- Transparency
- Reproducibility
- Controlled Environment
- Comparability

A CSEP group is developing a cloud-compatible testing structure that can meet these goals without a dedicated testing environment using cloud-based technologies:

<https://github.com/cseptesting/floatcsep>

## CSEP Floating Experiments



An application to deploy reproducible and prospective experiments of earthquake forecasting

docs passing tests passing pypi v0.1.4 codecov 77% DOI 10.5281/zenodo.7953817

- Set up a testing **experiment** for your earthquake forecasts using authoritative data sources and benchmarks.
- **Encapsulate** the complete experiment's definition and rules in a couple of lines.
- **Reproduce, reuse, and share** forecasting experiments from you, other researchers and institutions.

Developer and scientific team:

- Pablo Iturrieta - Deutsches GeoForschungsZentrum (GFZ), Germany - [pciturri@gfz-potsdam.de](mailto:pciturri@gfz-potsdam.de)
- William Savran - University of Nevada, USA
- Asim Khawaja - GFZ, Germany
- Toño Bayona - University of Bristol, United Kingdom
- Danijel Schorlemmer - GFZ, Germany
- Fabio Silva - Southern California Earthquake Center (SCEC), USA
- Phil Maechling - SCEC, USA
- Max Werner - University of Bristol, United Kingdom

# *Quakeworx: A Collaborative CSSI Proposal*

PI: Yehuda Ben-Zion (USC/SCEC)

PI: Alice-Agnes Gabriel (UCSD/Scripps Institution of Oceanography)

PI: Amit Chourasia (UCSD/San Diego Supercomputer Center)

PI: Ahmed Elbanna (UIUC/UIUC Civil Engineering)

Project Start: September 1, 2023

Project End: August 31, 2028

Project Funding: Awarded 500k/year to USC, UCSD, UIUC

# Quakeworx Public Project Website ([quakeworx.org](http://quakeworx.org))



A science gateway for seismic simulations

[Publications](#) [About us](#)

## Curated apps

Next generation earthquake simulators and Machine Learning apps such as Moose-FARM, QuakeNN, RSQSim, SeisSol, Tandem and others will be available for anyone to use.

## Curated data

Provides input configurations and output results for select simulations with curated apps.

## Curated pipelines

Preset simulation pipelines for select earthquake scenarios.

## Integrated computing

Interactive and batch computation resources will be integrated.

## Collaborative workspace

Projects and groups will enable and support collaboration.

## FAIR Publishing

Users can publish FAIR products such as executable software applications, reports, data and other content with DOIs.

## Community contribution

Users will be able to contribute apps, data and other material.

[Get project updates / early access invite](#)

## Upcoming event

CyberTraining for Seismology Workshop at the 2024 SCEC annual meeting.

## Work in progress

Moose-Farm, SeisSol, Tandem and UCERF3-ETAS applications are currently being integrated with the gateway.

## Reduce barriers to

Access, expertise, software, compute and data resources.

## Catalyze & empower

Community wide reuse & sharing of research products.

## Scale

Your research, education and workforce development efforts.

## Advance Science

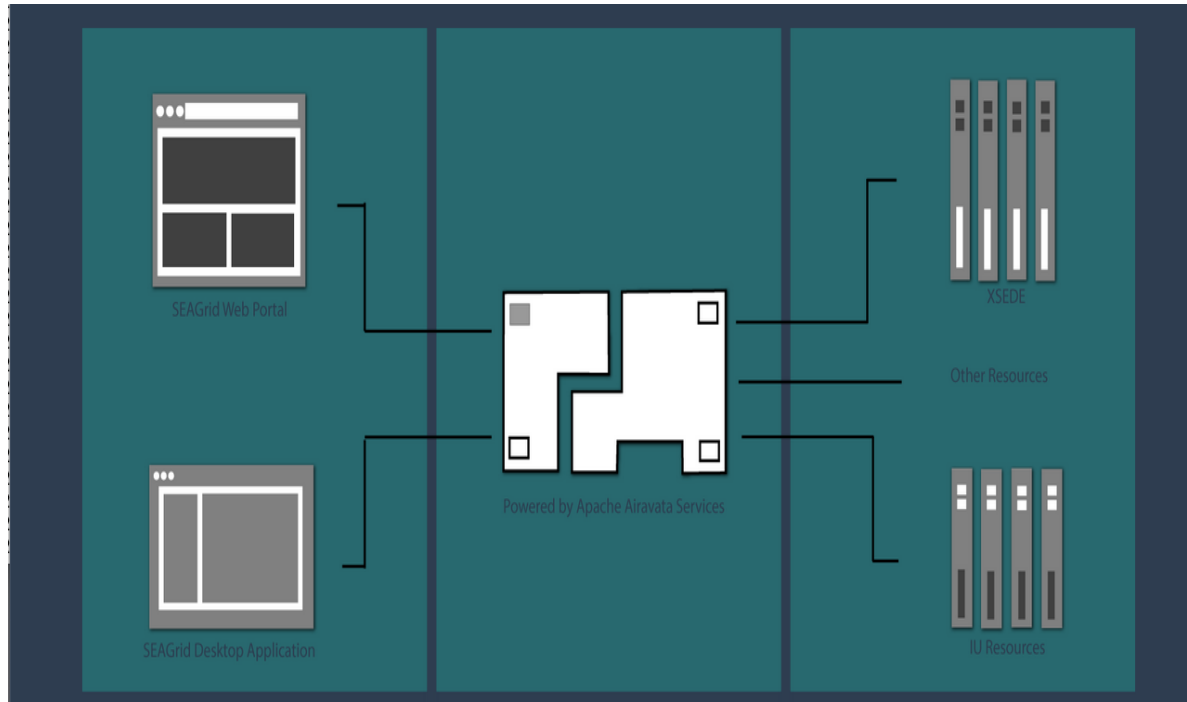
Rupture forecasts, Earthquake physics & Hazard estimates.

## Partners

Statewide California Earthquake Center  
University of Illinois, Urbana-Champaign  
University of Southern California  
Scripps Institution of Oceanography at UC San Diego  
San Diego Supercomputer Center at UC San Diego



# Technical Overview of Science Gateway



Quakeworx  
Web-browser  
User Interface

OneSciencePlac  
e Gateway  
Software SDSC

SDSC Expanse  
Computer and  
Storage

What's the  
difference  
between a  
website and  
a gateway?

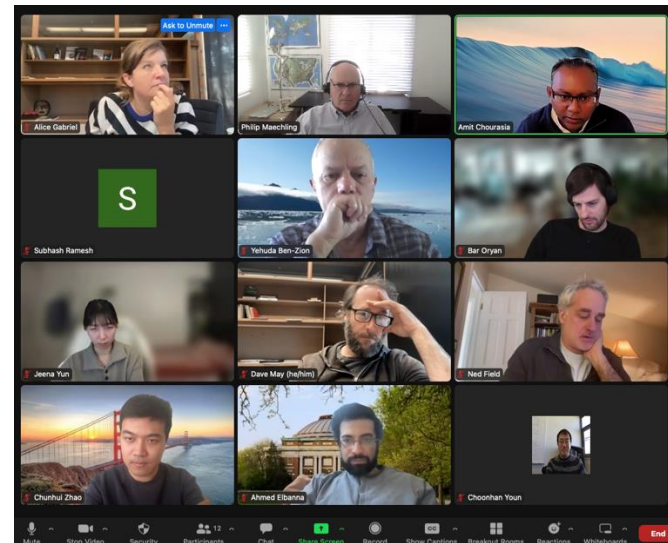
Primarily  
the  
computing  
and storage  
available  
through the  
gateway.





# Testing Essential In Scientific Process

Researchers with community-oriented computing needs should coordinate with the Quakeworx Pis (Ben-Zion, Gabriel, Elbanna, Chourasia) to install forecast experiments in Quakeworx.





**SCEC** ... all thanks to you!



The **Statewide California Earthquake Center (SCEC)** develops and shares cutting-edge earthquake system science to enhance California's resilience and to educate and inspire future scientists.

# Quakeworx Private Development Website (<https://qwx1.onescienceway.com/>)

The screenshot displays the Quakeworx Private Development Website interface. At the top left is the Quakeworx logo, and at the top right are search and login icons. The main content area is titled "All apps" and includes a list of available applications. A search filter is present with "App type" set to "- Any -" and an "Apply" button. The applications listed are:

- Jupyter Notebook**: WEB APP / DOCKER, ver. 0.1.0, system. AWS System (EC2), Jupyter Notebook
- Moose-FARM**: MOOSE BATCH APP / EXECUTABLE, ver. 0.0.2, system. Expance service, Moose simulator
- SeisSol**: SeisSol BATCH APP / EXECUTABLE, ver. 0.0.1, system. Expance service, 0.0.1
- Tandem 2d revised**: BATCH APP / EXECUTABLE, ver. main\_17c42dc9ae0ec519dccb5732681b2e4054666f1, system. Expance service, Tandem app
- Tandem 3d revised**: BATCH APP / EXECUTABLE, ver. main\_17c42dc9ae0ec519dccb5732681b2e4054666f1, system. Expance service, Tandem 3d app
- UCERF3-ETAS**: BATCH APP / EXECUTABLE, ver. 02b30e5, system. Expance service, UCERT3 ETAS application

At the bottom left, it states "Quakeworx project is powered by OneSciencePlace®".



# Testing Essential In Scientific Process

Proposed floatCSEP development (et al) addresses CSEP principles in a modern way:

Transparency – Software source in open-repository. Experiments are retrieved from open-source repos.

Controlled Environment – Scientific programs are retrieved from open-source repositories and build with specific codes into a docker-based images and run on external cloud resources. Results are written to persistent data storage systems and assigned persistent Digital IDs

Reproducibility – The processing are essentially fully defined, with source code, build environment, and output storage as persistent data.

Comparability – The input data is available archived and made available to other researcher to support comparisons.

Forecasts used in broad impact applications require more types of testing. The end-user of forecasts often define the types of testing required.

