SC/EC

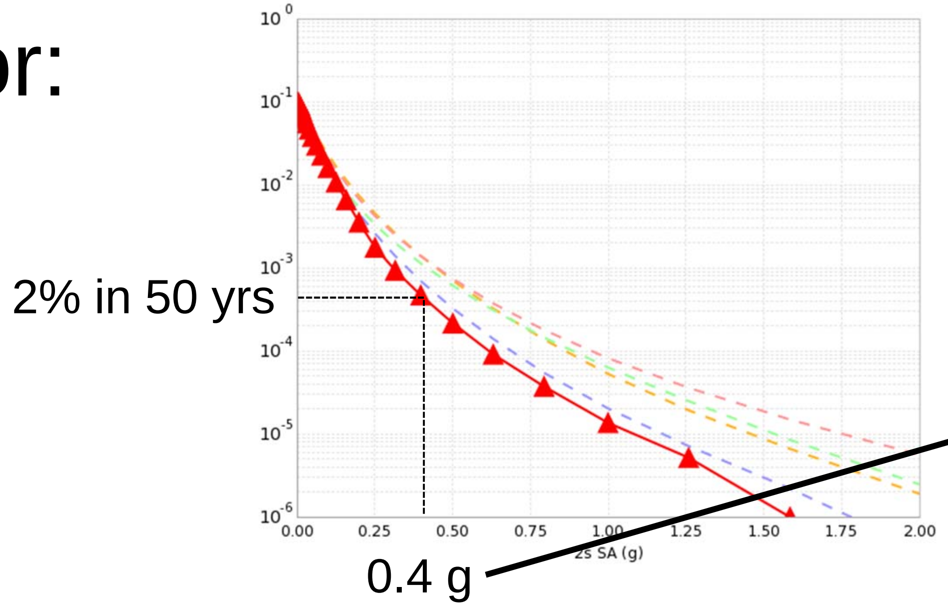
CyberShake Seismic Hazard Models for Central California

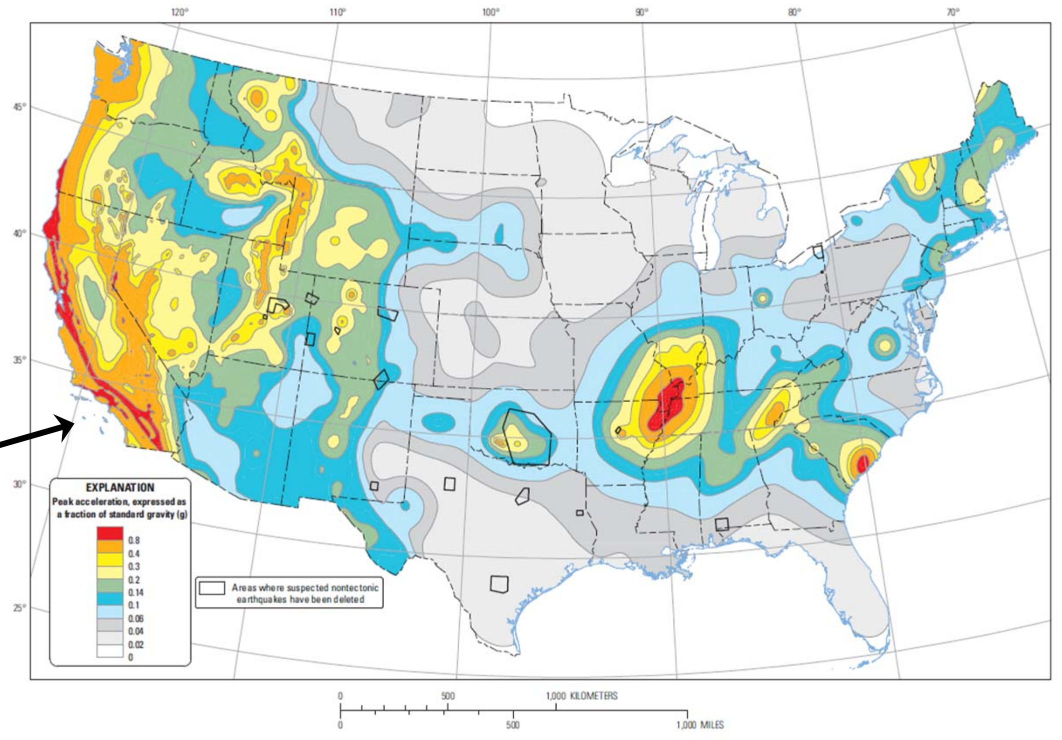
Scott Callaghan

SC17

Probabilistic Seismic Hazard Analysis (PSHA)

- What will peak earthquake shaking be over the next 50 years?
- Useful information for:
  - Building engineers
  - Disaster planners
  - Insurance agencies



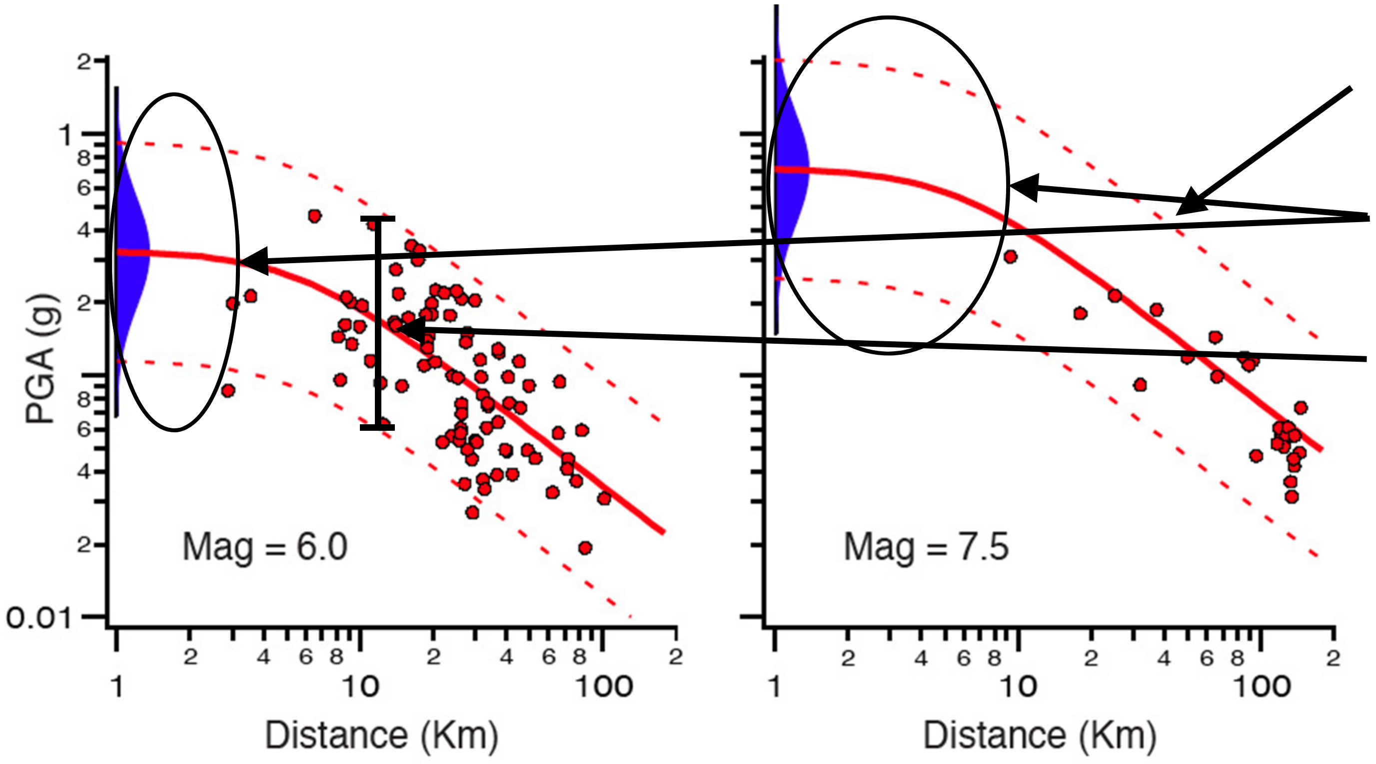


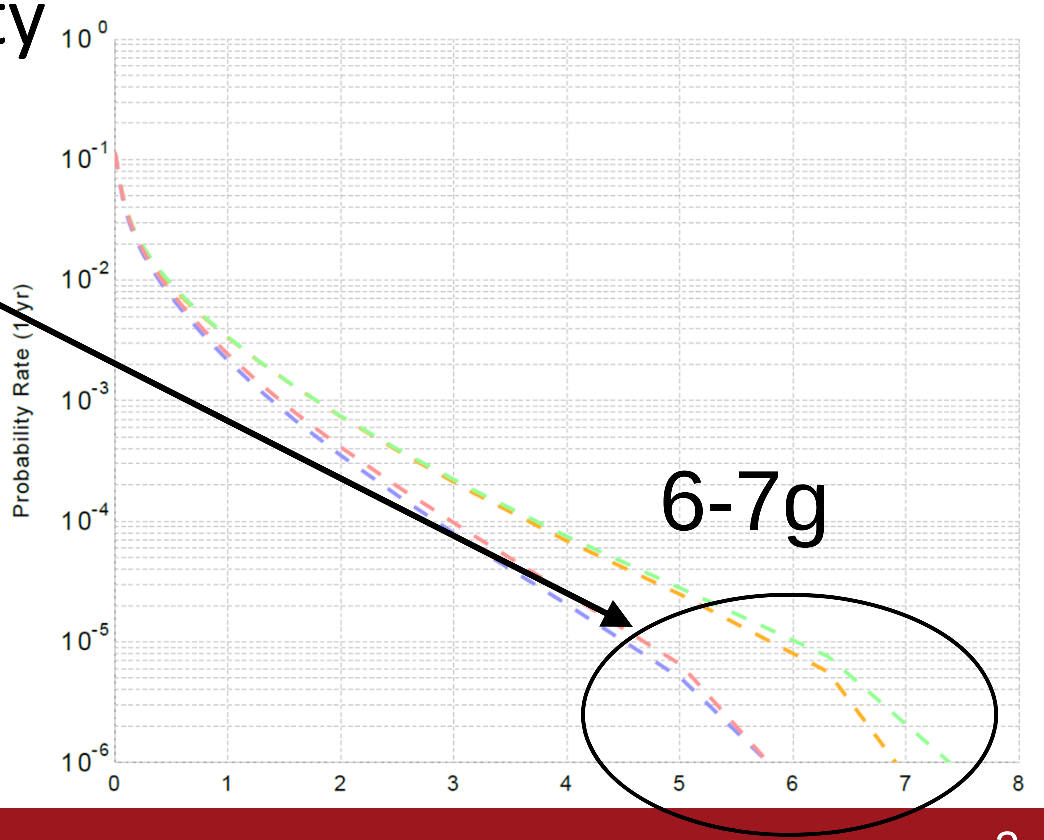
Two-percent probability of exceedance in 50 years map of peak ground acceleration

- PSHA performed by
  1. Assembling a list of earthquakes
  2. Determining how much shaking each event causes
  3. Combining the shaking levels with probabilities

PSHA Approaches

- Ground Motion Prediction Equations (GMPEs)
  - Equations derived from historical data

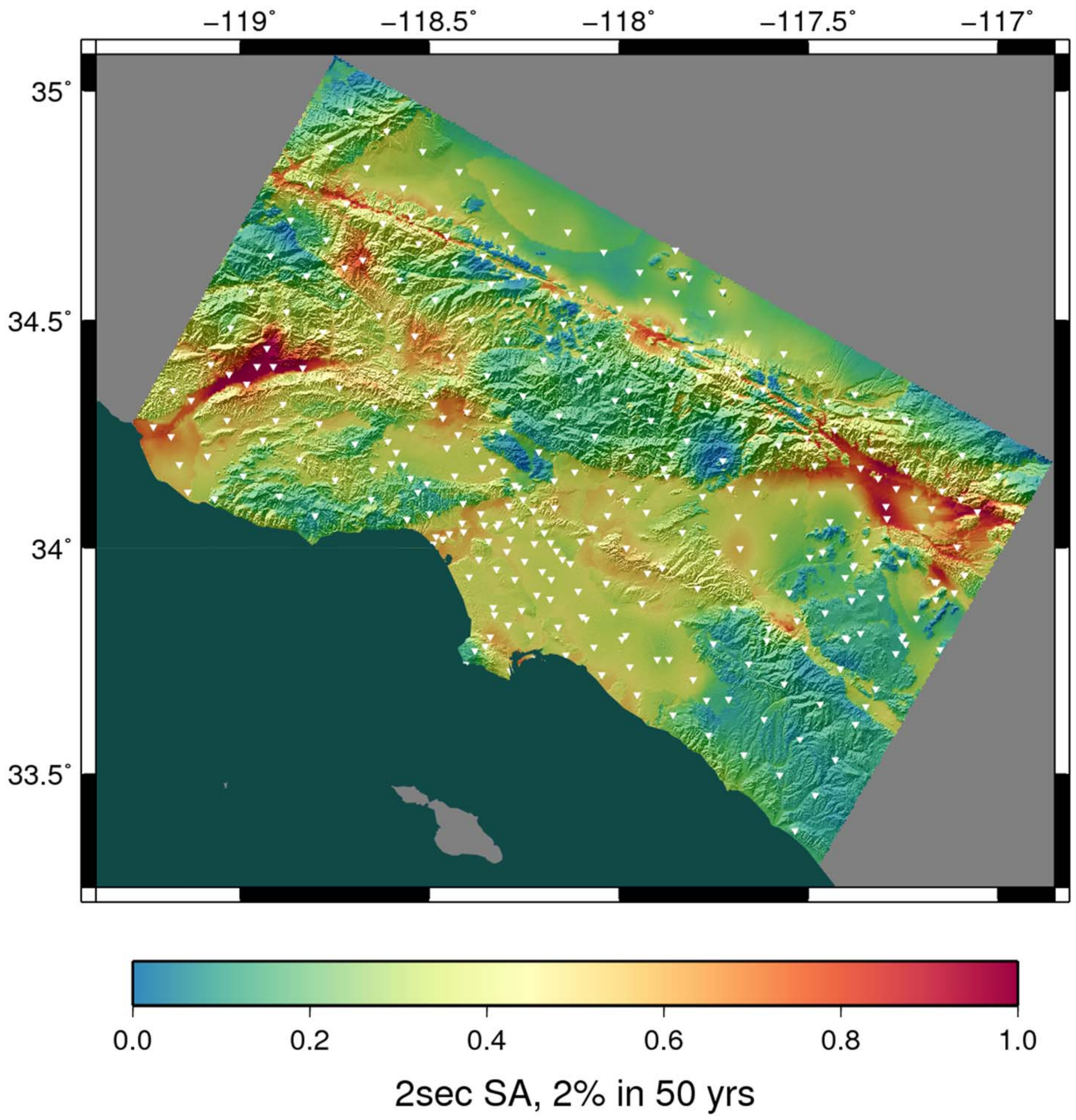




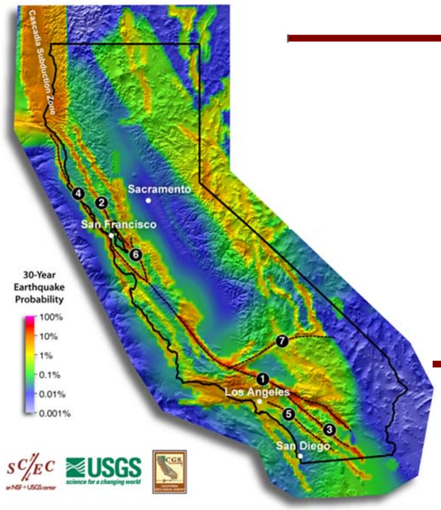
- Simulation-based approach
  - Reduce uncertainty by capturing complex physics
  - Computationally expensive

SCEC CyberShake Project

- 3D physics-based platform for PSHA
- For each site of interest:
  - Determine nearby (<200 km) earthquakes
  - Add variability to earthquakes
  - Simulate each of 500,000 earthquakes
  - Determine maximum shaking from each
  - Combine with probabilities
- Project began in 2007
- Continual improvement

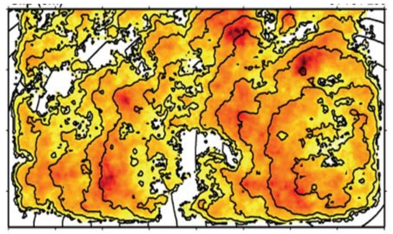


CyberShake Data Flow



Uniform California Earthquake Rupture Forecast

Rupture Generator



UCVM
Mesh generation
1 job per site
MPI, 1500-4000 cores

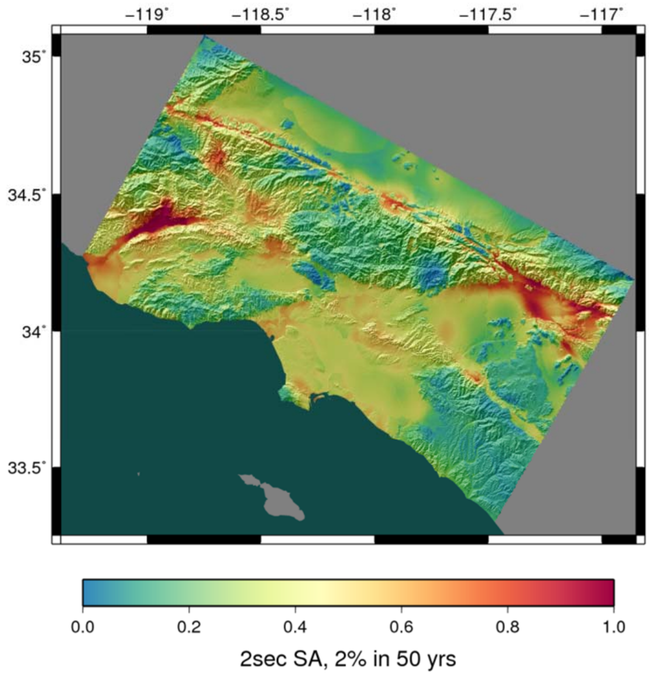
AWP-ODC
SGT computation
2 jobs per site
MPI, 200-800 GPUs

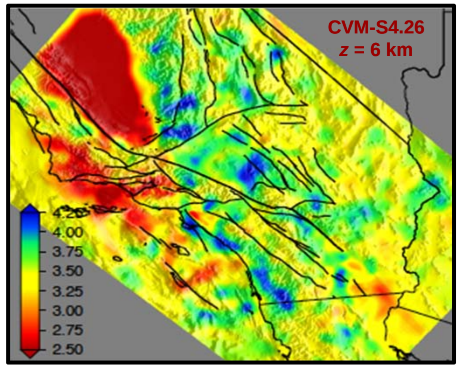
Seismogram Synthesis
Post-processing
500,000 jobs per site
MPI master/worker, 3712 cores

3 TB
data transfer

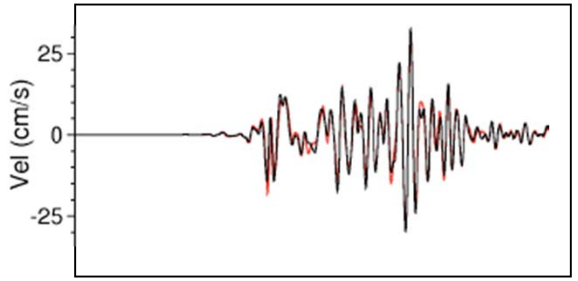
Data Product Generation
Populate DB, construct images
6 jobs per site

CyberShake Hazard Map

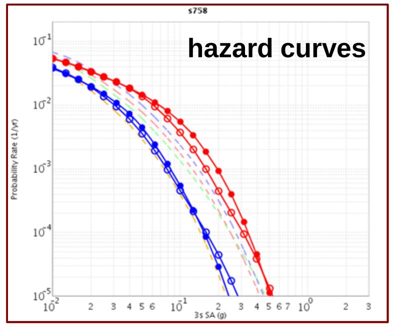




Community Velocity Model



500,000 Seismograms
75M intensity measures



Most recent CyberShake study took 4 weeks of real time and used 21 million core-hours

CyberShake Requirements

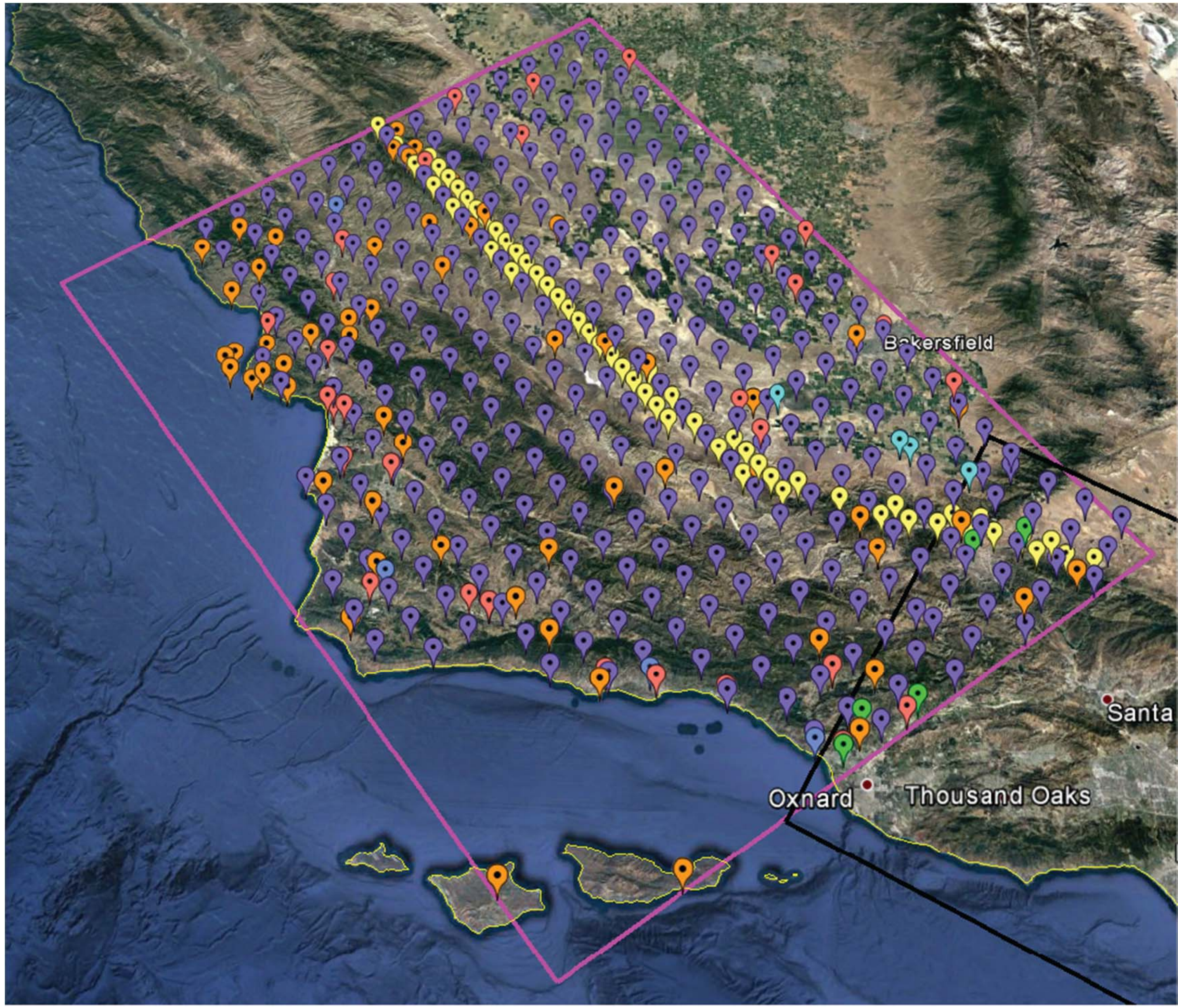
- 200-400 sites needed for a hazard map
- Execute jobs across multiple systems
  - USC HPC, NCSA Blue Waters, OLCF Titan
- Automated execution of jobs
- Data management
  - 28,000 files and 30 GB output per site
  - Migration of input and output files when needed
- Error recovery
- Decided to use scientific workflows

CyberShake Workflow Tools

- Pegasus-WMS (USC ISI)
  - Use API to describe workflow as tasks with dependency
  - Plan workflow for execution on specific resource(s)
  - Adds data transfer jobs and metrics wrappers
  - Intro to Pegasus Wed at 2; office hours Tues at 3 and Wed at 3
- HTCondor (U of Wisconsin)
  - Manages runtime execution of jobs
  - Resolves dependencies
  - Checkpoints workflows
- Globus (booth #373)
  - GRAM for communication between workflow host and remote system
  - GridFTP for file transfer

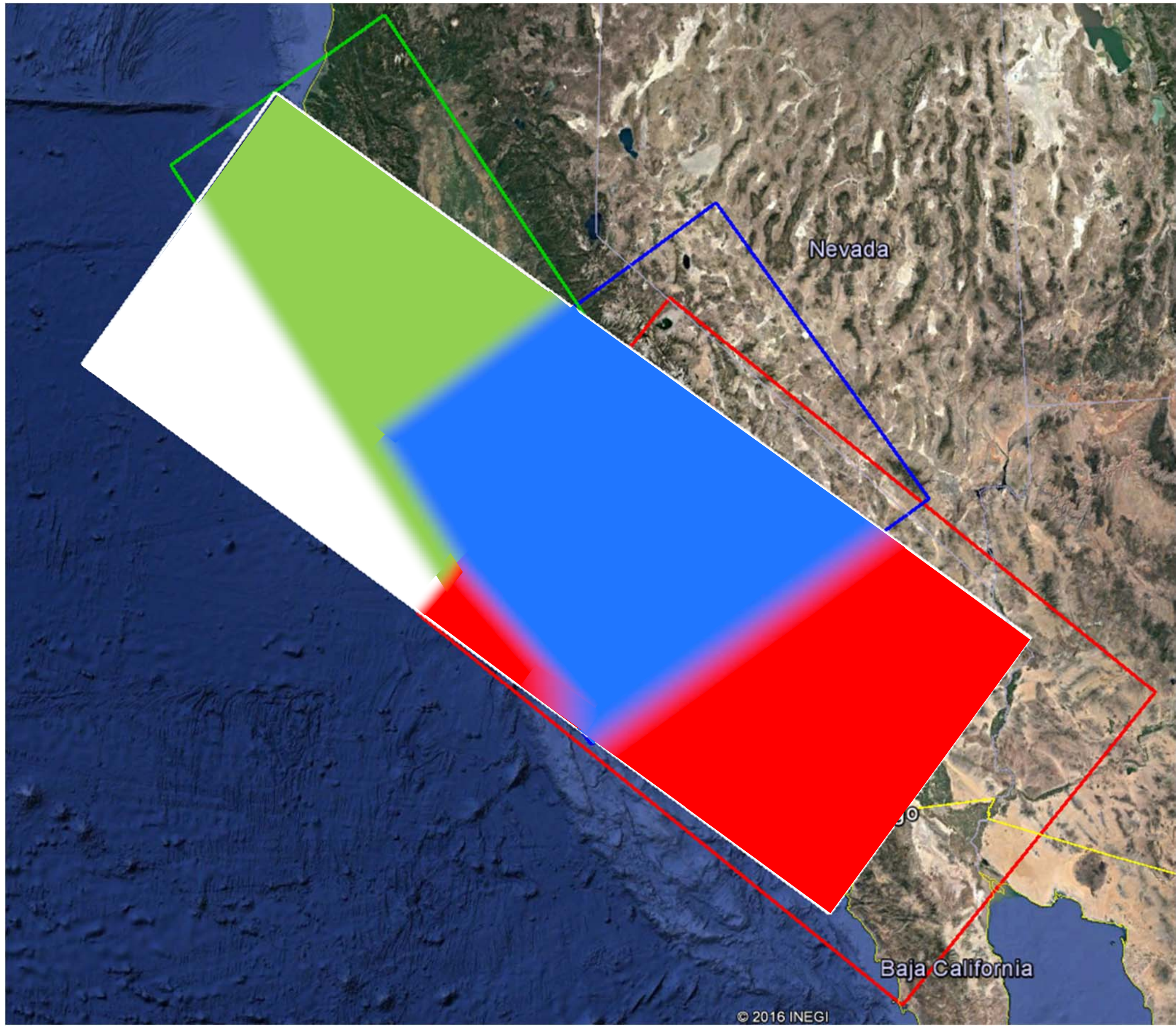
CyberShake in Central California

- 2017 science goal
- Proof-of-concept that CyberShake can be moved outside of Southern California
- New sites
- New velocity model of earth's crust
- New workflow approach



Central California Challenges

- Simulation volume
  - Had to combine multiple velocity models
  - Smooth across interfaces
- Targeted OLCF Titan for 200-node GPU jobs
  - Requires two-factor authentication
  - Difficult to automate job submission
  - Tried pilot job approach in 2015; successful but with 32% overhead
  - Looking for more efficient solution

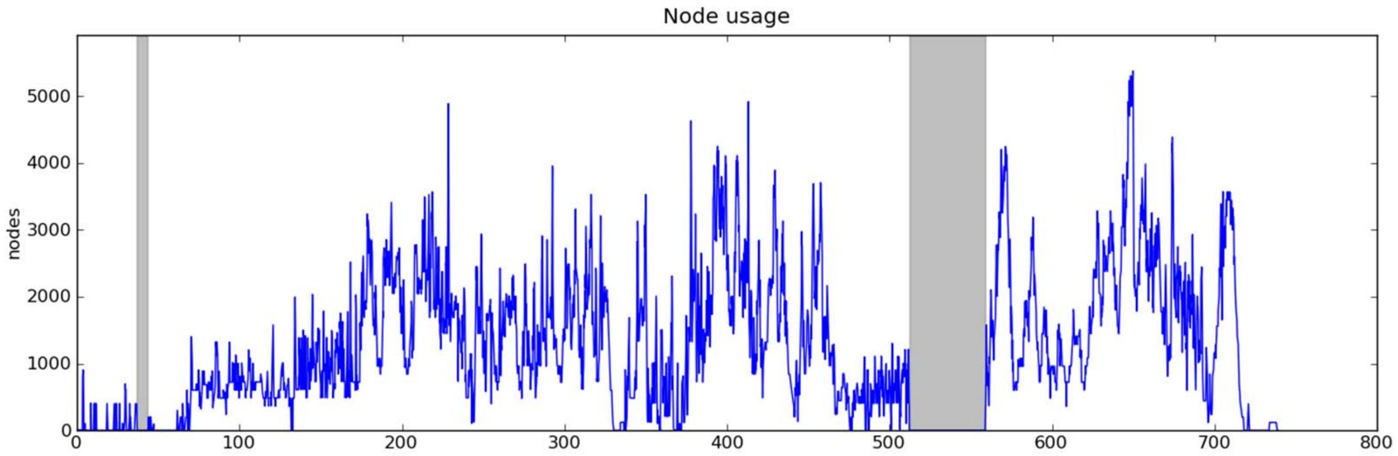


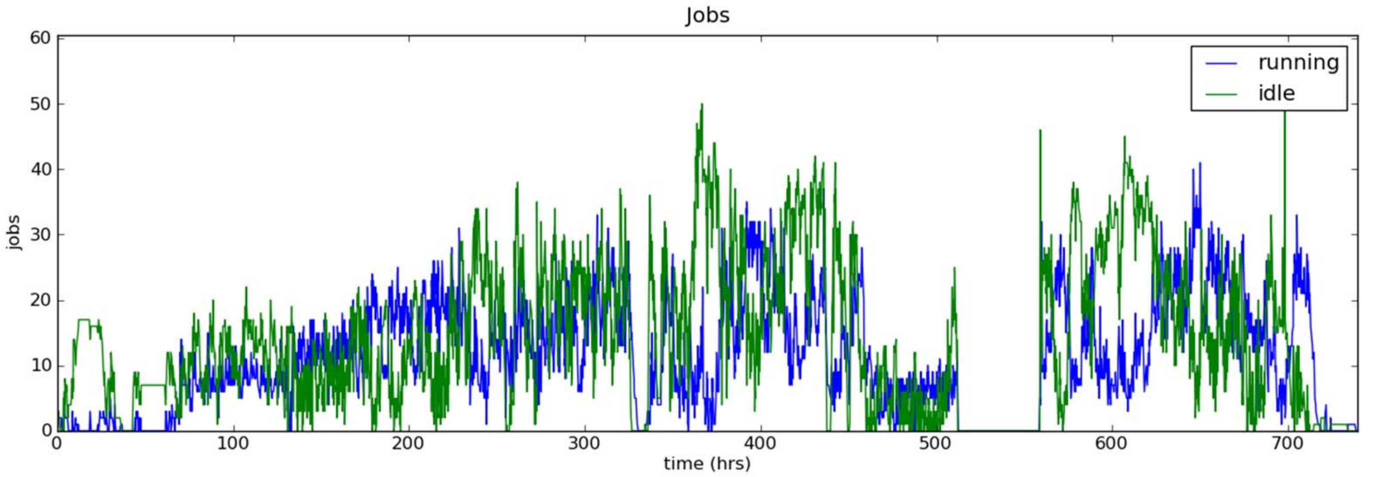
rvGAHP Workflow Approach

- “reverse GAHP” – approach using HTCondor protocol
  - Enables remote job submission without authentication to remote system
- Daemon runs on remote system
- Initiates connection to workflow submission host
- Workflow can then use connection to submit jobs to remote queue
- Opens up Titan to low-overhead workflows

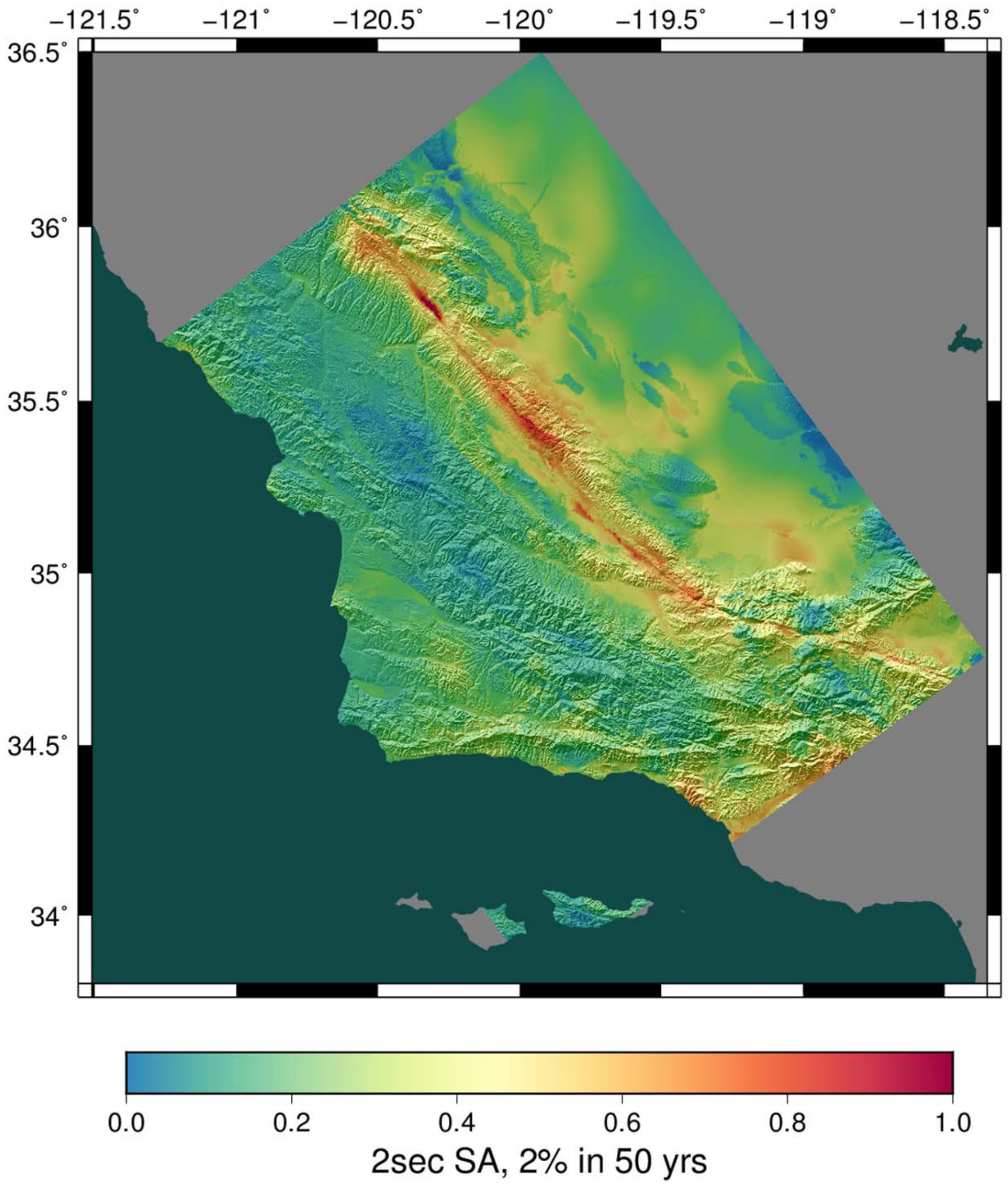
CyberShake Study 17.3

- Conducted over 31 days in March-April 2017 on NCSA Blue Waters, OLCF Titan, USC HPC
- 2 models at 438 sites
- Averaged 1295 nodes, max of 5374
  - 900,000 node-hrs (21.6M core-hrs)
- Workflow tools scheduled 15,581 jobs
- 777 TB of data managed
  - 308 TB transferred
  - 10.7 TB archived on USC disks
- Generated 285 million seismograms

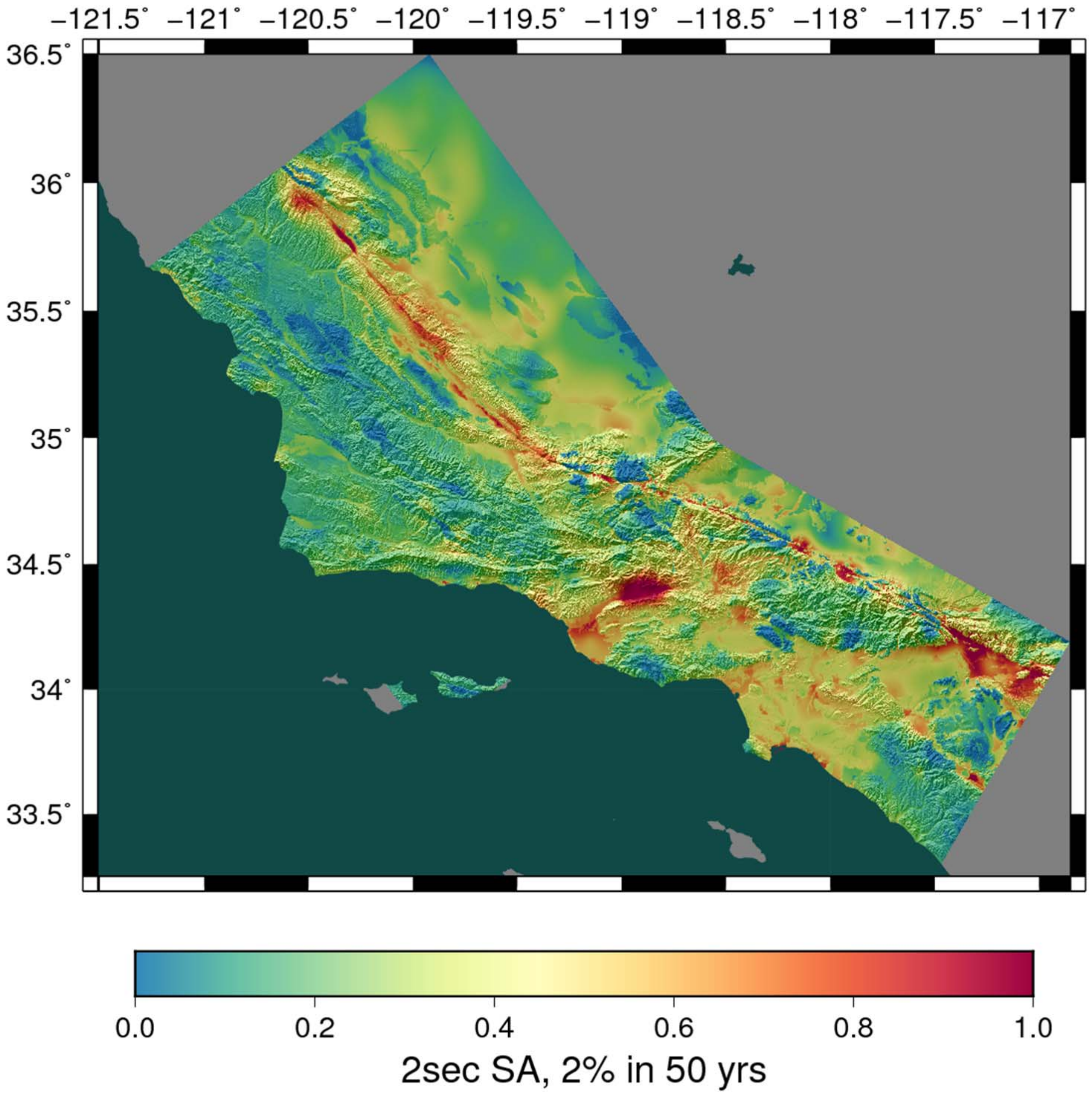




CyberShake Study 17.3 Results



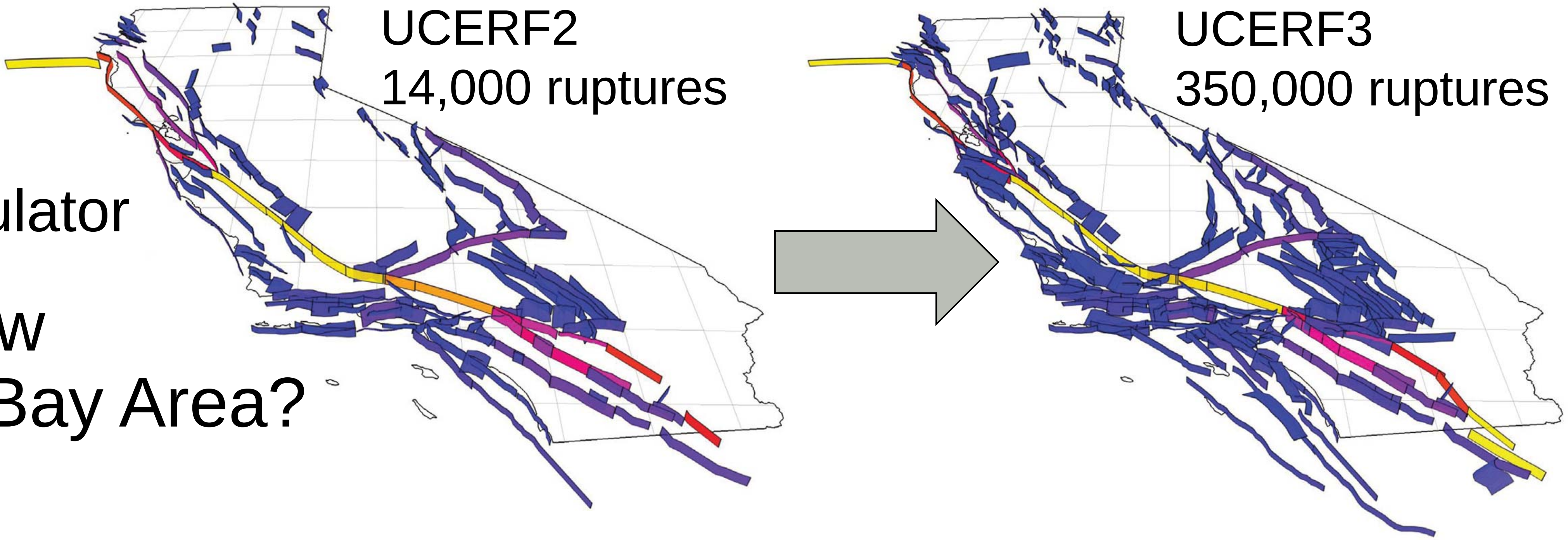
Study 17.3



Combined map

Future Directions

- Move to new forecast
  - UCERF 3
  - RSQSim earthquake simulator
- Run CyberShake in new regions – San Diego? Bay Area?



- Add new physics
  - Increase maximum frequency – applicable to more buildings
  - Velocity model heterogeneities, fault roughness, topography
- Optimize
  - Seismogram synthesis – compress data? Large shared-memory nodes?
  - Use machine learning to eliminate some earthquakes

Thanks!

SCEC
AN NSF+USGS CENTER

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

U.S. DEPARTMENT OF ENERGY

NCSA

BLUE WATERS
SUSTAINED PETASCALE COMPUTING

ISI
Information Sciences Institute

USGS
science for a changing world

TACC

U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

NSF National Science Foundation

XSEDE
Extreme Science and Engineering
Discovery Environment

Pegasus

USCGRID

OAK RIDGE National Laboratory
OAK RIDGE LEADERSHIP COMPUTING FACILITY