CyberShake Study 22.12 Technical Readiness Review

Scott Callaghan

Study 22.12 Technical Goals

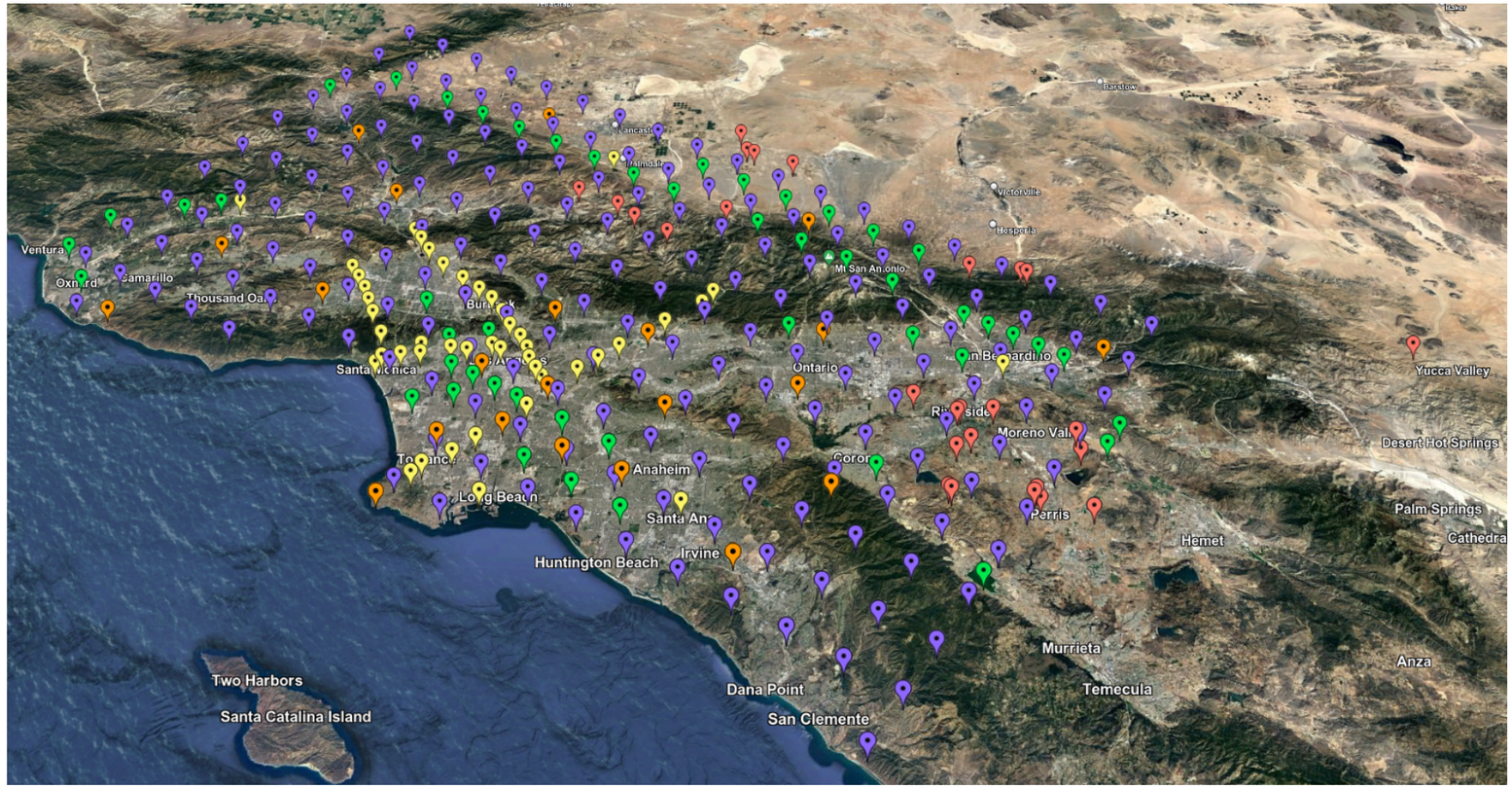
- Perform computationally largest CyberShake study to date
- Use job bundling to improve queue times for SGT and post-processing jobs
- Improve upon high-water mark of 46% of Summit

Study 22.12 Scientific Goals

- Perform an updated Broadband CyberShake calculation in Southern California, using BBP codes.
- Calculate CyberShake results with a modified velocity model.
  - Merged taper model, using CVM-S4.26.M01 with the Ely-Jordan taper to 700m
- Use an updated version of the Graves & Pitarka rupture generator
  - V5.5.2, same as is used in the BBP v22.4
  - Includes sampling of variability in rupture velocity
  - Denser hypocentral spacing (4.5 km to 4 km) → 31% more variations

Proposed Study sites

- Same 335 sites as were used in Study 21.12
- Site calculation order:
  - Standard 10 (USC, PAS, LADT, LBP, WNGC, SABD, SBSM, FFI, CCP, SMCA)
  - 20 km grid
  - 10 km grid
  - Additional POIs
  - 5 km grid sites



Study 22.12 Data Products

- Two-component seismograms (626k/site x 335 sites x 2 = 420 million)
- Intensity measures
  - Deterministic
    - RotD50 and 100 at 25 periods (1-10 sec) + PGV: 2, 3, 4, 5, 7.5, and 10 sec stored in DB
  - Stochastic
    - RotD50 and 100 at 66 periods (0.01-10 sec) + PGA, PGV: 19 periods, PGA, PGV stored in DB
- Duration metrics
  - Energy integral, Arias intensity, cumulative absolute velocity
  - Significant durations (5-75%, 5-95%, 20-80%) for velocity and acceleration: acceleration 5-75% and 5-95% for X and Y stored in DB
- Hazard curves for 335 sites (0.1, 0.2, 0.5, 1, 2, 3, 4, 5, 7.5, 10 sec)
- RotD50 hazard maps at 0.1, 0.2, 0.5, 1, 2, 3, 5, and 10 sec

Study 22.12 Parameters

- 1 Hz deterministic
  - 100 m grid spacing
  - dt = 0.005 sec
  - SGT nt = 40000 timesteps (200 sec); 60000 timesteps/300 sec if any source-to-hypocentral distances are greater than 450 km
  - Seismogram nt = 8000 timesteps (400 sec)
- 50 Hz stochastic
  - Seismogram nt = 40000 timesteps (400 sec)
- UCERF 2 ERF
- Graves and Pitarka (2022) rupture generator

Changes to SGT Workflow Software Stack

- Velocity mesh generation
  - UCVM 22.7.0
  - Implemented option for merged taper model
  - Modified implementation of minimum values
- SGT code
  - Te-Yang's kernel update applied to ensure reproducibility

Changes to PP Workflow Software Stack

- DirectSynth post-processing code
  - Added OpenMP implementation
  - Support for passing in random seed and rupture velocity
- Broadband code
  - CyberShake now links to BBP object files for consistency
  - Merge approach modified
- Rupture generator
  - Updated to v5.5.2 and optimized

Changes to Workflows

- SGT and DirectSynth jobs will be bundled to reduce queue times
  - Bin 2 jobs (922-2764 nodes) get +10 day aging boost
  - Bundles of 14 SGT jobs (938 nodes) x 40 minutes
  - Bundles of 10 DS jobs (1000 nodes) x 10 hours
- Other jobs run using rvGAHP, with daemon running on dtn36
- Use default pegasus-transfer to manage Globus Connect transfers
- Pegasus v5.0.3
- HTCondor v9.1.0

Computational Requirements

- SGT workflow
  - 78.8 node-hrs/site x 335 sites + 10% extra = 29k node-hours
- Deterministic PP workflow
  - 8.6 hrs/site x 92 nodes x 335 sites + 10% extra = 292k node-hrs
  - About 1/3 more rupture variations per site than previous studies
- Stochastic PP workflow
  - 8.8 hrs/site x 92 nodes x 335 sites + 10% extra = 297k node-hrs
  - High-frequency code is longer-running than in previous BB studies
- Currently 285k node-hours available on Summit

Computational Plan

- Unsure if OLCF will permit allocation overrun, as they have previously
- Perform 20-site stress test of all 3 phases: 33.6k node-hours
- Calculate SGTs for remaining 315 sites (24.8k node-hours)
- Calculate low-frequency post-processing for remaining 315 sites (249.2k node-hours) – this comes to ~20k node-hours over allocation
  - Best to calculate this entirely on Summit, so SGTs don't have to be transferred.
- While study is going, establish broadband capability on Frontera
- If OLCF cuts us off, run broadband calculations on Frontera
  - 375k node-hours (828k available)

Storage Requirements

- Summit (learned from Study 21.12 that no quotas are enforced on scratch)
  - 507 TB SGTs
  - 553 TB temporary data
  - 73 TB output data
- CARC
  - 73 TB output files to project storage
  - 66 TB free; will migrate additional pre-Study 15.4 data to OLCF HPSS
- Database on moment.usc.edu
  - 654 GB (7.1 billion rows)
  - Moment has 379 GB on disk + 846 GB in database free
- Shock-carc workflow submission host
  - 737 GB workflow logs (1.5 TB free)

Estimated Duration

- At Study 21.12 throughput rates:
  - 116 days to use up Summit allocation
  - 252 days to finish study
- Anticipate higher throughput on Summit due to job bundling
- At Study 21.12 percentage of system usage, 88 days to complete broadband calculations on Frontera
- Overall estimated duration of 90 days on Summit + 88 days on Frontera with 30 days overlap = 150 days.

Risks

- Storage on Summit
  - ~460 TB for SGTs for 315 sites
  - May need to temporarily store them elsewhere (OLCF HPSS, CARC scratch, TACC Ranch)
- Storage on CARC
  - Plan to migrate more data to OLCF HPSS (can move ~1.5 TB/day)
  - After study completes, investigate data compression for seismograms
- Limited compute time on Summit
  - Move calculations to Frontera
  - Additional resources include CARC, Expanse, Perlmutter, Delta
- Reduced support during holidays
- Dependency on moment

Action Items

- Tag code in github repo
- Test Pegasus with Globus file transfer (with v5.0.3)
- Notify OLCF contact of plans, and request additional bin 5 jobs