SCEC’s Research Computing and Software Ecosystem

Philip Maechling
SCEC Open Source Software and Data Access Workshop
8 September 2018
Palm Springs, California
Welcome

Welcome to the SCEC Open-source Software and Data Access Workshop.

• We appreciate your interest and participation!

We want this to be useful to SCEC research community:

• Please feel free to ask questions.

• We hope this workshop supports your use of SCEC software.
Workshop Goals Include

1. Widen the research communities knowledge about SCEC Software and Data Access Tools
2. Support SCEC researchers who’s research would benefit from available tools.
3. Provide a venue for users to ask questions about their advanced use or special cases.
4. Gather feedback on current software and on future desired capabilities.
Meeting Logistics

Agenda Overview (posted online): [https://www.scec.org/workshops/2018/software](https://www.scec.org/workshops/2018/software)

- 30 min of Introductory Presentations (including this one)
- 6 x 45minute Software/Data Access Tool Overview & Demonstration Sessions
- 15 min break ~10:15am
- 1hr Lunch break ~ 12noon (boxed lunches provided)
- 2.5hr Software Clinic ~ 2:30pm (run demos on VirtualBox, ask questions of developers, discuss special or advanced features)
Meeting Technical Logistic

Workshop organized to show SCEC software programs (demonstration oriented).

Workshop Users Guide developed so participants can run the demonstration using the VirtualBox version of our software on their own computers after the workshop. Latest version of Workshop Users Guide posted online (hardcopies in the room may lack last minute updates)

Presentations slides will be posted after the workshop

Technical support from SCEC Staff at meeting and via software@scec.usc at any time.
Running Demos on your Computer

SCEC Software installed in a VirtualBox image file so participants can run the demos on their own computers without installing SCEC software.

Workshop image contains simplified versions of the codes included to save space. Full install needed to obtain all features, models, apps.

VirtualBox image available on in-room USB (preferred) or online. Installation may take 30 minutes so we recommend you install during breaks, or at lunch.

Software Clinic will provide a small number (~5) of shared computers with the Workshop VirtualBox codes installed for use during Software Clinic.

Technical support from SCEC Staff at meeting and via software@scec.usc at any time.
Workshop Approach

Many of the software tools under discussion are complex. We will not attempt comprehensive descriptions, of the codes, just introductions.

We will hold six (6) sessions with similar formats, with a Motivation and Overview for each software tool, and then a demonstration.

We plan to combine demonstration and discussion times to allow for questions during demonstrations.

Will focus on the essential aspects of software. Details and advanced uses can be covered in the afternoon Software Clinic.
• Goal of research is to predict behavior of nature
  *Provide prospective information about earthquake behaviors useful in reducing seismic risks and preparing for earthquake disasters*

• Earthquake behaviors of societal interest
  – Fault rupture
  – Ground shaking
  – Tsunami
  – Liquefaction, landsliding, and other secondary effects
Seismic Hazard Analysis is a System-Level Problem

- Seismicity
  - Seismicity
  - Paleoseismology
  - Local site effects
  - Geologic structure

- Faults
  - Faults
  - Stress transfer

- Seismic Hazard Model
  - Crustal motion
  - Crustal deformation
  - Seismic velocity structure
  - Rupture dynamics

InSAR Image of the Hector Mine Earthquake
• A satellite generated Interferometric Synthetic Radar (InSAR) image of the 1999 Hector Mine earthquake.
• Shows the displacement field in the direction of radar imaging.
• Each fringe (e.g., from red to red) corresponds to a few centimeters of displacement.
SCEC Community Modeling Environment Research is Designed to Improve Ground Motion Forecasts

1. OpenSHA: Uniform California Earthquake Rupture Forecast (UCERF3)

2. CyberShake: physics-based seismic hazard analysis

3. TeraShake: ShakeOut, M8, Chino Hills simulations

4. DynaShake: dynamic rupture modeling

5. F3DT: full 3D waveform tomography
Office of Advanced Cyberinfrastructure (OAC) supports and coordinates the development, acquisition and provision of state-of-the-art cyberinfrastructure resources, tools and services essential to the advancement and transformation of S&E.
**SCEC NSF SI2 Awards**

SI2 means Software Infrastructure for Sustained Innovation. SCEC was funded through NSF SI2 (and other) programs to create software infrastructure needed by the geoscientific community. Both OAC and Geo Directorates have contributed to the funding for these software projects. The desired outcome is Sustainable Software that helps geoscientific research.

This workshop funded as part of SCEC SI2 Awards:

**SI2-SSI: A Sustainable Community Software Framework for Petascale Earthquake Modeling**  
PI: Thomas H. Jordan, Jacobo Bielak, Yifeng Cui, Kim B. Olsen  
Project Dates: 1 August 2012 through 31 July 2015  
Award Number: OCI-1148493

**SI2-SSI: Community Software for Extreme-Scale Computing in Earthquake System Science**  
PIs: Thomas H. Jordan, Yifeng Cui, Kim B. Olsen, Ricardo Taborda  
Project Dates: 1 September 2015 through August 31, 2019  
Award Number: ACI-1450451
SCEC Support for Free Open Source Software (FOSS)

SCEC community codes are provided as Free Open Source Software (FOSS).

Free – Academic geoscience community relies heavily on free software
1) Academic research budgets are typically very limited, and commercial software costs inhibit adoption of tools and methods (e.g. GoCAD, AutoMesh)
2) Much scientific software developed with University and Governmental research funds with requirements that software is made publically available without cost.

Open-Source - Scientific community depends on open-source because:
1) Scientific Transparency. Open Source allows others to examine results careful, and to identify exactly how results were produced. Results from proprietary software often difficult to accept without ability to examining code.
2) A public open-source version of the software helps to preserve a version of the code if current software maintainer stops supporting work.
From Research Code to Community Code

Many SCEC researchers develop software. Codes that become community codes tend to share basic characteristics:

Community codes imply:
- Software supports geoscience research purpose needed by multiple researchers
- At least one SCEC scientist supports the scientific use of the software.
- The code is contributed for community use, and for release as FOSS
- Can exchange inputs and outputs with other codes in the ecosystem
- Codes co-evolve (e.g. scale up) together.
Concept of SCEC Software Ecosystem
**Scientific Software Value Depends on Responsible Scientists**

<table>
<thead>
<tr>
<th></th>
<th>Role</th>
<th>Resp. Scientist</th>
<th>Resp. Software Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CVM-S4</td>
<td>MK/HM/RC</td>
<td>Mei-Hui Su</td>
</tr>
<tr>
<td>2</td>
<td>CVM-H</td>
<td>JS/AP</td>
<td>Mei-Hui Su</td>
</tr>
<tr>
<td>3</td>
<td>SORD</td>
<td>KO/SD/GE</td>
<td>…</td>
</tr>
<tr>
<td>4</td>
<td>OpenSHA</td>
<td>EF/PP/KM/…</td>
<td>Kevin Milner</td>
</tr>
<tr>
<td>5</td>
<td>UCVM/CXM</td>
<td>LH/SM/JS/AP/…</td>
<td>Mei-Hui Su</td>
</tr>
<tr>
<td>6</td>
<td>Broadband Platform</td>
<td>CG/CG/KO/AP/…</td>
<td>Fabio Silva</td>
</tr>
<tr>
<td>7</td>
<td>CyberShake</td>
<td>JV/TJ/RG/KO/</td>
<td>Scott Callaghan</td>
</tr>
<tr>
<td>8</td>
<td>CSEP</td>
<td>MW/TJ/WM/DR/DJ/…</td>
<td>William Savran</td>
</tr>
<tr>
<td>9</td>
<td>AWP-ODC</td>
<td>KO/DR</td>
<td>YC, DR, DM, OO</td>
</tr>
<tr>
<td>10</td>
<td>Hercules</td>
<td>JB/RT</td>
<td>D, RT, NK</td>
</tr>
<tr>
<td>11</td>
<td>RSQSim</td>
<td>JD/KR/JG</td>
<td>BS, KRD, KM, JG</td>
</tr>
<tr>
<td>12</td>
<td>SCEC-VDO</td>
<td>TJ</td>
<td>KM/JY</td>
</tr>
</tbody>
</table>

SCEC provides the scientific community needed to develop, evaluate, and use scientific software. These codes could not be produced by computer scientists working alone.
Productivity Gains from a Software Ecosystem

As a software developer, I recognize the great efficiencies that can be achieved with a collection of software utilities to support an activity. Groups with a effective software collection can produce results much more rapidly.

SCEC research has benefitted from the great efficiencies of the SCEC Computational Pathway software eco-system. This collection of software, including UCVM, OpenSHA, Broadband, AWP, Hercules, CyberShake, RSQSim are used together.

It is not individual codes, but the collection of codes, that has produced the achieved research productivity
Software Sustainability Challenge

For those working in the scientific software area, there is significant activity and interest and growing community interested in sustainable scientific software:

https://sciencegateways.org//united-states-research-software-sustainability-institute-urssi-conceptualization

http://wssspe.researchcomputing.org.uk/
Southern California Earthquake Center