

NSF Site Visit to the Computational Infrastructure for Geodynamics

Community Issues and CIG
Marc Spiegelman (LDEO/Columbia)



CIG and the Geodynamics community

- *Dual Mission* of CIG
 - Serve Community *and* Transform Community
 - Delicate Balance and Principal challenge
- Community Building Successes
 - workshops, SSC, software policy, software projects
- Toward Transformation:
 - Internal Guidance: Science Steering Committee
 - External Engagement: CIG Proposals and independent proposals
- Other issues/Open Questions

Fundamental Community Challenges for CIG

- **Goal #1:** Transform the way the solid-earth community computes (from the Strategic Plan):
 - Develop *reusable, well documented and open-source* geodynamics software;
 - Develop an *infrastructure layer* from which state-of-the-art modeling codes can be quickly assembled
 - Extend existing software frameworks to interlink multiple codes and data through a *superstructure layer*
 - Form strategic partnerships with the larger world of computational science and geoinformatics

Fundamental Community Challenges for CIG

- Goal #2: Be “community driven”: provide tools the community actually *wants* and will *use*
- A delicate balance and dynamic tension...
- Question: How do you give the community what it wants and change what it wants at the same time?
 - Answer #1: Start with what it wants...
 - Answer #2: Show it a better way through concrete examples...

Successes in CIG

Community Building

- Community elected Science Steering Committee
 - Represents all major Earth science disciplines and computational science
- Community Workshops
 - 4 completed in 2005 (Computational Seismology, Short-term Crustal Dynamics, Long-Term tectonic modeling, Mantle Convection)
 - 5 proposed or in progress for 2006
 - CIG-MC Compressible Convection: March 2006, Purdue
 - CIG-Fault Systems: June 2006, Golden CO
 - CIG-Magma Dynamics: August 2006, Columbia, NYC
 - CIG-SEIS: Fall 2006, Rice?
 - CIG-Scientific Computation: Fall 2006 Austin, TX
 - + software training workshops (SEDI, CIDER...)
- Workshops have been principal mechanism for education, community building and developing software priorities.

Successes in CIG Community Building

- Community driven software initiatives
 - GALE: direct response to long-term tectonics meeting
 - PyLith: Product of SCEC, then CIG workshops
 - “1-D” Geophysics: Seismology and Global dynamics
 - Compressible Convection: from mantle convection ws.
 - Contributed codes/[Software repository](#): CitcomS, CitcomCU, SpecFem3D
- CIG: [Open Software Policy](#)
 - Maintain repository with open checkout policy
 - Try wherever possible to be “open source” (i.e. establish copyright but distribute under the Gnu GPL)
 - Maintain flexibility to maximize community benefit

Successes in CIG Community Building

- Considerable activity, directly servicing needs of communities/disciplines
- But still tend to be disciplinary, stand-alone codes.
 - Actually, some infrastructure/reusable components exist or will be incorporated in current software initiatives, however, emphasis has been on getting specific functionality to work.
- How do we take the next step?

Toward Transformation

- How to move from initial CIG codes to a modular, reusable framework (that still delivers scientific utility)?
- The Science Steering committee
 - Community elected
 - “emphasis is to identify and balance common needs across disciplines” (strat. plan)
 - Guide Software development to balance Earth Science and computational science objectives
 - Ideally choose specific tasks that that provide both scientific utility and demonstrate new functionality.
- Still need constant source of new ideas from community: Two mechanisms
 - Proposals *to* CIG for new functionality
 - Independent NSF proposals: Interactions with CIG
 - Question: How to maximize and incorporate individual innovation into/with CIG?

Proposals submitted to CIG

- [General Policy](#) for proposal submission
 - #1 *Members of the scientific community are encouraged to provide input into the priorities of CIG by submitting proposals to CIG. Proposals are expected to span the range of activities from community building workshops...to specific software tasks and projects*
 - #6 *all proposals should take a balanced view between earth science and computational science objectives*
- 5 Proposals to date: all posted on [website](#)
 - Mantle Convection Workshop
 - A framework for Adaptive meshing (becomes CS workshop)
 - Crustal Deformation software (PyLith)
 - 3-D Magma Migration Demonstration (become MM workshop?)
 - Global Spectral flow codes and benchmarks ("1-D" geophysics)

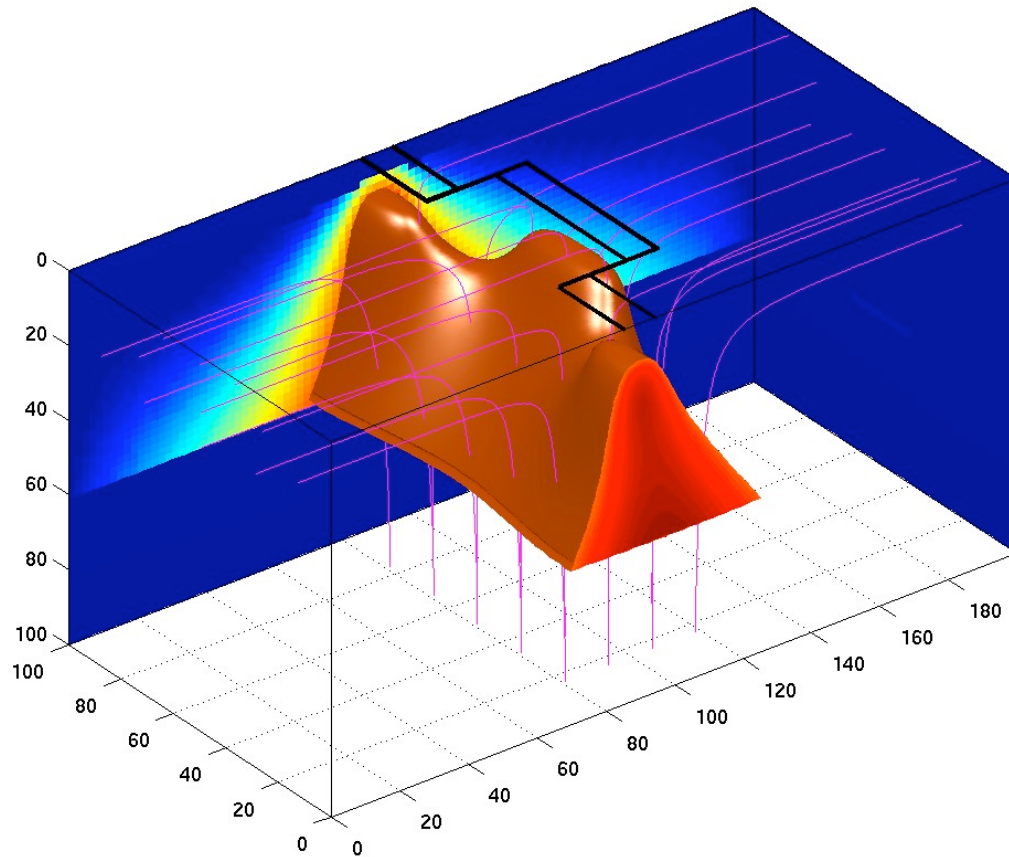
Proposals submitted to CIG

- Issues:
 - How to encourage more proposals?
 - Is there a minimum/maximum number?
 - *Iterative refinement and SSC feedback is crucial.*
 - *How to guarantee open channels for discussion?*
 - *Need an improved commenting mechanism?*
 - Clear definition of “demonstration” goals (SSC responsibility)
 - How to balance individual vs. community initiatives (e.g. Magma Migration proposal)?
 - Better community input mechanisms?
- Need more experience (and more proposal) but seems manageable

Interaction of CIG with independent researchers and external NSF proposals

- Beyond existing CIG projects is a community of talented/innovative scientists with some overlap, but no strong ties to CIG.
- Question: How to tap, co-opt, encourage this community without stifling innovation.
- Two Examples

Example #1: Other frameworks & approaches



- 3-D MOR calculation w/ melting region Richard Katz (LDEO)
- Fully, T - σ dependent viscosity
- Scalable Parallel (PETSc)
- 3 weeks development time (easier for Rich than learning CitComCU)

Questions:

- How to compare with CIG efforts (CitComCU, CitComS)? Is this problem currently doable within CIG?
- How to add to CIG if improvement (e.g. for magma migration)?
- How to motivate Scientists to contribute?

Example #2: CIG interactions with external software development proposals

- Scenario: Individual PI writes a proposal for software development independent of CIG but in-line with long-term CIG goals.
- Several flavors:
 1. Wants CIG endorsement but not workshopped or presented as a formal CIG proposal.
 2. Ignores CIG in proposal and gets negative reviews
 - “Not working with CIG”
 - Worst case: “CIG is already doing it” when CIG is not
- What are mechanisms to encourage innovation, not stifle it?
- We also want to leverage external funding from non EAR sources (OCE, CISE, DMS, DOE?)

Example #2: CIG interactions with external software development proposals

- Suggestions/mechanisms
 - Clear list of CIG Priorities on web-site
 - Short term [priorities](#) already available
 - Long term priorities in strategic plan but should probably be bulleted
 - Instructions to PI's on how to clarify their level of interaction
 - Provide clear guidance to NSF for evaluating external proposals that interact with CIG.
 - Expected to be a continuing issue as we make progress.

Other Community related issues

- Hardware support (not strictly CIG)
 - TeraGrid
 - Peta-Scale initiatives (NCAR, SCEC, UCSD...)
 - Need to firm up partnerships, as well as get better understanding of NSF commitment/strategy for hardware access.

Successes in CIG Community Building

- Community driven software initiatives
 - GALE: direct response to long-term tectonics meeting
 - PyLith: Product of SCEC, then CIG workshops
 - "1-D" Geophysics: Seismology and Global dynamics
 - Compressible Convection: from mantle convection ws.
 - Contributed codes/[Software repository](#): CitcomS, CitcomCU, SpecFem3D
- CIG: [Open Software Policy](#)
 - Maintain repository with open checkout policy
 - All CIG software "open source": ©CalTech distributed under Gnu General Public License (Gnu GPL)
 - All co-developed software shared copyright with GPL wherever possible
 - Modification of Pre-existing or contributed software: establish © and enter into agreement for open-source license. Minimum free for non-commercial use
 - Maintain flexibility to maximize community benefit