CIG Visit to the NSF Nov 28 2007

Marc Spiegelman
(Columbia Univ., chair ExComm)

Computational Infrastructure for Geodynamics

National Science Foundation
WHERE DISCOVERIES BEGIN

With

- Brad Aagaard (USGS Menlo Park) - Chair Science Steering Committee
- Omar Ghattas (UT Austin) - Science Steering committee
- Michael Gurnis (Caltech) - Director of CIG
- Peter Olson (JHU) - Executive committee (and previous Chair SSC)

Outline

- What is CIG? - Overview and Objectives
- A Brief History of CIG
- CIG Accomplishments
- Intermediate and Long-term Goals
- Achieving our Long-term goals
  - Challenges
  - Opportunities
- Additional Information (Committees, members, working groups, codes, pretty pictures)

What is CIG?

- A Community driven organization to leverage advanced scientific computation for the benefit of the geoscience community.
- Enables science for a wide range of problems including
  - Mantle convection
  - The Geodynamo
  - Magma Dynamics
  - Crustal and Earthquake Dynamics
  - Seismology
- A Strategic **partnership** between Earth Sciences and computational Sciences
- A structure for developing, supporting and disseminating advanced computational tools for both developers and end-users.

What is CIG?

Key Components of the Infrastructure being developed

- A central system for modern software development
  - software repositories, bug tracking, regression testing…
- A coordinated effort to develop *reusable, interoperable, well documented, open-source* codes
- Basic building blocks (infrastructure) for assembling advanced codes (e.g solvers, meshers, data structures)
- Extension of frameworks for linking multiple codes and data (superstructure)
- A Science Gateway to the Teragrid
- Strategic partnerships with the larger world of computational science and geoinformatics
- Training and workshops for Earth Science and computational science communities

What is CIG?

- An organizational structure including
  - A ***Software Development Team*** (managed by M. Gurnis @Caltech) working on behalf of the community providing *development*, *support* & *training*
  - Disciplinary **Working Groups** providing interaction between scientists and developers
  - An elected ***Science Steering Committee*** providing community input and guidance to the SDT by identifying and balancing common needs across disciplines.
  - An elected **Executive Committee** providing oversight
- A growing and evolving set of **Software products** and services.
- A web-site: geodynamics.org for coordination and dissemination of software, services and information

CIG Objectives

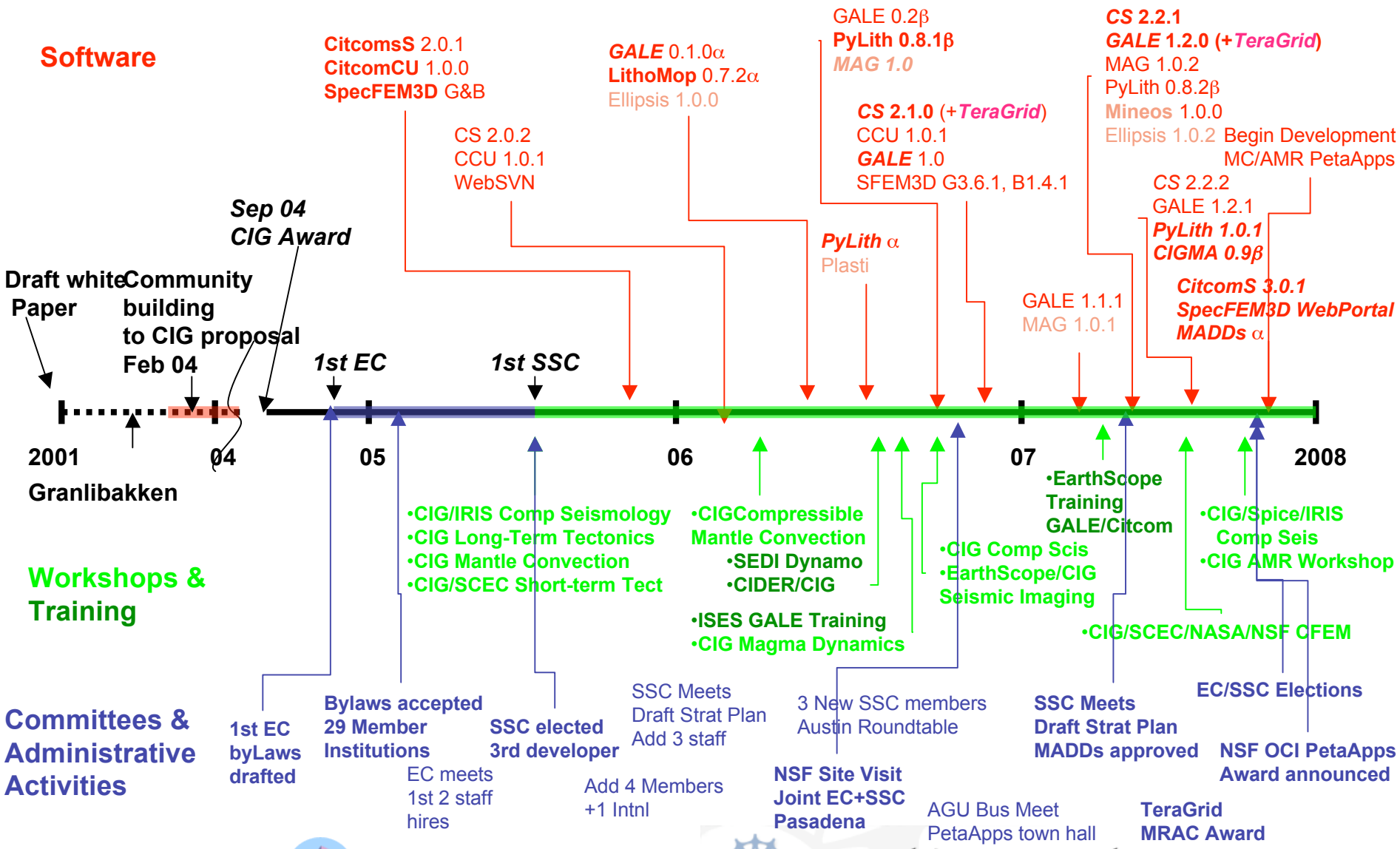
CIG Long-term goal:
to create a set of computational tools and data structures that

- Can be commonly applied within the geodynamics community.
- Promote more interaction between different geodynamic sub-disciplines.
- Enable the development of models of Earth evolution that intimately *couple* lithosphere, convecting mantle and core
- Provide the capability to eventually simulate, and *understand*, the planet as a whole.

This goal is ambitious -- but we believe we have made significant progress in the past few years towards realizing this goal and have identified several new objectives to move us into the future

The purpose of this meeting is to candidly discuss where we are, where we hope to go and what are the challenges and opportunities.

A brief history of CIG



CIG Accomplishments

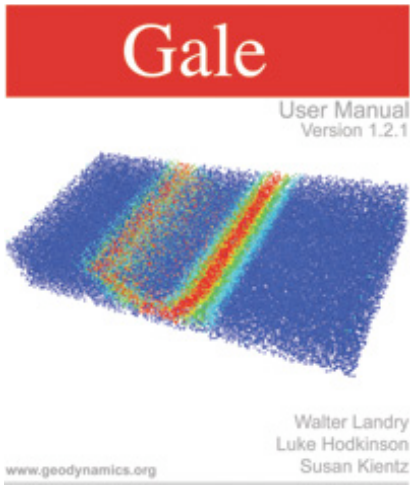
Software Development Infrastructure

- Software repositories (SVN and Mercurial)
- Web-based Bug tracking system (Roundup)
- Nightly Regression testing (CIG-Regresstor)
- CIG sponsored -- new software infrastructure
  - Sieve (Matt Knepely, Dmitry Karpeev ANL) for parallel unstructured mesh handling
  - CIGMA (Luis Almendierez) CIG-model analyzer (tool for FEM model intercomparison)

CIG Accomplishments

Long-Term Tectonics

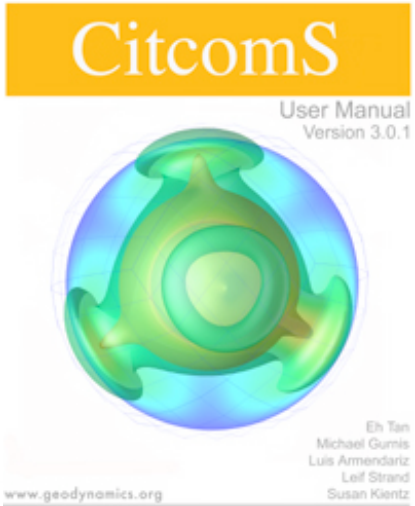
- Development of GALE
  - Solves problems related to orogenesis, rifting, subduction and coupling to surface erosion models
  - Initiated in response to 2005 Breckenridge workshop: first **publicly available** lithospheric deformation code
  - Current version 1.2.1 (release July 2007) solves
    - Shortening, extension, subduction and thermal evolution
    - 2-D/3-D scales to 100's procs (Underworld/stG *VPAC/Monash*)
    - Available through MRAC on the **TeraGrid**
  - Training sessions: Earthscope (March 2007), CFEM (June 2007), ISES Summer School
- Issues:
  - building a core user group
  - Benchmarking (with GEOMOD)
- Long-term tectonics working group approved Nov 2007



CIG Accomplishments

Mantle Convection

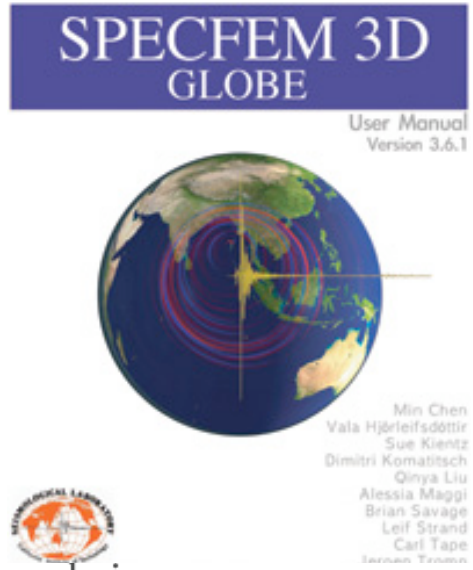
- Continued Development of CitcomS
  - Solves global and regional 3-D spherical mantle convection
  - Current version 3.0.1 (release Nov 2007)
    - Solves **Compressible** and Incompressible Thermal Convection (outcome of Compressible convection WG)
    - Thermo-chemical convection
    - Coupled global and regional models
    - Variable viscosity (T-dep, non-newtonian, pseudo-plasticity)
    - Available through MRAC on the **TeraGrid**
- Future development
  - Integrating thermodynamics (Stixrude & Lithgow-Bertelloni)
  - Integrating with seismic models
  - Continued development of Benchmarks



CIG Accomplishments

Computational Seismology

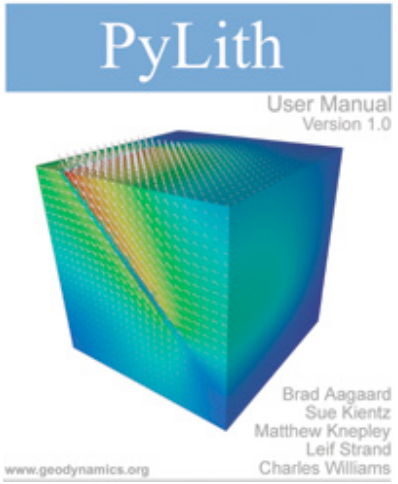
- Continued Development of SpecFEM3D Globe and Basin
  - Full 3-D Spectral Element wave-propagation codes for global and regional synthetic seismograms
  - Latest versions G3.6.1 B1.4.1 from Tromp but available through CIG
  - New Web-portal for on-demand synthetic seismograms from SpecFEM3D Globe developed by CIG with computation on the TeraGrid
- Support of Mineos
  - "1-D" fast normal-mode code
- Two workshops
  - Oct 2006 Earthscope/CIG Seismic Imaging
  - Oct 2007 CIG/SPICE/IRIS



CIG Accomplishments

Short Time-Scale Tectonics - EQ physics

- Release of PyLith 1.0
  - Significant rewrite of two legacy codes using new computational methods (Pyre, Sieve, PETSc) for parallel unstructured models
  - Solves both quasi-static visco-elastic crustal deformation and *dynamic* Earthquake rupture (EqSim)
  - Modern Modular code, built with regression testing
- Extensive use in 2006 and 2007 CFEM Workshops
- Strong and active working group
- Developing common Benchmarks for EQ codes (e.g. NASA GeoFest)



CIG Accomplishments

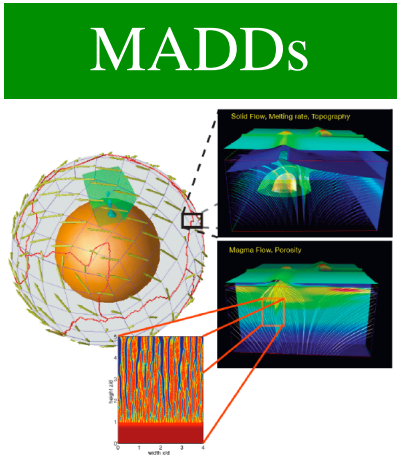
Geodynamo

- Community Dynamo code MAG
  - Contributed code new release with manual
  - Includes enhanced visualization tools
  - Currently serial code
- Issues (Olson)
  - Additional contributed code MOSST (Kuang)
  - Some difficulty benchmarking and releasing
  - Need new effort by Dynamo community
  - Bruce Buffett (SSC) may lead
- Dynamo issues discussed at SEDI meeting (2006)

CIG Accomplishments

Magma Dynamics

- Relative Late comer
- Aug 2006 Magma Dynamics Workshop (Columbia Univ)
  - Developed Strategy, Benchmark Document and Magma Dynamics Working Group
  - Proposed MAgma Dynamics Demonstration Suite (MADDs) for developing coupled mantle/magma dynamics codes
  - MADDs accepted (April 2007) partnership between MDWG and VPAC (same code base as GALE)
  - Hg Software repository implemented at CIG
  - First Benchmark Milestone Nov 2007 for accurate pressure solutions in Stoke's flow
- Issues
  - First "integrated code" requires interoperation with other CIG solids codes (MC, GALE)
  - Highly Coupled Multi-scale/Multi-physics problem
  - Coupling adds significant complexity
  - Integration with Thermodynamics and Chemical databases (e.g. EarthChem)



CIG Accomplishments

Computational Science

- Two Major workshops bringing together Earth Science and Computational Science Communities
  - Oct 2006 Austin Texas: *Challenges and Opportunities at the Interfaces of Scientific Computing and Computational Geodynamics*
  - Oct 2007 Boulder Colorado: CIG/AMR Workshop on Adaptive Mesh Refinement
    - *AMR* identified as major long-term goal for nearly all CIG projects
    - Workshop provided hands on experience with deal.ii libraries and discussed overall strategies for advances in both Earth Science and computational science

CIG Goals

	Short-Term Goals	*Intermediate-Term*	*Long-Term Goals*
Computational Seismology	✔ •Demonstrate Seismology Science Gateway using SPECFEM on TeraGrid. ✔ •Organize computational seismology workshop in New Hampshire with Spice and IRIS	✔ •Refine and deploy for general use Seismology Science Gateway. •Highest priority for two packages: a 1D package for Cartesian reflectivity, and a finite difference 3D wave propagation code. Goal will be more specifically defined after SPICE meeting.	•Continued refinement of Seismology Science Gateway with multiple codes. •Coordination of model databases.
Computational Science	✔ •Organization of a workshop in Boulder, October 2007, to familiarize the community with adaptive mesh refinement techniques and to jump-start the development of AMR-enabled codes.	•Incorporation of AMR into a new generation of codes. Further integration with existing libraries for parallel computations.	•Work on integration of computer sciences and geophysics into single framework.
Short Time-Scale Tectonics	~ •Complete merging of EqSim and PyLith 0.8 into PyLith 1.0.	•Direct support for computing Green's functions in PyLith. • Coupling short-term simulations to other simulations in order to accurately capture interactions among geodynamics phenomena.	•Adaptive mesh refinement capabilities in PyLith or a code with similar functionality. Integrate tools for formal data assimilation.
Long Time-Scale Tectonics	~ •Recruit members to a larger, more active Working Group in long-term tectonics. ? •Incorporate frictional BCs into Gale. ? •Hold Gale training session, Fall 2007	•Receive donation of SNAC, port to CIG build procedure and refine Users Manual for general use. •Start incorporation of AMR in code for long time-scale tectonics	Develop new code with adaptive mesh refinement.

CIG Goals

	Short-Term Goals	*Intermediate-Term*	*Long-Term Goals*
Magma Migration	•Implementation of Madds using StGermain, PETSc, and Comsol Multiphysics. •Calibration of pressure solvers and solitary wave benchmarks.	•Training of user base in CIG software with Madds. • Implementation of Madds with AMR.	•Integration of magma dynamics in global mantle convection. •Coupling of thermodynamics modeling and Madds-enabled software.
Geodynamo	•Benchmark serial geodynamo code, MAG • Development of user interfaces. •Preliminary community building discussions will be held during SEDI in July 2008.	•Community building workshop in 2008, to utilize and train users on the two serial codes, possibly partnering with other organizations.	•New dynamo code using common components with Mantle Convection codes including common Earth structure framework and grid and adaptivity.
Mantle Convection	•Release of compressible mantle convection CitcomS by September 2007. •Add benchmark cases with high Ra, internal heating and time- dependent solutions by the end of 2007. •Add compressibility and the geoid calculations to the HC code.	•Develop convection codes with AMR capability. **Begun development with NSF PetaApps Award**	•Develop convection codes with AMR capability that share basic components with other geodynamics codes and that are scalable to 100s and 1000s of processors.
Organizing Community Participation	•Further recruitment of members for working groups..particularly early career investigators •Continue support for workshops and committee meetings to increase community participation in CIG		
User Training	•Increase adoption of CIG developed software by the general Earth Science Community, in particular geophysicists •Increase training and documentation •Work on variable mechanisms for increasing buy-in: small workshops, visits to CIG, production of training manuals, web-site (web-portals)		

CIG Goals

	Short-Term Goals	*Intermediate-Term*	*Long-Term Goals*
Computational Seismology	•Demonstrate Seismology Science Gateway using SPECFEM on TeraGrid. •Organize computational seismology workshop in New Hampshire with Spice and IRIS	•Refine and deploy for general use Seismology Science Gateway. •Highest priority for two packages: a 1D package for Cartesian reflectivity, and a finite difference 3D wave propagation code. Goal will be more specifically defined after SPICE meeting.	•Continued refinement of Seismology Science Gateway with multiple codes. •Coordination of model databases.
Computational Science	•Organization of a workshop in Boulder, October 2007, to familiarize the community with adaptive mesh refinement techniques and to jump-start the development of AMR-enabled codes.	•Incorporation of AMR into a new generation of codes. Further integration with existing libraries for parallel computations.	•Work on integration of computer sciences and geophysics into single framework.
Short Time-Scale Tectonics	•Complete merging of EqSim and PyLith 0.8 into PyLith 1.0.	•Direct support for computing Green's functions in PyLith. • Coupling short-term simulations to other simulations in order to accurately capture interactions among geodynamics phenomena.	•Adaptive mesh refinement capabilities in PyLith or a code with similar functionality. Integrate tools for formal data assimilation.
Long Time-Scale Tectonics	•Recruit members to a larger, more active Working Group in long-term tectonics. •Incorporate frictional BCs into Gale. •Hold Gale training session, Fall 2007	•Receive donation of SNAC, port to CIG build procedure and refine Users Manual for general use. •Start incorporation of AMR in code for long time-scale tectonics	Develop new code with adaptive mesh refinement.

CIG Goals

	Short-Term Goals	*Intermediate-Term*	*Long-Term Goals*
Magma Migration	•Implementation of Madds using StGermain, PETSc, and Comsol Multiphysics. •Calibration of pressure solvers and solitary wave benchmarks.	•Training of user base in CIG software with Madds. • Implementation of Madds with AMR.	•Integration of magma dynamics in global mantle convection. •Coupling of thermodynamics modeling and Madds-enabled software.
Geodynamo	•Benchmark serial geodynamo code, MAG • Development of user interfaces. •Preliminary community building discussions will be held during SEDI in July 2008.	•Community building workshop in 2008, to utilize and train users on the two serial codes, possibly partnering with other organizations.	•New dynamo code using common components with Mantle Convection codes including common Earth structure framework and grid and adaptivity.
Mantle Convection	•Release of compressible mantle convection CitcomS by September 2007. ?Add benchmark cases with high Ra, internal heating and time- dependent solutions by the end of 2007. ?Add compressibility and the geoid calculations to the HC code.	•Develop convection codes with AMR capability. **Begun development with NSF PetaApps Award**	•Develop convection codes with AMR capability that share basic components with other geodynamics codes and that are scalable to 100s and 1000s of processors.
Organizing Community Participation	•Further recruitment of members for working groups..particularly early career investigators •Continue support for workshops and committee meetings to increase community participation in CIG		
User Training	•Increase adoption of CIG developed software by the general Earth Science Community, in particular geophysicists •Increase training and documentation •Work on variable mechanisms for increasing buy-in: small workshops, visits to CIG, production of training manuals, web-site (web-portals)		

CIG Long Range Goals

Specific Goals ~3-5year Term

- ***On-Demand/On-Request Seismology***
  - Develop web interfaces (Science Gateway) to High Performance Seismic wave propagation codes.
  - Development of common Earth Structure Model data-structures
- **Petascale Computing**
  - Position CIG to take advantage of PetaScale resources (both capacity and capability computing)
  - Provide computational challenges for petascale

CIG Long Range Goals

Specific Goals ~3-5year Term

- *** Coupled and Whole Earth Models ***
  - **Integration** and **Interoperation** of modeling capabilities to enable the exploration (and understanding of) multi-physics/multi-scale coupled problems such as
    - Plate Boundary processes - Coupled Lithosphere/Asthenosphere/Magma Dynamics
    - Interaction of small-scale localization on global dynamics (plate boundary interaction with global mantle convection)
    - Coupled Mantle convection- Geodynamo
    - Geochemical evolution of the planet
    - Long-term Earthquake cycle - dynamics of fault networks
    - Your problem here!
  - Significant algorithmic, software and Earth Science challenges but potentially revolutionary.

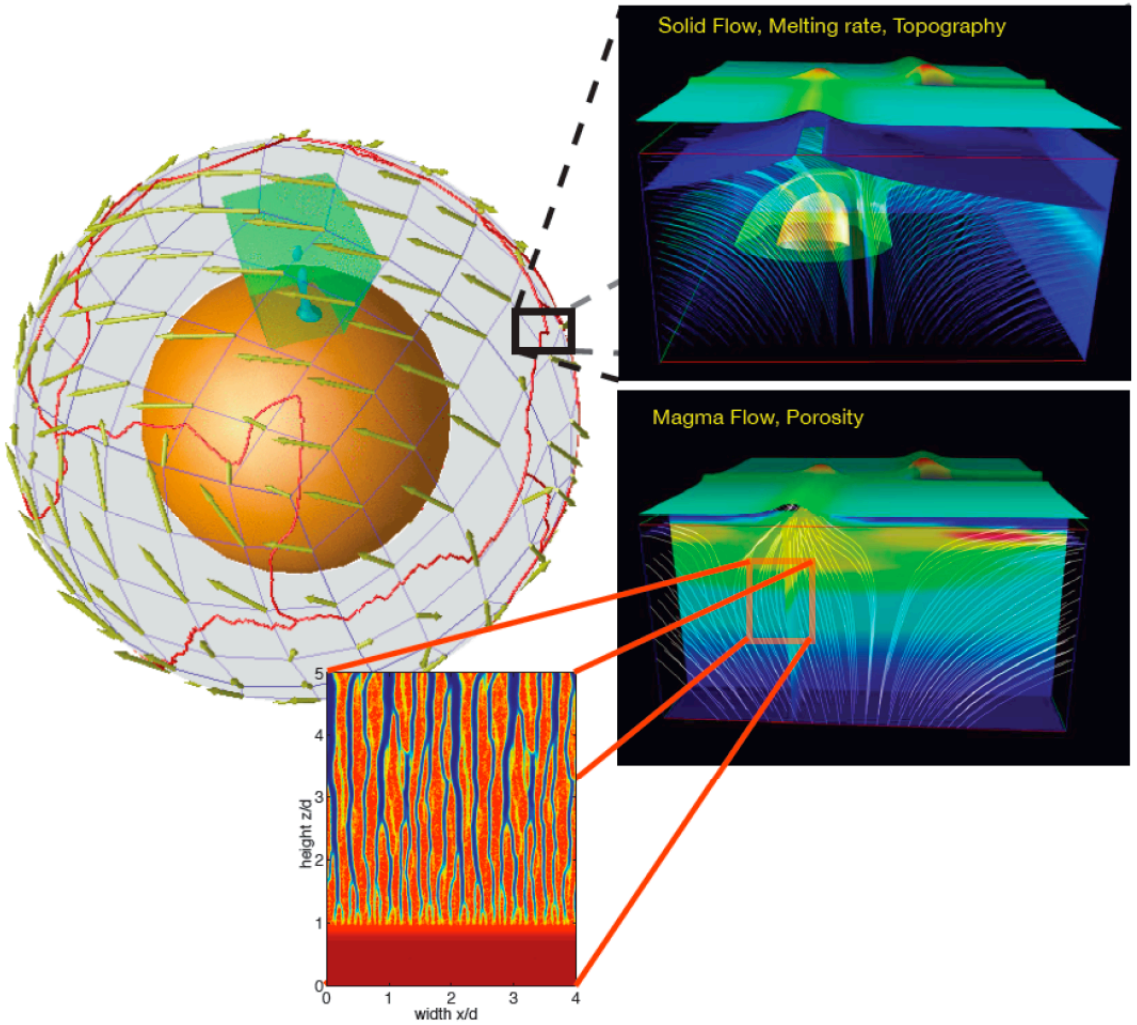
CIG Long Range Goals

Specific Goals ~3-5year Term

- **Earth Structure Model Frameworks**
  - Development of common interfaces to growing Earth Science Data and meta-data e.g.
    - Global seismic models
    - Thermodynamic data and models
    - Material properties/Experimental data
    - Geochemical/Petrologic data

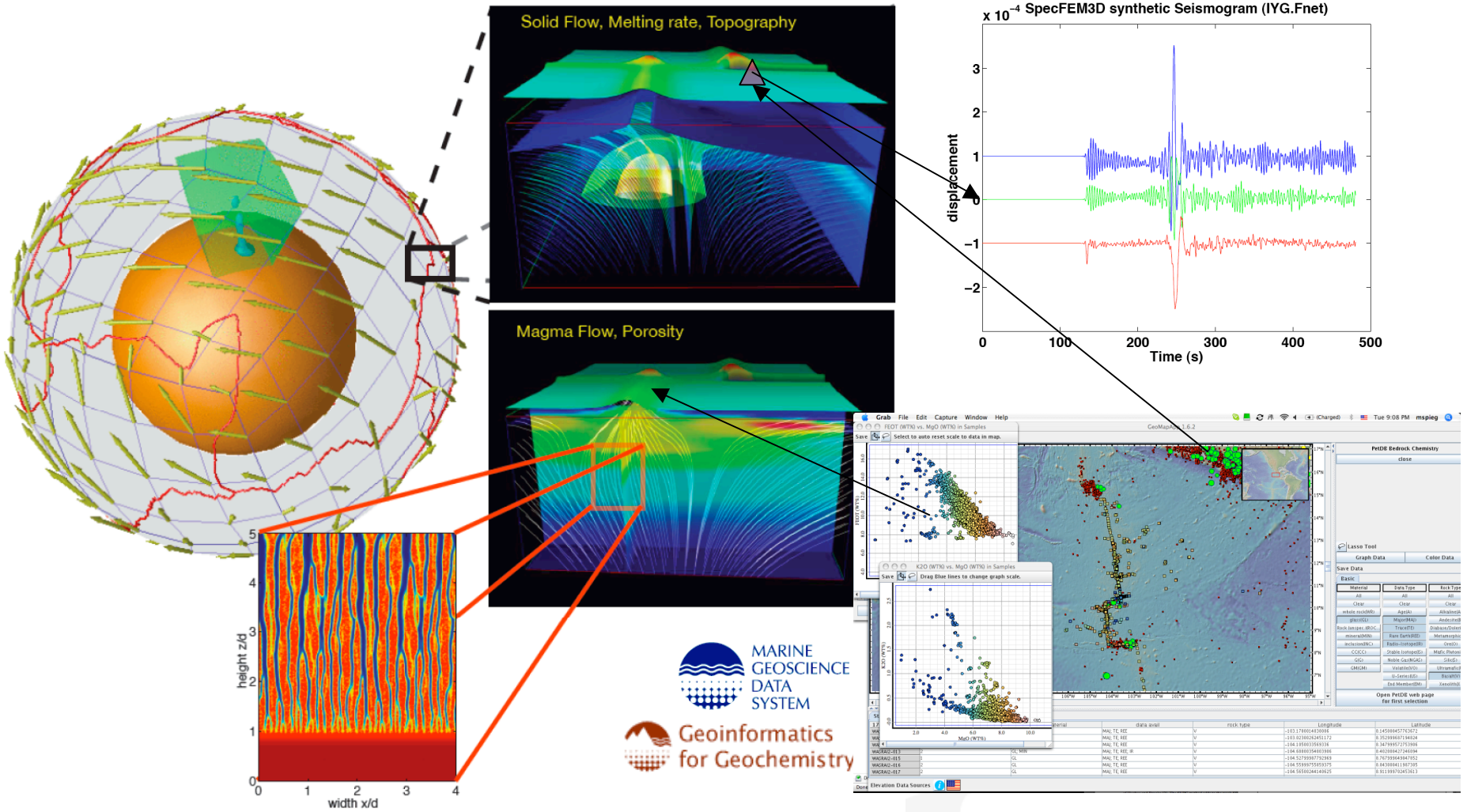
The Future of CIG?

Coupled Multi-Scale, multi-physics models



Future Directions

Coupled Multi-Scale, multi-physics models



Towards Achieving our long Term Goals

Last 2-3 years has partly been "experimental phase" in trying and adapting new technologies

- Initially in support of individual sub-disciplines
- Outwardly not obviously reusable/interoperable
- However many codes do share common components
  - Mantle Convection: CitcomS 3.0.1
    - Pyre, HDF5
  - Magma Dynamics: MADDsα
    - StGermain/Underworld (VPAC/Monash), PETSc
  - Short Term Deformation: Pylith 1.0.1
    - Pyre, PETSc, Sieve, VTK
  - Long-Term tectonics: Gale 1.2.1
    - StGermain/Underworld (VPAC/Monash), PETSc
  - Seismology: SpecFEM3D
    - Pyre, Custom F90 (+ web portal)
  - GeoDynamo: Mag1.0
    - custom

Towards Achieving our long Term Goals

- Need to evaluate current technologies as well as being open to new ones (Octree, AMR, Sieve)
- Need to begin to stress underlying reusable components
  - Example: Robust, scalable, AMR Variable Viscosity Stokes solver (with good pressures) is essential to at least
    - Mantle Convection
    - Magma Dynamics
    - Lithospheric deformation
  - AMR requested across all sub-disciplines
  - Regularization of IO and interfaces would be extremely useful
    - Additional Web Portals?
- Need to address fundamental scientific and computational challenges inherent in multi-scale/multi-physics problems (see Omar's talk)
- These are hard problems and will need a continuing collaboration between GeoScientists and Computational Scientists.

Towards Achieving our long Term Goals

Hurdles and potential pit falls

- Fundamental Complexity
  - Of coupled physics
  - Managing uncertainty (parameters, constitutive relationships, IC's, well-posedness)
  - Of modern hierarchical software
  - Algorithmic Complexity (e.g. generality of pre-conditioners for coupled problems?)
- Too high a cost of buy-in…
- Potential misunderstanding/mistrust of simulation
- Difficulty in mapping observations to process (really inverse theory).

Some potential steps towards avoiding these

- Minimize initial buy-in : WebPortals/CookBooks/PreBuilt TeraApps (at what cost to SDT)
- Significant Training and Workshops
- General Education in Computational methods for Earth Scientists to understand what these tools can and can't do
- Making models products be comparable to data…

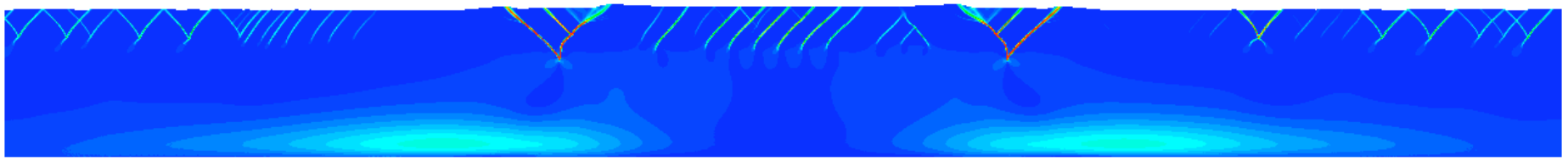
Final Thoughts

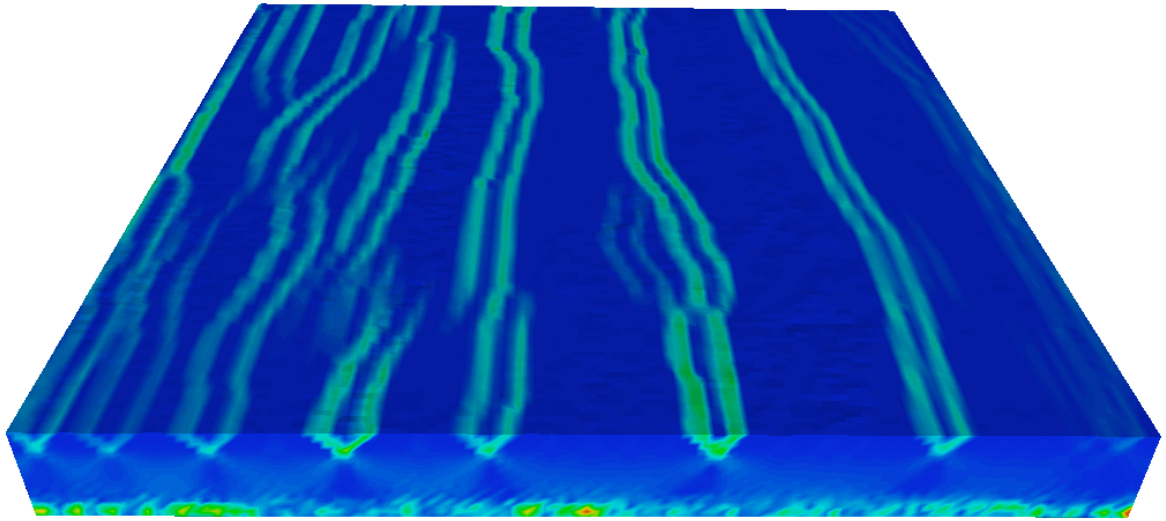
- In 2 years, CIG has accomplished quite a lot
  - Begun the transformation of computation across the Solid Earth Sciences
  - Has delivered *documented, accessible, maintained* code for a wide variety of disciplines
  - Formed serious collaborations between CS and ES
  - CS drives new ES capabilities, ES provides significant Computational and Mathematical Challenges
  - Eased access to significant computational resources
- The next phase is probably harder
  - But we've gained a lot of collective experience
  - Our understanding of the complexities has matured considerably
  - Will require increasing partnerships with Math, Computational Science, Earth Sciences
  - We're well positioned to move forward

Additional Slides

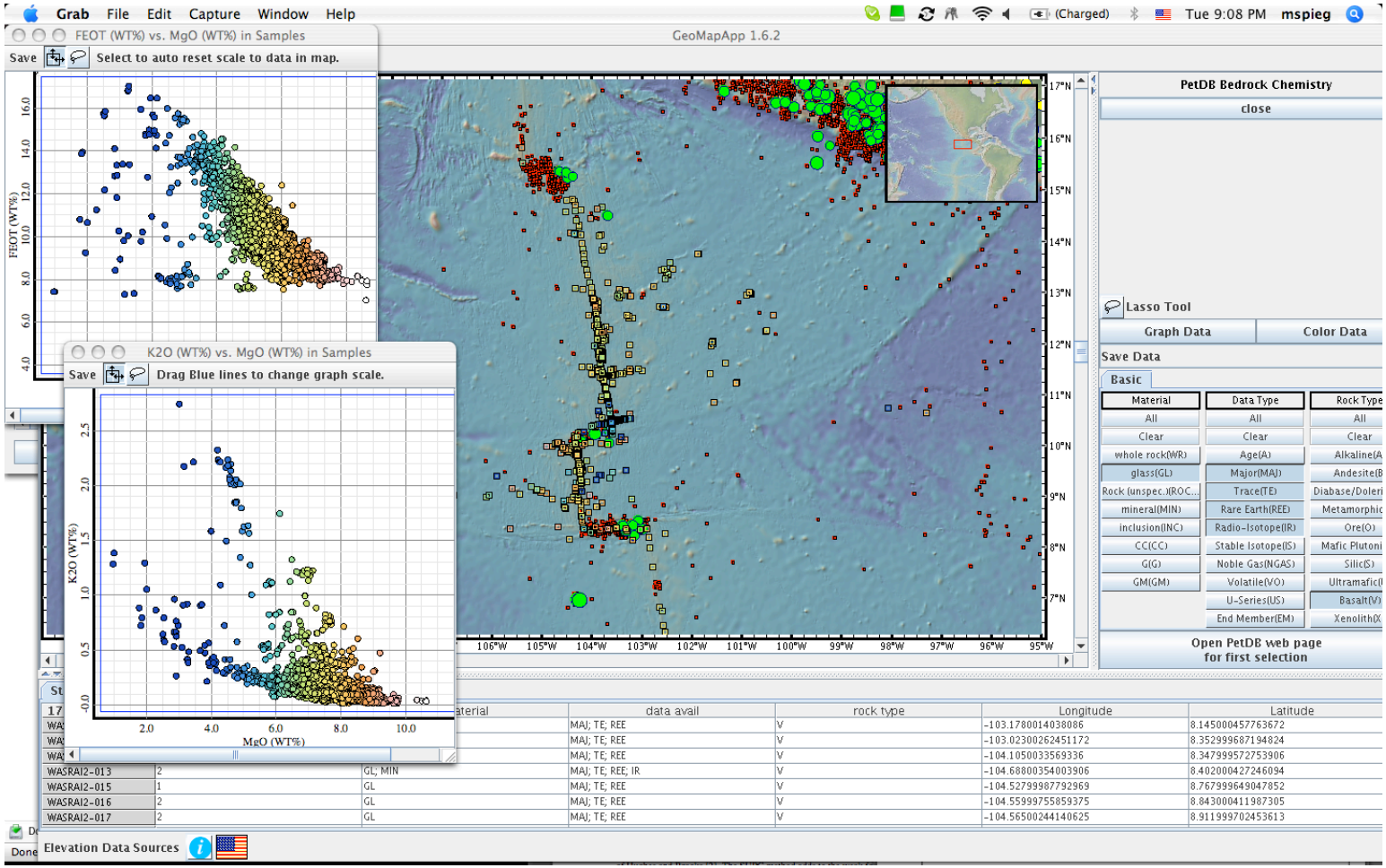
CIG Accomplishments

Long-Term Tectonics: Recent results from Gale 2-D/3-D extension Benchmarks





Interfacing Models with GeoData



MARINE GEOSCIENCE DATA SYSTEM

Geoinformatics for Geochemistry

CIG Institutional Members

CIG Institutional Members: Universities (37)

Arizona University
Brown University
California Institute of Technology
Colorado School of Mines
Colorado State University
Columbia University
Cornell University
Georgia Institute of Technology
Harvard University
Johns Hopkins University
Massachusetts Institute of Technology
Oregon State University
Pennsylvania State University
Princeton University
Purdue University
Rensselaer Polytechnic Institute
State University of New York at Buffalo
State University of New York at Stony Brook
University of California, Berkeley
University of California, Davis
University of California, Los Angeles
University of California San Diego
University of Colorado
University of Hawaii
University of Maine
University of Maryland
University of Michigan
University of Minnesota
University of Missouri-Columbia
University of Nevada, Reno
University of Oregon
University of Southern California
University of Texas at Austin
University of Washington
Virginia Polytechnic Institute and State University
Washington University
Woods Hole Oceanographic Institution

CIG Institutional Members: National Labs/USGS (4)

- Argonne National Laboratory (MSC)
- Lawrence Livermore National Laboratory
- Los Alamos National Laboratory (ES)
- U.S. Geological Survey (Menlo Park)

CIG Foreign Affiliates (9)

- Australian National University
- Geological Survey of Norway (NGU)
- GNS Science
- Monash University
- Munich University LMU
- University College London
- University of Science and Technology of China
- University of Sydney
- Victorian Partnership for Advanced Computing

Current CIG committees (Oct 2007)

- Executive Committee
  - Marc Spiegelman (Columbia) - Chair
  - Brad Hager (MIT) -Vice-chair
  - Alan Levander (Rice)
  - Carolina Lithgow-Bertelloni (UCL)
  - Peter Olson (JHU)
  - Michael Aivazis (Ex Officio)
  - Michael Gurnis (Ex Officio)
- Science Steering Committee
  - Brad Aagaard, , (June 2008), USGS
  - Wolfgang Bangerth, , (June 2008), Texas A&M
  - Bruce Buffett, , (Sept. 2010), U Chicago
  - Omar Ghattas, , (Sept. 2010), UT Austin
  - Louise Kellogg, , (Sept. 2009), UC Davis
  - Laurent Montesi, , (Sept. 2009), U Maryland
  - Jeroen Tromp, , (Sept. 2010), Caltech
  - Shijie Zhong, , (June 2008), UC Boulder
  - Marc Spiegelman, Columbia University, Ex officio

CIG Core Constituencies

- CIG is developing software for a wide range of problems in the Earth Sciences including
  - Mantle Dynamics
    - Mantle Convection
    - Magma Dynamics
  - Crustal and Lithospheric Deformation
    - Short Term Deformation & Earthquake physics
    - Long- Term tectonics
  - Seismology
  - The Geodynamo

- Each discipline supported by active working groups and software

Principal CIG Codes

- Mantle Dynamics
  - Mantle Convection
    - CitcomS 3.0.1 - parallel spherical -global and regional compressible convection
  - Magma Dynamics
    - MADDs (MAgma Dynamics Demonstration suite)
- Crustal and Lithospheric Deformation
  - Short Term Deformation & Earthquake physics
    - PyLith 1.0.1 - quasi-static and dynamic EQ simulation in 1,2,3D (parallel unstructured FEM)
  - Long- Term tectonics
    - GALE 1.2.1 - Complex rheology 3-D Visco-Plastic
- Seismology
    - SpecFEM3D Global and Basin - Spectral Element synthetic seismogram code (+ web portal)
- GeoDynamo
    - MAG 1.0 - Serial Spectral Geodynamo code