Computational Infrastructure for Geodynamics (CIG)

NSF Award Number EAR-0426271

Five-year Strategic Plan: Sept. 1, 2006 to Aug. 31, 2011

Submitted July 17, 2006

Table of Contents

Community Oversight

Executive Committee (EC)

Mark A. Richards, (Chairman, Oct., 2007), University of California, Berkeley

Marc Spiegelman (Vice Chairman , Oct., 2006), Columbia University

Bill Appelbe (At-Large, Oct., 2007), Victorian Partnership for Advanced Computing

Bradford H. Hager (At-Large, Oct., 2008), Massachusetts Institute of Technology

E. Marc Parmentier (*ex-officio*, June 2006) Brown University

Michael Gurnis (*ex-officio)* California Institute of Technology

Mchael Aivazis (*ex-officio)* California Institute of Technology

Science Steering Committee (SSC)

E. Marc Parmentier (Chairman, June 2006), Brown University

Brad Aagaard (June 2008), USGS, Menlo Park

Wolfgang Bangerth (June 2008), Texas A&M University

Roger Buck (June 2006) Lamont-Doherty Earth Observatory of Columbia University

Omar Ghattas (June 2007), The University of Texas at Austin

Peter Olson (June 2007), John Hopkins University

Jeroen Tromp (June 2007), California Institute of Technology

Shijie Zhong (June 2008), University of Colorado at Boulder

1. An Overview of CIG

The Computational Infrastructure for Geodynamics (CIG) develops, supports, and disseminates software for the geoscience community from model developers to end-users. The software is being developed for problems ranging widely from mantle and core dynamics, crustal and earthquake dynamics, seismology, and we plan to implement magma dynamics in the next phase of software development. With a high level of community participation, CIG leverages the state-of-the-art in scientific computing into a suite of open-source tools and codes. The infrastructure now under development consists of:

- a central system facilitating software development, including a repository, a bug tracking system, and automatic regression testing;
- a coordinated effort to develop reusable, well documented and open-source geodynamics software;
- the basic building blocks – an infrastructure layer – of software by which state-of-the-art modeling codes can be quickly assembled;
- extension of existing software frameworks to interlink multiple codes and data through a superstructure layer;
- strategic partnerships with the larger world of computational science and geoinformatics;
- specialized training and workshops for both the geodynamics and larger Earth science communities.

CIG has established a small team of dedicated software architects and engineers whose work is guided by scientific objectives formulated by the scientific community. The Software Development Team (SDT) provides software services to the community in terms of programming, documentation, training, and support. Guidance for the programmers comes from a Science Steering Committee (SSC) and an Executive Committee (EC) whose emphasis is to identify and balance common needs across disciplines.

Since the start of project in September, 2004, the CIG staff, committee members, and members of the community have been diligently moving the project forward. The staff has brought online a robust set of tools for our software repository, bug tracking, and automatic build system, all available through our web site (http://geodynamics.org). Already, we have been able to develop and release software for several different topics, including mantle convection, short-time scale tectonics and long-time scale tectonics. During the next year, we will have completed all of our short-term goals, made substantial progress on our intermediate term goals, and started work on moving toward our long term goal of developing new software for several communities simultaneously with common components.

This year's Strategic Plan (SP) was developed by the Science Steering Committee (SSC) with assistance provided by the CIG staff. Following the Feb. 28, 2006 site visit by NSF at CIG, the SSC held several telecoms at which an approach to the formulation of the SP was discussed. Each committee member polled different subdisciplines of the CIG communities and submitted written descriptions that included accomplishments and goals organized over the short-, intermediate-, and long-terms. A SSC meeting at CIG in Pasadena May 18-19 was held specifically to develop this SP. During the meeting, we reviewed our accomplishments and had extensive discussions on our goals while attempting to identify common themes in different disciplines. This document emerged from those discussions and was then made available for public comment June 28-July 10. This Strategic Plan was approved by unanimous consent by the CIG Executive Committee on July 10, 2006.

2. Our Long-range goals

The long-term goal of CIG is to create a set of computational tools and data structures that can be commonly applied within the geodynamics community. These tools and data structures will promote more interaction between different geodynamics sub-disciplines. A common set of computational tools will enable the development of models of Earth evolution that intimately couple lithosphere, convecting mantle and core, with the capability eventually to simulate the planet as a whole. Coupled models of the dynamics of lithosphere, the convecting mantle, and the core dynamo along with an understanding of how these relate to whole Earth structure is an ultimate goal of the geodynamics community. In striving toward these goals, the community has identified the following long term (3-5 year) goals.

On-demand/on-request seismology. With the availability of wave propagation codes such as SPECFEM3D, the simulation of seismograms with realistic whole earth models has become routine. SPECFEM3D is available to users from the CIG repository. A next step is the development of web interfaces that make these simulations easy and routine for the seismological community. Part of such an effort will be to incorporate seismic velocity and attenuation structures of the earth derived using various global tomography methods. The development of common data structures will be needed for representing these earth structure models. This may be regarded as a first step toward more general earth structure model frameworks discussed below.

Earth Structure Model Frameworks. Unified data structures for representing the physical And chemical propterties of the Earth will be of immediate value across all the CIG disciplines. We view such frameworks as essential first components for developing and interpreting coupled simulations of the whole earth. A variety of data types will be important; perhaps the first being global models of seismic velocities, density, and attenuation, plus the physical properties derived from these such as hydrostatic pressure and gravity. In addition we intend to add other Earth data such as thermodynamic and transport properties, composition models, plate motions, plate boundary type and location, seismicity, crustal stress distributions, properties that are essential for defining dynamical models and for comparing the output of our numerical models with the Earth. In addition to the three spatial dimensions, a fourth time dimension will also be important, for example geologic reconstructions of the history of plate boundaries.

The data structures listed above most commonly consist of sets of prescribed values of some parameter at fixed points in space and time. Another type of data structure may be the output of programs like MELTS in which an empirically derived database of free energies of chemical components is used to predict the amount and composition of melt that can be derived from a given solid composition at a temperature and pressure prescribed from other components of a coupled earth simulation.

It will be essential that CIG collaborate with other Earth science community efforts in this regard. Other initiatives, such as IRIS, SCEC, GEON, and others are more focused on data. However, CIG should work with others so that the methodologies of an Earth structure framework are appropriate for constructing dynamic models and linking those models with data.

*Coupled and whole earth (planet) models***.** The wide range of physical processes involved in studying the evolution of our planet as a whole has led to geodynamics studies that usually treat individually the dynamics of the lithosphere, the convecting mantle, and the core. While couplings between the behavior of different parts of the earth is recognized to be fundamental, for example the thermal structure of the core will be controlled by the convecting mantle and the dynamics of the lithosphere through the development of plate boundaries will control mantle convection; these couplings have been treated in only limited ways. As in many highly nonlinear problems, the couplings involve strong feedbacks: the dynamics of the lithosphere not only influences convection in the underlying mantle, but mantle convection exerts forces responsible for the dynamics of the lithosphere.

Processes that operate on a wide range of scales is another fundamental feature of the lithosphere, convecting mantle, and core. For example deformation of the lithosphere on large scales can be localized into narrow fault zones; mantle dynamics with strongly temperature-dependent rheology involves the instability of very thin thermal-rheological boundary layers. Magmatic and fluid processes are also examples of multi-scale coupled processes that significantly affect large scale planetary behavior, such as the influence of magmatism on the dynamics and rheology of narrow plate boundary regions. Fundamentally new approaches and codes are needed to treat such problems.

Adaptive mesh refinement is an established but still evolving approach to treat problems with processes operating at very different scales. Therefore, as current projects come to completion, CIG will undertake the development of a framework of codes that will address the full range of coupled problems needed to understand the earth. Such codes would be formulated with data structures suitable for adaptive mesh refinement and would take advantage, as much as possible, of existing software like equation solvers.

3. CIG accomplishments

3.1 Development Infrastructure. An important aim is to introduce good software design practices into the large-scale software development efforts at CIG. This includes, for example, techniques for automated build and test procedures, development of benchmarks and test cases, and documentation.

The software repository and attendant web site are central to CIG's objectives of facilitating collaboration and sharing of validated open-source software and reuseable components. The repository is critical to bring modern software engineering practices to our community and CIG's software development team. We now have a single repository for developer use that manages multiple developers working concurrently on modular software components shared through the repository. For the CIG software repository, we use the open source package Subversion (SVN). The entire contents of the repository are navigable from our Web site. Users can either directly check out the latest development version of a code, or they can download a "tar file". In addition, CIG provides a bug-tracking database (Roundup) to allow developers and external participants to register and comment on bugs and requests for new functionality in CIG software that can then be worked on by the developers of a program.

A key problem that faces any dynamic software repository is ensuring that "nothing breaks" despite frequent dynamic changes needed to meet the evolving scientific goals of the community. CIG uses state-of the art software engineering technology – agile computing to minimize the risks of software development for continuously evolving requirements. In particular, the repository uses unit and regression testing. Building and testing occurs either nightly or automatically in response to a software commit to the SVN repository using *CIG-Regresstor*, a collection of Python codes written by CIG engineer Luis Armendariz. This software uses Build-bot and the results of the testing are both stored in a database and made available interactively on our web site. Nightly regression testing generates an electronic report that contains the build and test failures (including the platforms on which they occurred). Regression testing allows the SDT to rapidly identify when a change in a repository component or platform has caused an error or inconsistency. Regression testing allows users of the repository to have confidence in the robustness of the repository.

As part of the general tool kit needed for the solution of many of the problems that CIG encounters, we supported the development of *Sieve* by Mathew Knepley at Argonne National Laboratory (ANL) in collaboration with Dimitry Karpeev (ANL). *Sieve* is infrastructure for storing and manipulating general finite element meshes and can be used so that a developer avoids many of the complexities associated with parallel processing. Seive is already being used in *PyLith 0.8* (see below).

Work continues on a suite of Python software for benchmark intercomparisons, *BM.py* *BM.py* will allow the results of geodynamic models to be compared against standard benchmarks and report back global and local mismatches in solutions, independent of the method of discretication. This code uses FIAT, a library of finite element basis functions that allow two geodynamic model results to be compared even if they use different meshes and basis functions. The code is initially being developed for mantle convection problems.

3.2 Long-term tectonics. After extensive correspondence with members of this community, we made available *Plasti*, a 2-D ALE code originally developed at Dalhousie University in Canada and later extensively modified and improved by Sean Willett of the University of Washington. CIG ported the code into our SVN repository and in June, 2006 released a tar file with documentation and test cases under the GNU GPL. The release of an existing 2-D ALE code with an open source license was a recommendation from the NSF workshop on Tectonic Modeling June, 2005 at Breckenridge Colorado.

Our most significant achievement in long-term tectonics has been the initial development of *GALE* (Geodynamics Augmented Lagrangian Eulerian) that solves problems related to orogenesis, rifting, and subduction with coupling to surface erosion models. This code is being developed in response to an additional recommendation from the Breckenridge workshop. In collaboration with the Victorian Partnership for Advanced Computing (VPAC), CIG engineer Walter Landry released a developers version of *GALE*; *GALE 0.1.0* was released in April and version 0.2.0 was released in June. The original goal was to develop an open source code that is at least as useful as some existing 2D codes, such as *Plasti*, with the addition of 3D capability. CIG development remains on target and we are expecting to release a fully working version and manual in October, 2006. Rob Bialas, a graduate student at LDEO of Columbia University, spent a week at CIG in Pasadena advising the SDT. We have installed the alpha release on the compute cluster at LDEO, and Bialis is working on comparisons with a FLAC based 2D code for problems of continental extension, which are among the simplest tectonic problems to numerically treat in terms of boundary conditions.

3.3 Mantle convection. CIG staff continued to maintain *CitcomS.py* a state-of-the-art mantle convection code within the *Pyre* framework. We released bug fixes *CitcomS.py 2.0.2* in February. Following the recommendation of the 2005 Boulder CO workshop, a 3D Cartesian/Regional spherical mantle convection code *CitcomCU* was released in December of 2005. Shijie Zhong worked with CIG staff engineer Luis Armendariz to make this code available. A bug fix *CitcomCU 1.0.1* was released in February. The code also has thermochemical mantle convection capability (particle-based). The release also included documentation and benchmarks of the code. CIG initiated work with HDF5, a general purpose library and file format for storing scientific data. HDF5 has extensive capabilities for parallel I/O and this allows us to solve the confusing proliferation of files resulting from large parallel runs. We have now standardized the output from our convection codes and versions *CitcomS.py 2.1.0* and *CitcomCU 1.1.0* were expected to be released on June x. Several benchmark (BM) cases for isoviscous, basal heating spherical shell convection cases were computed using *CitcomS* by Shijie Zhong and were added to an extensive list of BMs on our web site.

A working group on the coding of a flexible analytically based spectral flow code that treats radial variations in viscosity and density was formed with Thorsten Becker, Carolina Lithgow-Bertelloni, Rick O'Connell, Craig O'Neil and Bernhard Steinberger. An informal meeting was held at the 2005 Fall AGU, and a more detailed plan was prepared and sent to the CIG by Becker. Becker and Steinberger have made considerable effort in putting the code, *HC*, together and setting up a suite of benchmarks. Becker worked with the CIG staff to port the code into the SVN repository which is now being used for development and can be checked out by potential users.

Scott King and Shijie Zhong organized a special workshop on compressible mantle convection at Purdue University this Spring to help move forward the effort on compressible mantle convection code development (see the Purdue workshop report on our web site). A detailed work plan was developed in which a new code for compressible mantle convection would be developed by extending the capabilities of *CitcomS*.

3.4. Computational seismology. The first CIG/IRIS Computational Seismology Workshop was held June 8, 2005. Several community recommendations that followed from Workshop have been accomplished. One of these was Pyrization of the Masters mode catalog code MINEOS. Also spectral-element packages SPECFEM3D_BASIN and SPECFEM3D_GLOBE have been made available for download via the CIG web pages.

3.5. Short time-scale tectonics. Working group members have held a workshop each of the last four years cosponsored by various combinations of the Southern California Earthquake Center, NASA, Los Alamos National Laboratory, NSF, and CIG. These workshops have served to (1) establish a suite of benchmarks for testing codes and comparing modeling techniques, (2) train students, postdocs, and others in the use of a variety of modeling tools (including mesh generators and modeling codes), and (3) facilitate an exchange of ideas among modelers from academia, national laboratories, and government agencies.

While the proposal to form CIG was in its infancy, two members of the working group, Brad Aagaard (USGS) and Charles Williams (RPI) began working towards integrating their modeling codes (*EqSim* and a version of *Tecton*) into the *Pyre* framework with the ultimate goal of developing highly modular codes for the simulation of earthquake dynamics. A significant amount of commonality was identified between the codes and since then Aagaard and Williams have coordinated their development with a plan to merge their codes into a single suite of modules, *PyLith*. *PyLith* is now under development as discussed below. In addition, in year one CIG added Python bindings to the Portable, Extensible Toolkit for Scientific Computation (*PETSc*). This significantly reduces the amount of code required to write solvers in the *Pyre* superstructure framework using *PETSc*. Consequently, *PyLith* will use *PETSc* routines and solvers directly from the framework, an important step toward our vision of a framework with common components.

In June, 2006, CIG released *PyLith 0.8.0* as both a binary executable and a download from the SVN repository. This version can solve elastic and viscoelastic problems with faults with parallel processing afforded through the use of *Seive* and *PETSc*. *PyLith 0.8.0* was used extensively during the June 26-30 “Workshop on Community Finite Element Models for Fault Systems and Tectonic Studies” on the campus of the Colorado School of Mines that CIG partially supported.

3.6 Geodynamo. CIG started work in this area following the hiring of a new software engineer in June of 2006, Mi Wei. CIG received a donation of the *MAG* geodynamo code from Peter Olson along with post processing and visualization software. *MAG* is a basic Boussinesq convection dynamo code written by Gary Glatzmaier and adapted by Peter Olson and Uli Christensen for use on serial workstations. Peter has since been working with engineer Wei Mi. The code has been brought under source control using SVN and we expected to make a release of the software in July. In addition, we have been having discussions with the NSF TeraGrid to start developing a science gateway (web portal) for the geodynamo community using *MAG* and other software.

4. Details of our short, intermediate, and long-term goals

The Science Steering Committee has identified our short, intermediate-, and long-term goals. We hope that our short term goals would be completed within the 2006 calendar year, while making substantial progress on our intermediate goals during the project year that ends Aug. 31, 2007. Achieving our long-term goals will be more technically challenging and will involve unifying many of the efforts of the individual disciplines; our strategy for achieving our long-term goals is presented in Section 5. All of our goals are summarized in Table 4.1

4.1. Common Infrastructure. The routine comparison of model results with existing Web-accessible benchmarks is essential for increasing the overall quality of our science especially as it moves into the realm of complex, multi-physics and multi-scale simulations. Our short term goal is to complete *BM.py* for the mantle convection community. Our intermediate goal is to modify this code for the short-term tectonics community, and then the long-term tectonics community. A standard procedure by which geophysical codes can be compared requires the establishment of benchmarks that are sensible mathematically (verification) and geophysically (validation). A common and permanent repository will be created for these benchmark results. CIG will collaborate with individual communities in order to develop this procedure: CIG staff will adapt and refine this procedure as it gains experience. We expect that the *BM.py* cod to be used by individual investigators studying their own benchmarks, while also being incorporated within our automatic regression testing.

4.2 Long-time scale tectonics. Finite amplitude deformation of the crust and lithosphere occurring across a range of time scales results in the geologic structure of the Earth and other planets. This deformation occurs both at the fast rates of earthquakes and dike intrusions and at slower rates of mantle convection and glacial loading. Elastic, viscous and brittle plastic strains all contribute to large-scale deformation. These time scales and deformation mechanism presents significant challenges in developing numerical codes to simulate such deformation. Thus, very different approaches are being used and there is need for making different approaches available to a wide community and to encourage benchmarking of those codes.

The wide availability of a 3-D implementation of the ALE (Augmented Lagrangian-Eulerian) and PIC (particle-in-cell) method and implicit methods that include adaptive mesh refinement (AMR) are important long-term goals for lithospheric deformation/long-time scale tectonics. Both ALE and PIC methods use a combination of Lagrangian particles and Eulerian meshes. AMR is desirable for efficiency and is a necessity for solving large-scale problems with high spatial resolution in regions of strain localization and fluid/rock interactions.

Further development and testing of the 2D and 3D versions of the *GALE* code will be needed, particularly adding functionality in terms of rheology. As described above, CIG is collaborating with VPAC on the development of *GALE*, a 3-D ALE code using common components in the *StGermain* framework. Much of this functionality exists from the *SNARK* application. *SNARK*, developed by the geophysics group at Monash University in Australia, is a hybrid between a PIC method and an implicit Finite Element Method. It is suitable for long-time scale crustal dynamics and mantle convection. A benefit of this approach is that the ALE method would be able to use different equation solvers available through *PETSc*.

We expect to release *GALE 1.0* in October, 2006. In August, 2006 CIG will participate in s rheology summer school at Colorado College sponsored by ISES. A more extensive training session will be held just prior (or after) the EarthScope National Meeting in March of 2007.

An additional intermediate goal is the expected donation of *SNAC* to CIG in April of 2007. SNAC, also based on the StGermain framework, uses the FLAC (Fast Lagrangian Analysis of Continuum) method that is widely used for 2-D studies of plastic failure and shear zone development in the crust.

There is a great need for planning to carry out benchmarking of codes. This is of particular importance for codes that deal with localization phenomena such as those that simulate tectonic fault development. CIG cooperates with a European group lead by Suzanne Buiter in the Geodynamics Center at the Norwegian Geological Survey (NGU) to develop benchmark standards. In association with a GeoMod2008 meeting in Europe, Buiter is organizing a two-day pre-conference workshop aimed at discussing results of a new numerical benchmark, new analogue benchmarks, new numerical-analogue comparisons, and modeling techniques. CIG will closely cooperate with this group on comparing benchmarks from *GALE* with those from other codes. CIG will support US participation in these activities.

For our longer goals, incorporation of adaptive mesh refinement into our codes is essential. For AMR, we see considerable value in implementing *deal.II*,, the finite element differential equations analysis library, as components in the infrastructure framework. Wolfgang Bangerth (Texas A&M), the author of the *deal.II* library, and Luc Lavier (UTIG) have been collaborating on this problem and would collaborate with CIG to make the products of their development available through our infrastructure frameworks.

4.3. Mantle convection. Among several long-term goals identified by the mantle convection community, we judge that the development of compressible spherical convection codes as particularly important for a better integration of mineral physics and seismology with deep-earth dynamics. The community also identified development of methods to couple small-scale physics, especially lithospheric deformation and melting/melt migration, into large-scale mantle flow models, as well as modular tools for grid refinement, element types, and solvers.

An open source, widely available, and well-engineered spectral analytic flow code remains an important goal of the mantle convection community. This project is still in its early stage. Thorsten Becker and Bernhard Steinberger have made substantial progress in putting together a code, *HC* (already in the SVN repository), based on some existing code and are testing the code with some success. More group discussions are currently under planning, and benchmark results of several codes are expected to be archived for comparison. It is hoped that the project will be completed by the end of 2006. It should be pointed out that this effort involves a significant amount of programming and documentation. The benchmark code, *BM.py* will allow close comparison between *HC* and *CitcomS* results.

The Boulder workshop also identified that the community needs to maintain an open source and widely available 2D convection code. Although most of our convection codes use iterative solvers, it is important for us to maintain a code that uses a direct solver that may work better for certain convection problems such as those with rapid viscosity variations. *Conman* seems to be a natural choice, given its reputation and history. Scott King and his student are working to provide a cleaned-up version of their *Conman* code to the CIG software repository. The main task is to remove the features associated with vectorization and reorganize the code accordingly. It is hoped that this project will be completed by the end of 2006.

The establishment of benchmarks, especially those for spherical mantle convection and making them widely available increases the quality of the science. Several benchmark cases of isoviscous, basal heating spherical shell convection cases were already computed using *CitcomS*. We are trying to get more groups to participate in this effort so that the results can be made available on the CIG website. Although this effort is a continuous process, we hope that by the end of 2006 we should have some benchmark cases for most commonly seen model types (isoviscous and temperature dependent viscosity) posted on the CIG website. Intercomparison of BMs will be made substantially easier and routine with the release of *BM.py*.

Another continuing need of mantle convection are visualization scripts (e.g., for DX, AVS, etc.) for producing animations and still images as part of the software packages.

By the far, the most important intermediate goal of the mantle convection community is the development of a new compressible mantle convection code. The Purdue workshop on compressible mantle convection provided an excellent starting point for this project. At the workshop, a few groups presented the results on modeling formulations and numerics. With Eh Tan joining the CIG this summer to help this project and other mantle convection projects, we hope that a preliminary version of the 3D spherical compressible mantle convection code (e.g., *CitcomS* based or other platforms) becomes available by the end of 2006, and a more polished product by the summer of 2007. We may also need to consider incorporating some particle-based advection schemes (e.g., the feature that Allen McNamara added to *CitcomS*). However, it should be pointed out that, given that this project involves new efforts in coding and benchmarks and that the optimal algorithm to be used still needs more investigation, the actual completion time may differ somewhat from this estimated time line.

An important long-term goal for the CIG that was identified at the Boulder workshop is to develop portable, module convection modeling codes that can incorporate different solvers and mesh types, in particular, with features like adaptive mesh refinement and higher order elements. These are also long range goals identified in magma dynamics and long-term tectonics. Consequently, important synergies exist with these other areas and CIG must vigorously pursue them. In many ways, this is our next generation mantle convection modeling tool that will differ significantly from our existing ones. So far, there is very little effort in the CIG aiming at this particular long-term goal. This can be applied to the existing mantle convection codes in CIG software repository, as being done to *CitcomS*. It should play a more important role in developing next generation mantle convection modeling software.

4.4. Computational seismology. A central goal for computational seismology is the establishment of automated and on-demand simulations (e.g. seismic wave propagation and synthetic seismograms) through a seismology *science portal*. A science portal is a web site that launches a simulation on a remote machine using data gathered from web sites and databases and then returns the results to the user. There are several components of portals, some of which are beyond the CIG scope, such as seismic databases. In order for a science portal to operate effectively, there is the need to select between multiple simulation methods and codes with uniform specifications. Consequently, the short-term goal of this community is migrating their principal codes for simulating seismic waves into an accessible repository and then reengineering such codes to work with a superstructure framework, in this case *Pyre*, the science neutral Python framework. This will allow the seismic codes to be initiated in a standard fashion. We have not identified low-level components from seismic codes as reusable within an infrastructure framework, but continued development of the superstructure and control of simulations in the framework via the web is a generic CIG need. Moreover, we have identified the movement of data and results between geodynamic simulations and seismic wave propagation as a community need for the EarthScope and CSEDI program areas, for example. Science portals could be developed as collaborations with other organizations and CIG is nurturing such relationships with IRIS, GEON, and CACR.

The seismology working group, made up of J. Tromp, M. Wysession, M. Ritzwoller, and A. Levander, has been able to identify the following short and intermediate term goals for computational seismology. By the fall of 2006, we plan to add a mode summation code to the CIG repository. This is likely to be facilitated through a subaward made to the University of Colorado (M. Ritzwoller, PI). The programmer at CU, familiar with seismology methods will work closely with the CIG software engineers. During the summer of 2006 we will release an updated and Pyrized versions of SPECFEM3D_BASIN and SPECFEM3D_GLOBE. Through collaboration with CACR at Caltech, during the summer/fall of 2006 we will implement a rudimentary version of a seismology web portal which will put us on track for on-demand seismology using both the SPECFEM and normal mode codes.

We will hold the second CIG computational Seismology workshop jointly with the "Imaging Workshop" organized by Levander/Aster/Pavlis/Rondenay. This will be held

October 31- Nov. 2 at Washington University in St. Louis. Finally, this summer we will firm up our collaboration with the European SPICE effort.

The future directions of computational seismology will be major subjects of discussion at the imaging workshop in St. Louis, but we have been able to identify some of the long-term goals for this area. Automated/On-Demand Simulations: CIG will work to establish a Seismology Science Portal, involving both automated and on-demand simulations. Automated simulations would provide near real-time 1D and/or 3D synthetics to accompany IRIS data for all events over a certain magnitude threshold using past and emerging events in the Harvard CMT catalog. Seismic Model Database: There is the need for a database of seismic models, including structural models of the crust and mantle together with databases of topography and bathymetry. Various resolutions are needed to match the capabilities of codes being developed under CIG. Mechanisms for the contribution of models must be established.

Data Processing Tools: The SSC is currently considering whether CIG should investigate the feasibility of facilitating the development of data processing tools for field and laboratory use. These could include low-level routines for standard data manipulation (e.g., filtering, simple array analyses); higher-level functionality such as earthquake location, traveltime picking, and moment tensor analysis; and high-level functionality such as tomography, receiver functions (perhaps with migration), and shear-wave splitting.

Visualizations of 2D and 3D seismic models are increasingly important in seismology and present an area of great overlap with other CIG efforts that require coordination. Imaging/tomographic tools may be included productively within the CIG framework.

4.5 Short time-scale tectonics. Short-term crustal dynamics focuses on simulating crustal deformation associated with the accumulation and release of strain over the earthquake cycle. The spatial scales range from meters to thousands of kilometers and the temporal scales range from tens of milliseconds to hundreds of thousands of years. At the larger length scales and longer time scales, models blend into those of the long-term crustal dynamics working group. An important continuing goal is the development of software for the simulation of multiple earthquake cycles with sufficient resolution to capture the buildup of strain in the crust, strain release in propagating ruptures that radiate seismic waves, and postseismic relaxation of the crust. Additionally, infrastructure is needed to couple crustal dynamics software to other models of Earth processes, as well as allow data assimilation into crustal dynamics software. Easy data assimilation in crustal dynamics software will promote integration of the wide spectrum of EarthScope data now being collected.

During the June, 2005, workshop the short-term crustal dynamics working group set forth four priorities for the year: (1) easy installation of *PyLith* and other CIG developed software, (2) implementation of a parallel version of *PyLith* with increased modularity, (3) infrastructure for archiving and comparing output associated with the suite of working group benchmarks, and (4) development of simple, general, efficient data structures and routines to store and manipulate finite-element mesh information that would serve as a common denominator between mesh generators and modeling codes.

PyLith, in addition to other CIG developed software, is now using the GNU build system, which is used on nearly every Unix operating system. This has allowed binary packages to be created for three of the most widely used platforms: Linux, OSX, and Windows reducing installation time from more than an hour of painstaking setup and compiling to a couple of minutes. For less standard platforms where building from source is required, the installation process has improved significantly with this new build system that greatly facilitates configuration checking and more explanatory error messages.

A parallel version of *PyLith, version 0.8.0,* supporting modeling of quasi-static crustal deformation with viscoelastic material models (written primarily by Williams) was released in June, 2006. This release includes documentation with detailed explanations of all parameters and a tutorial involving one of the community benchmarks. This signifies a major step forward for this community driven effort to develop scalable, modular software for short-term crustal dynamics. Development of *PyLith* continues, including the addition of full dynamic modeling capabilities via merging features from Aagaard's *EqSim* code and further integration within the *Pyre* framework to improve its modularity. All of the basic material and fault friction rheologies in *Tecton* are currently being reimplemented using this modular approach. This also allows users to extend *PyLith* and add their own constitutive models by writing small, simple modules. Release of this more comprehensive version, including documentation, is expected in late 2006. As part of parallelizing Williams's code, Matthew Knepley (CIG/ANL) in collaboration with Dimitry Karpeev (ANL) has developed *Sieve*, which is infrastructure for storing and manipulating general finite-element meshes. Although *PyLith* is currently the only CIG code using Sieve, several of the other modeling codes will likely take advantage of *Sieve* in the near future.

An ongoing implicit task associated with the development of *PyLith* is the construction of all components necessary for using *PyLith* as a research modeling tool. This includes interfacing *Sieve* with a few of the most popular and capable mesh generators used by the community (e.g., Cubit, LaGriT, and TetGen) so that users can export meshes directly to *Sieve* for use in *PyLith*, and interfacing *PyLith* with a couple of the most popular visualization packages used by the community (e.g., ParaView and OpenDX).

The initial CIG effort on developing infrastructure for archiving and benchmarking output from community benchmarks has focused on the mantle convection working group. In May 2006, Luis Armendariz (CIG) began refining the *BM.py* software to meet the more general needs of the short-term crustal dynamics working group while the working group began defining standards for simulation output and metrics for comparing results. Setting up *PyLith* (and permitting other non-CIG supported short-term crustal dynamics modeling codes) to use this infrastructure is anticipated in late 2006.

In 2007, we expect a significant shift to occur in the development of *PyLith*. The focus will move from transforming the old, entangled legacy simulation codes to enhancing various components of the new, modular simulation code. The working group has identified several areas in which to target resources with four given highest priority. (1) Adding adaptive mesh refinement capabilities to *PyLith* in order to automatically tune the finite-element discretization for each problem, thereby improving the efficiency of the computation while incorporating error estimates. (2) Refine *PyLith's* interface with *PETSc* to expand the different kinds of solvers (e.g., nonlinear multigrid solvers) it can use in order to leverage advances in *PETSc* solver development and allow users to select solvers most appropriate for their computation. (3) Allow the use of spherical geometry, in addition to the currently supported Cartesian geometry, so that boundary conditions and parameters can be specified in a more natural coordinate system. (4) Develop robust, validated, open-source versions of codes for dislocations in layered elastic and layered viscoelastic domains. While a variety of codes are in use to do these types of computations, few are portable and none undergo continuous regression testing. Furthermore, these codes would serve as tutorials/introductions to the highly modular, portable software developed by CIG for the short-term crustal dynamics community and provide reference solutions for additional benchmarking.

Several priorities beyond 2007 have already been identified. (1) Extending *PyLith* to automatically adapt the timestep and solution scheme in order to resolve the rapid rupture propagation and deformation during earthquakes as well as the gradual pre- and post-seismic deformation between earthquakes. (2) Integrating tools for formal data assimilation in order to permit data from EarthScope and other sources to be included directly into simulations. (3) Coupling short-term crustal dynamics simulations to other simulations in order to more accurately capture interactions among geodynamics phenomena.

4.6. Geodynamo. Dynamo codes represent a powerful new tool for the quantitative study of a broad range of geophysical processes, ranging from short time-scale phenomena such as magnetic variations, rotational variations, and flow in the core, to long-term phenomena such magnetic excursions, reversals, superchrons, and the evolution of the core and its thermal and chemical interaction with the mantle. The primary objective of CIG in this area is to provide the Earth Science community with robust, reliable, efficient, flexible, state-of-the-art numerical codes for modeling dynamo processes in the Earth's core and in the interiors of other planets. Another CIG objective is to support graphical- and user-interfaces for these codes that allow Earth scientists to analyze, display, and interpret dynamo code results, and to compare results from the various codes that we support, as well as with geomagnetic, space magnetic, and paleomagnetic data. A longer-term goal of CIG in this area is to foster development of the next generation of dynamo codes, with emphasis on modular design, inter-operability, and compatibility with CIG framework standards.

SSC member Peter Olson has polled members of the US geodynamo and related communities (including paleomagnetism and geomagnetism) to determine priorities for CIG in this area. In addition to the *MAG* code already donated, the geodynamo community is ready to donate a second and possibly a third working dynamo code to CIG. The two committed codes are (a) *MAG*, a basic Boussinesq convection dynamo code written by Gary Glatzmaier and adapted by Peter Olson and Uli Christensen for use on serial workstations, and (b) *MoSST*, a serial dynamo code written and used by Weijia Kuang. *MoSST* includes capabilities for rotational interactions between the mantle and the outer and inner core. In addition, we intend to pursue negotiations to obtain a parallel dynamo code that involves inner core physics and mantle interactions. The two codes now available to us are pseudo-spectral, with *MAG* using Chebyscheff polynomials whereas *MoSST* uses a 4th-order finite difference scheme in the radial direction.

The following objectives were defined in the course of polling community members. All of these constitute essentially short-term goals, which could be accomplished in approximately one year's time following the hiring of a software engineer dedicated to geodynamo codes. Longer term goals identified by the same polling are listed in the next section. Standardization of existing dynamo codes to allow inter-comparison of results. Making benchmark comparisons of the three codes. Porting the dynamo codes onto TERRA GRID platforms to enhance community access and use. Development of a common user interface. Development of a common graphics interface. Development of one or more "user pacs", designed to translate between dynamo code output and standard data formats and definitions used in geomagnetism, paleomagnetism, and in the planetary and space sciences. Discussion of CIG objectives at upcoming SEDI meeting in Prague, with dynamo workshop to follow.

In spite of the recent notable successes, substantial challenges confront those trying to numerically simulate the geodynamo. Although existing dynamo models often produce "realistic" output, in terms of their magnetic and velocity fields, the parameter regime that is now accessible is for the most part very far removed from the Earth and other planets. Conversely, attempts to force existing models to use more realistic values of the critical control parameters typically causes the same numerical dynamos to fail. Not all the reasons for this limitation are yet clear, but this is now acknowledged as a major obstacle for future progress.

A part of the longer-term CIG strategy in developing the next generation of software in this area would to free dynamo codes from some of their current limitations. One limitation is that the dynamo codes now in use were not designed with a modular construction. Other disadvantages stem from their reliance on traditional spherical harmonic expansions. Advantages of the spherical harmonic representations include that divergence-free magnetic and velocity fields result automatically, the pressure term is eliminated using the vorticity equation, and the boundary conditions are also easily incorporated in the solution procedure for the interior. However, spherical harmonics waste resources in the polar regions, and the transformations between physical and harmonic spaces occupy an increasingly large fraction of cycle time as the model resolution increases. There is an obvious need for algorithm development, in particular a fast, stable Legendre transformation that allows parallelization, since spherical harmonics will continue to be used in dynamo codes, for satisfying magnetic boundary conditions and other specialized uses. But a local basis function formulation for the interior variables offers clear advantages for implementing sub-grid scale physics, variable properties, and uniform resolution on spherical surfaces.

In addition there is a developing consensus in the community on the need to understand turbulence in the core, so that sub-grid scale capability is an emerging priority for future dynamo codes. Similarly, most existing dynamo codes use the Boussinesq approximation, although thermodynamic considerations indicate that compressibility may be important, certainly in the gas planets, but possibly also in the terrestrial planets as well. A more locally-based, modular dynamo code structure may be required to overcome these problems as well.

In contrast to research in some of the fields represented by CIG, there is nearly universal agreement on the form of the most important elements in a dynamo code. The basis of all dynamo codes is classical MHD, with the Navier-Stokes equation expressed in Earth's rotating frame and the magnetic induction equation in its standard form. Although there are widespread differences in the forms of the driving forces, including thermal and compositional buoyancy plus rotational variations and possibly tidal effects, all of these are based on known equations from the realm of classical physics. Accordingly, the basic structure of present and even future dynamo codes is stable. This situation offers an opportunity to reverse-engineer existing dynamo codes, or alternatively, engineer new codes, with a modular structure, so that researchers can combine the individual modules in new and creative ways so as to explore new types of dynamo action in new parts of the parameter space.

Assimilating surface geomagnetic observations into numerical geodynamo models is a way to bring the use of dynamo models into the area of geomagnetic forecasting, and as well as providing an avenue to improve the models themselves. The goal would be to provide the capability to assimilate surface and satellite observations into the models to better constrain model parameters. Major challenges in assimilation of geomagnetic data into numerical dynamo models include the differences between the parameter domains used in numerical dynamo modeling and that appropriate for the Earth's core, the fact that surface observations only provide part of the poloidal magnetic field and none of the toroidal field, and the facts that geomagnetic observations are limited to a small fraction of one magnetic free decay time whereas paleomagnetic observations have limited spatial and temporal precisions.

4.7 Magma dynamics. Melt migration is an area of emerging CIG activity. Discussions within the SSC have identified important pedagogical and computational synergies with other CIG areas. This includes compressible mantle convection where energy conservation and compressible viscous flow algorithms being developed will be useful in melt migration modeling. Crustal fluid migration problems, including for example hydrothermal systems and basin evolution, are also highly complementary with melt generation and migration.

Melt migration, perhaps more directly than any other area within CIG, represents an area where physical and chemical phenomena on a wide range of scales interact in a very direct way. This would include the mechanisms of melt localization and the implicit coupling between melt generation, solid state mantle flow and temperatures, and physical/chemical state of the melting solid. Melt migration problems offer an opportunity to be a prototypical area for multiscale modeling that will be important in other CIG areas.

A large number of private research codes exist that explore many aspects of magma dynamics in the mantle and crust as well as their observable chemical and physical consequences. These codes range from simplified process models that explore the fundamental physics of melt and chemical transport to more applied models of magmatic regions, particularly mid-ocean ridges, subduction zones and mantle plumes. These codes have been written in a variety of languages, using a wide range of formulations and algorithms. Unlike mantle-convection or computational seismology, there is no obvious candidate "hero-code" for contribution in the same manner as CitCOM or SpecFEM. However, given that magma-dynamics is a direct extension of mantle-convection, it seems more sensible to begin directly re-engineering these codes from the algorithmic level using existing computational frameworks to develop a more flexible modeling toolkit for coupled fluid-solid problems. There has already been considerable success with developing scaleable, parallel magma-dynamics codes using PETSc and we have begun exploring interactions with VPAC and the St. Germain framework. All of these codes and approaches will be discussed in detail at the upcoming CIG workshop on Computational Magma Dynamics organized by Marc Spiegelman and Laurent Montesi to be held Aug 18-19 at Columbia University. The goal of this workshop is to review and discuss existing and future scientific and computational needs in magma dynamics and develop a specific set of short, intermediate and long-term goals for CIG similar to those produced by other workshops such as the long-term Tectonics Breckenridge Meeting.

While specific recommendations will be available only after community input from the meeting, the discussion will include at least the following topics:

- Integration of magma dynamics (multi-phase flow) with existing CIG projects in solid-state mantle convection and lithospheric deformation.
- Development of multi-scale/multi-physics models for integrating magma dynamics and geochemical evolution into large-scale mantle dynamics.
- Common computational needs for multi-scale localization problems (e.g. AMR)
- Interfacing CIG magma dynamics with ongoing efforts in geochemical databases (e.g. EarthChem) and computational thermodynamics (e.g. MELTs).
- Development of a regional plate-boundary scale modeling toolkit for developing region-specific models of mid-ocean ridges, subduction zones and mantle plumes with consistent solid and fluid mechanics. Such a toolkit should be of significant utility to the larger MG&G community and projects such as RIDGE2000 and MARGINS in need of dynamic models to compare with data.

After a thorough discussion of scientific priority and computational feasibility, the workshop will present specific short, intermediate and long-term goals for magma dynamics, ready for incorporation into the next phase of CIG software development activity.

4.8 Facilitating the use and access to the TeraGrid. CIG doesn't maintain its own major computational hardware, but we are starting to develop a relationship with the NSF TeraGrid so as to provide a resource for both the community and CIG staff. Perhaps most critical is that CIG software is being developed so that it works well on any platform from desktop Linux & Mac OSX platforms to Beowulf clusters and large supercomputers such as those available on the TeraGrid. Since many members of the community from smaller departments and universities do not have ready access to parallel computers, availability of CIG software on the TeraGrid provides an important resource for this part of the community. CIG currently has a moderate allocation from TeraGrid and the SSC will develop a strategy for a larger allocation this coming year. CIG needs for TG resources include: workshops, software testing, allocation to the geophysics community for production runs using CIG software, and, in collaboration with other groups, development of several science gateways (e.g. web portals) that utilize CIG developed and maintained codes. We will continue to develop expertise in this area and future improvements will include the follow:

- On our Web site, we will develop a set of pages so that members of the community can transparently get accounts to TG under CIG allocations. We will also provide instructions on accessing these resources.
- In our User Manuals, we will develop chapters and test examples that illustrate how to run the codes on the TG.
- For some of our software and for some TG platforms, we will develop the software so that it is pre-installed on the supercomputer for use by any user.

The allocations will give the CIG community and staff a concrete estimate of how our software performs and scales on TG resources. When CIG works on the build procedure, it is essential that we develop and exercise the software on as many platforms as possible and the TeraGrid provides us with a useful set of resources where we can accomplish this.

There are probably two significant impediments to the use of national supercomputer facilities by many segments of the scientific community. First, is the availability of software that effectively uses massively parallel supercomputers. Clearly, the fact that CIG software is designed for such purposes and is preinstalled on such platforms should largely overcome these problems for the greater geodynamics community. Second, is the need to write a proposal that documents the level of supercomputer resources needed and that the software effectively uses and scales on the platform in question. CIG will provide the necessary results showing the performance and scaling of our software on shared national resources. Such performance information will be a natural product of CIG software development. Consequently, we are likely to see increased usage of TeraGrid facilities by a broader range of geoscientists.

4.9. Organizing community participation. Centrally important for CIG, is the guidance from the scientific community on what this infrastructure should accomplish for their evolving research needs. This is accomplished through community oversight of CIG, committees, and workshops. Two key principles guide CIG's interaction with the community:

- *Openness*: All CIG reviews, meeting minutes, and other documents are openly available to the community in a timely fashion, for review and public discussion, unless this conflicts with individual or institutional privacy rights (such as salary level or personal information). Openness minimizes the risk of actual or perceived bias or conflicts of interest within CIG.
- *Interaction*: All CIG committees and workshops should have an open and balanced representation of both the scientific and software communities. Interaction minimizes the risk of balkanizing the community CIG serves.

Most of the organization of CIG for the community is now in place, however there is likely the need to recruit more members of the community to participate in working groups, especially early career investigators. Workshops and committee meetings will continue to be the mechanism by which we layout a software system that delivers the core functionality of geodynamics in an open and extensible fashion. As described already, CIG has been sponsoring or co-sponsoring a number of workshops this summer with the principal goal of increasing community participation in CIG.

The Web site is an important tool for community participation. It is not just used as a means to distribute software and documentation, as described above, but it is also used for members to communicate. CIG currently uses a technology called Plone, a user-friendly and powerful open source Content Management System, that allows users to edit the web pages directly. For example, each of the workshop reports has a comment section at the end and registered users of the web site can add comments and content to both their own individual area as well as special areas for each working group. In addition, CIG maintains List-Servs for both the entire community (cig-all@geodynamics,org) , each committee, and each subject area. These list servs are cross-linked so that a user can easily navigate from the Plone editable areas to the List-servs.

4.10. User training. A key objective of CIG is the widespread adoption of CIG-developed software by the general Earth science community and, in particular, by geophysicists. Both significant training and comprehensive documentation will be required. Potential users will want to know how to use CIG software as well as gain an understanding of the underlying algorithms and implementation details. Based on current research directions, sophisticated users who require a deeper understanding of the workings of the software will likely have different backgrounds and interests. One class of user consists of traditional computational scientists want to use several different computational components (such as a solver and a mesher) to create working codes that incorporate algorithmic innovations. Another class of user consists of scientists who attempt to string together several components of CIG software with non-CIG data analysis tools. In addition, there will be a class of less-technically-demanding users who will want to use

CIG codes in a standard manner (e.g., for pedagogical purposes or using standard codes on new input data sets).

Several mechanisms will be used to train CIG users: (A) small, focused workshops; (B) visits to the CIG site to work with CIG staff and other users; and (C) the production of training manuals and a web site. CIG continues to use these mechanisms to provide training for both the expert and non-expert user.

Table 4.1
Short-, Intermediate-, and Long-term goals of CIG

	Short Term Goals	Intermediate Term Goals	Long Term Goals
Short Term Tectonics	PyLith 0.8, a parallel version of code with visco-elasticity functions, due by June 26, 2006 Workshop. Goal is to have 20-30 users after workshop in June.	PyLith 1.0, which incorporates static and dynamic functions. Target date 1/1/07	Adaptive Mesh Refinement using PyLith. Future goals would be to allow code to couple with other geophysical codes. Also Adaptive Spherical geometry.
Computational Seismology	Have 1-D mode code supplied by Ritzwoller and given Pyre wrap, with mode catalogue and mode summation. Also, a fully Pyrized and normal version of the Global and Basin SPECFEM code will be ready by the next workshop (10/06). The stand-alone (and hopefully Pyre version) will have full documentation.	Highest priority for two packages: a 1-D package for Cartesian reflectivity, and a finite difference 3D wave propagation (?) code. Goal will be more specifically defined after SPICE meeting.	On demand/on request seismology. Community is currently working on a rudimentary web-interface. Another goal is for coordination of model databases.
Computational Science	When community meets for workshop in October 2006, specific intermediate and long-term goals will be discussed.	To be defined at Workshop	Work on integration of computer sciences and geophysics into single framework
Long Term Tectonics	GALE beta to be released for some users, with additional functionality, Aug., 2006 Reach out to community via a two-day tutorial in code, for Fall 2006.	Code with 3-D Spherical Elasticity, either via GALE or SNAC (available April 07). This code would tie into Short-Term Tectonics via reducing time-steps in 2D computations – beneficial to both communities. Host workshop in conjunction with a pre-meeting at Earthscope in Spring 2007.	Develop new code with mesh refinement, Adaptive Mesh Refinement.
Magma Migration	When community meets for workshop in August 2006, specific intermediate and long-term goals will be discussed.	To be defined at Workshop	TBD at workshop but possibly including, integration of magma dynamics in global mantle convection, development of a regional plate-boundary modeling toolkit for ridges, arcs and plumes, integration of fluid flow in lithospheric deformation.
Geodynamo	Benchmark and adapt two serial geodynamo codes, MAG and MOSST; development of user interfaces. Preliminary community building discussions will be held during SEDI in July 2006	Community building workshop in 2007, to utilize and train users on the two serial codes, possibly partnering with other organizations. Test user interfaces and web portal. Addition of parallel code in web portal.	New dynamo code using common components with Mantle Convection codes including common Earth structure framework and grid and adaptivity.
Mantle Convection	Deliver beta version of compressible spherical code (CitcomCS 3.0) by 12/06. Basic visualization package by 12/06. 1-D code (HC 1.0) version 12/06	CitcomCS 3.N with possible additions (self consistent thermodynamics) designed by community.	New code with adaptivity, and Adaptive Mesh Refinement.

Table 4.2
Cross-Disciplinary Goals - Infrastructure

Mostly completed	Long-Term
Web site (implemented Plone)	
Automatic Builds (implemented Buildbot)	On demand/on request computing using web-interface.
Benchmarking	Benchmarking (*BM.py*)
Regression testing (developed *CIG-Regresstor*)	
Launcher package (developed *Addendum.py*)	
Software repository (implemented SVN)	

Table 4.3
Cross-Disciplinary Goals – Scientific Computing

Mostly Completed	Long-Term
Sieve code	AMR
PETSc/Pyre framework for codes.	Early parallel mesh generator (TK of Texas) to be altered to become solution adaptive, and therefore provide possible AMR that can be scaled to billions of elements and thousands of processors.
HDF5 – output of codes in similar format for ease in benchmarking and for future visualizations.	Continue use of PETSc and Pyre to develop software frameworks

Table 4.4
Workshops Planned

WORKSHOPS	
Short-Term Tectonics	June 26-30, 2006 Workshop, to be held in Golden, Colorado. Workshop will focus on using available (Lagrit, Cubit) and new codes (PyLith 0.8), as well as community discussion on benchmarks, meshes, and case studies. Presumably would co-sponsor meeting in June, 2007.
Magma Migration	August 18-19, 2006, Columbia University. Meeting to discuss current software needs of community in magma dynamics and magma migration.
Computational Seismology	October 31-November 2, 2006.Washington University in St Louis. Second CIG computational Seismology workshop as part of “Seismic Imaging Workshop”.
Computational Sciences	October 16-17, 2006, Austin, TX. Cross-disciplinary workshop bringing together various geophysical and computer science experts to work on common theoretical and mathematical obstacles to scientific coding. Followed by Computational Science Round Table October 18.
Mantle Convection	Summer 2007 workshop to discuss science, current software development, and future plans for new code with adaptivity and AMR.
Geodynamo	Pre-meeting to begin community building in Spring 2007, as part of SEDI program, to be followed up by a full workshop later in 2007.
Long-Term Tectonics	Pre-meeting discussion to be held before EARTHSCOPE in March 2007.

5. Achieving our long-term goals

After detailed evaluation of CIG's goals, the Science Steering Committee has concluded that we will likely pass through an important transition in the coming year in which all of our short-term goals have been achieved. We believe that we will have also made substantial progress on many of our intermediate term goals. Consequently, we must seriously evaluate long-term objectives and articulate a plan for achieving them. The details of our plans are still under development, but in this section we described how we hope to move forward.

The long-term goals of CIG described generally in Section 2 and then in detail in Section 4, essentially involve two high level tasks: Adopting advanced numerical methods and coupling multiple methods or codes to solve multi-scale or multi-physics problems. The need for adaptive mesh refinement is an example of a need for an advanced numerical method. In both cases, this involves software and ideas developed at the forefront of computational science. In CIG's NSF proposal, the community articulated the view that our resources would always be limited and entirely new software would only be generated when needed. This remains true, and CIG must rely on lower level components developed in more science-neutral arenas, including some DoE supported projects, and higher-level frameworks to interlink software as much as possible.

The need to adopt these new computational science technologies was anticipated in the NSF proposal. The development of a software framework for geophysical modeling and its subsequent upkeep were viewed as central to CIG's strategy. We have articulated and used a framework that consists of software on two levels:

- An *infrastructure layer* of software by which community and state-of-the-art modeling codes will be assembled

- A *superstructure layer* of software by which multiple codes and data can be interlinked through the extension of existing software frameworks

CIG has already been working with software on these two levels. For example, *StGermain* is an example of an infrastructure framework, while *Pyre* is an example of a superstructure framework. We have developed *GALE* "upward" by using the existing components in *StGermain*. While in the case of *PyLith*, we have built the code "downward" by writing and adapting existing components and linking them with *Pyre*.

In addition to treating coupled problems, there is already a lot of functional commonality in the codes CIG has been working on. For example, *PyLith* and *GALE* in treating problems in short and long-term lithosphere dynamics, respectively, already have many common elements. Compressible mantle convection, as in the *Citcom*-based code currently being developed by CIG, and melt migration have the common elements of treating dilatational viscous flow and formulations of the energy equation that go beyond simple advection and conduction of heat. The next step beyond our intermediate term goals will be incorporating into our framework adaptive mesh refinement as a central common feature, thus providing the geodynamics research community a unified set of tools that will permit the treatment of intimately coupled, multiscale problems that will lead eventually to a whole earth evolution model.

In the past, CIG's software effort was mostly focused on integrating existing codes and bringing them up to the standards necessary for inclusion in CIG's software suite. In addition, several new codes have been developed, for example the arbitrary Langrangian Eulerian code *GALE*. In the long run, one of the goals is a cautious re-engineering of some of the codes in a number of directions. For example, several of the groups wish the ability to easily switch between different solver kinds to evaluate which solvers are most efficient, or to switch to more robust solvers if problems become too ill-conditioned or change character (for example from diffusion to transport dominated). Likewise, several of the groups intend to explore concepts of adaptive mesh refinement, higher order discretizations, or time step control.

Past experience in computational science has shown that in all these areas, robust solutions require tens or even hundreds of thousands of lines of code, and months or years of work. They are therefore best realized in supporting libraries such as *PETSc* (solvers) or *deal.II* (adaptive mesh refinement, discretizations), rather than implementing them anew for each specific project. *PETSc* has already been integrated into a number of the projects under CIG and allows the simple interchange of solvers. However, no attempt has been made thus far to integrate adaptive mesh refinement strategies into any of the codes because, among other reasons, of the pervasive changes this usually entails. In order to create greater synergy between the various disciplines using CIG software, we need to rationalize how we proceed. For example, do we only write codes with the *StGermain* framework, such as we did with *GALE*, or do we develop and migrate existing code and libraries into *Pyre*, as we have done with *PyLith*?

In order to seriously evaluate our plans, the Science Steering Committee will organize back-to-back meetings in Austin Texas. First, will be a two-day Computational Science workshop organized by W. Bangerth, O. Ghattas, and B. Aagaard October 16-17. The aim of the workshop is to bring together computational geodynamicists with computational scientists in other sciences, mathematics, and computer science, to discuss problems of common interest. Immediately following this workshop, the SSC will host a computional science “round table”. In this meeting, members of the SSC and others involved with the CIG will first individually describe how they think CIG should move forward with software development for their respective discipline. This would be followed by an open discussion on trying to find the best path(s) forward. The SSC would then use the results of this round table in the development of next year’s Strategic Plan.

The two-day workshop will feature speakers in all application areas represented in CIG, i.e. short-term and long-term deformation, geodynamo, seismology, mantle convection, and magma migration. Speakers will be encouraged to focus more on open computational issues, rather than scientific achievements. On the other hand, computational scientists will present results and numerical strategies for some of the problems that seem to appear again and again in various disciplines. In particular, this includes (i) the problem of meshing complicated geometries while respecting interior interfaces such as slip surfaces (faults) or paths of localized fluid (melt) migration, as well as adaptive mesh refinement techniques, (ii) iterative and direct, possibly distributed linear solvers and preconditioners, (iii) nonlinear solvers and operator splitting techniques for complex and coupled problems, (iv) data formats, data management, and parallel visualization for large data sets (terabytes) resulting from massively parallel computations, and (v) issues in packaging and quality assurance of Open-Source libraries supporting complex large-scale computations.

By bringing members of the community affiliated with CIG in contact with scientists in other areas, we hope to increase transfer of knowledge, cross-fertilization of ideas, as well as start collaborations between scientists working in different areas but encountering similar computational problems.

6. Annual Goals and Milestones (Sept. 1, 2006 – Aug. 31, 2007)

I. Common Infrastructure

a. Maintain LAN, servers, desktops, notebook computers
b. Maintain Plone Site (http://geodynamics.org)
c. Maintain repository (SVN)
d. Maintain and expand regression testing (*CIG-Regresstor*)
e. Expand *Sieve* software suite
f. Expand benchmark code (*BM.py*)

II. Core Computational Software

a. Computational Seismology code (normal mode, NM), include Pyrization
b. Continue migrating NM and *SPECFEM* codes under Pyre into Web Services
c. Compressible mantle convection code development (*CitcomCS3.0*)
d. Mantle convection support, *HC*
d. Long term tectonics code development (*GALE 1.0*)
e. Long term tectonics code migration into SVN (*SNAC 1.n* donation)
f. Short term tectonics code development *PyLith 1.0*
g. Geodynamo migration into SNV, release, benchmarking *MAG/MoSST*
h. Geodynamo pre-/prost-processing
i. Initiate development of magma dynamics software
j. Move toward incorporating AMR

III. Organizing Community Participation

a. Annual meeting of EC November, 2005, Columbia University, NYC.
b. Computational Science Workshop, Oct. 16-18 Held at the Institute for Computational Engineering and Science (ICES) in Austin Texas.
c. Computational Science Round Table, Oct. 18, 2006, ICES, Austin, TX.
d. Computational Seismology meeting in associate with Imaging workshop at Washington University, St. Louis, Oct. 31-Nov. 2, 2006.
e. Annual business meeting at Fall AGU, San Francisco, CA
f. Annual meeting of SSC, MAY, 2007, Pasadena CA.
g. Finite element modeling of fault interactions workshop (Co-sponsorship with SCEC, summer 2007)

IV. User Training

a. User manuals
    i. GALE manual
    ii. PyLith manual
    iii.SNAC manual
    v. CitcomCS manual
    vi. MAG/MoSST manual
b. Training sessions

i. Long-term tectonics training session at EarthScope Annual Meeting, March, 2007, Monterey CA.

7. Allocation of resources by goal

Sept. 1, 2006 – Aug. 31, 2007

Software Engineer #1 (WL,Numerical Analyst/Modeler) [Total 1 FTE]

- I.b (repository): 0.1 FTE
- II.d GALE development: 0.45 FTE
- II.i (AMR migration): 0.25 FTE
- III. b.c (computational sci. wks:): 0.05 FTE
- IV.a.i (GALE manual): 0.1 FTE
- IV.b.i (GALE training session): 0.05 FTE

Software Engineer #2 (LS, Software Integration) [Total 1 FTE]

- Ie (Build system and installation): 0.25 FTE
- II.a (1-D seismic codes): 0.25 FTE
- II.b (Pyre seismic codes and web services): 0.4 FTE
- III.d Computational seismology workshop: 0.05 FTE
- III.h CFEM workshop: 0.05 FTE

Software Engineer #3 (LA, Software Integration) [Total 1 FTE]

- I.c SVN repository: 0.1 FTE
- I.d CIG-Regresstor: 0.1 FTE
- I.g Benchmarking (BM.py): 0.5 FTE
- II.c.(Compressible Solver CitcomS.py): 0.15 FTE
- II.d support for HC: 0.05 FTE
- III. b.c (computational sci. wks:): 0.05 FTE
- IV.a.iv (MC manuals): 0.05 FTE

Software Engineer #4 (ET, Numerical Analysis) [Total 1 FTE]

- II.c (Compressible Mantle Convection development): 0.75 FTE
- II.i (AMR): 0.1 FTE
- IV.a.iv (convection manual): 0.15 FTE

Software Engineer and system admin. (WM, software integration) [Total 1 FTE]

- I.a. System administration: 0.25 FTE
- II.g (Geodynamo codes): 0.25 FTE
- II.h (Geodynamo pre-/post- processing): 0.4 FTE
- IV.a.v (geodynamo manual): 0.1 FTE

ANL Subcontract [Total 0.5 FTE]

- I.d Sieve development:0.05 FTE
- II.c (Compressible Solver in CitcomCS): 0.1 FTE
- II.f (PyLith): 0.30 FTE
- III.h (Short term Wkshp): 0.05 FTE

University of Colorado Subaward [Total 0.33 FTE]
II.d deliver fully function normal mode code to CIG: 0.33 FTE

VPAC Subcontract [Total 0.5 FTE]
II.d (GALE): 0.5 FTE

Director [Total 0.12 FTE paid by CIG]
Center Management 0.12 FTE

Chief Software Architect [Total 0.12 FTE]
I,II (common infrastructure oversight): 0.12 FTE

Secretary [Total 1 FTE]
III (community workshops): 0.3 FTE
Paperwork for Management: 0. 5 FTE
General Web Management: 0.2 FTE

Technical Writer/Web master [Total 1 FTE]
I.b (maintain web site): 0.25 FTE
IV.a (User Manuals): 0.75 FTE

Supplies and Expenses [$65K]
I, II, III, IV

Travel [$28K]
I, II, III, IV

Participant Costs [$91K]
III, IV

8. Membership

8.1 CIG Members and Member representatives:

Argonne National Laboratory (MSC)
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
Lawrence Livermore National Laboratory
Los Alamos National Laboratory (ES)
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
U.S. Geological Survey (Menlo Park)
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
Washington University
Woods Hole Oceanographic Institution

8.2 CIG Foreign Affiliates and representatives:

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

8.3 Strategy for keeping members informed

Member representatives and individuals within the larger CIG community (including those at member institutions) will be kept informed in several ways.

1. Through e-mail. CIG maintains several list servers through the CIG website including several for the main committees (e.g. Executive Committee, Science Steering Committee) as well as for working groups and general information (e.g. cig-all@geodynamics.org). A CIG Newsletter highlighting new developments and capabilities with appropriate links to the CIG website will be distributed by email on a regular basis.

2. Through the http://Geodynamics.org web site. The upcoming CIG calendar of events is posted and continuously revised. Nearly all CIG documents, including proposals submitted to CIG, the annual revision of the CIG Strategic Plan, By-Laws, etc., are posted of this site. The Web site is the principal means for standard software downloads, sharing of community benchmark, specifications of standards, and distribution of user & training manuals.

3. The annual CIG Business meeting. This meeting will be open to all and will be a forum for open discussions of the working of CIG, including past and upcoming activities & the Strategic Plan. In year one, this meeting will be held in conjunction with the Fall AGU meeting in San Francisco. Depending on the success of this meeting, this may be the venue for subsequent years.

4. CIG sponsored and co-sponsored workshops and training sessions. The current status of CIG will be presented at these workshops and we expect that CIG members will attend such workshops.

9. Five Year Management Plan

CIG will need the expertise, vision, and guidance of the community if it is to remain a nimble and evolving organization. Consequently, we are adopting a *community-centric* management structure that draws upon features of successful NSF-supported community infrastructure projects in the Earth sciences. The management plan, outlined here, has been codified in a set of by-laws available on our web site (http://Geodynamics.org).

9.1. Institutional Membership and Executive and Science Standing Committees.

CIG is an institutionally-based organization governed by an Executive Committee. The structure of CIG recognizes member institutions, which are educational and not-for-profit organizations with a sustained commitment to CIG objectives, and a number of foreign affiliate members. The Member Institutions will change over time because CIG is an *open organization*, available to any institution seeking to collaborate on the development of open-source software for computational geodynamics and related disciplines.

The Executive Committee is the primary decision-making body of CIG; it will meet at least twice per year to approve the annual science plan, management plan, and budget, and to deal with major business items, including the election of a Nominating Committee. With the Director, the Executive Committee will handle the day-to-day decision-making responsibilities through its regular meetings, teleconferences, and electronic mail. The Executive Committee will have seven members. It will have four voting members: the Chairman, the vice Chairman, and two members at-large. These members will be elected by representatives of member institutions for staggered three-year terms. The three nonvoting members are the Director, the Chief Software Architect, and the Chairman of the Science Steering Committee. The Executive Committee will have the authority to approve proposal submissions and contractual arrangements for CIG. The Executive Committee believes that having an odd number of voting members is prudent than an even and so will be proposing an amendment to the CIG ByLaws to increase the number of voting members from four to five.

CIG has a Science Steering Committee that consists of eight elected members including a chairperson. The committee has a balance of expertise in both geoscience and computational sciences and provides guidance within all of the sub-disciplines of computational geodynamics. Their principal duties are to assess the competing objectives and needs of all the sub-disciplines covered by CIG, provides initial assessment of proposals submitted to CIG, and will revise the Five Year Strategic Plan. Recommendations from the SSC are passed on the EC.

9.2. Administration.

The Director is the Chief Executive Officer of the organization and bears ultimate responsibility for its programs and budget. The Director's responsibilities include: (a) devising a fair and effective process for the development of the Strategic Plan, based on proposals or work plans such as those submitted to the Executive Committee by the

Science Steering Committee, and overseeing the plan's implementation, (b) acting as P.I. on proposals submitted by the core CIG facility, retaining final authority to make and implement decisions on grants awarded to the core facility and contracts, (c) ensuring that funds are properly allocated to various CIG activities, (d) overseeing the preparation of technical reports. The CIG ByLaws do not yet stipulate the term of the Director and so a discussion item at our future Business meetings will devising a mechanism for the orderly transition to subsequent Directors.

The Chief Software Architect (CSA) will serve as a non-voting member of the Executive Committee. His role is to provide advice and perspective to the Executive Committee on the overall composition, integration, and balance between software development activities of the organization. He provides frequent assessments of our software, identifies new opportunities in both computational science and methods for software development, and provides evaluations of prospective members of the Software Development Team. The Executive Committee retains the authority to appoint the CSA.

9.3. Formulating CIG Priorities and Management of its Resources.

Concepts and plans for CIG activities will come directly from the community, member institutions, working groups and their elected committees. Ideas and plans will move from members to the Science Steering Committee and finally to the Executive Committee. As part of the development of the Strategic Plan , the SSC formulates a prioritized list of tasks for software development for the coming year, how these tasks are both inter-related and related to the broader needs of the community, and then transmits this as a recommendation to the Executive Committee. On at least a yearly basis, the Executive Committee will allocate resources to specific software development tasks. Following this allocation of resources, the EC will periodically appoint small committees to interface directly with the software development team (SDT).

It is expected that members of the SSC will be fully engaged in a dialog with the user community and active users of CIG software. Besides the constant dialog that such committee members would naturally have with the community, CIG will have a formal process for bringing new ideas up from the community. On a continual basis, users from Member Institutions will be able to submit one-page proposals for new CIG software development tasks. These proposals can be submitted at any time and are posted on the web for the community to read and evaluate. There will be a comments page where members of the user community can add scientific comments and evaluation. Periodically, but at least once per year, the SSC will evaluate these proposals in light of other information obtained from the community, formulate a prioritized list of tasks, and then submit it to the Executive Committee.

By the end of the five-year CIG award from NSF, it will be important for CIG to understand clearly the scientific impact that we have had. CIG is a novel project for our Earth science community and therefore it will be essential that we understand how the CIG software has been used and what concrete scientific advances have been made. How has our community changed the way it does science and has this lead to scientific advances that could not have been made without CIG? In order to answer these and related questions, CIG will need to develop metrics on how our community is using the CIG codes. It is possible that these metrics could involve such things as, lists of papers and abstracts resulting from the use of CIG software, one-abstracts submitted by members of our community highlighting important scientific results, and statistics on software downloads and use. We do not yet know what these metrics will be, but for the coming year (starting Sept. 1, 2006), the Executive Committee will charge the Science Steering Committee with coming up with a set of metrics and a plan to implement these metrics within our activities.

At its disposal, the Executive Committee will have resources to respond to the evolving community needs expressed through these task lists, including the Software Development Team and funds for contracts. However, the Executive Committee will also put into place two mechanisms for generating new resources and funds for CIG.

- *Augmented funding*. CIG will agree to develop additional software upon receipt of augmented funding. For example, a PI at a Member Institution may submit a science proposal to a federal agency in which the proposed work is either wholly or in part dependent upon software not yet available. This software would presumably be more specialized than the highest priority and core CIG tasks, but still encompassed within the mission of CIG and needs of the community. Following submission of a one-page proposal as described above, the Executive Committee will determine whether or not CIG can develop this software. If CIG can develop the software, the EC will detail the resources and funding required on a form for attachment to the PI's proposal. If the proposal successfully passes through peer review and the federal agency agrees to fund the project with augmentation to CIG funding, we will develop the software.
- *Collaborative proposals*. CIG has a specialized staff with skills in software development, numerical analysis, information technology, and related fields, skills not readily accessible within the geoscience community. We believe that members of the community will formulate collaborative research projects with SDT members. If such collaborative projects are judged to be of high merit for CIG by the EC, CIG will develop collaborative proposals. We expect one target of opportunity to be federal programs that require collaboration between scientists from both information technology and the domain sciences, such as the geosciences. It would be expected that such projects would provide funding for both external PIs and members of the SDT.

Software developed through either of these two mechanisms will be open source and made available to the community without restriction, like all CIG software. During its first two years of operation CIG must by necessity focus on a core set of objectives, and would most likely be unable to respond to proposals through these two mechanisms of expansion and funding. However, these two approaches would likely play an increasingly large role within CIG after its formative period.

10. Annual CIG allocations and expenditures

see next page

11. Additional funding

none

	FY 2006-7 Year 3	FY 2007-8 Year 4	FY 2008-9 Year 5
Senior (Total)	*$55,000*	*$56,650*	*$58,350*
Engineers (Total)	*$413,232*	*$425,629*	*$438,397*
Support (Total)	*$97,000*	$99,910	$102,907
Fringe	$144,134	$148,458	$152,912
Overhead	$422,073	$434,735	$447,777
Total Salaries/Benefits	*$1,131,439*	*$1,165,382*	*$1,200,343*
CIG Travel (15 trips)	*$23,895*	*$24,899*	*$25,947*
CIG Visitors (10 Trips)	*$20,735*	*$21,517*	*$22,330*
Total Travel	*$44,630*	*$46,416*	*$48,276*
Total Material and Supplies	*$38,440*	*$40,853*	*$43,291*
Total Subcontracts (3)	*$215,488*	*$221,063*	*$226,806*
Participant Costs			
Workshops (5)	$70,000	$72,100	$74,263
Training Sessions (2-3)	$30,000	$30,900	$31,827
Total Participant Cost	*$100,000*	$103,000	$106,090

Total Budget Amount	($1,529,997)	($1,576,714)	($1,624,806)	Final Balance
Carry Forward	$231,516	$201,519	$124,805	0
Budget Request Amount	$1,500,000	$1,500,000	$1,500,000	
Balance	$201,519	$124,805	$0	