Computational Infrastructure for Geodynamics (CIG)

NSF Award Number EAR-0426271

Five-Year Strategic Plan:
Sept. 1, 2008 to Aug. 31, 2013

Table of Contents

Community Oversight

Executive Committee (EC)

Marc Spiegelman (Chair, Sept. 2009), Columbia University
Bradford H. Hager (Vice Chair, Oct. 2008), Massachusetts Institute of Technology
Alan Levander (Sept. 2010), Rice University
Carolina Lithgow-Bertelloni (Sept. 2010), University College, London
Peter Olson (Sept. 2010), Johns Hopkins University
Brad Aagaard (*ex-officio*), USGS, Menlo Park
Michael Gurnis (*ex-officio),* California Institute of Technology
Michael Aivazis (*ex-officio),* California Institute of Technology

Science Steering Committee (SSC)

Brad Aagaard (Chair, June 2008), USGS, Menlo Park
Wolfgang Bangerth (June 2008), Texas A&M University
Bruce Buffett (Sept. 2010), University of Chicago
Omar Ghattas (Sept. 2010), University of Texas at Austin
Louise Kellogg (Sept. 2009), University of California, Davis
Laurent Montési, (Sept. 2009), University of Maryland, College Park
Jeroen Tromp (Sept. 2010), Princeton
Shijie Zhong (June 2008), University of Colorado at Boulder
Marc Spiegelman (*ex-officio*), Columbia University

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 in the Earth sciences, including mantle convection, the geodynamo, magma, crustal and earthquake dynamics, and seismology. 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 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 assembled;
- extension of existing software frameworks to interlink multiple codes and data through a superstructure layer;
- a Science Gateway to allow users to start simulations on the TeraGrid via the web;
- 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) whose emphasis is to identify and balance common needs across disciplines. An Executive Committee (EC) provides overall oversight for all CIG activities.

Since the start of the 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 within several different areas, including mantle convection, short time-scale tectonics, long time-scale tectonics, seismology and the geodynamo. During the previous year, we completed all of our original 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.

New opportunities have emerged since the original CIG proposal to NSF, especially in the availability of computer hardware for both capability and capacity computing. Our NSF funding sponsors have also partially reorganized through the creation of an Office of Cyberinfrastructure that will likely influence any future initiative in computational geophysics. Given this backdrop, we are now seriously evaluating our long-term objectives and formulating a plan for continuing CIG beyond the current funding. These plans are described in Section 5 and will likely evolve over the next twelve months.

The Science Steering Committee (SSC) developed this year's Strategic Plan (SP) with assistance provided by the CIG staff. Each committee member polled different sub-disciplines of our respective sub-communities and submitted written descriptions that included accomplishments and goals organized over the short-, intermediate-, and long-terms. An SSC meeting in Pasadena on May 29-30, 2008, was held specifically to develop this SP. During the meeting, we reviewed our accomplishments, refined our objectives for the final year of current CIG funding, and had extensive discussions on the future of CIG beyond its currently funded lifetime. The CIG Executive Committee approved this Strategic Plan on June 30, 2008.

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 greater interaction among different geodynamic 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 to eventually 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 that extend beyond our currently funded lifetime.

Multi-Scale Physics. Processes that operate on a wide range of scales are a fundamental feature of the crust, lithosphere, convecting mantle, and core. For example, deformation of the lithosphere can be localized in space into narrow fault zones with brittle deformation localized in time via earthquake ruptures; mantle dynamics with strongly temperature-dependent rheology involves the instability of very thin thermal-rheological boundary layers. 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. CIG is beginning application of adaptive mesh refinement techniques to mantle convection and magma dynamics, but additional work is needed to improve scalability. Furthermore, alternative multi-scale techniques may be necessary to resolve some processes, particularly those operating on a wide range of temporal scales.

Coupled and Whole Earth 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 crust, lithosphere, convecting mantle, and core. While couplings between the behavior of different parts of the Earth are 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. 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 build upon the data structures used to implement adaptive mesh refinement and will take advantage, as much as possible, of existing software like equation solvers.

Inverse, Adjoint, and Data Assimilation Methods. Advanced geodynamics models that attempt to mimic behavior of Earth processes require quantitative evaluation using all pertinent observations. Models that incorporate multi-scale physics and coupling of multiple processes will require development of inversion techniques that can accommodate these advanced formulations. Furthermore, geodynamics codes will need to include both the solution and its adjoint in solving forward problems in order to provide quantitative measures of the resolution of the model. These developments will enable transition from deterministic inversions to statistical inversions, which include quantification of the probability density functions of the parameters.

Petascale Computing. Petascale computing will play an important role in future Earth science research in general, and for geodynamics in particular. Petascale machines open the possibility of exploring geodynamic phenomena over a much more extensive range of spatial and temporal scales than can be done at present. For the most part, approaching realistic conditions and parameters in geodynamic computations is linked to broadening the range of spatial and temporal scales that a model simulates, as well as incorporating physical and chemical processes over each of these scales. Petascale computing has the potential to transform geodynamics modeling if we can develop software that both scales well on computers with $>10^4$ processors with equation solvers capable of handling the extreme ill conditioning of geodynamic problems. The modeling, algorithmic, and software developments and the community training needs that will be required to make use of a petascale computer are daunting. There will be an enormous need to support activities (like CIG and others) that help provide researchers with the proper tools to move up from one scale of computing to the next. This is especially important at the expected bottlenecks for petascale computing where performance scaling becomes ever more of a problem.

Earth Structure Model Frameworks. Unified data structures for representing the physical and chemical properties of the Earth will be of immediate value across all the CIG disciplines. Although these issues transcend CIG (not being a major contributor to data generation or storage), such frameworks are 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. Other Earth data such as thermodynamic and transport properties, composition models, plate motions, plate boundary type and location, seismicity, and crustal stress distributions, are properties 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 is also 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 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 EarthRef, IRIS, GEON, etc., are more focused on data. However, CIG should work with others, especially those involved with planning for a National Geoinformatics system, so that the methodologies of an Earth structure framework are appropriate for constructing dynamic models and linking those models with data.

3. CIG Accomplishments

3.1 Infrastructure. An important aim is to introduce good software design practices into the 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 reusable 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. We use the open-source package Subversion (SVN) for the main CIG software repository, which contains most of CIG's codes, and a Mercurial repository (hg) for the magma dynamics code *MADDs*. The entire contents of both repositories 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 in the SVN repository occurs either nightly or automatically in response to a software commit using *CIG-Regresstor*, a collection of Python codes written by CIG engineer Luis Armendariz. This software uses *Buildbot*, extended by CIG engineer Leif Strand, 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 gives users of the repository confidence in the robustness of the software. Strand extended *Buildbot* so that executable binaries for common platforms are automatically generated.

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 Matthew Knepley at Argonne National Laboratory (ANL) in collaboration with Dimitry Karpeev (ANL). *Sieve* is infrastructure for parallel storage and manipulation of general finite element meshes and can be used so that a developer avoids many of the complexities associated with parallel processing. *Sieve* was used in *PyLith 0.8* and is an integral component of the *PyLith 1.x* release series (see Section 3.3, page 11).

Work continues on software for benchmark intercomparisons. *Cigma* (CIG Model Analyzer), originally implemented in Python, is now written in C++ for improved performance. *Cigma* allows the results of geodynamic models to be compared against standard benchmarks and reports back global and local mismatches in solutions, independent of the method of discretization. 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. *Cigma* beta version .9 was released in July 2007; version 0.9.1 is planned for July 2008. *Cigma* 1.0 will follow shortly thereafter.

The short-term goal for using *Cigma* for the short time-scale tectonics benchmarks (currently only the strike-slip benchmark) has been met. *Cigma* can now compare arbitrary unstructured (hexahedral or tetrahedral) datasets from *PyLith* in a reasonable amount of time. A unit testing infrastructure, using CppUnit, is in place for making sure changes to *Cigma* remain consistent. Since *Cigma* is also available as a library, other codes can use it to perform comparisons inside their respective test cases/unit tests.

The intermediate goal for using *Cigma* in the mantle convection and long-term tectonics communities has also been met. This required adding support for reading parallel XML VTK formats. Currently this support is limited to reading data as unstructured datasets, so very large datasets need to be broken up manually. Codes involved in these benchmarks are *CitcomS* and *Gale*.

3.2. Computational Seismology. A joint CIG-SPICE-IRIS computational seismology workshop was held in Jackson, NH, on October 9-11, 2007. The meeting had approximately 60 international participants. It was decided to work towards a joint SPICE-CIG software library. The future of SPICE (www.spice-rtn.org) is uncertain, but the hope and expectation is to continue this European-U.S. computational seismology collaboration. At the meeting, CIG launched the Seismology Science Gateway, which launches a simulation on a remote machine using data gathered from web sites and databases and returns the results to the user. The Seismology Science Gateway is now live and accommodates request for 1D mode synthetics and 3D SEM synthetics.

Currently available CIG seismology software includes new releases of 1D, 2D and 3D spectral-element method (SEM) codes for seismic wave propagation (collectively 379 downloads in the past six months), a 3D (an)isotropic finite-difference code (47 downloads), the *Mineos* normal-mode software (108 downloads), and links to the SPICE software library and the Princeton dynamic ray tracing software. We soon expect to add a LLNL finite-difference code for regional seismic wave propagation.

3.3. 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 tectonics 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 post-seismic 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.

Working group members have held a workshop each of the last seven 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. Since then Aagaard and Williams, with help from Matthew Knepley, formed a plan to merge their codes into a single suite of modules, *PyLith*. *PyLith* 0.8 was released at the June 2006 workshop and *PyLith* 1.0 was released at the June 2007 workshop. Since the June 2007 workshop, CIG has released *PyLith* 1.1 and 1.2 along with three bug fix releases. Each *PyLith* release includes binary executables for several common platforms, source code from the SVN repository, and a user manual with numerous examples. *PyLith* can solve 1D, 2D, and 3D quasi-static and dynamic elastic and viscoelastic problems with prescribed slip on faults, Neumann (traction) boundary conditions, or Dirichlet (displacement and/or velocity) boundary conditions with parallel processing afforded through the use of *Sieve* and *PETSc*.

During the June 2007 workshop, the short-term crustal dynamics working group set three main priorities for the coming year: (1) continue *PyLith* development, including adding output of fault and state-variable information, a generalized Maxwell viscoelastic material model, traction boundary conditions, and support for computation of 3D Green's functions; (2) infrastructure for archiving and comparing output associated with the suite of working group benchmarks; and (3) catalog available semi-analytic codes and assess their suitability for benchmarking 3D codes, such as *PyLith*. Discussions via the working group listserv further refined priorities for the year with greater emphasis on implementing *Tecton* features in *PyLith*, such as velocity boundary conditions, multiple earthquake ruptures on a fault, and gravitational body forces, rather than adding advanced new features, such as support for computing 3D Green's functions.

PyLith 1.1 provided a new output implementation with user-friendly specification of output of solution fields and state variables as well as boundary condition and fault information. This release also included a generalized Maxwell viscoelastic model, traction boundary conditions, velocity boundary conditions (as part of the Dirichlet boundary condition formulation), a simple absorbing boundary condition, and explicit time stepping for wave propagation problems. *PyLith* 1.2 focused on integrating efficiency improvements in *Sieve*, implementation of gravitational body forces, and support for multiple earthquake ruptures on a single fault.

Aagaard, Williams, and Knepley employ unit and regression testing throughout the entire development cycle, so that the functionality of each module of the code receives thorough testing continuously from initial implementation to release. These 700+ tests are run on multiple platforms whenever changes to the software source code are committed to the subversion repository, aiding the SDT in identifying bugs early in the development cycle. *PyLith* has been used extensively in the 2006 and 2007 “Workshops on Community Finite Element Models for Fault Systems and Tectonic Studies” and 2008 “Workshop on Numerical Modeling of Crustal Deformation and Earthquake Faulting,” which were held on the campus of the Colorado School of Mines and which CIG partially supported.

In 2006 the working group also began using the CIG benchmarking infrastructure for comparing output from *PyLith*, *GeoFEST* (a code similar to *PyLith* 0.8), and *COMSOL* (formerly called *Femlab*) for community established crustal deformation benchmarks. Luis Armendariz (CIG) implemented quantitative metrics to compared solutions across the codes in addition to solutions from semi-analytic methods. The benchmark comparisons demonstrated that *PyLith* 0.8 and *GeoFEST* produce essentially identical results for elastic problems. *PyLith* 1.x releases also produce the comparable elastic solution and demonstrates the greater efficiency (reduced runtime and smaller global error) for linear hexahedral finite-element cells compared with linear tetrahedral finite-element cells. The benchmarking effort is ongoing with an emphasis on identifying the source of differences in the viscoelastic time-dependent solutions between *PyLith* (0.8 and 1.0) and *GeoFEST*, and developing benchmarks more directly related to research problems to test the present and future capabilities of 3D crustal deformation modeling codes in a more realistic setting.

3.4 Long-Term Tectonics. The most significant achievement in the long-term tectonics group has been the continued development of *Gale*, which solves problems related to orogenesis, rifting, and subduction with coupling to surface erosion models. Development of this code was initiated in response to a recommendation from the “2005 Breckenridge Workshop.” The wide availability of a 3D code able to handle large deformations, while accurately tracking material properties, has been an important long-term goal for lithospheric deformation/long time-scale tectonics. In collaboration with the Victorian Partnership for Advanced Computing (VPAC) and Louis Moresi’s group at Monash University, CIG software engineer Walter Landry released a developer’s version of *Gale* in 2006. Subsequent versions of *Gale* have been released as development proceeds. *Gale* can simulate a range of geophysical phenomena related to long-term tectonics, including shortening, extension, and subduction, all coupled to thermal evolution, with deforming upper boundaries. *Gale* can incorporate a wide variety of boundary conditions in 2D and 3D in serial or parallel, and scales to hundreds of processors.

In June 2008, CIG released the most recent version, *Gale* 1.3. This code contains numerous improvements and documentation over previous versions. The current version includes results for all of the GeoMod 2004 benchmarks, which are also described in the manual. The short-term goals for further development include merging *Gale* back with the VPAC code to provide *Gale* with multigrid solvers, restart capabilities, and improved *PETSc* compatibility; improving normal stress boundary conditions; implementing deformed lower boundaries; and interfacing with custom surface process models in parallel. The working group and CIG staff also plan to implement the GeoMod 2008 benchmarks.

3.5 Mantle Convection. Through the hard work of Eh Tan and Wei Leng, a graduate student of Shijie Zhong's group at U. of Colorado (Boulder), *CitcomS* with compressible mantle convection was released in November 2007. The code passed important benchmarks for Stokes' flow. Shijie Zhong extended benchmark work for *CitcomS* and has submitted a paper documenting the benchmark results (coauthors include Allen McNamara, Eh Tan, Louis Moresi and Michael Gurnis). In the meantime, the working group on compressible mantle convection (2D/3D Cartesian models) consisting of Scott King (Virginia Tech), Peter van Keken (Michigan), Louis Moresi (Monash), and Shijie Zhong, continued making progress in formulating benchmark problems and implementing numerical methods for compressible mantle convection problems.

The working group for an analytic solutions code, comprised of Thorsten Becker (USC), Bernhard Steinberger (Norway), Rick O'Connell (Harvard), and Carolina Lithgow-Bertelloni (University College, London), continued to work on the HC code that is currently in the CIG software repository.

Omar Ghattas (UT Austin), Michael Gurnis (Caltech), and Shijie Zhong, under separate funding from the NSF-PetaApps program, worked together to develop a prototype next-generation mantle convection code, *RHEA*, that builds on the parallel mesh-adaptivity library *Octor* developed by Tiankai Tu. *Octor* incorporates distributed data structures and dynamic load balancing and has been scaled to several 10,000 processors. The design of the new mantle convection code is influenced by *CitcomS*, and extends the data structures, discretization, and solver components to accommodate adaptivity and the wide range of spatial scales involved. This new code has already been shown to solve mantle convection problems with acceptable scaling on up to 32k processors. Its components are currently improved to scale to even larger sizes. Future work will extend the codes' current capability from cube geometries to spherical shell geometries, either by embedding the sphere into a Cartesian octree, or else via decomposition into a forest of octrees, as employed in the *deal.II* library. One of the goals for developing this prototype code is to assess and confront the issues encountered in scaling up AMR methods so they can handle challenging geodynamics problems.

The Mantle convection community played an important role in organizing the workshop on adaptive mesh refinement (AMR) in Boulder, Colorado, in October 2007. The mantle convection community is also a major driving force behind the workshop on lithospheric and mantle dynamics to be held at UC Davis in July 2008.

3.6 Magma Dynamics. Magma dynamics is an extension of both mantle convection and long-term tectonics modeling to consistently include the dynamics of low-viscosity reactive liquids such as hydrous fluids and magma in strongly deformable media. As such it is one of the prime examples of a coupled multi-physics problem in the CIG suite of problems. Activity of the Magma Dynamics working group this year centered on development and implementation of the Magma Dynamics Demonstration Suite (*MADDs*), which is a collection of exercises designed to demonstrate the feasibility of modeling magma migration as described by the equations of McKenzie [1984] using a variety of software models. The five individual *MADDs* benchmarks were defined at the 2006 New York City meeting of the Magma Dynamics community and described in the document *An Introduction and Tutorial to the "McKenzie Equations" for Magma Migration*, by Spiegelman, Katz, and Simpson, available on the CIG web site. The five benchmark models (hereafter described as *MADDs*-1 to *MADDs*-5) are a systematic series of problems to demonstrate all the critical components for coupled solid-fluid dynamics. The first series implements solid-flow solvers for 2D and 3D Mid-ocean Ridge spreading geometries (*MADDs*-1) that have accurate advection schemes (*MADDs*-2) and accurate dynamic pressure fields (*MADDs*-3) sufficient for including the behavior of low viscosity fluids. The 4th benchmark tests the accuracy of the full fluid-transport codes against analytic solutions for magmatic solitary-waves in 2D and 3D (*MADDs*-4). The final benchmark combines all of the previous models into a demonstration model for magma dynamics at mid-ocean ridges with forced adiabatic pressure release melting.

The bulk of this year's activities has been spent implementing the first four benchmarks in the *StGermain* framework, which is the same underlying code-base from which *Gale* is derived. Most of this work was conducted by Dave Lee at VPAC (with assistance from Luke Hodkinson, David May and Marc Spiegelman) through a subcontract from CIG. The resulting implementation and new *StGermain* features required for magma migration are known as *stgMADDs*. In addition, community members Marc Spiegelman (Columbia), Laurent Montési (U. Maryland), and Wolfgang Bangerth (Texas A&M) have implemented various *MADDs* components directly using *PETSc*, *COMSOL Multiphysics*, and *deal.II*, respectively.

MADDs-1, a regular Stokes solver benchmark using a segmented mid-ocean ridge configuration, was implemented on all platforms. It necessitated the development of higher-order elements in *stgMADDs* and novel strategies for incorporating the relevant boundary conditions. The semi-Lagrangian advection scheme of Spiegelman *et al.* [2006] was implemented in *StGermain* for *MADDs*-2 (advection test). *MADDs*-3 (constant porosity ridge model) was handled as a postprocessing step. *MADDs*-4 (solitary waves) is now functional in *stgMADDs*, constituting the first generally available implementation of the coupled viscous porous flow theory. Development of *MADDs*-5 (coupled mid-ocean model with forced melting) is underway.

Other activities are proceeding more slowly, as they are not directly supported by CIG, but instead take the form of a CIG-facilitated community effort. *PETSc*-based software is available for *MADDs*-4. *COMSOL* presented difficulties with the advection test, which are being worked around, but *MADDs*-3 was straightforward. Implementation of *MADDs* in *deal.II* began in the spring of 2008 and is part of activities of the Computational Sciences working group effort to provide geodynamics-specific *deal.II*-based software with adaptive mesh refinement.

Working group members submitted a proposal to NSF-OCE to continue the involvement of the community in this effort and further CIG-mediate magma dynamics activities. The outcome of the proposal is not known at the time of this report.

3.7 Geodynamo. CIG software engineer Wei Mi worked with Peter Olson (Johns Hopkins) in the past year on several CIG projects, including the first release of the geodynamo code *MAG* in September 2006. This was followed by the release of a new version of *MAG*, including a new user manual, at the end of March 2007. The latest *MAG* version includes new capability for output of standardized Gauss coefficients of the external magnetic field as a function of time, plus enhancement of the *MAG* visualization tools for time series analysis, 3D rendering, and animations. In addition, Wei Mi and Peter Olson constructed a suite of example dynamo cases as part of a tutorial for *MAG* beginners on the CIG web site. A web portal for *MAG* was established in March 2008 with the goal of increasing the user base. Efforts to promote the use of the web portal have been targeted at the paleomagnetic community and geoscience educators. Wei Mi has also been working with the research group of Weijia Kuang at Goddard Space Flight Center in an effort to bring their dynamo code *MoSST* into the CIG repository and to benchmark this code against the published community dynamo benchmark cases. To date, the *MoSST* code has been donated to CIG and preliminary comparison cases have been run, but successful benchmarks will require that some additional code modifications be undertaken by the *MoSST* developers.

3.8 Computational Science. In 2007 and the first half of 2008, CIG made significant progress with regard to our long-range goals in computational science. This includes the development of codes that use adaptivity and the organization of two workshops on computational science and mathematics issues in the geosciences.

Adaptive mesh refinement (AMR) has been a goal for the software developed by CIG since its inception. However, the steep learning curve practitioners have to undergo as well as the virtual impossibility of re-engineering existing codes to support adaptive meshes have prevented more significant progress in this area. This changed in 2007 and 2008 with the development of two sets of codes. The first, developed by researchers with partial support from CIG (the main part being funded by an NSF PetaApps grant) is a project under the guidance of Omar Ghattas, Mike Gurnis and Shijie Zhong. This code uses a fully adaptive, time dependent, three-dimensional octree-based mesh to simulate thermal convection in the Earth mantle. It has been shown to scale with acceptable efficiency to up to several tens of thousands of processors on the NSF-funded Ranger system built at the University of Texas at Austin.

The second set of codes is a CIG-funded effort by Wolfgang Bangerth (Texas A&M) to develop a suite of six programs that solve a variety of geodynamics applications with adaptive finite elements based on the *deal.II* software library. A fully adaptive Stokes solver is already available, while an adaptive solver for thermally-driven convection will be available by the end of August 2008. All of these programs are and will be freely available as part of the extensively documented *deal.II* tutorial under an Open Source license.

To jump-start developing an AMR-enabled new generation of geodynamics codes, CIG held an AMR workshop and tutorial on October 24-27, 2007, in Boulder, CO. The goal was for the main lecturer, Wolfgang Bangerth, to spend the majority of the time going through the extensive tutorials of the *deal.II* library (of which he is the principal author) to explain how to implement AMR-enabled models. A small number of other participants presented talks on related topics. The rest of the time was spent in hands-on sessions, working individually or in small groups on building codes with *deal.II* that can solve some simple models of interest to the geodynamics community, and that can serve as starting points for later extension to more complete models. The workshop was attended by approximately 20 members of the community and has already led to a number of collaborations that will yield new AMR-enabled codes for the geodynamics community.

Finally, CIG is organizing a workshop on the interface between Computational Science, Mathematics, and Geodynamics, to be held September 15-17, 2008, in Santa Fe, NM. This workshop will bring together researchers from different disciplines to learn about progress made in related fields and foster new collaborations. The workshop will also help refine priorities for a future initiative after the current CIG funding expires.

4. Details of Our Short-, Intermediate-, and Long-Term Goals

The Science Steering Committee has updated our short-, intermediate-, and long-term goals. All of our original short-term goals and several of our intermediate-term goals have now been completed. We continue to make progress on our other intermediate-term goals. Achieving our long-term goals will be more technically challenging and will involve unifying many of the efforts of the individual disciplines; Section 5 presents our strategy for achieving these long-term goals. All of our goals are summarized in Table 4.1.

4.1 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. CIG engineer Luis Armendariz is completing a new version of *Cigma*, version 1.0. Short-term plans involve writing unit tests that make sure the comparisons involving quadratic or higher-order elements are taking place properly. Another short-term goal is to incorporate additional benchmarks from the other codes such as mantle convection benchmarks from *CitcomS* and the long-term tectonics benchmarks from *Gale*. We plan to run *Cigma* on data collected from the higher-order elements in the *MADDs* benchmark suite, and collect authoritative benchmark data for each of the benchmark cases defined on each of CIG's geodynamics codes. Automatic regression testing development continues, with more *CppUnit* examples being developed which other codes can use, with the *Cigma* library API automatically checking for consistency and adherence to community benchmarks between different versions of the code. CIG staff will collaborate with individual communities in order to adapt and refine this procedure. We expect that *Cigma* software will be used by individual investigators studying their own benchmarks, while also being incorporated within our automatic regression testing.

Mid-term goals include making it easier to break up the integration meshes by using existing partioning libraries such as Chaco (http://www.sandia.gov/~bahendr/chaco.html) and Metis/ParMetis (http://glaros.dtc.umn.edu/gkhome/views/metis/). This will allow automatic processing of larger datasets than are currently possible. We also plan to refine the current schema for representing numerical models in HDF5 files, to provide higher-level support for scientific datasets.

A remaining long-term goal is to make *Cigma* available through our Science Gateway. This will allow comparisons against the collected suites of community benchmarks.

4.2 Computational Seismology. CIG will work to enhance the Seismology Science Gateway by adding automated simulations to the current on-demand simulation capabilities. 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 CMT catalog. The Seismology Science Gateway is now live, and accommodates request for 1D mode synthetics and 3D SEM synthetics. Near-real time automated simulations should be available within the next six months.

The computational seismology community needs 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. The Seismology Working Group plans to implement mechanisms for the contribution of models to the Seismology Science Gateway.

The SSC is considering whether CIG should investigate the feasibility of facilitating the development of data processing tools for field and laboratory use. These might include low-level routines for standard data manipulation (e.g., filtering, simple array analyses); higher-level functionality such as earthquake location, travel-time 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 within the CIG framework.

4.3 Short Time-Scale Tectonics. Working group members have voiced strong support for continued development of *PyLith*. Community members requested that all features present in *Tecton* be transferred into *PyLith* by January 2009.

In response to these requests, CIG plans additional releases of *PyLith* (1.3, 1.4, and 1.5) to transfer all features present in *Tecton* into *PyLith*. Release 1.3 will add adaptive time stepping and support for specifying an initial stress state. Release 1.4 will include integration with *PETSc*'s nonlinear solvers to support nonlinear bulk rheologies and fault constitutive models. This addition of fault constitutive models will also mark the transfer of all features present in *EqSim* into *PyLith*. Another target for this release will be improved scaling for simulations using tens to one hundred processors for both quasi-static and dynamic problems. Release 1.5 targets large deformations and finite strain (without remeshing) and implementing a more generalized description of time dependence in Dirichlet boundary conditions and incorporating this into the Neumann (traction) boundary condition.

The current SDT does not have sufficient resources to meet the community's timetable of completing these releases by January 2009. CIG is working with the community to identify additional personnel with the necessary geoscience and computational science background to expedite *PyLith* development. Upon completion of these releases, *PyLith* will contain all of the functionality present in its ancestors as well as many additional features in the form of an open-source, modular, scalable, extensible software package with comprehensive documentation and thorough unit and regression testing.

Future releases beyond 1.5 will target more advanced features, such as support for automatic computation of 3D Green's functions and direct coupling of the quasi-static and dynamic problems, providing multi-scale resolution of the earthquake cycle.

Members of the short-term tectonics community are cataloging available semi-analytic codes that can be used for benchmarking *PyLith* and other quasi-static crustal deformation codes. The objective is to make these codes available via the CIG web site and use these codes in benchmark tests of 3D codes.

The short-term tectonics community envisions that the techniques and tools used to reengineer other CIG codes for adaptive mesh refinement will be applied to *PyLith*. Long-term goals also include integration of tools for formal data assimilation in order to permit data from EarthScope and other sources to be included directly into simulations.

4.4 Long-Term 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 mechanisms present significant challenges in developing numerical codes to simulate such deformation. Thus, very different approaches are being used, and there is the need for making different approaches available to a wide community and encouraging benchmarking of those codes.

One of our primary goals has been expand participation by the community through the long time-scale tectonics working group for the setting of priorities for future development of *Gale*, participating in the development of adaptive mesh refinement, and establishing long-term priorities for long-term tectonics. A working group has been formed with co-leads Dennis Harry (Colorado State) and Mousumi Roy (New Mexico), and includes Thorsten Becker (U. of Southern California), Todd Ehlers (Michigan), Noah Fay (Arizona), Ritske Huismans (Bergen U.), Carolina Lithgow-Bertelloni (University College, London), Dietmar Muller (U. of Sidney), Patrice Rey (U. of Sydney), and Jolante van Wijk (Los Alamos). This group held a meeting at the Fall 2007 AGU meeting and has worked by conference call to develop priorities for long-term tectonics.

There is a considerable need for benchmarking of codes. As mentioned in Section 3.2, running established benchmarks has been one focus of *Gale* development over the past year, and it remains a priority. This is of particular importance for codes that deal with localization phenomena such as those that simulate tectonic fault development. CIG is cooperating 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 GeoMod2007 meeting in Europe, Buiter organized a two-day pre-conference workshop aimed at discussing results of a new numerical benchmark, new analog benchmarks, new numerical-analog comparisons, and modeling techniques. The benchmarks established by this group focus on two experiments: extension and compression of a layered medium, in a configuration representing published laboratory experiments. The GeoMod2004 benchmarks have been computed on *Gale* and the GeoMod2008 benchmarks are next. CIG will support U.S. participation in activities related to the further development of these benchmarks.

For our longer-range goals, incorporation of adaptive mesh refinement into our codes remains a high priority. AMR provides the efficiency needed for solving large-scale problems with high spatial resolution in regions of strain localization and fluid/rock interactions. There is also an interest in coordinating the further development of *Gale* with the efforts of the Magma Dynamics Working Group to, for example, calculate the volumes of melts predicted in extensional tectonics models. In addition, there is an interest in extraction of information such as the predicted geobarometer and thermochronometer observables at the surface of *Gale* simulations. The long-term tectonics group has also identified the need for *Gale* to include the capability of specifying pressure-, temperature- and depth-dependence of material properties.

4.5 Mantle Convection. Among several long-term goals identified by the mantle convection community at the 2005 Boulder workshop, the most challenging ones are to develop compressible spherical convection codes and convection codes with adaptive mesh refinement. The compressible spherical convection is particularly important for a better integration of mineral physics and seismology with deep-Earth dynamics. A code with adaptive mesh refinement capability helps couple small-scale physics, especially lithospheric deformation and melting/melt migration, into large-scale mantle flow models.

CitcomS with compressible convection capability was released in November 2007, with its latest release, version 3.0.2, in March 2008. We do not anticipate further significant development, although additional benchmarks and optimizations are needed. Incorporation of more realistic thermodynamic formulations from the Stixtride and Lithgow-Bertelloni group into *CitcomS* with compressible convection is under consideration.

New development efforts will focus on codes with adaptive mesh refinement, in particular *RHEA*. In many ways this will be our next generation mantle convection-modeling tool. Many modules of these codes, such as meshing, element types, and solvers, may be shared with software used in long-term tectonics, short-term tectonics, magma dynamics, and seismic wave propagation. Consequently, important synergies exist with these other areas, and CIG will attempt to vigorously pursue them.

Additional future projects include (1) improving the HC code by adding compressibility and geoid modeling features; (2) adding a 2D convection code (e.g., *ConMan* and *Citcom*) to CIG's software repository; and (3) continuing benchmark efforts, including 2D convection and 3D high Rayleigh number cases with internal heating and time-dependence.

4.6 Magma Dynamics. The feasibility of including magma dynamics in *StGermain* and other codes is underway through the Magma Dynamics Demonstration Suite (*MADDs*), a series of benchmarking exercises with progressive levels of sophistication based on the McKenzie [1984] theory of magma migration in a viscously deformable porous medium. The sequence of benchmarks was defined at the August 2006 Magma Dynamics workshop and is described in the document, *An Introduction and Tutorial to the "McKenzie Equations" for Magma Migration*, by Spiegelman, Katz, and Simpson, available on the CIG web site in the Magma Dynamics Working Group's work area.

The primary tasks for the next year include: (1) finishing the implementation of the full two-phase mid-ocean ridge model with forced adiabatic melting (*MADDs*-5) and possibly the melt-band localization problems (Katz [2006]*) (MADDs-6)* using *StGermain*; (2) including *MADDs*-1 (segmented mid-ocean ridge), *MADDs*-4 (solitary waves), and *MADDs*-5 (coupled magma migration and mid-ocean ridge with force melting) in the *deal.II* geodynamics tutorials; and (3) determining a rigorous and complete description of benchmark setup and boundary conditions, and defining a quantitative evaluation strategy. Implementation of magma dynamics with *deal.II* will enable new models with localization phenomena such as channelization and reaction fronts. Furthermore, the interoperability of *StGermain*-based software should be tested via a test involving both ALE (*Gale*) and magma dynamics (*stgMADDs*).

An intermediate-term goal will be to build the user base and crystallize the magma-migration community around CIG activities. The OCE proposal has helped in this direction but further effort is needed. High on this list is the short-term goal of publishing detailed benchmark definitions. Furthermore, CIG would like to convene a community/ users meeting to publicize the capacity of the CIG software, train users, and identify research directions of interest to the magma dynamics community. This meeting could be held together with the long-term tectonics community, who expressed the desire of including melting in *Gale*. A joint meeting would stress the interchangeability of the software platform underlying *Gale* and *stgMADDs* and facilitate interaction between the two communities.

Further intermediate- to long-term goals are to further explore the possibility of code coupling. Beyond coupling with *Gale* (intermediate goal), it is important to integrate thermodynamics and geochemical modeling. The extent of progress in this direction will depend upon success of the OCE proposal, in which the role of CIG is principally to organize the science community. Further activities include a standard interface with geochemical databases (e.g., EarthChem), computational seismology (*SPECFEM*) and the modeling of localization phenomena (dissolution channels, dikes). Data assimilation and further model development in the context of magma dynamics will likely require significant interaction with computer sciences, including algorithm development for coupled, multi-physics models, and will be a topic discussed at the Santa Fe meeting.

4.7 Geodynamo. Dynamo codes represent a powerful 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 as 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, efficient parallelization, inter-operability, and compatibility with CIG framework standards.

Consultation with members of the U.S. geodynamo and related communities (initiated by Peter Olson) defined a set of priorities for CIG in this area. The first goal is to bring several dynamo codes into the CIG system. In addition to the *MAG* code already donated, benchmarked, and installed into the CIG repository, community members have donated a second serial code, *MoSST*, written and used by Weijia Kuang (Goddard). *MoSST* includes capabilities for rotational interactions between the mantle and the outer and inner core.

A number of additional priorities were identified in the course of polling community members. The most common responses included (1) standardization of existing dynamo codes to allow inter-comparison of results; (2) comparison of benchmark results between the serial codes in the CIG repository; (3) development of common visualization tools; (4) 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 the planetary and space sciences; and (5) development of educational applications using the web portal. All of these priorities are worthy short-term objectives for CIG. However, progress on these objectives is limited by the participation of community members in the Geodynamo Working Group. Efforts to foster and grow the Geodynamo Working Group will be a primary short-term of objective of CIG because the existing software will not fully exploit current advances in computational capabilities.

The long-term strategy for CIG is to facilitate the development of the next generation of dynamo models. A primary focus is to reduce or eliminate some of the limitations in existing codes. For example, the dynamo codes now in use were not constructed with a modular design. Other disadvantages stem from their reliance on traditional spherical harmonic and Chebyshev polynomial expansions. Advantages of the spherical harmonic and polynomial representations include that divergence-free magnetic and velocity fields result automatically, the pressure term is eliminated using the vorticity equation, the boundary conditions are also easily incorporated in the solution procedure for the interior, and they naturally result in better spatial resolution in boundary layer regions. Chebyshev polynomials are difficult to parallelize, 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 and a replacement for Chebyshev polynomials that allows for parallelization. Spherical harmonics will continue to be used in dynamo codes for satisfying magnetic boundary conditions and other specialized uses. A local basis function formulation to replace spherical harmonics, and perhaps more urgent, a finite difference or other local method of calculating radial derivatives, offer clear advantages for parallelization, implementing sub-grid scale physics, variable properties, and adaptive re-meshing. A first step in realizing this longer-term CIG strategy is to convene a workshop to discuss strategies for the next generation of software. Interested participants would be encouraged to form an active CIG working group.

4.8 Computational Science. CIG has achieved many of its original computational science goals regarding the creation of software based on reusable and well tested libraries. For example, several CIG codes use the *Pyre* framework to control program flow, and virtually all use *PETSc* for parallel linear algebra and solvers. This allows for simple plug-and-play experiments to optimize solver strategies since it builds on an extensive framework of widely used components.

Since last year's report, CIG has made significant progress integrating adaptive mesh refinement into geodynamics codes, a technique that has been consistently mentioned as one of the important goals in the field and that has been mentioned in every Strategic Plan so far. The reasons for the previously slow progress have varied and range from a lack of experience in the community to the fact that it is considered very difficult to modify existing codes for AMR since it typically leads to pervasive changes throughout programs, especially in the underlying data structures. Moreover, AMR remains at the forefront of research in computational science, and truly general solutions that scale well on parallel computers are not readily available.

A more promising, if more radical strategy is to completely rewrite those parts of the program that deal with discretized problems, and only keep the control structure. An important component is to reuse the experience gained with existing codes. This recognizes that the main investment in most computational science codes is not the code itself, but the knowledge of which linear and nonlinear solvers work, how to optimize them, and the conditions under which they are applicable. If sufficiently rich software frameworks for computational science are used, the effort to recreate these solvers under a new framework should be far smaller than the original implementation in legacy code.

One of CIG's medium-range goals is to encourage the incorporation of AMR into a new set of codes that use the algorithms, solvers, and experience from previous generations of codes but are built atop existing software libraries supporting AMR, following the strategy outlined above. This year a collaboration was begun among SSC members Shijie Zhong and Omar Ghattas and postdocs at U. of Texas at Austin as well as Mike Gurnis to develop a prototype next-generation mantle convection code that builds on the parallel mesh-adaptivity library *Octor* developed by Tiankai Tu. *Octor* incorporates distributed data structures and dynamic load balancing and has been scaled to several tens of thousands of processors. The design of the new mantle convection code is influenced by *CitcomS*, and extends the data structures, discretization, and solver components to accommodate adaptivity and the wide range of spatial scales involved. This new code has already been shown to solve mantle convection problems with acceptable scaling on up to 32k processors. Future work will extend the code’s current capability from cube geometries to spherical shell geometries, either by embedding the sphere into a Cartesian octree, or else via decomposition into a forest of octrees, as employed in the *deal.II* library. One of the goals for developing this prototype code is to assess and confront the issues encountered in scaling up AMR methods so that they can handle challenging geodynamics problems.

Current goals also include further development of demonstration codes for various applications that use adaptive finite-element methods based on *deal.II.* This project has already started, and will include the generation of at least five more extensively detailed tutorial programs that will serve as the starting point of experiments for interested parties.

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 a 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 lay out 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. Additionally, the Science Steering Committee is assessing the feasibility for a meeting of the entire CIG community in the spring of 2009. The meeting would provide an opportunity for scientists within the various disciplines to interact and gain a broader understanding of CIG software and the research being done with it. The format would likely combine hands-on tutorials with talks highlighting research accomplished using CIG software. Some of the planned workshops might immediately precede or follow the meeting.

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 communication with and among members. CIG currently uses a technology called Plone, a user-friendly and powerful open-source Content Management System, which 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 mailing lists for both the entire community (cig-all@geodynamics.org), each committee, and each subject area. These e-mail lists are cross-linked so that a user can easily navigate from the Plone editable areas to the mailing list archives.

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 are 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 users consists of traditional computational scientists who 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 users 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 datasets).

Several mechanisms are used to train CIG users: (1) small, focused workshops; (2) visits to the CIG site to work with CIG staff and other users; and (3) 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
Infrastructure	Write unit tests for *Cigma* to make sure quadratic or higher-order element comparisons are correctly executed. Run *Cigma* with *MADDs* benchmark suite data.	Facilitate break up of integration meshes to allow automatic processing of larger datasets. Refine schema for representing numerical models in HDF5 files.	Develop *Cigma* science gateway.
Computational Seismology	Build up a core group of users of Seismology Science Gateway using *SPECFEM* on the TeraGrid. Add more 3D models to the gateway. Coupling of *SPECFEM* and *Citcom* (Earth models).	Allow user-defined 3D models in the Seismology Science Gateway. Distribute LLNL WPP code via CIG web site. Continue work towards coupling *SPECFEM* and *Citcom* (Earth models).	Refine mechanisms for contribution of user-defined 3D models. Integrate data processing and visualization tools with CIG codes and the Seismology Science Gateway.
Short Time-Scale Tectonics	Continue *PyLith* development with transfer of all *Tecton* and *EqSim* features into *PyLith*.	Direct support for computing Green's functions in *PyLith*. Coupling of quasi-static and dynamic simulations for earthquake cycle modeling.	Adaptive mesh refinement capabilities in *PyLith* or a code with similar functionality. Integrate tools for formal data assimilation.
Long-Term Tectonics	Continue *Gale* development and benchmarking (GeoMod2008). Encourage construction of numerical benchmarks for *Gale* and other long-term tectonics codes. Explore coupling with magma generation.	Receive donation of *SNAC* code, port to CIG build procedure, and refine User Manual for general use. Start incorporation of AMR in code for long-term tectonics.	Develop new code with adaptive mesh refinement.
Mantle Convection	Continue development of *RHEA* and *deal.II* examples. Release compressible *CitcomCU*. Place donated *ConMan* code in repository with updated documentation. Release HC code. Coupling of *SPECFEM* and *Citcom* (Earth models).	Continue development of *RHEA* and *deal.II* examples. Coupling of *SPECFEM* and *Citcom* (Earth models).	Continue development of AMR codes including *RHEA*. Explore the possibility of integrating and sharing common components with other codes.
Magma Dynamics	Complete implementation of MADDS benchmarks with *StGermain*, *deal.II*, *PETSc*-based code, and *COMSOL Multiphysics*. Define precise descriptions of benchmarks.	Training of user base in CIG software with *MADDs*. Initiate coupling with thermodynamics (pending funding). Demonstrate interaction between *Gale* and *stgMADDs*.	Integration of magma dynamics in global mantle convection. Coupling of thermodynamics modeling and *MADDs*-enabled software.
Geodynamo	Advertise *MAG* web portal to recruit users and investigate potential partnership for developing educational module. Encourage formation of a working group via a workshop on 2nd generation geodynamic codes.	Community building workshop in 2009.	New dynamo code using components in common with mantle convection codes, including common Earth structure framework and grid and adaptivity.
Computational Science	Organization of a workshop in Santa Fe, Sept. 2008, on the interface between computational science, geodynamics, and mathematics. Continue development of *RHEA* and *deal.II*-based AMR codes.	Maintain code repository, documentation, and science portals. Improve scalability of *RHEA*. Integrate parallel-distributed meshes into *deal.II.*	Work on integration of computer sciences and geophysics into a single framework.

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 (working version of *Cigma* for short term tectonics)	Benchmarking (complete *Cigma* for multiple communities).
Regression testing (developed *CIG-Regresstor*)	
Launcher package (developed *Addendum.py*)	
Software repository (implemented SVN and Mercurial)	

Table 4.3
Cross-Disciplinary Goals – Scientific Computing

Mostly Completed	Long-Term
Sieve code for storing and manipulating finite-element meshes.	Use adaptive mesh refinement. Develop software components that can handle meshes distributed to thousands of processor cores and that adapt to the mesh based on the structure of the solution.
Use common components from *PETSc* and *Pyre* as the basis of most codes.	Develop software interfaces for coupled and whole-Earth models in support of simulations that couple multiple physical processes associated with the dynamics of the crust, lithosphere, convecting mantle, and core.
HDF5 – output of codes in similar format for ease in benchmarking and for future visualizations.	Use common data structures for representing the physical and chemical properties of the Earth.

Table 4.4
Workshops Planned

WORKSHOPS	
Computational Science	CIG Workshop on Mathematical and Computational Issues in the Solid Earth Geosciences, to be held Sept. 15-17, 2008, in Santa Fe, NM.
Geodynamo	Convene a workshop to discuss strategies for developing a new generation of geodynamo modeling codes. Participants will be encouraged to form an active CIG working group.
Magma Dynamics / Long-Term Tectonics	Joint meeting in late spring or early summer of 2009 focusing on user training and defining community priorities.
Mantle Convection	2008 Workshop for Advancing Numerical Modeling of Mantle Convection and Lithospheric Dynamics, to be held July 9-11, 2008, UC Davis, CA. Discussion items: Science, current software development, and future plans for new code with adaptivity and AMR.
Short Time-Scale Tectonics	Co-sponsor meeting in 2009, continuing emphasis on user training and application of CIG software in crustal deformation research.

5. Achieving Our Long-Term Goals and the Future of CIG

The Science Steering Committee has concluded that we have achieved the short-term goals originally described in the CIG proposal to NSF and have made substantial progress toward achieving many of our intermediate- and longer-term goals as well. New opportunities have emerged since the original proposal, especially in the availability of computer hardware for both capability and capacity computing. Our NSF funding sponsors have also reorganized and created an Office of Cyberinfrastructure and a Cyber-Enabled Discovery and Innovation program that will likely influence any future initiative in computational geophysics. Given this backdrop, we are seriously evaluating our long-term objectives and formulating a plan for achieving them through a future initiative, CIG-II. The details of our plans are still under development and will likely evolve over the next twelve months. This section describes our plans as they currently exist.

The long-term goals of CIG described generally in Section 2 and in detail in Section 4, essentially involve three grand challenge problems: mantle convection and lithospheric deformation (large-scale solid deformation with complex rheologies), magma dynamics (coupled fluid/solid flow in strongly deformable, reactive media), and crustal dynamics and the earthquake cycle (multi-scale brittle mechanics of the crust). Although quite diverse in their scientific scope, these grand challenge problems involve several common activities.

- Geodynamic modeling needs to embrace the many scales of heterogeneity present in nature, including both structural heterogeneities (e.g., the presence of plate boundaries or faults) and temporal heterogeneities (e.g., seismogenic motion and volcanic eruptions). Doing so requires both novel theoretical concepts, centered upon localization, and appropriate numerical techniques.
- Earth materials commonly display highly nonlinear and, in some cases, even discontinuous behavior. For example, plasticity and crack propagation affect development of plate boundaries, magmatic intrusions, and the development of fault networks. Including these two processing in geodynamics codes presents numerous numerical challenges related to discontinuous rheological descriptions, weakening, and the variety of available theories.
- Fluids are increasingly recognized as playing a crucial role in geologic phenomena. Obviously, magma (a fluid) is central to magma dynamics and also affects the geochemical evolution of the mantle and development of plate boundaries. Water and gases, such as CO_2, participate in seismogenesis and tectonics. Although fluids are often included in hydrogeology and petroleum engineering, they are often neglected in solid-Earth research.
- Geophysical data needs to be assimilated into geodynamic models for quantification of uncertainty and discrimination among competing processes. Data used to test geodynamic models is sparse, heterogeneous, and often incomplete. All three grand challenge problems face the difficulties of assimilating information from geochemistry and geology, in additional to the more uniform geophysical information from seismology and potential fields. Integration with databases and implementation of advanced data assimilation

paradigms for CIG codes would significantly benefit both the modeling and data gathering communities.

In undertaking these challenges while pursuing our long-term goals, CIG is rapidly approaching the limit of currently available methods, and further progress will require innovative techniques in scientific computing and their application to geodynamics. The Science Steering Committee believes CIG should seek proactive collaborations with the broader scientific computing community at a level far beyond that envisioned during the initial phase of CIG. This would ensure that new developments in scientific computing target these difficult obstacles in geodynamics, enlighten the scientific computing community of challenges in geodynamics, and bring new untapped talent to the solid Earth science community.

We have identified three principal crosscutting computational issues (summarized in Table 5.1) that impede current models in which progress will require development of new computational techniques. These include discretization and adaptivity for multi-scale physics associated with localization in space and/or time, scalable linear and nonlinear solvers for multi-scale and/or multi-physics problems, and integration of inverse, adjoint, and data assimilation techniques into geodynamic models.

We anticipate organizing CIG-II as two complementary efforts. One effort would focus on collaborative research and development among solid Earth scientists, mathematicians, computer scientists, and computational scientists to advance algorithms, modeling techniques, and software and apply them to geodynamics problems. The second effort would support a central software development team, likely smaller in scale than the one in CIG-I, to maintain the community infrastructure and assist the research and development efforts result by integrating new features into existing codes or distributing new codes to the broader community. These two efforts would work in tandem to further CIG's objectives of developing, supporting, and disseminating state-of-the-art software for the geoscience community from model developers to end users. We have identified five tasks (described in the following sections) associated with the three crosscutting computational issues to serve as the focus of the research and development effort. A small team of researchers composed of a few solid-Earth scientists from relevant working groups along with mathematicians, computer scientists, and/or computational scientists would work on each task. One of the primary objectives of the “Workshop on Mathematical and Computational Issues in the Solid Earth Geosciences” to be held September 15-17 in Santa Fe, NM, will be to refine the vision for these interdisciplinary collaborations. Table 5.2 shows the plan and timeline for seeking funding for CIG-II with the central software development team and support staff funded through the NSF Division of Earth Sciences (EAR) and the research and development effort funded through a Type III award from the NSF-wide Cyber-enabled Discovery and Innovation (CDI) program.

Table 5.1
CIG Grand Challenge Problems and Crosscutting Computational Issues

Grand Challenge Problems	**Crosscutting Computational Issues**		
	Multi-Scale Physics	**Solvers for Coupled Systems**	**Inverse Problems and Data Assimilation**
Mantle Convection and Lithospheric Deformation	Localized deformation due to boundary layers, moving internal boundaries, mantle plumes, and plasticity.	Solvers for large variations in physical properties and ill conditioning arising from boundary layers, plate margins, moving internal boundaries, mantle plumes, and plasticity.	Data assimilation and inversion using datasets, such as plate motions, Earth structure, metamorphism, thermal chronology, and mineral physics.
Magma Dynamics	Coupling of deformation, geochemistry, and petrology along with flow localization.	Solvers for large variations in physical properties and ill conditioning arising from localization, and fluid/solid coupling.	Data assimilation and inversion using datasets, such as geochemical data, thermochronology, seismic wavefields, and seismicity.
Crustal Dynamics and the Earthquake Cycle	Localized deformation due to rheologic variations as well as brittle deformation (earthquake rupture).	Solvers for large variations in physical properties, plasticity, and multiple time scales.	Data assimilation and inversion using datasets, such as geodetic observations from InSAR, LIDAR, and GPS, seismic wavefields, plate motions, and paleoseismic records.

Table 5.2 Timeline for Seeking Funding for CIG-II

Sept. 1, 2008	Final year of CIG starts
Sept. 1 – Nov 1, 2008	Prepare preliminary CDI proposal.
Sept. 15-17, 2008	Workshop on Mathematical and Computational Issues in the Solid Earth Geosciences
Sept. 30, 2008	Letter of intent for CDI proposal due
Nov. 4, 2008	Preliminary CDI proposal
Feb. 27, 2009	Full CDI proposal due
July 9, 2009	EAR Geoinformatics proposal due
Aug. 31, 2009	CIG-I ends. Fund transition to CIG-II with carry over, existing small OCI funding, and potentially MG&G funding.

5.1 Discretization for localized deformation. Approaching realistic conditions and parameters in geodynamic computations is linked to broadening the range of spatial and temporal scales that a model simulates, as well as incorporating physical and chemical processes over each of these scales. Incorporating this multi-scale and multi-physics behavior requires development of discretization techniques (e.g., adaptive mesh refinement schemes with error estimators) for problems with solutions that are localized in space and/or time. Spatially or temporally varying viscosities create localization in mantle convection and lithospheric deformation for boundary layers, plate margins, mantle plumes, moving internal boundaries, and plastic deformation. In magma dynamics, interactions of reactive, low-viscosity fluids with a strongly deformable matrix result in localization. In crustal dynamics, earthquake rupture causes very localized deformation in both space and time. Adaptive mesh refinement techniques show promise for resolving spatial localization, but the currently available software for distributing the dynamically changing meshes is straining to scale into the range of tens of thousands of processors. In particular, further improvements to the solvers for highly locally refined, distributed meshes are going to be a priority for CIG-II. Another area for potential innovation is developing efficient methods for simultaneously handling the localized deformation that occurs in both space and time, such as aseismic and coseismic fault slip in crustal dynamics.

5.2 Multigrid solver for adaptive mesh refinement. Develop a multigrid solver for parabolic and elliptic problems with adaptive mesh refinement for applications in mantle convection and crustal dynamics. Preliminary experience with the mantle convection code *RHEA* shows that scaling to several thousands of cores is feasible without too much complexity, but that the currently used algebraic multigrid preconditioner has difficulties to scale to several tens of thousands of cores. At the same time, we are confident that these problems can be overcome with further research, and that the experience so gained will provide valuable to similar efforts in other areas of geodynamics.

5.3 Inverse problems, adjoints, and data assimilation. Develop adjoint formulations and data assimilation techniques for discriminating among competing models, parameterizing probability density functions, and optimizing experimental design. Most geodynamic models do not quantify the resolution of the model or incorporate an estimate of the uncertainty of model parameters. Similarly, inversions tend to be deterministic rather than statistical with probability density functions describing the uncertainty in the model parameters. Integrating Bayesian techniques into inversions, computing solution adjoints, and applying data assimilation techniques to geophysical datasets is difficult due to the sparse, clustered, and incomplete spatial and temporal coverage of the multi-dimensional data. A particular problem for Bayesian techniques in geodynamics, as in many inverse problems for partial differential equations, is the extremely large dimensionality of the parameter space that can be in the range of thousands to millions; random or uniform sampling techniques are inefficient in such cases, and other methods like Markov chain Monte Carlo (MCMC) may also reach the limits of their ability to characterize uncertainty with a reasonable number of model solutions. New approaches are necessary to move geodynamic models forward with statistical inversions and data assimilation of the rapidly growing solid-Earth datasets.

5.4 Performance Scalability. Improve parallel scalability of state-of-the-art codes by identifying and removing bottlenecks, e.g., optimizing solvers, meshing, and input/output. Approaching realistic conditions and parameters in geodynamics problems is linked to expanding the range of spatial and temporal scales that a model simulates. Likewise, deciphering complex interactions among geophysical and geochemical processes necessarily entails many realizations of the same model to explore the parameter space and test model assumptions. Improving performance scalability would facilitate users moving up from one level of computing to the next. Petascale computing has the potential to transform geodynamic modeling through capacity computing and by permitting investigation of a much broader range of scales.

5.5 Mapping and parameterization of geophysical fields. Incorporating more direct coupling among geodynamic models with strong feedback and moving towards whole Earth models require mapping physical and geochemical properties from one model to another across different scales, discretization schemes, and parameterizations. For example, elastic properties in a seismic wave propagation simulation (using spectral finite-elements with discretization size proportional to shear wave speed) may be coupled to solid composition at a temperature and pressure prescribed from a mantle convection simulation (using lower-order finite-elements with discretization size proportional to the spatial extent of localized mantle flow features). Implementing this complex information exchange for various combinations of coupled geodynamic simulations running on hundreds to thousands of cores presents a formidable challenge due to differences in basis functions, physical parameters, and the large volume of data.

6. Annual Goals and Milestones (Sept. 1, 2008 – Aug. 31, 2009)

I. Common Infrastructure
  a. Maintain LAN, servers, desktops, notebook computers
  b. Maintain Plone Site (http://geodynamics.org)
  c. Maintain repository (SVN in general, Mercurial for magma)
  d. Maintain and expand regression testing (*Buildbot*; *CIG-Regresstor*)
  e. Expand *Sieve* software suite
  f. Expand benchmark code (*Cigma*)
  g. Science Gateway for benchmarking

II. Core Computational Software
  a. Finish development of Seismology Science Gateway
  b. Mantle convection support
    i. Minor enhancements and bug fixes (*CitcomS*)
    ii. Maintain repository (*HC)*
    iii. Migrate into repository (*ConMan*)
  c. Long-term tectonics code development (*Gale*)
  d. Long-term tectonics code migration into SVN (*SNAC* donation)
  e. Short-term tectonics code development (*PyLith)*
  f. Geodynamo code bug fixes (*MAG*)
  g. Magma development (*MADDs*)
  h. AMR: Develop suite of *deal.II* programs for several geodynamics problems.
  i. AMR: Mantle Convection Code (*RHEA*)
  j. Coupling of *CitcomS* to *SPECFEM*

III. Organizing Community Participation
  a. Annual meeting of EC, November 2008, Columbia University, NYC.
  b. CIG Workshop on Mathematical and Computational Issues in the Solid Earth Geosciences, to be held Sept. 15-17, 2008, in Santa Fe, NM.
  c. Annual business meeting at Fall AGU, San Francisco, CA.
  d. CIG-All meeting, Spring 2009 (under discussion).
  e. Annual meeting of SSC, May 2009, Pasadena CA.
  f. Joint Magma Dynamics and Long-Term Tectonics Workshop, Summer 2009.
  g. Numerical modeling of crustal deformation and earthquake faulting workshop (Co-sponsorship with SCEC, etc., Summer 2009).
  h. Geodynamo Workshop, Summer 2009.

IV. User Training
  a. User manuals
    i. *Gale* manual
    ii. *PyLith* manual
    iii. *SNAC* manual
    iv. *CitcomS* (compressible) manual
    v. *MAG* Manual

    - vi. *MADDs* Manual
    - vii. Seismology Science Gateway Manual
    - viii. *ConMan* manual
    - ix. *Cigma* manual
- b. Training sessions
    - i. Tutorials for all CIG-codes at the CIG-All meeting, Spring 2009.
    - ii. *PyLith* training session at the CIG/SCEC workshop, Summer 2009.
    - iii. *Gale* and *stgMADDs* training session at the Magma Dynamics and Long-Term Tectonics Workshop, Summer 2009.

7. Allocation of Resources by Goal

Sept. 1, 2008 – Aug. 31, 2009

Software Engineer #1 (WL, Numerical Analyst/Modeler) [Total 1 FTE]
- I.c Repository: 0.1 FTE
- II.c *Gale* enhancement, bug fixes and BMs: 0.5 FTE
- II.e *PyLith* Development 0.25 FTE
- II.d *SNAC* Build: 0.05 FTE
- IV.a.i *Gale* manual: 0.05 FTE
- IV.b.i *Gale* training session: 0.05 FTE

Software Engineer #2 (LS, Software Integration) [Total 1 FTE]
- I.d Build system and installation: 0.25 FTE
- I.g Science Gateway for BM 0.15 FTE
- II.a Seismology Science Gateway: 0.5 FTE
- II.j *CitcomS-SPECFEM* coupling: 0.05 FTE
- III.g CFEM workshop: 0.05 FTE

Software Engineer #3 (LA, Software Integration) [Total 1 FTE]
- I.c SVN repository: 0.1 FTE
- I.f Benchmarking (Cigma): 0.5 FTE
- I.g Science Gateway for benchmarking: 0.2
- II.b.ii Support for HC: 0.05 FTE
- IV.a.ix *Cigma* manual: 0.15 FTE

Software Engineer #4[1] (ET, Numerical Analysis) [Total 1 FTE]
- II.b.i Maintain *CitcomS*: 0.45 FTE
- II.j *CitcomS-SPECFEM* coupling: 0.05 FTE
- II.i AMR: 0.45 FTE
- IV.a.iv Mantle convection manual: 0.05 FTE

Software Engineer and System Admin (WM, software integration) [Total 1 FTE]
- I.a. System administration: 0.35 FTE
- II.f Geodynamo (MAG) bug fixes: 0.1 FTE
- II.a Science gateway infrastructure: 0.2 FTE
- II.b.iii *ConMan* migration: 0.3 FTE
- IV.a.v *MAG* manual updates: 0.05 FTE

---

[1] Costs shared between CIG Cooperative Agreement (NSF Grant No. EAR-0426271) at 40% and NSF PetaApps Award (NSF Grant No. OCI-0748898) at 60%.

University of Chicago Subcontract [Total 0.5 FTE]
    I.e Sieve development: 0.05 FTE
    II.e *PyLith*: 0.3 FTE
    II.h AMR Unstructured meshes: 0.1 FTE
    III.g CFEM Wkshp: 0.05 FTE

New[2] subcontract [Total: 0.5 FTE]
    II.e *PyLith* Development: 0.5 FTE

VPAC Subcontract [Total 0.75 FTE]
    II.c *Gale* maintenance: 0.4 FTE
    II.g Magma Dynamics (*MADDs*): 0.35 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

Administrative Assistant [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 [$30K]
    I, II, III, IV

Travel [$59K]
    I, II, III, IV

Participant Costs [$140K]
    III, IV

---

[2] CIG is currently searching for a subcontractor.

8. Membership

8.1 CIG Members and Member Representatives:

Argonne National Laboratory (MSC), Matthew Knepley
Arizona University, Noah Fay
Brown University, E.M. (Marc) Parmentier
California Institute of Technology, Jean-Paul Ampuero
Colorado School of Mines, Paul Sava
Colorado State University, Dennis Harry
Columbia University, Marc Spiegelman
Cornell University, Jason Phipps Morgan
Georgia Institute of Technology, Wenyue Xu
Harvard University, Jeremy Bloxham
Johns Hopkins University, Peter Olson
Lawrence Livermore National Laboratory, Arthur Rodgers
Los Alamos National Laboratory (ES), Carl Gable
Massachusetts Institute of Technology, Bradford Hager
Oregon State University, Gary D. Egbert
Pennsylvania State University, Kevin Furlong
Princeton University, *TBD*
Purdue University, *TBD*
Rensselaer Polytechnic Institute, Charles Williams
State University of New York at Buffalo, Abani Patra
State University of New York at Stony Brook, Lianxing Wen
U.S. Geological Survey (Menlo Park), Brad Aagaard
University of California, Berkeley, Mark Richards
University of California, Davis, Louise Kellogg
University of California, Los Angeles, Jonathan Aurnou
University of California, San Diego, Yuri Fialko
University of Colorado (Boulder), Shijie Zhong
University of Hawaii, Garrett Ito
University of Maine, Peter Koons
University of Maryland, Laurent Montési
University of Michigan, Todd Ehlers
University of Minnesota, David Yuen
University of Missouri-Columbia, Mian Liu
University of Nevada, Reno, John Louie
University of Oregon, Douglas Toomey
University of Southern California, Thorsten Becker
University of Texas at Austin, Luc Lavier
University of Washington, *TBD*
Virginia Polytechnic Institute and State University, Scott King
Washington University, Michael Wysession
Woods Hole Oceanographic Institution, Jeff McGuire

8.2 CIG Foreign Affiliates and Representatives:

Australian National University, Jean Braun
GNS Science, Susan Ellis
Monash University, Louis Moresi
Munich University LMU, Hans-Peter Bunge, Heiner Igel
University College London, Carolina Lithgow-Bertelloni
Geological Survey of Norway (NGU), Susanne Buiter
University of Science and Technology of China, Sidao Ni
University of Sydney, Dietmar Muller
Victorian Partnership for Advanced Computing, Bill Appelbe

8.3 Strategy for Keeping Members Informed

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

1. *Through e-mail.* CIG maintains several list servers through the CIG web site 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 web site is distributed by e-mail on a regular basis.

2. *Through the* http://geodynamics.org *web site.* A calendar of upcoming CIG events is posted and continuously revised. Nearly all CIG documents, including proposals submitted to CIG, the annual revision of the CIG Strategic Plan, Bylaws, etc., are posted on this site. The web site is the principal means for standard software downloads, sharing of community benchmarks, specifications of standards, and distribution of user and training manuals.

3. *The annual CIG Business Meeting.* This meeting is open to all and is a forum for open discussions of the working of CIG, including past and upcoming activities and the Strategic Plan. This meeting will again be held in conjunction with the AGU Fall meeting in December, which has successfully garnered member participation in previous years.

4. *CIG sponsored and co-sponsored workshops and training sessions.* The current status of CIG is 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 have adopted 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 bylaws available on our web site (http://geodynamics.org).

9.1. Institutional Membership and Executive and Science Steering 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, of which five are voting members: the Chairman, the vice Chairman, and three 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 policy. Therefore, the CIG Bylaws were amended after approval of the community at the previous Business Meeting to increase the voting members to five from the original four.

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, provide initial assessment of proposals submitted to CIG, and revise the Five-Year Strategic Plan. Recommendations from the SSC are passed on to 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: (1) 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, (2) acting as PI 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, (3) ensuring that funds are properly allocated to various CIG activities, and (4) 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 final Business meeting will be 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 interrelated 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.

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 has set up a formal process for eliciting new ideas 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 led to scientific advances that could not have been made without CIG? In order to answer these and related questions, CIG has begun to develop metrics on how our community is using the

CIG codes. Currently, CIG collects statistical data on downloads of CIG codes from the repository and additional metrics such as lists of papers and abstracts resulting from the use of CIG software. One-page abstracts submitted by members of our community highlighting important scientific results are also being considered as a metric.

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 be 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. Collaborative proposals, such as the PetaApps proposals and the OCE MG&G proposal, are examples of how we are now moving in this direction.

10. Annual CIG Allocations and Expenditures

	FY5 2008-09
Senior (Total)	$56,220
Engineers (Total)	$406,364
Support (Total)	$115,960
Fringe	$144,636
Overhead	$428,846
Total	$1,152,027
CIG Staff Travel (15 trips)	
CIG Travel Total	$30,586
CIG Visitors (10 trips)	
Visit Total	$28,610
Total Travel	$59,196
Total Material and Supplies	$30,362
Subcontract Total	$345,848
Total Participant Cost	$140,000

Total Budget Amount	($1,727,433)
Carry Forward	$227,433
Budget Request Amount	$1,500,000
Balance	*$0*

11. Additional funding

11.1 Petascale Code Development Grant

On August 29, 2007, NSF's Office of Cyberstructure (OCI) through the PetaApps competition funded the “Proposal for High-Resolution Mantle Convection Simulation on Petascale Computers.” This new project, led by PI Omar Ghattas (U. Texas at Austin), and co-PIs George Biros (Penn), Michael Gurnis (Caltech), and Shijie Zhong (U. Colorado, Boulder), will capitalize on upcoming petascale computing systems to carry out the first high resolution mantle convection simulations that can resolve thermal boundary layers and faulted plate boundaries, leading to breakthroughs in the understanding of the dynamics of the solid Earth. To accomplish this, *CitcomS* will be improved and scaled up to the petascale. Then parts of *CitcomS* will be incorporated with new parallel algorithms for adaptive mesh refinement and inverse solution that can scale to hundreds of thousands of processors. CIG will freely distribute the resulting software.

The project is for three years, with CIG’s portion funded at $205,824. CIG has already received $88,824, of which $68,677 is allowed for the current year. In Sept/Oct 2008, CIG expects to receive $117,000 for the remaining two years. This amount will be kept in a separate account from CIG mainline funding.