Computational Infrastructure for Geodynamics (CIG)

NSF Award Number EAR-0426271

Five-Year Strategic Plan:
Sept. 1, 2007 to Aug. 31, 2012

Table of Contents

0. Community Oversight 3
1. An Overview of CIG 4
2. Our Long-Range Goals 6
3. CIG Accomplishments 8
    3.1 Development Infrastructure 8
    3.2 Long-Term Tectonics 9
    3.3 Mantle Convection 9
    3.4 Computational Seismology 10
    3.5 Short Time-Scale Tectonics 10
    3.6 Geodynamo 12
    3.7 Magma Dynamics 13
    3.8 Computational Science 13
4. Details of Our Short-, Intermediate-, and Long-Term Goals 14
    4.1 Common Infrastructure 14
    4.2 Long Time-Scale Tectonics 14
    4.3 Mantle Convection 15
    4.4 Computational Seismology 16
    4.5 Short Time-Scale Tectonics 17
    4.6 Geodynamo 18
    4.7 Magma Dynamics 19
    4.8 Computational Science 20
    4.9 Organizing Community Participation 22
    4.10 User Training 23
5. Achieving Our Long-Term Goals and the Future of CIG 27
    5.1 Adaptive Mesh Refinement 28
    5.2 Opportunities in Petascale Computing 28
    5.3 The Future of CIG 30
6. Annual Goals and Milestones 32
7. Allocation of Resources by Goal 34
8. Membership 36
    8.1 CIG Members and Member Representatives 36
    8.2 CIG Foreign Affiliates and Representatives 37
    8.3 Strategy for Keeping Members Informed 37
9. Five-Year Management Plan 38
    9.1 Institutional Membership and Executive and Science Steering Committees 38
    9.2 Administration 38
    9.3 Formulating CIG Priorities and Management of its Resources 39
10. Annual CIG Allocations and Expenditures 41
11. Additional Funding 41

Community Oversight

Executive Committee (EC)

Mark A. Richards, (Chairman, Oct., 2007), University of California, Berkeley
Marc Spiegelman (Vice Chairman, Sept., 2009), Columbia University
Bill Appelbe (At-Large, Oct., 2007), Victorian Partnership for Advanced Computing
Bradford H. Hager (At-Large, Oct., 2008), Massachusetts Institute of Technology
Peter Olson (*ex-officio*, June 2007), John Hopkins University
Michael Gurnis (*ex-officio)* California Institute of Technology
Michael Aivazis (*ex-officio)* California Institute of Technology

Science Steering Committee (SSC)

Peter Olson (Chairman, June 2007), John Hopkins University
Brad Aagaard (June 2008), USGS, Menlo Park
Wolfgang Bangerth (June 2008), Texas A&M University
Omar Ghattas (June 2007), The University of Texas at Austin
Louise Kellogg (Sept., 2009), University of California, Davis
Laurent Montesi, (Sept., 2009), Woods Hole Oceanographic Institution
Jeroen Tromp (June 2007), California Institute of Technology
Shijie Zhong (June 2008), University of Colorado at Boulder

1. An Overview of CIG

The Computational Infrastructure for Geodynamics (CIG) develops, supports, and disseminates software for the geoscience community from model developers to end-users. The software is being developed for problems ranging widely 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 now under development consists of:

- a central system facilitating software development, including a repository, a bug tracking system, and automatic regression testing;
- a coordinated effort to develop reusable, well documented and open-source geodynamics software;
- the basic building blocks – an infrastructure layer – of software by which state-of-the-art modeling codes can be 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 starting to formulate a plan for achieving them both within the current CIG as well as any future initiatives or reincarnations of CIG. The details of our plans are still under development and will likely evolve over the next twelve months. These emerging plans are described in Section 5.

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 21-22, 2007, was held specifically to develop this SP. During the meeting, we reviewed our accomplishments, had extensive discussions on our goals while attempting to identify common themes in different disciplines, and started to consider the future of CIG beyond its currently funded lifetime. The CIG Executive Committee approved this Strategic Plan on July 12, 2007.

2. Our Long-Range Goals

The long-term goal of CIG is to create a set of computational tools and data structures that can be commonly applied within the geodynamics community. These tools and data structures will promote more interaction between different 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.

On-demand/On-Request Seismology. With the availability of wave propagation codes, the simulation of seismograms with realistic whole Earth models has become routine. A next step is the development of web interfaces (a Science Gateway) that make these simulations easy and routine for the seismological community. Part of such an effort will be to incorporate seismic velocity and attenuation structures of the earth derived using various global tomography methods and represent an important link between geodynamic process models and observational seismology. The development of common data structures will be needed for representing these earth structure models. This may be regarded as a first step toward more general earth structure model frameworks discussed below.

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.

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 is recognized to be fundamental – for example, the thermal structure of the core will be controlled by the convecting mantle and the dynamics of the lithosphere through the development of plate boundaries will control mantle convection – these couplings have been treated in only limited ways. As in many highly nonlinear problems, the couplings involve strong feedbacks: the dynamics of the lithosphere not only influences convection in the underlying mantle, but mantle convection exerts forces responsible for the dynamics of the lithosphere.

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

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

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, 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 Development 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. For the CIG software repository, we use the open source package Subversion (SVN). The entire contents of the repository are navigable from our web site. Users can either directly check out the latest development version of a code, or they can download a "tar file." In addition, CIG provides a bug-tracking database (Roundup) to allow developers and external participants to register and comment on bugs and requests for new functionality in CIG software that can then be worked on by the developers of a program.

A key problem that faces any dynamic software repository is ensuring that "nothing breaks" despite frequent dynamic changes needed to meet the evolving scientific goals of the community. CIG uses state-of-the-art software engineering technology – agile computing to minimize the risks of software development for continuously evolving requirements. In particular, the repository uses unit and regression testing. Building and testing occurs either nightly or automatically in response to a software commit to the SVN repository using *CIG-Regresstor*, a collection of Python codes written by CIG engineer Luis Armendariz. This software uses Build-bot, 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 allows users of the repository to have confidence in the robustness of the software. Strand extended Build-bot 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 Mathew 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 now released *PyLith 1.0* (see below).

Work continues on a suite of Python software for benchmark intercomparisons, *Cigma* (CIG Model Analyzer). *Cigma* (version 0.9 release in July) 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. The software was initially being developed for short time-scale tectonic problems.

3.2 Long-Term Tectonics. Our most significant achievement in long-term tectonics has been the initial 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.

In March, 2007, CIG released *Gale* 1.2.0. This included a manual, source code, and binary executables for popular architectures. The current release (version 1.2.1) can simulate the geophysical phenomena originally contemplated for *Gale* (shortening, extension, and subduction all coupled to thermal evolution), as well as others not initially contemplated, such as fault initiation in pre-weakened material and coronae formation on Venus. *Gale* can incorporate a wide variety of boundary conditions in 2D and 3D in serial or parallel, and scales to hundreds of processors. A training session was held just prior to the EarthScope National Meeting in March of 2007. At this workshop, 31 attendees (mostly graduate students and postdocs) received a hands-on tutorial on using *Gale*. *Gale* was also used during the June "Workshop on Community Finite Element Models for Fault Systems and Tectonic Studies" on the campus of the Colorado School of Mines that CIG partially supported. Further development and testing of the 2D and 3D versions of the *Gale* code will be needed, in particular for running the complete SOPALE (2D ALE code) integration benchmark.

3.3 Mantle Convection. Eh Tan joined the CIG as staff scientist in July 2006 and helped with a number of mantle convection related projects. Working with Shijie Zhong (Colorado) on the geoid calculations and Allen McNamara (Arizona State) on tracers, Eh Tan successfully added capabilities of the geoid calculations and tracer modeling to CIG's finite element code for spherical mantle convection, *CitcomS*. The version of *CitcomS* released in the last year can be used to compute the geoid and also thermochemical convection. Eh Tan also worked with Carolina Lithgow-Bertelloni and

Lars Stixrude (both at Michigan) to explore the possibility of including their thermodynamics code in CIG's mantle convection codes.

The working group on compressible mantle convection continued to make progress in formulating benchmark problems and implementing numerical methods for compressible mantle convection problems. The group [Eh Tan, Scott King (Virginia Polytech), Peter van Keken (Michigan), Louis Moresi (Monash), and Shijie Zhong] presented a poster at the 2006 Fall AGU meeting. Eh Tan finished coding to add a compressible convection feature to *CitcomS*, and he is currently performing benchmark calculations for the Stokes flow problem by comparing with propagator matrix solutions provided by Wei Leng, a graduate student from University of Colorado.

The working group for analytic solutions code [Thorsten Becker (USC), Bernhard Steinberger (Norway), Rick O'Connell (Harvard), and Carolina Lithgow-Bertelloni] continued to work on the *HC* code that is currently in the CIG's software repository. Becker and Steinberger performed benchmark studies by comparing solutions from their *HC* code with those from *CitcomS* and obtained satisfactory agreement.

With the help from his graduate student, Nan Zhang, and Eh Tan, Shijie Zhong generated a benchmark web page for *CitcomS*. The benchmark page included detail information for input and output from about 20 cases with different Rayleigh numbers and activation energy (from isoviscous models to stagnant-lid convection with viscosity variations of up to 10^7).

3.4. Computational Seismology. In the fall of 2006, CIG co-sponsored a computational Seismology “Imaging Workshop” organized by Levander/Aster/Pavlis/Rondenay where a rudimentary version of the CIG on-demand web portal was demonstrated.

In the past year, CIG made a sub-award to the University of Colorado (M. Ritzwoller, PI) which led to the development of a normal-mode summation package (*Mineos*) that is now available via the CIG seismology web page. A *Mineos* Users Manual was also developed by Ritzwoller’s group and CIG technical writer Sue Kientz. During 2006, we also
saw the release of new versions of the spectral-element packages SPECFEM3D_BASIN and SPECFEM3D_GLOBE developed by J. Tromp (Caltech), but made available through CIG with a CIG-developed users manuals. In addition to these CIG seismology software packages, the CIG seismology pages now also provide access to the SPICE (www.spice-rtn.org) software repository. Finally, a 3D ray-tracing package developed by Princeton University is also available for download.

3.5. Short Time-Scale Tectonics. Short-term crustal dynamics focuses on simulating crustal deformation associated with the accumulation and release of strain over the earthquake cycle. The spatial scales range from meters to thousands of kilometers and the temporal scales range from tens of milliseconds to hundreds of thousands of years. At the larger length scales and longer time scales, models blend into those of the long-term crustal dynamics working group. An important continuing goal is the development of software for the simulation of multiple earthquake cycles with sufficient resolution to capture the buildup of strain in the crust, strain release in propagating ruptures that radiate seismic waves, and 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 six 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, have coordinated their development with 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. In each case the releases included binary executables for several common platforms and source code from the SVN repository.

During the June 2006 workshop, the short-term crustal dynamics working group set forth four priorities for the year: (1) implement additional viscoelastic material models and traction boundary conditions in *PyLith* 0.8, (2) development of *PyLith* 1.0 with increased modularity and functionality over *PyLith* 0.8, (3) improved support for computation of 3D Green's functions using *PyLith*, and (4) infrastructure for archiving and comparing output associated with the suite of working group benchmarks.

Prior to the release of *PyLith* 1.0, CIG released two updates to *PyLith* 0.8 with additional viscoelastic material models, traction boundary conditions, efficiency improvements, and bug fixes. Version 0.8.1 (October 2006) featured primarily efficiency improvements and bug fixes. Version 0.8.2 (May 2007) included addition of several commonly used viscoelastic models and traction boundary conditions. The *PyLith* user manual was updated to document the additional functionality and provide more detailed examples and installation instructions.

A significant shift occurred in 2007 in the development of *PyLith*. The focus moved from transforming old, entangled legacy simulation codes to a new modular implementation that leveraged *Pyre*, *Sieve*, and *PETSc*. This new implementation, *PyLith* 1.0 (released in June 2007), merges the dynamic earthquake rupture capabilities present in EqSim with the quasi-static crustal deformation capabilities present in *PyLith* 0.8. This new version can solve 1D, 2D, and 3D quasi-static and dynamic elastic and viscoelastic problems with prescribed slip on faults or prescribed displacements on the boundaries of the modeling domain with parallel processing afforded through the use of *Sieve* and *PETSc*.

Aagaard, Williams, and Knepley employed unit and regression testing throughout the entire development cycle, so that the functionality of each module of the code received thorough testing continuously from initial implementation to release. These 500+ 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. The modular implementation of *PyLith* 1.0 with *Pyre* allows users to extend the code by adding their own modules (e.g., bulk and fault constitutive models, importers for additional mesh generators, exporters to additional visualization packages, and interfaces with physical property models) by writing simple, small modules. *PyLith* 0.8.0 and 1.0 were used extensively in the 2006 and 2007 "Workshops on Community Finite Element Models for Fault Systems and Tectonic Studies" held on the campus of the Colorado School of Mines 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* 1.0 also produces the comparable elastic solution. 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.

3.6 Geodynamo. CIG software engineer Wei Mi worked with Peter Olson (John 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 are building a suite of example dynamo cases to be part of a tutorial for *MAG* beginners on the CIG web site. 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.

3.7 Magma Dynamics. The magma dynamics community gathered at an initial workshop supported by CIG in New York City on August 18-19, 2006. The magma dynamics community expressed the need for generally available, documented, open-source software to explore how magma dynamics interacts with mantle convection and/or long-term tectonics and to explore novel description of magma dynamics, addressing in particular the transition from porous flow to diking.

A working group was formed to respond to these needs and the decision was made to develop a Magma Dynamics Demonstration Suite (*Madds*) to prove the feasibility of including magma dynamics in more general software. The basic formulation for magma dynamics, as well as a concrete set of benchmark problems culminating in 2D and 3D mid-ocean ridge models, is laid out in a set of notes developed as an outcome of the CIG Magma Dynamics workshop. In particular, this document lays out a systematic sequence of benchmark exercises for *Madds*: tests for accurate dynamic pressures in existing solid deformation codes; flow within a column with forced melting; solitary waves; shear bands developed through the coupling with shear flow in 2D; 2D ridge models driven by either buoyancy or shear flow in 2D, including forced adiabatic melting; imposed velocity field; self-consistent velocity fields; coupling with temperature solvers; and 3D ridge models with boundary conditions from global mantle flow model.

CIG has subcontracted 0.75 FTE to VPAC to facilitate this community effort, in addition to *Gale* code maintenance. The initial task of estimating the reliability of pressure solvers in existing software is due in July 2007. This task is well underway, with three different techniques for improving the reliability of pressure gradient obtained: quadratic elements, Recovery by Equilibrium in Patches (REP), and inner-element pressure nodes. Marc Spiegelman (Columbia) is providing guidance to VPAC on the implementation of *Madds*. He also shared a spectral 2D and 3D mid-ocean ridge flow code (SPECRIDGE) and is developing with Richard Katz a PETSc-based solver for *Madds*, starting with treatment of solitary waves. Other working group members are implementing *Madds* in other general software platforms (e.g., Comsol Multiphysics). Laurent Montesi (WHOI) shared a pressure-velocity corner flow benchmark using power law rheology (NNcorner). These *Madds*-related examples and others that will be developed are gathered in a Mercurial software repository at CIG.

3.8 Computational Science. CIG supported a workshop on "Challenges and Opportunities at the Interfaces of Scientific Computing and Computational Geodynamics" in Austin, TX, October 16-18, 2006. At this workshop more than 20 speakers from applied geosciences, computer science, and mathematics presented their view of large-scale computer simulations and related problems, in particular pre- and post-processing the large data sets involved. The workshop was structured to allow computational scientists to respond on the second day to the challenges laid out the day before by application scientists. Its goal was to increase awareness of challenges in computational geodynamics on the one hand, present existing solution approaches developed in other areas of the sciences and engineering, and finally to stimulate interdisciplinary collaboration.

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

The Science Steering Committee has identified our short-, intermediate-, and long-term goals. Our original short-term goals have now all been completed, and we have made substantial progress on our 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; our strategy for achieving our long-term goals is presented in Section 5. All of our goals are summarized in Table 4.1.

4.1 Common Infrastructure. The routine comparison of model results with existing Web-accessible benchmarks is essential for increasing the overall quality of our science, especially as it moves into the realm of complex, multi-physics and multi-scale simulations. CIG engineer Luis Armendariz has now completed a working version of *Cigma* (formally called *BM.py*) for the short time-scale tectonics community and demonstrated it at the recent CFEM workshop in Golden, CO. Our intermediate goal is to modify this code for the mantle convection community, and then the long-term tectonics community. A standard procedure by which geophysics codes can be compared requires the establishment of benchmarks that are sensible mathematically (verification) and geophysically (validation). A common and permanent repository is now being created for these benchmark results and will interface with *Cigma*. CIG will collaborate with individual communities in order to develop this procedure: CIG staff will adapt and refine this procedure as it gains experience. We expect that the *Cigma* software will be used by individual investigators studying their own benchmarks, while also being incorporated within our automatic regression testing. We are also exploring the possibility of making *Cigma* available through our Science Gateway.

4.2 Long Time-Scale Tectonics. Finite amplitude deformation of the crust and lithosphere occurring across a range of time scales results in the geologic structure of the Earth and other planets. This deformation occurs both at the fast rates of earthquakes and dike intrusions and at slower rates of mantle convection and glacial loading. Elastic, viscous and brittle plastic strains all contribute to large-scale deformation. These time scales and deformation 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.

With the release of *Gale*, one of our primary goals is to 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. We are attempting to recruit additional members to the working group; we have had discussions with Mousumi Roy (New Mexico) and Dennis Harry (Colorado State).

An additional intermediate goal is the expected donation of SNAC (from Eun-seo Choi, a graduate student at Caltech) in the fall of 2007. SNAC, developed with the same common components used in *Gale*, uses FLAC (Fast Lagrangian Analysis of Continua) that is widely used for 2D studies of plastic failure and shear zone development. SNAC will come with a users manual, but CIG will probably need to clean the manual up and rewrite the build procedure of this code so that it's consistent with that of *Gale*.

There is a considerable need for benchmarking of codes. 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. CIG is closely cooperating with this group on comparing benchmarks from *Gale* with those from other codes. CIG will support U.S. participation in these activities.

For our longer-range goals, incorporation of adaptive mesh refinement into our codes is essential. AMR provides the efficiency needed for solving large-scale problems with high spatial resolution in regions of strain localization and fluid/rock interactions. There are a number of freely available AMR packages, and the immediate task is to identify which package is suitable. A workshop on AMR is planned for Fall 2007 in Boulder, CO.

4.3 Mantle Convection. Among several long-term goals identified by the mantle convection community, the remaining goals include development of 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 is near its final testing and benchmark stage. The solvers for both the momentum equation and the energy equation have been completed. CIG staff member Eh Tan is currently working with Wei Leng (Colorado) on benchmarks of the Stokes flow for compressible problems. We expect that testing and benchmarking will be done and that the code will be released sometime in the summer of 2007. We may make an additional effort to incorporate more realistic thermodynamic formulations from the Stixrude and Lithgow-Bertelloni group into *CitcomS* with compressible convection. Full calculations of thermodynamic relations at different grid points may be computationally costly, and we will explore the possibility of using pre-calculated tables.

After the completion of a compressible convection version of *CitcomS*, there will be no significant new code developments for Citcom codes at CIG. CIG will focus its efforts on developing new convection codes to address the other important long-term goal: a convection code with adaptive mesh refinement, flexible element types, and portable and modular solvers. It should be pointed out that this new code would be significantly different from existing codes at CIG. In many ways, this will be our next generation mantle convection-modeling tool. Many parts of the codes include meshing, element types, and solvers that may be shared with software for problems 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. To initiate this effort, we will have a workshop on adaptive mesh refinement and its scalability in October 2007 in Boulder, Colorado.

Another important task to be considered is a workshop in the summer of 2008 for mantle and lithospheric dynamics. This would represent a continuation of the 2005 Boulder workshop. We hope that this workshop will be held in the U.S. every two years, independent of CIG, similar to the European mantle and lithospheric dynamics workshop. This workshop series is important for training our graduate students and discussing important scientific and technical problems in our field.

Five additional projects, that may not take much of CIG's resources but would be useful for the community, include (1) examining the possibility of speeding up Citcom for production runs; (2) improving the *HC* code by adding compressibility and geoid modeling features; (3) adding a 2D convection code (e.g., Conman and Citcom) to CIG's software repository; (4) continuing to develop benchmark cases, including 2D convection and 3D high Rayleigh number cases with internal heating and time-dependence; and (5) continuing to develop visualization scripts (e.g., for DX, AVS, etc.) for producing animations and still images as part of the software packages.

4.4 Computational Seismology. A central goal for computational seismology continues to be the establishment of automated and on-demand simulations (e.g., seismic wave propagation and synthetic seismograms) through a seismology Science Gateway. A science gateway is a web site that launches a simulation on a remote machine using data gathered from web sites and databases and returns the results to the user. In order for a science gateway to operate effectively, there is the need to select between multiple simulation methods and codes with uniform specifications. Consequently, the short-term goal of this community is migrating their principal codes for simulating seismic waves into an accessible repository based upon uniform standards. Our goal is to have a working prototype ready for the October 9-11, 2007, CIG-SPICE-IRIS workshop in Jackson, NH.

The future directions of computational seismology will be major subjects of discussion at the CIG-SPICE-IRIS workshop in Jackson, NH, but we have been able to identify some of the long-term goals for this area:

> Automated/On-Demand Simulations: CIG will work to establish a Seismology Science Gateway, involving both automated and on-demand simulations. Automated simulations would provide near real-time 1D and/or 3D synthetics to accompany IRIS data for all events over a certain magnitude threshold using past and emerging events in the Harvard CMT catalog.

> Seismic Model Database: There is the need for a database of seismic models, including structural models of the crust and mantle together with databases of topography and bathymetry. Various resolutions are needed to match the capabilities of codes being developed under CIG. Mechanisms for the contribution of models must be established.
>
> Data Processing Tools: The SSC is considering whether CIG should investigate the feasibility of facilitating the development of data processing tools for field and laboratory use. These could include low-level routines for standard data manipulation (e.g., filtering, simple array analyses); higher-level functionality such as earthquake location, 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 productively within the CIG framework.

4.5 Short Time-Scale Tectonics. The primary focus of the short-term tectonics working group is the development of software for the simulation of crustal deformation associated with the accumulation and release of strain over the earthquake cycle. Working group members voiced strong support for continued development of *PyLith* at the June 2007 "Workshop on Community Finite Element Models for Fault Systems and Tectonic Studies." Community members requested that all features present in *PyLith* 0.8 be transferred into *PyLith* 1.0 and additional features be added.

In response to these requests, CIG plans additional releases of *PyLith* (1.1 and 1.2) to transfer all features present in *PyLith* 0.8 and EqSim into *PyLith* 1.0, followed by additional releases over the next two years associated with implementation of new features. *PyLith* 1.1, with an anticipated release in September 2007, will focus on including absorbing boundary conditions for better support of dynamic modeling with kinematic earthquake ruptures, reimplementation of the Power-law viscoelastic model present in *PyLith* 0.8 making use of the more modular approach used in *PyLith* 1.0, and support for output of stress, strain, and displacements fields over various subsets of the modeling domain. This release also includes traction and velocity boundary conditions. *PyLith* 1.2, with an anticipated release in November 2007, adds fault constitutive models for both dynamic and quasi-static modeling. With this release, *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.

Two other releases of *PyLith* are planned for the spring and summer of 2008. The first of these releases focuses on supporting automatic computation of 3D Green's functions. The second of these releases allows 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.

Features planned for future releases include use of PETSc nonlinear solvers and implementation of large deformations. 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.

4.6 Geodynamo. Dynamo codes represent a powerful new tool for the quantitative study of a broad range of geophysical processes, ranging from short time-scale phenomena such as magnetic variations, rotational variations, and flow in the core, to long-term phenomena such 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.

Following SSC member Peter Olson's consultations with members of the U.S. geodynamo and related communities (including paleomagnetism and geomagnetism) the SSC 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. MoSST includes capabilities for rotational interactions between the mantle and the outer and inner core.

The following additional objectives were defined in the course of polling community members. All of these constitute essentially short-term goals, and most have been accomplished already or are in the process of being accomplished: (1) standardization of existing dynamo codes to allow inter-comparison of results; (2) making benchmark comparisons of two codes; (3) development of software manuals and web tutorial examples; (4) development of common graphics tools; (5) development of one or more "user pacs," designed to translate between dynamo code output and standard data formats and definitions used in geomagnetism, paleomagnetism, and in the planetary and space sciences; and (6) formation of a CIG geodynamo working group from participants at the upcoming Les Houches Summer School in August 2007.

A part of the longer-term CIG strategy in developing the next generation of software in this area would to free dynamo codes from some of their current limitations. One limitation is that the dynamo codes now in use were not designed with a modular construction. Other disadvantages stem from their reliance on traditional spherical harmonic and Chebyscheff 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. However, spherical harmonics waste resources in the polar regions, Chebyscheff 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 Chebyscheff polynomials that allows for parallelization. Spherical harmonics will continue to be used in dynamo codes for satisfying magnetic boundary conditions and other specialized uses. But a local basis function formulation 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.

4.7 Magma Dynamics. The feasibility of including magma dynamics in these codes is demonstrated through the Magma Dynamics Demonstration Suite (*Madds*), a series of benchmarking exercises of progressive levels of sophistication based on the McKenzie [1984] theory of magma migration in a porous viscous medium. The sequence of benchmarks was defined at the August 2006 Magma Dynamics workshop and is separated into short-term, intermediate-term, and long-term activities as follows:

Short-term goals for magma dynamics essentially are the implementation and testing of the McKenzie formulation. Some of the specific tests will include flow in a column with forced melting in a steady state (1D equation). Once tested, we will then test the codes for the development of solitary waves with the time-dependent equation (in 1D and 2D) and shear bands which come about through the coupling with shear flow in 2D. Next, essentially with small modifications to existing software, we will attempt 2D ridge models with buoyancy and shear flow with forced adiabatic melting. Models with imposed velocity field will also be tested, especially the widely used Batchelor solution for corner flow. However, we would then hope to move on to self-consistent velocity fields that arise through the coupling with temperature solver and buoyancy. Finally in the short term, using existing software tests for 3D ridge models with velocity (and maybe stress) boundary conditions returned from global mantle flow models will be explored. We hope to achieve most of these short-term goals within the current subcontract to VPAC using the *StGermain* framework.

Our long-term goal is to couple magma dynamics with other software. Problems of interest include 2D subduction models with buoyancy and shear flow, models with imposed velocity field (including the Batchelor solution), and the self-consistent velocity field arising through coupling with a temperature solver.

For the longer-term, models used for magma dynamics will need to be tested with geochemistry and so more advanced software features will be needed. We will need an advection benchmark to validate the proper tracking of heterogeneity. Software for the distillation column with specific reactions will be needed. Models in 2D that can track trace elements in a ridge environment and code capable of modeling the U-series observations will be needed, as well as reactive flow (2D reaction/infiltration instability). In the longer term, trying to interface magma dynamics software with cracking and brittle failure within the upper crust will be essential so that we can connect the deeper dynamics and geochemistry with the processes that occur at shallower depths.

Over the next year, *Madds* will be tested at least with Underworld (an expansion of the *StGermain* framework), a new PETSc-based code, and the commercial software Comsol Multiphysics (a closed-source serial code), through a combination of CIG-sponsored efforts and the personal involvement by working group members. While only the short-term benchmarks described above are to be expected in the first year, *Madds* will go a long way towards the long-term goal of enabling multiphysics coupling between magma dynamics and solid Earth deformation such as mantle convection and lithosphere dynamics. Once *Madds* is implemented within the *StGermain* framework and validated, it will not be particularly difficult to couple all of these features with *Gale,* used by the long time-scale tectonics community.

A particularly important aspect of magma dynamics is the development of multiscale and localization phenomena such as reaction channels and dikes. The challenges associated with localization are best met using Adaptive Mesh Refinement (AMR). Therefore, it is expected that some working group members will participate in the AMR workshop of October 2007 and attempt to modify deal.II tutorial examples to handle some of the requirements of *Madds*. Treating multiscale phenomena remains an intermediate- to long-term goal of the Magma Dynamics working group.

Finally, long-term efforts should address the interfacing with geochemical database (e.g., EarthChem) and computational thermodynamics (e.g., MELTs). External research project such as the development of the GyPSM software (at Caltech) revealed that thermodynamics and geodynamics could be coupled but that difficulties remain especially when it comes to efficiently solving such problems. Such coupling would benefit not only the magma dynamics but also other efforts within CIG such as mantle convection and long-term tectonics, as well as providing important input for structural seismology. Therefore, we recommend that thermodynamics coupling be an important aspect of any general modeling platform produced by CIG in the long term.

Under the auspices of CIG, a proposal will be submitted to the NSF Marine Geology and Geophysics program to support the magma activities beyond completion of our short-term goals. This will likely be a collaborative proposal between Columbia (Marc Spiegelman) and CIG and submitted for the August 15, 2007, deadline.

4.8 Computational Science. CIG already 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* to deal with 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.

On the other hand, CIG has not made significant progress integrating adaptive mesh refinement (AMR) techniques into geodynamics codes, despite the fact that it is consistently mentioned as one of the important goals in the field and has been mentioned in every Strategic Plan so far. The reasons are varied and range from a lack of experience in the community to the fact that it is considered very hard 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 have yet to be implemented in software tool kits.

A more promising, if more radical strategy appears to be to completely rewrite those parts of the program that deal with discretized problems, and only keep the control structure. An important component will be 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 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.

To jump-start developing an AMR-enabled new generation of geodynamics codes, CIG plans to hold an AMR workshop and tutorial Oct 24-27, 2007, in Boulder, CO. The idea is to spend some lecture time going through the extensive tutorials of the *deal.II* library to explain how to implement AMR-enabled models. The rest of the time will be 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. Wolfgang Bangerth (Texas, A&M), the lead author of *deal.II*, will be available to help with implementation questions.

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. With this motivation, a collaboration has begun between SSC members Shijie Zhong and Omar Ghattas and postdocs at UT-Austin 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 3000 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. Currently, *Octor* supports only Cartesian geometries, so this will be the initial target of the new code. Future work will extend the capability 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.

Further down the road, we have started to think about a workshop devoted to scalability issues during late 2008. A scalability workshop would cover issues such as code and algorithm scalability, across all the CIG codes to 10x, 100x, 1000x, etc., processors. This will dovetail will the development of ever larger capacity computers.

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

- *Openness*: All CIG reviews, meeting minutes, and other documents are openly available to the community in a timely fashion, for review and public discussion, unless this conflicts with individual or institutional privacy rights (such as salary level or personal information). Openness minimizes the risk of actual or perceived bias or conflicts of interest within CIG.

- *Interaction*: All CIG committees and workshops should have an open and balanced representation of both the scientific and software communities. Interaction minimizes the risk of balkanizing the community CIG serves.

Most of the organization of CIG for the community is now in place; however, there is likely the need to recruit more members of the community to participate in working groups, especially early career investigators. Workshops and committee meetings will continue to be the mechanism by which we 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.

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

4.10 User Training. A key objective of CIG is the widespread adoption of CIG-developed software by the general Earth science community and, in particular, by geophysicists. Both significant training and comprehensive documentation will be required. Potential users will want to know how to use CIG software as well as gain an understanding of the underlying algorithms and implementation details. Based on current research directions, sophisticated users who require a deeper understanding of the workings of the software will likely have different backgrounds and interests. One class of user consists of traditional computational scientists 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 user consists of scientists who attempt to string together several components of CIG software with non-CIG data analysis tools. In addition, there will be a class of less-technically-demanding users who will want to use CIG codes in a standard manner (e.g., for pedagogical purposes or using standard codes on new input data sets).

Several mechanisms will be used to train CIG users: (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**
Computational Seismology	Demonstrate Seismology Science Gateway using SPECFEM on TeraGrid. Organize computational seismology workshop in New Hampshire with Spice and IRIS	Refine and deploy for general use Seismology Science Gateway. Highest priority for two packages: a 1D package for Cartesian reflectivity, and a finite difference 3D wave propagation code. Goal will be more specifically defined after SPICE meeting.	Continued refinement of Seismology Science Gateway with multiple codes. Coordination of model databases.
Computational Science	Organization of a workshop in Boulder, October 2007, to familiarize the community with adaptive mesh refinement techniques and to jump-start the development of AMR-enabled codes.	Incorporation of AMR into a new generation of codes. Further integration with existing libraries for parallel computations.	Work on integration of computer sciences and geophysics into single framework.
Short Time-Scale Tectonics	Complete merging of EqSim and PyLith 0.8 into PyLith 1.0.	Direct support for computing Green's functions in PyLith. Coupling short-term simulations to other simulations in order to accurately capture interactions among geodynamics phenomena.	Adaptive mesh refinement capabilities in PyLith or a code with similar functionality. Integrate tools for formal data assimilation.
Long Time-Scale Tectonics	Recruit members to a larger, more active Working Group in long-term tectonics. Incorporate frictional BC's into Gale. Hold Gale training session, Fall 2007	Receive donation of SNAC, port to CIG build procedure and refine Users Manual for general use. Start incorporation of AMR in code for long time-scale tectonics	Develop new code with adaptive mesh refinement.
Magma Migration	Implementation of Madds using StGermain, PETSc, and Comsol Multiphysics. Calibration of pressure solvers and solitary wave benchmarks.	Training of user base in CIG software with Madds. Implementation of Madds with AMR.	Integration of magma dynamics in global mantle convection. Coupling of thermodynamics modeling and Madds-enabled software.
Geodynamo	Benchmark serial geodynamo cod, MAG; development of user interfaces. Preliminary community building discussions will be held during SEDI in July 2008.	Community building workshop in 2008, to utilize and train users on the two serial codes, possibly partnering with other organizations.	New dynamo code using common components with Mantle Convection codes including common Earth structure framework and grid and adaptivity.
Mantle Convection	Release of compressible mantle convection CitcomS by September 2007. Add benchmark cases with high Ra, internal heating and time-dependent solutions by the end of 2007. Add compressibility and the geoid calculations to the HC code.	Develop convection codes with AMR capability.	Develop convection codes with AMR capability that share basic components with other geodynamics codes and that are scalable to 100s and 1000s of processors.

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

Table 4.4
Workshops Planned

WORKSHOPS	
Short-Term Tectonics	Co-sponsor meeting in 2008, likely shorter than previous meetings with an emphasis on user training
Computational Seismology	The CIG/SPICE/IRIS Seismology Workshop will be held October 9-11, 2007, in Jackson, NH
Computational Science	Workshop and hands-on tutorial sessions on using adaptive mesh refinement techniques for scientific computing will be held October 24-27, 2007, in Boulder, CO
Mantle Convection	Summer 2008 Mantle Convection Workshop (discussion items: science, current software development, and future plans for new code with adaptivity and AMR).
Geodynamo	Pre-meeting to begin community building in Spring 2008, as part of SEDI program, to be followed up by a full workshop later in 2008.
Long-Term Tectonics	Gale training session in Pasadena, Fall 2007

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

The Science Steering Committee has concluded that most of our short-term goals originally described in the CIG proposal to NSF have been achieved. We 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 that will likely influence any future initiative in computational geophysics. Given this backdrop, we are now seriously evaluating our long-term objectives and starting to formulate a plan for achieving them, both within the current CIG as well as any future initiatives or reincarnations of CIG. The details of our plans are still under development and will likely evolve (possibly substantially) over the next twelve months. In this section, we described 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 two high-level tasks: Adopting advanced numerical methods and coupling multiple methods or codes to solve multi-scale or multi-physics problems. The need for adaptive mesh refinement is an example of a need for an advanced numerical method. In both cases, this involves software and ideas developed at the forefront of computational science. In CIG's NSF proposal, the community articulated the view that our resources would always be limited and entirely new software would only be generated when needed. This remains true, and CIG must rely on lower level components developed in more science-neutral arenas, including some DoE supported projects, and higher-level frameworks to interlink software as much as possible.

The need to adopt these new computational science technologies was anticipated in the NSF proposal. The development of a software framework for geophysical modeling and its subsequent upkeep were viewed as central to CIG's strategy. In the proposal, we articulated a framework that would consist of software on two levels: (1) an infrastructure layer of software by which community and state-of-the-art modeling codes would be assembled, and (2) a superstructure layer of software by which multiple codes and data would be interlinked through extension of existing software frameworks. In both cases, a key element is the reuse of common components. For about half of the CIG effort, we have essentially achieved this level of software development and re-use. For example, *StGermain* is an example of an infrastructure framework, while *Pyre* is an example of a superstructure framework. In the first case, we developed *Gale* "upward" by using existing common components in *StGermain*. We are now developing *Madds* (e.g., magma dynamics) using many of the same common components found in *Gale*. In the case of the now released *PyLith 1.0*, we have built the code both upward by using *Sieve, FIAT,* and *PETSc,* and "downward" by interfacing the geodynamics specific portions of the code using *Pyre*. It's important to emphasize that *PyLith* uses the same finite element infrastructure to solve two previously separate geophysical problems with mostly common components: pre- and post-seismic loading (the quasi-static problem) and dynamic earthquake rupture and seismic wave propagation (the dynamic problem).

5.1 Adaptive Mesh Refinement. Now that CIG has achieved some level of common components in modern, well-engineered geodynamics software, the next step beyond our intermediate-term goals will be incorporating adaptive mesh refinement (AMR) into our software suite. The SSC has repeatedly come to realize that AMR should be a key feature of our software. We have identified this need in mantle convection (e.g., thin boundary layers around plumes), in short-term tectonics (e.g., rupture on different faults during the earthquake cycle), in long-term tectonics (e.g., the initiation and growth of new fault systems), and in magma dynamics (e.g., localization through the growth of reaction channels). AMR could provide the geodynamics research community with a unified set of tools that will permit the treatment of intimately coupled, multiscale problems that will lead eventually to a whole Earth evolution model. Unfortunately, AMR is not a solved problem in computational science and many important questions remain open. Given the fact that AMR remains at the forefront of research, how will CIG respond within the final two years of its current existence and then beyond?

We are developing a multi-prong effort in AMR, involving training of our community and the inclusion of existing methods into the CIG framework and codes. In order to move forward, a better understanding of the issues by both the scientists in our community and the CIG staff is needed. In an attempt to familiarize the community with real examples using available tools, we will support a workshop and tutorial in Boulder in October 2007. While part of the workshop will be spent in lectures explaining the structure, ideas, and use of *deal.II* (a finite element tool kit with AMR capabilities), the majority of the time will be spent in hands-on sessions in which we will work on extending some of the existing tutorial programs in directions of particular relevance to geodynamics. We will also have a set of talks by developers of alternative AMR methods. We hope that this will jump-start the development of a new generation of codes by members of the community, as well as clarify the difficulties inherent in coupling geodynamics applications with adaptivity. An important goal for the workshop will be to identify at least one geodynamics application primed for AMR where CIG would work directly in a “research mode” with our computational science partners during the last two years of the currently funded CIG.

In addition, we have identified three concrete ways in which we can merge existing AMR software with CIG software. First, for the Stokes problem in general (and mantle convection in particular), we will collaborate with the group at the University of Texas, as described in Section 4.8 above, to integrate an octree-based parallel mesh adaptivity system (Octor) with some of the components of the *CitcomS* finite element code. Second, we will attempt to use methods in the *deal.II* library within the *PyLith* code for the short-term tectonics community. Finally, we will continue discussions we are having with our collaborators at VPAC who are attempting to incorporate AMR methods in their *StGermain* framework. This could directly impact the functionality available within the *Gale* and *Madds* software suites.

5.2 Opportunities in Petascale Computing. Research in computationally intensive geodynamics spans a broad spectrum of styles and needs for computational resources. In many areas, cutting-edge research addresses how best to parameterize complicated multi-scale physical and chemical processes. At present, most advances in geophysics are done at a modest level of computing, with most researchers using a mid-sized cluster. The reason for this is inherent in the nature of geodynamics research. Progress in most areas of solid-Earth geophysics centers on deciphering the complex interactions between many geophysical and geochemical processes. Experience has shown that the most effective computational approach is to explore these phenomena at different spatial and temporal scales, using the widest possible range of parameter values. This style of research necessarily entails many realizations of the same model, to determine parameter sensitivity, test model assumptions, and compare with observations. It will likely remain at the heart of computational geodynamics for the foreseeable future. Petascale computing could change geodynamics through capacity computing – computers with enormous capacity would be an extraordinary tool that would allow us to explore parameter space at a level never imagined possible in the past.

In a few areas, the underlying equations are well understood and the algorithms sufficiently stable that cutting edge research involves access to the largest computing resources to make possible calculations at sufficient spatial and temporal resolution. Through a detailed survey made during late 2006, the SSC found that there is a consensus opinion in the geodynamics community that 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.

From both the perspective of capacity and capability, we recognize that petascale computing has the potential to transform geodynamics modeling. However, we also recognize that achieving this lofty goal is a far greater challenge than stating it. For our community, success at petascale computing entails a broader commitment than simply joining a hardware consortium. In spite of the fact that some areas in geodynamics research are prime candidates for first-stage application of petascale computing and that nearly all of the disciplines represented in CIG could potentially benefit by computing at this scale, to date only a limited number of applications in our field are now ready, or even suitable, for this level of computing. To help the process forward, CIG is now involved with the development of two proposals that will be submitted to the NSF PetaApps program, administered by the Office of Cyberinfrastructure (Accelerating Discovery in Science and Engineering through Petascale Simulations and Analysis). The first will be a proposal led by Jeroen Tromp (Caltech) to enhance the SPECFEM code for seismic wave propagation so that global calculations can be made at 1 Hz. CIG will provide software engineering support and a strong linkage to the community. The second will be a proposal led by Omar Ghattas (U. Texas) to adapt an octree-based AMR method to global mantle convection. This will allow global time-dependent simulations of mantle convection with resolutions of up to 1 km where needed.

5.3 The Future of CIG. The Science Steering Committee has now started the process of planning for a community initiative beyond the current lifetime of CIG. Considerable time was spent on this issue at their May meeting in Pasadena, and the committee will present its ideas at the next CIG business meeting in December (held during the AGU meeting). CIG will have made considerable progress during its lifetime (implementation of modern software development techniques, development of different software packages using common components and/or frameworks, a suite of powerful open-source packages that scale well on large parallel machines, etc.).

Several of our focus groups have concluded that the major challenge to sustained progress in the development of new geodynamics models with significantly higher capability than the existing CIG suite is to incorporate better techniques for meshing, adaptivity, and solvers into our software. The computational science community has developed and is currently developing many such new techniques that potentially could be used to our advantage. However, it is often the case that the differences between real geodynamics problems and the more idealized problems for which these techniques were initially designed prevent them from being simply imported into geodynamics codes. For example, adaptive mesh refinement (AMR) has been identified as a critical, unifying technology in future geodynamics software. But despite the fact that many AMR software packages are now available or under development, it is commonly found that existing AMR packages are rarely suitable to the structure of real geodynamic problems, which typically involve enormous rheological and other property variations, a wide range of spatial complexity, and often long-duration runs.

Instead of passively waiting for the computational sciences community to develop new tools and software libraries that might fortuitously be directly used in geodynamics codes, the Science Steering Committee has concluded that it is now necessary for CIG to become more proactive, and to interface at a fundamental level with members of the computational sciences community involved in the development of these basic tools. Active participation at this level will ensure that such tools will be designed from the outset with the particular difficulties in geodynamics in mind. This approach can lead to significant new payoffs, in our judgment. Our community will benefit by “getting in on the ground floor” of new software development, rather than waiting for useful software to come to us, which is typically a slow and uncertain process. In addition, by actively seeking partnerships in software development, we will be in a better position to bring state-of-the-art geodynamic problems to the attention of the computational science community. This will bring to our community a largely untapped reservoir of new talent, and along with it, greater opportunity for cross-disciplinary collaboration.

One of the ideas for a follow-on to CIG over the next five years would be a broadening of its mission from software engineering to include research and development in the computational aspects of geodynamics. The CIG follow-on would have a more research-based approach that would address computational issues that limit progress in geodynamics (for example, in dynamic partitioning, parallel adaptivity, coupled problems, etc.). We believe that this approach would dovetail well with emerging initiatives within the NSF Office of Cyberinfrastructure (OCI). Given present fiscal realities, the SSC has concluded that any community initiative within the geodynamics community must seek out additional funding sources outside of Earth sciences-discipline focused programs. A community initiative that supported a broad range of traditional computational scientists in methods development could have considerable appeal to OCI as our program would lead to substantial results that cross-cut traditional disciplines.

6. Annual Goals and Milestones (Sept. 1, 2007 – Aug. 31, 2008)

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 (Build-bot; *CIG-Regresstor*)
- e. Expand *Sieve* software suite
- f. Expand benchmark code (*Cigma*)
- h. Science Gateway for benchmarking

II. Core Computational Software
- a. Finish development of Seismology Science Gateway
- b. Deploy Seismology Science Gateway on TeraGrid
- c. Compressible mantle convection code development (*CitcomS3.0*)
- d. Mantle convection support (*HC)*
- e. Long-term tectonics code development (*Gale*)
- f. Long-term tectonics code migration into SVN (*SNAC* donation)
- g. Short-term tectonics code development (*PyLith)*
- h. Geodynamo code bug fixes (*MAG*)
- i. Magma development (*Madds*)
- j. AMR: Generic solution of Stokes flow
- k. AMR: Mantle Convection Code (structured mesh)
- l. AMR: PyLith (unstructured mesh)

III. Organizing Community Participation
- a. Annual meeting of EC, November 2007, Columbia University, NYC.
- b. Computational Seismology workshop (with SPICE and IRIS), Oct 9-11, 2007, Jackson, NH.
- c. AMR Workshop, Oct 24-27, 2007, Boulder CO.
- d. Annual business meeting at Fall AGU, San Francisco, CA.
- e. Annual meeting of SSC, May 2008, Pasadena CA.
- f. Finite element modeling of fault interactions workshop (Co-sponsorship with SCEC, etc., Summer 2008).
- g. Mantle Convection Workshop, Summer 2008 (with other partners).

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

b. Training sessions

  i. Long-term tectonics training session at CIG, Fall 2007.
  ii. Seismology software and Science Gateway, at CIG/SPICE/IRIS workshop, October 2007.
  iii. CitcomS training session at 2008 Mantle Convection workshop.

7. Allocation of Resources by Goal

Sept. 1, 2006 – Aug. 31, 2007

Software Engineer #1 (WL, Numerical Analyst/Modeler) [Total 1 FTE]
  I.c repository: 0.1 FTE
  II.d Gale development: 0.2 FTE
  II.j,k AMR migration: 0.50 FTE
  II.f SNAC Build: 0.05 FTE
  III.c AMR wks: 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
  II .a,b Seismology Science Gateway: 0.5 FTE
  II.j,k,l AMR Library Integration: 0.15 FTE
  III.b Computational seismology workshop: 0.05 FTE
  III.f CFEM workshop: 0.05 FTE

Software Engineer #3 (LA, Software Integration) [Total 1 FTE]
  I.c SVN repository: 0.1 FTE
  I.g Benchmarking (Cigma): 0.50 FTE
  I.h Science Gateway for benchmarking: 0.20
  II.d support for HC: 0.05 FTE
  IV.a.iv Cigma manual: 0.15 FTE

Software Engineer #4 (ET, Numerical Analysis) [Total 1 FTE]
  II.c Compressible Mantle Convection development: 0.5 FTE
  II.j,k 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: 0.25 FTE
  II.h Geodynamo (MAG) bug fixes: 0.1 FTE
  II. a,b Science gateway infrastructure: 0.60 FTE
  IV.a.v MAG manual updates: 0.05 FTE

ANL Subcontract [Total 0.5 FTE]
  I.d Sieve development:0.05 FTE
  II.f PyLith: 0.25 FTE
  II.l AMR Unstructured meshes: 0.1 FTE
  III.c AMR Wkshp: 0.05 FTE
  III.f CFEM Wkshp: 0.05 FTE

VPAC Subcontract [Total 0.75 FTE]
  II.d Gale maintenance: 0.4 FTE
  II: 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 [$65K]
  I, II, III, IV

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

Participant Costs [$91K]
  III, IV

8. Membership

8.1 CIG Members and Member Representatives:

Argonne National Laboratory (MSC)
Brown University
California Institute of Technology
Colorado School of Mines
Colorado State University
Columbia University
Cornell University
Georgia Institute of Technology
Harvard University
Johns Hopkins University
Lawrence Livermore National Laboratory
Los Alamos National Laboratory (ES)
Massachusetts Institute of Technology
Oregon State University
Pennsylvania State University
Princeton University
Purdue University
Rensselaer Polytechnic Institute
State University of New York at Buffalo
State University of New York at Stony Brook
U.S. Geological Survey (Menlo Park)
University of California, Berkeley
University of California, Davis
University of California, Los Angeles
University of California, San Diego
University of Colorado
University of Hawaii
University of Maine
University of Maryland
University of Michigan
University of Minnesota
University of Missouri-Columbia
University of Nevada, Reno
University of Oregon
University of Southern California
University of Texas at Austin
University of Washington
Washington University
Woods Hole Oceanographic Institution

8.2 CIG Foreign Affiliates and Representatives:

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

8.3 Strategy for Keeping Members Informed

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

1. Through e-mail. CIG maintains several list servers through the CIG 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 will be distributed by email on a regular basis.

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

3. The annual CIG Business meeting. This meeting will be open to all and will be a forum for open discussions of the working of CIG, including past and upcoming activities 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 will be presented at these workshops and we expect that CIG members will attend such workshops.

9. Five-Year Management Plan

CIG will need the expertise, vision, and guidance of the community if it is to remain a nimble and evolving organization. Consequently, we 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 by-laws available on our web site (http://Geodynamics.org).

9.1. Institutional Membership, 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. The nominating process for candidates for election is currently underway.

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 future Business meetings 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. Following this allocation of resources, the EC will periodically appoint small committees to interface directly with the software development team (SDT).

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

By the end of the five-year CIG award from NSF, it will be important for CIG to understand clearly the scientific impact that we have had. CIG is a novel project for our Earth science community and therefore it will be essential that we understand how the CIG software has been used and what concrete scientific advances have been made. How has our community changed the way it does science and has this 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-abstracts submitted by members of our community highlighting important scientific results, are being considered.

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 planned MG&G proposal, are examples of how we are now moving in this direction.

10. Annual CIG allocations and expenditures

	FY4 2007-8	FY5 2008-9
Senior (Total)	$56,650	$58,350
Engineers (Total)	$413,632	$426,041
Support (Total)	$110,210	$113,516
Fringe	$148,025	$152,466
Overhead	$345,393	$446,472
Total	*$1,073,910*	$1,196,845
CIG Staff Travel (15 trips)		
CIG Travel Total	$29,032	$30,586
CIG Visitors (10 trips)		
Visit Total	$27,594	$28,646
Total Travel	$56,626	$59,232
Total Material and Supplies	$46,654	$38,327
Subcontract Total	$225,444	$289,812
Total Participant Cost	$105,000	$108,150

Total Budget Amount	($1,507,634)	($1,692,367)
Carry Forward	$200,000	$192,366
Budget Request Amount	$1,500,000	$1,500,000
Balance	$192,366	*$0*

11. Additional funding

none