CIG Science Gateway and Community Codes for the Geodynamics Community - MCA08X011

July 22, 2024

Project Overview

The Computational Infrastructure for Geodynamics (CIG), an NSF cyberinfrastructure facility, aims to enhance the capabilities of the geodynamics community by providing the infrastructure for the development of software that addresses many important unsolved problems in geophysics. CIG's strategy is to support:

1. the benchmarking and validation of its community codes,
2. the development of new and existing codes and ensure they achieve good performance and scalability, and
3. new users by providing technical support, training, and small allocations of computation time.

CIG supports the aforementioned efforts primarily in the following research areas: dynamo simulation, mantle dynamics, seismic wave propagation, and crustal and lithospheric dynamics on both million-year and earthquake time-scales.

In this proposal, we request the renewal of our ACCESS allocation to continue these activities and to test next generation, large-scale computational codes for use in geophysics. In the next section, we describe the major scientific questions and computing challenges that CIG focuses on. We then describe the codes and methodologies used and offer a justification of the requested resources.

Science Background & Objectives

Core Dynamics and Dynamos. It is widely accepted that the Earth's outer core consists of liquid iron alloy and that the geomagnetic field is sustained by its convection (so called geodynamo). Numerical simulations have played a large role in elucidating the fluid motion in the Earth's outer core and geodynamo processes. Although previous efforts (after Glatzmaier and Roberts, 1995) have successfully reproduced some spatial and temporal characteristics of the geomagnetic field, a large discrepancy still exists between the parameters used in geodynamo simulations and actual values associated with the outer core due to the small viscosity. The low viscosity results in a vast range of length scales of the flow required for a comprehensive simulation, ranging from the geometry of the outer core (L ~ 1000km) to the thickness of the boundary layer (L ~ 0.1m). Computational resources are still insufficient to achieve this level of resolution but the community is working to target a middle range (L ~ 100m) that can be achieved using the cutting edge numerical methods and high-end supercomputers available today.

Mantle and Lithospheric Dynamics. Mantle convection and lithospheric dynamics is at the heart of understanding how plate tectonics on Earth works. However, the processes governing plate tectonics remain poorly understood due the lack of direct observations from the deep part of the Earth and the highly non-linear nature of solid Earth deformation. Progress on fundamental questions, such as the dynamic origin of the tectonic plates that cover the surface, coupling between lithospheric deformation and convection deep inside the mantle, feedbacks between erosion and tectonic uplift, the drivers and dynamics of observed magmatism and volatile transport, and the evolution of plate boundary systems all require an interdisciplinary approach. Numerical models of mantle convection and lithospheric dynamics must therefore assimilate information from a wide range of disciplines, including seismology, geochemistry, mineral and rock physics, geodesy, and diverse geologic data sets. The technical challenges associated with modeling mantle convection and lithospheric dynamics are substantial. They are characterized by strongly variable (i.e., stress-, temperature-, and pressure-dependent) viscosities. The lithosphere exhibits processes such as elastic flexure and brittle shear zone deformation (strain localization) that are physically distinct from the viscous flow deeper in the mantle, and occur on fundamentally different (smaller) length scales. In addition, the mantle and lithosphere are chemically heterogeneous, replete with silicate melts and volatiles, and have numerous pressure- and temperature-induced structural changes that affect its dynamics.

While substantial progress has been made in recent years in large part due to improved numerical methods, scaling for massive 3D simulations, and new techniques for assimilating geophysical data sets, significant challenges remain across a wide range of topics. Here, we propose to develop and test new numerical methods in computationally expensive 2D and 3D simulations towards the following scientific goals:

1. Stabilization and accurate decomposition of non-linear viscoelastic-plastic deformation, which governs deformation within the lithosphere.

2. Implementation of new solver packages and improvement of new nonlinear solvers for models using matrix-free preconditioners (Geometric Multigrid) and two-phase flow (see below).

3. Improving large scale IO and code/model interoperability (loading of large datasets, MPI combined with large amount of input data, faster graphical output, generic intermediate exchange format).

4. Assessing the accuracy and efficiency of particles versus compositional fields for tracking viscoelastic stresses.

5. Testing of various new features for simulating lithospheric dynamics, including compressibility and coupled reactive fluid transport.

6. Improved methods for calculating crystal preferred orientations (CPOs) during dynamic simulations.

7. Further integration with external software packages that provide a platform for geologic data assimilation and assembly of complex initial conditions.

8. Improving state of the art high-resolution global mantle flow models incorporating plate boundary databases, databases of subducted plates, and a grain-size dependent rheology.

Work plans associated with these topics are outlined within the Resource Requirements section.

Coupling Core and Mantle Dynamics. The timing of critical events in Earth's evolution, such as the evolution of Earth's inner core, are highly debated, as large uncertainties of core conductivity impede the definition of a well-defined cooling history. Long-term (> 10^7 years) temporal variations in the geomagnetic field, as seen in paleomagnetic data, can be tied to deep Earth processes predicted from numerical geodynamo simulations, and are therefore widely studied as a proxy to study the evolution of Earth's deep interior. With this approach, multiple possible timings for the onset of inner core nucleation (ICN) have been identified from the observation of extreme variations in the ancient geomagnetic field at 550 Ma and between 1.5-1.0 Ga. With the currently available geodynamo simulations, it is difficult to ascertain which – if any – of these signals corresponds to the onset of ICN. This is because the influence of mantle convection, and the resulting changes in amplitude and pattern of heat flux at the core-mantle boundary (CMB), on the geodynamo are still poorly understood. New estimates for the expected range of CMB heat flux variability from mantle convection models run in ASPECT can now be systematically coupled to geodynamo simulations that have demonstrated the ability to produce Earth-like magnetic fields.

Crustal Dynamics: Earthquake time-scales. A rapidly advancing area of crustal geodynamics and one of great societal importance, is the problem of the physics of the earthquake cycle. The increasing volume of high-accuracy measurements of deformation of the Earth's surface in real time in a field long starved for both high density as well as spatially distributed regional and global observational data continues to advance our knowledge. Recent observations made with high precision space geodesy indicate that displacements caused by slow aseismic motions between and following earthquakes, demonstrating substantial post-seismic strain and stress, comparable to coseismic displacements.

It has recently been recognized that relatively modest changes in stress can trigger earthquakes. Theoretical advances in rock mechanics have led to algorithms relating temporal variations in stress to changes in earthquake activity, and are beginning to enable quantitative predictions of how stress changes from fault interactions and fluid injection influence seismic cycles. For example, a 3D finite element model of the Coulomb stress has addressed whether the 1999 Hector Mine earthquake was triggered by the 1992 Landers earthquake. Although results from models such as these have been impressive, more definitive tests require an order of magnitude finer nodal spacing, meshes incorporating the actual elastic structure of the region, the interaction of many faults, and more realistic rheologies.

Seismic Wave Propagation. Seismology provides the means to image the three-dimensional structures within the Earth's interior that are evidence of geodynamic processes. The foundation of computational seismology is the generation of synthetic seismograms and adjoint methods, used in the modeling and inversion for Earth structure, earthquake source, and wave propagation effects. CIG aids the community by supporting 3D codes that provide a more accurate representation of Earth's structure and properties including topography/bathymetry, anisotropy, attenuation, rotation and gravitational effects. Such 3D codes are now revolutionizing seismology, by allowing a direct investigation of countless geodynamic topics such as the fate of subducted lithosphere, existence of mantle plumes, the nature of ULVZ's, lithospheric structure, and plate boundary zone complexity as well as seismic hazard assessment and environmental seismology problems.

Infrastructure. An important role of CIG for the geoscience community lies in the maintenance of geodynamics software packages and in training and supporting the computational needs of their users. The ACCESS allocations enabled part of this work that pertains to large scale computations in two ways: First, running on the various leading edge systems allowed us to improve the software packages and their underlying software stack to run on ACCESS resources. Second, we developed installation and usage guides on ACCESS resources that allows users to get started quickly and without requiring extensive knowledge and research into the machine in question. See https://github.com/geodynamics/aspect/wiki for some examples.

Our work has enabled many scientists to be able to evaluate and use ACCESS resources who otherwise would not have been able to do so.

To that end we plan to:

- Port and tune the software to new ACCESS machines/compilers.
- Continuously maintain the whole software stack and report and fix bugs that appear.
- Perform performance tests and tuning of the codes.
- Create installation and usage guides.

Computational Experiments and Resource Requirements

Numerical Approaches

Rayleigh. Rayleigh is an open source community dynamo code initially launched and developed by the CIG community (with NSF support). It solves the fully nonlinear Magnetohydrodynamics (MHD) equations of motion for a compressible fluid in a rotating spherical shell under the anelastic approximation and employs a pseudo- spectral algorithm with spherical harmonic basis functions and mixed explicit/implicit time stepping (Adams-Bashforth/Crank-Nicolson). A poloidal/toroidal representation ensures that the mass flux and magnetic field remain solenoidal. This code has been performance tested extensively on Stampede2, NASA's SGI Pleiades system, and Argonne's Blue Gene/Q system, Mira.

ASPECT. ASPECT is a CIG developed code designed to solve the equations that describe thermally driven convection in the Earth's mantle and tectonic deformation in the Earth's lithosphere. It allows for both 2D and 3D models of arbitrary geometry (generally focused on segments of or whole mantle), adaptive mesh refinement in locations of scientific interest, easy modification of material, gravity, rheology and temperature models, and tracers to model geochemistry and material transport. Recent work has started investigating the effectiveness of GPU or other coprocessors in ASPECT simulations. Further details are available in [Kronbichler et al., 2012; Heister et al., 2017, Bangerth et al., 2023]. Based on improvements we made on ACCESS resources, we were awarded 161,120 SUs for TACC Frontera Pathway allocation from April 2023 to August 2024 for both ASPECT and Calypso. The award was used to optimize ASPECT's solvers (e.g., matrix-free geometric multigrid methods) on the Frontera system.

Leeds Spherical Dynamo (LSD) LSD solves the Boussinesq dynamo equations by representing velocity, and magnetic field, as poloidal and toroidal scalars [Willis et al., 2007]. It is pseudo-spectral; the variations in a sphere are expanded in spherical harmonics, and radial variations are discretized by finite differences with a nonequidistant grid using Chebyshev zeros as grid points. The nonlinear terms are evaluated by the transform method. Time stepping is using a predictor-corrector method and the time step is controlled using a CFL condition and error information from the corrector step. The LSD code is parallelized using MPI in both radial and in θ (meridional) directions [Matsui et al., 2016].

PyLith. PyLith is a 2D and 3D finite element code for modeling interseismic and seismic processes related to capturing the physics of earthquakes, including slow strain accumulation, sudden dynamic stress changes during earthquake rupture, and slow postseismic relaxation. Implicit time stepping provides efficient time integration for quasistatic (interseismic deformation) problems, and explicit time stepping provides efficient time integration for dynamic (rupture and wave propagation) problems. Key features of PyLith are its ability to accommodate unstructured meshes (which allows complex non planar fault geometry), implementation of a variety of finite element types, and implementation of a variety of fault and bulk constitutive models appropriate for the Earth's lithosphere. The bulk constitutive models include linear and nonlinear viscoelastic models in addition to linear elastic models. PyLith uses PETSc [Balay et al., 1997, 2001, 2004] to achieve fast, efficient, parallel solution of the partial differential equation.

SPECFEM3D GLOBE. Led by Princeton University and in collaboration with domestic and international partners (see https://specfem.org/about), CIG offers this software, which simulates global and regional (continental scale) seismic wave propagation using the spectral element method (SEM). The SEM is a continuous Galerkin technique, which can easily be made discontinuous; it is then close to a particular case of the discontinuous Galerkin technique, with optimized efficiency because of its tensorized basis functions. In particular, it can accurately handle very distorted mesh elements [Oliveira and Seriani, 2011].

SPECFEM3D GLOBE has very good accuracy and convergence properties [De Basabe and Sen, 2007]. The SEM approach admits spectral rates of convergence and allows exploiting hp convergence schemes. It is also very well suited to parallel implementation on very large supercomputers [Carrington et al., 2008] as well as on clusters of GPU accelerating graphics cards [Komatitsch, 2010].

Table 1: List of Websites

Code	Website
Rayleigh	https://geodynamics.org/resources/rayleigh
ASPECT	https://geodynamics.org/resources/aspect
PyLith	https://geodynamics.org/resources/pylith
SPECFEM3D GLOBE	https://github.com/SPECFEM/specfem3d_globe

Resource Requirements/Usage Plan

CIG researchers have only used a portion of the attributed resources and plan to continue using the remaining computation time on the ACCESS platform over the next allocation period for (1) scalability testing and code validation, (2) development of new numerical methods for better code performance, (3) science production models as described below, and (4) workshops, education, and training sessions to nurture new users on ACCESS resources using ASPECT, PyLith, Rayleigh, and SPECFEM3D GLOBE. New users will require additional SUs to conduct their research. CIG expects to support their feasibility testing and spin up, which will enable researchers to apply for their own allocations. Details of how the requested resources are going to be used are provided below.

SPECFEM3D_GLOBE: Full-waveform inversion workflows

We will demonstrate containerizing 2D and 3D full waveform inversion (FWI) workflows for training workshops and classroom teaching, as well as production multiscale FWI runs on a range of available resources such as personal laptops, clusters, and supercomputers. FWI workflows are well-defined and typically consist of three stages:

1. preprocessing stage (i.e., data download, filtering time series, selecting measurement windows, measurements, and computation of adjoint sources, etc.),
2. numerical forward and adjoint simulations to compute synthetic seismograms and data-sensitivity kernels, respectively (computationally the most expensive part),
3. post processing stage (i.e., summation of event kernels to compute gradients, smoothing and preconditioning gradients, line search, model updates, etc.).

The initial setup and maintenance of the entire FWI workflow, specifically the Python environments for the preprocessing stage as well as the effective configuration and compilation of 3D solvers require a nontrivial amount of work on different computer systems. Our goal is to use containers to reduce the complexity and increase the reproducibility of results. We have containerized a 2D FWI workflow and also demonstrated a sample case in 3D that can be run in workstation and HPC environments.

Containers are initially created and run on the workstations using Docker and on the TACC's Frontera system, where Singularity is used as the container platform. During the NSF-CSSI funded SCOPED project's training workshop in Seattle in May 2024, about 40 on-site participants demonstrated running two sample banana-doughnut adjoint kernels for P waves propagating through an oceanic and a continental domain. Each trainee performed their low-resolution (NEX=96) one-chunk simulations by the containerized SPECFEM3D_GLOBE using 36 cores on a reserved single node of TACC's Frontera system. For the current setup forward and adjoint simulations for 30 s seismograms take about 8.5 min and 24 min, respectively. The simulations also included the pre-processing stage as well as remote visualization of kernels with Paraview where all the steps run smoothly and successfully

As part of this allocation, we will also demonstrate the containerized FWI workflows on ACCESS Stampede3 to further tune the containers to multiple HPC platforms. We also plan to use the containerized workflows in future CIG-SCOPED training workshops. Based on our kernel calculation experience at the SCOPED workshop in Seattle, in future workshops we aim to run a complete FWI workflow including computation of event kernels (with ~3-5 earthquakes) and a model update. We request *5,000 SUs on Stampede3* for this effort including a training workshop with about 50 participants and preparation and tests of the containerized FWI workflows.

Rayleigh

Cian Wilson and Peter Driscoll at Carnegie Science Earth & Planets Laboratory, in collaboration with Nick Featherstone of the Southwest Research Institute, have extended the functionality of Rayleigh to include extra scalar fields that couple to the velocity field through buoyancy forces and to each other through coupled boundary conditions. Ongoing work is extending this functionality to include more potential couplings, such as through source terms. These modifications change the underlying matrix structure and size of the innermost solve on the radial grid for individual spectral harmonics. To ensure this does not affect Rayleigh's scaling we seek *10,000 SUs on Stampede3* to perform scaling and performance analysis as development proceeds.

Mantle Convection and Lithosphere Dynamics Modeling with ASPECT

Below, we outline distinct projects using ASPECT that will require ACCESS resources.

Continental Rifts Formation

Prof. Eunseo Choi with his student Chameera Silva will continue to develop models that can represent the development of passive continental rift margins. These models will help in better understanding of some of the most fundamental geological processes here on Earth. In the current allocation period, they have tested the implementation of their models for the rift driving forces, material properties, and geometry of the weak zone initiating the rifts using two-dimensional models.

Next, they plan to extend these models in three-dimension to study the rift reactivation scenarios considering mantle plumes. To do this, we estimate that they would require *75,000 SUs on Stampede3*.

Mantle Convection Studies

Prof. Juliane Dannberg and Prof. Rene Gassmoeller will continue their work on mantle convection modeling together with their graduate student Ranpeng Li. The next objective is to apply a recently developed method to accurately model mineral phase transitions [Dannberg et al., 2022] to large-scale global mantle convection simulations. These models can answer the question of how phase transitions have changed the pattern of mantle convection throughout Earth's history. The method itself is applicable to a wide range of geodynamic problems such as the dynamics of subduction and mantle plumes, material recycling through the Earth's interior, lithospheric dynamics etc. Since the new method can model phase transitions — a critical component of many geodynamic modeling studies — more accurately than the methods commonly used in the field. Testing its applicability to large-scale realistic problems will benefit the whole modeling community and will provide a basis for multiple groups to build on other research projects. Our initial tests show that we require about 60,000 core hours for models testing a specific parameter regime for a given time in Earth's history, and about 250,000 core hours for models covering whole Earth evolution after the end of the magma ocean period. We plan to run 20 models to map out the parameter space and then three production runs covering Earth's evolution from the early Earth to the present, amounting to a total of *40,625 SUs on Stampede3* .

New methods for two-phase fluid transport

Prof. John Naliboff and his PhD student Daniel Douglas at New Mexico Tech will test the recently implemented methods for comparing coupled versus uncoupled two-phase simulations in ASPECT, with a focus on applications to volatile transport in the convecting mantle and lithosphere. These comparisons will involve benchmarking in 2D and 3D, in combination with extensive testing of solver parameters. Following the completion of benchmarking against analytical or previously published solutions, a limited number of high-resolution 3D simulations of the Hikurangi subduction zone will be conducted.

For this, they request *33,600 SUs on Stampede 3*, which would allow extensive testing in both 2D [*4,800 SUs [40(models) x 5(nodes) x 24(hours)]]* and 3D [*28,800 SUs [20(models) x 15(nodes) x 24(hours)]]* .

Stokes Solver Improvements

Prof. Timo Heister and his student Quang Hoang have been working on improvements for the matrix-free Stokes solver in ASPECT. Specifically, they are replacing the preconditioner for the geometric multigrid solver to handle large viscosity contrasts in the models. In their initial implementation, they found that the new BFBT preconditioner is indeed nearly unaffected by an increase in viscosity ratio. They plan to test and benchmark the new solver against the existing one on different benchmark problems and require *20,000 SUs on Stampede3*.

Detachment Fault Formation

Dr. Moh Gouiza will continue to examine models of lithospheric rifting, examining the tectonic and physical processes that control the linkage between crustal-scale detachments faults. The models are constrained by high-resolution 3D seismic data, from the northern margin of the South China Sea, which were used to map the 3D geometry of a crustal-scale detachment fault system.

In the future, they plan to increase the resolution of the models, from 1.25x1.25x1.25 km to 0.625x0.625x0.625 km, to investigate the interaction between several overlapping detachments faults. To do so, they expect to run several 3D simulations on 30 nodes for 48-72 hours, which would require approximately *15,000 SUs on Stampede3*.

Tracking of material properties and stresses

Profs. John Naliboff, Bob Myhill, Juliane Dannberg, Rene Gassmoeller and Dr. Anne Glerum will continue their collaborative work on the implementation of a new viscoelastic-plastic (VEP) formulation in ASPECT. The majority of this work includes testing the relative efficiency and accuracy of using particles versus compositional fields to track lithologies, viscoelastic stress tensor components, and additional properties in high-resolution 2D and 3D simulations. While related development work for a range of analytical benchmarks has occurred in 2D, the trade-offs between efficiency and accuracy are likely to be different in high-resolution and more complex 2D/3D simulations. Over the past reporting period, the majority of work on this topic has been dedicated to the underlying numerical VEP formulation and assessing the tradeoffs between particles and compositional fields in complex (i.e., realistic) 2D simulations of continental extension and tectonic inversion. The work proposed here will focus on extending these tests to complex 3D simulations. Although the size and required run times of lithospheric dynamics simulations vary dramatically, we have found a reasonable average for high-resolution simulations on *Stampede3* is 15 nodes used over 24-48 hours (360-720 SUs). Here, we request *9,600 SUs [20(models) x 20(nodes) x 24(hours)]*

New Rheologic Formulations for Tectonics

Profs. John Naliboff, Bob Myhill, Cedric Thieulot, Timo Heister, and additional collaborators will continue their work on developing and testing new rheological formulations that will be of significant use to the ASPECT mantle convection and lithospheric dynamics community. This work will include further testing of currently implemented plasticity stabilization in combination with adaptive mesh refinement, composite rheological formulations, and compressibility. While the majority of development work for these topics will be completed using small simulations on local resources, the resources requested here will be used for limited proof-of-concept production models that will be highlighted in the ASPECT repository. As the underlying simulations are very similar in nature to those outlined to those in the previous section, here we also request *9,600 SUs on Stampede3 [20(models) x 20(nodes) x 24(hours)]*

netCDF and I/O testing

ASPECT recently gained support for loading large datasets using netCDF, but this has not been tested on large parallel computations. Prof. Timo Heister is planning to verify the parallel scalability when loading large files and perform other I/O performance tests to ensure ASPECT works correctly for high-fidelity simulations. For this, *5,000 SUs on both Stampede3 and Expanse* systems are requested.

Education and Training

An emphasis for CIG moving forward is the contribution to training and education efforts using Jupyter Notebooks and containers deployable from hosted resources (vs. virtual machines, binaries, or local installation). CIG has been already involved in training over 125 users annually on CIG community software. We anticipate increasing our training and education initiatives by developing education resources for geodynamics and computational modeling to be deployed in a similar fashion as well as expanding access for undergraduate research. Thus, we anticipate utilizing *1000 SUs annually on Stampede3*.

Resource Allocation Request

The following table summarizes our request for the period October 1, 2024 - September 30, 2025. We anticipated updating this request with our ACCESS renewal annually based on community needs. The CIG project is funded by NSF through January 31, 2028.

Table 2: Requested SUs for Stampede3

Software	Purpose	Requested SUs*
Rayleigh	Scaling & Performance Analysis on Extended Coupled Equations	10,000
SPECFEM3D	Full Waveform Inversion Workflows	5,000
	Continental Rifts Formation	75,000
ASPECT	Mantle Convection	40,625
	Two-Phase Fluid Transport	33,600
	Stokes Solver Improvements	20,000
	Detachment Fault Formation	15,000
	Tracking Material and Stresses	9,600
	Rheologic Formulations	9,600
	netCDF and I/O	5,000
ASPECT and PyLith	Education and Training	1,000
	TOTAL	**224,425 SUs**
	ACCESS Credits	**4,024,473**

* We assume here running on Skylake where 1 SU = 1 Node hour

Table 3: Requested allocations for Expanse (core hours)

Software	Purpose	Requested core hours
	netCDF and I/O	5,000
	TOTAL	**5,000**

ACCESS Credits Requested: 4,029,473