Computational Infrastructure for Geodynamics - Community Code Scaling

1. Summary of results for the early access period and Pathway allocation 2020 and 2021

The Computational Infrastructure for Geodynamics (CIG) is an NSF-funded organization dedicated to providing and maintaining a suite of high-quality, open-source software packages that are widely used in the geosciences to simulate the dynamics of the solid earth (commonly called "geodynamics"). CIG partners with U.S. and international researchers in the development of computational modeling software. An important mission of CIG is the development of computational capabilities in the scientific community we serve, particularly the use of modern numerical methods, software development methodologies, and high performance computing facilities.

Software developed by CIG and associated researchers has been shown to scale to thousands, tens of thousands, or even more cores. We aim to further the scalability of these codes to larger machines using the next generation of hardware architectures, so users of these packages can efficiently run their simulations once these machines become available to the wider research community.

To achieve this goal in 2021, we received an Frontera Pathway allocation of 140,386 SUs on Frontera to produce scaling results and example production runs using two CIG community codes: ASPECT and Calypso for mantle and core dynamics, respectively. As of June 10, 2022, we used 94,487 SU's and expect to finish the rest of allocations by Aug. 31, 2022. These codes are also used by a wide range of research groups on existing XSEDE resources (Stampede2, Comet). The geodynamics community anticipates using them on the next generation of XSEDE resources to tackle increasingly complex problems in geodynamics. We know from our years of experience gained dealing with parallel codes that the scaling runs we propose herein will point out current bottlenecks. Continued access will help us to further identify the root causes, improve the scalability of our codes, and test the scaling performance of new features that are developed concurrently with the work proposed herein.

Based on results we have achieved using a previous allocation, we would like to request a renewal of pathways allocation that will allow us to continue our work. We discuss our prior results and propose a future plan below.

1.1 ASPECT

ASPECT (the "Advanced Solver for Problems in Earth ConvecTion") is an open-source finite-element modeling software that simulates solid-state, creeping flow of material. ASPECT makes use of the libraries deal.II, p4est, and Trilinos libraries, all software packages that are widely used in computational science and that provide massively parallel finite element, meshing, and linear algebra capabilities (see http://www.dealii.org/, http://www.p4est.org/, https://trilinos.github.io/, and also https://aspect.geodynamics.org/).

1.1.1 Technical achievements

Access to Frontera enabled the development of necessary features in the finite element library deal.II that is used by ASPECT and major new developments in ASPECT itself to support fast, scalable computations and more recently scalable IO for checkpointing and graphical output.

We have been developing a framework for matrix-free, massively parallel geometric multigrid inside deal.II [1]. It is amenable to a variety of partial differential equations and starting from the early access period, we have used Frontera for medium test runs and very large scaling tests.

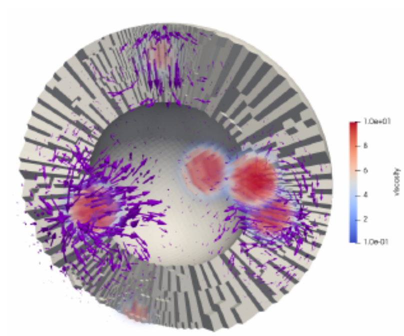
In 2020/2021, we focussed on implementing the matrix-free multigrid solver in ASPECT. Applying it to a Stokes problem with large viscosity contrasts is non-trivial but we were able to show great results on Frontera scaling to 100k MPI ranks and up to 200 billion degrees of freedom (100x larger than possible before). The new solver is 3x faster and uses 10x less memory, which allowed us to increase the total problem size dramatically, see [2, 3].

The last year / access period was used to refine and improve the linear solver developed beforehand and to dramatically speed up and scale up parallel input and output for input data sets, graphical output, and checkpointing.

First, we developed a new geometric multigrid strategy using global coarsening for better workload distribution for highly adaptively refined meshes. We report on the results in [4]. Medium-sized runs on Frontera show 2-3x better performance in ASPECT (see Figure 1).

Second, we worked on making the matrix-free solver in ASPECT applicable to many more complicated problems including models with free surface deformation, support for highly nonlinear problems using the Newton solver, and more flexibility regarding boundary conditions.

Third, with the ability to perform much larger computations, the need for improved IO came more and more apparent. Large input datasets that need to be accessible on each MPI rank can now span 10s of GB. To support this, we fixed several issues, improved performance, and implemented MPI shared fenced allocations, so only one copy is stored on each compute node. IO also appears when reading/writing checkpoint files and graphical output. We conducted various experiments and scaling tests on Frontera and were able (with the help of Frontera staff!) to now support very large checkpointing and graphical output files by reimplementing our MPI I/O support. Early performance tests show an improvement of about 10x to the old implementations (now 23 GB/s on /scratch1/). We had to work around Intel MPI large IO bugs and implement a new low-level MPI IO layer in deal.II, see [5, 6].



L	#DoFs [1e6]	global coarsening #it	global coarsening solve [s]	global coarsening V-cycle [s]	local smoothing #it	local smoothing solve [s]	local smoothing V-cycle [s]
5	10.0	25	1.38	0.003	25	1.15	0.006
6	20.7	25	1.61	0.006	25	1.84	0.008
7	43.2	25	2.26	0.009	26	2.63	0.013
8	88.0	22	2.39	0.013	27	4.15	0.022
9	178.0	25	4.62	0.020	28	7.87	0.047
10	355.0	26	6.47	0.039	28	13.66	0.109
11	715.7	27	10.65	0.069	27	23.67	0.214
12	1441.4	28	22.17	0.141	29	50.85	0.436
13	2896.7	28	37.38	0.266	26	99.07	0.971
14	5861.9	29	77.94	0.515	26	193.19	2.060

Figure 1: Stokes flow in a spherical shell using global coarsening GMG in ASPECT showing a 2-3x improvement over the old solver ("local smoothing"). Taken from [4].

1.1.2 Science achievements

In our previous phase of Frontera allocation, we used ASPECT to set up a series of global 3D instantaneous models of mantle convection with a goal of finding a model that could generate the present-day surface velocities. Each of these models are discretized into 3D hexahedron elements utilizing a nonlinear fourth-order mapping from unit cell to real cell, which accounts for the spherical curvature of each element. To be able to model narrow tectonic plate boundaries, we use an adaptive mesh in ASPECT and a resolution between 17 km and 82 km depending on the location in the model. This range of resolutions results in models with approximately one billion degrees of freedom, and a typical output size of 18 GB. We make use of the recently implemented matrix-free, geometric multigrid solver (see above).

In order to constrain a physical model that best matches with the observed surface velocities, and investigate the importance of the the different model components, we vary the geometry of plate boundaries (4 types), the prescribed strength along the plate boundaries and at the base of the lithosphere (9 - 25 values), the strength of the continental cratons (2 models), the temperature distribution in the model (3 models), and the viscosity of the subducted slabs (2 models). The combination of all these model components resulted in 62 models.
Through our parameter study, we find our best-fit model has a directional fit of 92% and can explain 85% of the average speed relative to the observed GPS velocities (Figure 2). We also find that the choice of plate boundary geometry is critical for the direction of plate motions and using a plate boundary model with closed polygons where all plate boundaries have the same strength — as done in the previous mantle convection studies — is not a good approximation of plate tectonics on Earth.

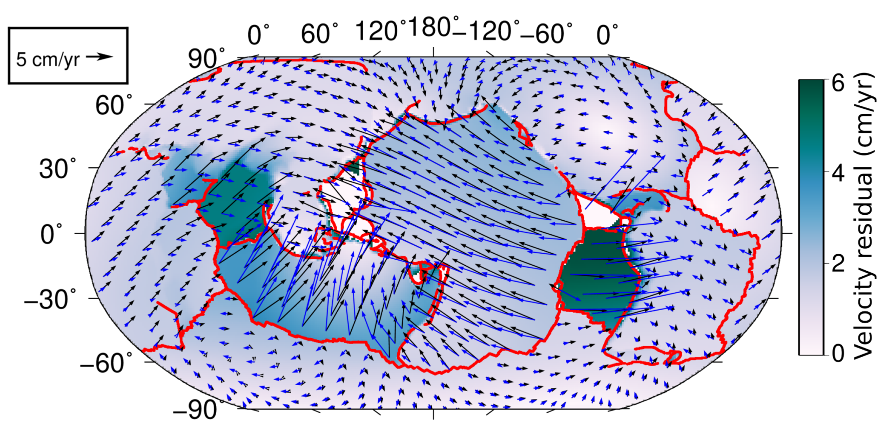


Figure 2: Velocity residual together with the modeled velocity at the surface (blue arrows) and the observed GPS velocities (black arrows) for the best-fit model.

1.2 Calypso

Calypso is a three dimensional magnetohydrodynamics (MHD) model to solve geodynamo problems. It uses a pseudo-spectral method and a finite difference method in the horizontal and radial discretization, respectively. Calypso is parallelized through both MPI and OpenMP. In MPI parallelization, the directions of domain decomposition are changed in the spherical harmonics transform. A parallel volume rendering module has been included in Calypso to enable visualization during the simulation runtime. In 2021, we also implemented the 3-D Line Integral

Convolution (LIC) module for visualization of vector fields such as the velocity or magnetic field (see Figure 3).

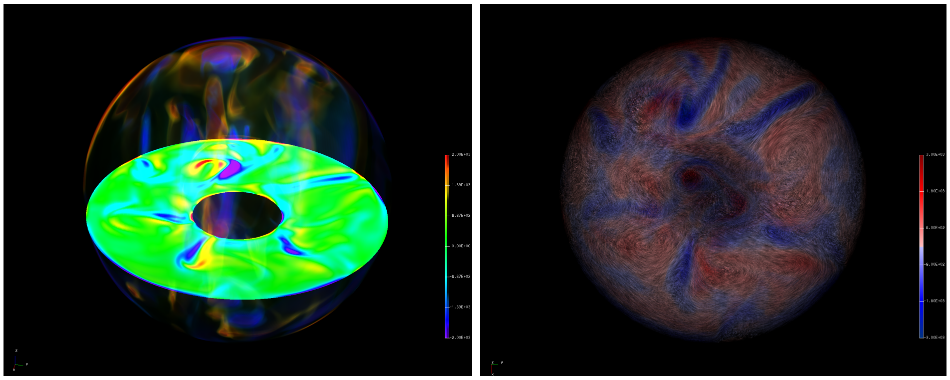


Figure 3: Visualization results by the volume rendering module (left) and by LIC module (right). In the left panel, The *z*-component of the vorticity $\boldsymbol{\omega}_z$ is displayed by the volume rendering and section at z = -0.3. In the right panel, the flow pattern in the southern hemisphere is displayed by LIC colored by the $\boldsymbol{\omega}_z$.

We tested the scalability and performance of Calypso on Frontera up to 512 nodes (28,672 cores) with two different spatial resolution cases. As seen in Figure 4, both simulation and visualizations including the LIC module keeps good scaling up to 28,672 cores. Through optimization of the loop orders in the simulation modules, simulation time decreases approximately 0.7 times as compared to the modules in Ver.1.2. In addition, the present optimized version (2022 version) keeps better scaling at the largest parallelization level (28,672 cores). To keep the simulation time to more than 0.9 times of the total elapsed time, the results in Figure 4 suggest that each LIC visualization will be performed for every 1,000 steps. The present optimization of the simulation module will be included in the next release of Calypso Ver. 2.0.

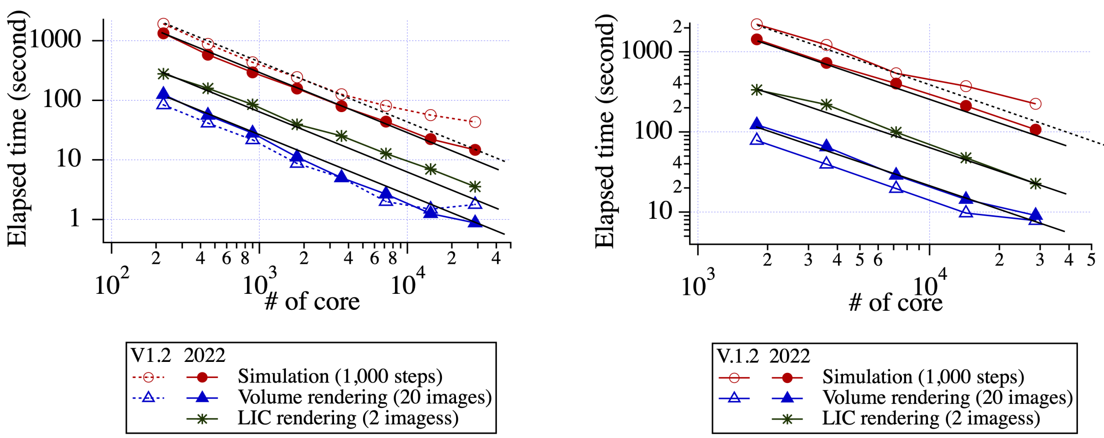


Figure 4: Strong scaling of Calypso on TACC Frontera for up to 28,672 processor cores. Simulations are performed along volume and LIC rendering. The simulation and visualizations are performed with (N_r, N_θ, N_ϕ) = (336, 384, 768) grids and (N_x, N_y) = (1600, 1200) pixels, respectively for the left panel. In the right panel, the simulation and visualizations are performed with (N_r, N_θ, N_ϕ) = (670, 768, 1536) grids and (N_x, N_y) = (3200, 2400) pixels, respectively. Results obtained by the Ver. 1.2 are plotted with the open symbols for comparison.

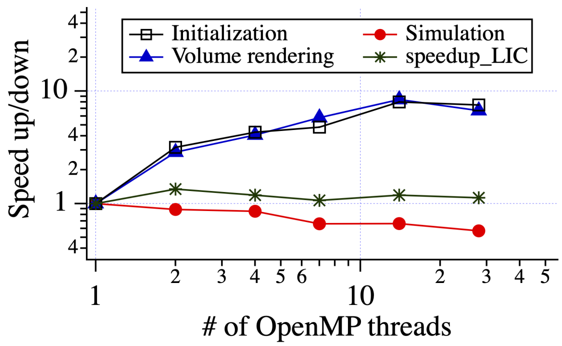


Figure 5: Elapsed time for simulation and visualizations with fixed resolution (N_r, N_θ, N_ϕ) = (336, 384, 768) and fixed number of processor cores (56x256 cores) as a function of OpenMP threads.

We also investigated the performance of OpenMP parallelization through additional testing. In these tests, we fixed the spatial resolution and total number of nodes, while changing the relative number of OpenMP threads and MPI processes. As seen in Figure 5, the performance for simulation decreases with increasing the number of threads, while the LIC module almost keeps its performance by increasing the number of threads. On the other hand, the performance volume rendering module increases rapidly with increasing the number of threads (i.e. decreasing the number of MPI processes).

1.3 Publications and other communications

Publications [4]-[11] and other citations are, in part, enabled by the past allocation on Frontera.

2. Proposed Work

2.1 ASPECT

As we will discuss in the science-driven goals below, we will develop the following features:

1. Integrating more large-scale data for subducted plate geometries (implementing readers for the Slab 2.0 model into ASPECT).
2. Integrate our existing grain-size evolution models into the 3D large-scale models developed during the last year.

2.1.1 Technical aspects

Our plan for technical improvements and scaling tests include:

First, improvements to the current linear/nonlinear solvers. This includes improved evaluation of coefficients in the multilevel hierarchy (currently Newton derivatives are ignored in the preconditioner for example) and more options for tangential flow boundary conditions on curved geometries.

Secondly, we want to properly integrate and test the currently experimental global coarsening GMG into ASPECT. The prototype for [4] only works for a very limited set of models.

Thirdly, we plan to perform performance and scaling tests of the new features. This will hopefully finally include runs at full machine scale on Frontera.

Finally, some technical work will be necessary for archiving our science-driven goals mentioned above and in 2.1. While none of the required features have been used in ASPECT models of the scale we now regularly run on Frontera, we do not anticipate large technical difficulties implementing them. Integrating additional data on the order of a few GB per dataset can build on the same framework for large parallel data input we have developed over the past years. We will build on our IO improvements including shared memory functionality to store data only once per node. Our grain-size models were not used on large scale models before either (>1000 MPI ranks), however their compute load is purely local, requiring no communication. Therefore, we do not expect an impact on scaling or parallel efficiency.

2.1.2. Science-driven goals

In the previous phase we developed global 3D mantle flow models of the Earth with large viscosity variations and narrow plate boundaries using an adaptive resolution of up to ~ 6 km. Through a parameter study, we found models that give us a good directional fit of up to 95% to the observed plate velocities. We will extend these models as laid out in the NSF proposal that funds this research (EAR-1925677). In particular we will:

1. Include a more realistic slab morphology as described in the Slab 2.0 database provided by the United States Geological Survey (USGS). Slab pull and suction are main driving forces for surface plate movements. In the previous phase of this project, we have shown that these forces depend heavily on the precise geometry of the subducting slabs. Therefore, we expect a more data driven approach to slab integration to result in more realistic surface deformation patterns, which are better able to explain the present-day stress state of the lithosphere.
2. Integrate an updated lithospheric structure model.
3. Improving upon the assumption of a constant mineral grain size by coupling to a physics based grain-size evolution model as we have done in an earlier study. The grain size of minerals heavily influence the viscosity of the material and therefore the flow pattern. Simple assumptions like constant grain size are common in geodynamic modeling studies, but can significantly mispredict the resulting flow. In addition, not knowing the approximate grain size leads to significant misinterpretations of seismic tomography models when used in geodynamic modeling studies as we have shown in an earlier study [12]. We have developed a dynamic grain-size evolution model, but so far lacked the resources to implement it in large-scale 3D models. Now that we have developed a large-scale 3D model on Frontera, we can extend this previous work to realistic 3D models to assess how much the current simplification of constant grain size limits our understanding of global plate motions.

2.2 Calypso

2.2.1 Technical aspects

We obtained good performance by using MPI and OpenMP hybrid parallelization for simulation and visualization modules for parallel volume rendering (PVR) and line integral convolutions (LIC). After the optimization of ordering of the loops, Calypso obtained approximately 30% better performance than the previous version.

Data in Calypso can be visualized during and post simulation runs. PVR is used for visualization of the scalar component, and Parallel Line Integral Convolution (LIC), newly developed, to visualize vector fields [13]. After implementing domain re-partitioning and sleeve extension modules for LIC, we obtained good parallel performance for the LIC module for up to 28,672 cores on Frontera. However, we also found that elapsed time for the volume rendering and initialization of the simulations rapidly increases with the number of MPI processes. We need more investigations to find the best MPI and OpenMP configurations using more than 28,672 cores for the simulation and visualizations.

2.2.2 Science-driven goals

We will develop and investigate a sub-grid scale model (SGS) model to model the effects of turbulence on the large scale convection and magnetic field generation of the Earth's core [14, 15]. In previous studies, we only investigated the characteristics of the SGS terms from resolved, direct simulations. To establish the validity of this approach, it is necessary to perform both large- and small-scale simulations for reference on Frontera. In addition, we will perform fully resolved simulations by using nonlinear terms which are obtained by filtered large-scale fields to investigate which turbulence process is the most important for generating and sustaining the geodynamo.

Calypso has 8 active users in 4 universities. Several projects are starting:

A. Geodynamo modeling for the past Earth. The early Earth had a smaller solid inner core than present.
B. Investigation of energy transfer between kinetic and magnetic energies during the geomagnetic dipole reversal.
C. Comparison of evolution of the dipole component between dynamo simulation and observed geomagnetic field using stochastic models.
D. Effects of aspherical growth of the inner core due to the thermal heterogeneity at the inner core boundary driven by the convection of the outer core [11].

The improvement of the performance on Frontera is expected to be applicable for Frontera and other supercomputers with similar architecture.

3. Resource usage plan

3.1 ASPECT

For the implementation and verification of these new features as well as answering the science questions mentioned above, we plan to run several mid-sized 3D models with minimum cell size of ~ 17 km – each utilizing approximately 250 SUs as for the models of the previous phase. We expect that the models utilizing the multi-physics grain-size evolution model the necessary compute time will increase by approximately a factor of 2, therefore calculating 500 SUs per model with dynamic grain-size evolution. We will require about 42000 SUs (distributed described in detail below) to run all the models with the additional improvements.

This work will be performed by Dr. Arushi Saxena, Dr. Juliane Dannberg, and Dr. Rene Gassmoeller at the University of Florida.

3.2 Calypso

We plan to test the Calypso including the volume rendering and LIC modules by using the large allocation class to investigate the performance under the larger parallel environment. We request 15,000 SUs for the performance test under the massively parallel environment.

We plan to perform simulations with the truncation of the spherical harmonics at L_{max} = 511 and 768 radial grid points on 256 nodes of Frontera. Based on Figure 4, we expect that we can perform 2,000,000 steps for 9 days of wall clock time. To establish the validity of this approach, it is also necessary to perform small scale simulations including the SGS model on Frontera. For the model including the SGS model, 32 nodes for 200 hours will be used for SGS model simulations with smaller spatial resolution.

We also plan to perform simulations to represent dipole reversals in the geodynamo simulation. To represent the dipole reversals, we need to run simulations to 50 magnetic diffusion times, which corresponds to 250 million time steps. We will perform the present model with L_{max} = 159 and 160 radial grid points on 64 nodes of Frontera for 15 days.

The work for Calypso will be performed by Dr. Hiroaki Matsui at University of California, Davis.

3.4 Resource table

The summary of the resource usage plan is given in the following table.

Type of computation	Resources per model	Total resources
ASPECT solver scaling tests, development, benchmarks, etc.	Highly variable (1 - 4,000 SUs)	10,000 SUs (various sizes)
ASPECT science models with improved slab data integration	250 SUs (mid-scale), 2,000 SUs (large-scale)	20,000 SUs (40 mid, 5 large)
ASPECT multi-physics grain size models	500 SUs (mid-scale), 4,000 SUs (large-scale)	22,000 SUs (20 mid, 3 large)
Calypso optimization and tests	2,000 - 10,000 SUs	10,000 SUs
SGS modeling for geodynamo simulations by Calypso	76,000 SUs for large scale DNS 6,400 SUs for SGS model	82,400 SUs
Calypso dipole reversal run	23,040 SUs	23,040 SUs

Total requested: 167,440 for Frontera.

3.5 Disclosure of Access to Other Compute Resources

CIG has access to the following resources from Apr. 1, 2021 to Sep. 30, 2022 under award TG-EAR080022N (*used/total SUs):*

TACC Ranch storage 0/2,048
TACC Stampede 31,690/50,000
SDSC expanse.ucsd 990,498/1,377,206
SDSC expanse-ps.ucsd 0/2,048
PSU bridges2-gpu 0/15,000
PSU bridges2 storage 0/2,048

CIG has access to the Peloton cluster at UC Davis with 109 compute nodes (5,536 cores).

T. Heister has access to the Palmetto cluster at Clemson University with currently 2,079 compute nodes (23,072 cores) and which allows scaling runs with up to about 2,000 cores.

A. Saxena and R. Gassmoeller have access to a 500 core allocation on the UF HiPerGator cluster.

H. Matsui has access to the cluster at the National Institute of Polar research of Japan with 18 computational nodes (720 cores) and allows each run on up to 8 nodes (320 cores).