CCSM Software Engineering Working Group Meeting Report
Wednesday, 25 January 2006
Via Vislab and Webcast, NCAR, Boulder, CO

Peter Gent (NCAR-CCSM) outlined general scientific priorities for CCSM4 development. These included the addition to CCSM3 of an interactive carbon/nitrogen cycle, the improvement of CCSM3 physics for reducing biases in the tropics (ENSO and MJO), the inclusion of the indirect effect of aerosols, and the addition of atmospheric chemistry. The atmospheric dynamical core will be the finite volume (FV) core. A high priority is to start control simulations as soon as possible using this core. Timelines for the next IPCC AR5 were also discussed.

Mariana Vertenstein (NCAR-CSEG) gave a summary of current CSEG development efforts. CSEG has successfully migrated the entire CCSM code base from CVS to Subversion, a next generation revision control system. Project tracking has also been improved via use of the new CCSM swiki (see http://swiki.ucar.edu/ccsm/2). A summary was also given of current CSEG efforts for creating a single-executable sequential CCSM and completely rewriting the current CCSM data models (the data model Version 7 project). The creation of a single-executable sequential CCSM will produce a model that permits plug and play functionality (i.e., a user can easily swap active and data components), keeps full backwards compatibility with current multiple executable/ concurrent CCSM functionality, keeps full backwards compatibility with all current stand-alone CAM functionality, and removes the need for stand-alone CAM. This new code base will also remove code duplication (no component specific models will exist) and standardize coupling interfaces. The current system is being designed such that alternative coupling mechanisms (e.g., MCT) can easily be utilized. Having more than one coupling system incorporated will be beneficial for both testing and system validation. The creation of a single-executable sequential CCSM is proceeding via two ESMF evaluation stages. In the first evaluation stage, ESMF will be used to couple CAM to its surface components via the creation of a top level application driver. An update was given on the progress that has already been made in the ESMF Stage 1 Evaluation Plan and work that still needs to be accomplished for its completion.

A summary was also given of the progress that has been made in both the development and release versions of the CCSM scripts to support new machines. The CCSM scripts have also been modified to provide a simpler user friendly tool to determine and specify optimal load balanced configurations.

Cecelia DeLuca presented an update on ESMF. There are now more than three dozen ESMF-compliant geophysical components whose status ranges from prototype to production. The ESMF team is currently working to help transition more components into production codes. The highest development priority for ESMF is creating data structures that can handle a wider variety of grids and making the software both easier to use and more robust. ESMF is working with other projects, such as the Space Weather Modeling Framework, the Center for Integrated Space Weather Modeling, and the PRISM European climate modeling infrastructure, to allow components from one framework to be used in others.

Forrest Hoffman discussed various aspects of biogeochemistry related software development efforts as well as simulations. He gave updates on three areas: the C4MIP-CASA' experiments, the SciDAC deliverable (a nine-year fully coupled BGC simulation), and the Terrestrial Biogeochemistry Intercomparison Project (now called C-LAMP). He summarized the C4MIP experimental protocol, where the primary objective is to examine the simulations of the 20^{th} century atmospheric CO_2 and the CO_2 fluxes at the land surface in the coupled atmosphere-terrestrial biosphere models and showed preliminary results of the CASA' component of the C4MIP experiments. He discussed the carbon pools that had been spun up and showed graphics describing the carbon pool turnover times from CASA'. He also gave a brief overview of the fully coupled climate/biogeochemical model configuration based on CCSM that was recently assembled by the DOE SciDAC Climate Consortium Project in collaboration with NCAR and run for nine model years. Finally, he gave an overview of the C-LAMP project. The goal of this project will be to provide an intercomparison of terrestrial biogeochemistry models running in the CCSM3 framework and is being organized by the CCSM Biogeochemistry Working Group. He described the three models that will be examined (CLM3-CASA', CLM3-CN, LSX-IBIS), the roles of participating institutions, the necessary BGC diagnostics, and the testing strategy for porting the BGC version of CCSM to the Cray X1E at Oak Ridge National Laboratory, where the simulations will be carried out.

Yu-Heng Tseng (LBNL-SciDAC) discussed the implementation of parallel netCDF in stand-alone CAM. To facilitate efficient and flexible input/output (I/O) in CAM, the Parallel netCDF library was combined with the ZioLib algorithm. This procedure remaps distributed arrays into a Z-decomposition on a subset of processors, and then writes to a disk file in parallel to obtain the maximum parallel performance. For a 1.1GB standard output field of CAM3.1 D- resolution run, the current procedure can speed up history I/O by a factor of over 13 on an IBM SP (Seaborg). The speed up depends on the machines and is highly related to the local memory. The next goal is to begin implementing this library in POP, and at the same time, bring this code base into the CCSM Subversion repository and utilize it as part of the standard CAM development code base.

Helen He (LBNL-SciDAC) summarized the status of work on the single-executable concurrent CCSM. Work is being done on both the development and release versions of CCSM. The single-executable CCSM is developed by redesigning the top level code structures. It maintains the coexistence of the original multiple executable and new single executable and ensures minimal impact to users with modified build scripts. The new CCSM single-executable improves the way of setting number of processors and threads for each model component from static to dynamic. One of the technical difficulties is the name conflicts originating from same name modules and subroutines, etc., used in different components within the single executable code. We proposed a module-based approach, using wrapper module to localize global symbols, to resolve this issue. The new method is based on the Multi-Program Handshaking (MPH) library.

Pat Worley (ORNL-SciDAC) gave a report on software engineering activities during the final year of the SciDAC project. These include the addition of interactive carbon and sulfur cycles to CCSM, the implementation of several ocean ecosystem trace gases through the coupler, the porting of CAM and subsequent performance optimization on the Cray X1E and XT3, the explicit typing of all variables and constants in CAM and CLM, the release of an improved MCT, load balancing and MPI/OpenMP hybrid testing in support of BGC production runs, the evaluation of parallel netCDF on ORNL systems, the analysis of scalability of FVCAM with respect to processor count, horizontal resolution, and tracer count, and, finally, working with CSEG to implement a single-executable version of CCSM

Mathew Rothstein (NCAR-CSEG) presented an overview of the migration of the CCSM revision control system from CVS to Subversion. CSEG migrated its CCSM code repository from CVS to Subversion in the first week of January 2006. This presentation summarized some of the major design decisions made along the way and then highlighted some of the important features introduced by Subversion and how the CCSM repository is taking advantage of them.

Erik Kluzek (NCAR-CSEG) provided an update of the ESMF Stage 1 Evaluation Plan. He provided an introduction to the new project pages and summarized the progress that has been made since June of modularizing the CAM surface models migrating their calls to top level of CAM. He outlined the goals of the sequential CCSM for use with CAM. He summarized the current work of creating a top level application driver that coordinates simulation time evolution for cam and the CAM surface components via the introduction of new application driver namelists and a top level synchronization clock and couples between CAM and its surface components (currently this only implies redistribution) via the introduction of top level couplers. The current design must at all steps support a flexible sequential coupling strategy that ensures future plug and play of active and data components, which permits alternative coupling frameworks to be examined and that is backwards compatible with all current stand-alone CAM and concurrent CCSM-CAM functionality.

Brian Kauffman (NCAR-CSEG) provided an update on the CCMS data model (Version 7) rewrite project. The basic goals of this effort are to rewrite the Version 6 code base to consolidate similar functionality into share code where possible, provide a uniform set of basic functionality, make the code base compatible with both MPMD and SPMD modes of running the models as well as different coupling frameworks (such as ESMF and MCT), and extend scientific functionality. The resulting code base will provide new functionality for both grid mapping and time series and time interpolation generality. A new atmospheric data model, datm7, has already been created that combines the functionality of both latm6 and datm6. Similarly, a new ocean data model, docn7, has been implemented and contains both prescribed data as well as slab ocean model functionality. Various new use cases for both datm7 and docn7 were discussed. Share code has been created to leverage as much code reuse as possible.

Brian Eaton (NCAR-CSEG) summarized CAM development efforts and milestones. Several new chemistry options have been added to CAM. They include a tropospheric MOZART chemistry package and both a simplified greenhouse gas chemistry and a full MOZART chemistry for the WACCM configuration of CAM. Options have been added to support carbon cycle experiments and new versions of the CLM have been added. A significantly expanded regression test set has also been implemented that currently includes over 150 full system tests that are run on three platforms before committing code to CAM's development trunk. Automated testing expands the test coverage on three additional platforms after new code has been committed to the trunk.

Significant progress has also been made in rewriting the coupling to the surface models in stand-alone CAM as part of the ESMF Stage 1 Evaluation Plan. CAM's surface models have been modularized and their interfaces are now called from the top level driver. A CAM component module has been created with initial, run, and final methods. The ESMF prototype code has been replaced by a Fortran90 implementation of the ESMF time manager interfaces taken from WRF. The CAM-CLM coupling layer has been replaced with an MCT based version. CAM no longer requires the use of compiler auto-promotion flags.

Finally, The CAM and GEOS5 versions of the FV dycore have been merged and integrated into CAM. Re-factoring work is ongoing to allow only arrays using the XY decomposition to be visible outside the dycore. Design work on generic dycore interfaces is just beginning. Work to eliminate CAM's implicit assumptions that the dycore works on a rectangular lat/lon grid has begun. The initial work has focused on removing this assumption from physics package code that does boundary data set interpolation.

Nancy Norton (NCAR-CSEG) summarized CCSM ocean model developments, including a brief mention of the HYCOM/MICOM coupling efforts, a more in-depth review of the POP2 integration project, and a summary of the new scientific features that are scheduled to be added to the CCSM POP2 model.

Attendees:

Patrick Worley, ORNL
Mat Rothstein, NCAR
Mariana Vertenstein, NCAR
Nancy Collins, NCAR
Lawrence Buja, NCAR
Gary Strand, NCAR
Erik Kluzek, NCAR
David Bailey, NCAR
Juli Rew, NCAR
Forest Hoffman, DOE
Brain Eaton, NCAR
Nancy Norton, NCAR
Jeff Lee, NCAR
Brian Kauffman, NCAR

Kevin Raeder, NCAR
Peter Gent, NCAR
Cecelia DeLuca, NCAR
Yu-Heng Tseng, LBNL
Helen He, LBNL