E3SM-v1.0 developments and simulations

Coupled model simulation plans

Low-resolution water cycle CMIP6 DECK experiments (expected completion April 1, 2018)

- Pre-industrial control simulation (a few hundreds years)
- Historical simulation 1850-2015 (multiple ensemble members)
- Abrupt 4xCO_2 simulation (150 years)
- 1% per year CO_2 increase up to doubling (150 years)
- AMIP simulation.

High-resolution water cycle experiment (expected completion April 1, 2018)

- Perpetual 1950 control simulation (minimum of a few decades).

BGC coupled simulation (low resolution) (expected completion date?)

Ocean-cryosphere simulation (variable resolution) (expected completion date?)

Atmosphere v1

1. A relatively high vertical resolution (72 layer) has been chosen with a top at about 60km to better resolve
   a. cloud features (water path, cloud thickness and altitude, cloud frequency of occurrence),
   b. aerosol plumes,
   c. vertical distribution of water vapor,
   d. dynamical variability in the upper troposphere
2. A one degree (ne30) horizontal resolution is used for model development and long simulations for BGC studies.
3. A quarter degree (ne120) horizontal resolution is used for the water-cycle and sea-level-rise science questions to better resolve dynamical features, precipitation intensity, climate variability
4. Many improvements were made to the aerosol and cloud microphysical treatments to provide more accurate representation, and improved consistency in aerosol cloud interactions. These improvements are important for simulating cloud feedbacks and the aerosol indirect effect.
   a. MAM3 to MAM4
   b. Revisions to convective transport, aerosol nucleation, and scavenging to transport more aerosols to high latitude
   c. Aerosol resuspension and evaporating aerosols added to the coarse model
   d. Numerics of Aerosol Nucleation
   e. Sea Spray Aerosol now contains an organic component
   f. Nucleation
   g. Conversion to precipitation
   h. Aerosols
   i. Ice Microphysics
   j. Cloud Microphysics - Morrison Gettelman version 2 (MG2)
5. The Shallow convection and turbulence parameterizations have been replaced with a unified treatment (CLUBB)
6. Improvements were made to the in-situ diagnostics for clouds (COSP) and aerosols (Aerosol Lidar Simulator) to facilitate better comparison with remote sensing products.
7. An efficient linearized ozone treatment was added to allow a consistent treatment of ozone and atmospheric dynamics in the upper troposphere and lower stratosphere. It will also provide variability of the ozone hole that is important for changes in the southern annual mode surface winds.
8. Added regional refinement capability to allow less expensive model development and evaluation over regions of interest.
9. Introduced a short-range hindcast capability (CAPT) to allow less expensive high-resolution tuning
10. Created a consistent set of simulation compsets that use updated datasets (CMIP5-era, and newer).

Land v1

1. A vertically-resolved reactive transport solver based on BeTR (Tang et al. 2013) that has the following capabilities:
   a. multi-phase (gaseous, aqueous, sorbed, etc.)
   b. multi-species (C, N, P, etc.)
   c. multi-reaction (decomposition, mineral surface interactions, etc.)
2. A variably saturated head-based hydrology model using PeTSC solvers. Global and site-level simulations have been completed and publications documenting the capability are underway.

3. Two different representations of the coupled C, N, and P cycles. Both approaches have global and site-level simulations completed and publications are being prepared to document the capabilities. Both approaches use the same global soil phosphorus datasets (Yang and Post 2011; Yang et al. 2013)
    a. One based on the Equilibrium Chemistry Approximation (Tang et al. 2013; Tang and Riley 2015; Zhu et al. 2015a, b) to resolve nutrient competition
    b. One based on Xiaojuan Yang's CNP model (Yang et al. 2014, Yang et al. 2016).
4. Crop model improvements, including a dynamic rooting depth distribution.
5. A river transport model (MOSART) (Li et al. 2013; 2015) to replace the current RTM framework
6. Generic interfaces implemented for coupling between vegetation and biogeochemistry codes, and for coupling between soil physics and biogeochemistry codes.
7. Three capabilities to perform sensitivity, functional unit, and benchmarking analyses have been integrated:
    a. A land model Bayesian UQ framework (Sargsyan et al. 2014, Safta et al. 2015)
    b. A land model benchmarking framework based on the ILAMB package and leveraging the BGC-Climate Feedbacks SFA work
    c. A functional unit testing framework for individual subroutines (Wang et al. 2014)
8. Improvements to the land biogeochemistry spinup procedures

Ocean, Sea Ice, and Land Ice v1

For V1, entirely new ocean, sea ice, and land ice model components have been introduced under the *Model for Prediction Across Scales* (MPAS) framework. The single biggest change relative to previous DOE ocean and ice models is the use of fully unstructured, variable resolution meshes based on centroidal Veronoi tessellations (CVTs). In V1, the ocean and sea ice models are fully coupled to E3SM and the ocean model is mass conserving. The land ice model is currently static when running the coupled model, although ocean circulation within static, sub-ice shelf cavities is supported (for Antarctica only). Other relevant models details are as follows:

1. MPAS-Ocean uses an Arbitrary Lagrangian Eulerian (ALE) vertical coordinate with monotone tracer advection and split-explicit time stepping.
2. MPAS-CICE is based largely on CICE4 and CICE5 physics ported to the MPAS modeling framework.
3. MPAS-Land Ice includes a new, finite element-based momentum balance solver for the 3d, 1st-order accurate Stokes approximation. This solver is built using the externally linked, *Albany* and *Trilinos* libraries, which provide access to advanced solution methods and tools. Conservation of mass and energy are based largely on finite volume dynamics and physics subroutines ported from the Community Ice Sheet Model (CISM) version 2.0.

Performance v1

1. An OpenACC version of atmospheric dynamics for use on Titan GPUs.
2. Hybrid MPI with OpenMP threading in all components.
3. Nested threading in the atmosphere for increased parallelism.
4. Load balancing in all components, especially new load balanced decompositions for the ocean/ice components.
5. Vectorization and many other local optimizations for this new model.

Software Engineering v1

1. CIME5 and its new python-based system was adopted for configuring, building, and testing the model. See https://github.com/ESMCI/cime/wiki
2. A gitworkflows-based development strategy and github, with "master", "next" ,and maintenance-integration branches are maintained by a small group of integrators. See https://www.kernel.org/pub/software/scm/git/docs/gitworkflows.html
3. The "master" branch is always-ready-to-release. Tests should never be broken on master.
4. Developers do their own testing on branches using the "acme_developer" test suite and branches should pass those tests before issuing a "Pull Request" on github for integration of their changes. When adding a new feature, a test of that feature should be included.
5. An "acme_integration" test suite runs nightly on several platforms, including target LCFS, using a Jenkins server. Results are reported automatically to a dashboard on cdash.org.
6. The tests are currently system tests that require nearly the full source code to compile/run.
7. We are finishing a transition to PIO2 for our scalable I/O.
8. The Software Engineering team has developed several modes of communication for doing our jobs:
    a. github issues for questions directly related to source code
    b. github pull requests for merging new code
    c. Confluence for archived conversations, decisions, roadmaps, and meeting notes
    d. Slack chat service for daily, informal, non-archived interactions

Workflow v1

1. Data Management
    a. ESGF: A multi-institutional effort to securely access, monitor, catalog, transport, and distribute petabytes of data for E3SM

research experiments and observations. After more than six months of downtime, LLNL and the other ESGF sites redeployed with the updated v2.X software stack in order to resume its role in serving E3SM, CMIP and various other data products. ESGF now uses the CoG frontend and gives users listings of other frontend nodes and their respective data projects which are available for search.

   b. Data publication: A data publication tool, with online as well as command-line interfaces, has been developed, tested, and deployed to support the publication of model data to the ESGF nodes at ORNL/CADES and LLNL/AIMS facilities. Version 3.1.0 of esg-publisher also was released. This version includes a complete rewrite of the utility to create publisher mapfiles (the driver of the publication process) and QC features, in support of upcoming E3SM V1 data.

2. Diagnostics, Analysis and Metrics
   a. UV-CDAT: A visualization and analysis platform was implemented that is extensible and customizable for high-performance interactive and batch visualization and analysis for E3SM. The diagnostics output can be seamlessly integrated, and the software is now easy to deploy and manage via the Anaconda package management system. This system, which has been deployed and managed on a variety of platforms across multiple sites, is being routinely used by the E3SM science teams.
   b. NCO: The enhancements to NCO vastly speed-up and improve the pre- and post-processing of simulation output. The NCO operators meet all applicable E3SM specifications, are well-documented, and are used daily to analyze CAM-SE, ALM, and MPAS-O/I simulations in support of v1 model development and analysis. Researcher feedback also has been mostly positive. Repeated use at multiple sites (ANL, LANL, LLNL, NERSC, ORNL) with various configurations (serial, intra-node and inter-node parallel, interactive, and batch) has improved the robustness and versatility of NCO tools.
   c. E3SM Diagnostics (UVCmetrics): As part of the underpinning workflow environment, a diagnostics, model metrics, and intercomparison Python framework known as UVCMetrics was developed. This tool employs parallel processing to rapidly evaluate E2SM simulations, providing fast output and customizable workflow diagnostics. UVCmetrics also expedites access to any of the hundreds of files produced by the diagnostics suite.
   d. Enhancements for coupled model development and validation: Interaction with the coupled simulation group resulted in development of a more flexible tool to quickly diagnose model runs and produce publication-quality plots for tracking model progress. The needs of the coupled model group were also addressed by customizing the metadiags package.
   e. Viewer: The E3SM model diagnostics suites generate huge amounts of output, making it difficult to analyze specific data. The viewer dramatically improves workflow by allowing access to model diagnostics suites without the need to rely on file names or post-hoc analysis of the file system.
3. Rule Engine: A rules-based expert system, the E3SM Rule Engine, is now under development. This system will draw on the experience of the coupled model team to eliminate combinations of parameters that generate run-time errors and inefficiencies. The Rule Engine will prevent invalid runs that aren't detectable at compile time, thus saving valuable facility cycles.
4. Provenance Capture: A provenance database to capture all workflow component-execution details is being deployed.
5. E3SM Dashboard/UI: The dashboard is a unified user interface that will provide the ability to configure, run, monitor, and analyze the E3SM model by using a browser, thus accelerating model development and analysis.
6. Request Hub: A "Request Hub" application has been created within Confluence in order to quickly respond to requests across various E3SM science and development teams. It has been extensively used by the workflow team to help resolve issues with diagnostics and metadiags. This is a continuing effort to help groups better communicate with one another.
7. Process flow
   - Workflow Integration
      a. The role of the E3SM Workflow Team is to develop, deploy, and integrate essential software systems and utilities across all major computational facilities. The E3SM process flow provides the infrastructure utilities that are necessary for the seamless integration of E3SM workflow components across multiple production sites. This facilitates simulation experiments by providing the necessary input data and archival of model output data, as well as by transferring data and analysis products.In addition, the E3SM process flow includes the publishing and sharing of output data that can also be incrementally regridded, reduced, or analyzed.
      b. The E3SM process flow allows the E3SM science teams to be more productive by eliminating the need to attend to detailed logistics of data management.
8. Resource Management
   a. The E3SM project has utilized multiple hardware and software resources and tools across multiple facilities, including NERSC, ALCF, OLCF, ORNL/CADES and LLNL/AIMS.
   b. The resource management has developed, deployed, and maintained the software stack that is essential for the productivity of E3SM science teams.