Accelerated Climate Modeling For Energy

Project Strategy and Initial Implementation Plan

Current Revision: July 11, 2014

ACME Council Members
David Bader
William Collins
Robert Jacob
Philip Jones
Philip Rasch
Mark Taylor
Peter Thornton
Dean Williams

Preface

Following peer review of the Accelerated Climate Modeling for Energy (ACME) project, program sponsors at the Department of Energy's (DOE) Office of Biological and Environmental Research tasked the project leadership to "Produce a succinct summary (25 pages or less) of the project that articulates a strategy for accountably achieving a very small number of high-impact scientific deliverables enabled by computing at extreme scale – first in the short term of 3 years and also in the longer term of 10 years; include prioritization and sequencing, and include both climate-science and technical (computational and workflow) elements."

The project leadership accepted this assignment as an opportunity to concisely articulate the project strategy, including the roadmap and initial implementation plan. The present document, which admittedly exceeds 25 pages, should be viewed as a supplement to the original proposal. In that context, this document will not repeat information contained in the proposal, but will often refer to it. ACME is an evolving project within a constantly changing environment. The project strategy and plan will be updated periodically in response to and in anticipation of these changes, with the first update occurring in approximately six months.

Contents

1.0 Introduction

The Accelerated Climate Modeling for Energy (ACME) project is a newly launched project sponsored by the Earth System Modeling program within DOE's Office of Biological and Environmental Research. ACME is an unprecedented collaboration among eight national laboratories, the National Center for Atmospheric Research, four academic institutions, and one private-sector company to develop and apply the most complete, leading-edge climate and Earth system models to the most challenging and demanding climate-change research imperatives. It is the only major national modeling project designed to address U.S. Department of Energy (DOE) mission needs and efficiently utilize DOE Leadership Computing resources now and in the future. While the project's capabilities will address the critical science questions articulated in this plan, its modeling system and related capabilities also can be flexibly applied by the DOE research community to address mission-specific climate change applications, such as those identified in the report, *U.S. Energy Sector Vulnerabilities to Climate Change and Extreme Weather* (http://energy.gov/downloads/us-energy-sector-vulnerabilities-climate-change-and-extreme-weather).

The remainder of this section provides an overview of the ACME project strategy. Section 2 articulates our science priorities and the short- and long-range plans to address them. Section 3 provides more detail on the computational research challenges the project faces as we move Earth system modeling onto the new disruptive architectures. Our scientific metrics of model performance are described in Section 4. Section 5 describes the initial construction of the software engineering and workflow infrastructure needed to support the ACME project. Finally, Section 6 addresses the changes in the ACME management structure and function that resulted from the peer review.

1.1 The ACME Vision

The Accelerated Climate Modeling for Energy Project is an ongoing, state-of-the-science Earth system modeling, simulation, and prediction project that optimizes the use of DOE laboratory resources to meet the science needs of the nation and the mission needs of DOE.

In this context, "laboratory resources" include the people, programs, and facilities, current and future. They collectively represent a unique combination of scientific and engineering expertise as well as leadership computing and information technologies required to construct, maintain, and advance an Earth system modeling capability that is needed by the country and DOE. A major motivation for the ACME project is the coming paradigm shift in computing architectures and their related programming models as capability moves into the exascale era. DOE, through its science programs and early adoption of new computing architectures, traditionally leads many scientific communities, including climate and Earth system simulation, through these disruptive changes in computing.

1.2 The ACME Ten-Year Goal

Over the next 10 years, the ACME project will assert and maintain an international scientific leadership position in the development of Earth system and climate models at the leading edge of scientific knowledge and computational capabilities. With its collaborators, it will demonstrate its leadership by using these models to achieve the goal of designing, executing, and analyzing climate and Earth system simulations that address the most critical scientific questions for the nation and DOE.

1.2.1 The ACME Ten-year Roadmap, Initial Priorities, and the ACME Project Learning Curve

ACME will achieve this goal through four intersecting project elements:

1. a series of **prediction and simulation experiments** addressing scientific questions and mission needs;
2. a well documented and tested, continuously advancing, evolving, and improving **system of model codes that comprise the ACME Earth system model**;
3. the ability to use effectively **leading (and "bleeding") edge computational facilities** soon after their deployment at DOE national laboratories; and
4. **an infrastructure** to support code development, hypothesis testing, simulation execution, and analysis of results.

Figure 1 depicts the ACME Project Roadmap, showing the relationships among the first three major project elements: the simulations, the modeling system to perform those simulations, and the machines on which they will be executed. Unlike the other three elements that have distinct but overlapping phases, the fourth element, the infrastructure, will evolve continuously based on the requirements imposed by project needs.

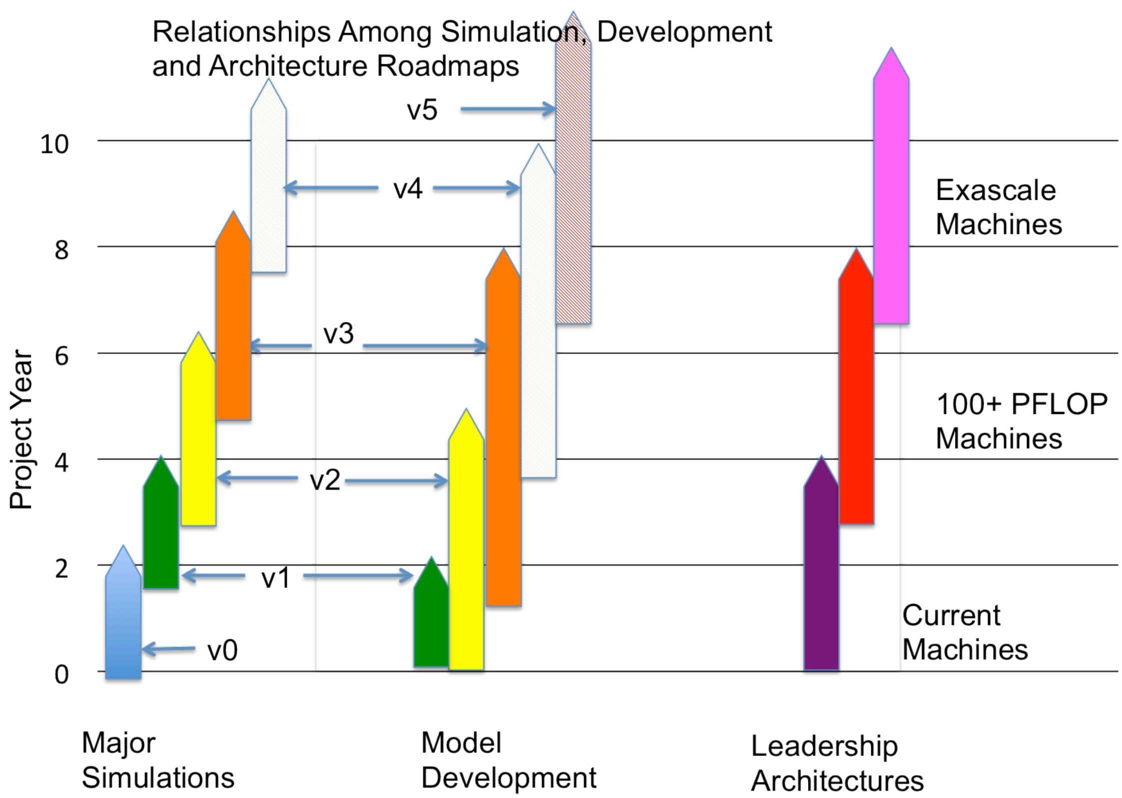

Figure 1. *The ACME Project Roadmap, showing the relative sequencing of major simulation campaigns, model version development, and machine deployment.*

1.2.1.1 Element 1: Major Simulations

Reference and benchmark experiments with model version v0, described in section 5.1.1.1 of the proposal, serve as the starting point for the ACME project. Mission and science needs defined our three grand challenge science questions and the simulations envisioned to answer those questions in 10 years. The grand challenge simulations are not yet possible with current model and computing capabilities. Nevertheless, we developed a set of achievable experiments that make major advances toward answering the grand challenge questions using a modeling system, v1, which we can construct to run on leadership computing architectures available to the project over the next three years. Like all research projects, early results will be used to refine science questions and develop new testable hypotheses to be addressed with subsequent versions of the modeling system. As shown on the left side of Figure 1, ACME envisions simulation campaigns of three to four years each with successive versions of the modeling system. Every campaign will inform the next, and we envision four to five campaigns with successive versions of the modeling system leading to the grand challenge simulations in approximately 10 years.

1.2.1.2 Element 2: Model Development

The core of the ACME project is model development. This element connects the science needs with computing power provided by the DOE Office of Science (DOE-SC). ACME is the only Earth system modeling project to target DOE Leadership Computing Facilities (LCFs) as its primary computing architectures. The ACME model development path currently envisions five development cycles for its modeling system over the next 10 years, although only three will be both started and completed over the time span. The staged nature of model development is depicted in the center of Figure 1, which recognizes that multiple versions are undergoing different stages of development and testing at any given time. The project is committed to fully documenting and testing each major model release, including benchmarking runs for performance and control simulations for scientific evaluation. Many pieces of version v1, which will be used for the initial experiment campaign, are complete, but the full system still requires further integration and testing. Developments estimated to take longer to complete than is possible for the v1 simulation campaign serve as the starting point for v2, the first full development cycle under the ACME project. This version will bridge the current LCF architectures with the next generation, 100 × 1015 (or greater) floating-point operations per second (PFLOP) systems just now being configured. Scientifically, version v3 initial requirements will be informed by the completion of the v0 experimental campaign as well as the early v1 simulations.

Because the ACME project requires additional flexibility to adapt to an uncertain and disruptive computing technology roadmap, experience with v2 on the next-generation machine will be a major factor in defining the computational science and performance requirements for v3. A similar strategy will be employed for v4 and v5 as the project gets closer to its goal of building models to answer the ten-year science questions on the LCF exascale machines. Depending on progress, the project may consolidate development for versions v3 and v4, and move up the schedule for v5.

1.2.1.3 Element 3: Leadership Architectures

We cannot over-emphasize the interdependence of the ACME project and DOE LCF resources. The project envisions that versions v2 and beyond will include elements of the co-design process developed by DOE. Co-design refers to a computer system design process where scientific problem requirements influence architecture design and technology constraints inform formulation and design of algorithms and software *(http://science.energy.gov/ascr/research/scidac/co-design/).* While the ACME project is not a co-design center, it aims to have an impact on the facility roadmap through a close partnership with LCF scientists and engineers. The architecture roadmap depicted on the right side of Figure 1 was produced using information in the report, *APPROACHING EXASCALE: Application Requirements for OLCF Leadership (www.olcf.ornl.gov/wp-content/uploads/2013/01/OLCF_Requirements_TM_2013_Final.pdf).*

There is considerable uncertainty about the capabilities of the next two generations of machines in their design and support, including tools, libraries, compilers, and even programming models. The three-laboratory "Collaboration of Oak Ridge, Argonne and Livermore (CORAL)" is currently negotiating with potential vendors on the design, cost, and delivery of the 100+ PFLOP LCF systems to be delivered in 2017, and more will not be known until those discussions are complete. Achieving grand challenge science goals necessitates that the ACME project adapt to the unforeseen changes in the computing landscape.

1.2.1.4 Element 4: ACME Infrastructure

The priority science drivers and resulting three-year experiments were used to define the functionality of the initial simulation system. Initial infrastructure design was based on the requirements to facilitate hypothesis-testing workflows (configuration, simulation, diagnostics, analysis). As mentioned above, the infrastructure element is continuously evolving and does not have the distinct phasing of the other three elements. The infrastructure will maintain a disciplined software engineering structure and turnkey workflows to enable efficient code development, testing, simulation design, experiment execution, analysis of output, and distribution of results within and outside the project.

2.0 Science Priorities

This project is designed to accelerate the development and application of a fully coupled, state-of-the-science Earth system model (ESM) for scientific and energy mission applications. The model will be optimized for deployment on systems in the DOE's LCFs and the National Energy Research Scientific Computing center (NERSC). While the goal is to build a state-of-the-science Earth modeling system for the full complement of DOE's mission-related goals and needs, initial scientific development of the system will be dictated by the team's focus on three climate change thematic areas (called "science drivers" hereafter). These science drivers cover broad, important areas of science requiring an accurate treatment of climate system processes for more accurate simulation and prediction of climate change. The drivers align closely with the Climate and Environmental Science Division's (CESD) mission objectives to advance DOE's leadership in the computational formulation, empirical evaluation and calibration, and use-inspired application of climate models.

The three drivers (and each driver's summary term used hereafter) are:

1. **(Water Cycle)** How do the hydrological cycle and water resources interact with the climate system on local to global scales?
2. **(Biogeochemistry)** How do biogeochemical cycles interact with global climate change?
3. **(Cryosphere Systems)** How do rapid changes in cryospheric systems interact with the climate system?

Our full proposal included 10 top-level science questions motivated by these science drivers. We retain that full suite of science questions as guidance for our long-term (10-year) model development strategy, but have undertaken an additional level of prioritization in response to review, focusing our near-term (three-year) model development on a single main question under each science driver. The three resulting questions are characterized by their centrality to major open issues in climate science, their relevance to DOE's mission objectives, and their potential to advance the state of Earth system simulation using ACME capabilities.

A pair of near-term and long-term experiments accompanies each question. These experiments represent simulation targets for v1 and v5 of the ACME system planned for a three-year term and a 10-year horizon. These time intervals correspond to the first period of performance and to a longer, decadal-length vision for this Scientific Focus Area. The first, near-term member of each pair is designed to be both a "stretch" goal for the initial round of ACME development and a preparatory first step toward a grand challenge objective suitable for the extreme-scale computing resources anticipated for 2020 and beyond. The second, long-term member of each pair is designed to be an experiment that revolutionizes our predictions connected with each driver and that utilizes the full power of these extreme systems leveraged with transformative advances in model physics, numerical formulation, analytics, and computational implementation.

2.1 Version 1 Model Configurations

The ACME project will target at least two base model configurations. The primary configuration is a global coupled high-resolution version that can provide sufficient detail to address science questions and can further provide improved decision-relevant simulation results for DOE mission needs. This configuration will also serve to produce benchmark and control simulations to evaluate model scientific and computational performance. A second configuration will utilize standard climate model grid resolution with refined higher resolution over regions of interest for better representation of physical and dynamical process. This coarse resolution version will provide more rapid turnaround for development purposes while also testing the regional focusing capability to address specific science questions.

2.1.1 Global Uniform High-resolution Configuration

A high-resolution model must be able to meet the needs of the scientific experiments proposed while also fitting within the affordable computational performance constraint. Our v1 target configuration will include an atmosphere model at ¼-degree (~25-km) resolution coupled to an eddy-resolving ocean model. The atmosphere component will utilize the spectral-element dycore on a cubed-sphere mesh with nominal 25-km resolution. The atmospheric physics will be based on the version 5 physics of the Community Atmosphere Model (CAM) with the additional new ACME developments included. The land model will be based on the Community Land Model (CLM) with additional ACME improvements described in the proposal. The ocean component will be the Model Prediction Across Scales (MPAS)-Ocean model using a grid with 15-km resolution in equatorial regions, decreasing to 5 km in polar regions to capture the reduction in eddy size with decreasing Rossby radius of deformation. The ocean grid will extend southward into Antarctic embayments to address the Antarctic ice sheet experiments and, if affordable, will attempt to utilize even higher resolution within those embayments. The sea ice component will be the new MPAS-Community Ice Code (CICE) model on the same ocean grid. An ice sheet component will be added for the Antarctic experiments that spans the Antarctic continent with a variable-resolution grid that extends to 500-m grid spacing in regions surrounding current and likely future positions of the grounding line.

A question asked by the reviewers is the timing of the addition of different components. The highest priority improvements were identified in Section 5 of the proposal as version v0.1, and are repeated here:

- CAM will upgraded to include the cloud microphysics and macrophysics from the Biological and Environmental Research (BER) "Polar" project; additional infrastructure to separate physics from dynamics (solver for the Euler equations) to enhance the extensibility and flexibility of CAM; and provisions for coupling to the land ice sheets via the surface mass balance linked to treatments of sub-grid orographic variability.
- CLM 4.0 will be replaced with CLM 4.5, including the full suite of new biochemistry and Variable Infiltration Capacity (VIC) hydrology developed under the IMPACTS project, and is further enhanced with dynamic land units and extensions to couple to land ice sheets.
- CICE is replaced by CICE 5, including iceberg parameterizations.
- Parallel Ocean Program (POP) is replaced by MPAS-Ocean.
- The Community Ice Sheet Model (CISM) version 2.0 is elevated to the status of core component of the Earth System Model.

The development of the paramaterizations and new components is complete, so the primary effort is on the replacement of the ocean component and the addition of CISM.

2.1.2 Variable-resolution Configuration

The high-resolution configuration described above will not be useful for routine model development and testing, therefore a coarse-resolution version must be supported. However, we can utilize the new variable-resolution capabilities to test high-resolution formulations in limited regions. For the atmosphere, a configuration using the spectral-element dycore on a 1-degree (~100-km) cubed-sphere mesh is desired. However, to test new water cycle improvements and subgrid orography schemes, additional focused resolution (as high as 10 km) near regions with heterogeneity (e.g., high topography) is desired. Similarly, for model validation exercises, focused domains near ARM sites or other field experiments would be useful.

In the ocean, we plan to support very-low-resolution configurations (grids of 120-km resolution with 60-km resolution around the equator needed for effective representation of equatorial current systems and variability). We also plan an eddy-permitting configuration using a mesh with 30-km resolution near the equator, with reductions to 10-km resolution in polar regions. This is likely to serve as a better development platform than the very-low-resolution version. For the Antarctic experiments, we would further enhance resolution in the Southern Ocean and Antarctic embayments, the details of which will be determined based on computational affordability. The ice sheet model described in the previous section may still be affordable within the overall coupled system.

2.2 Proposed Near-term and Long-term Simulations

Table 1 defines initial and long-term experiments that represent major advancements in predictive climate science. The design and implementation of these experiments will require a true systems integration of ACME, its computational environment, its analytical workflow, and its rigorous diagnostics against measurements of the whole Earth system.

Table 1. Summary of primary science driver questions, as well as the questions to be answered by near- and long-term experiments.

Questions	Near-term (3-yr) experiments	Long-term (10-yr) experiments
Water Cycle		
What are the processes and factors governing precipitation and the water cycle today, and how will precipitation evolve over the next 40 years?	How will more realistic portrayals of features important to the water cycle (resolution, clouds, aerosols, snowpack, river routing, land use) affect river flow and associated freshwater supplies at the watershed scale?	How will the integrated water cycle, ranging from bedrock to the tropopause, evolve in a warmer climate with changes to land and water use, and changing forcing agents (aerosols, greenhouse gases)?
Biogeochemistry		
What are the contributions and feedbacks from natural and managed systems to current greenhouse gas fluxes, and how will those factors and associated fluxes evolve in the future?	How do carbon, nitrogen, and phosphorus cycles regulate climate system feedbacks, and how sensitive are these feedbacks to model structural uncertainty?	How will coupled terrestrial and coastal ecosystems drive natural sources and sinks of carbon dioxide and methane in a warmer environment?
Cryosphere Systems		
What will be the long-term, committed Antarctic Ice Sheet contribution to sea level rise (SLR) from climate change during 1970–2050?	Could a dynamical instability in the Antarctic Ice Sheet be triggered within the next 40 years?	How will regional variations in sea level rise interact with more extreme storms to enhance the coastal impacts of SLR?

2.2.1 Experiments at the Three-year Horizon

2.2.1.1 Water Cycle

How will more realistic portrayals of features important to the water cycle (resolution, clouds, aerosols, snowpack, river routing, land use) affect river flow and associated freshwater supplies at the watershed scale?

We hypothesize that changes in river flow over the last 40 years have been dominated primarily by land management, water management, and climate change associated with aerosol forcing. During the next 40 years, greenhouse gas (GHG) emissions following a Representative Concentration Pathways (RCP) 4.5 or 8.5 scenario will produce changes to river flow with signatures that dominate those other forcing agents in at least one of the domains our experimental framework examines below.

The goal is to simulate the changes in the hydrological cycle, with a specific focus on precipitation and surface water in orographically complex regions such as the western United States and the headwaters of the Amazon. The experiment will exploit the combination of 10-km resolution atmosphere with a similarly high-resolution land surface that is further enhanced with sub-grid orographic parameterizations in the v1 configuration. The combination of high land-surface resolution with the sub-grid orographic treatment will enable much more realistic representations of upslope and orographically forced precipitation, high-altitude snowpacks, rainshadows on the downwind slopes of mountain ranges, and other interactions of precipitation and the landscape critical for the target river systems. Improved resolution, and improved parameterizations of clouds, aerosols, and their interactions, should produce a more realistic portrayal of the precipitation location, frequency and intensity, and aerosol deposits on snow and surface ice—all factors that influence runoff, snowpack, and snowmelt.

In this study, we will explore the role of various physical processes and their treatment in climate models in influencing river flow and fresh water supply, with a goal of producing an accurate portrayal of present-day river flow for major river basins on the planet. The Mackenzie and the Mississippi in North America, the Amazon and La Plata in South America, and the Ganges and Yangtze in Asia, are archetypes of major river basins dominated by very different climate and hydrologic regimes. Sea-surface temperature (SST) is a primary driver of regional precipitation patterns and, thus, the regional water cycle. As a result, the fidelity of the simulated regional water cycle will likely depend strongly on fidelity of near- and far-field simulated SST.

We will systematically explore the sensitivity of atmospheric and surface water budgets simulated in these river basins to ACME model improvements in resolution and the treatments of clouds, aerosols, subgrid orographic effects, and surface/subsurface hydrology. We will explore how water availability in the major river basins responds to anthropogenic forcing, including emissions, land use, and water use. A preliminary simulation plan is as follows:

1. Using prescribed SSTs from the last 40 years, the effects of atmospheric and land-surface resolution can be explored by comparing the global 1° resolution version of the model to companion experiments with the global 0.25° (or possibly 0.125°) atmospheric resolution/1-km land resolution model. We will also compare the global high-resolution results with (a) coarse resolution simulations with regional refinement over the river basins only, and (b) versions with and without the new sub-grid orography scheme.
2. For the same 40-year period, we will run fully coupled atmosphere–ocean–land–sea-ice simulations at both high and low resolutions to identify the feedback mechanisms prevalent

in the two different dynamical regimes and their impacts on systematic model biases, particularly in the SST fields.

3. Given the results of (2) above, we will continue the experiments for 40 additional years using the RCP 4.5 forcing scenario to test the hypothesis. This set of simulations is specifically identified in the letter requesting the ACME proposal.

2.2.1.2 Biogeochemistry

How do carbon, nitrogen, and phosphorus cycles regulate climate system feedbacks, and how sensitive are these feedbacks to model structural uncertainty?

This experiment has dual objectives: first, examining how more complete treatments of nutrient cycles affect carbon–climate system feedbacks, with a focus on tropical systems; and second, investigating the influence of alternative model structures for below-ground reaction networks on global-scale biogeochemistry–climate feedbacks. The focus on nutrients will center on the addition of phosphorus (P) to existing models of soil biogeochemistry, testing the hypothesis that P availability limits tropical ecosystem production and plays an important role in regulating global-scale feedbacks connecting CO2 concentration, temperature, and the hydrologic cycle. The P cycle is introduced on top of existing vertically resolved carbon-nitrogen (C-N) biogeochemistry within the CLM 4.5 framework. Our near-term experiments will be run using the P cycle in conjunction with two independently developed C-N reaction networks: the converging trophic cascade (CTC) model (Thornton et al., 2007) and the CENTURY model (Parton et al., 1987). These two reaction networks are accessible within CLM 4.5 via a single compile-time switch. By evaluating the introduction of P dynamics in both networks, we will assess the influence of model structural uncertainty on the sign and magnitude of biogeochemistry–climate feedbacks in the coupled global system. Since the success of the experiment does not critically depend on ultrahigh resolution, these simulations can commence using current well-tested, and considerably less expensive, model grids.

The structural uncertainty dimension of this experiment (testing alternative below-ground reaction networks) is expected to have important consequences across the globe. Of particular interest, as we look toward the much more mechanistically complete representation of biogeochemistry processes that characterizes our 10-year vision, is how these structural differences manifest in high-latitude systems. We will use the experiment described here to explore the influence of choice of C-N reaction network on thermal hydrology and carbon fluxes in Arctic tundra and other high-latitude systems.

A possible experimental protocol consists of four pairs of simulations, exercising the fully coupled climate and biogeochemistry system, using the low-resolution model configuration (Section 2.1) with modifications as described here. The two members of each pair are conducted with alternate representations of the below-ground reaction networks (CTC versus CENTURY), enabling a first-order resolution of the influence of model structural uncertainty. The four simulation pairs are as follows:

1. A pair of fixed-forcing control simulations, using preindustrial (circa 1850 AD) boundary conditions. Those boundary conditions include land cover, aerosols, nitrogen deposition over land and ocean, and atmospheric concentrations of non-CO_2 greenhouse gases. This pair of simulations uses prognostic land and ocean biogeochemistry as implemented in the CMIP6 protocols, including a dynamic nitrogen cycle, but not including phosphorus dynamics in the land model. Atmospheric CO_2 concentration is prognostic in this simulation, but is expected to remain quite stable, based on the results of a biogeochemistry spin-up procedure (described below). Duration of this experiment is a minimum of 250 model-years. If small

drifts in climate and/or CO_2 concentration remain after spin-up, this simulation will be used as a reference case for the removal of long-term residual trends from the following transient-forcing simulations.

2. A pair of transient-forcing control simulations, using historical forcings for the period 1850–2004, and a high radiative forcing scenario (RCP8.5, from the CMIP6 protocol) for the period 2005–2100. Transient forcings include fossil fuel and industrial emissions of CO_2, land use and land cover change, atmospheric nitrogen deposition over land and oceans, aerosols, and non-CO_2 greenhouse gas concentrations. This pair of simulations uses the same configurations of below-ground reaction networks as the fixed-forcing controls, with a dynamic nitrogen cycle on land but no representation of phosphorus dynamics.
3. A pair of fixed-forcing C-N-P simulations, configured exactly as the fixed-forcing controls, but replacing the standard carbon-nitrogen reaction networks with new carbon-nitrogen-phosphorus (C-N-P) networks, implemented in each of the reaction networks (CTC and CENTURY). This simulation requires a spun-up state, which is independent of the spin-up used for the fixed-forcing control. The spin-up procedures are discussed below.
4. A pair of transient-forcing C-N-P simulations, configured exactly as the transient-forcing control experiments, but replacing the C-N networks with their C-N-P equivalents.

The ACME team has at its disposal spun-up states for fully coupled climate–biogeochemistry simulations using C-N dynamics in the CTC reaction network. Since some model components are slated for rapid modification in the initial six months of the project, it is likely that additional model spin-up time will be required to meet typical steady-state metrics. Spun-up offline simulations (land driven by data atmosphere) are available for the C-N dynamics in the CENTURY reaction network, and we will commence a standard sequence of partially coupled simulations to equilibrate the biogeochemistry with this configuration. We expect these spin-up simulations to be ready within six months of project launch, allowing execution of the fixed-forcing control simulations in Q3/Q4 of the first project year.

Spun-up offline C-N-P simulations in the CTC reaction network are available today, and these will be used to initiate the fully coupled spin-up sequence immediately upon project launch. C-N-P integration within the CENTURY reaction network is not yet underway; this is the highest-priority model development task, slated for completion in Q2 of the first project year. Fixed-forcing and transient-forcing C-N-P simulations can then initiate in Year 1 Q3.

Our experimental results will be made more robust if we can carry out ensembles for the transient forcing experiments. We have set a target minimum of four ensemble members for each of these simulations. The fixed-forcing controls do not require ensemble members, but could be extended in length if long-period variation emerges.

2.2.1.3 Cryosphere System

Could a dynamical instability in the Antarctic Ice Sheet be triggered within the next 40 years?

The objective is to examine the near-term risk of initiating the dynamic instability and onset of the collapse of the Antarctic Ice Sheet due to rapid melting by warming waters adjacent to the ice sheet grounding lines. The experiment would be the first fully coupled global simulation to include dynamic ice shelf–ocean interactions for addressing the potential instability associated with grounding line dynamics in marine ice sheets around Antarctica. It will utilize several significant advances in the new ACME model, including the ability to enhance spatial resolution in both the ice sheet and ocean model to resolve grounding-line processes while still maintaining global extent in a coupled system and throughput for decadal simulations. The simulation will include an eddy-resolving Southern Ocean as well to better represent Circumpolar Deep Water (CDW) and dynamics associated with bringing this water onto the continental shelf under the ice sheet. Including the sea ice model captures the process of buttressing at the ice shelf–sea ice boundary. Finally, a fully coupled system is able to simulate changes in atmospheric forcing (e.g., poleward displacement of jets) that could influence the behavior of the Southern Ocean and sea ice.

The specific experiment will be a fully coupled simulation from 1970–2050 to explore whether rapid ice sheet instability is triggered in this time frame. An ensemble would be desirable to address the likelihood of such an event, though this is not likely to be affordable in our configuration in this timeframe. The model configuration for this experiment will be a modified version of the standard high-resolution ACME configuration described below. The base configuration includes the atmosphere/land on a 0.25° cubed-sphere grid using the ACME-modified CAM5-SE atmosphere model. The subgrid orography modifications will be needed to resolve Antarctic surface mass balance at the ice sheet margins. The ocean component will be MPAS-O on a Spherical Centroidal Voronoi Tesselations (SCVT) mesh with 15-km grid spacing at the equator, decreasing to 5 km in the Southern Ocean region. The default mesh will be extended southward to include critical Antarctic embayments and the resolution in these regions will be further enhanced if affordable. The vertical grid will be a hybrid coordinate with 100 vertical levels. The sea-ice component will be MPAS-CICE on the same ocean grid. Finally, we will add an Antarctic Ice Sheet model with resolution of 0.5–1 km near likely grounding-line locations and coarser resolution (~10 km) throughout the interior. For initial conditions, we will follow a similar spin-up procedure as with previous high-resolution simulations, with an ocean/ice state from an ocean/ice reanalysis-forced spin-up. For the ice sheet, an optimized initial condition should be available from the PISCEES project.

This first-of-its-kind coupled simulation will be focused largely on the ocean–ice shelf feedbacks and potential for dynamical instability and rapid SLR. It represents a first step toward a comprehensive SLR and impacts capability needed by the DOE to assess threats to coastal facilities. As work proceeds toward the more comprehensive experiments planned in the 10-year timeframe, we will be incrementally adding additional features. For example, work will begin under this project to develop an initial implementation of icebergs and primitive calving laws to capture the transport and distribution of ice and other material as the ice sheets flow into the ocean. Work also continues (as part of related projects) on a Greenland Ice Sheet model so that we can capture SLR contributions from both major ice sheets. We will also begin to include isostasy and ice-sheet self-gravity that can have a first-order effect on the regional SLR signature around the coastal U.S. We anticipate all of these effects to be included in a following ACME version. Further releases will begin to include wave models, further focusing of resolution in coastal and storm-track regions, and other capabilities needed to further refine SLR impact at regional scales.

2.2.2 Experiments at the Ten-year Horizon

2.2.2.1 Water Cycle

How will the integrated water cycle, extending from bedrock to the tropopause, evolve in a warmer climate with changes to land and water use, and changing concentration of atmospheric radiative forcing agents (aerosols, greenhouse gases)?

The goal is to understand how the coupled hydrological cycle will change and how these changes will affect the local, regional, and national supplies of fresh water. The novel aspects of this experiment are major extensions along four critical axes:

1. In the vertical dimension, from subsurface aquifers through the troposphere
2. In the horizontal dimension, from the scales of rain clouds and river headwaters to the size of the storm tracks
3. In process fidelity, ranging from cloud-scale physics and detailed subsurface flow and transport to nonhydrostatic, 1-km treatments of atmospheric dynamics
4. Through improved connections with anthropogenic agents of change (land use, land cover, water use, emission of aerosol, aerosol precursors, and greenhouse gases)

The simulation leverages DOE's world-leading capabilities in subsurface modeling and plans for deployment of the extreme-scale machines required for simulations at cloud-system scale.

2.2.2.2 Biogeochemistry

How will coupled terrestrial and coastal ecosystems drive natural sources and sinks of carbon dioxide and methane in a warmer environment?

The goal of this simulation is to determine the feedbacks from ecosystems on land and in the littoral zone as a single coupled system. The combination of coastal-zone biogeochemical cycling and its interaction with the silt, nutrients, and other substances transported by rivers and runoff has been identified as one of the major unsolved grand challenges in carbon-cycle simulation. The transformative aspects of this experiment are (1) the addition of detailed chemistry, biogeochemistry, and transport in rivers and other surface waters, a critical suite of processes missing from current Earth system models; (2) the addition of coastal ecosystems; and (3) the adoption of community-based representations of terrestrial and aquatic microorganisms and their interactions with the carbon cycle. This experiment will examine high-latitude feedbacks as affected by the implementation of major process enhancements, including the addition of root hydraulic redistribution, a reaction network for soil organic matter integrated in a one-dimensional, multiphase and multicomponent reactive transport solver, and a thermodynamically based decomposition model for soil organic carbon. This experiment leverages the advanced and efficient treatments of tracer transport developed through the SciDAC Applying Computationally Efficient Schemes for BioGeochemical Cycles (ACES4BGC) climate application and the prototype community-based biogeochemical cycling that is under development with TES and RGCM support.

2.2.2.3 Cryosphere System

How will regional variations in sea level rise interact with more extreme storms to enhance the coastal impacts of SLR?

The aim of this simulation is to determine the potential impacts on the nation's coastal zones due to SLR exacerbated by regional variations in SLR and extreme storm surges. The novel aspects of this simulation are:

1. Fully coupled models of the cryosphere, including both major land ice sheets, the floating ice shelves surrounding Antarctica, the interactions with surrounding sea ice, and icebergs calved from Antarctica and Greenland
2. Complete treatments of the impacts of time-evolving isostasy and ice-sheet self-gravity on SLR
3. Addition of wave models to the ocean component
4. Deployment of enhanced resolution in all components to resolve dynamics at ice-sheet margins, sea ice behavior, and the effects of severe weather on sea state in the major storm tracks

This experiment is based upon DOE's advances in dynamic and adaptive ice-sheet modeling combined with the capacity for ultrahigh resolution of the land ice sheets and surrounding oceans using upcoming advances toward extreme-scale computing.

3.0 Computational Research Path

ACME's 10-year computational research path must align with the science vision described in Section 2, as well as with the risk/reward profile of anticipated technical innovations. Our envisioned advancements in science require accessing the capabilities of the next two generations of LCF architectures provided by DOE-SC.

Figure 2 (included in Appendix G of the proposal) provides examples of how the computational science roadmap will align with versions of ACME science goals.

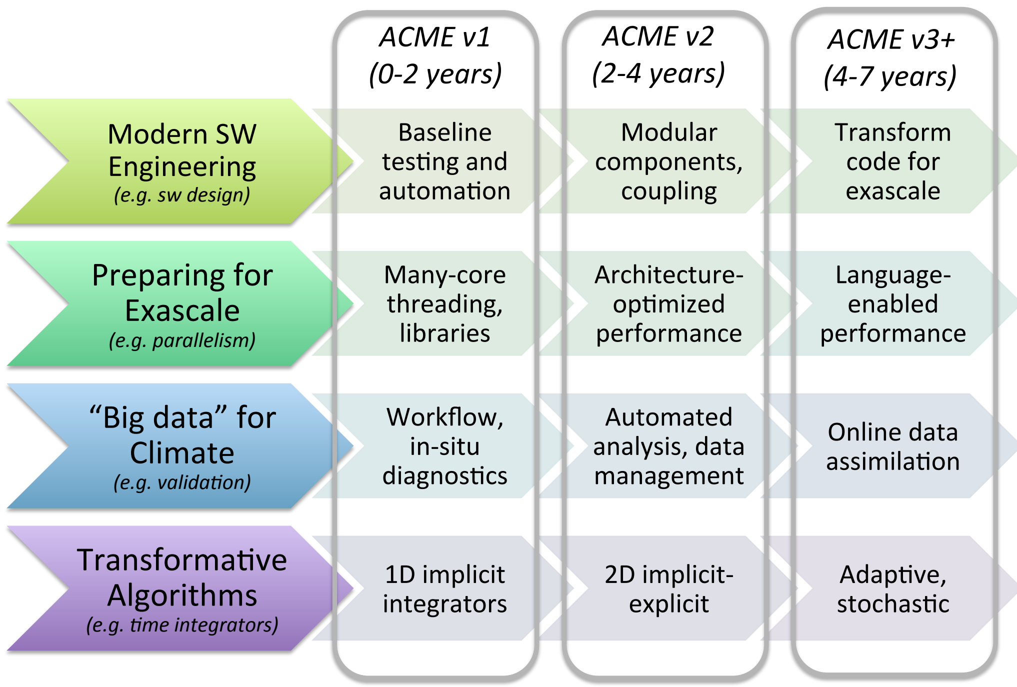

Figure 2: *Examples of potential stages in a technical roadmap for ACME over seven years.*

3.1 Computational Trends

The climate community has always been an early adopter of DOE's leadership-class computational facilities, often providing early feedback on new and experimental architectures. Climate modelers seek to accomplish better science in three ways: throughput (strong scaling), larger problems (weak scaling), and more complex climate systems (such as more complex physical representations). All of these must be explored simultaneously, creating a complex technical environment that requires significant investment across all the areas identified in Figure 2.

3.1.1 Preparing for Exascale

Future processor architectures will be constrained primarily by limited power budgets. This has two implications: more cores per processor, and less memory per core. To mitigate this dramatic change, it is expected that ever-more complex processor architectures will emerge (see Figure 3). The implication for climate codes is that new programming models will be required to support stricter memory management and more complex thread management. In addition, the power and latency cost of communication across the memory hierarchy will significantly affect performance and throughput. To mitigate these performance costs, new processors will likely expand compensating features, such as vectorization, instruction-level optimization, and in-place calculations on data within memory.

In response, ACME will need to position itself as an early tester for DOE exascale systems. CESM is already using OpenMP for threading on DOE's Argonne Leadership Computing Facility (ALCF) machine, Mira. We are exploring the use of Compute Unified Device Architecture (CUDA) and OpenACC for graphics processing unit (GPU) accelerators on Titan. Multithreading is underway within the SciDAC Multiscale project and in other CESM-focused efforts. In v3 of ACME and beyond, we will need to explore dynamic autotuning and load balancing, even work stealing, to minimize latency and make the application resilient to system disruptions from the higher failure rates and aggressive power management strategies anticipated on exascale architectures.

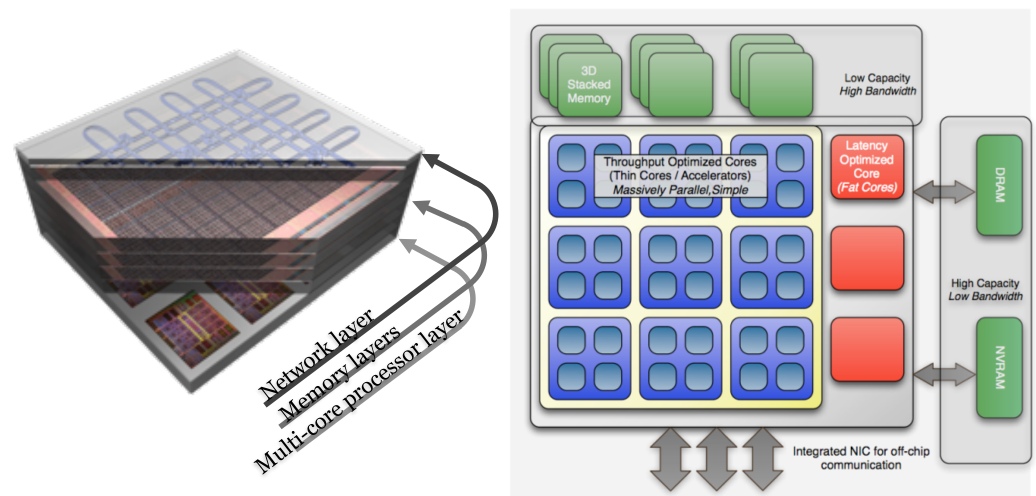

Figure 3: *Conceptualization of a future exascale many-core processor, with complex memory hierarchies and communication* (credit: Ang / Shalf – Computer Architecture Laboratory).

Compiler support poses another potential problem. It is feasible that in the next 10 years, ACME will need to abandon FORTRAN as the primary driver program and adopt C++ to take full advantage of compilers for advanced architectures, and to integrate with DOE investments in HPC libraries and frameworks. This will be a major transformation that will require extensive testing but will extend the longevity of DOE climate science applications.

3.2 Modern Software Engineering

Trends in HPC software engineering are likely to drive dramatic change in approaches to how DOE scientific applications are developed. Large distributed teams are adopting more agile methodologies, which allow for distributed code contributions, test-driven development, and extensive automation of routine software development tasks. This includes regression testing, tools for code coverage, correctness analysis, debugging at scale, and traceability of code back to scientific requirements. A recent report on HPC software productivity in DOE (www.orau.gov/swproductivity2014/SoftwareProductivityWorkshopReport2014.pdf) calls for greater use of libraries, frameworks, and tools that can be shared across the extreme-scale scientific computing community. In the near future, to maximize HPC performance, programmers will need to address many run-time issues, such as fault tolerance, power management, and complex memory hierarchies.

3.3 Big Data from Climate Models

Our current simulation capability already taxes existing analysis workflows that will not scale up to the volumes of data that will be produced by ACME simulations. Individual components often rely on separately developed analysis tools. Under ACME, we will transition to a unified approach combining DOE's existing efforts, including the Climate Science For a Sustainable Energy Future (CSSEF) and the Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT). Our workflows will make use of other Office of Science capabilities, such as Earth System Grid Federation (ESGF), Globus Online, and the AKUNA framework. These tools address data transfer, synchronization, sharing and publication, model configuration, and simulation provenance management, as well as providing a shared high-performance analysis engine. This advanced workflow and analysis tool chain will be augmented with increased use of in-situ diagnostics.

3.4 Transformative Algorithms

Climate applications will continue to evolve into more complex multiscale, multiphysics simulations. For example, near-term trends within the ACME project include watershed simulations and more complex BGC. Future Earth system models will require algorithmic advances to improve throughput (such as multitracer efficient schemes and increased use of implicit methods), to deal with the new architectures (such as the migration from global spectral methods to local element-based methods), and also to address errors from feedbacks among parameterizations that bridge across vastly different time and space scales.

4.0 Scientific Evaluation and Metrics

Evaluating model output at each stage of the project is critical for ensuring model credibility, for tracking improvement, and for addressing our driving questions. Much of our analysis will be based on well-established metrics developed by leading global modeling centers. Availability of new observations, a focus on our driving questions, and emphasis on high resolution require development of new diagnostics and metrics; this effort will be a major thrust of the ACME project. The following sections provide concrete descriptions of the analyses we will apply to the atmosphere, ocean, land, and ice components of the ACME. Since the emphasis of ACME is on the coupled system, when feasible, component metrics will be evaluated in the coupled system. In addition, we present metrics as well as methods for evaluating the realism of the coupled system and tracking improvement over time. The evaluation of the model is envisioned to be an ongoing collaborative activity with other Office of Science climate programs, especially the Regional and Global Climate Modeling, Atmospheric Systems Research, and Terrestrial Ecosystems Sciences programs. A large number of ACME participants are also supported by projects in these programs for work closely related to their ACME development tasks, resulting in strong, working-level connections that will likely lead to joint planning in the near future.

4.1 Using Metrics to Inform Development and Track Progress

Deciding which model developments make it into a model release is traditionally one of the most contentious and least objective aspects of climate modeling. In principle, we would prefer an objective process that maximizes a few important measures of model skill. In practice, identifying and selecting the few measures is not feasible, because it is impossible to provide, a priori, an exhaustive list of what matters. To make the expert judgment process transparent, ACME will establish Code Review Boards as ad hoc subcommittees of the Council to make informed recommendations on model release options, both component and fully coupled. The metrics listed here, as well as more extensive diagnostics and metrics, will inform that process. Based on past experience, it is highly likely that the choices will depend upon the target question of the simulations. Nevertheless, the following metrics are useful for tracking model improvement for versions of the fully coupled system in nearly all cases, as they represent broad, integrated measures of model performance:

- Climatology and trends of zonal precipitation
- Climatology and trends of zonal top-of-atmosphere incoming and outgoing radiation
- Climatology and trends of the timing and extent of sea ice
- Climatology and trends of two-dimensional SST fields
- Climatology and trends of zonal ocean heat content from the surface to a depth of 700 m between 65° N and 65° S
- Climatology and trends of northward annual zonal ocean heat transport by basin

4.2 Atmosphere

For the atmospheric model, our starting point will be the mean climate diagnostics and metrics compiled by the World Meteorological Organization's Working Group on Numerical Experimentation (WGNE); see www-pcmdi.llnl.gov/projects/amip/OUTPUT/WGNEDIAGS/WGNE_Mean_Climate_Diags.pdf. This document provides the “scorecard” for tracking atmospheric model progress and improvements against observations. In brief, this list includes top-of-atmosphere (TOA) and surface radiation budgets and fluxes (including all-sky and clear-sky partitioning), the surface water budget (evaporation and precipitation), measures of atmospheric moisture, measures of surface and atmospheric temperature and geopotential height, and mean and eddy transport of momentum, heat, and moisture. Evaluation includes maps of surface values and quantities at fixed pressure levels as well as global and zonal averages and time series of global average quantities. We will expand this list by considering the mean and seasonal cycle of modeled aerosols. This will initially be done using the PNNL Aerosol Diagnostics Package; see www.cesm.ucar.edu/events/ws.2013/presentations/Chemistry/ma.pdf.

We will improve our implementation of these traditional metrics relative to existing packages by using only the very best, most modern observations and by including error measures on observational uncertainty wherever possible. We expect to make particular use of the CFMIP Observation Simulator Package (COSP) for analyzing clouds in an accurate way. Because our team includes scientists heavily involved in the compilation and analysis of DOE's Atmospheric Radiation Measurement (ARM) data, we also expect that our diagnostics package will make better use of this resource than previous efforts. A unique opportunity is the development metrics based on the probability distribution of precipitation in the vicinity of ARM sites. Evaluating model skill at polar latitudes requires us to create several new metrics using recent advances in observations of liquid-ice phase partitioning.

We will also emphasize relationships between variables in the current climate, which may have bearing on future trends (commonly referred to as "emergent constraints"). Examples include relationships between subtropical stratocumulus and SST (Qu et al., 2013), estimated inversion strength (Wood and Bretherton, 2006), and vertical velocity (Meyers and Norris, 2013). Partitioning between shallow and deep convection, which is thought to affect climate sensitivity (Sherwood et al., 2014) will also be considered. Other studies emphasize relationships between variables as proxies for process interactions that are otherwise hard to capture (e.g., Gettelman et al., 2013; Lebsock et al., 2013; Suzuki et al., 2013); we will explore these relationships as well.

4.3 Ocean

Our basic ocean diagnostics and ocean model "scorecard" for tracking ocean model progress and improvements against observations are based on those developed by the CLIVAR Working Group on Ocean Modeling (www.clivar.org/organization/wgomd/resources/reos/metrics) and analyses done by the Global Ocean Data Assimilation Experiment (www.usgodae.org). This set includes a number of quantities that can be compared with observations, including hydrography (sea surface and interior fields), mixed layer depths, and equatorial jet strength and position. Poleward heat transport will be compared to direct oceanic observations and to reanalysis. We will compare mass transports against observations for approximately a dozen section lines and choke points, with acceptable bias defined initially via comparison against CESM.

Our use of a strongly eddying ocean component in our climate system model affords us the scientifically interesting opportunity to further develop metrics that take advantage of the near-global observational oceanic data sets collected over the past decades. These include sea surface height anomalies from altimetry, high-resolution (0.25°) SST from satellites, and surface velocities from drifting buoys. The advent of Argo has provided unprecedented numbers of observations of potential temperature and salinity over the top 2000 m of the water column. Metrics based on these data sets, together with recent time series measurements of Atlantic meridional overturning circulation (i.e., RAPID and MOVES) and a robust set of conventional diagnostics, will provide an increasingly solid basis upon which to gauge model performance and identify any unexpected issues that may arise.

It is important to realize that the development of effective diagnostics for a strongly eddying ocean climate model is still in its infancy, so we expect the above diagnostics and metrics to be augmented or superseded by the results of future research. A diagnostic especially appropriate to a high-resolution ocean model is evaluation of transport within deep western boundary currents, particularly in the North Atlantic. This diagnostic is particularly important because ocean eddies couple the energetics of the upper and deep ocean. We will evaluate deep western boundary transport north of the Gulf Stream based on Bryan, Hecht and Smith (2007) and references therein, and to the south (at approximately 15° N) based on Send et al. (2011). Deep flow associated with a number of major jet systems, including the Gulf Stream, Kuroshio and Antarctic Circumpolar Current, will also be compared with observations.

Other high-resolution-specific diagnostics will be constructed from more direct measures of mesoscale variability. For example, basin-wide values of eddy kinetic energy provide a simple measure of eddy activity. We will also compare the distribution of eddy length scales to observations and will construct Taylor diagrams indicating the degree of improvement in eddy variability and correlation of those features with observed variability, as in McClean et al. (2008). Much of the basis for comparison is provided by the two-dimensional view of altimetry, while moorings add information in depth, and Argo floats offer aspects of both. To correct for discrepancies between what is observed by Argo floats and what is output by climate models, an Argo float simulator will be written for the MPAS ocean model. This will come with other sophisticated diagnostic capabilities such as Eliasen-Palm eddy flux tensors (Ringler et al., 2014).

4.4 Ice Sheets

The ice sheet model brings in a new set of interactions that must be understood and monitored for development or reduction of biases. The state of the atmosphere is essential to the surface mass balance, which in turn must be compared against other mass loss terms (iceberg calving, submarine melting). Along with monitoring of changes in grounded ice area and volume and these related mass balance terms, some of which may prove useful as metrics in a coupled context, a research problem we intend to take on is the development of Circumpolar Deep Water (CDW) metrics. Where waters are all near freezing, the relatively warm CDW, often found below colder and fresher waters, has a strong potential to accelerate submarine melting and therefore to strongly impact the overall mass balance of the Antarctic Ice Sheet. Stronger winds can drive greater upwelling and bring these warm waters in contact with the ice shelf. Research to understand the controls on CDW state and variability will be undertaken, facilitating the establishment of metrics focused on CDW mean state and variability, ensuring that transitions to or from rapid submarine melt states are not modeling artifacts but are robust and well understood. This effort will be critical for successfully answering the cryospheric driving question.

4.5 Sea Ice

The areal extent of sea ice is well constrained through remote sensing and is an essential metric for evaluating the fully coupled system. Comparisons of seasonal ice extent to satellite observations from the special sensor microwave/imager (SSM/I) (en.wikipedia.org/wiki/Special_sensor_microwave/imager) are one of the primary realism metrics performed on coupled simulations. Sea ice thickness and its distribution is also available, and though it is more poorly constrained from remote sensing it is important to the preconditioning for extreme minimum events, and hence provides another sea ice metric that is important within the coupled system. We will also analyze sea ice velocity, which can be observed via remote sensing and Lagrangian measurements and tends to be biased low in climate models. We have recent experience with comparison of modeled and observed sea ice age, and have found it to be a useful constraint on model configuration. Export across several transport lines will be monitored. A number of surface properties will be monitored, including albedo, distribution of melt ponds, and snow depths. The timing of freeze-up and onset of melting must also be monitored.

4.6 Land

The International Land Model Benchmarking (ILAMB) activity provides the basic "scorecard" for tracking land model progress and improvements against observations (www.ilamb.org/benchmarks). The ILAMB goals are to assess and improve the performance of land models through international cooperation and to inform the design of new measurement campaigns and field studies.

Important classes of observations that we will use in our analysis include DOE Ameriflux observations of energy and carbon exchange, and NASA remote sensing observations of land and ocean ecosystem characteristics. An important improvement over previous benchmarking approaches will be the explicit application of functional response metrics, which relate a predicted state or flux to a forcing or static system property. Examples of functional response metrics that we will apply in the ACME development effort include:

1. Tropical, temperate, and high-latitude net primary productivity (NPP) responses to precipitation, temperature, and solar radiation
2. Soil moisture, runoff, and river discharge responses to precipitation and snowmelt
3. Soil nitrogen and carbon storage and export responses to NPP, precipitation, and temperature
4. Evapotranspiration (ET) response to precipitation, temperature, and NPP
5. Gross primary productivity (GPP) and ecosystem restoration (ER) as functions of leaf area index (LAI) in Arctic ecosystems
6. Vegetation biomass as a function of precipitation, temperature, and nutrient availability

4.7 Coupled System

Applying the described analyses to coupled simulations is already a stringent test of model skill. Here, we focus on diagnostics that relate to coupling between model components, or which only make sense to test in coupled mode. The unique aspects of our effort require the development and implementation of several new metrics. Because of our focus on the water cycle and the expectation that our high-resolution model will better capture propagation of continental mesoscale convective systems (such as those over the American Midwest) as well as tropical cyclones, we will look at the tracks, shapes, count, and intensity of these storms. Another important coupled-model diagnostic is the position of the Inter-tropical Convergence Zone (ITCZ), which couples tropical precipitation and SST. We will also pay attention to the model's ability to reproduce observed modes of internal variability (ENSO, PDO, AMO, MJO, etc.). Capturing these sources of variability is critical to understanding predicted precipitation variability and changes in the ACME model. Our starting point will be the recently released variability diagnostics package developed by Adam Phillips and colleagues at the National Center for Atmospheric Research (NCAR) (www2.cesm.ucar.edu/working-groups/cvcwg/cvdp). Our emphasis will be on metrics developed from the observed power spectra for these modes, and observed teleconnections associated with these oscillations.

5.0 Computational and Workflow Infrastructure

The ACME project is ambitious in its science objectives and simulation scale, and requires an equally ambitious technical infrastructure that will enable its success. In the short term, this means prioritizing investments that enable fast "time-to-science," including automation of development and verification, throughput of the coupled model, and verification of results. Figure 4 is a conceptual view of this iterative cycle and its key processes.

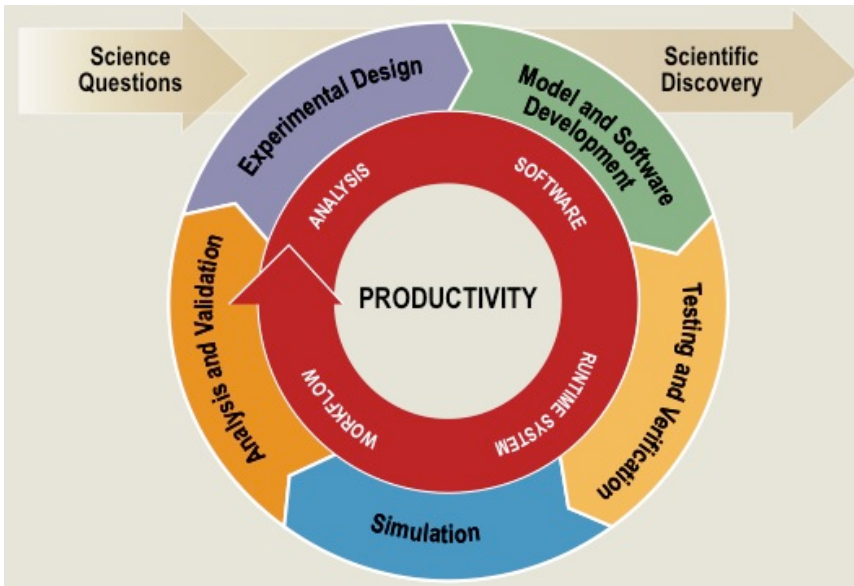

Figure 4: *Conceptual view depicting how science productivity is enabled by a continuously evolving infrastructure.*

- Experimental Design: Setting the scientific objectives for a particular simulation, creating a feasible experimental design, and establishing the criteria for assessing the results.
- Model and Software Development: Consistent configuration management, design and code reviews (against both scientific and technical criteria), and productive development tools and processes.
- Testing and Verification: Testing both new code and the coupled system, early and often, making sure the current code base satisfies new science requirements without unanticipated negative impacts.
- Simulation: Running the ACME application at scale, with required throughput and data to accomplish the scientific objectives of that simulation.
- Analysis and Validation: Analyzing simulation results and comparing them with observational and other results, with the confidence that simulation ensemble statistics can be reproduced.

Technical infrastructure tools focused on consistency and automation must support each of these processes. This is a key distinction between ACME infrastructure (for example, software engineering and workflow) and more speculative research tasks (numerical algorithms, speculative coding for HPC architectures, and major code refactoring). Together, these efforts must establish a solid technical capability, while maintaining agility and the flexibility to adapt ACME to future climate science goals.

ACME will prioritize infrastructure investments in the first 12 months, making a critical assessment of the project's ability to accelerate science results. In addition, these infrastructure investments must enable the project to function well with a large, distributed team, including tools for project management and communication, feature/issue tracking and release management, and dashboards that provide visibility into the state of ACME. The following three sections break these down into topics for Software Engineering, Workflow-Diagnostics, and Performance.

5.1 Software Engineering Infrastructure

The ACME software engineering team has prioritized the following Year 1 metrics:

>(SE1) *The ACME development branch will be in a releasable state 90% of the time.*

The definition of "releasable state" can be defined by the software engineering team, but includes: building on about five important platforms (ranging from development platforms to LCFs); automated tools to check out the current release configuration from repositories; and automated unit and component tests passing. Although this metric does not include assessment of the releasable simulation statistics, a basic check of the climate (conservation, growth rates) with relatively short runs (~1 climate day) must pass. Similarly, runtime regression tests must pass (indicating no performance, memory, or I/O degradation), and code test coverage and performance metrics should be gathered automatically.

>(SE2) *The development build/test cycle will be automated and as fast as possible.*

ACME must make sure that scientific and technical teams can develop and debug in a fast, agile, and easy workflow. For example, isolated source code modifications should take less than 10 minutes to fully regression-test, or development testing branch commits will be slowed significantly. Compiler output should emphasize warnings and errors over informational messages in log files. In addition, component diagnostics should happen quickly when run on model output during debugging and development.

With the ACME focus on delivering a coupled model, the code repository needs to expedite the merging and testing of the fully coupled system. The repository must support a distributed development environment where separate features are being codeveloped at different sites, and must also support ongoing development in separate branches with sophisticated tools for merging. Since some components are co-developed by those external to ACME, there must be an ability to tailor access controls on a component level, and to facilitate branching and merging with external component repositories.

The start of ACME provides an opportunity to make a change toward industry best practices. These are software engineering tools and processes that have been shown to increase developer productivity over the course of a project. By adopting commonly used open-source tools and development workflows wherever possible, a large community becomes available to draw upon for guidance and debugging, and the barrier to new developers becoming productive on the project becomes lower.

Based on the requirements and opportunities presented above, the ACME software engineering team chose the git version control tool. git is the tool of choice for large distributed projects ranging from DOE's Trilinos, PETSc, and CASL projects, to LINUX and Facebook development. The tool's functionality, as well as the culture that comes with it, promotes frequent branching and merging, and the ability to share ongoing development between distributed developers. We are developing the ACME usage of git so that it both supports the emphasis of a fully coupled model and also the requirements for some components to support external developers. In this second aspect, we are developing scripts that will track component development at an external site and reliably merge it into ACME in full compliance with our development workflow. A guiding principle in our design is to avoid (if possible) the diverging of active development branches, which leads to time-consuming merges, redundant development, or outright losses of capability. ACME software will be hosted on the github website (under the ACME-Climate organization name). This provides a rich software environment surrounding git for a wide range of tools to improve the collaborative code development process. This includes a myriad of plug-ins for project communication and management tools, including development statistics, visualization of code development branches, integrated issue tracking, continuous integration tools and dashboard, mailing lists, wiki pages, etc. The choice of github brings tremendous leverage from the broader software engineering communities in terms of tools and experience, and the environment will continue to improve without any additional input or investment from us.

5.2 Workflow-Diagnostics

An important part of the transition from multiple projects to an integrated ACME project is the consolidation of the independent workflows currently used for the coupled system among the separate BER projects. This is not just a technical issue, but also one of scientific productivity. Our plan to standardize ACME workflows involves the identification of the redundant, and possibly inefficient, current workflows across and within projects and collaborations and developing transition plans for them. Early tasks are geared toward enabling easier and more reproducible model development in ACME as the project begins.

Three general workflows were outlined in the proposal: developmental (DEV), experimental (EXP), and production (PROD). Currently, the largest issues facing the ACME team for workflows are maintaining provenance of large-scale coupled model runs, data volume and sharing, and the creation, sharing, and manipulation of diagnostic plots of model runs. There is currently no automatic process to verify the complete setup and settings of a model run before execution. The volumes of data generated from even short high-resolution coupled runs are too large to transfer outside the LCF centers, so data sharing requires that we move beyond requiring all scientists maintain an account on a single machine and transition to storing data on servers. Then scientists can perform analysis and post-processing locally and share plots and smaller processed data.

The workflow-diagnostics Year 1 priorities will focus on automation of manual, time-consuming processes for model development and validation of the coupled system. One key metric is:

>**(WD1)** *Starting from a completed simulation, minimize the wall-clock time and person-hours required to complete a diagnostic assessment of the coupled model.*

The goal of this metric is to improve the time to produce climate data files and related diagnostics even while the simulation is running, and to be able to assess diagnostics such as annual and seasonal averages, and monthly fields, which can be accumulated as they are produced. Automating the system for publishing these on ESGF and generating analysis that contributes to a science-driven coupled system scorecard will be key for demonstrating ACME's automated workflow tool environment.

Within six months, we will complete a number of foundational tasks that will focus on the metric above, including:

1. Interpolation of unstructured grid output from all components from a script that is automatically or manually executed after model runs

2. Build/run of UV-CDAT tools to visualize data in all components, including key static plots, configuration for the model run, and comparisons to sets of observations or other model results
3. Auto-publishing of data to ESGF on CADES from model runs (from a script)
4. Support reproducibility from model setup/configure information of a model run and experimental setup (including scripts to create clones, plus any changes)
5. High-performance I/O and file management to allow quicker use of UV-CDAT

Follow-on from the initial priorities will include training and expanding the user base, as well as growing support for more plots, metrics, and interactive analysis. Remote sharing and retaining history will be implemented through ESGF, as will support for a CMIP5-friendly format. When ACME science results are ready for public release, the workflow team will be responsible for creating a public site.

5.3 Computational Performance Improvement

During the first six months of the project, the performance engineering effort will assess ACME v0 application performance and identify the target metrics for each component in terms of its throughput and scaling behavior in the coupled system. Then improvements in on-node (using OpenMP or OpenACC to expose more parallelism and leverage features of the different LCF architectures) and between-node (with communication-hiding implemented over MPI) can be focused on improving the coupled model throughput. Our performance strategy for the v1 model was chosen to increase performance on the existing LCFs and position us to exploit new codesign-inspired approaches in the v2 model. We will focus on two key tasks: exposing increased concurrency throughout the model and increasing the on-core performance of key computational kernels. In v1, we will be using conventional approaches, such as threading and MPI for the first task and increased use of accelerators for the second task. Increases in concurrency and much of the work needed to refactor and modularize kernels for accelerators will be beneficial for most any possible exascale architecture. The throughput-based priority performance metric is:

> **(Perf1)** *Maximum simulated years per wall-clock day of the coupled system running without I/O.*

This includes the ACME target of five simulated years per wall-clock day (SYPD), and the required performance by each component that is needed to achieve that target. Because I/O performance is simulation-dependent (based on science objectives), it will be dealt with in a separate task that focuses on the most critical issues.

The performance metric is only based on the speed-up of the coupled system, so that we avoid focusing effort on irrelevant components. For example, if the project focused performance resources only on the atmosphere component and made it 100× faster, that would be an amazing achievement, but the coupled system would not run any faster and this metric would correctly show that little progress had been made. Because of the expense of adapting and tuning such a large code base to different architectures, the performance team will assess and determine which LCF platform will be the focus of the largest simulations, assuming that resources on that system are available through the ALCC and Innovative and Novel Computational Impact on Theory and Experiment (INCITE) allocation processes. I/O performance will be tracked relative to the cost increase as between the coupled system without I/O and the coupled system with a predefined “typical” set of high-frequency 3D output.

6.0 Management

6.1 Organizational Structure

Based on reviewer suggestions and early experience, we have altered the ACME project structure by eliminating the two standing subcommittees, reducing the number of Council members by one, hiring a full-time project engineer, and adding a seventh task team for coupled model simulation tasks. These changes are reflected in Figure 5.

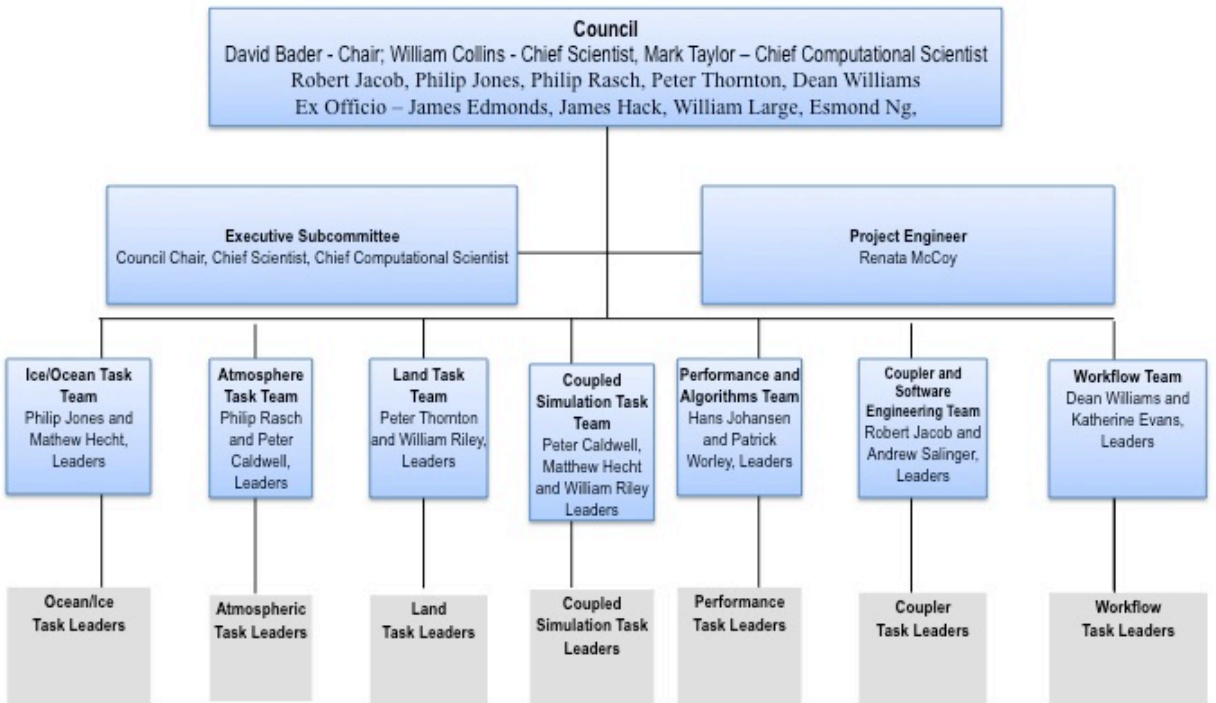

Figure 5: *New ACME Project organization.*

We have installed an executive function comprising the Council Chair, the Chief Scientist, and the Chief Computational Scientist. The Executive Subcommittee will be responsible for closely monitoring the project and identifying both the short- and long-term issues that may require full Council action. Final executive authority rests with Council chair, but only upon the advice of, and consultation with, the other members of the Executive committee, which nominally will meet weekly between the weekly full Council meetings. With the creation of the Executive Subcommittee, the standing Science and Computational Science subcommittees were disbanded. The flexible management approach we have adopted works better in a flatter organization. We have retained the concept of ad hoc, or Deep Dive, subcommittees comprising at least one Council member and any number of project participants, as needed. These Deep Dive committees will be chartered and tasked by the Council to address specific project issues. A concern raised by the reviewers was that the Council would not be actively involved in the project work. We note that six of the seven task teams are lead by Council members, and the dual roles provide essential coordination and integration functions.

The most significant change is the addition of Dr. Renata McCoy of LLNL as the ACME Project Engineer. Dr. McCoy is already instituting several of the reforms suggested by the reviewers, including the implementation of a formal project management and tracking system (Jira, www.atlassian.com/software/jira, is currently being evaluated). We retained the concept of organizing the project by task, not by people. This flexible approach allows for tasks to be accomplished by semiautonomous teams in relatively short amounts of time. It is a feature of the agile software development methodology described earlier and is used by organizations that require rapid development in a changing technological environment. Our tasks (Appendix C in the proposal and modifications) are still organized to fall uniquely under one task team, which allows project staff and even task leaders to work on multiple tasks in multiple areas.

6.2 Project Metrics

The project management tools enable the project participants, leaders, and sponsors to efficiently collect data about the project and its products and deliverables. Metrics will be developed at multiple levels, including tasks and subtasks, which will provide current and continuing information about project status. At the highest level, the following metrics provide overall measures of the project's status. Where practical, the top-level metrics will be evaluated and reported semiannually.

6.2.1 Scientific Quality of Simulations

In Section 4, we presented a list model fields for the coupled system and its components that provide quantitative information about the model's ability to simulate the observed climate. Our baseline values against which scientific improvements will be measured will be diagnostics of these fields from the ACME v0 model. Metrics will be computed as a fractional reduction of model errors compared to the v0 baseline values.

6.2.2 Model Capabilities

ACME is primarily a model development project, and its success is determined by improving model capabilities to answer critical science questions. A useful metric to measure progress in this area is the number of additional model capabilities added as a function of time compared to those planned. We would only count those improvements that have completed the unit-testing and preliminary evaluation process and subsequently have been approved for inclusion by a code-review board.

6.2.3 Model Performance

As was stated in section 5, a straightforward performance metric for the coupled system and its components is simulated years per wall-clock day of the coupled system running without I/O.

6.2.4 Scientific Productivity

The number of publications resulting from the ACME project will be tracked and reported. Additionally, the number of simulations completed compared to the number planned and the amounts of simulation output released to others for analysis are useful metrics of productivity.

6.2.5 Scientific Impact

Good measures of scientific impact are ACME publication citations and the number of independent downloads of released simulation results and formal data citations.

6.2.6 Mission Relevance

The number of new projects supported by DOE program offices outside of the Office of Science, as well as the number of USGCRP agency projects that utilize models or simulations produced by the ACME project, will be tracked as a measure of relevance to DOE and national mission needs.