

An Integration and Evaluation Framework for ESPC Coupled Models

PI: Ben Kirtman
University of Miami – RSMAS
Atmospheric Sciences
4600 Rickenbacker Causeway
Miami, FL 33149
Phone: (305) 421-4046 Fax: (305) 421-4696 email: bkirtman@rsmas.miami.edu

Award Number: N00014-13-1-0508
<http://www.earthsystemcog.org/projects/espc-infrastructure/>

LONG-TERM GOALS

To realize its potential, a U.S. Earth system modeling and prediction capability must encompass a network of agencies and organizations that contribute model components, infrastructure, and scientific and technical expertise. The model component contributions must be integrated using coupling software, the coupled systems optimized for emerging computing platforms, and the predictive skill of the resulting models assessed using standard metrics. We propose to provide these integrative functions for the Earth System Prediction Capability (ESPC), using as a reference application a version of the Community Earth System Model (CESM) running an optimized version of the HYbrid Coordinate Ocean Model (HYCOM).

OBJECTIVES

The objectives of our work are:

- To establish an ESPC Coupling Testbed that allows for collaborative research into coupling technologies, and use it to prototype multi-model optimization techniques focused on computing systems with accelerator technologies.
- To support migration of optimization strategies from the ESPC Coupling Testbed to infrastructure packages and coupled model applications, and provide support for coupling of optimized components in the ESPC program.
- To update a newly coupled CESM-HYCOM model configuration with an optimized version of HYCOM and assess the coupled system performance.
- To evaluate the CESM-HYCOM model using standard measures of predictive skill, and promote the usage of standard metrics by other models in the program.
- To extend ESPC-related computational committees to address new requirements driven by changes in computing architectures and program needs, and to initiate scientific committees.

APPROACH

A one-year seed project entitled *Optimized Infrastructure for the Earth System Prediction Capability (OI for ESPC)* initiated several technical elements of this project: updating numerical libraries in the Earth System Modeling Framework (ESMF) to use the DOE-supported MOAB finite element mesh; exploring how to optimize a component architecture like ESMF for accelerators; and coupling the

HYbrid Coordinate Ocean Model (HYCOM) to the Community Earth System Model (CESM). The *OI for ESPC* project will end on October 31, 2014.

The team's technical approach relies on making incremental changes and validating work frequently against established versions and known quantities. This is applicable to tasks associated with the integration of MOAB into ESMF (where MOAB is compared to the existing finite element mesh in ESMF), the integration of ESMF-based software into CESM (where the ESMF-based code is compared to native CESM and its resident MCT - Model Coupling Toolkit code), and the coupling of HYCOM to CESM (where HYCOM in CESM is compared to HYCOM standalone versions and will be compared to the Parallel Ocean Program – POP model in CESM). Where the work is more research oriented, the team has explored ideas using prototypes and toy codes. This is the approach being used for the exploration of accelerators in a modeling framework.

WORK COMPLETED

Year I milestones and their current status are shown below.

Admin **Initiate the project website, repositories, and communication schedule.**

Collaborators engage in a weekly technical telecon and are subscribed to an esp_tech@list.woc.noaa.gov mailing list.

The project is continuing to use the *OI for ESPC* site on CoG oG as a primary hub for wiki communications: <http://www.earthsystemcog.org/projects/esp-infrastructure/>

Progress is tracked using a set of roadmaps, which include:

MOAB finite element mesh in ESMF:

http://www.earthsystemcog.org/projects/esp-infrastructure/moab_in_esmf

Optimize Component Architectures for GPUs:

<http://www.earthsystemcog.org/projects/couplingtestbed/acceleratorplans>

Couple HYCOM to CESM:

(Part 1) Introducing NUOPC conventions into CESM

http://www.earthsystemcog.org/projects/esp-infrastructure/nuopc_in_cesm

(Part 2) HYCOM in CESM scientific validation

http://www.earthsystemcog.org/projects/esp-infrastructure/hycom_in_cesm_science

Some of the *OI for ESPC* project activities are closely associated with a Coupling Testbed, which is a collaborative area for groups working on research problems involving model coupling. A site for the Coupling Testbed was also initiated and is here:

<http://www.earthsystemcog.org/projects/couplingtestbed/>

However, the Coupling Testbed as a concept has not been developed or promoted in any significant way yet.

The CoG environment allows team members to show project and roadmap linkages among activities.

Repository strategies for working jointly among CESM, ESMF, and test codes have been established and are working effectively.

Admin **Get computer accounts and access to codes and repositories.**

This has been completed as needed.

Collaborative **Initiate the ESPC science working group and set up communication schedule.**

These interactions have managed informally so far.

Collaborative **Collect basic metadata for the ESPS Suite and create web presence.**

A website for the ESPS (Earth System Prediction Suite) was created here: <https://www.earthsystemcog.org/projects/esps/>

This site includes just a few metadata elements for each code. Many but not all codes include a statement from their primary organizational affiliation about how to apply for access to the NUOPC/ESPS version of the code. Work is continuing on collecting these access paths for all codes, with the underlying idea that accessibility and usability are essential for realizing an ESPS vision. Work began in September 2014 on developing a questionnaire and display for additional metadata elements.

In addition, participants from this ESPC project and developers of other ESPS models drafted a paper describing the ESPS suite and its current and projected impacts. A proposal for an article in the Bulletin of the American Meteorological Society (BAMS) was submitted and approved in July 2014, and the invitation to submit is open for six months following acceptance. Submission is waiting on the validation of HYCOM coupled to CICE using the NUOPC version of CESM. The reason for waiting for this is that CESM has a clear process for model access, even for development versions, relative to other ESPS codes. Having a CESM version for examination would give interested readers and/or reviewers a potential path to obtain a coupled ESPS code. A validated version of HYCOM-CICE in the NUOPC version of CESM is expected before the end of the *OI for ESPC* project on October 31, 2014.

Other ESPS coupling projects have advanced over the last year, including the NOAA Environmental Modeling System (NEMS). NEMS now supports an initial two-way coupling of atmosphere and ocean (HYCOM and Modular Ocean Model – MOM) components using NUOPC software to transfer data fields. The chart on the ESPS CoG page (URL above) provides more information about anticipated ESPS models and components.

Collaborative	<p>Coordinate with NUOPC leads to extend or supplement technical committees.</p> <p>The NUOPC Content Standards Committee (CSC) and Common Model Architecture (CMA) Committees have been extended to include key members of this collaboration and others working in areas relevant to global model processes and infrastructure. New members of the CSC include Tony Craig (involved in CESM, NUOPC, and other coupling projects), and new members of CMA are David Gochis (NCAR - land/hydrology, NOAA water center), and Tim Schneider (High Impact Weather Prediction Project, NOAA water center, hydrology). These committees meet approximately monthly.</p>
CESM-HYCOM	<p>Complete the CESM-HYCOM coupled system under the <i>OI for ESPC</i> award.</p> <p>This should be simplified by the use of the MCT datatype in ESMF. Make it available to Navy and other collaborators and establish baseline performance.</p> <p>As noted under the “Collect basic metadata ..” milestone, an initial version of the NUOPC-CESM coupled system configured for active HYCOM-CICE is expected to be validated by the end of the <i>OI for ESPC project</i>, which is at the end of October, 2014. Work on this activity is described in the <i>OI for ESPC</i> annual report.</p>
Model evaluation	<p>CESM-HYCOM simulations at low and high resolution (coupled system under the <i>OI for ESPC</i>); science application evaluation and comparisons with existing CCSM4 and CESM simulations – global analysis, mean errors, assessment of variability, air-sea coupling characteristics.</p> <p>CESM-HYCOM at high resolution not available yet.</p>
ESPC Testbed: Basic optimization	<p>Implement MCT datatype in ESMF and include in ESMF release.</p> <p>This was not yet started.</p>
ESPC Testbed: Basic optimization	<p>Complete initial optimization work under the <i>OI for ESPC</i> award and release as updates to ESMF/NUOPC Layer.</p> <p>Jayesh Krishna of University of Chicago led the effort to develop a suite of experimental codes that examine how accelerators might be used with a component framework like ESMF. Gerhard Theurich of the ESMF team provided guidance and expertise on the ESMF representation of machine hardware.</p> <p>Since there are a variety of programming paradigms and a variety of accelerator hardware options, an initial task was determining how much resource visibility and control over data movement these options would offer when called by a framework like ESMF. Such visibility and control would be</p>

necessary for ESMF to participate in accelerator resource negotiation or allocation. Other questions that were explored were whether components that used different paradigms could be coupled together, and whether there was a paradigm that fit most naturally into ESMF and could become its “native” implementation.

To begin answering these questions, Jayesh implemented the following examples:

1. An OpenCL application that provides information about the accelerator devices in the system
2. An ESMF application with two components - an MPI non-accelerated component, an OpenCL accelerated component - and a coupler
3. An OpenACC application that provides information about the accelerator devices in the system
4. An ESMF application with two components - an MPI non-accelerated components, an OpenACC accelerated component - and a coupler
5. An ESMF application with two components - an MPI non-accelerated component, an OpenACC accelerated component - and a coupler. The OpenACC accelerated component toggles (at runtime) between running the accelerated part of the code on the host and the accelerator.
6. An application that uses Intel utilities to provide information about the accelerator devices in the system
7. An ESMF application with two components - an MPI non-accelerated component, a component accelerated using Intel MIC directives- and a coupler
8. An ESMF application with two accelerated components - an OpenCL accelerated component, an OpenACC accelerated component - and a coupler
9. An ESMF application with two accelerated components - an OpenCL accelerated component, a component accelerated using Intel MIC directives - and a coupler

These prototypes were used to construct the following table:

Programming Models / Accelerator Framework Features				
Feature	OpenCL	OpenACC	Intel MIC Directives	Notes
Find number of accelerator devices	YES	YES	YES	
Can ESMF applications use the accelerator framework?	YES	YES	YES	
Turn off offloading to device at runtime	YES	YES	NO	Need to test building native Intel MIC ESMF apps
Check if current thread is executing on the device	YES	YES	YES	
Mix accelerator frameworks in the same application?	YES	YES	YES	

An initial conclusion from the prototypes was that OpenCL would be a reasonable choice for implementation of accelerator-aware capabilities in ESMF. It has features that will be needed and its non-directive based approach is closer to the implementation of ESMF's Virtual Machine or VM (representation of machine hardware) than the others. Note that model components using ESMF can still use the other approaches internally.

Working on a remote branch of the ESMF repository, Jayesh has started integrating an awareness of accelerator resources into the ESMF source code. The initial steps here were to add support for ESMF Persistent Execution Threads (PETS) to query the number of accelerator devices via an ESMF_VMGet() call to the ESMF Virtual Machine (VM). Next, a prototype application was developed that uses ESMF_VMGet() to query the number of accelerator devices attached to a PET and allocates all the devices attached to a node to one (root) process running on the node. This application is now available in the list of prototype applications, and is being tested on an Argonne computing resource (Fusion@ANL).

ESPC
Testbed:
Technical
evaluation **Begin performance comparison tests of Testbed infrastructure with CPL7 dead and data models.**

This work has not yet started.

MOAB mesh **Initial updates to MOAB to accommodate requirements of ESPC codes. Release updated MOAB.**

ESPC codes did not yet have specific requirements for MOAB. However, there were general changes that needed to be made to MOAB this year in order to integrate it into ESMF and use it to implement grid remapping. These changes are described in the *OI for ESPC* annual report.

ESPC
Testbed:
Interactive
ensemble **Initial prototype of multi-model interactive ensemble coupling infrastructure.**

Initial development of the interactive ensemble was done in the CCSM framework rather than the NUOPC framework, as the focus of the latter this year was to get HYCOM integrated.

Enhanced the interactive ensemble so that multiple atmosphere, land and ice component models can be simultaneously coupled to a single ocean component model.

This code development serves two purposes: (1) removes energetic inconsistency over the land-atmosphere interface and over the ice-atmosphere and ice-ocean interfaces; and (ii) allows for additional flexibility when implementing the multi-model interactive ensemble coupling. Figure 1 and the associated discussion below demonstrates the effect of this enhanced

interactive ensemble.

HYCOM – Platform Independent Code Development – The primary work revolved around designing a library capable of performing the operations needed to integrate the HYCOM model forward. The aim is to produce a tunable library that can be ported to the different high-performance platforms available while shielding the main part of the code from the changing hardware. This will reduce HYCOM's maintenance overhead by keeping a single code for the main portions of the model, and by restricting the modifications to the compute-intensive parts. A first version of the library has been produced which is capable of performing all the operations needed to compute averages and gradients on an Arakawa C-Grid, and which has been used to develop a simple shallow water code. The library is now being retooled to accept arrays laid out according to the HYCOM convention, and to expand the library's capabilities to include more low-level code commonly found in HYCOM.

RESULTS

Initial communication and collaboration infrastructure was mostly carried over from the *OI for ESPC* project, and expanded for new goals and partners. It includes a regular meeting schedule, wikis, and coordinated repository strategies. Definition and promotion of the ESPS has advanced, with a paper proposal accepted into BAMS and the submission waiting on completion of a validated initial version of the HYCOM-CESM coupling. Ongoing efforts to bring MOAB into ESMF and HYCOM into CESM are approaching completion of initial deliverables, with the target the end of the *OI for ESPC* project in October, 2014.

An initial set of prototype applications was delivered by the University of Chicago team, enabling ESMF collaborators to make an initial decision about implementation of accelerator-aware capabilities in ESMF. Integration of these capabilities into ESMF began with modifications to its hardware representation (Virtual Machine) to query for the presence and number of accelerator devices.

A forerunner, and indeed, the software underpinning the multi-model interactive ensemble as part of this project is the so-called interactive ensemble. For example, the interactive ensemble implementation developed for the NCAR family of models (CCSM3, CCSM4 and CESM1) uses multiple realizations of the atmospheric model (CAM) coupled to a single realization of the ocean model (POP), a single realization of the sea-ice model and a single realization of the land-surface model. The coupling of the multiple realizations of CAM to the single realizations of the other component models is accomplished through the CCSM coupler. The purpose of this coupling strategy is to significantly reduce the stochastic forcing of the ocean due to internal atmospheric dynamics. Ensemble averaging of fluxes of heat, momentum and fresh water produced by the individual CAM ensemble members before they are passed to POP effectively filters the noise in the fluxes due to internal atmospheric dynamics.

Over the ocean this approach works well since the coupling is entirely through the fluxes and the ensemble averaging of the fluxes is found to do little damage to the mean state of the coupled system. However, over the ice and land the coupling includes fluxes and state variables that are used by the component models to calculate fluxes. The ensemble averaging, therefore, has a disproportionately large effect on state variables compared to flux fields, which ultimately leads to differences in the

climatology of ice and land that are too large. For instance, over ice and land wind-dependent fluxes are calculated using the ensemble mean *state* (as opposed to flux) variables leading to strong biases over land and ice.

The above implementation of the interactive ensemble that only allows a single land and ice component model not only has bias issue noted above but is also a major difficulty in the development multi-model interactive ensemble. The problem for the multi-model interactive ensemble is that we need the flexibility to include additional atmospheric component models that may not include their own land or ice component models or may use fluxes and state variables as noted above. *To solve this problem, we enhanced the interactive ensemble so that multiple atmosphere, land and ice component models can be simultaneously coupled to a single ocean component model.*

Figure 1 shows how this enhanced version of the interactive ensemble affects variability over sea-ice in the the Northern Hemisphere. Here we plot the surface temperature variance over ice for the original interactive ensemble (i.e., the version with ten atmosphere instantiations coupled to one instantiation sea-ice; left panel) to be compared with ten instantiations of the sea-ice model coupled to ten atmosphere instantiations (right panel). The larger surface temperature variance in the original interactive ensemble compared to the enhanced interactive ensemble is readily apparent. Much of this larger variance is erroneous compared to the control version of CESM.

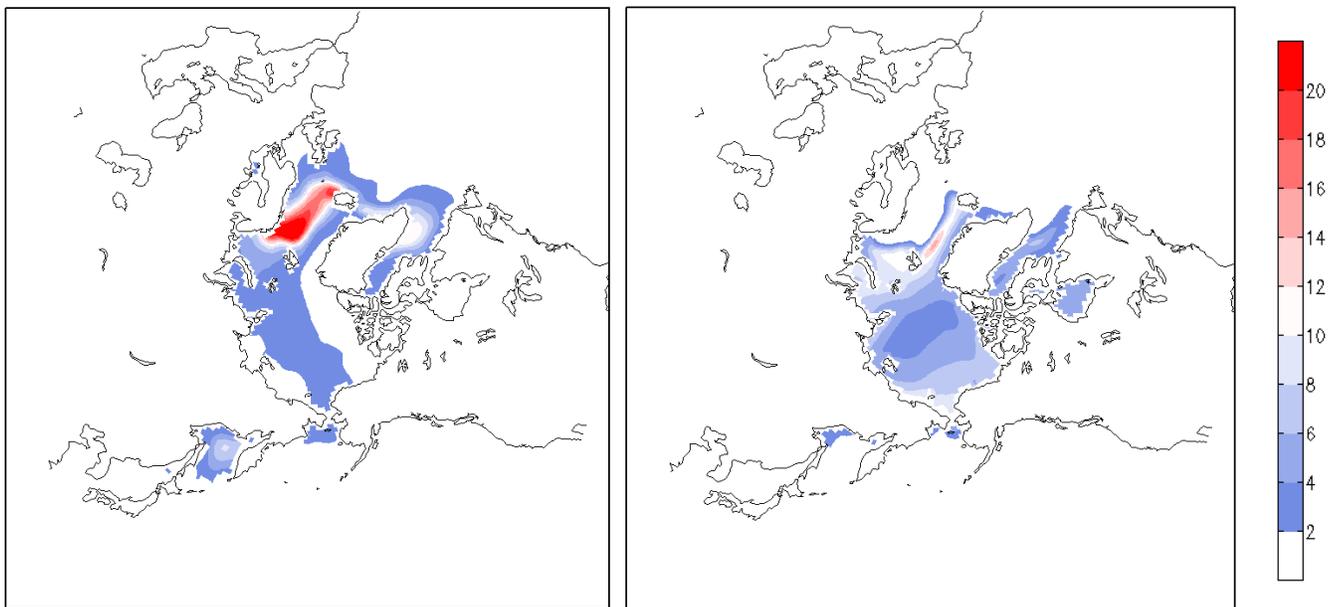


Figure 1: Interactive ensemble surface temperature variance. The left panel shows the variance for the original interactive ensemble with ten atmospheric instantiations coupled to one sea-ice instantiation. The right panel shows the variance for with ten atmospheric instantiations coupled to ten sea-ice instantiations.

IMPACT/APPLICATIONS

Through the actions of a succession of infrastructure projects in the Earth sciences over the last two decades, a common model architecture (CMA) has emerged in the U.S. modeling community. This has enabled coupled models to wrap high level model components in community-developed ESMF and NUOPC interfaces. The ESPS Suite, a collection of multi-agency coupled weather and climate systems that complies with these interfaces, is a tangible outcome of this coordination. It is a direct response to the recommendations of a series of National Research Council and other reports recommending common modeling infrastructure. Integration of the HYCOM model with CESM contributes to development of the ESPS, and prototyping of interactive and interactive multi-model ensembles looks to a future where experimental model configurations are more easily assembled.

RELATED PROJECTS

Optimized Infrastructure for ESPC, funded by ONR, is a one-year seed project for this effort. Partners also include projects under the program *Advancing Atmosphere-Ocean-Land-Ice Global Coupled Prediction on Emerging Computational Architectures*, described here:
<http://coaps.fsu.edu/aoli/projects>

REFERENCES

AOLI: <http://coaps.fsu.edu/aoli>
CESM: <http://www2.cesm.ucar.edu/>
CoG: <http://earthsystemcog.org/>
ESMF: <http://www.earthsystemmodeling.org/>
HYCOM: <http://hycom.org/>
MCT: <http://www.mcs.anl.gov/research/projects/mct/>
MOAB: <https://trac.mcs.anl.gov/projects/ITAPS/wiki/MOAB>
NUOPC Layer: <http://earthsystemcog.org/projects/nuopc/>