Strategic Plan for 2017-2020

Last updated: April 27, 2017

The ESMF Mission

- Earth system models that can be built, assembled, documented, and reconfigured easily, using shared toolkits and standard interfaces.
- A growing pool of Earth system modeling components that, through their broad distribution and ability to interoperate, promotes the rapid transfer of knowledge.
- Earth system modelers who are able to work more productively, focusing on science rather than technical details.
- An Earth system modeling community with cost-effective, high performance, shared infrastructure and many opportunities for scientific collaboration.
- Accelerated scientific discovery and improved predictive capability through the social and technical advances ESMF represents.

1 INTRODUCTION

The Earth System Modeling Framework (ESMF) is community-developed software for building and coupling Earth system modeling components.[1] It includes a superstructure for representing model and coupler components and an infrastructure of commonly used utilities, including grid remapping, time management, metadata management, I/O, and data communications. It is governed by a set of partners that includes NASA, NOAA, the U.S. Department of Defense, and the National Science Foundation.

ESMF can be used in multiple ways: 1) to create component-based coupled modeling systems, 2) as a source of libraries for commonly used utilities, 3) as a file-based offline generator of interpolation weights, and 4) as a Python package for grid remapping.

The way that ESMF is used to create component-based modeling systems has evolved over time. The original ESMF library provided component interfaces, data structures, and methods with few constraints about how to use them. This flexibility enabled it to be adopted by many

[1] See the ESMF home page: http://www.earthsystemcog.org/projects/esmf

coupled modeling systems, but it limited the interoperability across these systems. To address this issue, the National Unified Operational Prediction Capability (NUOPC) consortium developed a set of coupling conventions and templates called the NUOPC Layer.[2] The NUOPC Layer has been or is being adopted by most of the organizations that have implemented ESMF architectures, forming a pool of compliant applications called the Earth System Prediction Suite (ESPS).[3] The NUOPC Layer and ESPS are described in Theurich et al. 2016.[4]

Looking forward, the main challenge for the ESMF project is to anticipate the needs of the applications that it supports, so that it is ready with solutions. The ESMF team and its partners are preparing for the future by developing strategies for managing the increasing diversity, complexity, and interdependence of modeling systems, providing solutions for component hierarchies and concurrently running physics parameterizations, and exploring how programming and optimization approaches for emerging fine-grained architectures are best applied to multi-component systems. Developing a comprehensive training program is another area that will be increasingly important in the next few years.

2 BACKGROUND

The first phase of the ESMF project began in 2002 with a 3-year NASA award to a team representing major U.S. modeling centers, and culminated with the delivery of a prototype framework and a set of newly coupled demonstration applications. Also during this period, the NASA Global Modeling and Assimilation Office (GMAO) created the GEOS-5 atmospheric general circulation model, a wholly new model that used ESMF extensively throughout.

The second phase of the ESMF project began as other agencies and programs decided to adopt the framework. These included the DoD Battlespace Environments Institute (BEI), the NASA Modeling Analysis and Prediction (MAP) Program for Climate Variability and Change, the NOAA Environmental Modeling System (NEMS), and a number of smaller projects in related domains such as space weather and sediment modeling. These efforts introduced additional requirements, both technically and organizationally. The grid software delivered with the initial ESMF prototype was completely redesigned to accommodate unstructured meshes. The composition of the core team changed substantially as well, in order to bring in staff with greater expertise in mathematics, numerical methods, and the physical sciences. The management structure of the project was reworked and formalized in order to accommodate multi-agency sponsorship. This organization is described in the ESMF Project Plan.[5] The second phase culminated with the delivery of a production quality, portable, scalable, software package, ESMF 5.2.0r, in 2011.

---

[2] See the NUOPC Layer home page: https://www.earthsystemcog.org/projects/nuopc/
[3] See the ESPS home page: https://www.earthsystemcog.org/projects/esps/
[4] Theurich, G. and co-authors, 2016. The Earth System Prediction Suite: Toward a Coordinated U.S. Modeling Capability, *Bull. Amer. Meteor. Soc.* 97, 1229-1247.
[5] ESMF Project Plan, https://earthsystemcog.org/site_media/projects/esmf/paper_1004_projectplan.pdf

The emergence of the NUOPC Layer marked the third phase of ESMF. NUOPC, a consortium of operational weather prediction centers and their research partners, aims to promote research to operations through interoperability standards. This group recognized that ESMF could be implemented in a number of different ways, resulting in modeling systems that were ESMF compliant but not interoperable. The NUOPC Layer adds additional rules about how ESMF components interact and increases their interoperability. It addresses the specification of build dependencies, the standardization of initialization phases, field matching, and a compliance checker. ESMF 7.0.0, released in 2016, contained a mature version of the NUOPC Layer software. The ESPS is the collection of modeling components and applications from NOAA, Navy, NASA, NCAR, and other modeling centers that use the NUOPC Layer conventions.

2.1 Outcomes of the Previous ESMF Strategic Plan

The previous ESMF Strategic Plan, from 2012-2015[6] focused on five goals:

- To complete a stable, high-quality, complete ESMF product, which includes software, documentation, and a support function. This was achieved.
- To develop strategic partners and technological connections that enable ESMF applications to operate in increasingly heterogeneous environments and multiple contexts. To this end, the ESMF team and partners introduced options that enable ESMF components to operate as web services, created prototype links to frameworks in other domains such as hydrology, and built ESMPy, the widely used Python interface to ESMF grid remapping.
- To support the continued adoption of ESMF by modeling applications. When the previous 2012-1015 Plan was written, ESMF had about 4000 downloads. That number is currently at about 7000 downloads. ESMF is now used in most major climate and weather models in the U.S., and has a growing user base world-wide.
- To build a common model architecture (CMA) that increases interoperability among ESMF applications. The CMA can be defined as an emerging similarity in the native coupling architectures at major U.S. centers. This similarity has enabled these centers to implement NUOPC Layer wrappers around their existing model codes without undue difficulty. This goal was completed with the delivery of a production version of the NUOPC Layer software in 2016.
- To create self-documenting models and support end-to-end workflows. Increasing scrutiny of model outputs and derived products, the widespread use of ensembles, and the growing complexity of individual models make careful documentation of model provenance especially critical. Progress was made toward this goal via advancements in the “Attribute” or metadata class in ESMF, including the ability to extract grid and other information from ESMF data structures and write it out using standard schema.

---

[6] See the ESMF Strategic Plan for 2012- 2015:
https://earthsystemcog.org/site_media/projects/esmf/plan_1202_esmfstrat.pdf

3 KEY CHALLENGES, GOALS, AND STRATEGIES

Enable ESMF, ESMPy, and the NUOPC Layer to remain viable through maintenance, user support, and ongoing development and optimization in response to new computational and science requirements.

The continued viability of ESMF and related products requires that they be ported to new computing platforms, updated for emerging technologies and new versions of supporting libraries, augmented to address new scientific and predictive challenges, and supported by a dedicated team. Continued optimization of coupled ESMF-based applications on both current systems and emerging fine-grained architectures is a high priority. The ESMF Virtual Machine class, which is an abstraction of the underlying hardware, was updated over the last three years to recognize the presence of accelerators, to enable components to be mapped to heterogeneous computing resources (e.g. systems with both GPU and CPU processors), and to perform a basic negotiation for resources. An objective over the period of this *Plan* is to encapsulate and extend this work in a new ESMF mapper class, which will assist modelers in determining how best to map multi-component simulations to computing hardware.

Upcoming functional development will include additional interpolation methods and chunking techniques for handling increasingly high resolution grids. The ESMF team is also committed to enabling models to become more “self describing,” and continues to improve metadata handling.

Extend the CMA and NUOPC Layer for component hierarchies, nested domains, and ensembles, and ensure interoperability with MAPL

The Common Model Architecture (CMA) defined during 2012-2015 in collaboration with the NUOPC consortium was limited to a set of high level components, such as atmosphere, ocean, ice, and wave models. However, there is increasing interest throughout the community in extending the CMA to include additional types of components, and implementing component hierarchies and regional-in-regional and regional-in-global nested domains. There is also interest in extending the CMA to accommodate research directions in ensembles, including interactive ensembles (Kirtman et al. 2009, 2011).

Extending the CMA will leverage the work of community partners. In particular, the NASA modelers that constructed GEOS-5 employed a hierarchical paradigm based on ESMF, adding conventions for how such components interact in a Modeling Analysis and Prediction Layer (MAPL). MAPL supports hierarchies of 50+ components, and will be used as the design basis for NUOPC hierarchies. An objective over the period of this *Plan* is to ensure that the NUOPC Layer and MAPL are merged to the extent possible and are interoperable. NUOPC supports a greater variety of options than MAPL (e.g. options for component concurrency, more general grid remapping), and the idea is to produce a merged capability that provides the best features of both.

Support the continued adoption of ESMF component interfaces by modeling applications and the expansion of an Earth System Prediction Suite (ESPS)

ESMF-based interfaces are now implemented in most major climate and weather models in the U.S. It has a growing user base world-wide and in related communities such as space weather. However, opportunities remain for more extensive use even at centers that are already using parts of it. One key opportunity is the implementation of a new Community Earth System Model (CESM) coupler based on NUOPC Layer interfaces, work that is currently underway. The resulting coupler will be fully featured and open source, and will inherit a large user base, including university collaborators. It has the potential to evolve into a community coupler, a new concept for U.S. modeling, which could be shared by multiple centers or agencies. Other key opportunities are the continued development of the NUOPC Layer-based unified modeling system at NOAA, and of Navy global and regional coupled systems.

In order to increase awareness of the components available with ESMF-based interfaces and interaction among participating centers, the ESMF team and its collaborators introduced the Earth System Prediction Suite (ESPS) during 2012-2015. The ESPS is a set of components with NUOPC-compliant interfaces that satisfy a set of usability criteria, including associated metadata and standard tests. A strategic goal is to expand the ESPS by adding new instances of components, adding new component types, and increasing conformance of all ESPS components with usability criteria. A related objective is to simplify the process of bringing components into ESPS. Strategies include making it easier for component developers to implement NUOPC Layer interfaces and providing standard documentation templates for those interfaces.

Leverage ESMF software for disciplines related to weather and climate, including data assimilation and hydrology.

ESMF has been successful in leveraging the core capabilities of the framework and creating ancillary products such as ESMPy that appeal to other disciplines and communities (in ESMPy's case, the Earth system data analysis and visualization community). A strategic goal is to continue to do this for other disciplines. Data assimilation is one such area. At its inception, ESMF included both modeling and data assimilation teams, and offers a range of supporting data structures and tools. Emerging community data assimilation projects such as the Joint Effort for Data assimilation Integration (JEDI) offer an excellent opportunity for the ESMF to continue developing its data assimilation capabilities. From the extension of existing constructs such as mesh and location streams to specific regridding methods relevant to data assimilation, the ESMF can provide a robust foundation for advancing the US data assimilation infrastructure.

Hydrological modeling applications offer additional opportunities for leveraging ESMF software. The new National Water Model, which uses WRF-Hydro as the central forecast component, is now operational and plans suggest that it will introduce coupling to a range of different types of components as it evolves. NASA, NOAA, and Navy centers are currently implementing coupled systems that include land surface and hydrologic components, using NUOPC Layer-based coupling. In all three projects, the land surface component is NASA's Land Information System (LIS) and the hydrologic component is WRF-Hydro. These projects seek to improve representations of the hydrologic cycle by establishing physically consistent feedbacks. Coupling requirements include support for nested, high-resolution domains and communication between nests, and the NASA project explores coupled data assimilation using a NUOPC Layer interface for the data assimilation component. Building on these projects, an objective is to understand and anticipate future software requirements related to hydrologic modeling. This includes pre- and post-processing needs as well as coupling requirements. It will be critical to continue to combine ESMF capabilities with GIS formats such as shapefiles, so that modelers can easily move and remap data between local phenomena expressed using structures like catchment basins, and the regional and global processes represented using gridded modeling systems.

Develop a comprehensive ESMF and NUOPC Layer training program, leveraging usability tools.

Adoption of ESMF and the NUOPC Layer across multiple agencies has increased the need for model user and developer training. A goal is to improve materials and expand an in-person training initiative started in 2015-2016, during which several hundred people at NOAA and the Navy received instruction on ESMF and the NUOPC Layer in a series of face-to-face, one-day sessions. In addition, a regular schedule will be established for webinar-based training that can reach a broader national and international community. The current webinars can support an audience up to about 50, and have been fully subscribed each time they were offered.

New tools are helping to make learning about ESMF and the NUOPC Layer easier and more appealing, and these will continue to be improved and integrated into training. They include the Component Explorer, which acts as a generic driver for any NUOPC-compliant model, outputting key information about the component, the Compliance Checker, which produces verbose output in the ESMF log files and reports on compliance issues, and the Cupid Integrated Development Environment, a plugin for the widely used Eclipse package. Cupid offers a friendly interface through which developers can examine code structure, generate NUOPC-compliant code automatically, and identify compliance issues.[7]

4 STRATEGIC VISION

At the end of the period covered by this Strategic Plan, ESMF and the NUOPC Layer will be the primary coupling infrastructure in major U.S. modeling systems, including the Community Earth System Model (CESM). The NUOPC Layer and MAPL will have merged to the extent possible, and their components will be interoperable. Collaborations with the JEDI data assimilation effort and the National Water Model will be well defined. Documentation of the pool of Earth System Prediction Suite components, including information about code access and testing, will be improved and publicly available. The new ESMF mapper capability and related optimizations for fine-grained architectures will start to be tested at a number of modeling centers, and a comprehensive training program, combining in-person and webinar options, will be established.

[7] See the Cupid home page: https://www.earthsystemcog.org/projects/cupid/