# Contents

Table of Contents i

(I) Science, Technology and Management 1

1 Introduction 1

2 Scientific Objectives and Rationale 2
  2.1 Objectives of the ESMF ............................................. 2
  2.2 Scientific Benefits .................................................. 2
    2.2.1 Climate Modeling and GCMs .................................. 2
    2.2.2 Numerical Weather Prediction ............................... 3
    2.2.3 Data Assimilation ............................................. 3
  2.3 Community Delivery ............................................... 4

3 Technical Plan 5
  3.1 ESMF Functionality ................................................ 5
  3.1.1 Scope .................................................................. 5
  3.1.2 Requirements .................................................... 6
  3.2 Architecture ........................................................ 6
    3.2.1 Strategy for Code Reuse and Interoperability ........... 6
    3.2.2 Strategy for Performance Portability and Ease of Use 7
    3.2.3 Preliminary Design ............................................. 8
  3.3 Framework Development Plan ..................................... 9
  3.4 Test Application: the ESMF Validation Suite ................. 10
  3.5 Implementation Strategy ........................................... 11
    3.5.1 Use of Existing Tools and Frameworks ................. 11
    3.5.2 Implementation Issues ...................................... 12
  3.6 Expertise in Scalable Grand Challenge Applications .......... 12
  3.7 Vendor Participation .............................................. 13

4 Management Plan 13
  4.1 Management Strategy ............................................... 13
  4.2 Investigator Team .................................................. 13
  4.3 Management Structure and Implementation Team ............ 14
  4.4 Qualifications ..................................................... 15

(II) Software Engineering Plan 16

1 Software Teams and Management 16

2 Software Process 16
3 Software Tools and Techniques

(III) References

(IV) Biographical Sketches

(V) Milestones, Schedule and Cost

(VI) Endorsement Letters

(VII) Education and Public Outreach Addendum
(I) Science, Technology and Management

1 Introduction

This is one of three linked proposals to construct an Earth System Modeling Framework (ESMF) and implement using it high performance, interoperable codes for weather prediction, climate simulation, and data assimilation. This proposal (Part I) addresses core development of the ESMF. Part II focuses on implementing modeling applications using ESMF and Part III on implementing data assimilation applications. The institutions proposing this work are the NSF National Center for Atmospheric Research (NCAR), DOE Argonne National Laboratory (ANL), DOE Los Alamos National Laboratory, Massachusetts Institute of Technology (MIT), NASA Goddard Space Flight Center Data Assimilation Office (DAO) and NASA Seasonal to Interannual Prediction Project (NSIPP), University of Michigan, NOAA Geophysical Fluid Dynamics Laboratory (GFDL) and NOAA National Centers for Environmental Prediction (NCEP).

The ESMF is a concerted response to several current challenges in science and technology: the growing scientific complexity of Earth system models and larger collaborative communities, the increased demand for improved weather and climate prediction capabilities, and the challenge of developing applications in a volatile computing environment. The ESMF will address the technical aspects of these issues, laying the groundwork for addressing the more difficult scientific aspects, such as the physical compatibility of components, in the future.

The ESMF will consist of a superstructure for coupling and exchanging data between component models (e.g., atmosphere, ocean) and model subcomponents (e.g., physics, dynamics); and an infrastructure consisting of data structures for representing grids and fields and an optimized, portable set of low-level utilities. The data constructs and low-level utilities will be used by the coupling superstructure and may also be used separately to compose scientific applications. Thus our work will promote software reuse as well as interoperability. A unique feature of our design is an integrated approach to data assimilation. The ESMF will be designed in an object-oriented, layered manner to isolate machine dependencies, and to offer an application programming interface natural for the Earth sciences.

The “critical mass” of institutions participating in this proposal has also been active in the Common Modeling Infrastructure Working Group (CMIWG). For the past two years, this group has explored ways of enhancing collaborative Earth system model development. The CMIWG has attracted broad participation from major weather and climate modeling groups in research and operational centers. At a CMIWG workshop in February 2000, the core group of scientific CoIs on this proposal committed to developing a coordinated response to this CAN. Since then, working with computer scientists and software engineers, the group has established requirements and a preliminary design for a common framework. Several of our CoIs are also key participants in the Accelerated Climate Prediction Initiative (ACPI) Avente Garde project, a broadly based DOE-centered effort focused on restructuring the coupling and atmospheric component of the NCAR Community Climate System Model (CCSM). This work will be coordinated with ESMF development. Multi-agency and multi-institutional alliances of the sort represented by the CMIWG and collaboration on these ESMF proposals are a key recommendation of recent reports that discuss the deficiencies in our national climate modeling program [11, 34].
2 Scientific Objectives and Rationale

2.1 Objectives of the ESMF

One of the great strengths of atmospheric, oceanic and climate modeling in the U.S. is the variety, availability and wide use of models. But this diversity has also led to duplication of effort, a proliferation of models and codes which, due largely to technical reasons, cannot interoperate and have been unable to keep up and exploit advances in computing technology.

The specific objectives of the ESMF are to 1) Facilitate the exchange of scientific codes (interoperability) so that researchers may more readily interface with smaller-scale, process modeling efforts and can share experience among diverse large-scale modeling efforts; 2) Promote the reuse of standard, non-scientific software, the development of which now accounts for a substantial fraction of the software development budgets in many institutions; 3) Focus community resources to deal with changes in computer architecture; 4) Present the computer industry and computer scientists with a unified, well defined and well documented task to address; 5) Share the overhead costs of the housekeeping aspects of software development, such as documentation; and 6) Provide greater institutional continuity to model development efforts, by distributing support for modeling infrastructure throughout the community.

2.2 Scientific Benefits

In this section we describe ways in which the ESMF will increase scientific productivity and encourage new research in a range of Earth science domains. These include climate modeling and general circulation modeling, numerical weather prediction, and data assimilation.

2.2.1 Climate Modeling and GCMs

The ESMF will provide ready-made solutions to standard climate model issues such as grid representation and transforms, utility software and high-level scheduling. Standard interfaces will help keep the diverse community of developers coordinated.

Climate models are increasingly being used to guide policy decisions. The predictive requirements are becoming more stringent and data assimilation a crucial issue. The demand for interoperability of climate model components has intensified as growing evidence of anthropogenic climate change has focused scrutiny on the capabilities of the current generation of climate models. Without exchangeable model components, it is often difficult to point to a particular component as a clearly identifiable cause of divergent results when one model is compared against another or against the observations. The ESMF will facilitate just this sort of analysis.

The computational demands of coupled climate models necessitate the use of scalable parallel computational platforms. To effectively utilize such platforms codes must be based on flexible data structures and communication libraries that are performance portable. The ESMF will provide this infrastructure.

The most resource-intensive components of climate models are typically oceanic and atmospheric general circulation models (GCMs). Within a GCM physics and dynamics can be treated as separate coupled components themselves. This allows researchers to readily
exchange dynamical cores, thus offering the opportunity to explore more efficient algorithms. NSIPP and GFDL models are already structured this way using the GEMS [29] and FMS [16] frameworks, respectively. Plug-compatible dynamical cores for the the atmospheric portion of the NCAR CCSM are under development.

The climate models that will be implemented using the ESMF in Part II are the NCAR CCSM [9], the MIT general circulation model (MITgcm) [18], and climate models implemented under the GFDL FMS and the NSIPP GEMS frameworks.

### 2.2.2 Numerical Weather Prediction

Operational forecasting centers have strict requirements for performance and reliability. By specifically optimizing parts of the ESMF, some of the hand-tuning optimization burden at operational centers can be shared with the rest of the modeling community. Also, sharing utility code for error and signal handling, and having the whole community develop and test code, will contribute to the robustness of code overall.

The forecasting code we will use as a testbed for ESMF is the NCEP atmospheric global forecast code [21, 22, 23, 39]. It is a key component of the NCEP Global Data Assimilation System (GDAS) that provides the backbone of all numerical weather prediction at NCEP. It also is used to make the 4 times per day 120-hour Aviation forecast, the daily 384-hour Medium Range forecast and the 22 times per day 384-hour Ensemble forecasts at NCEP.

### 2.2.3 Data Assimilation

As in the modeling community, analysis software for data assimilation systems has been largely developed in isolation, making it difficult for groups to take advantage of activities taking place outside their home institutions. Transition of assimilation technology to the rigorous, stable standards of the operational environment is particularly challenging. The interoperability afforded by ESMF will enable the sharing of high level components such as on-line quality control systems, minimization algorithms and management of observational databases.

The data assimilation applications that will be implemented using ESMF in Part III span a variety of dynamical models and assimilation algorithms. Two different global 3D-VAR atmospheric data assimilation systems used operationally at the NASA DAO and NCEP [13, 14, 25, 31] will be implemented under ESMF, affording a great degree of interoperability between the two centers. Ocean products are only now emerging into the quasi-operational arena, so the inclusion of ocean state estimation tools within the same framework as atmospheric state estimation tools will be beneficial to speed development, portability, and interoperability. Two ocean data assimilation systems from NSIPP, the optimal interpolation and the Ensemble Kalman filter [28], will be implemented under the ESMF, as will a complementary approach from MIT. The MIT effort will develop the infrastructure under ESMF for 4D-VAR ocean data assimilation using the MITgcm and their tangent-linear and adjoint model compiler [17].
2.3 Community Delivery

The ESMF will be developed by and delivered immediately to a significant portion of the U.S. climate, weather prediction and data assimilation communities, and will be superbly positioned to migrate to closely related disciplines.

The variety of groups, institutions and disciplines that will implement applications using ESMF in this proposal set constitutes delivery to a "critical mass" of the U.S. climate and weather prediction communities. Our Investigator Team includes representatives of all the relevant disciplines and this proposal has the endorsements of the CMIWG leadership and most major modeling and data assimilation centers in the country. These groups include the following: the NASA/GSFC DAO and NSIPP, both central to NASA's mission to support data acquisition, dissemination, analysis and assimilation into models; GFDL, one of the world's leading academic and climate modeling centers; MIT, with expertise in model development, adjoint modeling, and innovative cluster technologies; NCAR, the internationally renowned climate modeling and research center, and home to the extensively used CCSM; and NCEP, the primary numerical weather prediction center in the US, with vast experience in the development of models and data assimilation systems.

Many Earth science application developers will adopt ESMF through relation to our Investigators' activities. This includes additional groups at participating institutions, potential collaborators who wish to interoperate with our applications, and groups active in the CMIWG. The NCAR CCSM model alone offers several key mechanisms for wider community support. We will promote the ESMF at the yearly CCSM workshop, a widely attended, international forum for the discussion of climate issues. The CCSM is also the testbed for the DOE/NCAR ACPI Avante Garde project, a current effort to restructure the atmospheric component and coupling tools of the CCSM for greater efficiency and modularity. In addition to several of the institutions represented on this proposal, the ACPI project includes researchers and computer scientists from Oak Ridge National Laboratory, Lawrence Livermore National Laboratory, and Lawrence Berkeley Laboratory. We have been coordinating our ESMF effort with them closely and will continue to do so, through Investigators who are participating in both ACPI and this proposal (Boville, Jones, DeLuca, Larson).

Investigators Stout, DeLuca and Killeen are key participants in several space weather community efforts to construct a space weather modeling framework that will leverage the ESMF. Compatible Earth/space frameworks will facilitate projects such as the University of Michigan NSF Knowledge and Distributed Intelligence (KDI) project [12], which is combining space weather models with Earth system models to study Sun-Earth interactions.

The ESMF effort will draw in varied groups who are already using framework-friendly design strategies and may wish to influence and leverage our work. For example, there are likely to be opportunities for the ESMF activity to interact with a group at Colorado State University developing a modular Terrestrial Carbon Model under the NSF Integrated Research Challenges Program.

Anyone with an interest in Earth system modeling will have access to the ESMF source code and documentation via the web. We will encourage moderated open source development after a prototype version of the ESMF has been released.

We are investigating the possibility of expanding the CMIWG website, which includes a soft-
ware repository for plug-in column physics parameterizations, to include ESMF distribution and a repository of scientific components for Earth system modeling.

*A workshop is planned at the end of the proposed work to NASA to introduce the ESMF to a broad community.* NCAR has committed to an ongoing support role for ESMF that will include the coordination of regular community workshops.

Finally, *we plan to present our work at conferences and publish our experiences and results in technical and Earth science journals.*

### 3 Technical Plan

The community collaborating on this proposal has already made significant progress in defining the structure of the ESMF. In August 2000, participating groups met at NCAR to review strawman documents each had prepared describing their application requirements, an ESMF scope and architecture, and implementation strategies (see http://www.scd.ucar.edu/css/NASACAN.htm). Here we describe our preliminary conclusions: the appropriate scope, design and implementation of the ESMF.

#### 3.1 ESMF Functionality

##### 3.1.1 Scope

The scope of the initial ESMF addresses two critical needs: 1) robust, optimized, non-scientific *infrastructure* libraries with which to build models and model subcomponents and 2) a *superstructure* for coupling scientific components. The first of these promotes code reuse, the second, interoperability. GUIs, optimized math libraries for purposes other than regridding, and differential operators are secondary needs and are *not* part of the initial ESMF proposed here. These functions can be introduced to the ESMF in the future.

The coupling superstructure will perform highly optimized regridding, interpolation and communication of gridded, distributed data. The data may represent multiple fields or a single field, may be in the same or different executables, may be in code segments executing serially or concurrently, and may be distributed among nodes and/or partitioned among multiple threads. The interfaces for components and couplers must be natural for our problem domain.

The software necessary to support the above capabilities includes infrastructure for describing a wide variety of grids and decompositions, and for performing high-level manipulations of fields discretized on those grids. The software for specifying decompositions will interface to a mechanism for performing dynamic load balancing. Operations on grids and fields must implement corresponding methods for the construction of tangent linear and adjoint models.

Both the coupling mechanism and application codes will use common utility routines. ESMF utilities will include communication libraries, I/O, performance profiling, time management, and signal and error handling.
3.1.2 Requirements

In addition to achieving the critical ESMF goals of code interoperability and reuse, the ESMF software will conform to a set of functional requirements. These include:

*Performance portability.* Portability and computational efficiency over a wide range of platforms are essential. Optimized performance on scalable architectures for moderate numbers of processors (16-500) is the highest priority.

*Flexible usage.* The application writer must be able to choose how much or how little of the ESMF to use. For example, some application developers may want to use ESMF for coupling models while others may want to build component models using lower-level tools provided by ESMF.

*Ease of use.* The ESMF must adopt a straightforward approach to implementing our large and complex codes using the framework, and must have a natural interface for our problem domain.

*Extensive grid support.* ESMF must be able to couple components that are discretized on: logically rectangular grids (including bipolar and cubed-sphere grids); reduced (cut-out or Kurihara) grids; unstructured grids; phase space grids (e.g., spectral, Fourier); nested and adaptive grids; and icosahedral grids. It must be simple to add additional grids. We also require support for describing lateral grid boundaries, masked regions and halo regions.

*High performance, extensible, multi-format I/O.* The ESMF utilities must support a generic interface for I/O of self-describing data in netCDF, binary, GRIB, BUFR and EOS HDF data formats. There must be support for high-performance parallel I/O.

*Multiple language bindings.* ESMF utility and coupling software must be usable by applications written in C/C++ and F77/F90.

*Other requirements.* These include robust error and signal handling, runtime configurability, and an efficient, low maintenance implementation (e.g., auto-documentation from code).

3.2 Architecture

We begin by describing our strategies for achieving the major design goals of the ESMF: interoperability, reuse, performance portability and ease of use. An object-oriented design (OOD) [38] and software layering are our basic tools.

3.2.1 Strategy for Code Reuse and Interoperability

OOD enhances code reuse since the class structure encourages well-defined, general purpose, encapsulated code segments that can work in a variety of contexts. Both ESMF parallel utility classes and ESMF procedural system-level utilities will be used repeatedly in a variety of more complex classes in the framework and can also be used to compose many scientific applications. We will follow the model of the NWChem/ECCE [30] framework, which offers a useful separate toolkit.

We are concerned with several levels of interoperability in ESMF. The most fundamental level involves rewriting Earth science applications in a modular manner, and defining a standard interface for coupling based on the constructs and terminology natural for our problem domain.
We will address this level of interoperability by building the “inheritance” feature of OOD into the ESMF design. Classes that inherit a core set of attributes and operations can be queried and can communicate with other classes in a standard way. A second key feature of OOD is polymorphism, the ability to create a variety of interfaces and implementations for the same operation. This enables different classes with related functions to be handled generically. Together polymorphism and inheritance promote interoperability (see [15]). We note that inheritance and polymorphism can be implemented in C, C++ or F90, with varying ease.

In practice, certain classes within an application will be supplied by the ESMF and others (mainly scientific components) will be customized by the application developer. For example, an ocean model component of a climate model might need to contain a “getGrid” method that returns a description of its data grid and distribution. However, this method might not have the same arguments as a method returning equivalent information from a land model component (polymorphism). Coupling tools supplied by ESMF would call the “getGrid” method to understand how data should be routed to and from each type of component.

A secondary level of interoperability is the purely mechanistic issue of language compatibility. There are many options for handling this. First, using pointers to pass “opaque” data structures across language boundaries, and relying on accessor methods to return information about those data structures, is common in C-based packages (e.g. netCDF [35], PETSc [8]) Second, calling Fortran routines from C++ frameworks in order to use high-performance numerical kernels is standard practice (e.g. SAMRAI [19], KeLP [7], DAGH [27], others). Here the C++ data structures typically do not extend into the Fortran code. Climate and weather codes can be structured for this type of interoperability; for example, the Weather Research and Forecast Model (WRF) [26] has a “mediation” layer that breaks down data structures from a top-level “driver” layer and passes the decomposed information to a high-performance “model” or kernel layer. A third option is executing components via an interoperability environment, such as CCA [5] or Cactus [4]. Here multi-language components utilize a new, common language or mechanism such as CCA’s Scientific Interface Definition Language (SIDL) in order to interoperate in a very generic environment. CCA is the most appealing mechanism of this type since it is being developed by a very large community as a standard.

We intend to study the strategies above and others further and present the work in an Implementation Report that will extend the work already done by the “ESS Integrator.” The study will include prototyping key ESMF functions using different approaches to see if there are significant efficiency, ease of use, or other factors, as well as addressing issues such as compiler quality, utilization of existing staff, and ease of hiring new staff.

A third type of interoperability is the ease with which a given framework or toolkit can combine with others. We cite the excellent example of PETSc, co-developed by CoI Smith, which interoperates efficiently with a multitude of other packages. This is accomplished largely through careful treatment of data as it is brought into PETSc and an interface that makes minimal demands on the user, both of which we regard as sensible strategies.

### 3.2.2 Strategy for Performance Portability and Ease of Use

We will achieve performance portability and ease of use by layering code and by using generic, carefully designed interfaces.
Like object classes, layers help to organize code, though at a larger scale. The ESMF software will consist of five distinct software layers, each with well-defined functions and interfaces. The “lowest” layers are non-scientific and general in function. The “highest” layers describe functions that are more specific to the Earth system domain. Methods in higher layers call methods from lower layers, and classes in higher layers are often composites of classes found in lower layers.

Machine-dependent code will be isolated to the lowest level in the framework by wrapping it in a generic interface. This improves performance portability by reducing the scope of changes that need to be made to run on or optimize for a new architecture, and allowing us to optimize primitive operations by using vendor-specific libraries.

Likewise, calls to higher-level communication functions such as transposes are isolated to a layer in the framework so that an application developer does not need to manage the details of distributed data transfers. This makes the application code easier to write and use.

3.2.3 Preliminary Design

We describe here a layered ESMF architecture that will fulfill the requirements laid out in Section 3.1.2. A complementary description of ESMF functionality from an application developer’s viewpoint is provided in Part II. The Part II ESMF description focuses on the data transformations necessary to support modeling applications. In this Section we focus on the structure of the framework itself.

The five layers in the ESMF architecture are shown in Figure 1. Since the middle three layers are closely related, we cluster them together and refer to them collectively as “Fields and Grids.”

At the bottom are low-level utilities. These may be machine-dependent and may be coded procedurally for efficiency. Some examples of functions provided in this layer include: calendar management and alarms, lightweight performance profiling, and methods to support disk I/O. It is essential that these utilities be easily usable to compose applications independently.

The second layer is a set of parallel utility classes. This layer includes a retrievable description of a machine model, and a layout class that specifies the portion of the machine model over which a data object is distributed. An example of a method in this layer is the specification of a simple topology for a given layout.

Classes representing distributed grids are in the third layer, which contains a substantial portion of the ESMF. Distributed grids contain a layout that describes the grid decomposition, and a grid specification that describes the grid coordinates and connectivity. The methods here include index-space and physical-space methods. Examples of index-space methods include data transposes on logically rectilinear grids. Physical-space methods are those which additionally require knowledge of the spatial metric of the distributed grid: these include more complex regridding methods, such as computation of grid overlaps.

Fields are in the fourth layer. Fields contain a distributed grid and a field specification that describe attributes related to the physical field (“metadata”). We will also support a field set class, for fields that are discretized on the same distributed grid, since this is often the form of data communicated between model components. Operations in this layer will include using field metadata information to write header information for the self-describing
data formats of Section 3.1.2.

The **coupling and components** layer includes large scale components, such as atmosphere and land models, and the classes used to simplify the transfer of data between them; for example, **boundary state vector** classes that comprehensively describe the portion of a component model state that is necessary for coupling. This layer also includes high-level control interfaces for scheduling different modes of component execution.

![Figure 1: ESMF Layered Architecture](image)

We show the application developer’s view of the ESMF in Figure 2. The figure includes both numerical components and core framework software, shaded grey, which is used by the components. Components operate under the direction of a high-level control program, whose role is to orchestrate a sequence of computation.

### 3.3 Framework Development Plan

ESMF development will involve intensive interaction between a core Implementation Team located at NCAR and a dispersed set of Earth scientists, computer scientists and software engineers located at collaborating institutions. Together these participants will define an **open standard** for ESMF software component interfaces - an evolving standard that will be open to members of our entire community to extend. The evolution of the standard from an interface design into a production-quality software package by the Implementation Team is the bulk of the work in this Part I proposal.

Early on in the project the Implementation Team will undertake a number of preparatory activities. They will work with application developers to create a validation suite from representative codes; they will collaborate closely with computer science CoIs and the ESMF Integrator to perform an implementation study (see Section 3.2.1); and they will begin work on creating a set of low-level utilities so that these can be used in the development of more complex classes.

Development on the core ESMF will proceed as detailed in the Software Engineering Plan. An appropriate sequence for class development will be identified and documented. As classes are completed and unit tested, they will be integrated into an evolving prototype of the ESMF. This work will be accompanied by efforts on the part of application developers to incorporate preliminary prototypes of the ESMF into their codes. We plan to use the existing frameworks GEMS and FMS to assist with rapid prototyping of implementation.
strategies. As the ESMF itself is implemented, both the application groups and Implementation Team will coordinate with the ESS Evaluation Team to undertake exhaustive testing and performance studies and will communicate their results to the wider community.

3.4 Test Application: the ESMF Validation Suite

The application that this Part I proposal will focus on is a synthetic ESMF Validation Suite (EVA) consisting of a set of representative segments from the Earth system codes described in detail in Parts II and III. EVA will evolve, be maintained, and be distributed with the framework, and will provide a straightforward way for developers and eventually installation sites to test and assess it. It will also serve as a focus to engage the computational science research community to boost the performance and scalability of our codes.

EVA performance and functionality will be baselined using the ESS Testbed (Milestone E). The metrics will be time to solution, scalability, and functionality in several areas: mode of execution (SPMD/MPMD), number of distributed grid types supporting optimized operations, an integrated hybrid (distributed/multithreaded) programming model, and support for C++ and F90 components. These metrics reflect both practical performance issues and an awareness of current trends. Milestone F focuses on functionality while Milestone G focuses on performance and scalability; this is consistent with the standard software engineering practice of “get it right before you make it faster,” see for example [37]. The interoperability of EVA codes with the application codes in Parts II and III will be demonstrated in Milestones I and J. We will run on a variety of platforms throughout development, and the portability of the ESMF and EVA will be verified in Milestone K.

Composition of the Validation Suite

Based on the prototype framework design presented in Section 3.2.3, and initial dialogue among our Investigator Team we anticipate developing a suite organized into the following sub-areas:

Synthetic Components. Test components such as stand-alone solvers for Helmholtz problems, simple shallow water models, and standalone river runoff simulations have proven useful at GFDL, MIT and GSFC, and will be included in EVA. The set of synthetic components will be instrumented to allow performance and correctness to be evaluated rapidly. The components will test parallel primitive operations, including exploring the performance of primitives bound to different software and hardware technologies (e.g., IPC, Quadrics, Myrinet). The components will also serve to refine an ESMF machine model that can map efficiently to the diverse mix of parallel machine architectures the ESMF will target. We will use preliminary EVA components to test language interoperability options as part of the implementation study described in Section 3.5.2.

Synthetic Drivers. The test suite will also include a number of EVA drivers to test framework support and compatibility with different modes of execution (SPMD/MPMD executables; sequential/concurrent sequencing). The synthetic driver tests will verify that the framework works reliably in all of these cases.

Synthetic Couplers Tests using synthetic EVA couplers will check correctness and the performance and scaling of the ESMF interpolation and regridding tools. The EVA suite will validate treatment of aspects of Earth science component coupling such as subgrid scale variations, ungridded observational network datasets, and enforcing global conservation
constraints. We will also create test codes that exercise the framework support for temporal accumulation and averaging. These tests will provide clear examples of how a component interfaces with a coupler.

**Interoperability Tests** Other framework functions that support interoperability will also be tested with the EVA couplers. These functions include mechanisms to allow one component to query parameters of another component, mechanisms to manage the representation of time and notification of periodic events within components in a standard way, and mechanisms to coordinate synchronized termination and subsequent resumption of multi-component, multi-CPU experiments. This work will also validate interoperation with software external to the framework.

### 3.5 Implementation Strategy

#### 3.5.1 Use of Existing Tools and Frameworks

Frameworks with a significant subset of the capabilities desired for ESMF have been co-developed by members of our Investigator Team (FMS, GEMS, STAPL). Each of these offers support for data parallelism within components plus a higher level of large-scale task parallelism for coupling components. Each of these also offers a set of data constructs and general utilities. STAPL is a C-based framework for composing radar applications consisting of multiple subsystems; it is notable for its hybrid programming paradigm and real-time performance. FMS, from GFDL, and GEMS, from NSIPP, are modular, F90-based frameworks developed for use within GCMs and for coupled climate modeling. They have achieved institutional support and are relied on for production modeling. However, they do not have all the flexibility sought by the broader Earth science community, as identified in the goals and requirements laid out for the ESMF. Specifically, these systems lack: a hybrid programming model (see Section 3.5.2); the ability to represent and perform efficient data transfers on nested grids; highly optimized regridding (especially for unstructured grids); the ability to couple multiple executables; integrated error and signal handling for complex multiple-executable applications; and extensive multi-language support.

The coupling capabilities of FMS and GEMS are being extended in the CCSM Next Generation Coupler (NGC), part of the current NCAR/DOE ACPI effort that involves several of our CoIs (see Section 2.3). The NGC will include optimized regridding for a wide range of grids and a SPMD/MPMD option.

It is tempting to consider using one of the tools or frameworks described above as the basis for the ESMF. However, we feel strongly that the ESMF standard should be a community effort. ESMF will represent a powerful synthesis of what we have learned from these previous projects.

Members of our collaboration have also investigated other frameworks, such as POOMA [36], Overture [10], and Cactus [4].

Much of the functionality in the C++-based POOMA and Overture packages lies in fine-grained mathematical operations. We find these frameworks inappropriate models for ESMF since data parallelism is inefficient for many of our multi-component applications, and by writing application codes in Fortran or C we retain good performance.

The term “framework” can also refer to an effort such as the CCA project or Cactus, which provide tools for component coupling and language interoperability. These tools may
be useful for ESMF. However, language interoperability is a relatively minor aspect of the interoperability we are concerned with, and there are alternative solutions which may be more appropriate (see Section 3.2.1).

We find that none of the frameworks we have examined are focused on our central needs, as described in Section 3.1.1. Therefore we would seriously question whether trying to modify an existing framework would benefit us more than adapting its more useful features into a wholly new ESMF implementation. Also, while we recognize the advantages of leveraging code from other groups, we feel that it is essential that the Earth system community assume significant control over the framework that our codes are built on. We cite as cautionary examples cases of frameworks that attempted to be general for whom support has largely vanished, such as POET [6] and POOMA. It seems that the more successful frameworks and problem solving environments (for example, the NWChem/ECCE effort) are those that are quite domain specific and are developed and maintained by their own user communities. The success of existing framework efforts within our community, such as GEMS and FMS, indicates that this approach is both possible and appropriate.

We have identified a multitude of low-level utilities that could be extended or directly incorporated into the ESMF, such as documentation generators, timing libraries, and communication libraries. The SCRIP [20] package for conservative regridding, developed by CoI Jones, is especially relevant to our work. We are very open to investigating other packages.

3.5.2 Implementation Issues

Language: We anticipate implementing the low-level utilities of the framework mainly in C. An implementation study at the beginning of the proposed work will help to assess alternative approaches to developing upper levels of the framework (see Section 3.2.1).

Microprocessor architectures: Many of the climate and weather prediction codes in use today were developed for platforms based on vector processors rather than the dominant microprocessor-based architectures. The optimization strategies for these platforms can be radically different. A charge of the ESMF will be to make it easier for developers to obtain better performance on these platforms without changing significant amounts of application code.

Hybrid programming strategy: Clusters of shared memory systems are an increasingly prevalent platform, and the current trend is towards more processors per chip and per node. Including in ESMF a well-integrated capability to distribute data over nodes and computation over threads will position the framework to adapt to a range of future architectures.

3.6 Expertise in Scalable Grand Challenge Applications

Currently all participating institutions have highly optimized parallel implementations of at least some of their production codes. Figure 3 shows speed-up curves for a sampling of these applications. These were chosen to highlight the breadth of applications and platforms being used. Shown in the figure are results for both atmospheric (NCEP, NSIPP, CCM) and ocean (POP) models. Two of the models are spectral and two grid point; two are operational codes (NCEP and NSIPP) and two are research codes (CCM and POP). We show results on platforms from four of the leading U.S. manufactures: SGI, Cray, IBM, and Compaq. Most
of these models, however, run on multiple platforms. Additional results appear in Parts II and III.

![Graphs](image)

**Figure 3:** Speedup curves.

### 3.7 Vendor Participation

IBM, Cray Research, Sun Microsystems, Api Networks, and High Performance Technologies have expressed interest in partnering with us on the development of the ESMF by participating in design reviews, code reviews, and benchmarking.

### 4 Management Plan

#### 4.1 Management Strategy

The key challenge in creating a management plan for the ESMF is to entrain broad expertise in the framework’s development while ensuring that work can proceed efficiently, and that decisions can be made in an unambiguous manner. In order to accomplish this we will: 1) engage a broad spectrum of the Earth system modeling community in the specification of requirements and the overall design of the framework, maximizing expert input and user buy-in; 2) utilize groups with more focused interests to oversee the design and implementation of specific framework components in order to achieve timely, informed decisions; 3) delegate much of the work of design drafts, prototyping and production coding to a closely integrated, central team of software engineers; and 4) resolve inconsistencies and differences of opinion throughout the project by allowing final software engineering decisions to be made by a central software manager and selected “Oversight Team” leaders.

#### 4.2 Investigator Team

The Principal Investigator of the proposal is **Tim Killeen**, the Director of NCAR.
The CoInvestigators are:

**Jeffrey Anderson**, Head, Experimental Prediction Group, NOAA Geophysical Fluid Dynamics Laboratory; **Byron Boville**, Senior Scientist and Head, Climate Modeling Section, NCAR; **Arlindo da Silva**, Meteorologist, Data Assimilation Office, NASA/GSFC; **Cecelia DeLuca**, Software Engineer, Scientific Computing Division, NCAR; **Robert Johnson**, Director of Education and Outreach, University Corporation for Atmospheric Research; **Philip Jones**, Staff Member, Theoretical Fluid Dynamics Group at Los Alamos National Laboratory; **J. Walter Larson**, Assistant Computer Scientist, Mathematics and Computer Science Division, Argonne National Laboratory; **Stephen Lord**, Director, Environmental Modeling Center, National Centers for Environmental Prediction; **John Marshall**, Professor of Atmospheric and Oceanic Sciences, Massachusetts Institute of Technology; **Barry Smith**, Computer Scientist, Mathematics and Computer Science Division, Argonne National Laboratory; **Quentin Stout**, Professor, Computer Science and Engineering, University of Michigan; **Max Suarez**, NASA Seasonal to Interannual Prediction Project; NASA/GSFC.

The Investigator Team we have assembled possesses the required balance of scientific and computational expertise.

### 4.3 Management Structure and Implementation Team

As Principal Investigator, Dr. Killeen will serve as the primary contact and administrator of the proposed work. Dr. Killeen will negotiate agreements with NASA and among Investigator Team members, and will arrange for disbursement of funds after payment. He will supervise the overall activities of the Investigator Team and promote the ESMF project to the wider community. A software engineering manager (DeLuca) will oversee much of the day-to-day activity of the project.

Three **Oversight Teams** including the CoIs of the project will closely track and guide the design and implementation of the ESMF software. Each Oversight Team will consist of a mix of physical scientists, computer scientists, and software engineers, with some individuals on multiple Teams. The Oversight Teams correspond to different layers of the ESMF software (low-level utilities, fields and grids, coupling), and reflect different interests and expertise. For example, the Oversight Team for coupling will consist primarily of physical scientists, since the coupling layer contains the high-level interface that scientists will utilize when composing applications. Oversight Teams will remain in close contact with an **Implementation Team** located at NCAR through regular teleconferences and meetings. The Oversight Teams will participate in requirements analysis, design reviews, prototyping and testing over the course of the project.

The Implementation Team will consist of five software engineers hired specifically for the ESMF project, and a software engineering manager (DeLuca). One member of the Implementation Team will be an **integrator**, whose duties will include setting up the configuration management system and documenting procedures early in the project, and, once the project gets underway, assimilating newly developed classes and functions into an evolving prototype of the ESMF. The Implementation Team will also include **two software engineers with experience in OOD** and high-performance computing to implement the upper levels of the framework; an **application engineer** with mathematical and Earth system simulation experience who will begin assembling the validation suite immediately, and will serve as a
primary interface to application developers at other institutions; and a systems engineer, to implement relatively straightforward utility functions.

We also request funds for a graduate student at the University of Michigan, to address issues in core ESMF development relevant to the space weather community; a research scientist at MIT, to assist with development of a variety of basic functions; and one-half FTE at Los Alamos to work on conservative regridding, building on the SCRIP package.

We propose to fund staff for core framework development partly through NSF cosponsorship and partly through NASA HPCC. Management positions, including the Principle Investigator and software engineering manager, will be funded through NSF cosponsorship.

4.4 Qualifications

Our Investigator Team possesses the combination of skills and backgrounds necessary to develop the ESMF and establish it as a standard throughout the climate and weather communities. The most critical of these skills are:

*Expert understanding of the way in which Earth science applications and components are constructed and combined to support scientific research.* The quality of the scientific expertise on our Team is demonstrated by the extensive publication lists, awards and honors, and prominent positions held by our scientific CoIs.

*Previous experience creating and supporting parallel, high-performance software frameworks and toolkits for a large user community.* Dr. Anderson was a co-developer of FMS at GFDL and Dr. Suarez a co-developer of GEMS at NSIPP, both frameworks for climate and general circulation modeling. Ms. DeLuca was a co-developer of the parallel STAPL framework for real-time signal processing, a software tool currently used in multiple military radar systems. Dr. Boville was a co-developer of the CCSM, a model that is distributed to and supports hundreds of researchers internationally. Dr. Smith has been central in developing and providing ongoing support for the PETSc package for solving PDEs on parallel platforms, which has a user base in the hundreds.

*Experience in technical management and development in a CMM Level II+ environment.* Most of the Investigators in this project include technical management as part of their job. Ms. DeLuca served as the technical lead for new development on the STAPL framework project at MIT Lincoln Laboratory, a large, DoD-sponsored project that employed approximately 15 FTEs over the course of about 4 years, and was targeted at CMM Level III.

*Proven skill in constructing algorithms and data structures to enable very high-performance codes.* All of our investigators have experience constructing scalable, high-performance codes on parallel platforms; for example, Dr. Stout was a NASA HPCC Round II Participant in a space weather project that achieved 343 GFLOPS on 1490 processors of a Cray T3E-1200 for the MHD code BATS-R-US.

*Extensive contacts to promote the ESMF.* Established FMS and GEMS users at GFDL and NSIPP, respectively, are accustomed to working with a framework and will readily adapt to the ESMF. Hundreds of researchers will be introduced to the framework through their involvement with the NCAR CCSM. Finally, the involvement and advocacy of Dr. Killeen, Director of NCAR, and other internationally recognized scientists on our Investigator Team will help to generate interest in and support for our project.
(II) Software Engineering Plan

In this section we present a software engineering plan for the ESMF, including software team structure and management, a software process that extends from system specification through distribution and maintenance, and tools and techniques to support development, collaboration, and distribution.

1 Software Teams and Management

An Implementation Team will be established at NCAR consisting of five software engineers supervised by a software engineering manager. One of the software engineers will be an integrator (no relation to the NASA “ESMF Integrator”) who is responsible for Team support functions such as configuration management and defect tracking. The Implementation Team will draft design specifications, prototype and implement ESMF components, test and validate the framework, and distribute releases. Software development will be guided by three partially overlapping Oversight Teams focused on different aspects of the framework: utilities, fields and grids, and coupling. The members of the Oversight Teams will include the CoIs of this proposal, and will consist of appropriate mixes of software engineers, application scientists and computer scientists. Each Oversight Team will designate a lead. Responsibilities of the Oversight Teams will include reviewing software design and tracking implementation progress.

The software engineering manager and integrator will maintain a system view and ensure that development is coordinated. The Oversight Teams and Implementation Team will coordinate with staff at other institutions working on applications for Parts II and III.

2 Software Process

The Implementation Team will follow a structured software process commensurate with CMM Level II [32, 33]. The process will include many of the procedures recommended by the Software Best Practices Initiative [3]. Our documents and reviews will be simplified versions of those described in standard references [24, 40, 41]. We will aim for an effective process free of extraneous overhead.

Staged Software Development The major milestones described in Part V are the result of the coordinated completion of many smaller events, each of which has a “completion gate,” such as approval of a document. Table 1 shows the progression of events in ESMF software development, and the product and gate associated with each event.

The initial set of events, labelled “ESMF Definition” is focused on specifying the ESMF system and procedures as a whole. The second group of events, “Class Implementation” describes the development steps applied to individual software classes. As classes are completed they will be integrated into an evolving prototype of the ESMF. The final development stage, “Integration and Distribution”, involves the integration of classes leading to a software release. The ESMF will have three major software releases corresponding to milestone I,J, and K; smaller releases and demonstrations will be scheduled to ensure that the project is on track.
### Table 1 ESMF Software Event Progression

<table>
<thead>
<tr>
<th>Event</th>
<th>Product</th>
<th>Completion Gate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ESMF DEFINITION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requirements specification</td>
<td>Requirements Document</td>
<td>Document review</td>
</tr>
<tr>
<td>Outlines ESMF functional scope and requirements.</td>
<td>Prepared by: all collaborators</td>
<td>Reviewed by: all collaborators</td>
</tr>
<tr>
<td>Architectural description</td>
<td>Architecture Document</td>
<td>Document review</td>
</tr>
<tr>
<td>Describes layering strategy, function and interaction of major components.</td>
<td>Prepared by: all collaborators</td>
<td>Reviewed by: all collaborators</td>
</tr>
<tr>
<td>Software process definition</td>
<td>Developer’s Guide Document</td>
<td>Document review</td>
</tr>
<tr>
<td>Implementation study</td>
<td>Implementation Report</td>
<td>Document review</td>
</tr>
<tr>
<td>Assesses existing software, optimal language, threading strategy, more.</td>
<td>Prepared by: implementation team</td>
<td>Reviewed by: all collaborators</td>
</tr>
<tr>
<td>Software implementation and test plan</td>
<td>Software Impl. and Test Plan</td>
<td>Plan review</td>
</tr>
<tr>
<td>Plan for Implementation and testing based on class dependencies, milestones.</td>
<td>Prepared by: software mgr.</td>
<td>Reviewed by: all collaborators</td>
</tr>
<tr>
<td><strong>CLASS IMPLEMENTATION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class design</td>
<td>Class Design Document</td>
<td>Design review</td>
</tr>
<tr>
<td>Includes requirements, function, and interface specification.</td>
<td>Prepared by: class developer(s)</td>
<td>Reviewed by: Oversight Team, software mgr.</td>
</tr>
<tr>
<td>Class implementation</td>
<td>Prototype code</td>
<td>Code review</td>
</tr>
<tr>
<td>A class may be stubbed or partially implemented for a given release.</td>
<td>Prepared by: class developer(s)</td>
<td>Reviewed by: Oversight Team, software mgr.</td>
</tr>
<tr>
<td>Class unit test</td>
<td>Unit test code</td>
<td>Unit test</td>
</tr>
<tr>
<td>Class is tested stand-alone with a variety of inputs.</td>
<td>Prepared and tested by: class developer(s)</td>
<td>Verified by: software mgr.</td>
</tr>
<tr>
<td><strong>INTEGRATION AND DISTRIBUTION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class integration</td>
<td>ESMF system prototype</td>
<td>System test and benchmarking</td>
</tr>
<tr>
<td>Unit tested class is integrated into an evolving prototype of the ESMF.</td>
<td>Prepared by: class developer(s), integrator</td>
<td>Verified by: software mgr.</td>
</tr>
<tr>
<td>User documentation updated</td>
<td>User’s Guide &amp; Reference</td>
<td>Review before software release</td>
</tr>
<tr>
<td>Class design documentation is updated and converted to user documentation.</td>
<td>Prepared by: class developer(s), integrator</td>
<td>Reviewed by: software mgr.</td>
</tr>
<tr>
<td>System release</td>
<td>System test</td>
<td>ESMF system release</td>
</tr>
<tr>
<td>Code and documentation is released. Defects and requests for features are tracked and incorporated into future releases.</td>
<td>Prepared by: integrator, software mgr.</td>
<td>Evaluated by: ESS Project, user community</td>
</tr>
</tbody>
</table>
Documents, Reviews and Verification

Documents: Documents to be prepared for this project are outlined in Table 1. We will structure design documents so they can be easily converted to user documentation. All documents will be prepared in a format that easily generates both hardcopy and web-friendly html. We intend to use a documentation generation tool to automatically create and update portions of our documentation.

Reviews: As shown in Table 1, a variety of reviews will be held, including requirements, design and code reviews. The outcome of a review will either be a pass or, if significant changes are required, another review iteration. The software engineering manager will attend all reviews for coordination.

Verification and benchmarking: Class and system tests will be designed so that they verify that the code being tested fulfills its requirements. The validation suite to be developed is described in Section 3.4. The ESS Evaluation Team and vendors will assist with performance evaluation.

Source Availability and Distribution We plan to develop our code in an open source development environment such as SourceForge [2]. After a prototype is released, we plan to engage the broader Earth science community in contributing to the ESMF. Community contributions will be integrated into new releases by the core team at NCAR.

Source code and documentation will be distributed via an ESMF website. The website may be an extension of that currently maintained by the CMIWG, as noted in Section 2.3. We plan to hold a series of workshops to introduce the broader community to the ESMF.

Software Maintenance NCAR is committed to offering an ongoing program of user support, maintenance, promotion, and research into improved and extended capabilities for the ESMF. We plan to retain some or all members of the Implementation Team at NCAR as core staff to carry out this work.

NASA ESS Team Role We welcome coordination with NASA technical staff, and would anticipate detailing our interactions during the negotiation process.

3 Software Tools and Techniques

Configuration management: We will likely use CVS for configuration management since it is mature, freely available, and the current community standard. We anticipate maintaining code at the following acceptance levels: Active (untested), Unit tested, and Integrated (code is part of a working ESMF prototype). These levels reflect the completion gates applied to code development shown in Table 1.

Software metrics: The software engineering manager will track a simple set of software metrics throughout development to evaluate progress and predict schedules.

Defect tracking: A tool such as Bugzilla [1] will be used to maintain a database of defects and new feature requests.

Collaborative tools: We plan to employ weekly teleconferences to keep Oversight Teams in close touch with the Implementation Team, as well as quarterly face-to-face meetings. We will continue to maintain project mailing lists and discussion forums.
(III) References

References


(IV) Biographical Sketches
TIMOTHY L. KILLEEN
National Center for Atmospheric Research,
P.O. Box 3000, Boulder CO 80307

PROFESSIONAL INTERESTS
My interests span the environmental and space sciences, collaboration science and technology, and undergraduate educational reform.

EDUCATION
B.Sc. 1972 1st Class Honors in Physics, University College, London
PhD. 1975 Atomic and Molecular Physics, University College, London

SELECTED HONORS AND AWARDS
1996-2000 President and President-Elect, American Geophysical Union, Space Physics and Aeronomy Section
2000 Excellence in Teaching Award, University of Michigan, College of Engineering
1998 NASA Achievement Award, Polar Spacecraft
1993 Excellence in Research Award, University of Michigan, College of Engineering

EMPLOYMENT
2000–Present Director National Center for Atmospheric Research
1997–2000 Associate Vice President for Research University of Michigan
1997–2000 Director Global Change Laboratory, University of Michigan
1993–1998 Director Space Physics Research Laboratory, University of Michigan
1987–1990 Associate Professor Atmospheric, Oceanic and Space Sciences University of Michigan
1984–1987 Associate Research Scientist University of Michigan

SELECTED PUBLICATIONS


JEFFREY L. ANDERSON
NOAA/Geophysical Fluid Dynamics Laboratory,
P.O. Box 308, Princeton, NJ 08542

PROFESSIONAL INTERESTS
My research interests include stochastic data assimilation and prediction, seasonal-interannual prediction, and software engineering to support climate system models.

EDUCATION
B.S. 1984 Meteorology and Computer Science, University of Utah
M.S. 1986 Computer Science
University of California, Berkeley
Ph.D. 1990 Atmospheric and Oceanic Sciences, Princeton University

EMPLOYMENT
1995–Present  Head  Experimental Prediction Group,
Geophysical Fluid Dynamics Laboratory
1992–Present  Lecturer  Atmospheric and Oceanic Sciences,
Princeton University
1992-1995  Meteorologist  Geophysical Fluid Dynamics Laboratory
1990-1992  Postdoctoral Scientist  Climate Analysis Center,
National Meteorological Center

SELECTED PUBLICATIONS


BYRON A. BOVILLE
National Center for Atmospheric Research,
P.O. Box 3000, Boulder CO 80307

PROFESSIONAL INTERESTS
My research has concentrated on developing and applying general circulation models of the lower and middle atmosphere for studies of atmospheric dynamics and climate. I have been one of the central figures in both the scientific and computational development of 4 generations of the NCAR atmospheric general circulation model. More recently, I have concentrated on coupled ocean-atmosphere modeling and was co-chair of the team which developed the NCAR Climate System Model (CSM). I am currently interested in the climate impact of solar variability and the role of the middle atmosphere in climate variability and climate change.

EDUCATION
B.Sc. 1975 1st Class Honors in Meteorology,
McGill University, Montreal, Canada
Ph.D. 1979 Atmospheric Sciences,
University of Washington, Seattle, Washington

EMPLOYMENT
1999–Present Head Climate Modeling Section,
Climate and Global Dynamics Division
National Center for Atmospheric Research
1992–Present Senior Scientist National Center for Atmospheric Research
1979–1981 Postdoc Advanced Study Program
National Center for Atmospheric Research

SELECTED PUBLICATIONS


CECELIA DeLUCA
National Center for Atmospheric Research,
P.O. Box 3000, Boulder CO 80307

PROFESSIONAL INTERESTS
My interests include the design of large, high-performance software systems, particularly those relating to atmospheric science; parallel algorithms; real-time systems, and software engineering tools and processes. I was a design lead on the development of the Space-Time Adaptive Processing Library (STAPL) parallel framework for real-time radar applications. STAPL is an integral part of multiple operational next-generation radar systems and has been ported to several platforms. It extends the serial Vector Signal and Image Processing Library (VSIPPL) standard to SMP-cluster architectures. Previous projects have included the development of parallel codes for the simulation of middle atmospheric dynamics, atmospheric chemistry, and remote sensing of atmospheric temperatures.

EDUCATION
A.I.B. 1992 Liberal Arts/Social Sciences, Harvard University, Cambridge, MA
M.S. 1994 General Engineering, Boston University, Boston, MA
M.S. 1996 Meteorology, Massachusetts Institute of Technology, Cambridge, MA

AWARDS
1994 Boston University College of Engineering Outstanding Achievement Award, first in graduating class

EMPLOYMENT
1999–Present  Software Engineer  Scientific Computing Division, National Center for Atmospheric Research
1998–1999  Lead Software Engineer  MIT Lincoln Laboratory
1996–1998  Software Engineer  MIT Lincoln Laboratory
1993–1994  Manager, Technical Support  Omnet Communications

SELECTED PUBLICATIONS


PROFESSIONAL INTERESTS

My current research interests include techniques for global atmospheric data assimilation, physical-space analysis systems, error covariance modeling, bias estimation and correction, quality control, land-surface, precipitation and aerosol data assimilation, and efficient methods for assimilation of remotely sensed data. Other research interests not in the area of data assimilation include aerosol forcing of climate, hydrological cycle of the subtropics, estimation of fluxes of heat, momentum and fresh water over the global oceans for observational studies and forcing ocean models.

EDUCATION
B.S. 1982 Physics, Catholic University of Rio de Janeiro, Brazil
M.S. 1984 Physics, Catholic University of Rio de Janeiro, Brazil
Ph.D. 1989 Meteorology, Massachusetts Institute of Technology

EMPLOYMENT
1994-Present Meteorologist Data Assimilation Office, NASA Goddard Space Flight Center
1990-1993 Assistant Professor University of Wisconsin-Milwaukee
1989-1990 Visiting Scientist Program in Atmospheric and Ocean Sciences Princeton University

RELATED PUBLICATIONS


27
PROFESSIONAL INTERESTS

My research interests include modeling and analysis of aspects of the coupled magnetosphere/ionosphere/thermosphere system, paleoclimatology, isotope geochemistry and atmospheric chemistry, education, and multimedia.

EDUCATION

B.S. 1980 Geophysics and Space Physics, University of California at Los Angeles
M.S. 1984 Geophysics and Space Physics, University of California at Los Angeles
M.S. 1987 Geophysics and Space Physics, University of California at Los Angeles

EMPLOYMENT

2000–Present Director Education and Outreach University Corporation for Atmospheric Research, University of California, Boulder
1993–2000 Associate Research Scientist Space Physics Research Laboratory University of Michigan
1998–1999 Adjunct Associate Professor Dept. Atmospheric, Oceanic and Space Science, University of Michigan
1989–1993 Assistant Research Scientist Space Physics Research Laboratory, University of Michigan

SELECTED PUBLICATIONS


PHILIP W. JONES
Theoretical Fluid Dynamics (T-3),
Los Alamos National Laboratory,
Los Alamos, NM 87545

PROFESSIONAL INTERESTS
Current interests involve the use of massively parallel computers to study problems in
gеophysical and astrophysical fluid dynamics, including atmosphere, ocean and coupled cli-
mate modeling, middle atmosphere dynamics and fully-compressible thermal convection.

EDUCATION
B.S. 1985 Physics and Mathematics with distinction,
Iowa State University
Ph.D. 1991 Astrophysical, Planetary, and Atmospheric Sciences,
University of Colorado

EMPLOYMENT
1993–Present Staff Member
Theoretical Fluid Dynamics (T-3),
Los Alamos National Laboratory
1991–1993 Post-doctoral Research Associate
Geoanalysis Group (EES-5),
Los Alamos National Laboratory
1986–1991 Research Assistant
Joint Institute for Laboratory Astrophysics
and Center for Applied Parallel Processing,
University of Colorado, Boulder

SELECTED PUBLICATIONS
   Proceeding of the Second International Workshop on Software Engineering and Code
   Design in Parallel Meteorological and Oceanographic Applications, ed. M. O’Keefe and
   on MPP and Clustered SMP Computers,” Making its Mark: Proceedings of the 7th
   ECMWF Workshop on the Use of Parallel Processors in Meteorology, ed. G. R. Hoff-
   mann and N. Kreitz (Singapore: World Scientific Publishing).
   Sci., 54, 1107-1116.
5. Jones, P.W., Kerr, C.L. and Hemler, R.S. 1995 “Practical Considerations in Develop-
   ment of a Parallel SKYHI General Circulation Model,” Parallel Computing, 21,
   1677-1694.
J. WALTER LARSON
Mathematics and Computer Science Division,
Argonne National Laboratory
9700 South Cass Avenue, Argonne, IL 60439-4844

PROFESSIONAL INTERESTS
My interests include the design and development of modular, high-performance software for use in data assimilation, regional and global climate modeling, and statistical analysis of climate model output. In the past I have performed research in a number of areas, including: soliton theory; chaos and dynamical systems; global climate change; regional climate modeling.

EDUCATION
B.A. 1984 Physics and Mathematics,
Drake University, Des Moines, IA
M.S. 1986 Physics,
College of William and Mary, Williamsburg, VA
PhD. 1992 Physics,
College of William and Mary, Williamsburg, VA

EMPLOYMENT
1999–Present Assistant Computer Scientist
Mathematics and Computer Science Division,
Argonne National Laboratory
1996–1999 Research Associate
Meteorology Dept, University of Maryland and
NASA Data Assimilation Office
1994–1996 Postdoctoral Fellow
Centre for Resource and Environmental Studies,
Australian National University

SELECTED PUBLICATIONS


STEPHEN J. LORD
Environmental Modeling Center
National Centers for Environmental Prediction
NOAA Science Center, Rm. 207
Washington, DC 20233

PROFESSIONAL INTERESTS
My interests are in managing and participating in all aspects of data assimilation and numerical model development for weather and seasonal climate forecasts. As Director of the Environmental Modeling Center, National Centers for Environmental Prediction, I oversee a staff of 90 who are dedicated to improving operational weather, ocean and climate modeling products to support the National Weather Service mission.

EDUCATION
B.S.  1969  Physics
    Yale University (cum laude)
M.S.  1975  Atmospheric Sciences
    University of California at Los Angeles
Ph.D. 1978  Atmospheric Sciences,
    University of California at Los Angeles

HONORS AND AWARDS
1997  AMS Fellow
1996  NOAA Dept. of Commerce Gold Medal for Implementation of the GFDL Hurricane Model
1993  NOAA Dept. of Commerce Bronze Medal for Applied research on hurricane track prediction

EMPLOYMENT
2000–Present  Director
              Environmental Modeling Center,
              National Centers for Environmental Prediction
1993–2000  Acting Director/Deputy Director
           Environmental Modeling Center,
           National Centers for Environmental Prediction
1989–1993  Meteorologist
           National Meteorological Center
1980–1989  Meteorologist
           Hurricane Research Division,
           Atlantic Oceanographic and Meteorological Laboratory

SELECTED PUBLICATIONS


JOHN MARSHALL
Department of Earth, Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

PROFESSIONAL INTERESTS
My research is directed at understanding key components of the general circulation of
the atmosphere and ocean and the development of models to study them. I am interested in
a variety of problems in geophysical fluid dynamics and their role in climate, ranging from
rotating convection, the global circulation of the ocean and air-sea interaction. I use and
develop numerical models of the atmosphere, ocean and climate.

EDUCATION
B.S. 1976 First Class Honors in Physics,
Imperial College, London
Ph.D. 1980 Physics
Imperial College

EMPLOYMENT
1992–Present Professor Massachusetts Institute of Technology
1992 Associate Professor Massachusetts Institute of Technology
1991–1992 Reader in Physics Imperial College
1984–1990 Lecturer in Physics Imperial College
1982–1983 Post-doctoral fellow University of Oxford

SELECTED PUBLICATIONS
and nonhydrostatic ocean modeling, J. Geophysical Res., 102(C3), 5733–5752.
incompressible Navier Stokes model for studies of the ocean on parallel computers, J.
shaved cells in a height coordinate ocean model, Mon. Wea. Rev., 125, 2293–2315
algorithms, Journal of Marine Systems, 18, 115–134
parison of implicitly parallel multi-threaded and data-parallel implementations of an
ocean model based on the Navier-Stokes equations. J. of Parallel and Distributed
Computing, 48:1, 1–51
PROFESSIONAL INTERESTS
My research interests include parallel computing and the numerical solution of partial differential equations.

EDUCATION
B.S.  1986  Mathematics
      Yale University
Ph.D. 1990  Mathematics
      New York University

HONORS AND AWARDS
1993  Co-winner, Householder Prize for best dissertation in numerical linear algebra during the preceding 1991  Second Prize, Fifth Leslie Fox Prize Meeting, international prize in numerical analysis offered every 1990  First Prize, Student Paper Competition, Copper Mountain Conference on Iterative Methods

EMPLOYMENT
1995–Present  Computer Scientist  Mathematics and Computer Science Division, Argonne National Laboratory
1994–1995  Assistant Computer Scientist  Mathematics and Computer Science Division, Argonne National Laboratory
1992–1994  Assistant Professor (Visiting)  University of California at Los Angeles
           Argonne National Laboratory

SELECTED PUBLICATIONS


PROFESSIONAL INTERESTS
My research interests include parallel computing and scientific computing, algorithms and data structures, adaptive statistical designs, and discrete mathematics.

EDUCATION
B.A. 1970 Mathematics,
Centre College, Danville, Kentucky
Ph.D. 1977 Mathematics,
Indiana University, Bloomington, Indiana

RECENT HONORS AND AWARDS
1999 College of Engineering Team Excellence Award, University of Michigan
1999 Partnership Award, IBM
1995 College of Engineering Service Award, University of Michigan

EMPLOYMENT
1997–Present Director Center for Parallel Computing,
University of Michigan
1992–Present Professor Department of Electrical Eng. and Computer Science
University of Michigan

SELECTED PUBLICATIONS


MAX J. SUAREZ
NASA Seasonal to Interannual Prediction Project
NASA Goddard Space Flight Center
Greenbelt, MD 20771

PROFESSIONAL INTERESTS
Large-scale atmosphere/ocean interactions, climate modeling, numerical methods, parameterization of subgrid-scale processes in atmospheric models, maintenance of the atmospheric general circulation, climate sensitivity.

EDUCATION
B.S.  1971  Engineering Science, University of Florida
M.E.  1972  Engineering Science, University of Florida
M.A.  1974  Geophysical Fluid Dynamics, Princeton University
Ph.D. 1976  Geophysical Fluid Dynamics, Princeton University

EMPLOYMENT
1983–Present  Meteorologist  NASA/Goddard Space Flight Center
1976–1983  Assistant Professor  UCLA

SELECTED PUBLICATIONS


(V) Milestones, Schedule and Costs

The total budget for this proposal is $4,462,800.
The dates on these milestones assume that the first payment is received June 2001.

<table>
<thead>
<tr>
<th>Milestone Number</th>
<th>Milestone Label</th>
<th>Milestone Title</th>
<th>Expected Completion Date</th>
<th>Advance Payment Amount</th>
</tr>
</thead>
</table>
| 1                | A               | Software engineering plan completed  
| 2                | E               | Code baseline completed  
*Assemble the EVA Suite and baseline it on the ESS Testbed. The metrics will be 1) time to solution 2) functionality and 3) scalability. Documented source code will be delivered via the Web.* | Nov. 2001                | $467,800               |
| 3                | H               | Design policy for interoperability and community delivery agreed on  
| 4                | B               | First Annual Report delivered  
*Submit FY01 Annual Report via the Web.*                                        | June 2002                | $100,400               |
| 5                | F               | First code improvement completed  
* Demonstrate EVA usage of the ESMF. At least one EVA component (not necessarily the same one) must show increased functionality in each of these areas: an additional mode of component execution (SPMD or MPMD), optimized communication for at least 2 additional grid types, transition to an integrated hybrid (multithreaded/distributed) programming model, C++/F90 language interoperability. Deliver documented source code via the Web.* | July 2002                | $567,800               |
(Milestones, continued)

<table>
<thead>
<tr>
<th>Milestone Number</th>
<th>Milestone Label</th>
<th>Milestone Title</th>
<th>Expected Completion Date</th>
<th>Advance Payment Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>I</td>
<td>Interoperability prototype from Milestone “H” tested with improved codes</td>
<td>Nov. 2002</td>
<td>$601,200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstrate interoperability of EVA test codes with applications: an EVA coupler, component and driver each interoperaing with 1 of 4 modeling applications using the framework: CCSM, GFDL, NSIPP, mitGCM; and an EVA component and coupler each interoperaing with 1 of 3 data assimilation schemes. Source code delivered via the Web.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>Second Annual report delivered</td>
<td>June 2003</td>
<td>$100,400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Submit FY02 Annual Report via the Web.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>G</td>
<td>Second code improvement completed</td>
<td>July 2003</td>
<td>$367,800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstrate full functionality of EVA with improved performance, using time to solution as the metric. No test code in the EVA suite will be worse than 10% of that measured as the baseline. At least one test code in the validation suite will demonstrate a 50% or better efficiency of scaling on 128 processors of the ESS Testbed. Deliver documented source code via the Web.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>J</td>
<td>Full interoperability demonstrated using improved codes</td>
<td>Nov. 2003</td>
<td>$401,200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstrate full interoperability of the EVA test codes with applications: 2 or more EVA couplers, components and drivers with 2 of 4 modeling applications using the framework: CCSM, GFDL, NSIPP, mitGCM; and 2 or more EVA components and couplers with 1 of 3 optimized data assimilation schemes. Source code delivered via the Web.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(Milestones, continued)

<table>
<thead>
<tr>
<th>Milestone Number</th>
<th>Milestone Label</th>
<th>Milestone Title</th>
<th>Expected Completion Date</th>
<th>Advance Payment Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>K</td>
<td>Customer delivery accomplished Port the framework to a new platform within 1 month and demonstrate its operation. Deliver a user-friendly website where the ESMF source code, including the EVA Suite, the ESMF Developer’s Guide, and the ESMF User’s Guide and Reference can be obtained. Hold a workshop on ESMF usage.</td>
<td>May 2004</td>
<td>$401,200</td>
</tr>
<tr>
<td>11</td>
<td>D</td>
<td>Final Report delivered Submit Final Report to ESS via the Web.</td>
<td>June 2004</td>
<td>$100,400</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>Education and Public Outreach Complete series of professional development workshops for middle school and high school educators</td>
<td>June 2004</td>
<td>$453,000</td>
</tr>
</tbody>
</table>
(VI) Endorsement Letters
(VII) Education and Public Outreach Addendum

Objectives and Planned Activities

The objective of the education and public outreach addendum to the suite of Earth System Modeling Framework proposals is to develop and implement professional development opportunities on the Earth system sciences for middle and high school educators. Building upon and extending existing resources, a series of nationally-advertised professional development workshops for middle school and high school educators will elucidate concepts associated with Earth system science and modeling. This project will be implemented by UCAR and collaborating scientists, education specialists, and nationally recognized experts in professional development in the geosciences for middle and high school educators. The ESMF Investigator Team will provide expertise in their various disciplines over the course of the project.

Workshops

The most effective way to improve the scientific literacy of students is to work to improve the training of those responsible for educating them. The focus of this effort will be to build upon our extensive experience and existing resources to develop and implement a series of professional development opportunities for middle and high school educators that provides the training they need to enrich the learning of their students. In order to ensure the highest leverage to educators, we propose to target our training activities on educators with a proven track-record of providing training opportunities in their own regions or districts.

The workshops proposed here will extend upon the highly successful model used for the Michigan Space Grant Consortium (MSGC) Geoscience Education Workshop (see below). The yearly workshop series will target 20 educators per year, and will include an intensive 2-week summer semester at NCAR in Boulder, followed by 2 additional 2-day professional development opportunities during the school year.

In addition to providing background content on the Earth system that highlights issues in climate modeling and weather prediction, the workshops will provide opportunities to explore modeling concepts through state-of-the-art Earth system models (the science focus of this suite of proposals) as well as simplified STELLA models of components of the Earth system relevant to curriculum needs of educators as expressed in the National Science Education Standards. Educators will be trained on easy to implement hands-on, inquiry-based classroom activities that illustrate concepts of modeling and aspects of the Earth system. The workshop will include opportunities for educators to gain field experience - for observations and data collection that will be utilized in their studies and also expand their appreciation of the Earth sciences. Finally, the workshop will provide guidance on techniques participants can use to bring the training they have received back to educators in their own districts.

The workshop opportunities will be advertised nationally. Participants will be selected from the applicant pool based on their previous experience and success in regional or local professional development for science educators, as well as our programmatic desire to have a diverse and geographically distributed set of workshop participants. Travel and living expenses for workshop participants will be covered during the workshops, and each participant will receive a stipend to support their efforts to bring the training they have been provided
back to educators in their own districts. In return for the training and support they have 
received, participants will be expected to train an additional 20 educators per year in their 
own districts. Materials needed for this training will be provided from the project.

A dedicated project web site will be developed that provides access to information about 
the project, all lectures, labs, and other training resources, and also provides a venue for 
communication among project participants.

**Previous Relevant Education and Outreach Activities of Project**

Personnel Science and education personnel participating in this proposal have extensive ex-
perience in development of highly-valued educational resources and successful professional 
development opportunities for K-12 educators. Over the past decade, UCAR has developed 
a number of professional development resources with funding from NSF for middle and high 
school educators on global change, climate change, and the ways in which components of 
the Earth system are interconnected. In addition, NCAR has been responsible for the de-
velopment of an influential series of documents *Reports to the Nation* that highlight our 
understanding of the Earth system and the potential contributions of humans in changing 
our climate. With funding from NSF, UCAR developed Project LEARN, in which science 
education faculty, scientists, and professional development specialists provided training to 
over 2000 middle school educators on the atmospheric sciences. After six years of fund-
ing, resources developed through this program are being made available on-line to support 
educators nationwide.

Prior to moving to UCAR, Dr. Roberta Johnson (the Director of the UCAR Office for 
education and Outreach) developed, in collaboration with colleagues, a highly successful 
workshop for middle and high school Earth science educators while she was the Director 
of the MSGC (funded by NASA). This workshop utilized the concept of global change as 
a vehicle for exploring the breadth of the geosciences, covering topics ranging from re-
cyling to anthropogenic effects on ozone and greenhouse gases to the role of plate tectonics 
in determining planetary climates. Workshop activities included a combination of lecture, 
computer-based training and activities (with an emphasis on modeling illustrated through 
STELLA), training on hands-on inquiry-based activities, field experience, and project de-
velopment. Resources utilized in the workshop included content and activities available through 
the NASA-sponsored Windows to the Universe project, also PI'd by Dr. Johnson.

**Management and Coordination**

Dr. Johnson will be responsible for the directing the educational activities described in 
this addendum. She will be assisted by the staff of the UCAR Office of Education and 
Outreach, including Ms. Susan Foster, who will manage the project and assist in design 
and implementation of the workshop. Dr. Killeen and other selected science CoIs on the 
project will participate in the educational effort by giving lectures to project participants 
and participating in other workshop activities. Mr. David Mastie and Mr. Parker Pennington 
IV, both recognized experts in teacher training in the Earth sciences at the middle and high 
school levels, will collaborate in the effort and continue in the roles they have had in the 
MSGC Geoscience Education Workshop as lead teacher trainers, working with local UCAR 
EO staff.