The Earth System Modeling Framework (and Beyond)

Fei Liu

NESII

NOAA Environmental Software Infrastructure and Interoperability

http://www.esrl.noaa.gov/nesii/

March 27, 2013

GEOSS

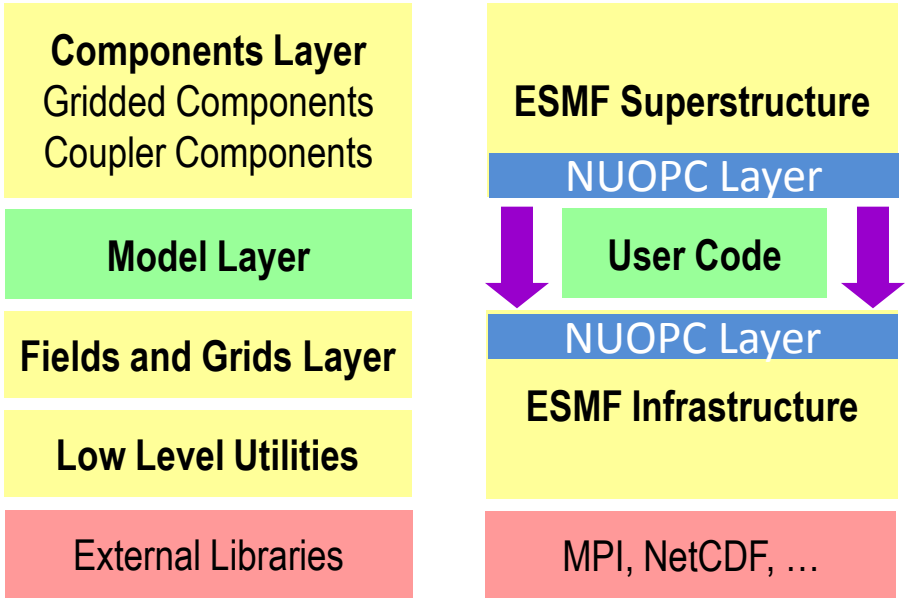
Community

- ESMF is an established component framework (started 2002)
- Began in the global weather and climate modeling community, now includes coastal, hydrological, and space weather applications
- Distributed, diverse set of operational and research customers who expect to incorporate tools into their local modeling system, not use a particular coupler or integration facility
- Established multi-agency sponsorship and governance
- Active development of new products, and exploration of new coupling paradigms for emerging applications

Plus many user organizations …

ESMF Basics

- ESMF provides a **superstructure** for assembling geophysical components into applications.
- ESMF provides an **infrastructure** that modelers use to share common utility code for functions such as grid remapping and time management
- The **NUOPC Layer** is a small set of calls that
  - Aggregate ESMF functions
  - Encode metadata, sequencing, and other conventions to improve interoperability and make using ESMF simpler



NUOPC is the National Unified Operational Prediction Capability

Standard Interfaces

- All ESMF components have the same three standard methods:
  - Initialize
  - Run
  - Finalize
- Each standard method has the same simple interface:

```
call ESMF_GridCompRun (myComp, importState,
exportState, clock, …)

Where:
myComp points to the component
importState is a structure containing input fields
exportState is a structure containing output fields
clock contains timestepping information
```

- Interfaces are wrappers and can usually be set up in a non-intrusive way

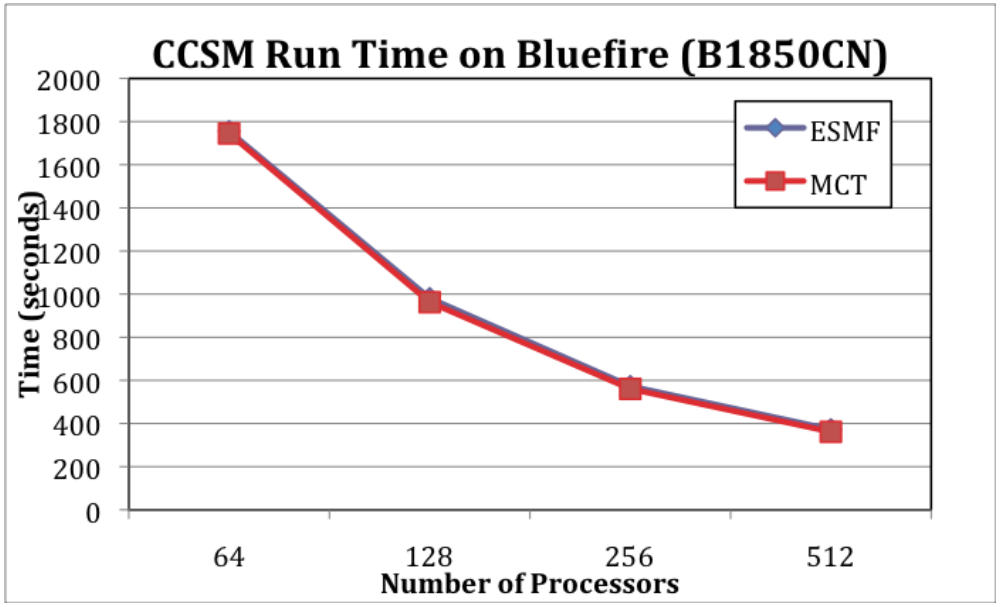
Efficient Data Structures

- ESMF represents data as Fields that can be built on four parallel discretization types:
  - Logically rectangular grids, which may be connected at the edges
  - Unstructured meshes
  - Observational data streams
  - Exchange grids, which are overlaid grids, used for conservative remapping
- ESMF can transform data among these representations, using an underlying parallel 3D finite element unstructured mesh framework (in place since ~2007)
- May replace finite element engine with MOAB if it's faster or more flexible

Component Overhead

- Test of the overhead for ESMF wrapped native CCSM4 component
- For this example, ESMF wrapping required no code changes to scientific modules

- **No significant performance overhead** (< 3% is typical)
- Few code changes for codes that are modular



- Platform: IBM Power 575, bluefire, at NCAR
- Model: Community Climate System Model (CCSM)
- Versions: CCSM_4_0_0_beta42 and ESMF_5_0_0_beta_snapshot_01
- Resolution: 1.25 degree x 0.9 degree global grid with 17 vertical levels for both the atmospheric and land model, i.e. 288x192x17 grid. The data resolution for the ocean model is 320x384x60.

Performance studies are here (about the same as MCT for mat mul out to 16K):
http://www.earthsystemmodeling.org/metrics/performance/

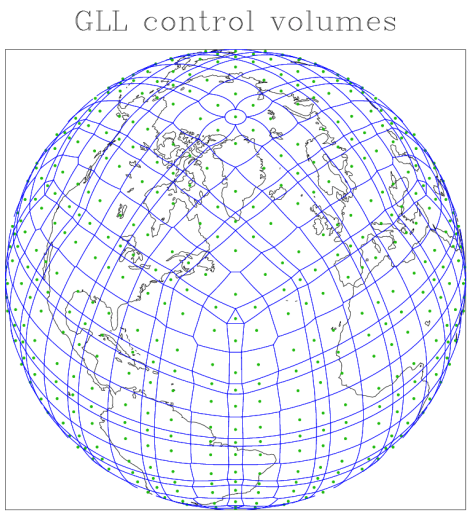
ESMF Parallel Regridding

ESMF offers extremely fast parallel regridding with many options (regional, global; bilinear, higher order, first order conservative, and nearest neighbor methods; logically rectangular grids and unstructured meshes; multiple pole options; 2D or 3D; invocation as **offline** application or **during model run**)

Summary of features:

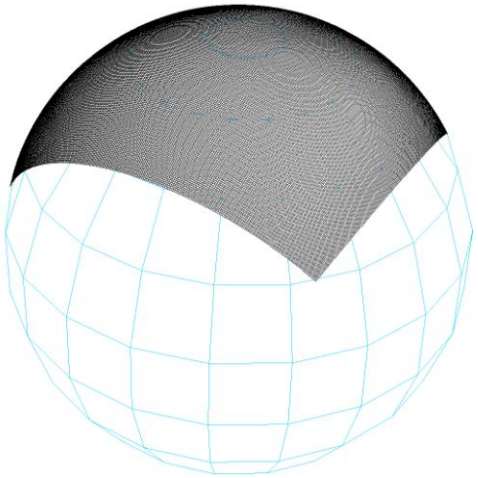
http://www.earthsystemmodeling.org/esmf_releases/non_public/ESMF_5_3_0/esmf_5_3_0_regridding_status.htm

IMPACT:
"use of the parallel ESMF offline regridding capability has reduced the time it takes to create surface datasets from hours to minutes" - Mariana Vertenstein, NCAR



HOMME Cubed Sphere Grid with Pentagons
Courtesy Mark Taylor of Sandia

FIM Unstructured Grid



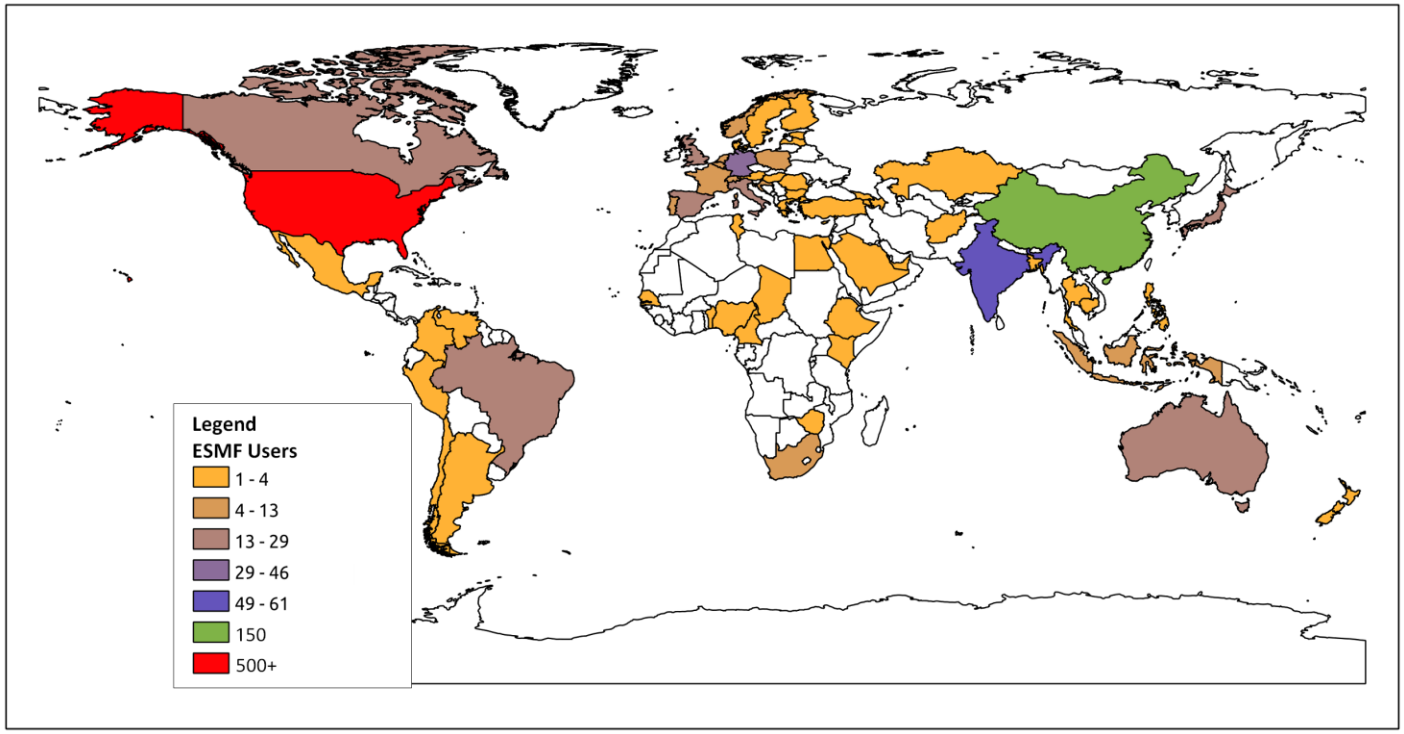
Regional Grid

ESMF supported grids

Metrics

- 4500+ downloads
- 4600+ regression tests on 30ish tested platforms
- 2700 registered people on informational list
- 85 components; 12 major modeling systems
- 700,000 SLOC

ESMF Download Locations

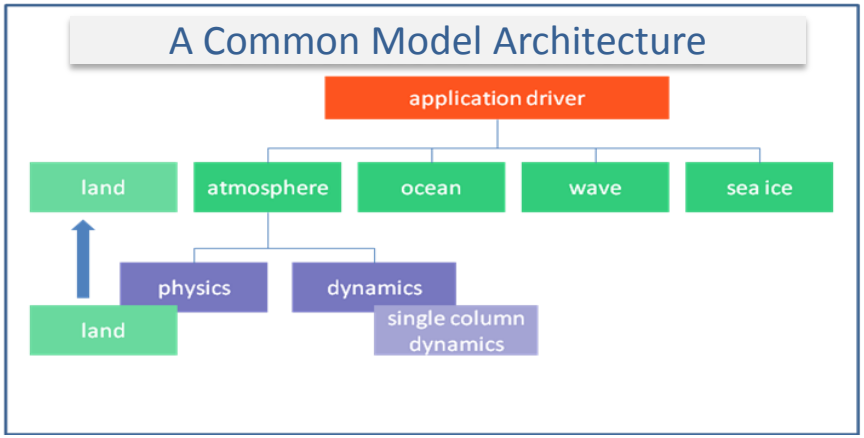


Metadata

- ESMF has a class called Attributes that can be used to output model descriptions using the Common Information Model (CIM) XML format used in CMIP5 (5th Coupled Model Intercomparison Project)
- There are CIM tools that were developed by an international consortium (ES-DOC) to create forms, display, and evolve this schema, and these are continuing to evolve,
  see: http://www.earthsystemcog.org/projects/es-doc-models/

  These tools are also being used in Model Intercomparison Projects (MIPs) such as DCMIP: http://www.earthsystemcog.org/projects/dcmip-2012/
- ESMF partners developed CF String Syntax for semantic mediation, a proposed extension to the globally accepted Climate and Forecast (CF) conventions
  - This syntax enables additional information to be attached to CF while still remaining backward compliant with the widely accepted CF conventions
  - Allows for use of data systems such as ESGF that rely on CF

A Common Model Architecture for Interoperability



- A technical coupling mechanism is necessary but not sufficient for interoperability
- A Common Modeling Architecture (CMA) defines typical component domains, behavior, and relationships
- A CMA has emerged in the U.S. with atmosphere, land, ocean, ice and wave as sibling components running in a single executable system –with variations
- The NUOPC Layer encodes elements of the CMA for ESMF

ESMF-enabled systems include:

Navy Coupled Ocean Atmosphere Mesoscale Prediction System / Wavewatch III
Navy Operational Global Atmospheric Prediction System
Hybrid Coordinate Ocean Model – CICE Sea Ice

NASA GEOS-5 Atmospheric General Circulation Model
NASA GISS Model E (in progress)

NOAA Environmental Modeling System (NEMS)
NOAA GFDL MOM4 Ocean Model

NCAR Community Earth System Model
Weather Research and Forecast Model

HAF Kinematic Solar Wind-GAIM Ionosphere

pWASH123 Watershed-ADCIRC Storm Surge Model

Features and Benefits:

- Interoperability promotes cross-agency knowledge transfer and collaboration
- Cost savings from code reuse

National Unified Operational Prediction Capability

Partners

- Air Force, Navy, and NOAA partnership to address common operational global NWP needs/requirements
- Includes research partners from NASA, NCAR, DOE and other agencies and universities

Products

- Creating managed multi-model ensemble system
- Supporting ESMF-bassed Common Model Architecture and NUOPC Layer infrastructure

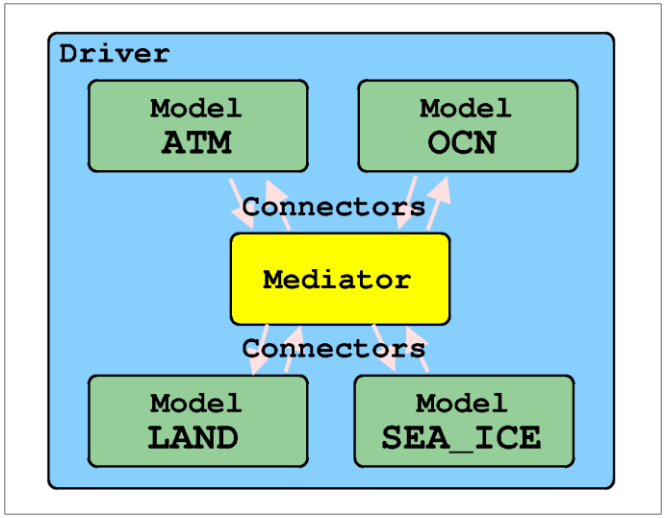
Goals

- Accelerate transition of new technology to operations
- Demonstrably improved level of component interoperability and capability
- Interoperability of NASA, NOAA, Naval Research Laboratory modeling systems using ESMF

Website: http://www.earthsystemcog.org/projects/nuopc/

NUOPC Software Building Blocks

- **NUOPC_Driver** Provides a harness for Models, Mediators, and Connectors, coordinating their initialization and driving them during the application time loop.
- **NUOPC_Model** Typically implements a specific physical domain, e.g. atmosphere, ocean
- **NUOPC_Mediator** - Is used for custom coupling code (flux calculations, averaging, etc.) between multiple Models.
- **NUOPC_Connector** Simple transforms that may perform straightforward operations such as grid remapping or data redistribution.



Architecture of a typical Earth System Model using the NUOPC Layer: Models are connected via simple Connectors o a central Mediator that performs merges and more complex operations.

Hierarchies, nesting, explicit, semi-implicit, and implicit coupling, ensembles, and other configurations are supported.

NUOPC Layer Status

- NUOPC Layer code is being finalized for release in early spring 2013
- Early versions are already being implemented in applications
- In 2013, implementation of the NUOPC Layer in applications is ramping up significantly
  - Development funding shifting over to applications
  - Network of plans for bringing in many new models (essentially all major U.S. NOAA, NASA, and DoD Earth system models): http://www.earthsystemcog.org/projects/nuopc/roadmap/

ESMF

Operational and Research Modeling Systems Implementing NUOPC Interoperability

- Global Forecast System (GFS)
- Global Ensemble Forecast System (GEFS)
- North American Mesoscale Model (NMM)
- Finite Element Icosahedral Model (FIM)
- NOAA Environmental Modeling System (NEMS)
- Climate Forecast System (CFS)
- Weather Research and Forecasting Model (WRF)
- NASA GISS Model E
- NASA GEOS-5
- GFDL MOM5-LANL CICE
- Naval Operational Global Atmospheric Prediction System (NOGAPS) and NavGEM
- Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS)
- Navy Coastal Ocean Model (NCOM)
- Hybrid Coordinate Ocean Model (HYCOM)
- Wave Watch 3 (WW3)
- Community Ice Code (CICE)
- Ensemble Forecast System (EFS)
- Simulating Waves Near Shore (SWAN)

Why does ESMF need a Python Interface?

- Enables ESMF regridding to be used with very little effort, in an object oriented way:
  - Regridding applied as a callable Python object
  - Numpy array access to distributed data
  - Some users report computation times reduced from hours to minutes
- Enables ESMF regridding to be used in other scientific packages with Python-based workflows – current users include:
  - UV-CDAT (PCMDI) – Ultrascale Visualization Climate Data Analysis Tools
  - PyFerret (NOAA) – Python based interactive visualization and analysis environment
  - Community Surface Dynamics Modeling System (CSDMS) – tools for hydrological and other surface modeling processes
- See: http://www.earthsystemcog.org/projects/esmp/
- Next steps: time management classes, component classes

CUPID: An IDE for Model Development and Modeler Training

Motivation

Tools for development and use of Earth System Models tend to be either highly flexible and targeted for experts (e.g. FRE at GFDL) or GUIs that enable novice users to run models but are not designed to support code changes (e.g. CESM Portal at Purdue).

Integrated Development Environments (IDEs) offer an appealing middle path that can enable students and other users to change and test model source in an environment that provides guidance, task automation, and code generation.

CUPID: An IDE for Model Development and Modeler Training

Development Paths

- Development and deployment of CUPID
- Adaptation of GISS Model E for ESMF/NUOPC
- Integration of CUPID with reworked Model E

Team

- Georgia Institute of Technology (Dunlap/Mark/Rugaber)
- University of Colorado/CIRES (DeLuca)
- NASA Software Systems Support Office (Clune)
- NASA GISS (Schmidt)

Impacts

- More modular, flexible, interoperable, and extensible version of GISS Model E
- Development environment for Model E for students, new users, and scientists/programmers
- Demonstration of IDE viability for the ESM community
- Extensible platform for other models (e.g. GEOS) to adopt

Web Services Pilot: Climate-Hydro Coupling

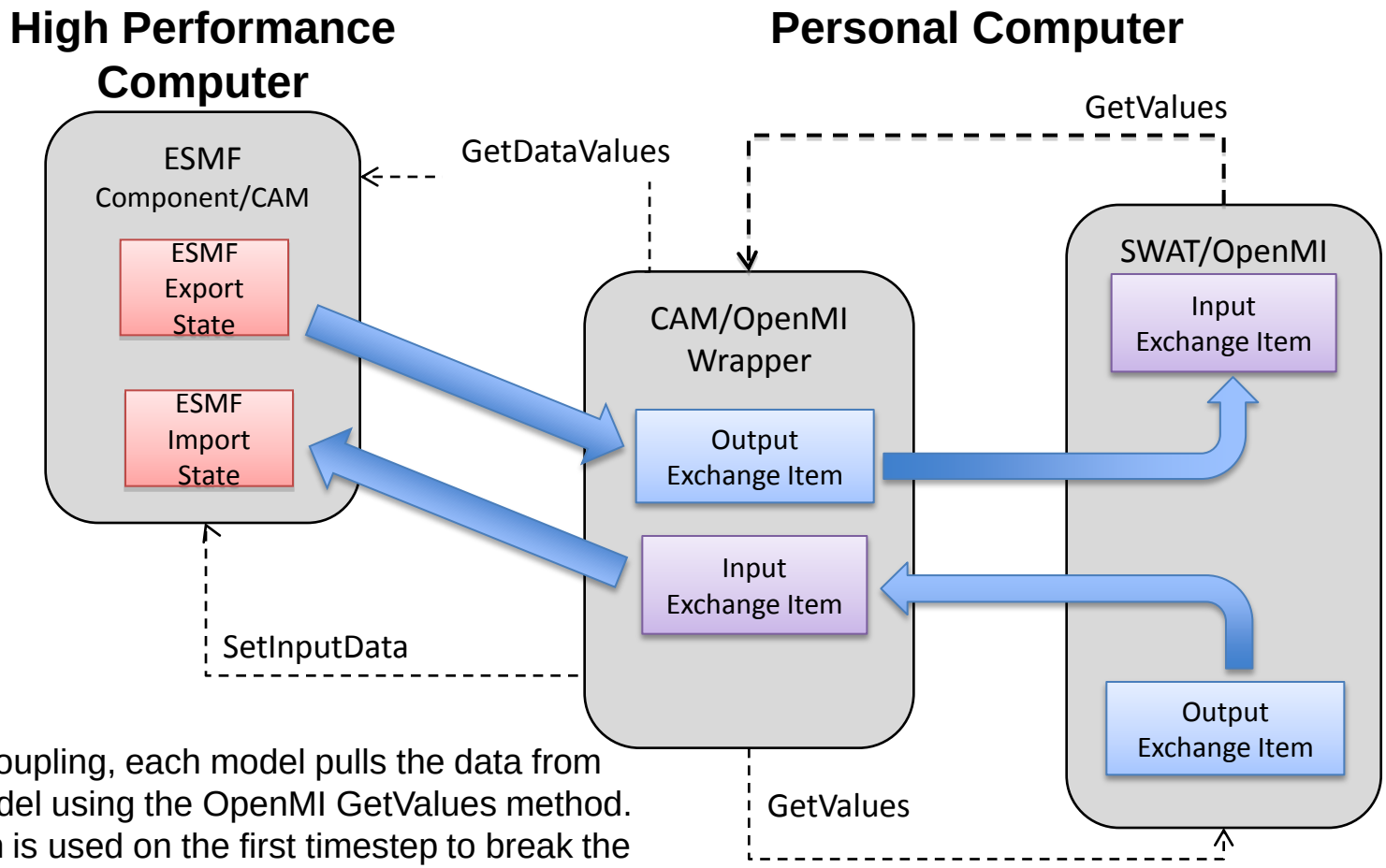
GOAL: Develop a two-way coupled, distributed, ***service-based*** modeling system comprised of an HPC atmospheric climate model and a hydrological model, utilizing component interfaces from each domain (ESMF, OpenMI).

IDEA: Support **distributed, heterogeneous service networks** for prediction that incorporate high performance computing platforms

- Two way technical coupling completed during 2012 uses the Community Atmospheric Model (with active land) and the Soil Water Assessment Tool (SWAT)
- Utilized simple switch in CAM's ESMF component wrapper that enables web service operation
- Coupled system is run using the OpenMI configuration editor, a web service driver developed in the hydrological community

STATUS: Prototype completed - next steps: using WRF, both models invoked as services, utilization of cloud services

Data Flow
Two-Way Coupling



- In two-way coupling, each model pulls the data from the other model using the OpenMI GetValues method. Extrapolation is used on the first timestep to break the deadlock between the two model requests.
- OpenMI Input and Output Exchanges items are again used to exchange and translate the data.

ESMF in Context

The **NESII** team builds on and integrates a set of community-developed, established technologies, satisfying emerging requirements by combining, extending, and transforming these core elements

Core technologies
- Major community resources
- 5-10 years development

modeling tools

data services

metadata services

Building on the core
- Using ESMF

- Using ESGF and ES-DOC

- CoG: Wiki and data environment to link and keep track of our various efforts, and also to conduct model intercomparison projects (e.g. DCMIP)
- OpenClimateGIS: Tool for subsettinng, calculating indices, and converting data to GIS compatible formats, beginning integration with ESMP