Update on the Earth System Modeling Framework (ESMF)

Rocky Dunlap[1]

Ben Koziol[2], Peggy Li[3], Fei Liu[4], Bob Oehmke[1], Ryan O'Kuinghttons [1], Himanshu Pillai[1], Dan Rosen[2], Gerhard Theurich[4], Ufuk Turuncoglu[1]

1 National Center for Atmospheric Research

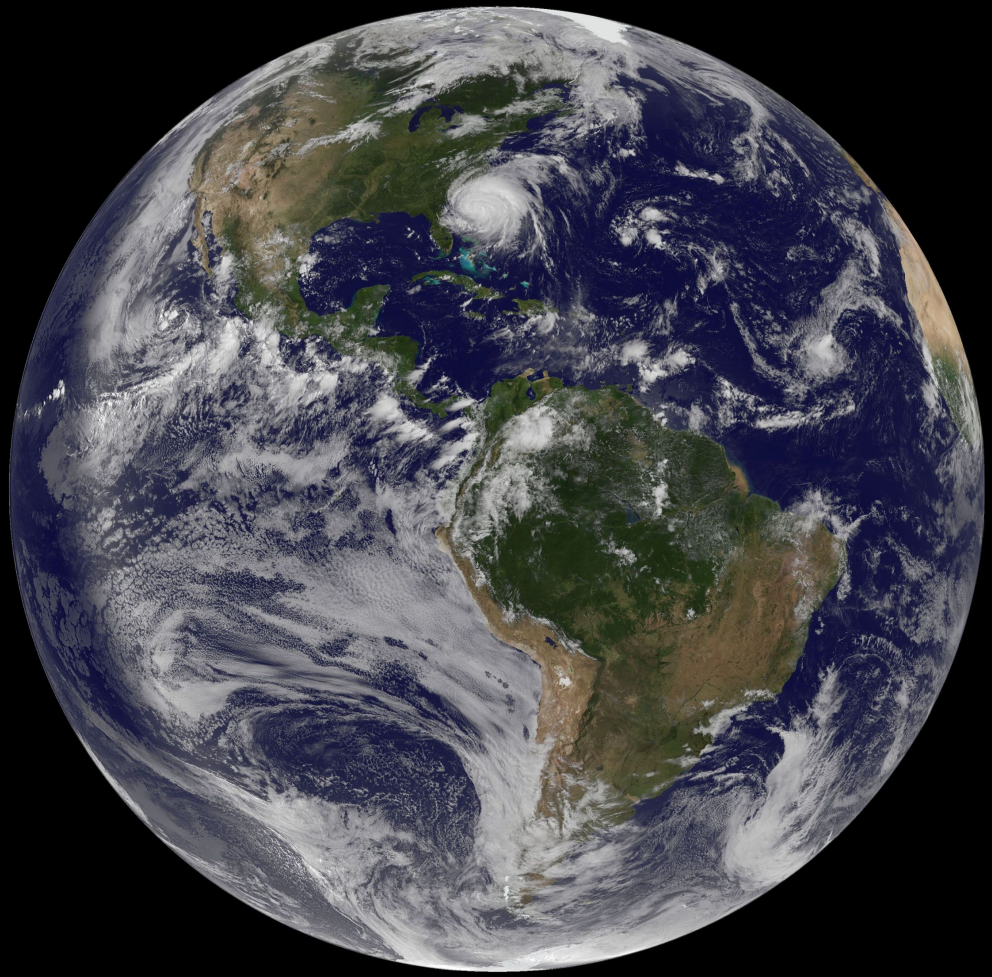
2 NOAA Global Systems Laboratory / U. of Colorado/CIRES;

3 NASA Jet Propulsion Laboratory;

4 Naval Research Laboratory

September 21, 2020

5th Workshop on Coupling Technologies for Earth System Models (CW2020)



Hurricane Irene/NASA GOES-13 satellite image/August 26, 2011

Outline

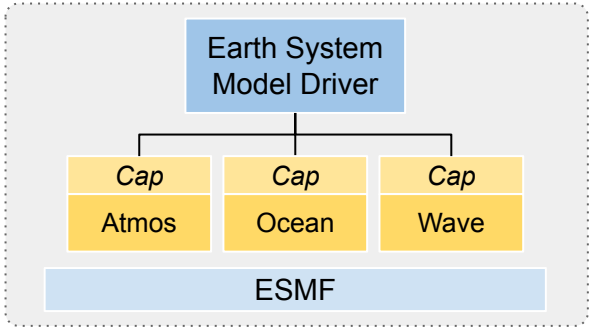
- Overview of ESMF
- The NUOPC Interoperability Layer
- Coupled data assimilation interface with ESMF/NUOPC
- Update on ESMF releases

Earth System Modeling Framework

Overview

The Earth System Modeling Framework (ESMF) is high-performance software infrastructure used in coupled Earth science applications.

There are different ways to use ESMF:



Coupling infrastructure *in a modeling system (includes the NUOPC Layer)*

Command line

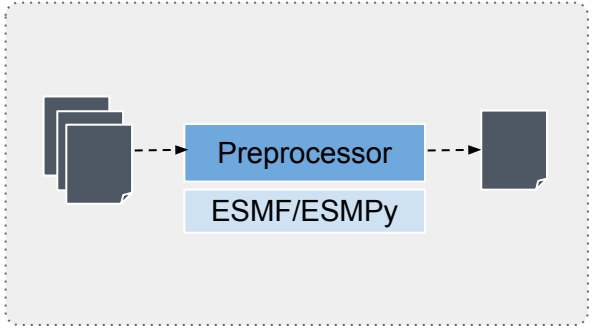
```
$ ./ESMF_Regrid \
  -s src.nc \   # source
  -d dst.nc     # destination
```

Python

```
import ESMF

ESMF.Regrid(src, dst, …)
```

Offline tool *for* ***grid remapping*** *and* ***interpolation weight generation*** *(command line and Python)*



Library *used to construct* ***custom tools****, such as preprocessor or postprocessor*

NUOPC Interoperability Layer

National Unified Operational Prediction Capability

NUOPC is a software layer included with ESMF that includes four **generic components** that provide **standard coupling protocols** and **increase interoperability** of model components.

- **NUOPC generic components**
  A NUOPC component is an ESMF component with specified rules of behavior depending on the component's role in the coupled system.

Driver

Driver

A Driver has one or more child components and is responsible for coordinating their initialization sequence and driving them through a customizable run sequence.

Connector

Connector

A Connector performs standard communication operations, in parallel, between other components, such as grid remapping and redistribution of data. Connectors have a built-in field matching algorithm based on standard names.

Model

Model

A Model "cap" wraps a geophysical model code with standard initialization and run methods so it can be plugged into a Driver.

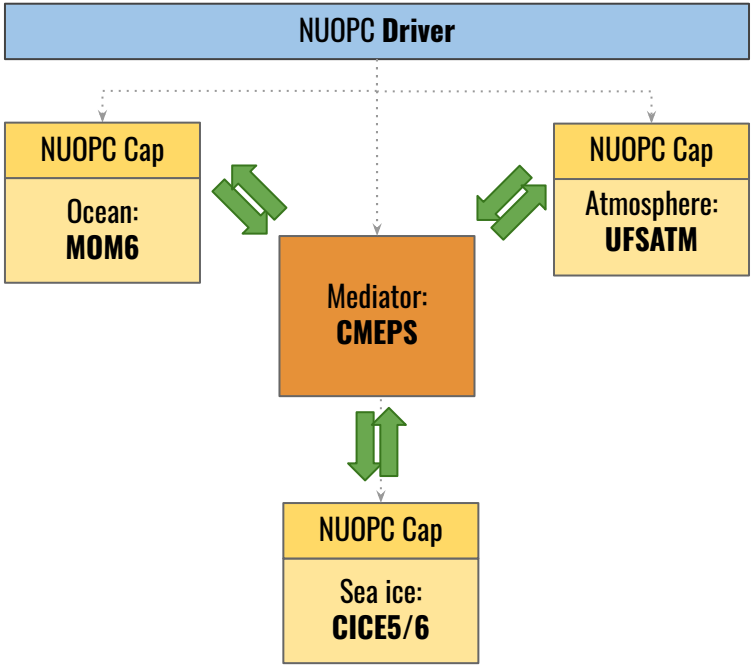
Mediator

Mediator

A Mediator contains custom coupling code such as flux calculations, accumulation/averaging, and merging of fields among several components.

Driver: Flexible Control Structures

- ESMF/NUOPC applications are typically compiled into a single executable with a single top-level driver
- Run sequences specified in small, user-friendly text file (no code!); this can be modified without recompiling
- Model components are optionally included/excluded during initialization, allowing different configurations without code changes
- Drivers are easily extended with new components, and can support multiple models of the same type (e.g., multiple ocean models)

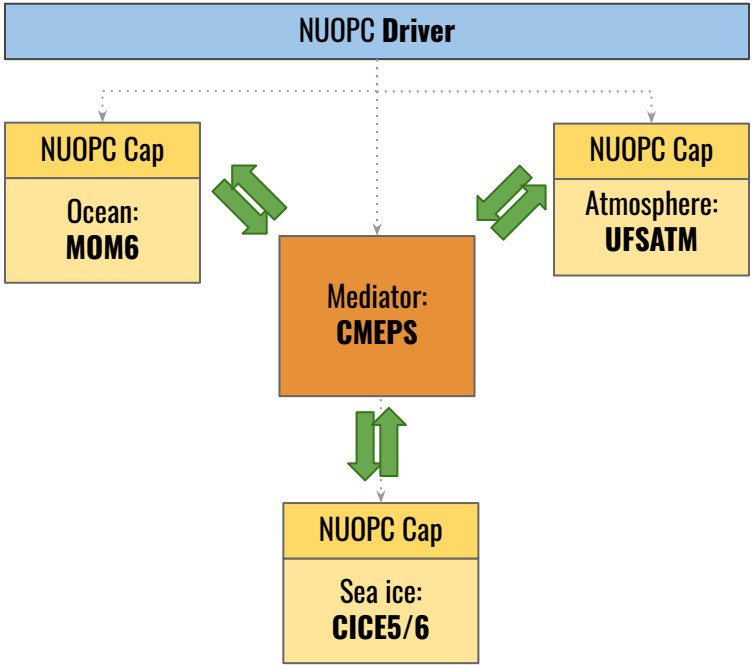


A configuration of the Unified Forecast System (UFS) S2S Application

Driver: User-friendly Run Sequence

```
runSeq::
@1800
  MED med_phases_prep_ocn_accum_avg
  MED -> OCN :remapMethod=redist
  OCN
  @600
    MED med_phases_prep_atm
    MED med_phases_prep_ice
    MED -> ATM :remapMethod=redist
    MED -> ICE :remapMethod=redist
    ATM
    ICE
    ATM -> MED :remapMethod=redist
    ICE -> MED :remapMethod=redist
    MED med_fraction_set
    MED med_phases_prep_ocn_map
    MED med_phases_prep_ocn_merge
    MED med_phases_prep_ocn_accum_fast
    MED med_phases_profile
  @
  OCN -> MED :remapMethod=redist
@
::
```

Run sequence syntax replaces hundreds of lines of Fortran code



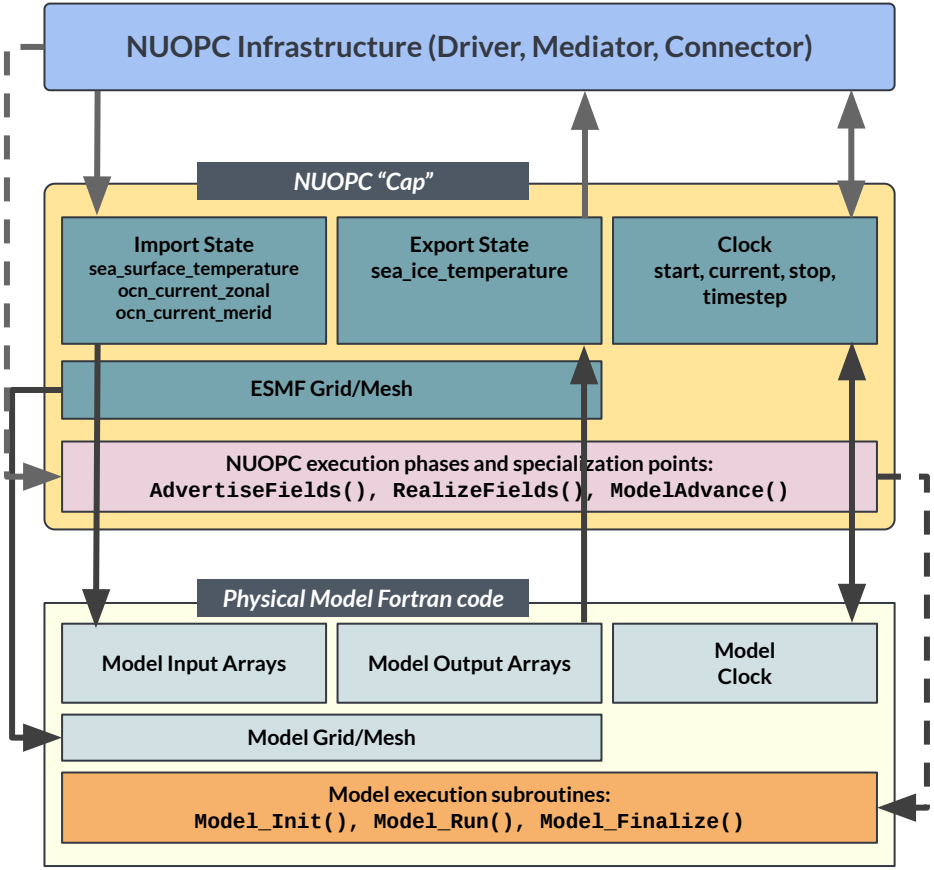
A configuration of the Unified Forecast System (UFS) S2S Application

NUOPC “Caps”

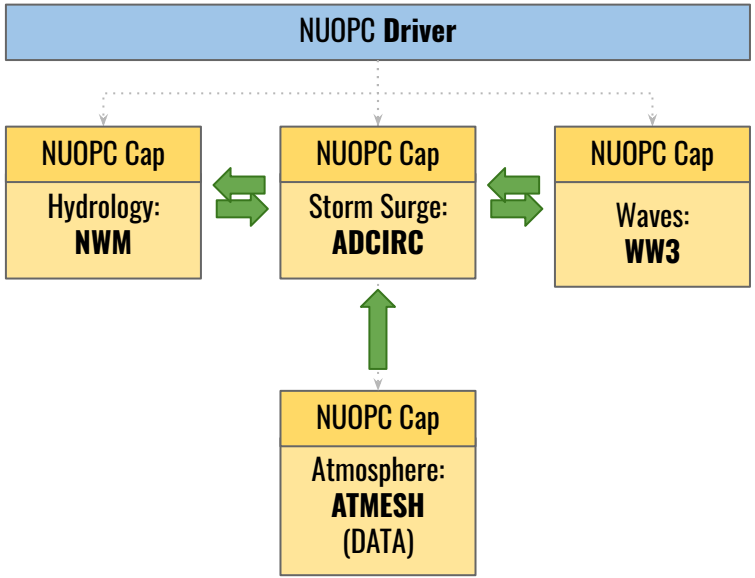
- NUOPC “caps” are non-intrusive - a small translation layer; usually a single source code file
- A NUOPC “cap” lives in a model component’s authoritative repository and is shared across different community modeling systems -- i.e., only one NUOPC “cap” per model
- Import and export coupling fields are identified by standard names; model internal names do not have to change
- Supports 1D, 2D, and 3D coupling fields and a wide variety of structured grids and unstructured meshes; global and regional
- Adapts to native memory layouts already used by the underlying model

Example “cap” code from MOM6 ocean model:
code, **docs**

See *Earth System Prediction Suite* for list of NUOPC-compliant models:
https://earthsystemprediction.gov/About/ESPS



Connectors: Optimized Inter-model Communication

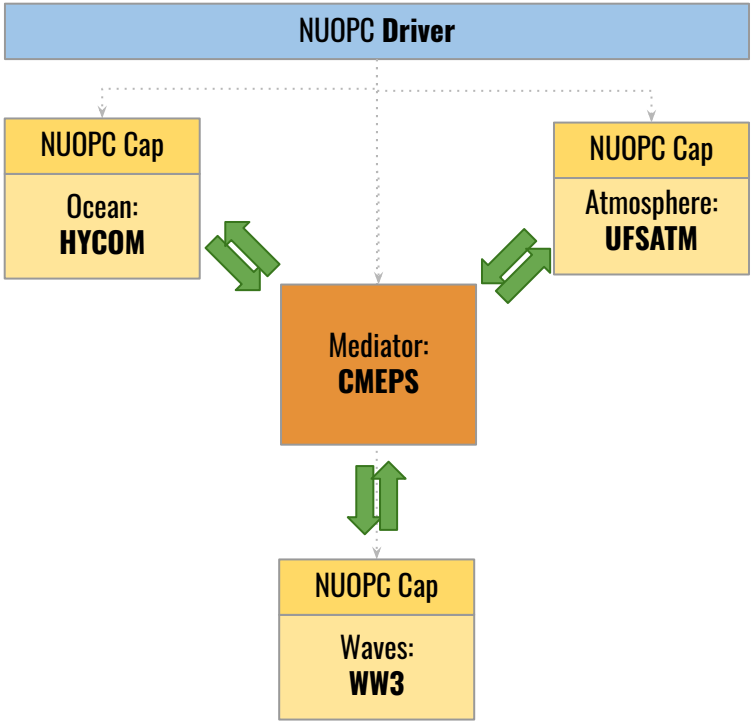


Named Storm Event Model (NSEM) configuration of the UFS Coastal Application

- NUOPC "Connectors" are generated automatically by the Driver - no user code!
- Connectors determine at runtime which coupling fields need to be exchanged, removing hard-coded field mappings
- Connectors provide fast parallel communication options, including online generation and applications of interpolation weights
- Large number of interpolation options: bilinear, patch, conservative (1st and 2nd order), nearest neighbor; extrapolation
- Connectors negotiate the most optimized connection possible, allowing tight coupling (shared memory) to loose coupling (grid remapping)
- Grids can be transferred through a Connector (e.g., a model component can inherit a grid from another model component)

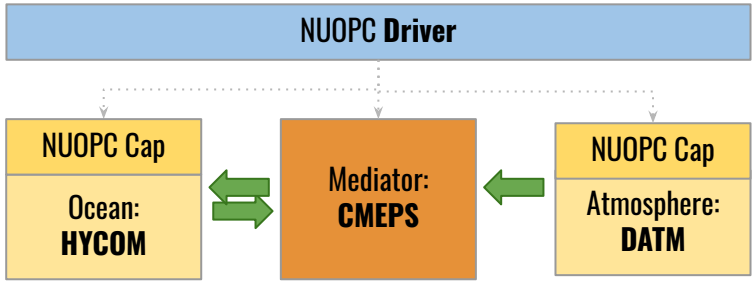
Mediators: Multi-Model Coupling Interactions

- NUOPC “Mediators” manage complex interactions involving several models
- Mediators can be used to create a “hub and spoke” architecture with a centralized coupler
- Each model component tells the Mediator what fields it requires and can provide
- Mediators leverage the full power of ESMF parallel communication and online regridding
- User has full control over how to map and merge fields among the components, e.g., combine source fields from multiple models and send to destination model
- Mediators organize code and promote “separation of concerns”



Hurricane Application

Incremental Building and Testing of Coupled Configurations

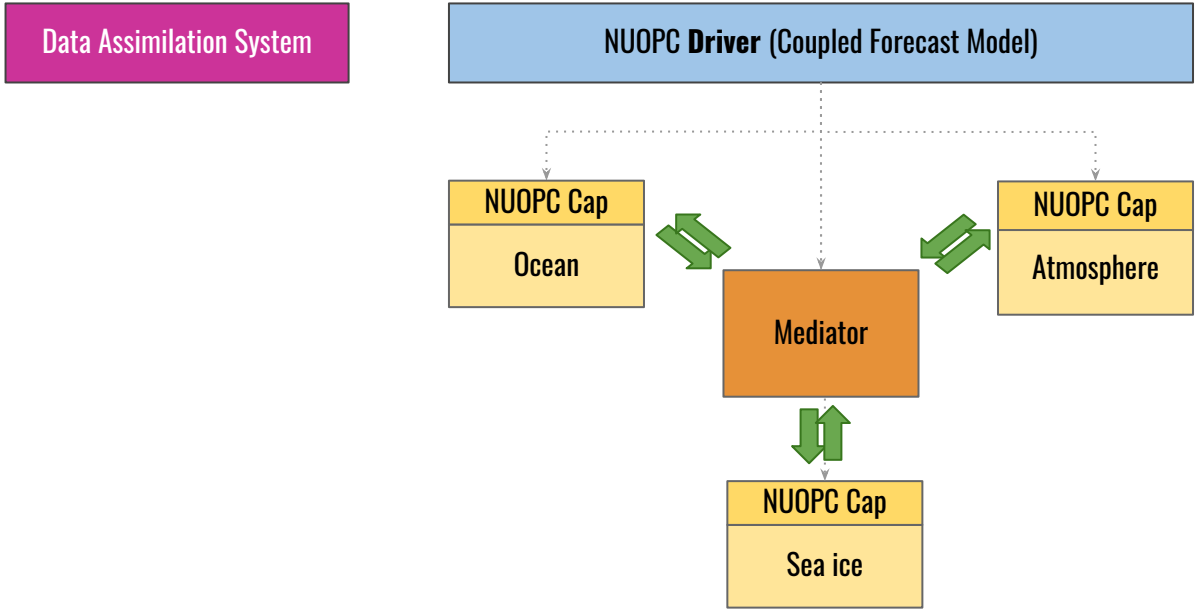


Configuration of a Hurricane Application with a “Data Atmosphere”

- ESMF/NUOPC supports incremental building and testing of coupled configurations
  - the NUOPC “cap” approach simplifies the process of substituting a “data” component for an active model
  - enables “hierarchical model development” by isolating coupling feedbacks
  - Mediator interfaces do not change when connecting data models
  - “data” models are typically less expensive to run than active models, speeding up the development process

Coupled Data Assimilation Architecture

How should data assimilation systems interface with coupled models using the ESMF/NUOPC architecture?



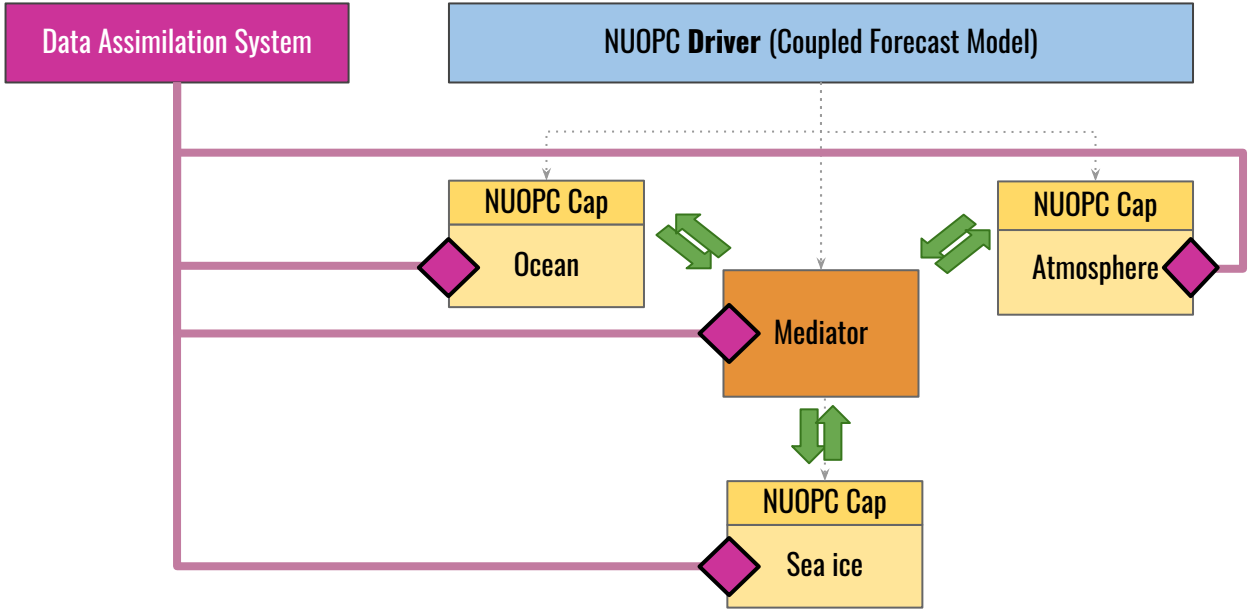
Coupled Data Assimilation Architecture

How should data assimilation systems interface with coupled models using the ESMF/NUOPC architecture?

<u>Design Option 1:</u>

Develop a set of *separate model interfaces* that can be called by the data assimilation (DA) system.

- Difficult to retain the physics coupling in the same way as the coupled forecast
- DA system needs to reproduce the coupled model driver and run sequence
- Potential redundancy in model interfaces, e.g., access to model state and control methods



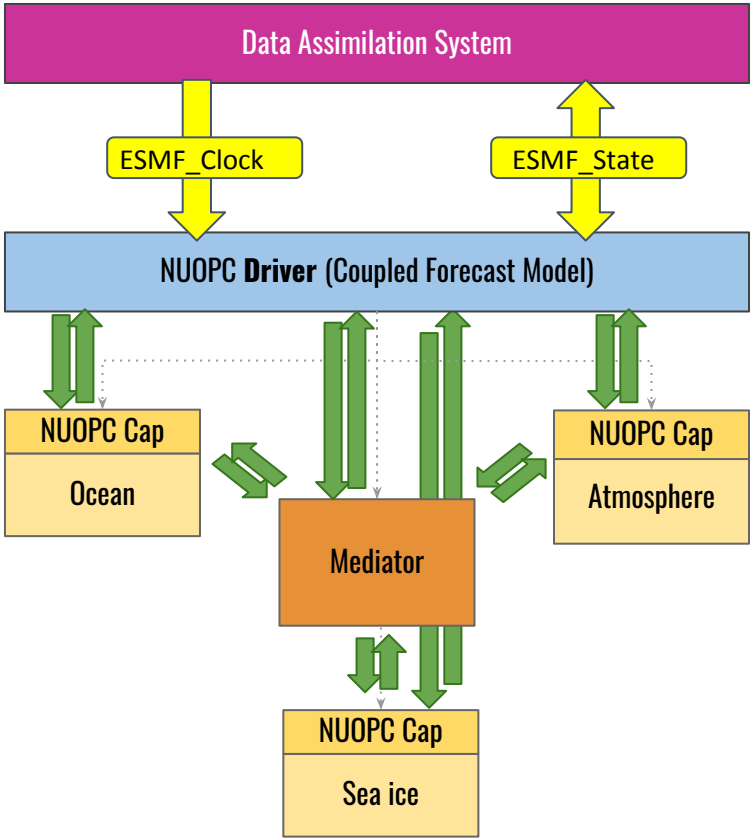
Coupled Data Assimilation Architecture

How should data assimilation systems interface with coupled models using the ESMF/NUOPC architecture?

<u>Design Option 2:</u>

DA system accesses the NUOPC-based coupled model through the Driver. “Vertical connectors” provide access to coupled model state.

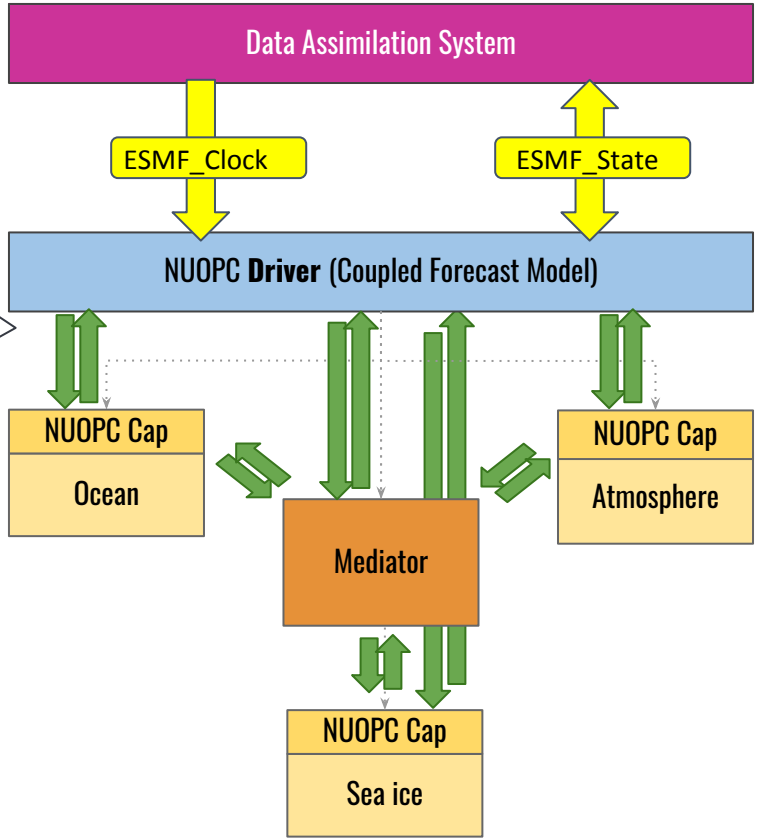
- Single entry point for accessing the forecast model
- NUOPC Driver is retained including the run sequence of the component models
- Access to individual model and coupled model state
- ESMF regridding used to interpolate between model state and observation locations
- Shared memory references through Connectors reduce high-volume communication
- ESMF redistribution moves model data to DA system memory layouts
- Coupled model state can be updated by the DA system



Coupled Data Assimilation Architecture

```
runSeq::
@1800
  MED med_phases_prep_ocn_accum_avg
  MED -> OCN :remapMethod=redist
  OCN
  @600
    MED med_phases_prep_atm
    MED med_phases_prep_ice
    MED -> ATM :remapMethod=redist
    MED -> ICE :remapMethod=redist
    ATM
    ICE
    ATM -> MED :remapMethod=redist
    ICE -> MED :remapMethod=redist
    MED med_fraction_set
    MED med_phases_prep_ocn_map
    MED med_phases_prep_ocn_merge
    MED med_phases_prep_ocn_accum_fast
    MED med_phases_profile
  @
  OCN -> MED :remapMethod=redist
@
::
```

Return to DA layer as requested



“External” NUOPC Interface

- A new API has been added to support “external” applications (such as data assimilation) that need to interface with a top-level NUOPC Driver
- Provides a set of “hooks” into key points of the initialization, run, and finalize
- For DA, this allows the DA system to:
  - request a set of fields from the coupled model
  - send a set of updated fields
  - run the forecast for a set duration
  - reset the forecast clock and request the model components to reinitialize

methodFlag	phaseLabel	Meaning
ESMF_METHOD_INITIALIZE	label_ExternalAdvertise	Called after the external component has set up the import- and exportStates with fields that it plans to interact with. On the NUOPC application side this call will got through the complete advertise cylce.
ESMF_METHOD_INITIALIZE	label_ExternalRealize	Called after the external component has been informed about the connected status of the fields in the import- and exportState. On the NUOPC application side this call will finish setting up RouteHandles between all components involved.
ESMF_METHOD_INITIALIZE	label_ExternalDataInit	Trigger a complete data initialize throughout the NUOPC application. The expectation is that all components reset their data consistent with the `clock` argument.
ESMF_METHOD_RUN		The default Run() method steps the NUOPC application forward in time according to the `clock` argument.
ESMF_METHOD_FINALIZE	label_ExternalReset	Inform the NUOPC application about a `clock` reset.
ESMF_METHOD_FINALIZE		Completely finalize and shut down the NUOPC application.

See NUOPC Reference Manual

ESMF 8.0.1 Release

8.0.1 was released May 21, 2020 and is the latest release

See 8.0.1 release notes for full list of release items:
https://www.earthsystemcog.org/projects/esmf/download/

Highlights in ESMF (since 7.1.0):

- Added support for **GRIDSPEC Mosaic files** and different UGRID stagger locations to the **ESMF_Regrid** application
- The **ESMF_FileRegrid** API supports multi-tile mosaics stored in separate files, 2nd order conservative, UGRID corner stagger, and multiple variable regridding.
- Added **"creep fill" extrapolation** method
- **RouteHandles can be written to file** and read in from file
- **Irregular decompositions** supported for **multi-tile** grids
- Support for **"packed" Field Bundles**, with fields interleaved or packed in memory
- **Sharing of decomposition elements between PETs** on the same node
- Output **component timing profiles** to per-PET or summary text files

New in NUOPC:

- **Field dictionary** can be read in from **YAML file**
- **NUOPC Driver can be called from higher level** application, such as a **data assimilation system**
- **Switch between run sequences** during execution
- Connector improvements for **sharing Field and GeomObjects**, **reuse of redist RouteHandles, handling multiple nests**

New in ESMPy:

- **In-memory weight generation** option, allowing re-use of weight vectors (NumPy or dict) without writing them to NetCDF.
- Ability to **write/read RouteHandles** to/from file, a kind of cache to speed up regridding

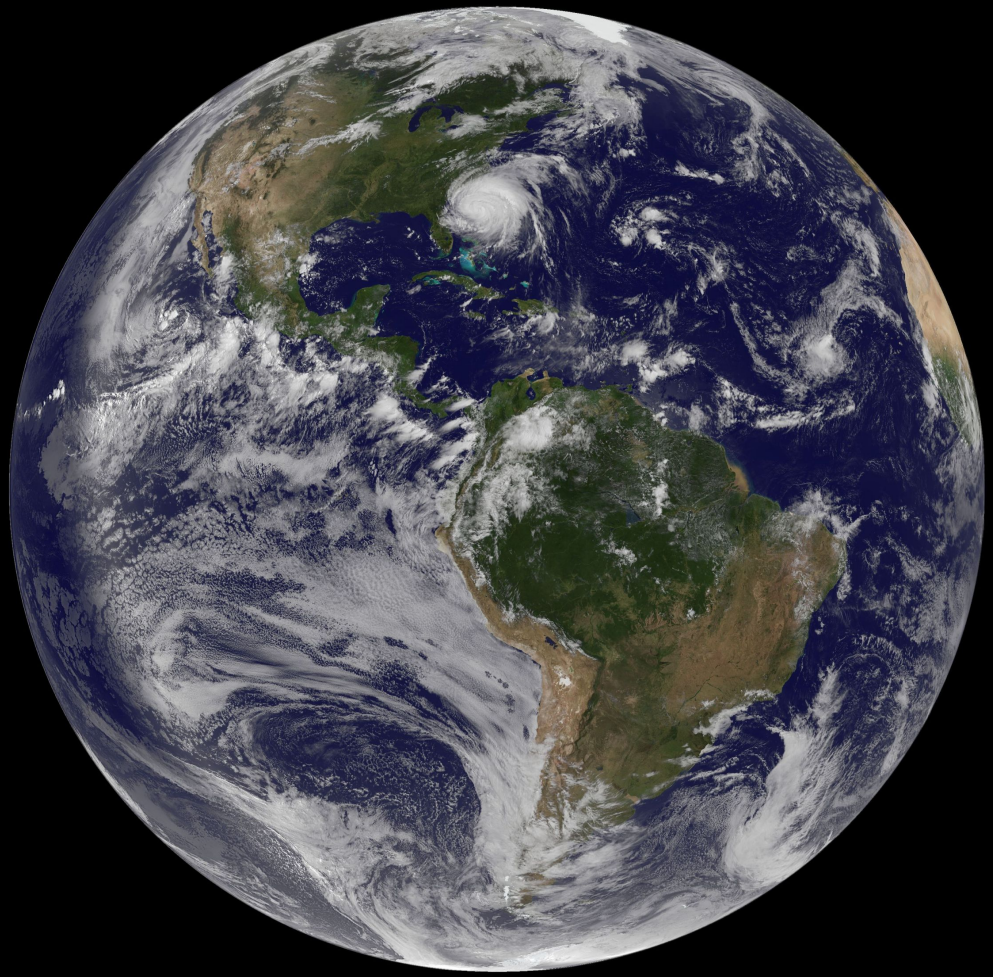
New development priorities for upcoming release: ESMF v8.1 (March 2021)

- Improved “creep fill” extrapolation that dynamically determines creep level parameter
- Full support for all communication methods on “packed” ESMF_FieldBundles
- MOAB mesh database fully integrated and all regridding methods supported. (Will replace native mesh implementation.)
- Optimizations:
  - removal of redundant “proxy objects” used to represent non-local grids/fields
  - NUOPC Connector reuse of transferred grid/mesh/locstream objects
  - testing/profiling of high resolution global grids
- Refactor ESMF_Attribute with new JSON-based implementation
- Move to a newer version of PIO and remove redundant data redistributions in I/O layer
- Implement ArrayBundle as core data structure for all major I/O operations (FieldWrite, GridWrite, MeshWrite and XGridWrite)
- Add flag to validate quality of grids/meshes at runtime (i.e., that could lead to regridding problems)
- Support for ESMF-aware hybrid parallelism via resource idling in NUOPC (switch between coarse- and fine-grained threading)

All tasks not listed. See full list here:
https://www.earthsystemcog.org/projects/esmf/schedule_8_1_0

Thank you!

- ESMF on GitHub:
  https://github.com/esmf-org/esmf
- ESMF Home Page:
  https://earthsystemcog.org/projects/esmf
- NUOPC Layer Home Page:
  https://earthsystemcog.org/projects/nuopc/
- Support:
  esmf_support@ucar.edu



Hurricane Irene/NASA GOES-13 satellite image/August 26, 2011

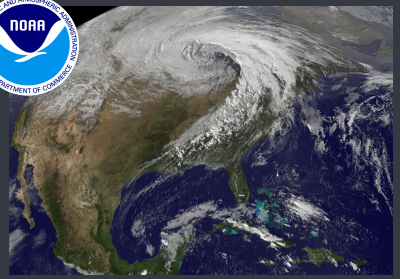
Extra Slides

Modeling Systems using ESMF/NUOPC

Multi-agency collaboration on modeling infrastructure

UFS

NOAA's Unified Forecast System



Next-generation operational prediction for weather through seasonal time scale

UFS architecture is based on ESMF/NUOPC and supports **multiple coupled modeling applications** with different model components and different coupling configurations.

COAMPS & NavGEM

Navy Regional and Global Forecasting

Research and operational weather forecasting in support of military operations and national security

Regional and global systems use ESMF/NUOPC interfaces.

Support for specialized coupling requirements with telescoping **nested domains and nest-to-nest coupling**.

GEOS & Model E

NASA Modeling and Data Assimilation

Data assimilation, utilization of satellite measurements, seasonal to climate forecasting, creation of reanalysis datasets

GEOS features a large number of ESMF components, each handling different physics, **organized into a deep hierarchy**.

CESM

Community Earth System Model

Research into all aspects of the climate system, including participation in the Intergovernmental Panel on Climate Change assessment reports

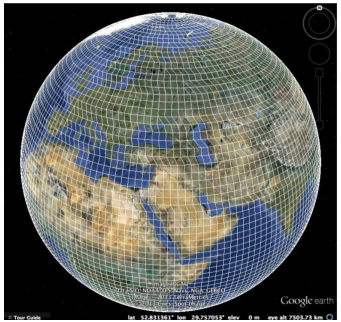
CESM's next generation coupler is based on ESMF/NUOPC, including a shared, community-developed NUOPC-compliant Mediator (CMEPS).

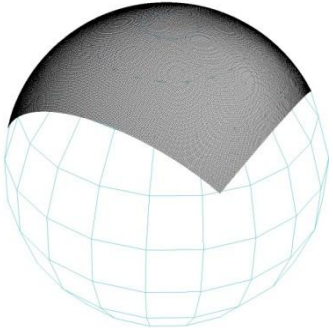
ESMF Download, User Support, and Training

- **Download/Install**
  - GitHub: https://github.com/esmf-org/esmf

- **User support** for bug fixes, technical questions, and features requests: esmf_support@ucar.edu

- Regular **training opportunities** are provided in webinar format. See training page for upcoming dates and materials from previous events: https://earthsystemcog.org/projects/esmf/tutorials



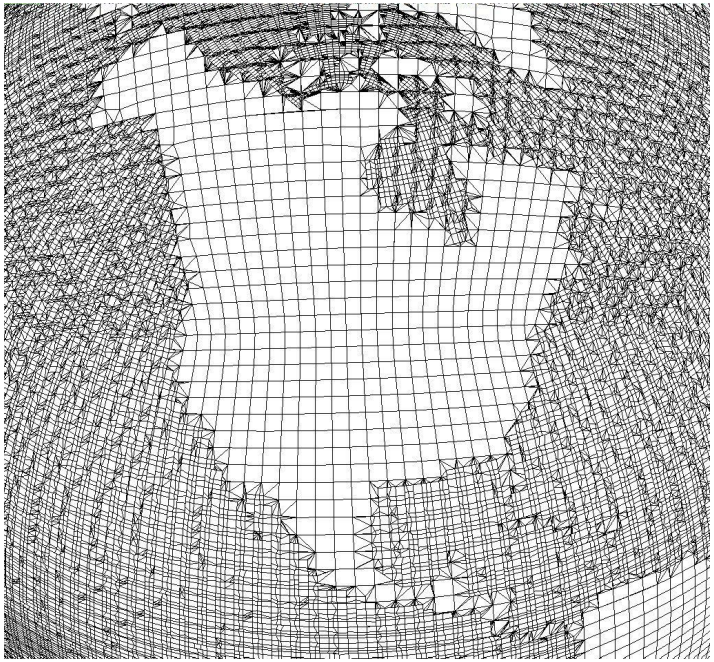


ESMF Regridding

Fast, flexible interpolation of gridded data

- ❖ High-performance
  - ➢ Interpolation weight matrix is generated in parallel in 3D space and applied in parallel
- ❖ Wide range of supported grids
  - ➢ Logically rectangular and unstructured grids in 2D and 3D, observational data streams (point cloud), global and regional grids, Cartesian and spherical coordinates
- ❖ Multiple interpolation methods
  - ➢ Bilinear, higher-order patch recovery, nearest neighbor, first order conservative, second order conservative
- ❖ Options
  - ➢ Masking, multiple pole treatments, straight or great circle distance measure
- ❖ Multiple interfaces
  - ➢ **Fortran API** - generate and apply weights during a model run
  - ➢ **Python API** - generate and apply weights using ESMPy
  - ➢ **File-based** - generate and apply weights from grid files using ESMF command line utilities

Implicit Coupling with NUOPC Components

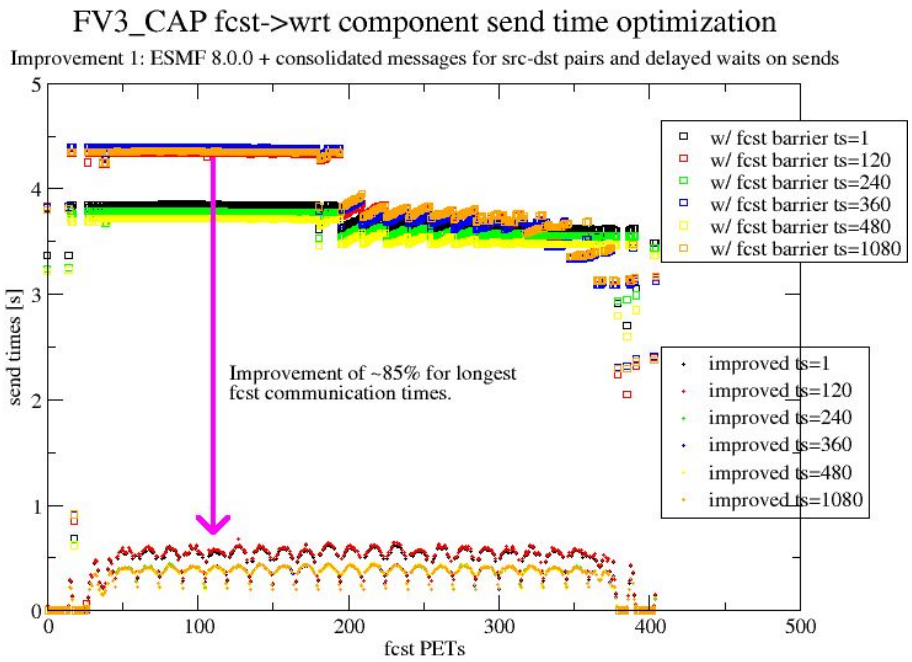


Visualization of exchange grid generated from C48 cubed sphere (atmosphere/land) and 720x400 tripolar grid. Image courtesy Bob Oehmke.

- The modularity provided by NUOPC components does *not* limit applications to explicit coupling.
- *Implicit coupling* has advantages such as improving model stability, flexible time-stepping, and conserving fluxes.
- An *exchange grid* (XGrid) is a data structure that represents the surface boundary layer usually between the atmosphere on one side and ocean/land on the other.
- Within NUOPC, a *Mediator* component creates an exchange grid and a multiphase run sequence with atmosphere, mediator, and land allows for an implicit solver through the atmosphere-land column.
- Code example demonstrates atmosphere-land implicit coupling through a Mediator using NUOPC components: https://github.com/ESCOMP/ImplictCouplingXGrid
- Report on performance and accuracy comparison between ESMF XGrid and GFDL exchange grid implementations.

Asynchronous I/O

- ESMF/NUOPC provides asynchronous I/O in the UFS Atmosphere
  - Hides time to write model output by overlapping disk access with model forecast
  - UFS Atmosphere divided into two sub-components: *forecast* and *write*
  - Treat as a coupling problem, and leverage ESMF regridding to send data
  - Recent optimizations in ESMF 8.0.1 reduce communication time for asynchronous I/O and other communication methods with a disproportionate number of sending/receiving processes



Reduction in send times from UFSATM forecast component to write component were required for GFSv16 to meet operational 8min/day requirement. ESMF 8.0.0 vs. 8.0.1.