ESMF Strategies to Address HPC Challenges

Raffaele Montuoro

NOAA/CIRES
NOAA Environmental Software Infrastructure and Interoperability

May 18, 2018

5th ENES Workshop on High Performance Computing for High-Resolution Climate and Weather Modelling
Officine Cantelmo, Lecce, Italy

ESMF NUOPC

ESPC
EARTH SYSTEM PREDICTION CAPABILITY

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Horizon 2020 Programme

Outline

- **Trends in Earth System Modeling lead to new HPC challenges**
  - Increasing model complexity
  - Heterogeneity in computing systems
  - Increasing use of community models and infrastructure
- **The role of ESMF/NUOPC**
  - ESMF 7.1.0r provides advanced building blocks for HPC applications
  - The NUOPC Layer standardizes architectural options for interoperable coupled model components
- **Addressing HPC challenges from the bottom up**
  - Focus on fast, flexible tools for next-generation prediction systems to address model complexity and leverage heterogeneous computing
  - Deliver development tools and training programs to support growing user community

Outline

- **Trends in Earth System Modeling lead to new HPC challenges**
  - Increasing model complexity
  - Heterogeneity in computing systems
  - Increasing use of community models and infrastructure
- **The role of ESMF/NUOPC**
  - ESMF 7.1.0r provides advanced building blocks for HPC applications
  - The NUOPC Layer standardizes architectural options for interoperable coupled model components
- **Addressing HPC challenges from the bottom up**
  - Focus on fast, flexible tools for next-generation prediction systems to address model complexity and leverage heterogeneous computing
  - Deliver development tools and training programs to support growing user community

Trends in Earth System Modeling

Increasing model and hardware complexity lead to new HPC challenges

- **Increasing model complexity**

- More processes modeled, more components
- Growing demand for architectural options – coupled systems, hierarchies, ensembles, nested domains
- Interest in modeling systems that span multiple temporal and spatial scales
- Increasingly high resolution, unstructured and dynamic grids

Trends in Earth System Modeling

Increasing model and hardware complexity lead to new HPC challenges

- **Heterogeneity in computing systems**

- Emergence of fine-grained architectures
- Emergence of cloud computing and elastic resources
- Connections between traditional high performance computing and other platforms (e.g. GIS-based, mobile)
- Optimization challenges

Trends in Earth System Modeling

Increasing model and hardware complexity lead to new HPC challenges

- **Increasing use of community models and infrastructure**

- Demand for user-friendly development environments

- Need for tools and training programs

Outline

- **Trends in Earth System Modeling lead to new HPC challenges**
  - Increasing model complexity
  - Heterogeneity in computing systems
  - Increasing use of community models and infrastructure
- **The role of ESMF/NUOPC**
  - ESMF 7.1.0r provides advanced building blocks for HPC applications
  - The NUOPC Layer standardizes architectural options for interoperable coupled model components
- **Addressing HPC challenges from the bottom up**
  - Focus on fast, flexible tools for next-generation prediction systems to address model complexity and leverage heterogeneous computing
  - Deliver development tools and training programs to support growing user community

The Role of ESMF/NUOPC

The latest ESMF release provides improved building blocks for HPC applications

- The Earth System Modeling Framework (**ESMF**) is community-developed, community-governed software for building and coupling model components

- The National Unified Operational Prediction Capability (**NUOPC Layer**) is a set of extensions to ESMF that increases component interoperability and adds architectural options

- **ESMF version 7.1.0r** was released on March 8, 2018

- Download from:

  earthsystemcog.org/projects/esmf/download_710r

ESMF Metrics:

7000 downloads

150 components in use

4000 subscribers to info mailing list

40 platform/compilers regression tested nightly

8000 regression tests

1M SLOC

Outline

- **Trends in Earth System Modeling lead to new HPC challenges**
  - Increasing model complexity
  - Heterogeneity in computing systems
  - Increasing use of community models and infrastructure
- **The role of ESMF/NUOPC**
  - ESMF 7.1.0r provides advanced building blocks for HPC applications
  - The NUOPC Layer standardizes architectural options for interoperable coupled model components
- **Addressing HPC challenges from the bottom up**
  - Focus on fast, flexible tools for next-generation prediction systems to address model complexity and leverage heterogeneous computing
  - Deliver development tools and training programs to support growing user community

Addressing HPC challenges

ESMF/NUOPC Toolset

- **Increasing model complexity:**
  - **Support for many component configuration and coupling options**
    - NUOPC Layer architecture creates reconfigurable systems out of *generic, standard components*: driver, mediator, connector, model
    - Component hierarchies hide sub-components complexity
    - Implicit, semi-implicit, and explicit coupling
    - Pool of interoperable components (ESPS)
  - **Support for many grid types and grid remapping options**
    - Many interpolation options: from bilinear to conservative, including masking
    - Exchange grid approach to representing 2D boundary layers

Addressing HPC challenges

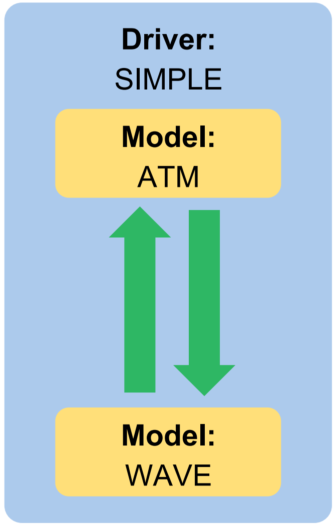
ESMF/NUOPC Toolset

- **Increasing model complexity:**
  - **Support for many component configuration and coupling options**
    - NUOPC Layer architecture creates reconfigurable systems out of *generic, standard components*: driver, mediator, connector, model
    - Component hierarchies hide sub-components complexity
    - Implicit, semi-implicit, and explicit coupling
    - Pool of interoperable components (ESPS)
  - **Support for many grid types and grid remapping options**
    - Many interpolation options: from bilinear to conservative, including masking
    - Exchange grid approach to representing 2D boundary layers

Flexible/Hierarchical Component Configurations

NUOPC Layer architectural options

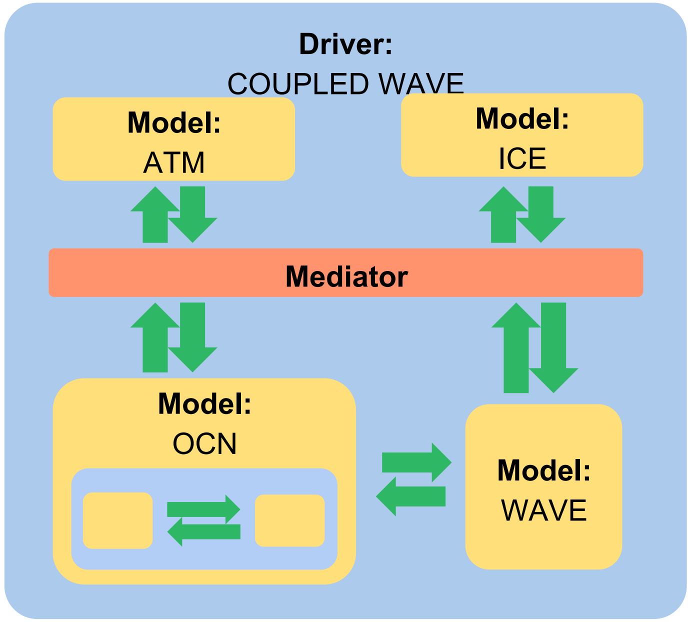
- NUOPC Layer architecture creates reconfigurable systems out of generic component structures following standard design patterns: driver, mediator, connector, model



Coupled system with a Driver, two Model components, and two Connectors

This configuration creates a coupled system that allows a two-way feedback loop between ATM and WAVE.

- Component hierarchies in the NUOPC Layer allows for hiding complexity of sub-components



A Driver with four Models and a Mediator

The OCN and WAVE components communicate directly while other components receive data only after processing by the Mediator.

The OCN component is hierarchical with an embedded driver for components representing subprocesses.

Flexible, Parameterized Run Sequences

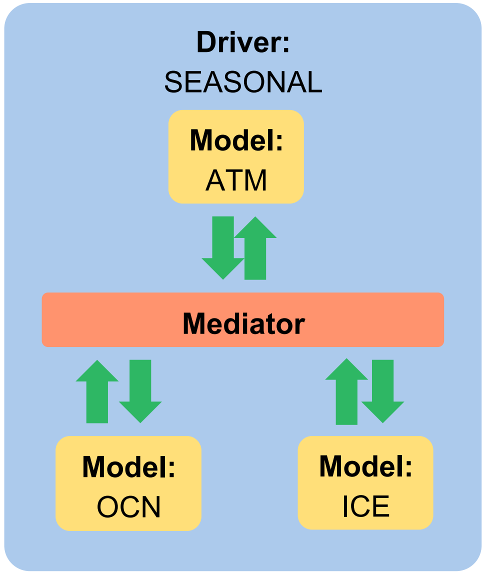
Easy to change layout, component, order of operations at run time

Colors show actions performed by:
- Connectors (->)
- Mediator (MED)
- Models

(@) indicates coupling interval

```
#######################################
# Run Time Configuration File  #
#######################################
# EARTH #
EARTH_component_list: MED ATM OCN ICE

# MED #
med_model:    nems
med_petlist_bounds: 60 65

#ATM#
atm_model:    fv3
atm_petlist_bounds:         0 31

# OCN #
ocn_model:    mom5
ocn_petlist_bounds:         32 55

# ICE #
ice_model:    cice
ice_petlist_bounds:         56 59
```

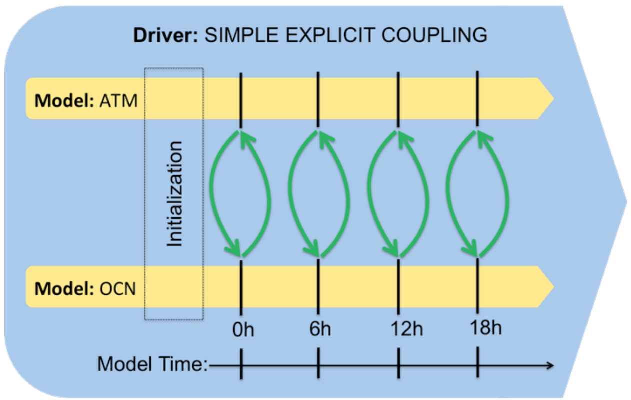
```
# Run Sequence #
runSeq::
 @1800.0
   MED MedPhase_prep_ocn
   MED -> OCN :remapMethod=redist
   OCN
   @600.0
     MED MedPhase_prep_ice
     MED MedPhase_prep_atm
     MED -> ATM :remapMethod=redist
     MED -> ICE :remapMethod=redist
     ATM
     ICE
     ATM -> MED :remapMethod=redist
     ICE -> MED :remapMethod=redist
     MED MedPhase_atm_ocn_flux
     MED MedPhase_accum_fast
   @
   OCN -> MED :remapMethod=redist
 @
::
```



Coupling Options

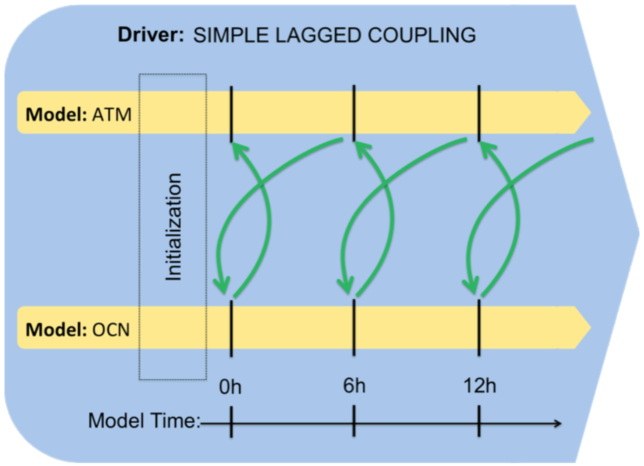
Implicit, semi-implicit, and explicit coupling

- Different predictive time scales and problems may require different types of coupling schemes:



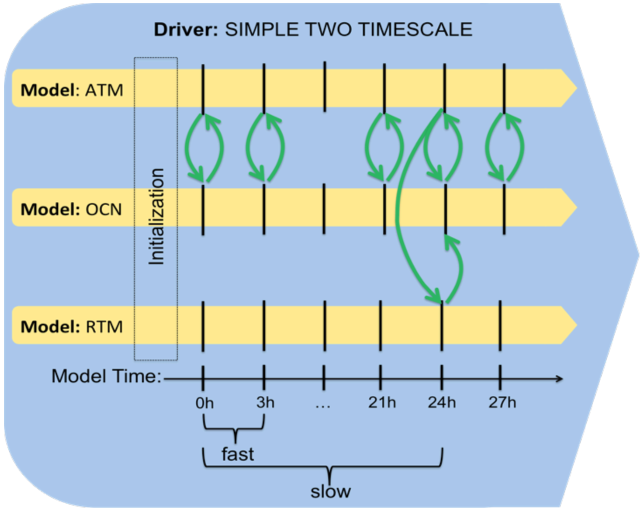
Simple explicit coupling loop

All components exchange data at the same coupling interval.



A lagged scheme

Some components may run ahead or lag behind to satisfy numerical constraints.



Multiple timescales

Some components execute and communicate less frequently.

The Earth System Prediction Suite

Pool of interoperable components

- **Interoperable components**

  The ESPS is a collection of federal and community weather and climate model components that use the Earth System Modeling Framework (ESMF) with interoperability conventions called the National Unified Operational Prediction Capability (NUOPC) Layer.

- **Model components are more easily shared across systems**

  The standard component interfaces enable major modeling centers to assemble systems with components from different organizations, and test a variety of components more easily.

COUPLED MODELING SYSTEMS

● **NUOPC-compliant**
○ In progress

	NEMS	COAMPS and COAMPS-TC	NESPC	GEOS-5	GISS ModelE	CESM
Driver(s) and Coupler(s)	●	●	●	○	○	●
ATMOSPHERE MODELS						
CAM						●
COAMPS atmosphere		●				
GEOS-5 FV atmosphere				○		
GSM	●					
ModelE atmosphere					●	
NavGEM			●			
NEPTUNE		○				
NMMB	●					
OCEAN MODELS						
HYCOM	●		●			●
MOM	●			○		
NCOM		●				
POP						●
(Additional rows not shown)						

Complete table at: earthsystemcog.org/projects/esps/

Addressing HPC challenges

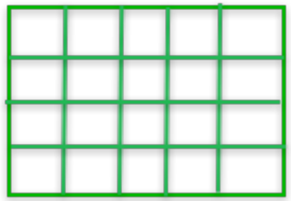
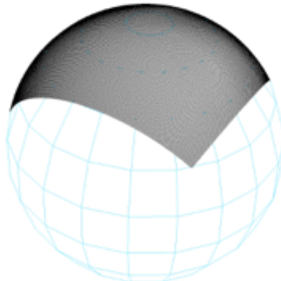
ESMF/NUOPC Toolset

- **Increasing model complexity:**
  - **Support for many component configuration and coupling options**
  - NUOPC Layer architecture creates reconfigurable systems out of *generic, standard components*: driver, mediator, connector, model
  - Component hierarchies hide sub-components complexity
  - Implicit, semi-implicit, and explicit coupling
  - Pool of interoperable components (ESPS)
  - **Support for many grid types and grid remapping options**
  - Many interpolation options: from bilinear to conservative, including masking
  - Exchange grid approach to representing 2D boundary layers

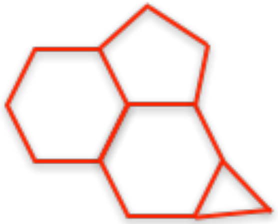
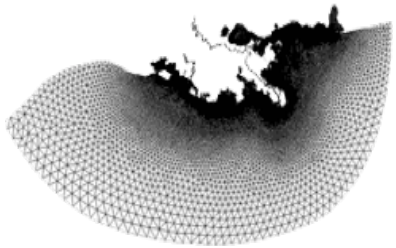
Grids, Meshes, and Location Streams

ESMF natively supports a wide range of geometry types

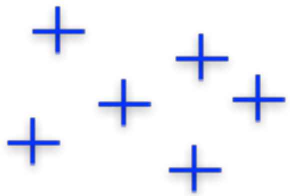
- **Grid**
  - Structured representation of a region
  - A logically rectangular tile


- **Mesh**
  - Unstructured representation of a region
  - In 2D: polygons with any number of sides (including concave)
  - In 3D: tetrahedrons & hexahedrons

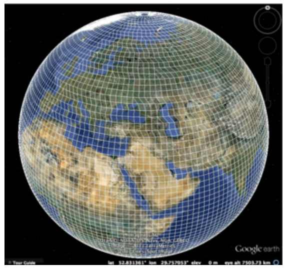



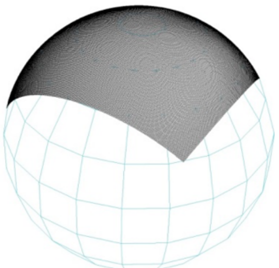

- **LocStream (Location Streams)**
  - Set of disconnected points (typically observations)
  - Very flexible and efficient
  - Can't be used with every regrid method



High Performance Grid Remapping

ESMF supports fast interpolation across a wide range of model grids





- **High-performance**
  Interpolation weight matrix is generated in parallel in 3D space and applied in parallel
- **Wide range of supported grids**
  Logically rectangular connected tiles, unstructured meshes, observational data streams (point cloud), 2D and 3D, global and regional grids, Cartesian and spherical coordinates
- **Multiple interpolation methods**
  Bilinear, higher-order patch recovery, first and second order conservative, nearest neighbor methods
- **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
- **Standalone tools** - generate and apply weights from grid files using ESMF command line utilities

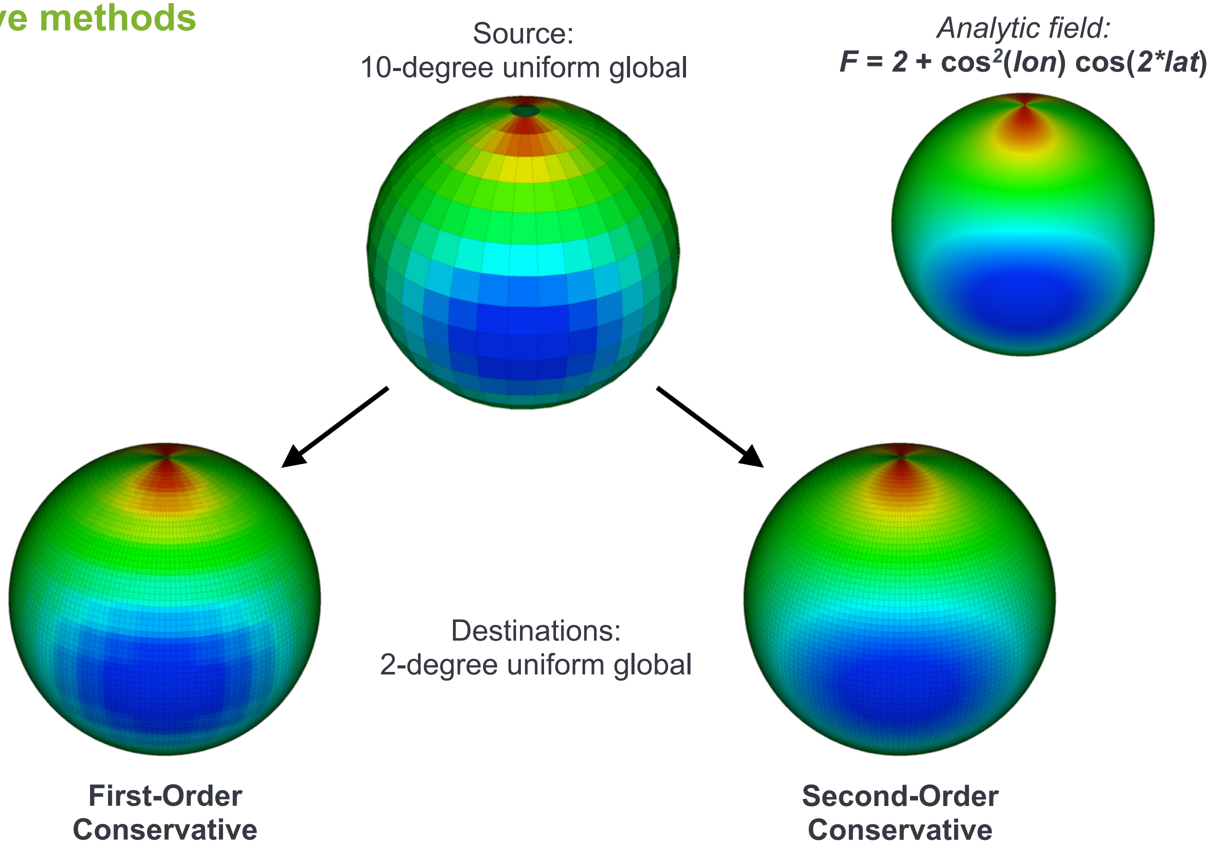
earthsystemcog.org/projects/esmf/regridding

High Performance Grid Remapping

ESMF supports multiple conservative methods

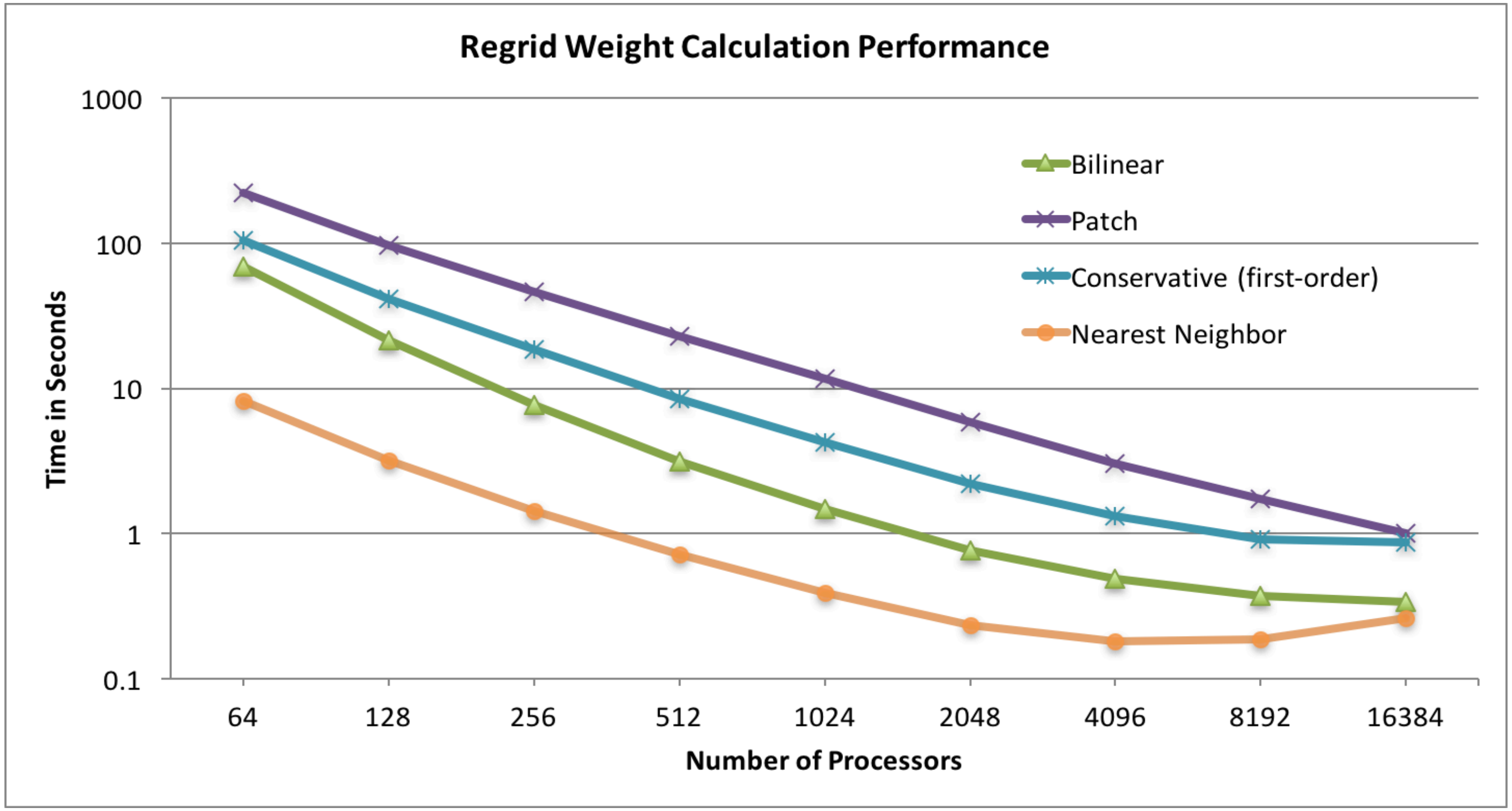
Second-order conservative:

- Destination cell value is the combination of values of intersecting source cells modified to take into account the **source cell gradient.**
- Requires a **wider stencil** and more computation, so more expensive in terms of memory and time than first-order
- Preserves integral of field across interpolation, but gives **smoother results than first-order** (especially when going from coarse to fine grids)



Interpolation Weight Generation Performance

ESMF regridding from regular cubed-sphere to lat/lon



Source:
regular cubed sphere grid
(~25 million cells)

Destination:
uniform latitude longitude grid
(~17 million cells)

Platform: IBM iDataPlex cluster
(Yellowstone at NCAR)

ESMF version:
ESMF_7_1_0_beta_snapshot_2
4

Courtesy: Bob Oehmke

Addressing modeling complexity

A look into the future

- Run sequence syntax that includes ensembles
- Ongoing feature additions to grid remapping
- Large grid performance and memory optimizations (e.g. 64 bit representation of sequence index)

Outline

- **Trends in Earth System Modeling lead to new HPC challenges**
  - Increasing model complexity
  - Heterogeneity in computing systems
  - Increasing use of community models and infrastructure
- **The role of ESMF/NUOPC**
  - ESMF 7.1.0r provides advanced building blocks for HPC applications
  - The NUOPC Layer standardizes architectural options for interoperable coupled model components
- **Addressing HPC challenges from the bottom up**
  - Focus on fast, flexible tools for next-generation prediction systems to address model complexity and leverage heterogeneous computing
  - Deliver development tools and training programs to support growing user community

Addressing HPC challenges

Getting closer to the hardware

- **Heterogeneity in computing systems:**
  - **Support heterogeneous computing**
    - CPUs + GPGPUs + MICs, threading and SIMD parallelism on CPU and device cores, multiple programming models
  - **Explore fine-grained computing with a goal of *smart* resource mapping**
    - Focus on longer-term goal of efficient mapping multi-component, coupled systems to heterogeneous resources through abstraction of those resources and the mapping process
    - Optimization includes components which may use different programming models and optimization strategies
    - Approach designed to preserve kernel-level optimizations and adapt to multiple computing platforms

Addressing HPC challenges

Getting closer to the hardware

- **Heterogeneity in computing systems:**
  - **Support heterogeneous computing**
  - CPUs + GPGPUs + MICs, threading and SIMD parallelism on CPU and device cores, multiple programming models
  - **Explore fine-grained computing with a goal of *smart* resource mapping**
  - Focus on longer-term goal of efficient mapping multi-component, coupled systems to heterogeneous resources through abstraction of those resources and the mapping process
  - Optimization includes components which may use different programming models and optimization strategies
  - Approach designed to preserve kernel-level optimizations and adapt to multiple computing platforms

Recognizing heterogenous platforms

Getting closer to the hardware

- ESMF and NUOPC components rely on the ESMF Virtual Machine (VM) to interact with and manage hardware resources
- The ESMF VM provides generic, abstract representation of diverse HPC resources
- The ESMF VM was extended to recognize heterogeneous hardware platforms, opening the way to further resources optimization
- The ESMF VM supports multiple software stacks: OpenCL, OpenACC, Intel MIC, OpenMP

Work supported by the ONR National Oceanographic Partnership Program (NOPP)

Addressing HPC challenges

Getting closer to the hardware

- **Heterogeneity in computing systems:**
  - **Support heterogeneous computing**
    - CPUs + GPGPUs + MICs, threading and SIMD parallelism on CPU and device cores, multiple programming models
  - **Explore fine-grained computing with a goal of *smart* resource mapping**
    - Focus on longer-term goal of efficient mapping multi-component, coupled systems to heterogeneous resources through abstraction of those resources and the mapping process
    - Optimization includes components which may use different programming models and optimization strategies
    - Approach designed to preserve kernel-level optimizations and adapt to multiple computing platforms

Accelerator awareness and resource negotiation

Getting closer to the hardware

ESMF NUOPC

Smart mapping sequence (funded under ONR NOPP)

- Capture information about the current performance of component run phases, and the requirements and capabilities of each component (e.g., can utilize accelerators
- Generate dependency graphs that capture relationships among components, building on an existing syntax that organizes component run phases into run sequences
- Analyze the dependency graphs to generate a set of potential scenarios that satisfy the dependencies
- Assess the potential scenarios using a minimization algorithm and make a selection
- Test the new configuration
- Iterate over this process to build up a knowledge base

Overall Approach to Performance Optimization

The ESMF philosophy

- **Do no harm**
  - Overhead of ESMF/NUOPC component interfaces is small (for ESMF, ~ μs)
  - Set of NUOPC prototypes demonstrates preservation of accelerator and other component-specific optimizations

- Key methods must **scale well**
  - key methods (e.g. sparse matrix multiplication) tested to ~16K processors by ESMF team, 40K+ processors by customers (*e.g.* NASA)

- Component interfaces and sequential/concurrent modes support **increasing task parallelism and optimized mappings to hardware**

Overall Approach to Performance Optimization

The ESMF philosophy

- Data communications between components can **preserve locality:**
  - Components with the same grid and decomposition:
    - **direct reference sharing**
    - **local memory-to-memory copy**
  - Components on disjoint processor sets:
    - **redistribution**
    - **parallel grid remapping**

- To learn more about ESMF performance, visit:

earthsystemcog.org/projects/esmf/performance/

Addressing Heterogeneous Computing

Future Steps

- Smart mapping implementations in model applications
- Increasingly sophisticated additions to smart mapping (e.g. optimized mediator distribution)
- Continued exploration of web, cloud, and mobile service connections, both downstream and as an interactive (steering) problem

Addressing HPC challenges

Getting closer to the users

- **Increasing use of community models and infrastructure:**
  - **Demand for user-friendly development environments**
    - The Cupid IDE for development, training, and optimization
  - **Need for tools**
    - Component Explorer
    - Compliance Checker
  - **Need for training programs and material**
    - Development of a cross-agency training program and training tools
    - Ongoing improvements to examples and documentation
    - Reduce barriers to adoption: growing number of NUOPC prototypes; non-intrusive approach

Addressing HPC challenges

Getting closer to the users

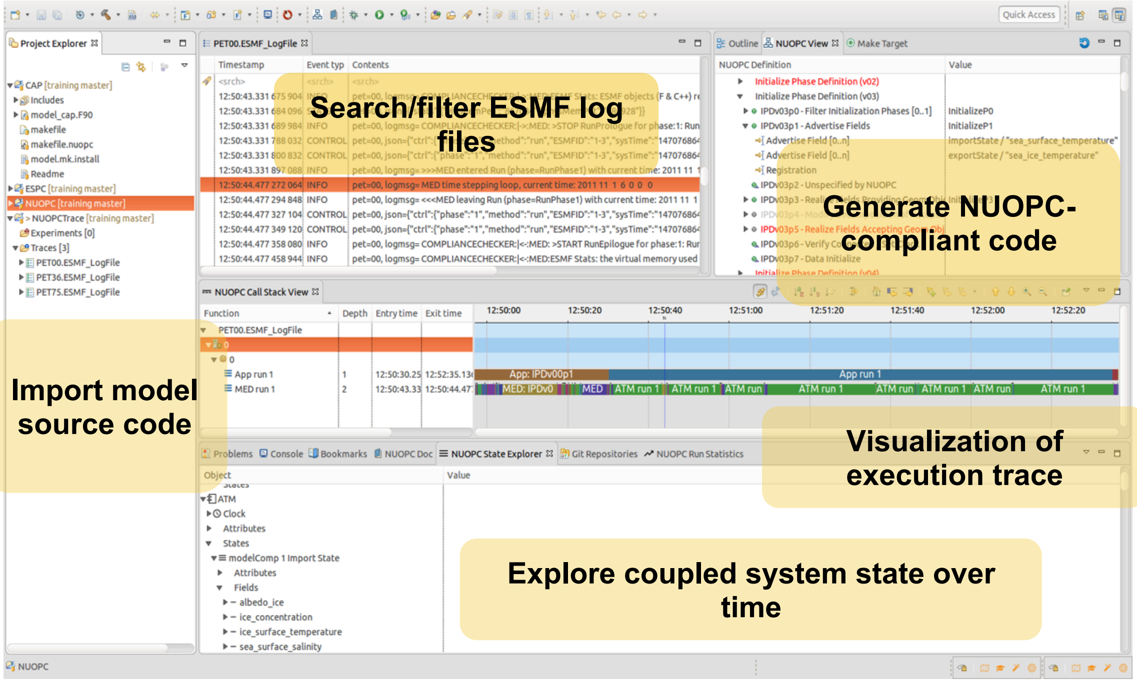
- **Increasing use of community models and infrastructure:**
  - **Demand for user-friendly development environments**
    - The Cupid IDE for development, training, and optimization
  - **Need for tools**
    - Component Explorer
    - Compliance Checker
  - **Need for training programs and material**
    - Development of a cross-agency training program and training tools
    - Ongoing improvements to examples and documentation
    - Reduce barriers to adoption: growing number of NUOPC prototypes; non-intrusive approach

Integrated Development Tools

IDE Tool for Development, Compliance Checking, and Optimization

Cupid is an Eclipse-based development environment (GUI) with special tooling for building and analyzing NUOPC applications.

Cupid supported under NASA CMAC 2011 and 2014



- automatically generate NUOPC infrastructure code (including “caps”)
- assist developers in understanding code structures in NUOPC components
- compliance check NUOPC “caps” for technical correctness
- streamline debugging by visualizing coupled model execution sequences and providing a dynamic view of the coupled system state over time

- **earthsystemcog.org/projects/cupid/**

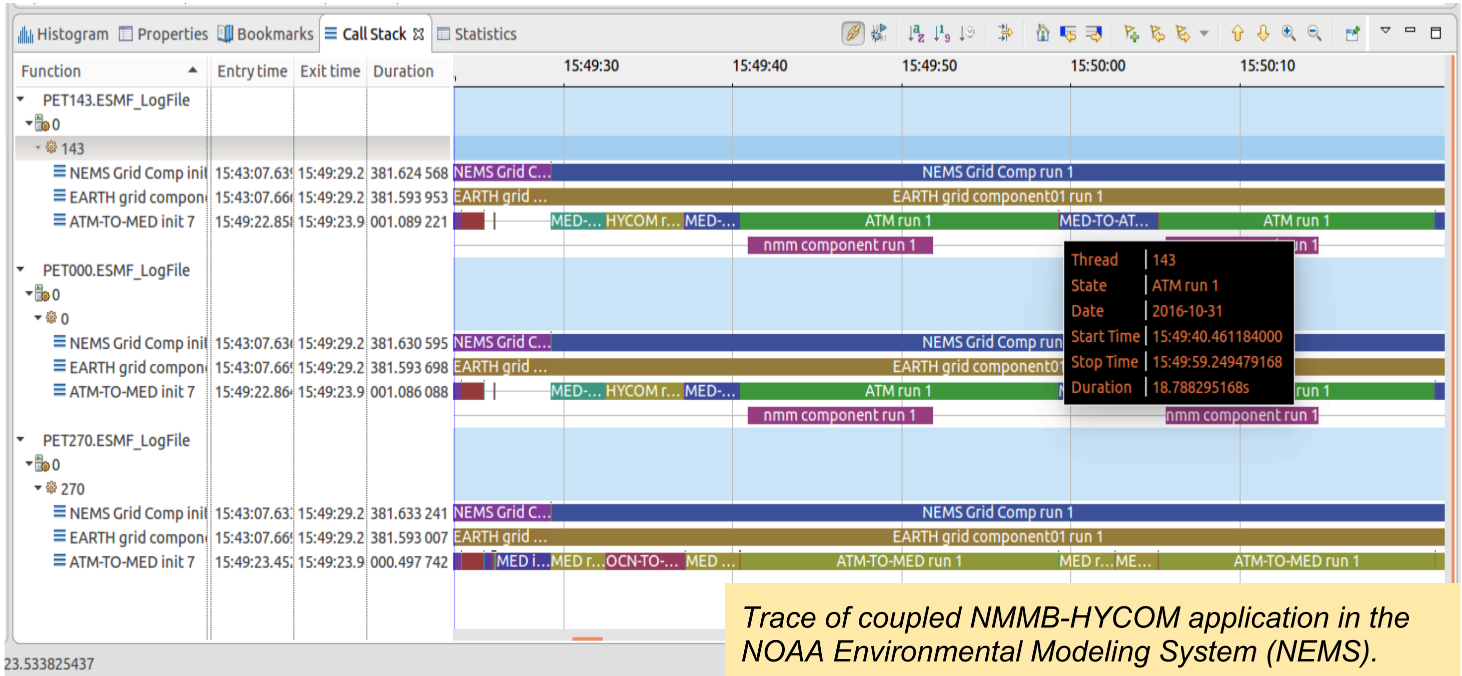
Integrated Development Tools

NUOPC Call Stack View

NUOPC Call Stack View

Comprehensive visualization of multiprocess component execution

- Shows entry/exits of NUOPC execution phases including timing information
- Stack process traces to see concurrency
- Advantage over existing tracing tools:
  - built into ESMF (no additional setup)
  - display reflects the organization of a NUOPC application



Trace of coupled NMMB-HYCOM application in the NOAA Environmental Modeling System (NEMS).

Addressing HPC challenges

Getting closer to the users

- **Increasing use of community models and infrastructure:**
  - **Demand for user-friendly development environments**
    - The Cupid IDE for development, training, and optimization
  - **Need for tools**
    - Component Explorer
    - Compliance Checker
  - **Need for training programs and material**
    - Development of a cross-agency training program and training tools
    - Ongoing improvements to examples and documentation
    - Reduce barriers to adoption: growing number of NUOPC prototypes; non-intrusive approach

Development and Training Tools

IDE Tool for Development, Compliance Checking, and Optimization

Compliance Checker

Activated by an environment variable, the Compliance Checker intercepts all NUOPC phases and writes out extensive compliance diagnostics to the ESMF log files. Any compliance issues found are flagged as warnings.

Component Explorer

A generic Driver that links to a single NUOPC component and outputs information to standard out such as registered phases and import/export fields.

Addressing HPC challenges

Getting closer to the users

- **Increasing use of community models and infrastructure:**
  - Demand for user-friendly development environments
    - The Cupid IDE for development, training, and optimization
  - Need for tools
    - Component Explorer
    - Compliance Checker
  - **Need for training programs and material**
    - Development of a cross-agency training program and training tools
    - Ongoing improvements to examples and documentation
    - Reduce barriers to adoption: growing number of NUOPC prototypes; non-intrusive approach

Training Program Overview

Topic	Duration	Frequency	Format
ESMF/NUOPC overview tutorial	1.5 hours	quarterly	webinar
ESMF/NUOPC overview tutorial	1 day	as requested, at least yearly at NRL, EMC, and NASA	on-site
Advanced topics: Grids and grid remapping	1.5 hours	as requested	webinar
Advanced topics: Building a NUOPC cap	1.5 hours	as requested	webinar
Site-specific tutorials	4 hours	as requested	webinar
Cupid tutorial	1 day	as requested, at least yearly at NRL, EMC, and NASA	on-site

Upcoming Webinars	
May 22 - 11 MT / 1 ET	Overview of ESMF and NUOPC
July 24 - 11 MT / 1 ET	Building coupled applications with NUOPC

Visit **earthsystemcog.org/projects/esmf/tutorials** to register and for up-to-date training schedules

Increasing use of community models and infrastructure

Future steps

- Establish community approaches to managing component code development and repositories (e.g. CICE consortium)
- Continued implementation of ESPS code usability criteria
- Expansion of ESMF training program and tool improvement

Summary

ESMF Strategies to Address HPC Challenges

- ESMF/NUOPC is rooted in the research and operational community, with a large user base and established collaborations
- Development strategies in response to HPC trends and challenges:
  - Support for many different model configurations (hierarchies, ensembles, etc.) and innovative coupling modes
  - Smart, increasingly automated resource mapping to enable efficient use of diverse hardware resources by coupled modeling systems
  - Continue to expand the ESMF training program and advance domain-specific, modern development tools
- Collaborate closely with the research and operational communities to anticipate, understand, and deliver new HPC requirements in a timely way

ESMF 7.1.0r Release

March 8, 2018

Changes since the last public release (6.3.0rp1) include:

- Implementation of higher (2nd) order conservative grid remapping method (Bob Oehmke/CIRES)
- New shortcuts for cubed sphere grid creation (Bob Oehmke/CIRES)
- Support for component hierarchies in the NUOPC Layer (Gerhard Theurich/NRL)
- Implementation of dynamic masking during the application of interpolation weights (Gerhard Theurich/NRL)
- Extrapolation of points that lie outside the source grid during grid remapping (Bob Oehmke/CIRES)
- Optimizations for performance and memory (Gerhard Theurich/NRL)
- Extension of the ESMF Virtual Machine to recognize heterogeneous resources such as accelerators - part of a broader effort to automate optimized mapping of components to hardware (Jayesh Krishna/ANL)

See full description of release:

earthsystemcog.org/projects/esmf/download_710r