NESII

Update on ESMF, Earth System Curator, and Earth System CoG

Cecelia DeLuca and the ESMF team
CCSM Software Engineering Working Group
June 2, 2011

Outline

- NESII Overview
- NESII Projects
  - Earth System Modeling Framework (ESMF)
  - National Unified Operational Prediction Capability (NUOPC)
  - Earth System Curator
  - Earth System Commodity Governance (CoG)
  - Curator Hydrology
  - NOAA Climate Projection Pilot (NCPP)
  - TeraGrid Environmental Science Gateway
  - Global Interoperability Program

The Vision

- Develop interoperable modeling components that can connect in multiple ways
  Improve predictions and support research
- Build advanced utilities that many models can use
  Enable research, promote efficiency
- Enable models to be self-describing
  Increase understanding and defensibility of outputs
- Create workflows that automate the modeling process from beginning to end
  Improve productivity
- Build workspaces that encourage collaborative, distributed development of models and data analysis
  Leverage distributed expertise

The Team

Person	Role	Location
Cecelia DeLuca	Technical Manager	ESRL
Sylvia Murphy	Operations & Project Manager	ESRL
Silverio Vasquez	Test and Integration Lead	ESRL
Gerhard Theurich	Senior Developer – architecture	CA
Bob Oehmke	Senior Developer - grids	ESRL
Luca Cinquini	Senior Developer - database	ESRL/CA
Peggy Li	Developer - performance	CA
Allyn Treshansky	Developer - metadata	CA
Walter Spector	Developer – porting	CA
Ryan O'Kuinghttons	Developer - everything	ESRL
Fei Liu	Developer - applications	New Jersey
Kathy Saint	Developer - web	Florida
Earl Schwab	Developer - utilities	CO

NESII visitors:
Tony Wong, intern
and
Jay Hnilo, NCDC

Earth System Modeling Framework

Started: 2002
Collaborators: Co-developed and used by NASA (GEOS-5 climate model), NOAA (NCEP weather models), Navy (global and regional models), Community Earth System Model, others
Sponsors: NASA MAP, NOAA NWS and CPO, NSF SEIII, DoD HPCMP

- ESMF increases code reuse and interoperability in climate, weather, coastal and other Earth system models
- ESMF is based on the idea of components, sections of code that are wrapped in standard calling interfaces

ESMF as an Information Layer

NESII

Applications of information layer

- *Parallel generation and application of interpolation weights*
- *Run-time compliance checking of metadata and time behavior*
- *Fast parallel I/O (PIO from NCAR/DOE)*
- *Redistribution and other parallel communications*
- *Automated documentation of models and simulations (new)*
- *Ability to run components in workflows and as web services (new)*

Structured model information stored in ESMF wrappers

ESMF data structures

Standard metadata

Attributes: CF conventions, ISO standards, METAFOR Common Information Model

Standard data structures

Component | *Field* | *Grid* | *Clock*

User data is referenced or copied into ESMF structures

Native model data structures

modules | *fields* | *grids* | *timekeeping*

ESMF Regridding

Fully parallel, portable

Working towards CF compliance for structured and unstructured grid formats

Two ways to invoke regridding:

- ESMF Offline:
  - Application which can be automatically built as part of ESMF
  - Application generates a netCDF weight file from two netCDF grid files
  - Supports SCRIP format grid files, and a custom ESMF unstructured format

```
mpirun –np 32 ESMF_RegridWeightGen –s src_grid.nc  –d dst_grid.nc –m bilinear –w weights.nc
```

- Integrated:
  - ESMF library subroutine calls which do interpolation during model run
  - Can get weights or feed directly into ESMF parallel sparse matrix multiply
  - Can be used without ESMF components

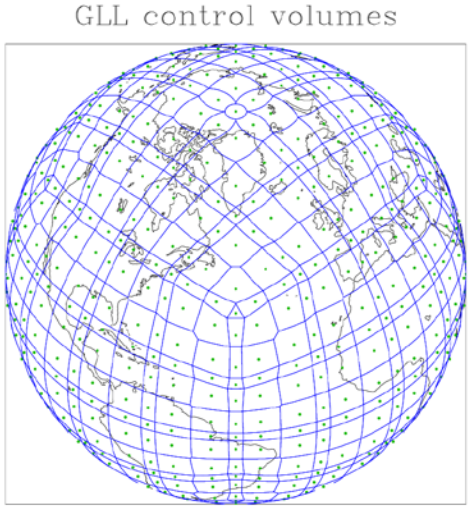
```
call ESMF_FieldRegridStore(srcField=src, dstField=dst,
                           regridMethod=ESMF_REGRID_METHOD_BILINEAR, routehandle=rh)

call ESMF_FieldRegrid(srcField=src, dstField=dst, routehandle=rh)
```

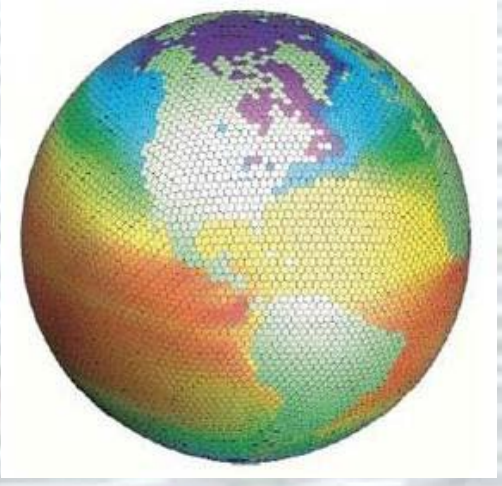
ESMF Offline Supported Grids

- Grids with spherical (lon, lat) coordinates – any pair of:
  - Global 2D logically rectangular grids
  - Regional 2D logically rectangular grids
  - 2D unstructured meshes composed of polygons with any number of sides: triangles, quadrilaterals, pentagons, hexagons,…
  - Multi-patch grids (e.g. cubed spheres) currently supported via unstructured format

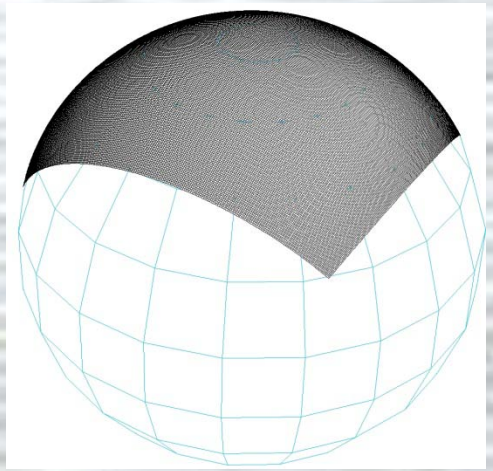
RESULT:
"use of the parallel ESMF offline regridding capability has reduced the time it takes to create CLM surface datasets from hours to minutes" - Mariana Vertenstein, NCAR



HOMME Cubed Sphere Grid with Pentagons
Courtesy Mark Taylor of Sandia



FIM Unstructured Grid
Courtesy ESRL GSD

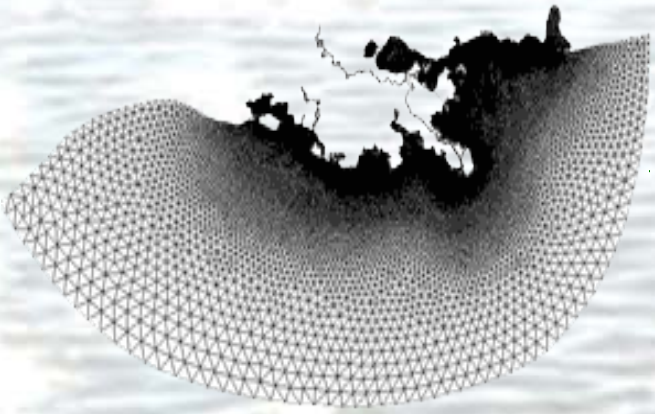


Regional Grid

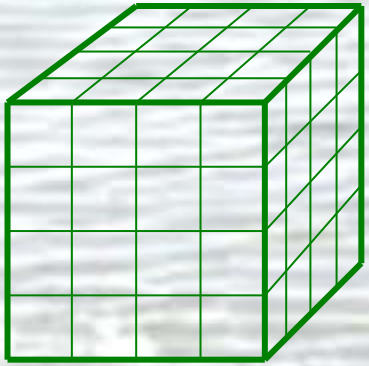
Integrated Supported Grids

In addition, integrated regridding supports Cartesian (x,y) coordinates:

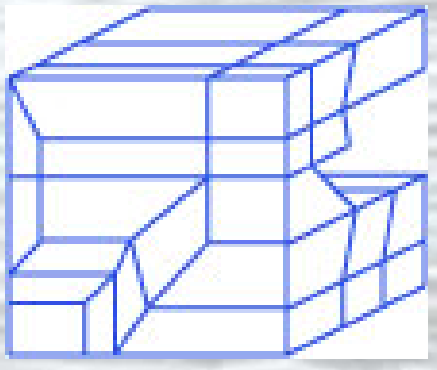
- Regridding between any pair of:
  - 2D meshes composed of triangles and quadrilaterals
  - 2D logically rectangular Grids composed of a single patch
- Bilinear regridding between any pair of:
  - 3D meshes composed of hexahedrons
  - 3D logically rectangular Grids composed of a single patch



2D Unstructured Mesh
From www.ngdc.noaa.gov



3D Grid



3D Unstructured Mesh

ESMF Offline Regrid Use

- Enable CLM land model to run on any ustructured grid
  - Grids: Land lat/lon grid to unstructured grid including HOMME cubed sphere
  - ESMF parallel bilinear mapping from lat/lon to HOMME cubed sphere allowed investigation of high resolution land model to move forward for CESM
- Reduce noise in interpolated wind stress values
  - Grids: CAM atmosphere lat/lon to POP ocean displaced pole lat/lon
  - ESMF patch interpolation reduced imprint of coarser resolution atmosphere grid on ocean for interpolated wind stress values.
- Enable interpolation of POP ocean and HOMME
  - Grids: HOMME cubed sphere atmosphere to POP ocean grid
  - ESMF conservative regridding enabled integration of a high resolution dynamical core into CAM, reduced distortion near the pole
- Improve pole treatment for geodesic to lat/lon remapping
  - Grids: MPAS unstructured grid to POP ocean grid
  - ESMF conservative interpolation solved problems with negative weights at the pole
- *Also: geodesic to geodesic and other interpolation for CSU is working, NCL*

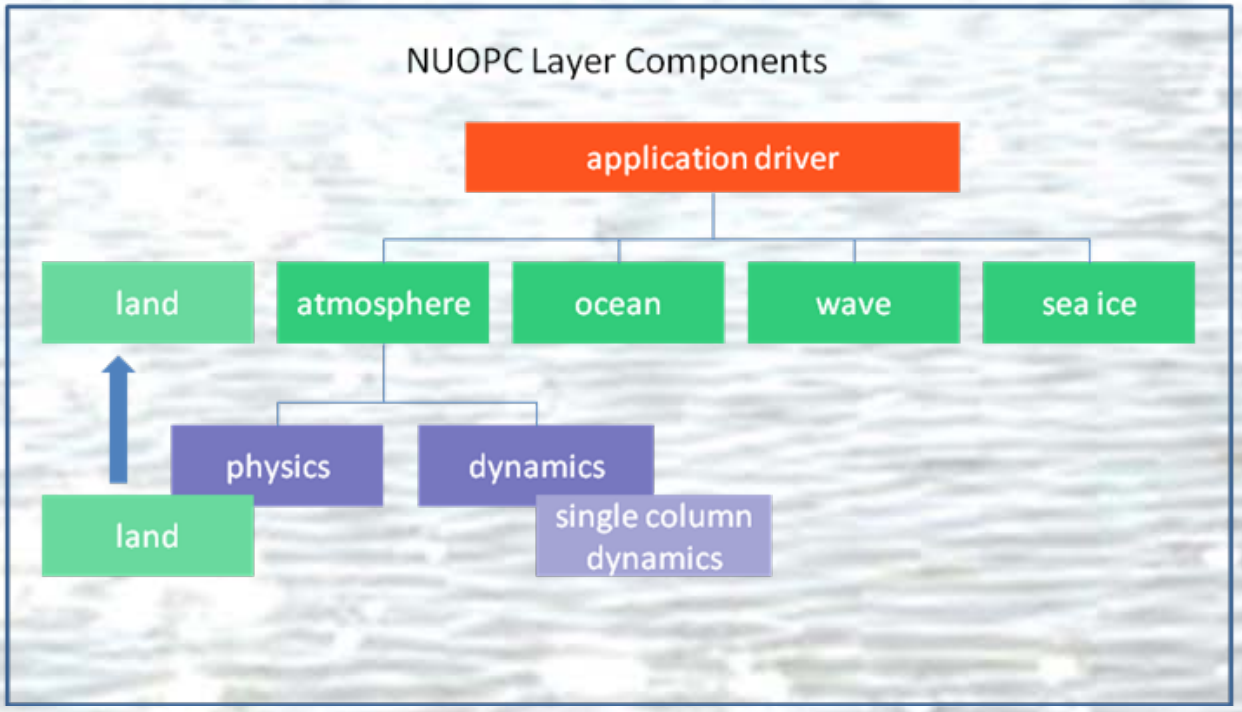
National Unified Operational Prediction Capability (NUOPC)

Started: 2010
Collaborators: Tri-agency (NOAA, Navy, Air Force) consortium of operational weather prediction centers, with participation from NOAA GFDL and NASA modelers
Sponsors: NOAA NWS and Navy

- ESMF allows for many levels of components, types of components, and types of connections
- In order to achieve greater interoperability, usage and content conventions and component templates are needed
- This collaboration is building a "NUOPC Layer" that constrains how ESMF is used, and introduces metadata and other content standards
- The initial pilot project (to be delivered June 2011) focuses on atmosphere-ocean coupling in NCEP NEMS and Navy NOGAPS and COAMPS codes

A Common Model Architecture

NUOPC partners have agreed on a subset of components whose interactions will be standardized



NUOPC Layer prototype

- Establish an architecture in which major components are siblings. The initial design supports explicit coupling and concurrent or sequential execution of components.
- Allow inheritance from generic component templates.
- Couple components with pre-fabricated connectors.
- Standardize the number and function of phases of initialize, creating a standard setup pattern.
- Constrain the data that may be sent between components with standardized field data structures and a CF-based field dictionary.
- Implement a compliance checker to provide feedback during component development.
- Use compatibility checking to determine if required import fields for a component were supplied. Other run-time reporting alerts users to any issues with compliance.

NESII

Interoperability impact

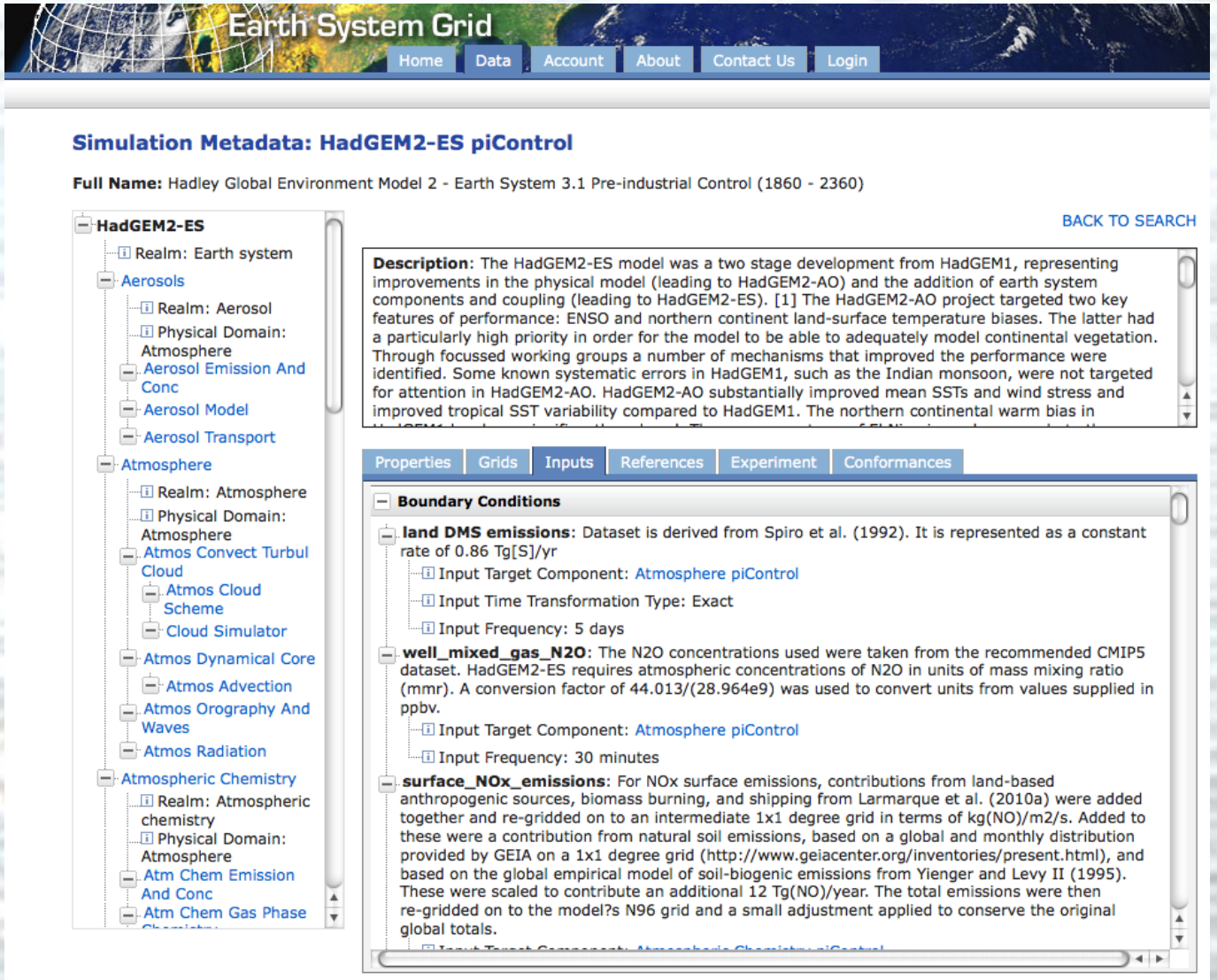
- Component templates and generic connectors make it easier for modelers to create compliant systems since they can reuse existing code and patterns
- Constraining and clearly defining initialize phases ensures that components fit into standard drivers.
- Constraining Field data structures and metadata ensures that the data communicated between components is understood
- The ESMF layer ensures that components fit together into an executable application. Detailed reporting provides the model developer with the insight necessary to debug and rectify code quickly.

Earth System Curator

Started: 2005
Collaborators: METAFOR , NCAR, DOE PCMDI, Earth System Grid Federation, NOAA GFDL, Georgia Institute of Technology
Sponsors: NSF SEIII, NASA MAP, NOAA GIP

- Intergovernmental Panel on Climate Change (IPCC) assessments rely on data generated by Coupled Model Intercomparison Projects (CMIPS), where different scenarios are tested across many models
- For the last IPCC assessment, there was little metadata available about the runs performed
- The Curator project collaborated on a comprehensive metadata schema for climate models, and implemented a metadata display in the Earth System Grid data distribution portal

ESG Metadata Display



This screen shot shows a real CMIP5 run as it appears in an ESGF portal

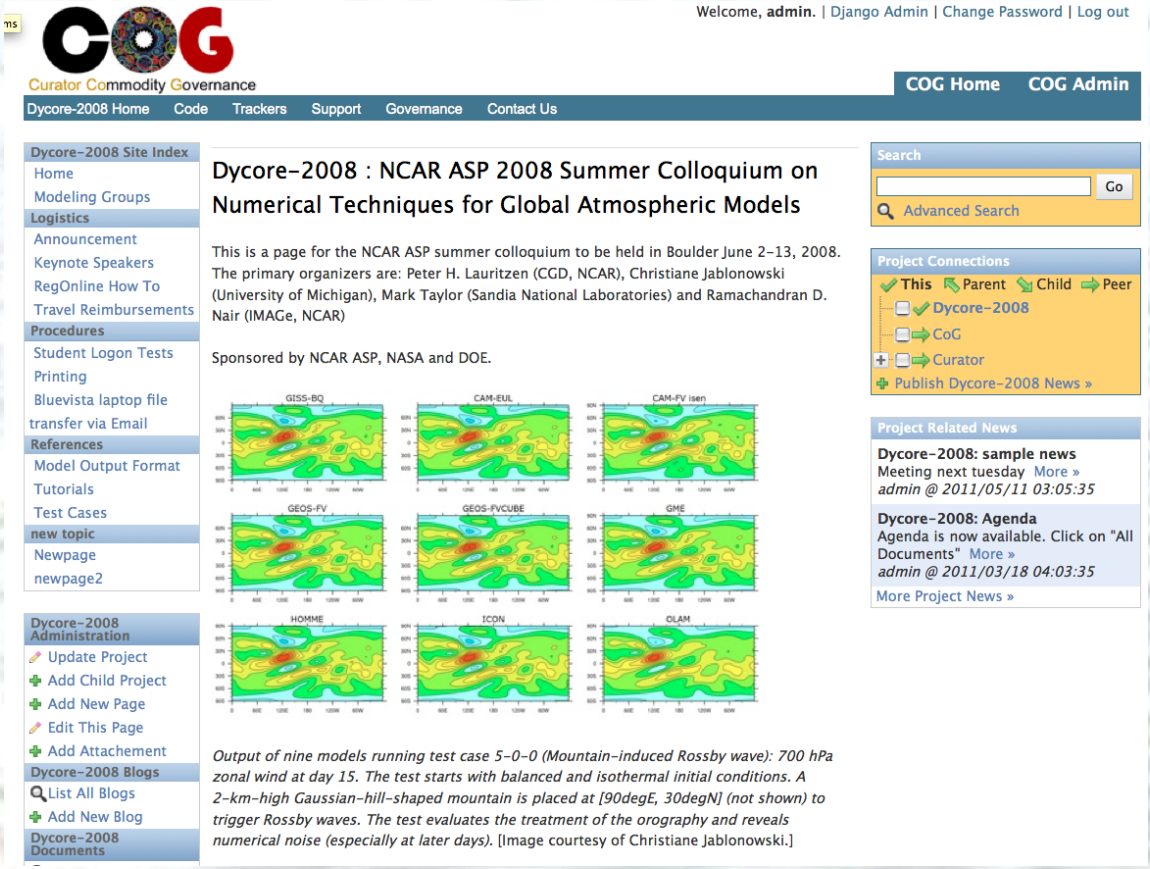
RESULT: *MUCH more information about climate models used in assessments, in browsable, searchable form*

Earth System CoG

Started: 2009
Collaborators: NCAR, Eart System Grid Federation, University of Michigan, CU Community Surface Dynamics Modeling System
Sponsors: NSF CDI

- Project hosting and indexing with connections to data and analysis services through the Earth System Grid Federation (ESGF)
  - Workspaces for collaborative model building, evaluation and analysis
  - Templates for project layout so information is easy to find
  - Peer or parent/child connections between projects
  - Multiple modes of communications between projects (e.g. broadcast news to all children)
- Pilot project is 2012 workshop on comparison of atmospheric dynamical cores (previously supported 2008 workshop)

2008 Dynamical Core Colloquium on CoG



CoG prototype includes data search, wikis, and communications

NEXT: *Users will be able to save ESGF datasets, LAS analysis and visualizations and other artifacts back to the workspace*

Curator Hydrology

Started: 2009
Collaborators: University of South Carolina, University of Michigan
Sponsors: NOAA GIP

- A new perspective on climate impacts modeling
- Instead of what do we “put in” the climate model …
- How do we create a linked network of models that multiple communities can use?

Design Goals

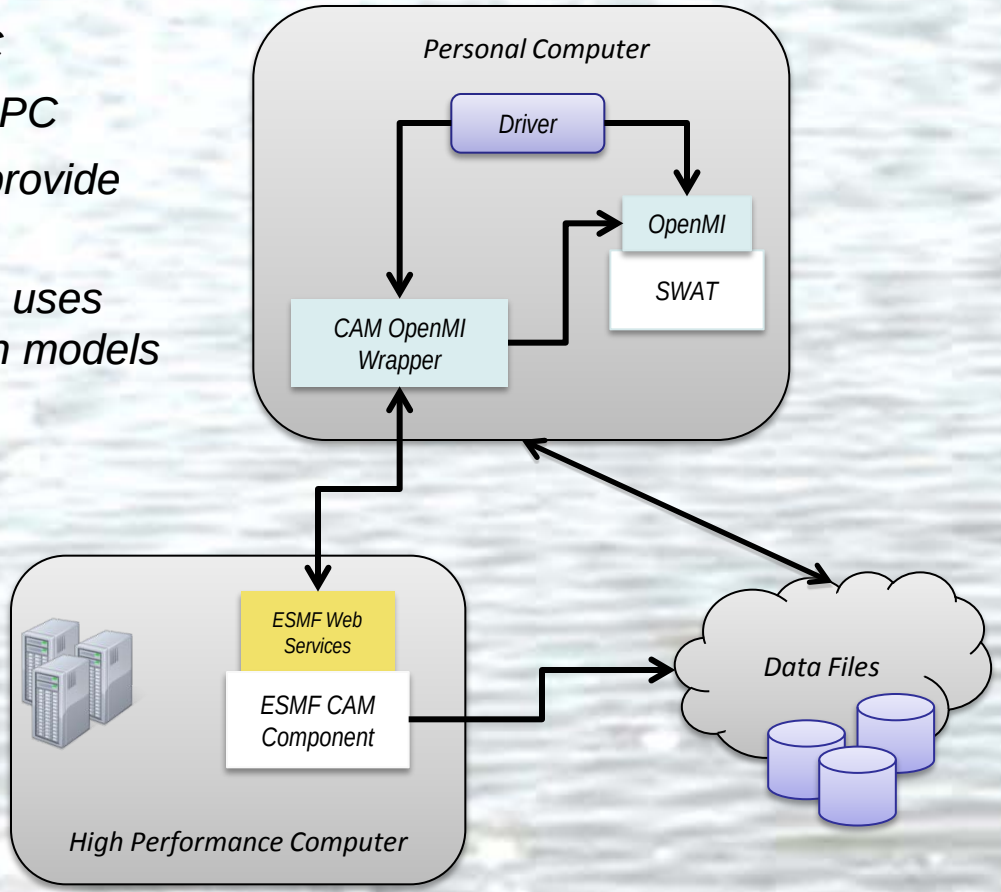
Goals	Strategies
Modeling systems can be reconfigured easily for including different models or solving different problems	Leverage model interface and data standards
Modeling systems are highly accessible and can be integrated into workflows that include analysis, visualization, and other processing of outputs	Service oriented architecture
Communities formed around local/regional modeling and climate are able to utilize the social and technical structures that have evolved in their domains	Models retain their native codes, computing platforms, and data formats as much as possible

Climate-Hydro Coupling

- Hydrological impact studies can be improved when forced with data from climate models [Zeng *et al.*, 2003; Yong *et al.*, 2009]
- Ideally the coupling would be two-way
- A technology gap exists:
  - Many hydrological models run on personal computers
  - Most climate models run on high performance supercomputers
- Existing frameworks: ESMF (climate/weather) and OpenMI (hydrology) can connect these types of models
  - ESMF and OpenMI components can be operated as web services that can be used to communicate across a distributed network
  - Both ESMF and OpenMI are widely used

Prototype Climate-Hydro System

- *SWAT (hydrology model) runs on PC*
- *CAM (atmospheric model) runs on HPC*
- *Wrappers for both SWAT and CAM provide OpenMI interface to each model*
- *Driver (OpenMI Configuration Editor) uses OpenMI interface to timestep through models via wrappers*
- *Access to CAM across the network provided by ESMF Web Services*
- *CAM output data written to NetCDF files and streamed to CAM wrapper via ESMF Web Services*
- *Using prototype to explore feasibility of 2-way coupling*



From Saint, iEMSs 2010

Target Coupled System

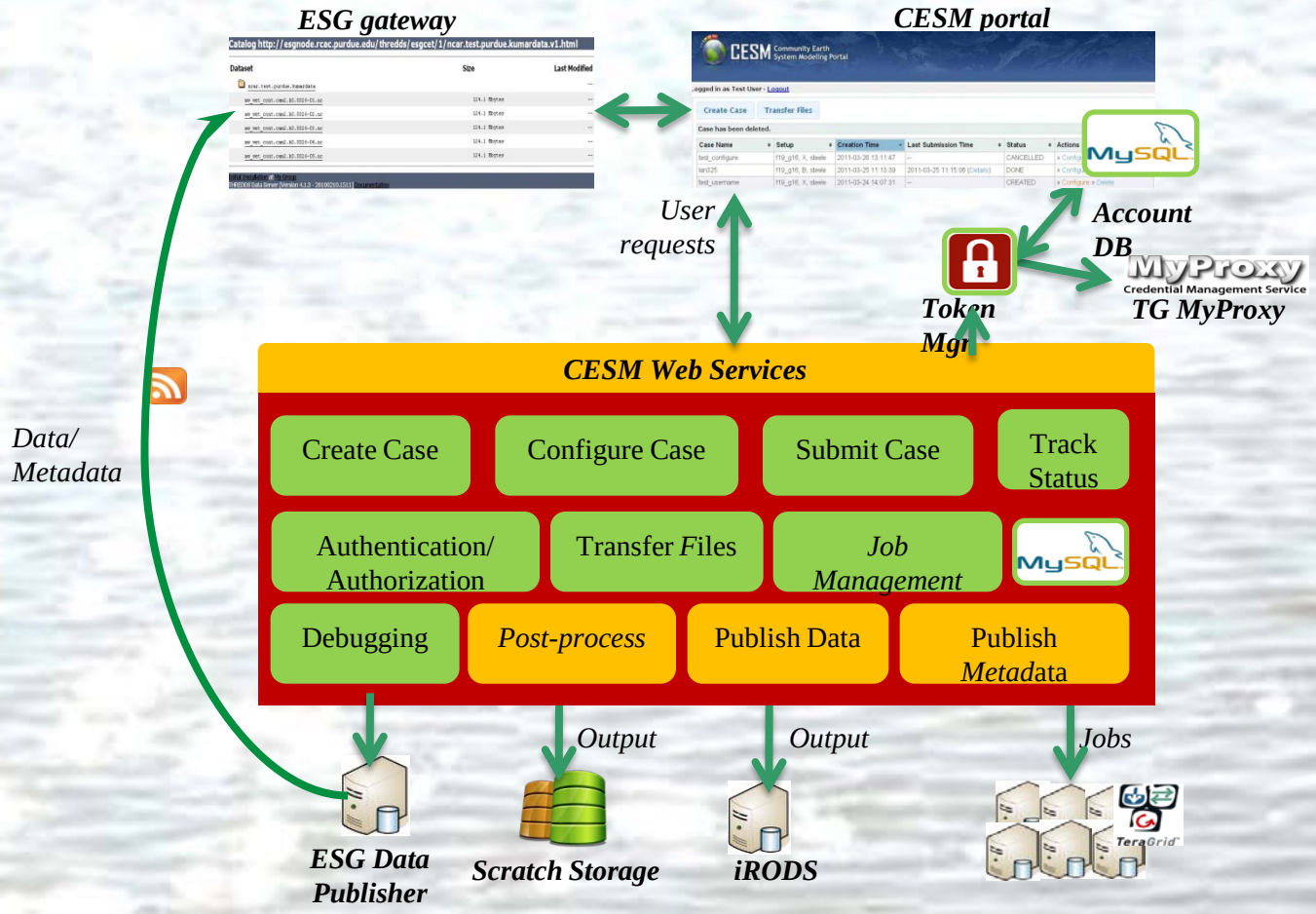
- Target system informed by exploration of parameter space for different strategies (estimated SWAT and CAM run times and transfer times)
- SWAT covering southeast U.S. coupled to CAM/CLM – purple region
- Restricting finest SWAT resolution to watersheds of interest (Neuse and Savannah) makes calibration somewhat easier
- SWAT forced by CAM fields (precip, temperature, wind speed, etc.); ET from SWAT nudges values in CAM

TeraGrid Environmental Science Gateway

Started: 2008
Collaborators: NCAR CISL and CESM, Purdue University
Sponsors: NSF TeraGrid

- Creates an end-to-end, self-documenting workflow for running the Community Earth System Model (CESM)
  - GUI configuration and submission of runs through the Purdue CESM portal
  - ESMF is used within CESM to organize and output extensive model metadata
  - Data and metadata is archived back to an Earth System Grid Federation Gateway, where it can be searched and browsed
  - Currently have a working prototype

Gateway Architecture



NOAA Global Interoperability Program

Started: 2009
Collaborators: NOAA GFDL, PMEL, GSD, and NCDC, Unidata, NCAR, CSU, University of Michigan
Sponsors: NOAA CPO

- GIP builds software infrastructure that
  - can be used in the weather, water, and climate disciplines, and for training modelers
  - integrates and automates workflows
- NESII lead DeLuca coordinates the project

Building Along Workflows

	Climate Simulations	Application of Climate Information	Weather and Water Forecasting	Training Modelers
Model Utilities and Coupling	• Standardized analysis workflows for climate models • Metadata display for CMIP5 • ESMF in CESM	• The NOAA Climate Projection Pilot	• A common model architecture for operational weather centers (NUOPC) • NOAA Environmental Modeling System (NEMS) • Geodesic grids in NEMS	• Summer School in Atmospheric Modeling • The Art of Climate Modeling course
Metadata Standards				
Data Services and Workflows				

Building Across Disciplines

	Climate Simulations	Application of Climate Information	Weather and Water Forecasting	Training Modelers
Model Utilities and Coupling	• ESMF core support • Hydrological-climate coupling with ESMF and OpenMI modeling frameworks			
Metadata Standards	• Gridspec integration into the Unidata LibCF library			
Data Services and Workflows	• Merger of Ferret and CDAT analysis services			

RESULT:
Better coordination of infrastructure development across disparate groups

The Vision

- Develop interoperable modeling components that can connect in multiple ways
  Improve predictions and support research
- Build advanced utilities that many models can use
  Enable research, promote efficiency
- Enable models to be self-describing
  Increase understanding and defensibility of outputs
- Create workflows that automate the modeling process from beginning to end
  Improve productivity
- Build workspaces that encourage collaborative, distributed development of models and data analysis
  Leverage distributed expertise

Questions?

- For more information, links and references, see our newish group page: http://esrl.noaa.gov/nesii/