



Towards a new NorESM

Mariana Vertenstein

Nov. 28, 2023

Outline

- Coupling between model components:
CMEPS
- Hierarchical modeling with data components:
CDEPS
- Where are we in NorESM3 development?

What is ESMF/NUOPC?

The **Earth System Model Framework (ESMF)** is a software for building numerical weather prediction and Earth system modeling applications.

- ❖ **high-performance capabilities such as run-time grid remapping and optimized parallel communication**

The **National Unified Operational Prediction Capability (NUOPC) Layer** is part of ESMF and simplifies the interoperability of ESMF components:

- ❖ Introduces four generic, extensible components with built-in capabilities
- ❖ Driver, Connector, Model, Mediator

What earth system models are now using ESMF/NUOPC?

USA:

- CESM
- NOAA Unified Forecast System (UFS)
- Earthworks (3.75 km global MPAS grid)

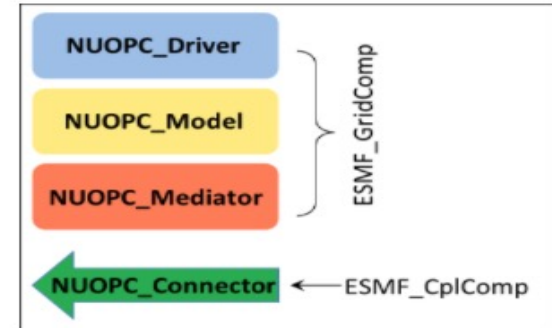
Outside USA:

- **NorESM**
- **CMCC-CM3 (Italy)**
- COSIMA (for upcoming ACCESS-OM3)
(Australia)

ESMF/NUOPC Provides New Coupling Capabilities

NUOPC Driver:

- **Data driven run sequence** - can easily see lags in model evolution

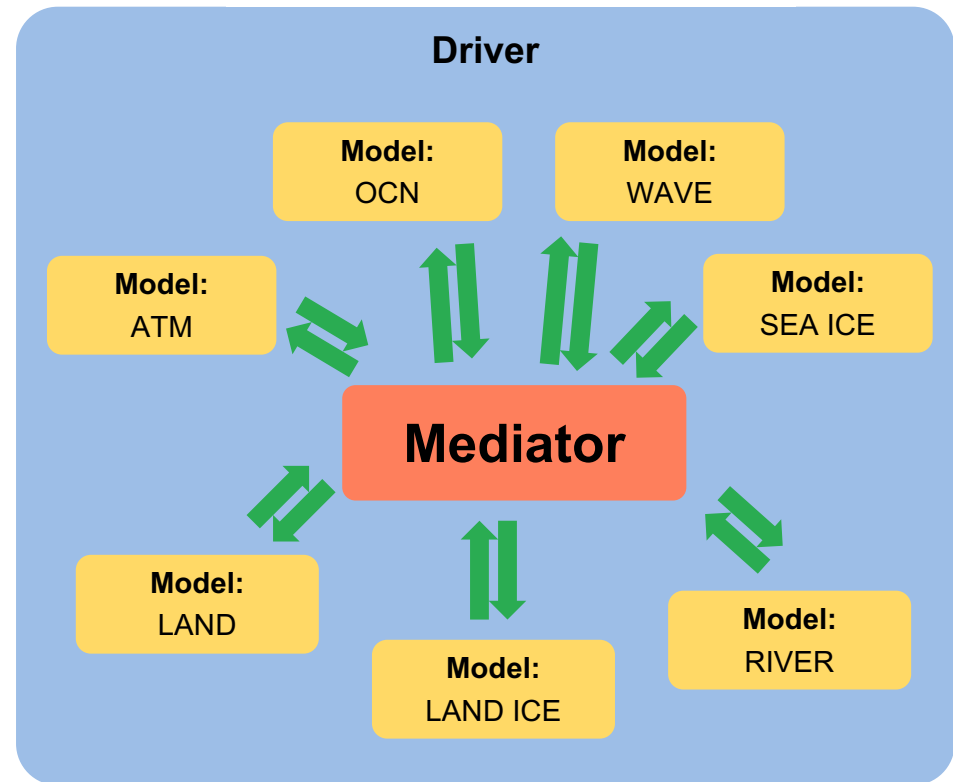


NUOPC Mediator:

- **Parallel online generation of remapping weights** - no more mapping files!

NUOPC Connectors:

- **Automatic** transfer of component data to and from the mediator



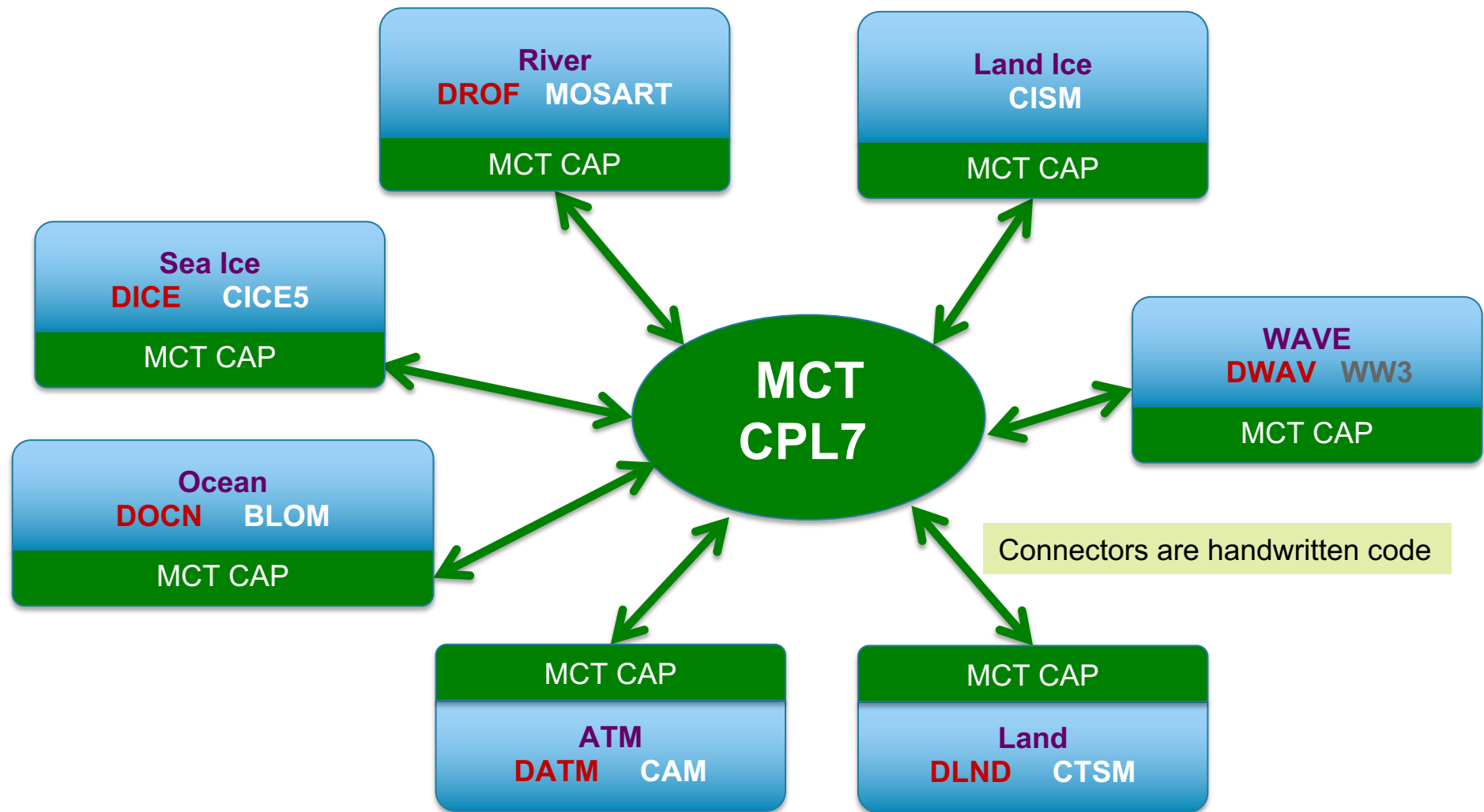
Coupling between components

Community Mediator for Earth Predictive Systems

CMEPS

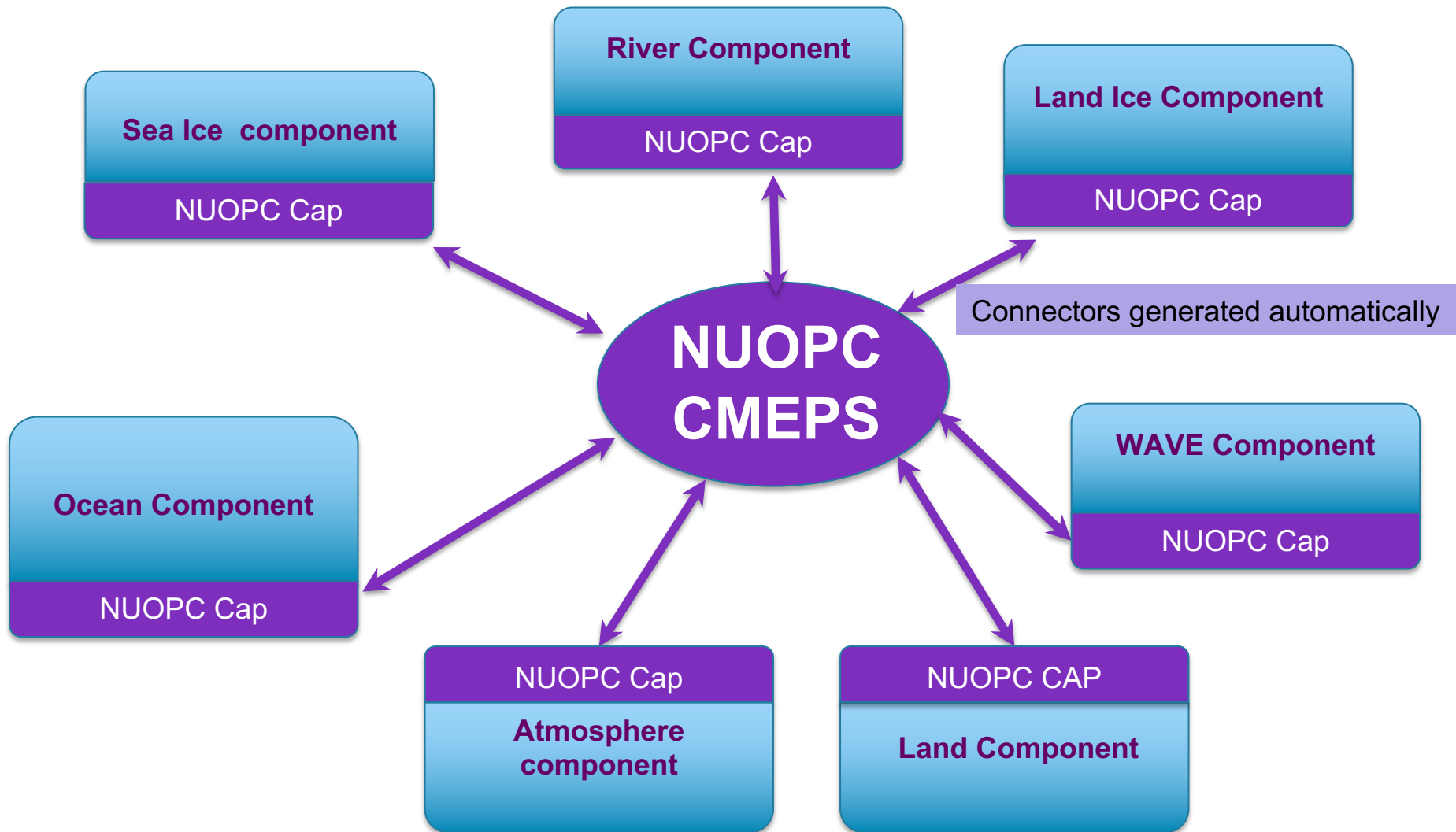
NorESM1/NorESM2 Coupling Framework

No clear separation between driver and “hub”



Temporal evolution of system comprises more than 4K lines of custom code

NorESM3 architecture

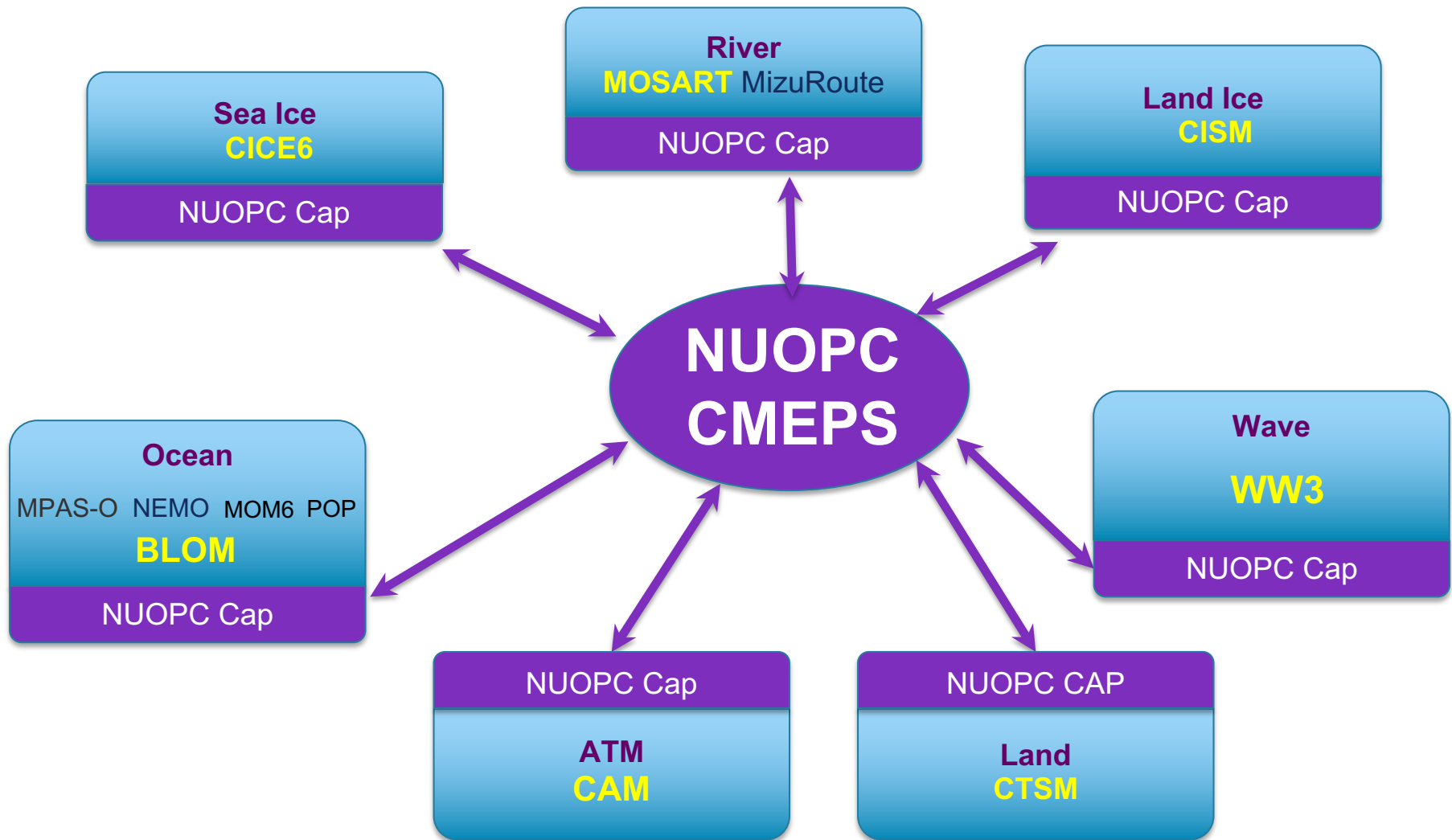


Temporal evolution contained in a small ascii file that can be modified

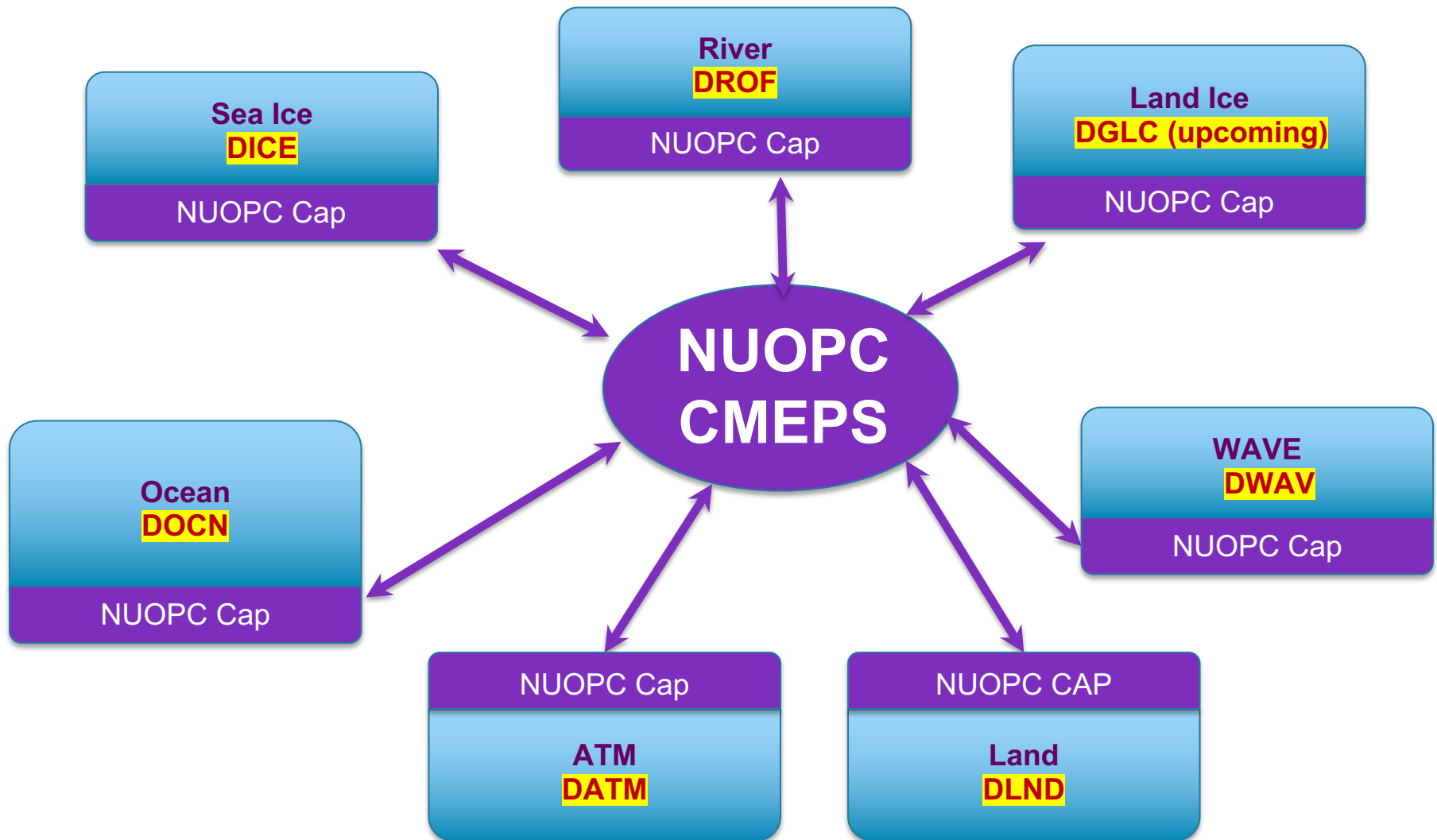
Prognostic CMEPS compliant components

NUOPC is now the coupling infrastructure in NorESM

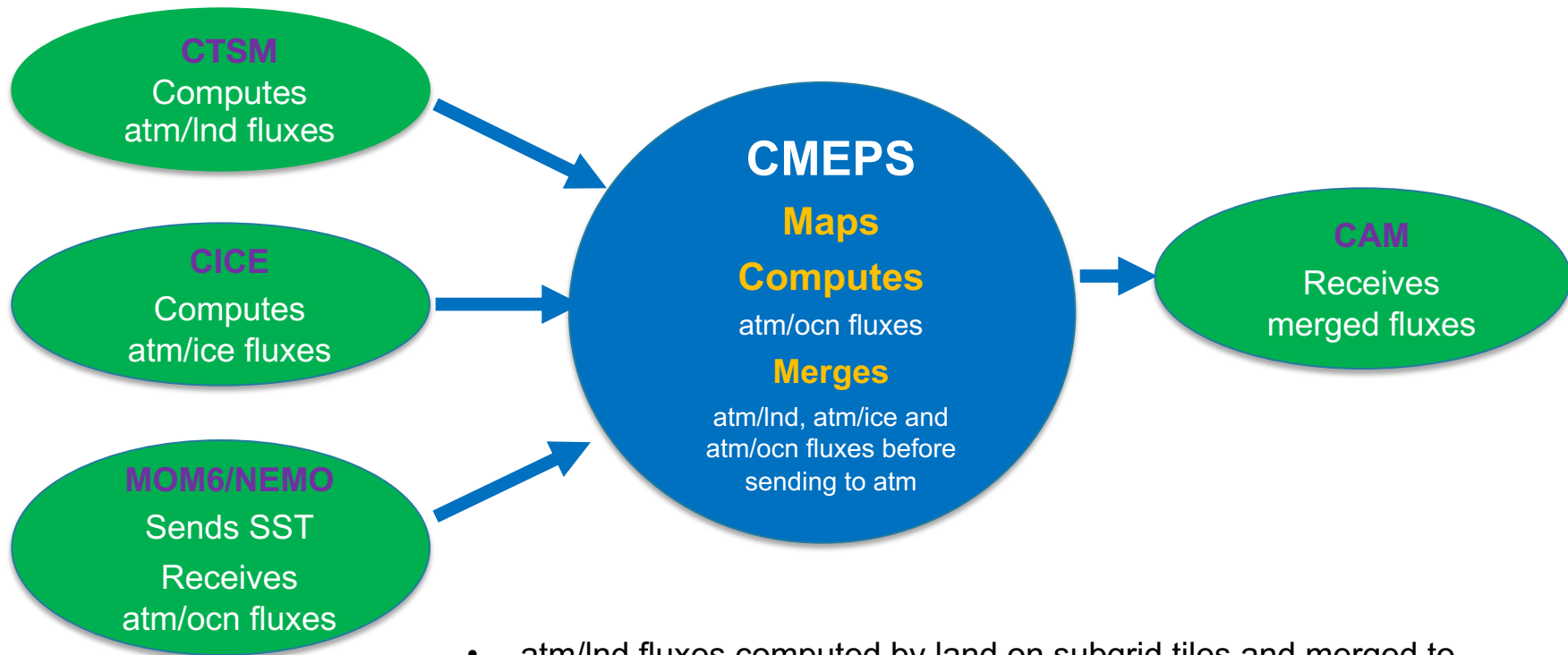
Extensive validation done in multiple configurations of CESM, including multiple century long fully coupled simulations!



CMEPS compliant data components (CDEPS)



atm/Ind, atm/ice and atm/ocn fluxes



- atm/Ind fluxes computed by land on subgrid tiles and merged to gridcell by land
- atm/ice/fluxes computed by ice on subgrid ice categories and merged to gridcell by ice.
- atm/ocn fluxes computed on the atm, ocn OR exchanged grid in cmeps (mapping and time averaging of atm/ocn fluxes done in cmeps)

Benefits of CMEPS – Introducing new Grids

- **Easier to introduce new grids (1)** – no longer need offline mapping files
 - **Before:** all inter-component mapping files were created offline
 - **25 mapping files** needed for a fully coupled pre-industrial control
 - **Now:** all non-custom mapping files!!! are generated at run time.
 - **Only 4 mapping files are** needed
- **Easier to introduce new grids (2)** – no longer need to generate offline land fraction files
 - Land and ocean fractions on atm/land grid is determined by mapping ocean mask conservatively to land grid
 - **Before:** each new component grid required generating new offline fraction files and updating CIME configuration files.
 - **Now:** land and ocean fractions are generated at runtime during model initializations!.

Benefits of ESMF – Introducing new Grids (cont)

- **Easier to introduce new grids (3)** – land surface dataset generation is now parallel!
 - **Before:** needed to create 17 offline mapping files and use these as input to a surface dataset generation code that ran on one processor. **Took over 2 days** to generate a surface dataset at 7.5 km grid.
 - **Now:** all mapping is done at run time and all I/O is parallel. Now **takes 10 minutes** to generate a surface dataset for a 7.5 km grid.
 - **Now:** Enable mapping of 30 second (724M points) soil texture dataset.
- Creation of new land surface dataset capabilities has used new ESMF features like dynamic masking for calculating the standard deviation of surface elevation

Benefits of CMEPS – New Land-Ice Capabilities

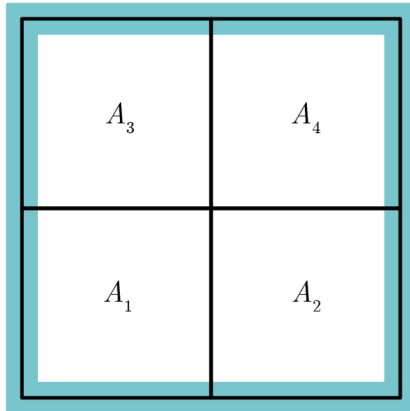
- Running both Antarctica and Greenland in one simulation (1)
 - **Before:** proposed approach was to create a unified global grid. New global grids would have to be created for every combination. Results in combinatorial explosion!
 - **Now:** create a 'nested state' where each ice sheet in CISM couples to a corresponding ice sheet in the mediator. Very extensible and user friendly approach – new glaciers can be added easily.
 - **An arbitrary number of ice sheets can be coupled at run time.**
- Enabling Antarctic ocean <-> land-ice coupling (2)
 - Requires regridding ocn->cism fields at multiple levels. Each level has different mask due to different bathymetry.
 - **Before:** a different mapping file for ocn->cism mapping was required for each ocean level. Each ocean level field was passed separately.
 - **Now:** can do regridding leveraging ESMF **dynamic masking functionality** in the mediator. Only one field with multiple levels is passed.

Benefits CMEPS – new exchange grid capability for calculating atm/ocn fluxes

- Exchange grid is the union of atm and ocn grids.
- Traditionally in CESM atm/ocn fluxes were computed on the ocean grid – problems arise if the atm grid is much higher resolution

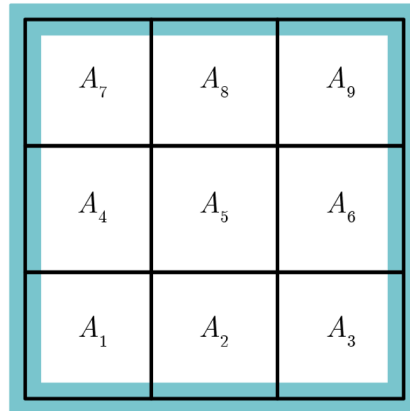
(a.)

$A_k^{(pg2)}$



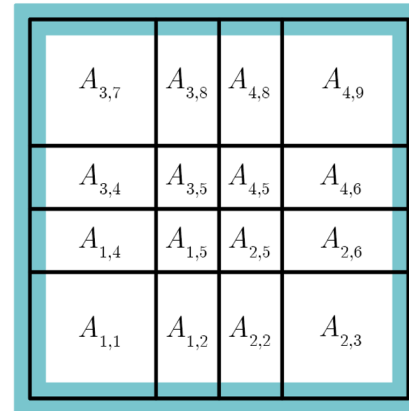
(b.)

$A_l^{(pg3)}$



(c.)

A_{kl}



CMEPS now can compute atm/ocn flux calculation on either the ocean grid, the atm grid OR the exchange grid

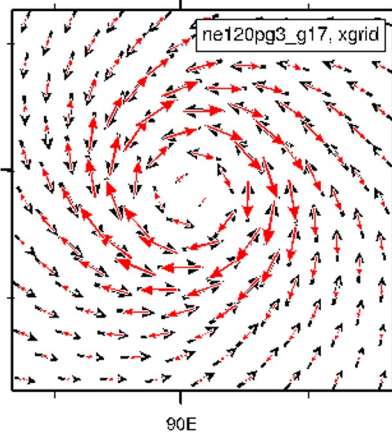
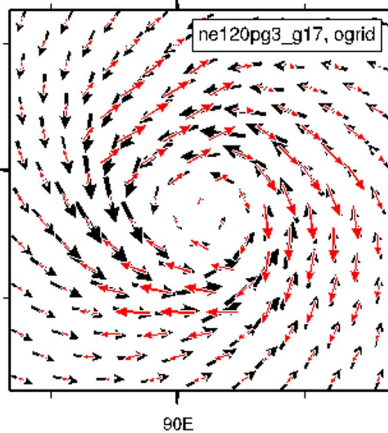
Idealized Tropical Cyclone Experiments (Adam Herrington)

“ne120pg3_g17” grid alias

xgrid maintains the important identity, that the **wind stress** is 180° to the winds (**UBOT,VBOT**)

default ('ogrid')

new xgrid



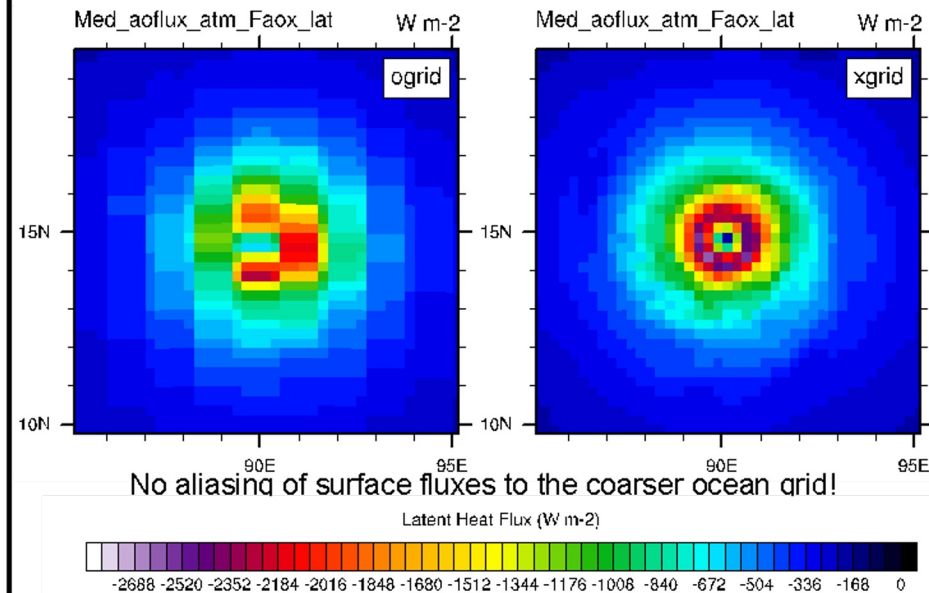
Vectors on ATM grid

lowest level wind (UBOT, VBOT) \rightarrow wind stress (TAUX, TAUY)

LHFLX after being mapped to the ATM grid

default ('ogrid')

new xgrid



Extensive simulations, including fully coupled 100 year runs, have been done to validate the exchange grid in CESM!!! This will be the default in CESM.

Easy to See and Modify Run Sequence

```
@1800  
MED med_phases_prep_ocn_accum_avg  
MED -> OCN :remapMethod=redist  
OCN  
@900  
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_aofluxes_run  
MED med_phases_prep_ocn_merge  
MED med_phases_prep_ocn_accum_fast  
MED med_phases_history_write  
@  
OCN -> MED :remapMethod=redist  
MED med_phases_restart_write  
@
```

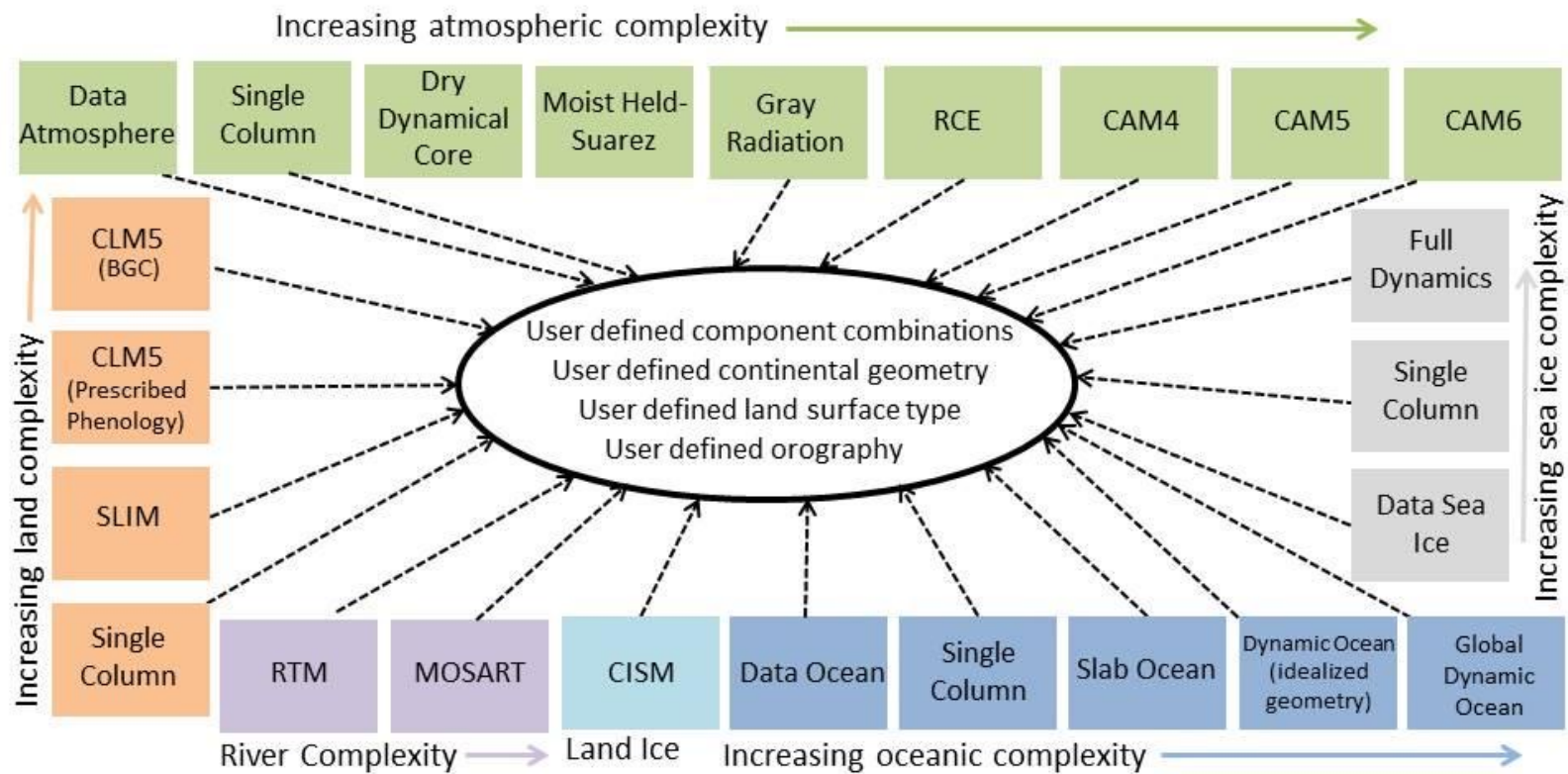
- Simple syntax for driver looping structure - component coupling frequency and order of component execution
- Connectors transfer data between mediator and components and are generated automatically – **no user code**
- Can bypass the mediator by simply introducing a connector between two components
- Components can have multiple named phases
- Run sequence can be changed without recompiling
- Sequential and concurrent execution in separate runtime configuration
- **Run sequence generated automatically by CIME/CCS**

Benefits of CMEPS – Greater Computational Efficiency

- Components sharing cores can now have different threading levels using **ESMF-managed threading**
- **Before:**
 - If component A is threaded 4 ways and component B is not threaded, if they are to share the same nodes, component B can **only** use 1/4 of the cores in a node
 - This leads to idle cores and poor HPC resource utilization
- **Now:**
 - If component A is threaded 4 ways and component B is not threaded, if they are to share the same nodes, component B can use **ALL** of the cores in a node
 - This greatly increases the efficiency of the overall model
- **Pre-industrial, fully coupled run (2° atm/1° ocn)**
 - **Now** Model Cost: 2531 pe-hrs/simulated_year
 - **Now** Model Throughput: 35 simulated_years/day
 - **Before** Model Cost: 3140 pe-hrs/simulated_year
 - **Before** Model Throughput: 31 simulated_years/day

Community Data Models for Earth Prediction Systems

CDEPS



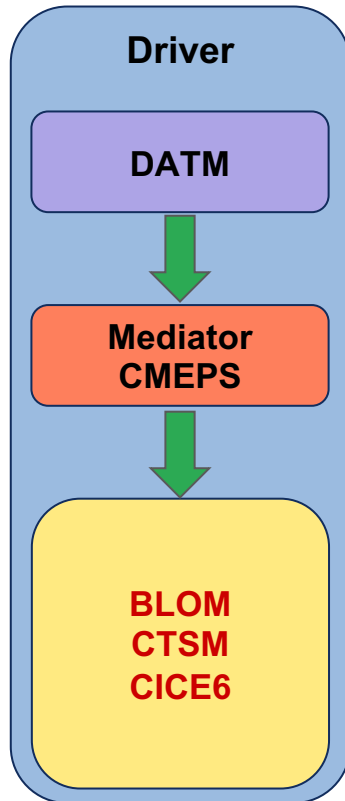
***Hierarchical Model Development:
A simple-to-more-complex comprehensive approach to
identify systematic biases and improve models.***

CDEPS

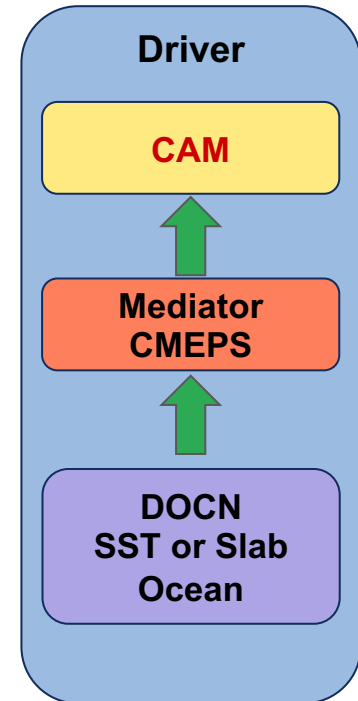
- **CDEPS contains ESMF/NUOPC compliant data components that are modular and flexible:** Can be used in any ESMF/NUOPC compliant modeling system
- **CDEPS** handles the ability to ingest multiple data sources with different spatial and temporal resolutions. Also provides ability to customize the ingested data (e.g. unit conversions)
- **All data is read with parallel IO (PIO2) – can easily ingest 2d or 3d fields!**
- **Automated regridding capability:** **1)** online regridding of 2D/3D fields, **2)** support for different regridding types such as conservative, patch, **3)** extrapolation and **4) various** time interpolations (coszen, bilinear, etc)
- **Inline data models:. CDEPS share code provides an interface that can be called directly from prognostic components and is used throughout CESM (future targeting of aerosol ingestion, nudging)**

CDEPS (cont)

CDEPS provides many different forcing scenarios out of the box



Data atmosphere CESM forcings
CORE2, JRA forcings for ocean
GSWP3, CRU, NLDAS forcings for land

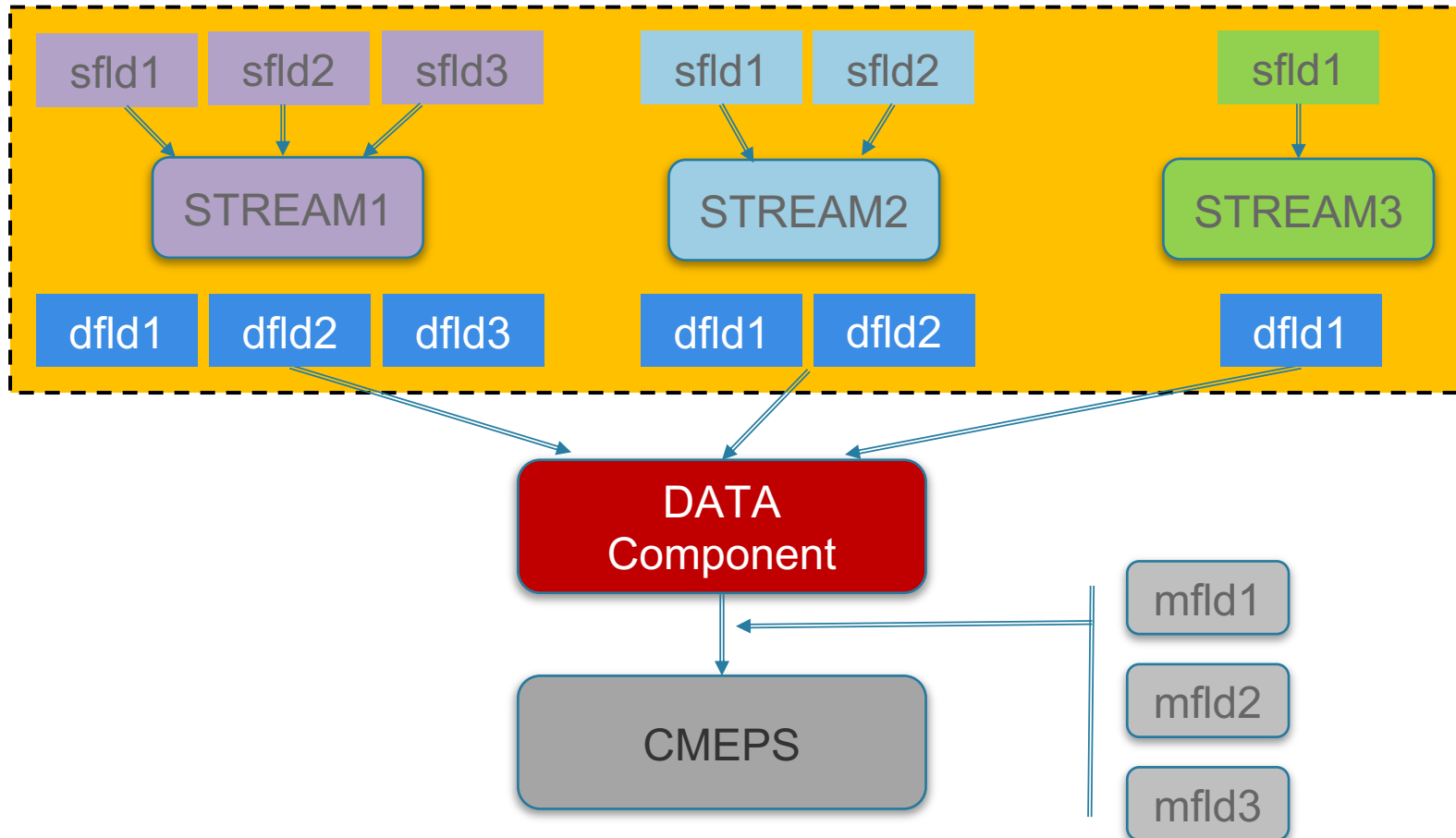


CDEPS Data Flow

sflds – stream fields on stream grid at stream time

dflds – stream fields mapped to data model mesh at data model time

shared code



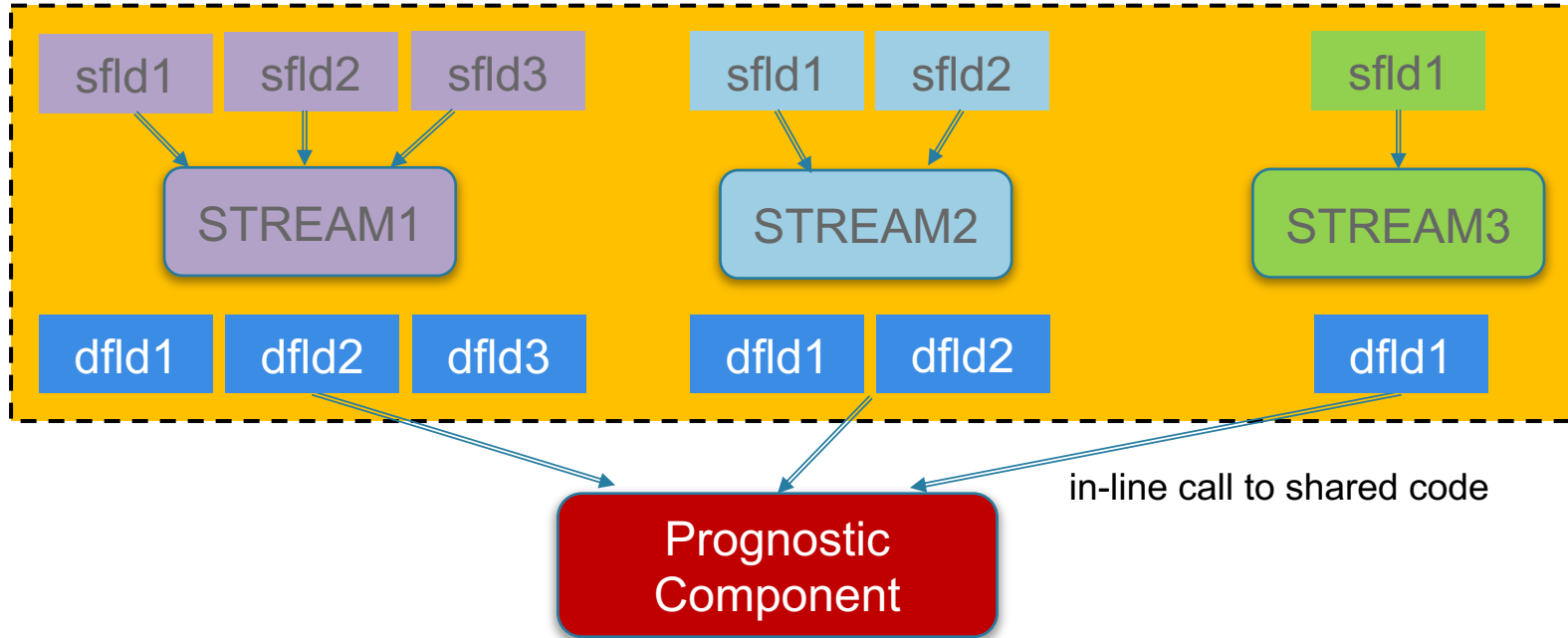
mflds – dflds manipulated to be fields sent to CMEPS

CDEPS In-line Functionality

sfld - on stream grid at stream time

dfld – on data model mesh at data model time

shared code



NOTE: no CMEPS mediator – this is just used internally by the prognostic component

Data Models Support Hierarchical Model Development

Hierarchical model development capability enables systematic model development

Provides ability to turn feedbacks on and off

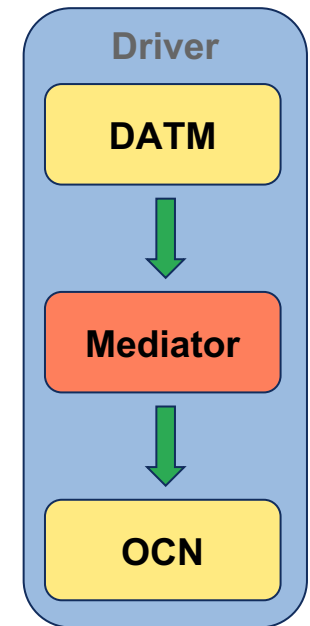
- Using forcing data eliminates coupling feedbacks

Reduces computational overhead

- Enables faster development cycle

Reduces time spent in debugging and testing

- Debugging can be done by isolating desired components
- Lightweight reproducer/s for problems can easily be setup



Example of CDEPS inline functionality

- Use case:
CAM ability to read in NDEP if its not produced by WACCM
- New module in CAM nuopc cap directory:
\$SRCROOT/components/cam/src/cpl/nuopc/atm_stream_ndep.F90
- Create a new stream data instance
call shr_strdata_init_from_inline(sdat_ndep,....
- Interpolate stream data to model grid and model time
call shr_strdata_advance(sdat_ndep, ymd=mcdte, tod=sec,...
- Map interpolated stream data to CAM model data
cam_out(c)%nhx_nitrogen_flx(i) = dataptr1d_nhx(g)
- **The CDEPS inline functionality has tremendous potential to be used throughout NorESM3 components to ingest forcing data as well as nudge data**

NorESM2.5 development status

Questions you might have.....

What is the status of CMEPS/CDEPS in NorESM3_alpha?

- All NorESM components have NUOPC caps

What is status of CMEPS/CDEPS in CESM?

- Extensive validation with multiple multi-century runs
- MCT/CPL7 no longer supported
- Some components (e.g. CICE6) ONLY have NUOPC caps
- ALL new science in CESM is being done with CMEPS/CDEPS

What about spinup capabilities in NorESM3_alpha?

- Still have ability to write coupler history files that can be used for spinup runs
 - e.g. coupler history files containing high frequency data from CAM that can then be used to spinup CTSM or BLOM
- Specification of data forcing is now a CMEPS namelist input rather than hard-coded as was done in the older MCT/CPL7 driver code
- CMEPS/CDEPS spinup capabilities have been carefully validated in CESM

New NorESMhub repositories:

- New prognostic component repositories:
 - forks of CAM, CTSM, CICE6, WW3DEV, MOSART, CISM
- New NorESMhub OSLO_AERO repository!
- New coupling and data component repositories:
 - forks of CMEPS and CDEPS
- New repository for machines and grids:
 - ccsn_config_noresm repository
 - Grids and machines are specified that are NorESM specific
- New scripting repository:
 - fork of CIME – latest new testing infrastructure capability

All NorESM development code runs and can be tested on betzy

More on this later!

Thank you!

Questions?