Update on CESM Software Engineering

CESM Breckenridge Workshop
June 2013

Mariana Vertenstein
CESM Software Engineering Group
(CSEG)

Thanks to NSF and DOE for their long-term support of the CESM Project
Over the last year

- New Scripting and Testing Infrastructure
  - New mechanism to specify compsets and grids
  - Simpler approach to port to new machines
  - New Testing/Porting Database
- New coupling infrastructure
  - Separation of river runoff into its own component
  - Introduction of new wave component (WaveWatch III)
  - Multi-instance support in release code
- New major component developments
  - CAM5.3 and CLM4.5
Over the last year (cont)

- New PIO capabilities and documentation
- New machine and compiler support
  - yellowstone, bluewaters, edison, mira, titan, stampede
  - NAG support added
- New CESM Forum Bulletin Board
- Three releases to community
  - CESM1.2.0, CESM1.1.1, CESM1.0.5
- Plan to support most F2003 features in all post CESM1.2 development code
New grid naming convention

Each model resolution can be specified by its alias, short name and long name.

Example of equivalent alias, short name and long name:
- alias: f19_g16 (atm/Ind_ocn/ice)
- short name: 1.9x2.5_gx1v6
- long name = a%1.9x2.5_l%1.9x2.5_oi%gx1v6_r%r05_m%gx1v6_g%null_w%null

atm Ind ocn/ice river Ind Ind-ice wave

New compset naming convention

Each model compset can be specified by its alias, short name and long name.

Example of equivalent alias, short name and long name:
- alias: B1850
- short name: B_1850
- long name = 1850_CAM4_CLM40%SP_CICE_POP2_RTM_SGLC_SWAV

time atm Ind ice ocn river Ind-ice wave
New Scripting Infrastructure

- New xml-based mechanism for specifying model grids and component sets
- Can now add new compset on the “fly” by using `user_compset` argument to `create_newcase`

http://www.cesm.ucar.edu/models/cesm1.2/cesm/doc/modelnl/compsets.html
- New grids can now be supported more easily

http://www.cesm.ucar.edu/models/cesm1.2/cesm/doc/modelnl/grid.html

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
</table>
| WAV_GRID | null
| WAV_NX   | 0           |
| WAV_NY   | 0           |

null is no grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
</table>
| CISM_GRID | null
| null      | is no grid: |

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM_GRID</td>
<td>ne30np4</td>
</tr>
<tr>
<td>ATM_NX</td>
<td>48602</td>
</tr>
<tr>
<td>ATM_NY</td>
<td>1</td>
</tr>
</tbody>
</table>

ne30np4 is Spectral Elem 1-deg grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>LND_GRID</td>
<td>ne30np4</td>
</tr>
<tr>
<td>LND_NX</td>
<td>48602</td>
</tr>
<tr>
<td>LND_NY</td>
<td>1</td>
</tr>
</tbody>
</table>

ne30np4 is Spectral Elem 1-deg grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLC_GRID</td>
<td>ne30np4</td>
</tr>
<tr>
<td>GLC_NX</td>
<td>48602</td>
</tr>
<tr>
<td>GLC_NY</td>
<td>1</td>
</tr>
</tbody>
</table>

ne30np4 is Spectral Elem 1-deg grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICE_GRID</td>
<td>gxlv6</td>
</tr>
<tr>
<td>ICE_NX</td>
<td>320</td>
</tr>
<tr>
<td>ICE_NY</td>
<td>384</td>
</tr>
</tbody>
</table>

gxlv6 is Greenland pole v6 1-deg grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCN_GRID</td>
<td>gxlv6</td>
</tr>
<tr>
<td>OCN_NX</td>
<td>320</td>
</tr>
<tr>
<td>OCN_NY</td>
<td>384</td>
</tr>
</tbody>
</table>

gxlv6 is Greenland pole v6 1-deg grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>MASK</td>
<td>gxlv6</td>
</tr>
</tbody>
</table>

gxlv6 is Greenland pole v6 1-deg grid:

<table>
<thead>
<tr>
<th>Grid Type</th>
<th>Grid Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROF_GRID</td>
<td>ro5</td>
</tr>
<tr>
<td>ROF_NX</td>
<td>720</td>
</tr>
<tr>
<td>ROF_NY</td>
<td>560</td>
</tr>
</tbody>
</table>

ro5 is 1/2 degree river routing grid:
New Testing Infrastructure

• New xml based testing infrastructure
  – easier to query existing tests for target component configuration and platform and determine test coverage

• New namelists tests
  – able to now separate answer changes due to code changes versus namelist changes

• New testing and porting database
  – much easier to track testing, problems and upcoming plans
### CESM_3.0 Testing and Porting Database

#### Planned Tags

<table>
<thead>
<tr>
<th>Tag Name</th>
<th>Last Update</th>
<th>Planned Tag Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>cesm1_3_beta01 (white board)</td>
<td></td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
<tr>
<td>cesm1_3_alpha01d</td>
<td>2013-06-15</td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
<tr>
<td>cesm1_3_alpha01c</td>
<td></td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
<tr>
<td>cesm1_3_alpha01b</td>
<td>2013-06-16</td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
<tr>
<td>cesm1_3_alpha01a</td>
<td>2013-06-15</td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
<tr>
<td>cesm1_3_alpha01b</td>
<td>2013-06-16</td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
<tr>
<td>cesm1_3_alpha01a</td>
<td>2013-06-15</td>
<td>Fix for WACCM and strat chem bug in which negative constituent concentrations cause the model to crash in a CLY/BRY family adjustment routine.</td>
</tr>
</tbody>
</table>

**CESM_3.0 "white board"**
CPL7 Infrastructure – this year

• New river component
  – Previously RTM was imbedded in CLM
  – Flooding can now be sent back to land
  – Easier to incorporate new capabilities in river component (riverine nutrient transport)

• New multi-instance component capability in release code
  – Used by DART to run fully coupled data assimilation for first time
  – Also being used by SPCAM to couple unique land to every subgrid column

• New wave component
  – Wind and surface waves -> Langmuir circulations in the upper ocean
  – Mixing in the upper ocean
  – Add Wave Watch III to provide needed inputs to POP2 Langmuir parameterization – now in development code base
CPL7 Infrastructure – next year

• Enable on-line regridding capability
  – working with ESMF
  – need to establish standard for how components send vertices as well as cell centers
  – support for regionally refined grids (MPAS, CAM-SE)

• Prototype new mechanism for components to exchange fields
  – only exchange what we need – components tell coupler what they will send and what they need

• Generalize routing capability
  – New possible flow of information between components requires reworking current code
Feedbacks from CISM

**CAM**

- 10 glacier elevation classes

**CLM**

- Surface Elevation
- Icesheet Area
- Surface Mass Balance
- Surface Temperature

**CISM**

- Ice Calving

**CICE** (Sea Ice)

Area of cropland, urban, etc.
CAM Software Engineering – last year

• **New prescribed aerosol module extensions for modal aerosols (in CAM5.3)**

• Refactored MG microphysics for MG2 development

• Improved debugging capabilities
  – New validation of physics_state objects w/ namelist control.

• New subcolumn infrastructure

• CHEM infrastructure:
  – improved robustness of chemistry preprocessor

• WACCM infrastructure:
  – new support for WACCM-SE and unified treatment of WACCM4/WACCM5

• Many new "use cases" to support new science.
CAM Software Engineering – this year

• Refactor physics driver layers to allow more flexibility in switching parameterizations on/off.
  – having a single code base for many parameterization development efforts
  – "unit" (subcomponent) testing
  – integrating the offline radiation driver

• Finish subcolumn infrastructure.
  – Begin implementing subcolumn schemes.

• Infrastructure extensions to allow the physics and dynamics to run on distinct grids
  – Using “non-scale” aware parameterizations – dynamics could be at high resolution and physics could maintain a coarser/tuned grid

• Add grid metadata (CF), and model metadata (CIM)
  – Working with ESMF
  – CIM is the model metadata standard used for the CMIP5 output
CLM/RTM Software Engineering – last year

- **New CLM4.5 capability – with backwards compatibility option for CLM4.0**
  - New soil biogeochemistry (including methane emissions), new crop, lake, fire and prognostic wetland models, plus other broad set or changes

- **RTM**
  - No longer imbedded in CLM – now a separate component at the driver level and new capability to return "flooding" back from RTM to CLM
  - Support for new 1/10° RTM

- **Testing**
  - Moved CLM test framework to new CESM testing framework
  - New streamlined testing capability for CLM developers

- **Datasets**
  - New high resolution raw datasets (for surface dataset creation)
  - Capability to run offline (I compsets) with CRUNCEP forcing

- **New BGC spinup capability**
  - Replaced complex CPPs with namelist input – much more robust
CLM/RTM Software Engineering - next year

- Large refactoring effort underway
  - clmtype refactoring – only one level of indirection (much simpler code)
  - Remove pointers in routines – use F2003 associate feature.
  - Remove ALL CPP ifdefs!
- Introduce dynamic landunit capability
  - landunit fractions (glacier, crop, etc.) are fixed after initialization – want ability for landunit areas to change over time – changing crop area or changing ice sheet area
  - involves revisiting many CLM design decisions and implicit assumptions
- Incorporate new CLM Ecosystem Demography model
  - represents vegetation successional processes post-disturbance
- Transient CLM configurations to include time-varying CO2 streams by default.
- RTM functionality
  - add riverine transport of nutrients and new DLND capability that allows for stand-alone testing of RTM
CISM Software Engineering – last year

• New build capabilities
  – added capability to include C++ code in the CESM build – allows linking to Trilinos and other libraries using their native C++ interface
  – added support for building component models with cmake
• Added parallel capabilities to CESM’s glacier component
• Added CISM2 code to CESM development trunk
  – includes higher-order dynamics, parallel capabilities
  – current status: builds but does not run successfully
CISM Software Engineering – this year

• Finish integration and testing of CISM2
• Enable feedbacks from changing ice sheet area to the rest of CESM
  – changing landunit areas in CLM
  – liquid and ice runoff
  – changing topographic heights in CAM
• Improve CISM's mapping capabilities
  – map fields conservatively; challenge in the presence of multiple glacier elevation classes
  – allow for unstructured land grids
  – perform mapping in parallel
  – handle multiple ice sheet grids (e.g., Greenland and Antarctica)
CVMIX
Community Ocean Vertical Mixing Parameterization

• Multi-lab collaboration (LANL, GFDL, NCAR) to create an easy to use library of vertical mixing parameterizations that can be shared among multiple ocean models
  – POP, MPAS-O, MOM6
• No change in POP interface – but changes under the hood
• Ability to run stand-alone single-column vertical mixing tests
• Should be available by end of summer
POP2

• New blockone decomposition added
  – 1d decomp (like spacecurve) but no constraint on the block size

• New general point-to-point (P2P) communication (Allison Baker)
  – updated global reductions with local communication
  – applied to overflow code (14x speedup on 480 procs)

• New ecosystem module for CESM 1.2.0 (Keith Moore and Keith Lindsay)
MPAS–O
Doug Jacobsen and Todd Ringler (LANL)

- Efforts to couple are underway – a lot of progress has already been made
- Currently, running “C” compset with MPAS–O rather than POP
  - DATM (T62) + MPAS–O (120km)
- Can be run for 5 years within CESM on LANL linux cluster
  - Climate looks good!!!
- Goal is to have fully coupled MPAS–O, CAM, CLM, MPAS–I (sea ice) running by April
Other Activities – this year

• New Architectures and CESM
  – Collaborate in effort to migrate CESM to upcoming architectures – GPU, MIC, bluegene

• New CESM Workflow
  – Need to reduce total data volume stored on long term archive – exploring several paths including lossy compression (John Dennis, Jim Edwards, Mariana Vertenstein)

• Performance
  – Automated load balancing across components, new decompositions and communication algorithms within components (ANL, ORNL, NCAR)