CESM Software Engineering Group (CSEG) Update

Mariana Vertenstein
NCAR Earth System Laboratory
CESM Software Engineering Group (CSEG)

CESM is primarily sponsored by the National Science Foundation and the Department of Energy
We have been very busy!

- New model infrastructure
  - grids, model validation, namelists, build
- New model components
- New model run database
- Component developments
  - infrastructure and science
- CESM1.0.4 and new release policy
New Infrastructure

- **Utilities**
  - new regridding capability
  - new offline utilities greatly simplify and speedup addition of new model grids
  - new port validation functionality

- **Scripts**
  - New xml based build functionality (simplifies porting and debugging)
  - New unified namelist generation capability
    - unified mechanism for component namelist modification
    - access to all component namelists from easy to view web page
  - New heirarchical approach to specify compsets
New Infrastructure (cont)

- CPL7
  - New multi-instance component capability
  - New implementation for component field exchange
    - can now specify new fields in driver namelist
  - New mapping capability for vector fields
    - 2 non-conservative mapping options available
Challenge of New Grids

- Past grids were global and logically rectangular
- Ocean/Sea-Ice grids
  - displaced pole, tripole, new MPAS regionally refined icosohedral
- Atmosphere/Land grids
  - lat/lon, cubed sphere, new regionally refined cubed sphere, new regionally refined MPAS
- Grids span wide range
  - low res (3 degree) to ultra-high res (0.125° ATM/LND and 0.1° OCN/ICE)
- Regridding done in parallel at runtime using mapping files generated offline
  - both conservative and non-conservative mapping files needed
New Mapping Capabilities

- In past new mapping files were rarely generated – this has changed!!!
- New simplified tool chain to create mapping files (and other grid specific datasets) as a pre-build step
  - leverage new ESMF offline regridding capability (parallel and robust)
  - workflow went from hours to minutes
  - new support for user specified grids in new top level mapping directory
New Component Grid Capabilities

- CLM, CICE (prescribed), DOCN can now be run on any unstructured grid.
- Match target CAM grid (including locally refined grids).

- Can now carry out CAM prescribed forcing runs _with all models on the same grid_ (traditional approach).
  - Significant speed up for hi-res CAM-SE prescribed forcing runs.
CSSEF: Can MPAS grid be supported?

with Michael Duda (NCAR), Todd Ringler (LANL)

CAM, CLM
1° FV

Pres–CICE/DOCN
120 km MPAS 10x
refined on Cal Coast

SST/ICE-COV Forcing
10 km POP/CICE
coupled data
Next Steps

- ESMF online regridding
  - Collaborating with ESMF to look at leveraging on-line regridding in driver
  - ESMF now also supports 3d conservative mapping

- MCT/MOAB (CSSEF)
  - Longer term effort that will enable support for adaptive mesh refinement and fault recovery
New Model Build System
(Jim Edwards)

- Huge reduction in number of machine specific files!
  Compiler options, machine specific paths, model default decompositions in 3 xml files
- `config_compilers.xml (new)`
  - compiler options now set by hierarchy of compiler vendor, OS, machine
  - more consistency across machines
  - easier to define new machines
- `config_machines.xml (extended)`
  - machine specific info and *new list* of supported compilers go here
  - one machine name – different compilers!
- `config_pes.xml`
New Build System (cont)

• machine name no longer includes compiler
• machine specific environment and batch system files
  - mkbatch.machine
    • machine batch system interface is here
  - env_mach_specific.machine
    • module or dotfiles and machine specific environment variables go here
New Validation Process
(Mike Levy, Brian Eaton, Andy Mai)

- CAM error growth test no longer works with CAM5 physics
- “Vague” model validation process
  Carry out a 20-30 year 1.9x2.5_gx1v6 resolution, B_1850_CN compset simulation and compare the results with the diagnostics plots for the 1.9x2.5_gx1v6 Pre-Industrial Control (see the CESM diagnostics)
- Need easily adaptive process that keeps up with changes to model physics, dynamics and resolutions and returns Success or Fail.
What will CSEG do?

• Generates baseline - run ensembles (size TBD) of 12-month runs
  - e.g. B compset, CAM5 physics, and CAM spectral element dynamics

• Difference between initial conditions is perturbation to atmospheric temperature

• For each CAM output variable (e.g. U) at each grid point compute
  - Ensemble mean
  - RMSE (of all grid points) of each member relative to ensemble mean – this will be ONE number (score)
  - Mean and standard deviation of the RMSE of each ensemble member
What do you do?

- Run one single 12-month simulation with same setup as CSEG baseline
- Compute RMSE for each variable (compared to ensemble output mentioned above).
- See where RMSE falls in the distribution of the baseline RMSEs
- Success -> not being 3 standard deviations over the ensemble mean
- CSEG will provide a utility to do this!
New Namelist Generation

- **Uniform** xml-based build-namelist functionality across all model system
  - CAM, CLM, CICE, POP2, DATA models, (DRV)
- In case directory, each component (xxx) has a file, user_nl_xxx, where all namelist changes are made
  - user namelist mods are now in **ONE** place
  - new error checking functionality
Plus.....

- Directly call component build-namelist from Buildconf/ shell scripts – user never touches Buildconf/ files

- New utility (preview-namelists) in case directory to preview model namelists

- New auto-generated web pages document all model namelists for each development tag
Each development tag (and release version) now has a page for all model namelist and script xml files

**CESM Component Models Namelist Definitions (cesm1.1_beta15)**

DRV Namelists
CAM Namelists
CICE Namelists
CISM Namelists
CLM Namelists
POP2 Namelists
DATM Namelists
DICE Namelists
DLND Namelists
DOCN Namelists

**CESM xml files (cesm1.1_beta15)**

Supported Grids (config_grid.xml)
$CASEROOT Settings (env_case.xml, env_configure.xml, env_build.xml, env_run.xml)
Available Machines (config_machines.xml)
Each component has its own searchable namelist page

### DRV: Component Field Exchange

<table>
<thead>
<tr>
<th>Namelist Variable</th>
<th>Type</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>cplflds custom</td>
<td>char*1024</td>
<td>seq_cplflds_userspec</td>
</tr>
<tr>
<td>flida_o2 dmea</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>flida oo2a</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>flida oo2b</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>flida oo2c</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>glo nuc</td>
<td>integer</td>
<td>seq_cplflds_inparm</td>
</tr>
</tbody>
</table>

### DRV: Experiment Definition

<table>
<thead>
<tr>
<th>Namelist Variable</th>
<th>Type</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>bfbflag</td>
<td>logical</td>
<td>seq_infodata_inparm</td>
</tr>
<tr>
<td>branch retain casename</td>
<td>logical</td>
<td>seq_infodata_inparm</td>
</tr>
<tr>
<td>budget month</td>
<td>integer</td>
<td>seq_infodata_inparm</td>
</tr>
<tr>
<td>case desc</td>
<td>char*256</td>
<td>seq_infodata_inparm</td>
</tr>
<tr>
<td>case name</td>
<td>char*128</td>
<td>seq_infodata_inparm</td>
</tr>
</tbody>
</table>
**DRV: Component Field Exchange**

<table>
<thead>
<tr>
<th>Namelist Variable</th>
<th>Type</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>cplflds_custom</td>
<td>char*1024</td>
<td>seq_cplflds_userspec</td>
</tr>
</tbody>
</table>

New fields that are user specified can be added as namelist variables by the user in the cpl namelist seq_flds_user using the namelist variable array cplflds_custom. The user specified new fields must follow the above naming convention.

As an example, say you want to add a new state 'foo' that is passed from the land to the atm - you would do this as follows:
```plaintext
seq_flds_user
    cplflds_custom = 'Sa(foo->a2x)', 'Sa(foo->x2a')
```

This would add the field 'Sa foo' to the character strings defining the attribute vectors a2x and x2a. It is assumed that code would need to be introduced in the atm and land components to deal with this new attribute vector field.

Modify user_nl_cpl to edit this.

<table>
<thead>
<tr>
<th>flds_co2_dmsa</th>
<th>logical</th>
<th>seq_cplflds_inparm</th>
</tr>
</thead>
<tbody>
<tr>
<td>flds_co2a</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>flds_co2b</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>flds_co2c</td>
<td>logical</td>
<td>seq_cplflds_inparm</td>
</tr>
<tr>
<td>glc_nec</td>
<td>integer</td>
<td>seq_cplflds_inparm</td>
</tr>
</tbody>
</table>
New CESM Components

- CESM/WRF
  - Capability to have CESM run as a regional model with an active atmosphere

- Upcoming
  - ROMS
  - Wave Watch III
  - NEMO
  - RTM as a separate new component
WRF version off of 3.2 integrated into development version of CESM1.1
most work done by RASM (regional arctic system model)
CESM/WRF collaboration to review implementation, independently validate science and develop a long term strategy
work still needed for full compatibility with CESM scripts/build/IO
multi-decadal runs underway for regional US wrf/clm/docn and ongoing for multi-decadal arctic wrf/vic/pop/cice
CPL Update

- New specification of fields transferred between components
  - can now add fields (based on specified naming convention) to be exchanged as namelist variables
  - cpl code does not have to be modified to pass new fields between components

- New mapping capability
  - non-conservative mapping supports patch and bi-linear
  - cart3d for vector mapping

- Multi-component instance support
CPL Update (cont)

Multiple Components Instances and DART data assimilation
All instances run concurrently on non-overlapping processor sets –
number of instances set by ONLY 1 line of an xml file
Used for CMIP5 ocean data assimilation – 48 POP instances

DART data assimilation in CESM

- Obs
- DART
- CAM
- POP
- Coupler
- CICE
- CLM
- 3D restart
CISM Update

- New compset (CISM/DLND) for spinup
  - Run model with annual time step
- Upcoming addition of CLM dynamic landunits
  - Permits 2-way feedbacks between CLM and CISM
  - Permits specified changes in crop and urban areas, e.g., for linking with an integrated assessment model
- Upcoming incorporation of CISM2
  - Includes higher-order ice-sheet dynamics
CAM Update

- **Infrastructure**
  - Refactored physics buffer -- removed restrictions on buffer dimensions and improved API.
  - Radiative constituent module -- supports modal aerosols (previously restricted to bulk aerosols).
  - Prescribed aerosol module -- extended for modal aerosols.
  - Backwards compatibility -- support for older physics and chemistry packages as new ones are added.

- **Scientific functionality**
  - WACCM-X (CESM1.0.4)
  - New chemistry configurations
  - CARMA (Community Aerosol and Radiation Model for Atmospheres)
    - Flexible and extensible microphysics package for clouds and aerosols
CAM upcoming priorities

- Finalize prescribed modal aerosols.
- CLUBB (unified cloud macrophysics, shallow cumulus, and vertical diffusion package)
- WACCME-SE and validation of WACCMM-5
- Non-hydrostatic MPAS dynamical core
- SPCAM option (embedded cloud resolving model)
- Grid metadata (CF), and model metadata (CIM) collaboration with ESMF (CIM is the model metadata standard being used for the CMIP5 output)
CLM Update

Infrastructure

- New surface dataset generation and new CLM→RTM mapping
  - Uses offline generated mapping files
  - Faster, supports high-res and non lat/lon grids
  - High resolution option for RTM (1/10th degree 5X increase)
- Interpolation of non-lat/lon output to lat/lon grids
CLM upcoming priorities

- Major new science changes - CLM4.5 - targeted for CESM1.1
- Ability to run with CLM4.0 physics OR CLM4.5 physics (backwards compatibility for first time!)
- A few examples
  - New lake model with variable depths, and modeling soil layers beneath it, and snow layers above
  - New methane model
  - Allow irrigation to work for ALL crop types
  - Improvements to prognostic crop model and soil hydrology
  - New urban multi-density types
Releases

- CESM1.0.4 this February
- Upcoming:
  - New plan for releases
  - New plan for performance database