Workflow in CESM2

Jim Edwards
NCAR/CGD
jedwards@ucar.edu

Building reproducible workflows for earth sciences
ECMWF Oct 14-16, 2019
Overview

1. CESM2, CIME and the Case Control System.
2. Expanding CIME workflow capability
3. Projects using CIME workflow:
   a. Basic experiment workflow
   b. Ensemble workflow using CYLC
   c. NCAR/ICCC CMIP6 Large Ensemble
   d. Seasonal Forecast Experiment
CESM2 Components
**CESM2**

- **Version Control**
  - Each component has its own version control and governance
  - Top level repo is a collection of hashes referring to component repositories
  - Component repos may use git or svn
  - Managed via an NCAR developed tool manage_externals

- Supports a diverse user community and range of hardware
### CESM Development Management database

#### CESM2_1_alpha01
- **Tag Name:** cesm2_1_alpha01 (white board)
- **Status:** Done
- **Date:** 2018-12-11
- **Description:**
  - There are known issues with the pgI and cray compilers.
  - Intel 18 issues on hobart with larger tests.
  - PGi on hobart fails to run across multiple processors.

#### cesm2_1_alpha01b
- **Tag Name:** cesm2_1_alpha01b
- **Status:** Done
- **Date:** 2018-09-18
- **Description:**
  - f19_g17_BHST missing 2 degree CO2 flux fuel file.
  - f9_g17_BHST Diagnostic CO2 answer changes.
  - There are known issues for the pgI and cray compilers.

#### cesm2_1_alpha01c
- **Tag Name:** cesm2_1_alpha01c
- **Status:** Done
- **Date:** 2018-10-04
- **Description:**
  - _RT1_NM3_Ld7_f19_g17_BHSTWs.cheyne_intel.allicative-defaultio still failing.
  - There are known issues with the pgI and cray compilers.

#### cesm2_1_alpha01d
- **Tag Name:** cesm2_1_alpha01d
- **Status:** Done
- **Date:** 2018-10-16
- **Description:**
  - _SM5_Ld1_f09_g17_B18150.cheyne_intel.allicative-cmp16 is expected to fail.
  - _RT1_NM3_Ld7_f19_g17_BHSTWs.cheyne_intel.allicative-defaultio still failing.
  - There are known issues with the pgI and cray compilers.

#### cesm2_1_alpha01e
- **Tag Name:** cesm2_1_alpha01e
- **Status:** Done
- **Date:** 2018-11-15
- **Description:**
  - _SM5_Ld1_f09_g17_B18150.cheyne_intel.allicative-cmp16 is expected to fail.

#### cesm2_1_alpha01f
- **Tag Name:** cesm2_1_alpha01f
- **Status:** Done
- **Date:** 2018-12-03
- **Description:**
  - There are known issues with the pgI and cray compilers.
  - Intel 18 issues on hobart with larger tests.
  - PGi on hobart fails to run across multiple processors.
CESM2 Experiment Database

CESM Case Working Title & Details

CMIP6 20th century experiments (1850-2014) with CAM6, interactive land (CLM5), coupled ocean (POP2) with biogeochemistry (MARBL), interactive sea ice (ICE5.1), and non-evolving land ice (CISM2.1)

CMIP6 Project Details

Case Workflow Process Status

Caseroot Settings

Case Notes

Case Diagnostics, Process and Journal Publication Links

Data Publication Options

CESM Case Working Title & Details

CMIP6 20th century experiments (1850-2014) with CAM6, interactive land (CLM5), coupled ocean (POP2) with biogeochemistry (MARBL), interactive sea ice (ICE5.1), and non-evolving land ice (CISM2.1)

CMIP6 Project Details

Case Workflow Process Status

Caseroot Settings

Case Notes

Case Diagnostics, Process and Journal Publication Links

Data Publication Options

CESM Case Working Title & Details

CMIP6 20th century experiments (1850-2014) with CAM6, interactive land (CLM5), coupled ocean (POP2) with biogeochemistry (MARBL), interactive sea ice (ICE5.1), and non-evolving land ice (CISM2.1)

CMIP6 Project Details

Case Workflow Process Status

Caseroot Settings

Case Notes

Case Diagnostics, Process and Journal Publication Links

Data Publication Options

CESM Case Working Title & Details

CMIP6 20th century experiments (1850-2014) with CAM6, interactive land (CLM5), coupled ocean (POP2) with biogeochemistry (MARBL), interactive sea ice (ICE5.1), and non-evolving land ice (CISM2.1)

CMIP6 Project Details

Case Workflow Process Status

Caseroot Settings

Case Notes

Case Diagnostics, Process and Journal Publication Links

Data Publication Options

CESM Case Working Title & Details

CMIP6 20th century experiments (1850-2014) with CAM6, interactive land (CLM5), coupled ocean (POP2) with biogeochemistry (MARBL), interactive sea ice (ICE5.1), and non-evolving land ice (CISM2.1)

CMIP6 Project Details

Case Workflow Process Status

Caseroot Settings

Case Notes

Case Diagnostics, Process and Journal Publication Links

Data Publication Options
CIME: Common Infrastructure for Modeling the Earth

A collection of earth system model infrastructure tools

- Case Control System
  - Create, build and run an experiment
  - Unit testing
  - System testing
  - Record experiment provenance
- Inter-component Coupling Infrastructure
- Data Models (permits control of feedbacks)

CIME is used by CESM, E3SM, NORESM and is being considered for use in UFS
The CIME Case Control System (CCS)

The CCS is an extensible, coherent and coordinated set of object-oriented python scripts which uses a library of python objects along with data in xml to design, build and control the run of an earth system model.

The CCS has been designed to facilitate and encourage community collaboration!
Basic climate model workflow

1. Create an experimental case
2. Build the model
3. Run the model
4. Archive model output
5. Postprocess
CCS workflow generator

The CCS provides a basic workflow generator which uses queueing system native dependency tools to schedule jobs in a workflow.

Limitations:
- all jobs are submitted to queues
- no submission clock or calendar support
- limited to a single case

Use `preview_run` to view the current workflow.
XML Elements of a workflow definition.

- `workflow_jobs` `{case, [prepend], [append]}`
  - `job` `{name}`
    - `template` (script template to submit)
    - `dependency` (other job that must complete first)
    - `prereq` (logical to include in workflow)
    - `runtime_parameters`
      - `task_count`
      - `tasks_per_node`
      - `walltime`
Adding CYLC

CCS provides a script `generate_cylc_workflow.py` to translate a CCS workflow to a CYLC suite.rc

- Provides support for ensembles
- Allows the user to customize workflows with all of the extensive feature set of CYLC

https://cylc.github.io/
CESM2 Large Ensemble Experiment

- Cooperative project with ICCP South Korea
- 100 member ensemble climate study running from 1850-2100
- Will be run on the ICCP system Aleph with postprocessing and data storage at NCAR
NCAR/ICCP CESM2 CMIP6 Large Ensemble Experiment Workflow
Subseasonal to Seasonal prediction using CESM and CYLC

Running in hindcast mode for each Monday between 1999 and 2019.

Currently capable of 132 simultaneous model simulations.
Foundational CESM2 workflow experience:

We would like to acknowledge and credit the work done by NCAR’s ASAP group, especially Sheri Mickelson, in instrumenting CESM2 with CYLC for the CMIP6 experiments.

- CMIP6 Experimental Status (since August 2018)
  - Have run 979 different CESM cases.
  - Published 690 cases.
  - Generated \(~1.3\) PB of compressed (lossless) time series files.
  - Published \(~310\) TB of compressed CMIP6 files to ESGF.
CESM Ensemble verification test

CESM2 Ensemble Verification

Please see CESM2 Python Tools for details or for help with this form see DiscussCESM
CIME

github repository:  
https://github.com/ESMCI/cime

documentation:  
esmci.github.io/cime

developers guide:  

Questions?

Thank You