Workflow in CESM2

Jim Edwards
NCAR/CGD
jedwards@ucar.edu

CESM Software Engineering Working Group Meeting June, 17 2020
Workflow

The sequence of steps involved in moving from the beginning to the end of a working process.

Many CESM workflows are mostly manual processes which require human intervention at multiple points.

The CIME workflow tools are designed to reduce human intervention to the extent possible.
Workflow control within case.run

- **PRERUN_SCRIPT**
- model.exe
- **DATA_ASSIMILATION_SCRIPT**
- **DATA_ASSIMILATION_CYCLES**
- **POSTRUN_SCRIPT**
Workflow control in the case control system

RESUBMIT: While RESUBMIT>0 upon completion of a case.run->archive cycle, the job is resubmitted to the queue and the RESUBMIT value is decremented. CONTINUE_RUN is set to True (unless RESUBMIT_SETS_CONTINUUE_RUN=False)

The --resubmit-immediate option to the case.submit script will cause all RESUBMIT jobs to be submitted at once with queueing system dependancies.
Basic climate model workflow

- create an experimental case
- build the model
- run model
- archive model output
- postprocess

RESUBMIT
Extending the workflow

Workflows are defined in:
cime/config/cesm/machines/config_workflow.xml

./create_newcase --case mycase01 --compset FHIST --res f19_f19_mg17 --workflow timeseries

This will add a timeseries generation step.
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
Controlling the Workflow

- `case.submit` (submits end-to-end workflow)
- `case.submit` --resubmit-immediate
- `case.submit` --job `case.st_archive`
- `case.submit` --only-job `case.run`
Ensemble Workflow

create_clone --clone mycase --case mycase01 --ensemble 8 --keepexe

Will create 8 new cases
mycase01 - mycase08 all using the cesm.exe from mycase.
Adding CYLC

CCS provides a script `generate_cylic_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://cylic.github.io/
cime/scripts/Tools/create_cylc_workflow.py
/glade/scratch/jedwards/mycase01 --ensemble 8 --cycles 14

Will create a CYLC suite.rc file to set RESUBMIT=8 and run the ensemble.
Interglacial LandIce experiment

10 Year Model Run

Topography adjust

archive
MR Weather App

- chgres input interpolation
- GFS Model Run
- archive
- postprocess

https://github.com/ufs-community/ufs-mrweather-app
CESM2 Large Ensemble Experiment

- Cooperative project with ICCP South Korea
- 100 member ensemble climate study running from 1850-2100
- Running 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

- get initial data
- generate initial perturbation files
- build and configure model
- Archive history data
- run 11 member ensemble for 45 days
- postprocess
- dispose of data

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

Currently capable of 132 simultaneous model simulations.
CIME

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

documentation:  esmci.github.io/cime


Questions?

Thank You
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.