Run CESM with Chemistry on Derecho

***** PAGE actively in development *****

You will need to ssh into

```
derecho.hpc.ucar.edu
```

> **Tips from CISL:**
>
> **Tips for moving from Cheyenne to Derecho on https://ncar-hpc-docs.readthedocs.io/**

> **Comprehensive slides of information from NCAR/CISL about the Derecho HPC:**
>
> https://docs.google.com/presentation/d/1BAYI9NdyAnZzU3KLat0CK2sQj34XE_K5Wan9zL8rsQ0/edit?usp=sharing
>
> Old instructions for running on Cheyenne (the retired NCAR HPC) are here: Run CAM-Chem on Cheyenne

CESM2.2 on Derecho

Your $HOME and $WORK directories (on glade) are the same path on cheyenne and derecho. Therefore, it is recommended to create a new directory (in $HOME or $WORK) for cases you will run on derecho, to not confuse them with cases built on cheyenne (which will not run on derecho). Derecho and cheyenne have separate $SCRATCH directories.

(A) Download a copy of the model source code to your own directory

1. The first time, in your $HOME or $WORK directory, setup a folder to hold all your derecho cases.

```
> mkdir /glade/work/$USER/derecho_cases
```

Make a copy of the new version of the code to set up for use on derecho:, the latest code can be downloaded following the instructions at: http://www.cesm.ucar.edu/models/cesm2/release_download.html

This directory is referred to as $CESM_ROOT below. This should only need to be done once for each code base update of CESM. Note that when running on derecho you do not need to get any of the input data files.

> **Tip for New Users:**
>
> Having problems? Try the Troubleshooting page.

Downloading the source code will be completed by using the git clone feature. For example, first download the latest release (e.g. CESM2.2.0, with recent bug fixes):

```
> cd /glade/work/$USER/derecho_cases
> git clone https://github.com/ESCOMP/CESM.git cesm2.2.2
> cd cesm2.2.2
> git checkout cesm2.2.2
```

You should see in the $CESM_ROOT directory:

```
ChangeLog           cime_config         describe_version  Externals.cfg  manage_externals  tools
ChangeLog_template  CODE_OF_CONDUCT.md  doc               LICENSE.txt    README.rst
```

Download the model component source code.

```
> ./manage_externals/checkout_externals
```

You should now have in the $CESM_ROOT directory:

```
ChangeLog           cime_config         describe_version  Externals_cime.cfg  README.rst
ChangeLog_template  CODE_OF_CONDUCT.md  doc               LICENSE.txt
cime                components          Externals.cfg     manage_externals
```

with the new directories cime and components. The tools directory has been moved to within cime.

Get the optimized history-writing code for CLM with these commands: (Not sure this is still neeeded???)

```
> cd components/clm
> git fetch origin
> git checkout release-cesm2.2.01/hist_opt
```

If you want to run a regional refined model simulation on derecho, you may use the following sandbox in order to provide the same performance on derecho as it did on cheyenne:

```
/glade/work/fvitt/cesm/cesm2.2.2_musica
```

(B) Create a new case

2. Use the code in the model directory `$CESM_ROOT` to create a new case called `$CASENAME:`

```
> $CESM_ROOT/cime/scripts/create_newcase --case <your_path/$CASENAME> --res f09_f09_mg17 --compset FCnudged --
run-unsupported
```

From section **A**, `$CESM_ROOT` would be `/glade/work/$USER/my_cesm_sandbox`

A new directory `$CASEROOT = <your_path+$CASENAME>` is created. You may need to add `--run_unsupported` to the call line if you are not running with a scientifically validated compset and resolution combination.d

Note 1: If you need to specify a project on derecho, the command is --project $PROJECT_NUMBER.

Note 2: For additional help and options, type `./create_newcase -h`

Note 3: To find the possible compset types `$CESM_ROOT/cime/scripts/query_config --compsets cam.`

Note 4: The above call uses the new nudging scheme on 32 model layers, to use the old nudging scheme with 56 model layers use the compset: --compset FCSD, and adjust met files accordingly.

> **Tip for New Users:**
>
> Information on the compsets can be found here: http://www.cesm.ucar.edu/models/cesm2/config/compsets.html

(C) Set up your case

3. From within `$CASEROOT`

```
>./case.setup
```

4. Make changes to the model configuration using the `*.xml` files; you can edit the files directly or use the `xmlchange` tool (in your case directory).

Changes to `env_build.xml` must be made before building, or you will need to re-build.

See advanced pages to change the chemistry source code.

(D) Build the Executable

> **Tip for New Users:**
>
> You can use script xmlquery to query a variable in the xml files before modifying a variable with xmlchange command. For example:
>
> `>./xmlquery CALENDAR`

5. Compile and build the model in `$CASEROOT`

> **Note**
>
> First address any current bug-fixes: Bugs and Updates

```
>qcmd -- ./case.build


Note: you cannot run ./case.build interactively from the derecho prompt because it uses too much memory: you
must use 'qcmd'.
```

Note 1: You may need a project number to run qcmd: `qcmd -A $PROJECT_NUMBER -- ./case.build`

6. (Optional) Make changes to the model runtime setup: Changes to `env_run.xml` can be made at any time:

a. If not starting in the default 2005, change dates: see Changing Dates of a Run (also see relevant namelist changes)

b. The default option for biogeochemistry is “specified phenology” (satellite LAI) with CLM5 physics:

```
<entry id="CLM_BLDNML_OPTS" value="-bgc sp">

<entry id="CLM_PHYSICS_VERSION" value="clm5_0">
```

Other options are available, such as irrigated crops - please see the CLM documentation.

c. The default simulation is a test run for 5 days. Change `STOP_OPTION` and `STOP_N` to alter this to desired values.

7. (Optional) Check namelist settings in the namelist files `user_nl_cam` and `user_nl_clm`. Most CAM-chem related namelist variables are in `CaseDocs/atm_in`, but MEGAN and drydep are in `CaseDocs/drv_flds_in` (these files are created during build). To modify any of these, copy the appropriate lines to `user_nl_cam` and edit there.

> For example, if the startdate has been changed in `env_run.xml`, you have to also change the date of the initial meteorology file in `user_nl_cam` to start at the corresponding date.

For other changes see namelist changes or advanced options page.

> **Tip for New Users:**
>
> There are many namelist variables. You can find their definitions at: http://www.cesm.ucar.edu/models/cesm2/settings/current/cam_nml.html

After adding changes to `user_nl_*` files, optionally run:

```
>./preview_namelists
```

NOTE: most changes in `user_nl_*` files do not require re-building. However, during a run (CONTINUE_RUN = TRUE) no changes can be made to history output (fincl lists). If you want to change history output, create a new or branch run.

> **Note**
>
> **The pe-layout of existing compsets in the new code base of CESM2 for derecho (cesm2.2.2) has not been adjusted to the new computer, and running out of the box can lead to large differences in computer costs compared to running on cheyenne.**
>
> New PE-layouts are still being developed. We are working on updating these compsets. One way to increase performance is to use a namelist setting. To improve improve the performance of derecho, you can add the following to your `user_nl_cam` file.
>
> ```
> phys_loadbalance = 1
> ```

8. Check the run setup. In the `env_batch.xml` file make sure to have your project added correctly: `<entry id="PROJECT" value=$YOUR_PROJECT_CODE>` Depending on the version of CESM, you may instead find the entry id for PROJECT in `env_workflow.xml`

Change your walltime if desired: <entry id="JOB_WALLCLOCK_TIME" value="12:00:00"> (12 hours max)

9. If you make changes to `env_build.xml` variables or SourceMods after setting up and building, you may have to clean your setup and build again:

```
>./case.setup --clean
>./case.setup
```

OR

```
>./case.setup --reset
```

(which cleans, then does setup)

followed by

```
>./case.build --clean
>qcmd -- ./case.build
```

(E) Run the Model

10. Submit run to queue:

```
>./case.submit
```

While running, output is written to `<run_dir>: /glade/derecho/scratch/<username>/$CASENAME/run`

After the run completes successfully, output files are moved to the short term archive: `/glade/derecho/scratch/<username>/archive/$CASENAME/atm/hist` (similar directories exist for other model components: lnd, etc.).

Restart and initial conditions files are written to: `/glade/derecho/scratch/<username>/archive/$CASENAME/rest`

Note: long term archiving is currently not working

> **Tip for New Users:**
>
> For more options regarding the submission and running options type: case.submit --help

> **Tip for New Users:**
>
> Default output is monthly, so if you run a test 5-day simulation with monthly output, you will not see any files in the `atm/hist` location. However, restart files will have been created.

11. Useful commands while model is running:

check your run progress:

```
>qstat –u <user>
```

If you find an issue and need to delete your run:

```
>qdel <JobID
```

If your model run doesn't complete, try some of the suggestions on the troubleshooting page.

12. To continue a run from restart files, for example after an initial start up, change CONTINUE_RUN to `TRUE` in the `env_run.xml` file.

> **Tip for New Users:**
>
> High performance computing systems often have maximum wall times (e.g. JOB_WALLCLOCK_TIME = 12 hours), meaning a long run will need to be split into several smaller runs. In this case, change the “RESUBMIT” value in env_run.xml file to greater than zero. For example, you can simulate 10 years by changing STOP_OPTION=nyears, STOP_N=1 and RESUBMIT=9. This will perform an initial run of 1 years + (9 resubmits x 1 years per job) = 10 years.