WACCM 3.5.48 Model Output Archiving

From version 3.5.07 to version 3.5.48 of CAM/WACCM a procedural change occurred in the way model output is archived. Below is the documentation of the new method to archive CAM/WACCM output taken from a README file given with the model distribution, /waccm/fvitt/waccm03_cam3_5_48/scripts/ccsm_utils/Tools/archiving/README

Documentation for using the archiving scripts in this directory
---

Archiving of model output should be handled as part of model run post-processing. This takes place in two parts: short-term archiving and long-term archiving. These scripts are described below.

To date, these scripts have only been tested with stand-alone CAM and are not yet recommended for use elsewhere. Eventually we'd like for these scripts to be used for CCSM runs as well. For the moment, the list of supported machines is short, but will be expected to grow with machines and archival systems inside and outside of NCAR. The supported machines are Bangkok, Calgary, Bluevista, Blueice, Tempest, and Lightning.

****
st_archive.sh
****
This is the short-term archive script. Its primary goal is to move model output out of the run directory to a temporary directory (the short-term archive directory) where the long-term archiver can process it later. This clears the run directory for the output of subsequent iterations of a model run. You'll see in the template run scripts (run-ibm.csh, run-lightning.csh etc) that it is called in-line, immediately after the model run. Required arguments are the root of the short-term archive directory (STA_ROOT) and the archiving casename (ARCH_CASE), passed in as environment variables.

STA_SAVE_INTERIM_RFILES
FALSE (default) will only archive restart files from the end of the last model run - any interim restart files generated will be deleted; these will be stored in tar files that contain all the files required for a model restart (just stand-alone CAM for now); this file will be called something like:
<casename>.all.r.yyyy-mm-dd-sssss.tar

Any output messages will appear in the log of the run script.

****
lt_archive.sh
****
This is the long-term archive script. This is where most of the post-processing related to archiving takes place. Although it appears to be called in-line just after the short-term archiver, it actually spawns a batch job by default, so as to not delay a potential model resubmit. It requires STA_ROOT and ARCH_CASE as does the short-term archiver, but also the root of the long-term archive, LTA_ROOT. The following environment variables are available as options:

LTA_INTEGRITY
NORMAL (default) validates by file size, HIGH validates by retrieval and comparison
LTA_ANN_TARRING
TRUE (default) enables the bundling of monthly history files (secondary history files excluded) into yearly tar files
LTA_RETENTION
365 (default) days the file will be retained in the long-term archive
LTA_COS
"rel=ec" (default) requests an economy class of service, meaning only one copy will be saved; note this will likely require some modification when non-NCAR systems are supported
LTA_WPWD
specifies a write password (blank by default) the read password is also blank by default and there is no option for one currently
LTA_PROJ
specifies a project number for storage charges (uses the default assigned to each user by default)

Any output message of lt_archive.sh will appear in the log file of the run script itself, but messages from the spawned job will appear in its own log file, residing in the directory it was invoked from. Also in that directory will be the generated script used for the spawned job, and it'll be named: lta_<casename>_<machine>.sh. Once generated, it can be used to run the long-term archiver manually and interactively, if desired.

The spawned job will return with a non-zero return code if all files made it to the long-term archive without error, and local copies were deleted. Note, that copying to the long-term archive is done synchronously, and verification is done immediately after the copy.

Files remaining in the short-term archive directory after completion of the long-term archive script can be handled in the following ways. Users can wait for a subsequent iteration of the run script, and the files will be processed with the next batch of output files. Or, users can manually invoke the long-term archive script (described elsewhere in this document) which is recommended when no more model runs are required. The exception to this comes with "partial" files. These are tar files created during bundling of history files into yearly tar files when output is incomplete for a particular year. Normally, they just wait in the short-term archive directory until the remaining files for that year arrive, but if the model run ends mid-year and no subsequent iterations are expected, these partial files will require manual processing for them to reach the long-term archive.

<<<<more documentation likely to follow