2019-12-05

Yannick opened the meeting announcing, due to the holidays and the AMS annual meeting, that the schedule for these weekly meetings will be somewhat interrupted. We will take two weeks off during the holidays (Dec 26th and Jan 2nd), and take off the week of AMS (Jan 16th). We then went around a collected updates from the group.

EMC/GMAO

Dan has a PR submitted in fv3-jedi that enables FV3-GEOS support for CRTM. Up to this point, FV3-GEOS did not natively provide all of the fields CRTM requires so Dan has added the necessary conversion and interpolation to fill in the missing data.

Dan is enabling using FV3-GEOS in the real-time H(x) flow. He reported issues with sensitivities to the order of variables, issues with convergence, and is working to get these resolved.

Hamideh reported some confusion with ioda testing where the input AMSU-A netcdf file has nrecs set to 1, but the test passes when the expected value for nrecs is set to 100 in the YAML configuration. Steve H noted that the nrecs specification is no longer needed in the file and can be deleted, and the current testing should be ignoring the nrecs value in the file. (Hamideh and Steve followed up on this discussion after the meeting and determined that the test is operating properly. The ObsSpace constructor is now responsible for determining the sets of locations that comprise records and therefore is also responsible for setting nrecs. This means that the specification of nrecs in the input file is now obsolete.)

Ryan has a PR in ufo-bundle under review that adds the capability to report build and test results in CDash. CDash is a very nice web interface that displays build/test results and enables easy navigation through those results.

Sergey will submit a PR in soca soon that adds a grid iterator function of which will enable LETKF functionality. Sergey would like access to the FV3 repos so he can add the grid iterator to FV3. Mark M will help him get the access to the FV3 repos. Travis noted that while the grid iterator will enable LETKF, there is still a significant amount of work that needs to be done to attain scientific validity. JJ asked how much work it is get LETKF working (in MPAS). It was noted that you need to add a routine that processes localized grids, via the grid iterator, and then the LETKF module should be able to handle the rest.

Met Office

Marek introduced a new hire, Oliver Lomax, who will be working on the JEDI interface for the UM model.

Boulder

Mark M asked if all models have moved off of the old Fortran variables module in oops since he has a PR under review that removes this module. Travis responded that SOCA is seeing some issues with the new module where Fortran access is not properly synchronized with an object created in C++. He is actively debugging this issue. All other models are ready.

Mark M has prepared new Docker, Singularity and CharlieCloud containers that contain updates as well as a new clang-mpich configuration. The containers are now stored on the Sylabs Cloud, which is managed by Singularity. This means that the source has moved when doing a pull of a new image. Instructions are provided in the JEDI discussion group and a PR is under review that will add these instructions to jedi-docs. Mark can help with questions about the new containers.

Clementine has merged in new ensemble applications which have broken existing enshofx and ensforecast tests. Instructions on how to repair your tests are posted in the JEDI models discussion group. Clementine can help with questions about the new ensemble applications.

Maryam has added Clang compilers to the AWS CodeBuild testing. A few repositories are successfully using the new Clang testing now, and more will updated in the near future.

Maryam also has PRs in ioda and saber under review that move their test data to Amazon S3. These PRs include new tests that download the test data if it is not already present on the local machine. Maryam noted that this scheme is not necessarily tied to S3, but that is where we decided to start. These PRs need to be coordinated with recent updates to ecbuild. The required ecbuild updates have already been installed in the containers.

Mark O and Mark M added that there are more pending ecbuild updates needed for purposes other than using test data stored on S3 that include fixing the "hidden symbol" issue, fixing the netcdf find script, and enabling the use of srun for submitting MPI jobs. Yannick advised everyone to submit PRs directly to the EMCWF ecbuild repository for these issues.

Xin is working on including additional data in the variational bias correction that is required for QC checks. This work came from discussions with Emily. One example is the scan angle bias correction term (i.e., scan angle multiplied by a b.c. coefficient).

Chris S asked for updates on ODC and on the real time H(x) flow.

Steve H reported that he will submit two more PRs to ioda for ODC. The first will contain the ODC reader, and the second the ODC writer. Getting the reader in first will enable the collection of timing statistics which can be done in parallel with the development of the ODC writer. Chris expressed interest in having Steve V help with the ODC writer, which offers advantages in having more than one person be familiar with the IDOA IO interface and allowing for the creation of converters between netcdf and ODB. Having Steve V help with the writer will be the plan moving forward.

Yannick reported that the fv3-bundle GFS H(x) test is spending ~90% of it's time writing the netcdf file at the end of the test. This is being seen on multiple hardware platforms and with 6 MPI tasks. It was decided to discuss this further offline from the meeting.

Yannick also reported that the real time H(x) work will ramp up as soon as the workflow is in place. This will probably occur in January, 2019. NCAR responded that they are working on enabling MPAS for the workflow, and asked if GFS background data will be available. The plan is to provide background data to cover the entire month of July 2019. Mark O mentioned that GEOS background data is available, and GFS will be available next. It was noted that we are running into cost limitations with S3 transfers of data. Yannick mentioned that the plan is to place copies of the full July 2019 data on each HPC system so that costly S3 data transfers can be avoided. A copy of the GEOS and GFS backgrounds, and obs data will be available on S3 for use by jobs running on the AWS EC2 instances (which won't incur data transfer cost). Then the data will be downloaded (once) to Discover, and then distributed to the the other HPC systems from Discover. Yannick announced that if there are other models that want to participate in the real time H(x) flow, the developers should let us know and we'll work out a way to provide access to the July 2019 data.

Hailing is running comparisons between JEDI and GSI H(x) using October 2019 COSMIC 2 data. She should have some results soon that can be shown to the group.

Junmei is also verifying JEDI flows and has run into an issue where JEDI, using MPAS, hangs on Cheyenne when utilizing more than one compute node. Running on a single node successfully completes. It was noted that turning off IODA writing for a one cycle experiment made the hang go away, but more analysis needs to be done before the root cause can be determined.

NRL

Nancy reported that they are running into an issue with their ioda converter where some observations are being dropped during the conversion. Some discussion ensued and it was decided to send Steve H a test case exhibiting the symptom so he can help debug the issue.