

2021-07-29

Yannick opened the meeting with only one general announcement: There will be no meeting next Thursday because many people are on vacation. There will be a meeting on Thursday, August 12.

JEDI1

Mark M proceeded with a JEDI 1 update:

The main news of the week is that the UCAR DASH data server has been down for UCAR NWSC maintenance. It went down on Monday and came back up yesterday. Most of the participants in this meeting work with the develop and feature branches of JEDI test data, which are hosted on git-lfs and unaffected by DASH downtime. However, most bundles use a release tag of crtm and this data *is* hosted on DASH so fresh downloads did not work this week. Maryam is working on an option to bypass DASH and retrieve the crtm coefficients from AWS S3 when DASH is down. This will also fix the duplication of crtm coefficients in current bundles.

Maryam is also working on an enhancement to the CI to allow authors of pull requests to see build warnings, failures, and errors on the public CDASH site. Previously only runtime test status was presented. This is currently only implemented in saber. She has implemented a beta version in oops but is concerned because including the build logs in the cdash output greatly increases the build time. She is investigating the causes.

IODA post-release: [We're refactoring ioda to more fully integrate ioda-engines](<https://github.com/JCSDA-internal/ioda/issues/351>). This starts with the unification of ObsSpace and ObsGroup, the development of a Fortran API, and the replacement of get_db and put_db calls with more direct ioda-engines access. Then it will require changes in the variable names (not appending the channel number) and yaml files (ioda v2 directory-style naming conventions).

Datetime optimization: processing datetimes as strings for each observation is time-consuming and appears to be contributing to the [ObsSpace IO inefficiency that has been discussed in previous meetings](<https://github.com/JCSDA-internal/ioda/issues/252>). We agree that the best way forward is to express the date and time as integers instead of strings, possibly with additional metadata so users can see the time interval covered in a readable format when they look at the file contents with, e.g. ncdump. Steve is starting to look into this. When completed, this will require changes to the ioda files. But, we can coordinate these file changes with the file changes that will be needed when the effort to standardize naming conventions is completed. Ideally, both will be implemented in a future release.

We're considering a design question now: how to make auxiliary text/yaml files for ioda available for distribution. This has come up in the context of a yaml file from the UKMO to map ODB variable numbers to ioda variable names. But, it is a broader question; if anyone has ideas, [we'd like to hear them](<https://github.com/JCSDA-internal/ioda/issues/385>).

Work is still underway to get fv3-bundle to work with PGI/NVIDIA compilers. Because of slow queues, debugging on Summit is extremely inefficient. So, Mark M created an AWS instance and installed the latest PGI (now NVIDIA SDK) compilers there. This has accelerated progress: most tests are passing for oops, saber, ioda, crtm, fms, and fv3-jedi, except for some I95 and lam-cmaq tests. There are still about 45 test failures in ufo, however, that likely need to be fixed before we can run benchmarks. The hope is that once things work on AWS with PGI, they will also work on Summit.

Maryam has been making good progress on running MPAS with ewok. The plan is to get this running on Orion first because we know the ewok /r2d2 infrastructure is working well there. Then she will try on Cheyenne.

Maryam is working on adding error messages from the build step on CDASH. This is implemented in the Saber repo, and more repos will be updated soon.

In other news, JCSDA was just granted \$10k in resources for Microsoft Azure cloud. This is in part intended for benchmarking and performance comparisons with other platforms, but is also for exploration: to see what Azure has to offer in performance, cost, and diversity of services compared to AWS.

Orion, Hera, Discover, and S4 have been updated with a new version of pybind11 (used for the ioda python API). We'll get to Cheyenne next.

Also: this just in: our proposal for a JEDI short course (half-day) was just approved for the next AMS meeting.

There were a few questions after this report. David S pointed out that the UKMO has recently announced plans with Azure to provide computing resources and he wondered if the JCSDA Azure allocation might benefit this. Mark responded that Microsoft is very interesting in return on investments and one of the conditions of this award is to provide public machine images and configuration files to allow Azure users to launch a cluster and run JEDI at their own expense. These resources might indeed be leveraged for use at the UKMO.

David S also asked about the issue of making auxiliary text/yaml files available in ioda and there was some discussion. One approach would be to have a collection of similar ODB mapping files in ioda for different centers, as needed. Another approach would be to distribute a generic ODB mapping file with ioda and then have centers easily override this with their own customizations.

Steve V asked "I'm curious, why aren't there other classes/types we could use for representing date/time besides ints or strings?". Steve H responded that we do use C++ datetime classes to represent strings in memory. This has to do with how they are represented in the files. The refactoring we are describing will not change the representation of datetimes in the code.

JEDI2

JEDI 2 update given by Dan

Jedi 2.3 - Use of NUOPC driver with FV3-JEDI (Mark Potts)

JEDI can run the forecast and return some states in memory.

Model outputs in stand alone are different from what the model outputs to disk.

PR soon to develop of FV3-JEDI. Changes are in UFS develop.

Jedi 2.8 - Regional DA (Ting, Ming)

Comparison of GSI and JEDI.

Added convert state from cube to lonlat. Still some issues to resolve.

Differences in max increment location between GSI and FV3-JEDI.

Jedi 2.9 - MPAS general updates (JJ, BJ, Maryam)

Cycling some regional runs with MPAS.

Working on MPAS-EWOK.

Jedi 2.10 - LFRic general updates (Kristin, Steve, Steve)

PR for 4D h(x)

Jedi 2.11 - UM general updates (Kristin, Steve, Steve)

PR for 4D h(x)

Jedi 2.12 - Neptune general updates (Sarah, Nancy)

Evolving model at the same time as the DA. Updating to Neptune 1.2.

Working on the faster h(x) PRs.

Next up in IODA v2, after Neptune 1.2.

New developers coming on board.

A lot of work with the Cylc suite for cycling.

Adding ctests.

Alex R wants to update the build system. Use sub modules instead of a bundle approach.

Jedi 2.14 - Cubed sphere grid into Atlas (Dan, Marek, Olly, Steve V)

Steve V has created a cell centered version of the cube sphere. Some tidying required to have PR to ECMWF.

Generalized the tile ordering.

Marek added periodicity function for cube.

LFRic wants a cell based partition. Layout the cells and add nodes around the cells instead.

Jedi 2.15 - VADER (Steve V, Dan)

PR in VADER.



JEDI3

JEDI3 update from Anna

JEDI-Algorithms ran a two-day code sprint on OOPS interface classes (July 27-28), resulting in multiple PRs refactoring and documenting OOPS interface classes. Thank you to all the participants for contributing.



JEDI4

JEDI 4 update from Yannick

- Good progress in adopting different models in the H(x) suite, and with porting the workflow to different HPC platforms
- Currently adding user documentation for the workflow
- We are preparing for coupled DA flows by adding the capability to execute with multiple models in one run.
- Rahul is making progress in having EWOK run an FV3 forecast



OBS

OBS report from Ben

OBS:

Increasingly critical need for time-interpolated hourly 4D H(x) capability to finish UFO acceptance requirements with EMC. Cory Martin/Hui Shao can provide necessary requirements.



OBS1

OBS1 report from Hui

- Bug fixes to UFO: tier 2 UFO tests (obs files regenerated using upgrader); FPE for windspeed, variable transform, averaging kernel
- New PR: SSMIS filter, obs bound check updates, using parameter classes, pressure-height conversion
- JEDI_GDAS validations: focusing on conventional data currently, not expected to fully replicate GSI in terms of obs errors but try to get close enough QCs with reasonable results; also working with EMC to update GSI obs files for R2D2 (with additional obs files/matching files)
- BUFR converters are in progress; a few obs types are not using GSI obs and testing is ongoing

Chris S asked about a comment Hui made that there are some aspects of GDAS that JEDI will not duplicate and he asked which ones. Some discussion ensued that some aspects of the pre-QC and obs inflation will not be exactly the same as GDAS. Some qc algorithms from the UKMO will be used in lieu of those from EMC. The focus will be on validation rather than replication. But Greg said they are doing as many things as similar as possible. Rahul added that the inflation factor is not a unique number - it can depend on the grid (e.g. cubed sphere or GSI's gaussian grid) so it would be hard to reproduce exactly.

From the chat:

Zhiqian Liu 9:37 AM

both GSI and UKMO's DA system have Variational QC, is there ongoing work for VarQC?

Greg Thompson 9:40 AM

@ChrisSnyder: I met with JJ last week about QC steps now part of GDAS. Hopefully he and I can make some tests for use within MPAS soon.

Sergey asked if we are still working with GSI obs or does everything go through the converters. Hui and Greg answered that converters are used in many places but work is still proceeding on the prep bufr and WMO converters.

Ron McClaren

Sorry didn't chime in directly. The work is doing with NCEPLIB-bufr/ioda-converter is a little involved, but I'm seeing the light at the end of the tunnel.

Basically I'm extending the NCEPLIB-bufr with support for a query syntax to make it sane to get arbitrary data elements out of a BUFR file. Take a look at the README in <https://github.com/JCSDA-internal/ioda-converters/tree/feature/query/src/bufr> if you'd like a preview for whats coming. Code works, but lots of stuff to do before merge.

Chris Snyder 9:53 AM

@GregT: I'd be interested to see a check on how JEDI-processed conventional obs perform in cycling for an extended period. We could compare against results using conventional obs from GSI ncdiag files. Based on discussion here, I think the idea is that the JEDI-processed obs should give statistically indistinguishable (or perhaps better!) results.



OBS2

Francois then reported on OBS 2

Work continues on the diagnostics but Rachel is out this week. Travis & JJ agreed to help develop the diagnostics tools and in particular to help make the r2d2 data transfer into sagemaker studio more efficient.

Sergey asked for more information about how the diagnostics will be coordinated and some discussion ensued. Francois offered to give access for Sergey to the AWS diagnostic tools. Also, the broader perspective with Travis and JJ on the team will help produce a more generic diagnostics tool. They proposed to continue collaborating on this for some time and then discuss their approach in a JEDI focused topic meeting at some point in the future. Cory added that the framework should be co-developed and shared but many centers will have their own customized tools.



OBS3

Ryan reported on OBS 3

Greg, David, and Ryan are working on the IODA Conventions Document, and a draft will be out relatively soon. This is needed to kick off the ioda converter upgrades, as developers should know how IODA ObsSpaces are organized in the IODA-v2 format. This document will be a living document, and we will update it as needed.

We have a special topic discussion on QC and our plans for the future, and this is scheduled for August 19. To prepare for this, David S. has worked on a 10-page draft document, entitled QC Flags: A New Hope. We hope to work through this and see how we can fit this into the OOPS and UFO infrastructure.

We do have one general oops PR of interest for everyone: when we get a floating point exception in our tests, we can now report the line number where this exception occurred. This will save substantial debugging time for UFO developers.

Besides that, we have the usual enhancements to filters and operators, and continued improvements to the BUFR and ODB-reading codes.



CRTM

CRTM update from Ben

v2.4.1 release delayed until mid-August, although a release date was never set, I was hoping for the end of July. A transmittance coefficient generation roundtable meeting is set for early August, enabling expert users to coordinate on transmittance coefficient development for CRTM.

Patrick is developing a way to enable partners to contribute their own coefficient. Let him know if you are interested.



SOCA

SOCA update from Guillaume

Min is working on a coupled ocean & sea ice static B

Travis is working with JJ and Rachel on diagnostics as noted above

work proceeding on LETKF