SAWG Briefing to UFS-SC: The Graduate Student Test

30 March 2018

System Architecture Working Group

About the SAWG

- **The SAWG is a SIP working group to define and implement a vision for a community-based unified forecast system**
- **Initiated in October, 2016**
- **Members**

Jim Kinter, GMU/COLA
Cecelia DeLuca, CIRES/ESRL
Thomas Auligné, JCSDA
V. Balaji, Princeton University/GFDL
Rusty Benson, GFDL
Ligia Bernardet, CIRES/ESRL
Arun Chawla, NCEP EMC
Tony Craig, NESII and NCAR
Arlindo DaSilva, NASA GMAO
John Derber, NCEP EMC
Jim Doyle, NRL
Jean-Francois Lamarque, NCAR CGD
Mark Iredell, NCEP EMC
John Michalakes, NCAR/NRL
Phil Rasch, PNNL
Suranjana Saha, NCEP EMC
Vijay Tallapragada, NCEP EMC
Gerhard Theurich, NRL
Sam Trahan, NCEP EMC
Mariana Vertenstein, NCAR CGD
Jun Wang, NCEP EMC
Michael Farrar, NCEP EMC

SAWG Charge

To provide the NCEP/EMC Director with prioritized recommendations for the advancement of a system architecture that meets operational needs and enables and encourages collaboration with external model development partners and the broader research community.

https://esgf.esrl.noaa.gov/projects/sawg/

The Graduate Student Test

The Graduate Student Test generates requirements for the UFS:

- *Get code*.
- *Run code*.
- *Change code*.
- *Test code*.
- *Evaluate code*.
- *Transition code*.
- *Train students in using code*.

The Graduate Student Test

Get code.

- Easily identify most appropriate code for a given application and set of numerical experiments, including which options are available
- Easily distinguish between versions of code (capabilities, readiness, limitations) and identify an appropriate version
- Easily obtain code
- Easily

The Graduate Student Test

Run code.

- Easily obtain workflow (script) for given experimental setup
- Understand and access setups with:
  - Active or passive (data) components
  - Cold-start or DA-cycling runs
- Access code on HPC system(s) available to the public

The Graduate Student Test

Change code.

- Alternative or modified parameterizations
- Alternative or modified components (models)
- Alternative or modified coupling strategies

The Graduate Student Test

Test code.

- Have access to standard unit/system tests
- Have access to functional tests
- Easily obtain test data sets

The Graduate Student Test

Evaluate code.

- Easily obtain and use standard diagnostics
  - General behavior (climatologies, error statistics, etc.)
  - Individual processes (process-oriented metrics)

The Graduate Student Test

Transition code.

- Clear definition of pathway for R2O transition
- “Rules of the road” (governance)

- What about evolving nature of public releases and operational codes?
- Who will take responsibility? … Not necessarily the responsibility of the graduate student

The Graduate Student Test

Train students in using code.

- A course or mini-curriculum, possibly online, on how to use ESMF/NUOPC/CIME/CMEPS codes and workflows

- This can be developed by a university or consortium of universities

The Graduate Student Test

Overarching these requirements are requirements for governance:

- Software engineering best practices and governance procedures
- Facilitate free flow of information
- Facilitate feedback between GST-class users and sophisticated super-users

… and user support

- Users' guide
- Documentation
- Help desk

Scenario

- Consider the recent weather extremes:
  - Cold air outbreaks – North America and Europe (snow in Rome)
  - Negative NAO values
  - Atlantic blocking
  - Explosive cyclogenesis off the East Coast of North America (4 nor'easters in March 2018)
  - Sudden stratospheric warming (SSW)
    - Total breakdown of polar night jet - wave #2 pattern in second week of February
    - ~ 2-3 weeks prior to the beginning of sharply different winter weather in NA and Europe

Scenario - 2

- Hypothesis: SSW cause surface weather features that followed it, which is pretty well established statistically.
  - Leaving aside the possibility that tropical convective activity about a week earlier (precursor of Cyclone Gita) was responsible for the SSW
- Suppose a graduate student (GS) wishes to test the hypothesis that the SSW-surface-weather relationship enhances sub-seasonal predictability, specifically by evaluating the predictability in a sophisticated forecast model of such a sequence of events.

Scenario - 3

1. GS gathers data from the past to see how often (large) SSWs are followed by wave #2 in the stratosphere, cold air outbreaks near the surface in the Northern Hemisphere, mid-tropospheric North Atlantic blocking, (multiple) nor'easters etc.
2. GS gathers data on forecasts at 1, 2, 3, and 4 weeks lead time.
    - GS must make a choice of model, so why not use the public release UFS?
    - Past operational forecasts might not be suitable, because the operational model changes over time, so re-forecasts have to be used.
3. Requirement: large ensemble re-forecast experiment (LERF)
4. GS evaluates prediction skill in LERF and uses the ensemble members to evaluate the predictability
    - Target: stratosphere-troposphere sequence described above

Scenario - 4

- GS finds only modest, barely significant skill in the predictions, but the ensemble analysis suggests that there is a lot more predictability
- Need to investigate:
  - UFS unable to represent the necessary physics and dynamics?
  - UFS is over-confident?
- Possible experiment: Sensitivity of results to spatial resolution
  - Public release includes versions with 13, 25 and 50 km grid spacing
  - GS needs training in how to run experiments with UFS

Scenario - 5

- Possible experiment: Does detecting SSW influence require very high vertical resolution in the upper troposphere and lower stratosphere?
  - Can’t be done with the public release with single setting - 64 levels
- GS changes code to have a lot more vertical levels, bunched if possible near the tropopause.
  - GS needs training in how to change vertical structure in the atmosphere

- Possible experiment: Suppose process-oriented diagnosis suggests that gravity wave interaction with large-scale flow is poorly represented
- GS tests whether or not the gravity wave drag (GWD) parameterization can be improved, e.g. using a new and improved version developed elsewhere
  - GS needs training in how to modify the code to swap in alternative parameterization of GWD

Scenario - 6

- GS finds sensitivity in prediction skill to vertical resolution and/or to GWD
- GS shows that increasing resolution and using the alternative GWD scheme makes the 4-week forecasts of nor'easters more accurate and reliable
- EMC needs to be shown that the idea works and moves the needle …
  - May not be the GS' job: R2O transition not needed to publish and graduate
  - New version must be subjected to a standard functional test harness

Scenario - Notes

- The whole process has taken 2-3 years (typical GS intensive research time frame)
  - EMC has moved on to a new version (maybe even several implementations in 3 years),
  - Repeat tests with latest operational version

- GS wasn't asked to improve UFS - unlikely to be enough for a dissertation (or publishable).
- GS (w/advisor) simply made choice to use UFS at the moment when it was clear that a good forecast model would be needed to test the dissertation hypothesis.
- If the UFS is in shape for a GS to use it, that choice can be facilitated, and R2O transition of the results of a PhD student's model development labor can be effected

Scenario - Notes

- Scenario success depends on:
  - How well things are set up – code, data, workflows, documentation, testing, governance etc.
  - Training and documentation
- Failure mode:
  - UFS not in shape for GS use → UFS won't be natural choice of model
  - Results obtained with model other than UFS → no smooth R2O transition
  - Relevant findings in GS' publications must be re-invented, presumably by EMC