Results of Phase 1 Dynamical Core Testing for the NOAA Next-Generation Global Prediction System: Executive Summary

Jeffrey S. Whitaker
NOAA Earth System Research Lab
<jeffrey.s.whitaker@noaa.gov>

John Michalakes
NOAA National Centers for Environmental Prediction
<john.michalakes@noaa.gov>

Motivation

In order to explicitly simulate moist convection on the global scale, and take advantage of advances in high performance computing, NOAA is evaluating candidate non-hydrostatic dynamical cores with a battery of tests to culminate in the selection of a new dynamical core for the Next-Generation Global Prediction System (NGGPS). The current NOAA operational global forecast model is spectral and hydrostatic; therefore, it cannot explicitly simulate moist convection and may not be able to scale up to take advantage of peta- and exa-scale high-performance computing (HPC) systems. The first round of tests for solution accuracy and computational performance and scaling has concluded. This is a high-level summary of the results – more details are included in the three attached reports.

Candidate dynamical cores

The five candidate dynamical cores are listed below, with sponsors in parentheses.

- FV3 (GFDL) – Cubed sphere grid, finite-volume discretization (non-hydrostatic version of the hydrostatic core described in Lin (2004) and Donner et al (2011)).
- MPAS (NCAR) – Unstructured grid, finite-volume discretization with C-grid variable staggering (Skamarock et al, 2012).
- NEPTUNE (NRL) – Cubed sphere or icosahedral grid using a spectral element discretization with the Non-hydrostatic Unified Model of the Atmosphere (NUMA) core (Giraldo et al. 2014).
- NIM (ESRL) – Non-hydrostatic Icosahedral Model (unstaggered finite-volume A-grid implementation).
- NMMUJ (NCEP) – Finite-difference cubed-sphere grid version of the B-grid lat/lon mesh core used for operational regional modeling and described in Janjic and Gall (2012). The construction of the 'uniform Jacobian' cubed sphere grid is described in NCEP Office Note 467, available at http://www.lib.ncep.noaa.gov/ncepofficenotes/2010s/.

Idealized test case results

Three sets of idealized tests were run: a baroclinic wave on the full sphere, orographic mountain waves on a reduced radius sphere, and an idealized supercell thunderstorm on a reduced radius sphere. FV3 and MPAS produced very similar solutions for all the tests. FV3 did a slightly better job at representing the discontinuous nature of the front embedded in the baroclinic wave, and MPAS did a slightly better job of representing the details of downdrafts associated with the supercell thunderstorm. NEPTUNE produced different orographic mountain wave solutions than the other models, due to the fact that it does not make the shallow atmosphere approximation. NEPTUNE's baroclinic wave solution was more damped than MPAS and FV3, and it was unable to fully capture the splitting of the convective updrafts by 90 minutes into the solution at 4-km resolution. NMMUJ's baroclinic wave solution was heavily damped near the surface, and it was unable to produce realistic solutions for either the orographic mountain wave or the supercell test cases. NIM produced stronger 4Δx oscillations than the other models behind the front in the baroclinic wave test, but its orographic mountain wave and supercell thunderstorm solutions were qualitatively similar to MPAS and FV3. However, NIM (and NEPTUNE) required a much shorter time-step than FV3 or MPAS to generate split updrafts in the supercell test case.

3-km, three-day full physics simulations

To 'stress-test' the dynamical cores (and the development teams), three-day forecasts were run for two select cases at global cloud-permitting resolution (~ 3 km) with full physics, high resolution orography and initial conditions derived from an operational data assimilation system. NEPTUNE did not submit results for this test. NIM produced highly damped solutions for both cases, with very little variability at scales less than about 50 km. MPAS, NMMUJ and FV3 produced a shallowing of the kinetic energy spectra consistent with observations and three-dimensional turbulence theory. NMMUJ produced smoother and weaker vertical circulations than MPAS or FV3. The reasons for this are not clear but we hypothesize that it may be due to divergence damping and/or the presence of parameterized deep convection (which was absent or suppressed in MPAS and FV3). NMMUJ, FV3 and MPAS all were able to represent some observed aspects of severe convection during a Great Plains tornado outbreak. MPAS and FV3 both simulated realistic fine-scale detail within tropical cyclones. NMMUJ simulated realistic detail in the rotational wind field, but the vertical velocity field within the tropical cyclone vortices was weak and noisy at the grid scale. MPAS, NIM and FV3 conserved dry mass, NMMUJ did not. NMMUJ produced increasing levels of global precipitation and total water vapor during the forecasts, and NIM exhibited high levels of external gravity wave activity.

Benchmark results

The Advanced Computing Evaluation Committee (AVEC) evaluated performance and scalability of the five candidate models. Benchmarks based on the baroclinic wave test case with 10 added passive tracers were run at 13 and 3-km during two sessions of dedicated access to Edison, a 130-thousand processor supercomputer at the U.S. Department of Energy's

National Energy Research Scientific Computing Center (NERSC). Since NCEP did not provide a version of the Global Forecast System (GFS) model for testing, the hydrostatic ECMWF Integrated Forecast System (IFS) model was included as a proxy. Each candidate model used the same formulation and configuration as was used to run the aforementioned idealized test cases. Only one of the candidate models (NIM), was ready for testing on novel processor architectures such as NVIDIA's Graphics Processing Units (GPU) and Intel's Many Integrated Core (MIC) architecture. Therefore, testing was done on a supercomputer using a close-to-latest version of a conventional multi-core processor.

Performance was measured as the number of processors needed to reach an operational speed requirement – using a workload sized to represent global forecast domains today and in the near future. The rankings, from fastest to slowest were: IFS, NMM-UJ, FV3, NIM, MPAS, and NEPTUNE. MPAS and NEPTUNE were significantly more expensive than the other candidate models.

Scalability was measured as strong scaling efficiency running on high (greater than 100-thousand) numbers of processor cores with workloads planned to be in operational use in ten years. The rankings by scalability were essentially the reverse of the performance rankings: NEPTUNE scaled the best followed by MPAS, NIM, FV3 and NMM-UJ (tied), and IFS.

Summary

FV3 and MPAS produced solutions that were generally of higher quality than the other models, although FV3 was markedly faster than MPAS in the performance tests.

NEPTUNE, the only candidate to use spectral-element numerics, was the slowest model in the performance tests and was unable to perform the full-physics, real-data 3-km global simulations.

NIM produced acceptable idealized test solutions, although it took several iterations by the development team. The code is relatively fast, only slightly slower than FV3 in the performance tests. The real-data forecasts were highly damped, and used very smooth orography. Like MPAS, it was configured to use an icosahedral mesh for the tests. However, the MPAS mesh is considerably more flexible. There has been significant progress on optimizing the NIM dynamical core on novel processor architectures, although we expect much of this work can be transferred to other candidate dynamical cores.

Although NMMUJ was the fastest non-hydrostatic model in the performance tests, the solution quality in both the idealized and real-data simulation tests was substantially below that of the other models. It was learned, subsequent to completion of the AVEC benchmark report, that a shorter time step was used to run NMM-UJ for the 3 km full-physics stress test than was used for the idealized AVEC benchmark workloads, suggesting NMMUJ computational performance may have been overstated in the AVEC report by that factor (16.7%). The ranking of the models by order of performance does not change, however.

In its consensus report, AVEC noted the emergence of performance and scaling as competing value objectives with respect to solution accuracy, depending on whether one viewed future computing resources for NCEP operations in terms of scarcity or abundance. One view emphasizes streamlined numerical formulations that do the most with limited computing resources. The other view argues for best-possible numerical formulations that meet operational speed requirements by scaling with advances in HPC that are largely the result of increased parallelism. Although additional work is needed for NGGPS to fully exploit these next-generation systems, we believe growth in computing capability and improvements in model efficiency will be sufficient to support an NGGPS down-selection based primarily on solution quality.

Mesh refinement capabilities were not exercised in the first round of tests. However, this is an important consideration since globally uniform grids at cloud-permitting resolutions may not be operationally feasible for a long time. The MPAS mesh may be refined seamlessly (but only statically) anywhere in the domain. FV3 allows for grid refinement with a combination of nesting and grid stretching. The one and two-way nesting capabilities of the NMMB lat/lon dynamical core have not yet been ported to the non-orthogonal cubed-sphere grid in NMMUJ. NIM does not currently support nesting or mesh refinement but the developer states that there are plans to implement static (non-moving) nesting. Tests utilizing mesh refinement will be a part of Phase 2 testing.

References

Donner L. J, and coauthors, 2011: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *J. Climate*, **24**, 3484–3519. doi: http://dx.doi.org/10.1175/2011JCLI3955.1

Giraldo, F., Kelly, J., and E. Constantinescu, E. M., 2013: Implicit-explicit formulations of a three-dimensional non-hydrostatic unified model of the atmosphere (NUMA). *SIAM Journal on Scientific Computing*, **35(5)**, B1162-B1194. doi: 10.1137/120876034

Janjic, Z., and R.L. Gall, 2012: Scientific documentation of the NCEP non-hydrostatic multistage model on the B grid (NMMB). Part 1 Dynamics. NCAR Technical Note NCAR/TN-489+STR, doi: 10.5065/D6WH2MZX.

Lin, S.-J., 2004: A vertically Lagrangian Finite-Volume Dynamical Core for Global Models. *Mon. Wea. Rev*., **132**, 2293-2307. doi: http://dx.doi.org/10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2

Skamarock, W, M. Duda, L. Fowler, S.-.H Park and T. Ringler, 2012: A Multi-scale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tesselations and C-Grid Staggering. *Mon. Wea. Rev.*, **240**, 3090-3105, doi: http://dx.doi.org/10.1175/MWR-D-11-00215.1