Operational

Building Global, Kilometer-Scale Prediction Models

Mark Govett

NOAA Global Systems Laboratory

ECMWF HPC Workshop

September 20, 2021

Introduction

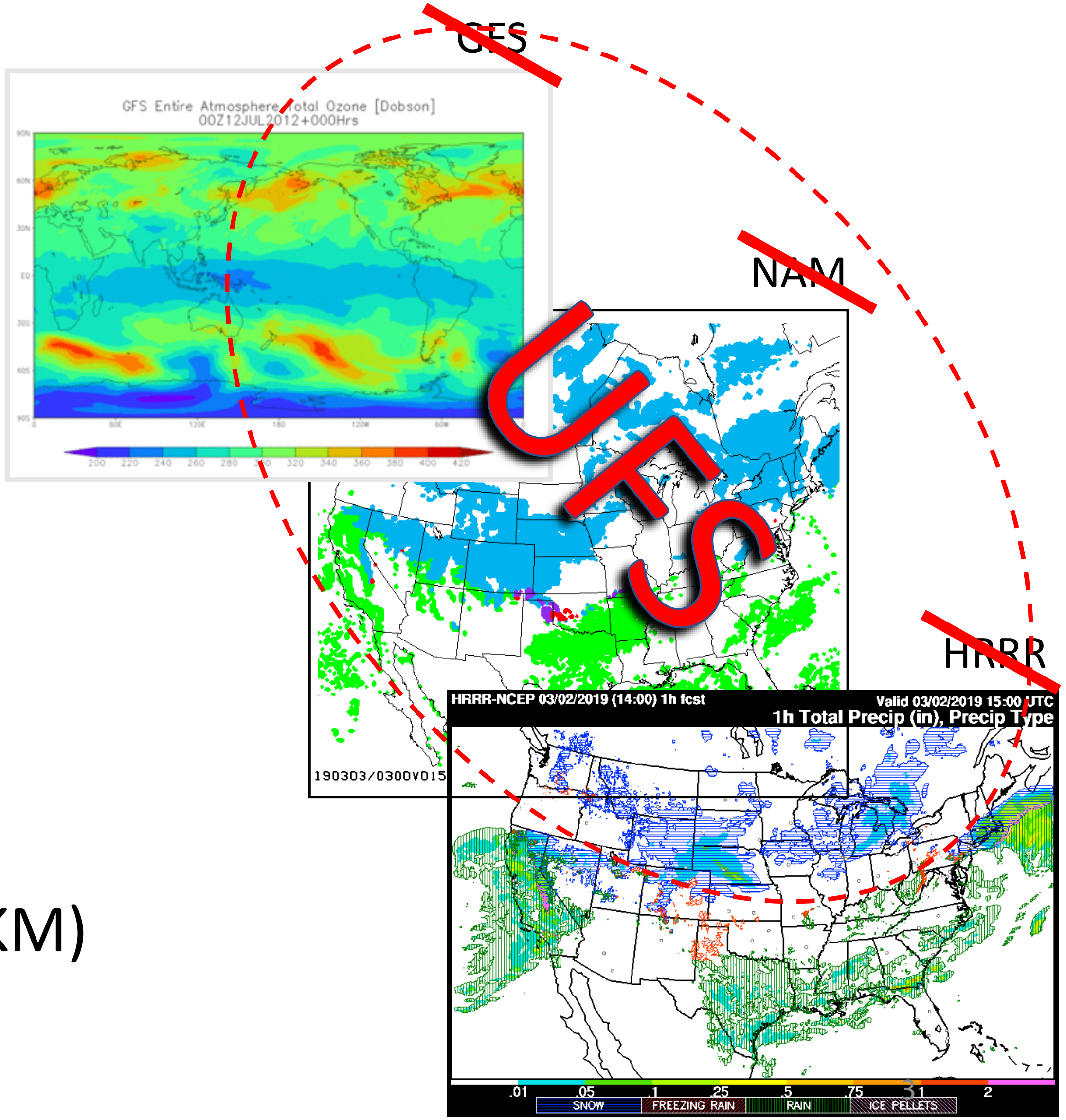
- Climate and weather models becoming increasingly similar
  - Weather – adding ocean, chemistry, hydrologic, ice, land   0 hours – 100 days
  - Climate – incorporating increasingly fine scale processes   100 days – 1000 years
- Kilometer-scale models may be an intersection point
  - Computing: diversity, uncertainty, cost, energy
  - Data handling: storage, analysis, visualization, distribution
  - Software: design, development, developers, collaboration
  - Science: processes, algorithms, accuracy
  - Society: risk, adaption, investments, planning

  - Time-to solution
    - Weather: 8 minutes / forecast day
    - Climate: 1 – 5 Simulated Years Per Day (SYPD)

Model Domain – Global, Regional, Local

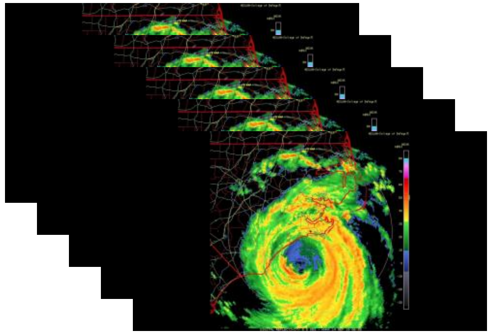
Higher resolution *means* **smaller domain** *and* **shorter forecasts**

- <u>Global</u>: Global Forecast System (GFS) (28 KM)
  - Weeks: 0 - 16 day forecasts, 4x / day
- <u>Regional</u>: North American Model (NAM) (12KM)
  - Days: 84 hours, 4x/day
- <u>Regional</u>: High Resolution Rapid Refresh (3KM)
  - Hours: 36 hours, 24x/day



Improved Prediction Models

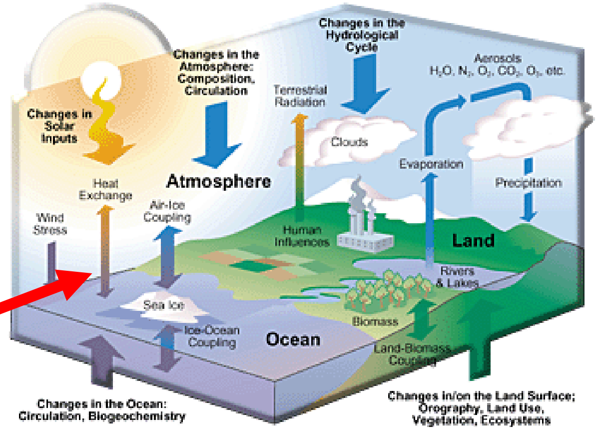
10-100s of members



Ensembles

Model complexity

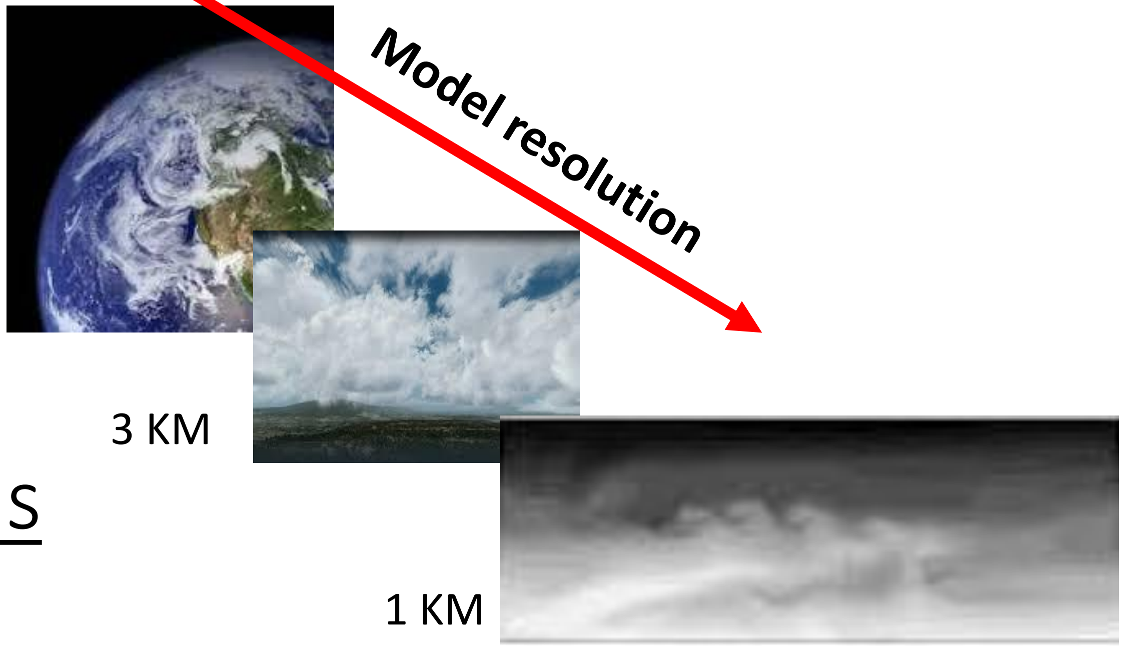
Global Climate System Components



is a tradeoff between

- Computing
- Accuracy
- Time-to-solution

Model resolution



13 KM

3 KM

1 KM

Alternatives to Uniform Grids

- nesting, stretch grids
- reduced precision

Outline

- National and International Efforts
  - WMO Activities
    - Exascale and Data Handling Concept Note
  - ICAMS HPC WG
    - Under the subcommittee on Earth System Modeling and Prediction
- Development Strategy

WMO Activities: Concept Notes

- Science for Services
- Advancing Earth System Modeling
- Advancing Earth System Observations
- Innovation in the Regions
- Exascale Computing, Data Handling and AI

Develop a Concept Note in the context of the numerical prediction value chain to inform future directions in these areas throughout WMO programs

Two sub-teams – bi-weekly meeting since March 2020

- **Exascale Computing and Data Handling** producer perspective
    - **Released for Comment – August 2021**
- ESM Data Exploitation and AI and Data Handling user perspective
    - To be released for comment: ~ November 2021

Exascale Computing and Data Handling

Recommendations for the Research Board and WMO members in particular

- We recommend **urgency** in dedicating efforts and attention to disruptions associated with evolving computing technologies that will be increasingly difficult to overcome, **threatening continued advancements in weather and climate prediction capabilities**. The general assumption in the research community that computing will take care of itself is no longer possible given the severity of the technology, software, and scientific challenges.
- The increasing scientific complexity along with the increasing complexity of the computing environment will require **major efforts to adapt or rewrite earth system prediction models**. In addition to scientific accuracy, models must be developed for performance, portability, and productivity.
- The **cost** of computing resources, power consumption, and the related **carbon footprint** must be considered along with the benefit of improved predictability such models may provide.
- Scientists, model developers, computer scientists and software engineers need to work as equal partners on design, development, and maintenance of the applications to overcome scientific, computing, and data challenges. This approach, often referred to as **co-design**, can also be extended to hardware-software co-design efforts with industry.
- A **data-in-place** strategy is needed to support an anticipated 1000 times increase in data volume primarily from observations, model and ensemble output, and post processing. Such a strategy will require co-location of HPC and data, with methods to access, extract, analyze, visualize, and store data by requesting processes and users.

Exascale Computing and Data Handling Concept Note

Recommendations for the Scientific Advisory Panel in particular

- Over time, **few organizations even within developed countries, will be able to fully address the software and data handling challenges**, let alone provision the necessary supercomputing to continue to increase the scientific performance of their codes. A common, shared center could strengthen collaborative research on science, tools, software, and other development activities. Impact and implications for research and operational centers needs to be investigated.
- Scientists from regions lacking access to HPC resources face additional difficulties in adapting to this evolution. Large centers should be encouraged **to open access to some shared resources** as the most effective way for the community to collaborate, foster training and make improvements in all aspects of the prediction system.
- We believe international coordination will enable a more equitable exchange of knowledge so the impacts of scientific research, weather and climate predictions, and data sharing will have broad benefit. Opportunities such as a **WMO education program, or a scientist exchange** could be much easier to implement and more effective with shared access to data.

Interagency Council for Advancing Meteorological Services (ICAMS)

- Restructuring of Federal Efforts
  - DoE, NASA, NOAA, Navy, Air Force
- Priorities
  - Strengthen Collaborations on Climate and Weather
  - Advance Meteorological Support to support Fire Weather
  - Address gaps in Services for Underserved Communities
- Develop workplans for FY2022 and a roadmap for a long term strategy
  - Coupled Global Modeling
  - Physics interoperability
  - Common Modeling Architecture
  - High Performance Computing

FT(Focus Team) / High Performance Computing

Description: This group supports the Earth System Model (ESM) Community through interagency discussions of both R&D and operational compute needs and adaptation to current and future HPC systems and capabilities. It builds broad consensus on coming technologies and discusses best practices for others to consider in their efforts.

High-level priorities, objectives, and actions

- Determine best approaches for developing future earth-system prediction models for increasingly diverse exascale HPC systems
  - Evaluate performance, portability, accuracy, software
  - Explore alternative techniques, languages, models
- Estimate computing needs and cost to develop, test and run KM-scale earth system models for real-time prediction
  - Communicate consistent message to Congress, agencies
  - Investigate feasibility of a large, cross-agency HPC center for earth system prediction

Anticipated deliverables (both short- and long-term)

Deliverable	Due Date	Status
HPC needs study and cost assessment for ESMs	January 2022	TBD
ESM readiness level report for exascale HPC	June 2022	TBD
Exascale Modeling Center feasibility report	December 2022	TBD

Coordination and collaborations with ICAMS and other external groups

- Improvements in models (software, efficiency, algorithms) enable scientific development and reduce HPC requirements and costs
  - Evaluate scientific accuracy and computational efficiency of models in coordination with ICAMS Science Teams
  - Determine the cost of computing to meet science goals in collaboration with ICAMS science and infrastructure teams
  - Explore alternative HPC technologies (AI, Quantum) in coordination with CyFi team
- Collaborate and coordinate with WMO, modeling centers, academia and industry partners

Membership

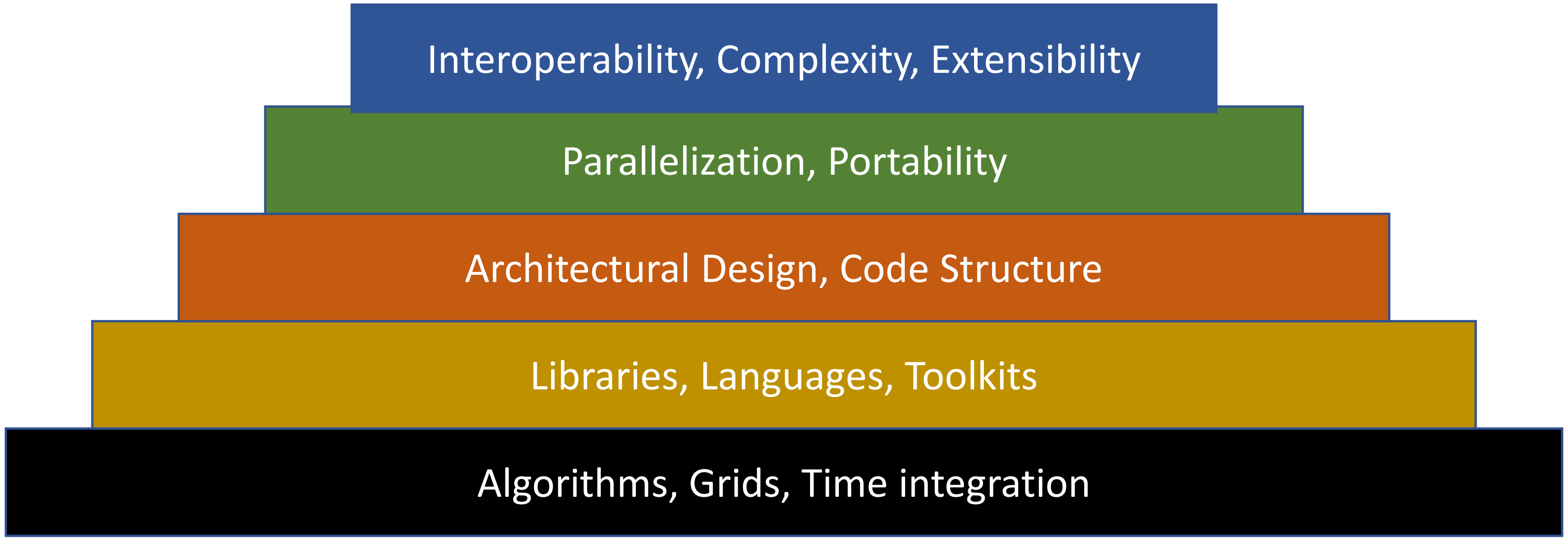
- Mark Govett (NOAA)
- Dave McCarren (Navy)
- Dan Duffy (NASA)
- John Michalakes (UCAR)
- Anke Kamrath (NCAR)
- Brian Comstock (Navy)
- Frank Giraldo (Navy)
- Frank Indiviglio (NOAA)
- Leslie Hart (NOAA)
- Tsengdar Lee (NASA)
- Christine Cuicchi (Navy)
- Phil Jones (DoE)
- Tim Whitcomb (Navy)
- Tom Clune (NASA)

Updated: July 22. 2021

Development Strategy

- Evaluate performance & scalability of models @ target km scale
- Estimate costs – hardware, power, facilities
- Explore alternatives

Model Performance, Accuracy, Usability



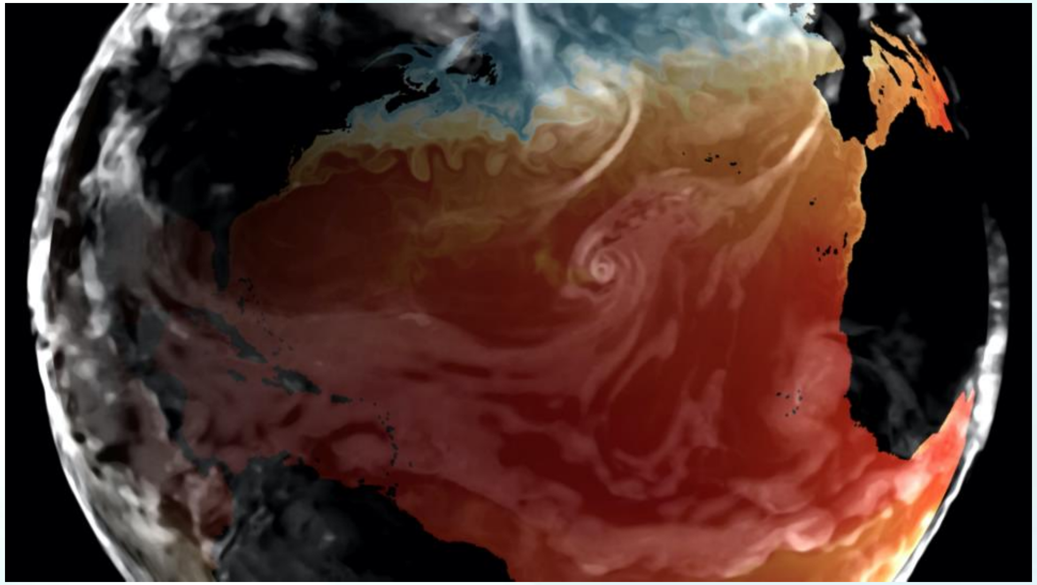
E3SM's GPU Strategy

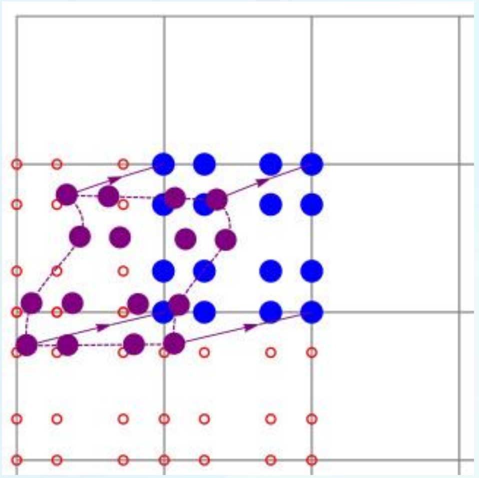
Mark Taylor
Sandia National Laboratories

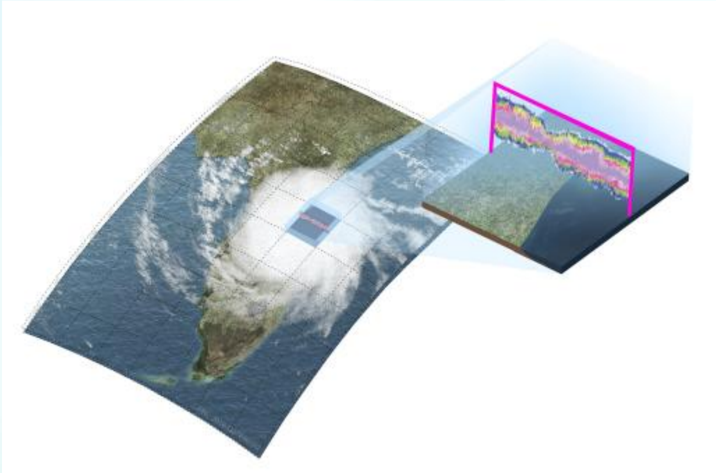
E3SM Nonhydrostatic Atmosphere NGD (Peter Caldwell)
E3SM-MMF Exascale Computing Project. (Mark Taylor)
E3SM Performance Group. (Phil Jones)

Presentation to ICAMS-HPC (Virtual)
September 10, 2021

E3SM Model Development







- **BER: E3SM Project**
- ~70 FTEs, 8 labs + Universities
- **E**nergy **E**xascale **E**arth **S**ystem **M**odel
- DOE-SC science mission: Energy & water issues looking out 40 years
- Ensure E3SM will run well on upcoming DOE exascale computers

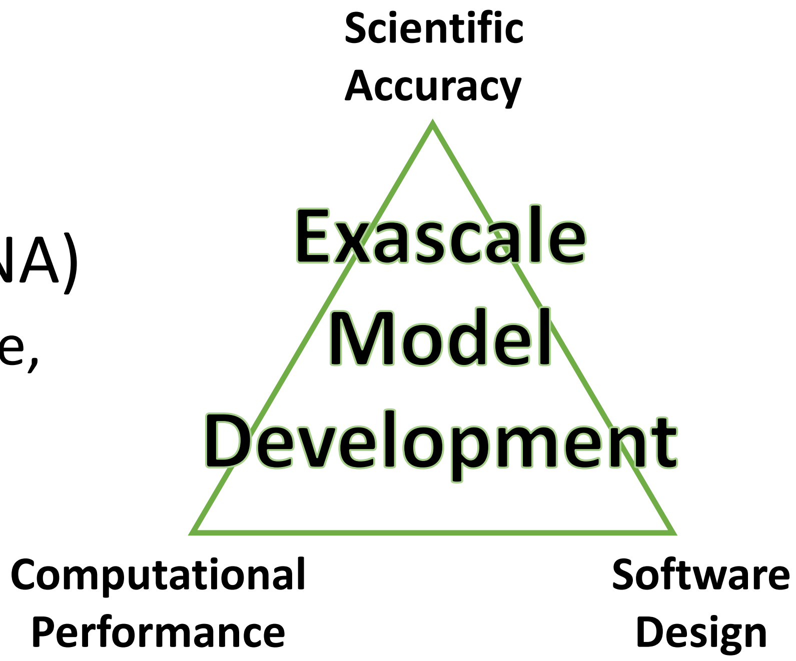
- **ASCR/BER SciDAC**
- ~10 FTEs over multiple projects
- Large focus on new algorithms

- **ASCR ECP Project**
- ~10 FTEs
- E3SM-MMF: "superparameterization"

SAND2021-11206 O

NOAA GSL Exascale Activities

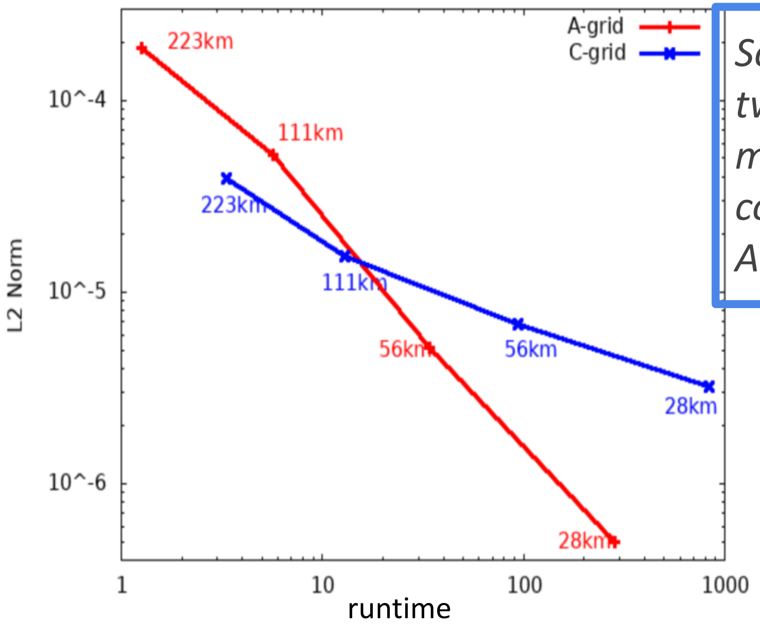
- Software Environment for Novel Architectures (SENA)
    - Improve capabilities of operational codes – performance, portability
    - Explore alternatives – languages, frameworks
- Evaluate scientific accuracy, computational efficiency of algorithms
    - Y. Yu, N. Wang, J. Middlecoff, P. Peixoto, M. Govett, Comparing the Numerical Accuracy of the Icosahedral A-grid and C-grid Schemes in Solving the Shallow Water Model, accepted for publication: **Monthly Weather Review**
    - J. Middlecoff, Y. Yu, M. Govett: Evaluating the Computational Performance of the A-grid and C-grid Shallow Water Model, paper in progress
- Development of GeoFLOW



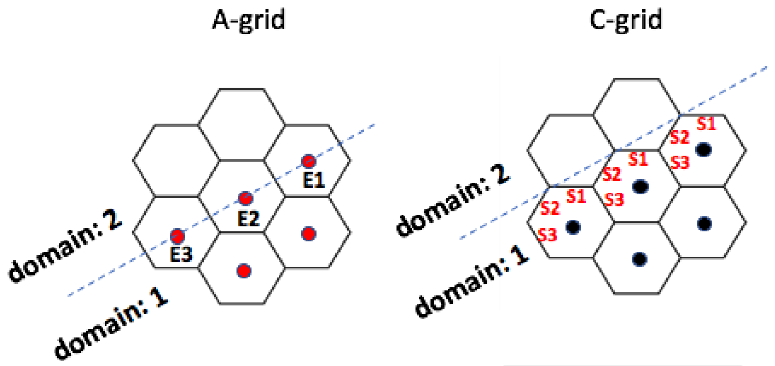
Shallow Water Model: A-grid versus C-grid staggering

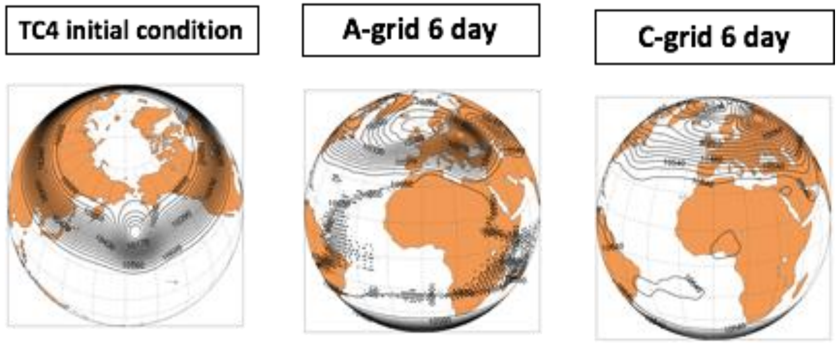
Yonggang Yu, Ning Wang, Jacques Middlecoff, NOAA ESRL, 2018-2019

- Comparison of numerical accuracy for shallow water icosahedral A vs. C grid
- Improved algorithm resulted in better stability and higher accuracy
- Evaluation shows that A-grid is 3X faster than C-grid with similar accuracy



Solution error vs forecast time for two different grids for the SWM model, showing higher order convergence (larger slope) for the A-grid formulation.

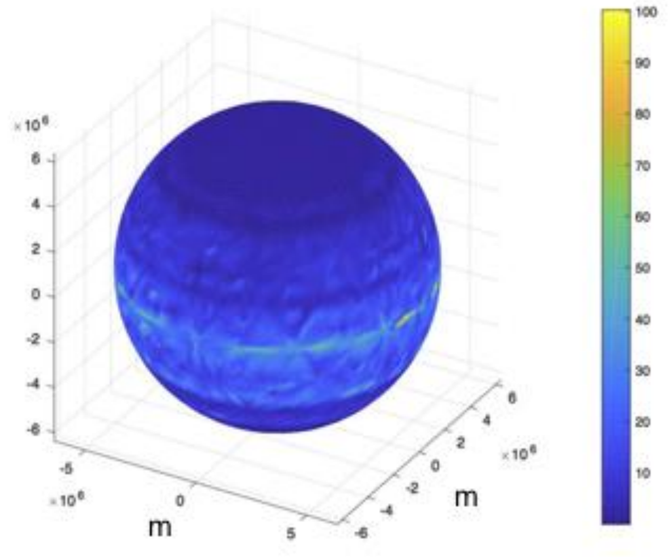


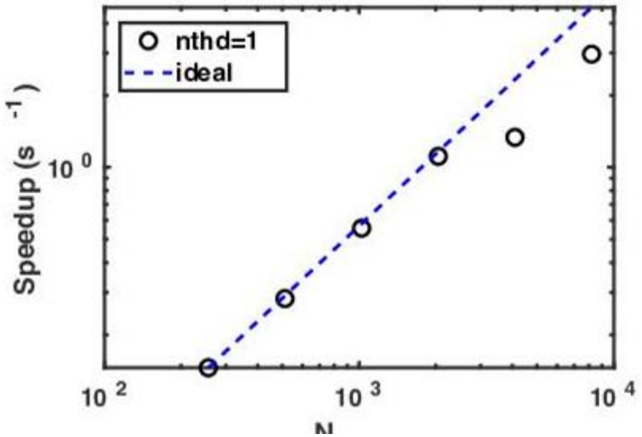


Yu, Y. G., Wang, N., Middlecoff, J., Peixoto, P. S., & Govett, M. W. (2020). Comparing Numerical Accuracy of Icosahedral A-Grid and C-Grid Schemes in Solving the Shallow-Water Model. Monthly Weather Review, 148(10), 4009-4033.

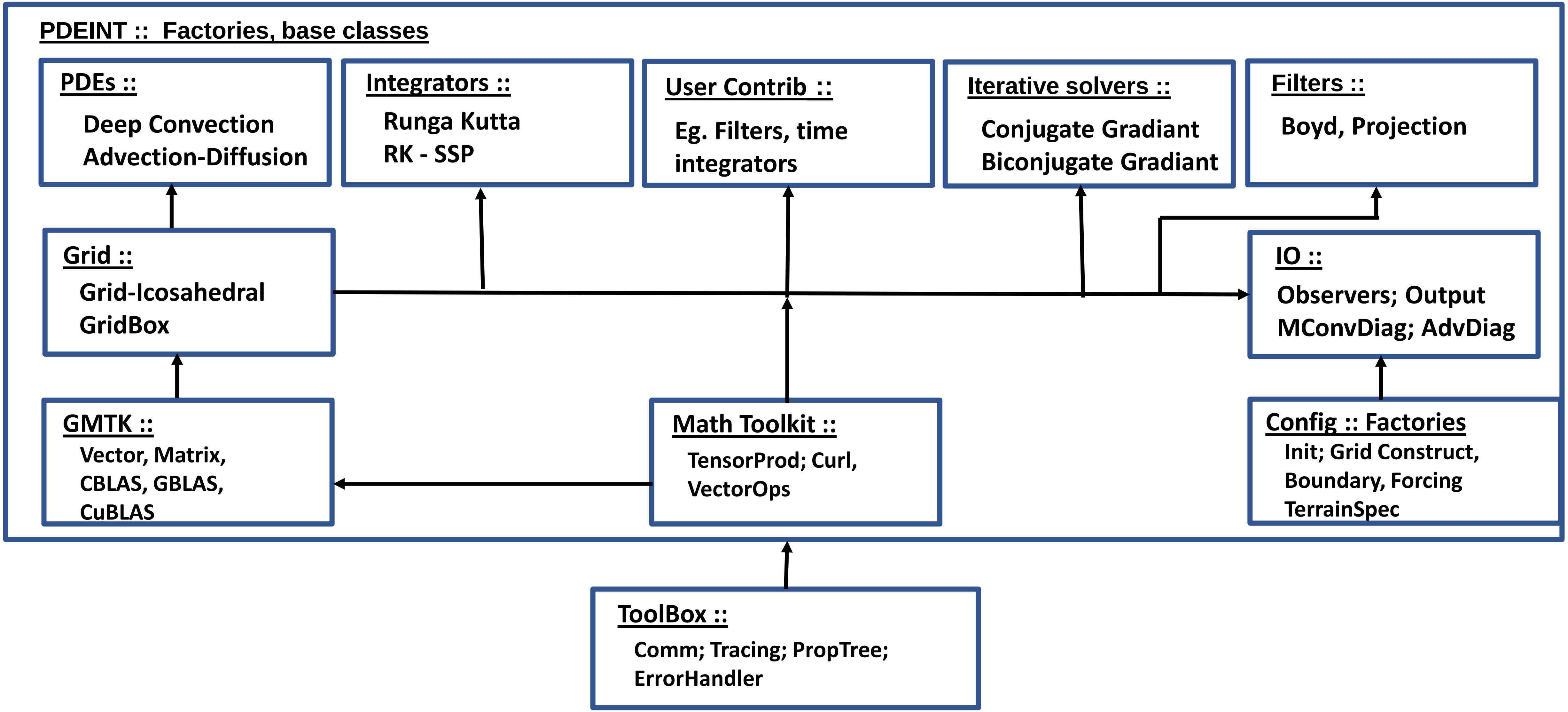
GeoFLuid Object Workbench (GeoFLOW)

- Development began in 2019
- Framework for Kilometer-scale Model Development
  - Evaluate algorithms in terms of scientific accuracy and computational performance
  - Evaluate languages, portability, performance
- Grids supported
  - 2D & 3D Cartesian, 2D & 3D icosahedral grids
  - Cube-sphere grid
- Support spectral element and finite volume methods
- Preliminary testing has begun
  - traditional dycore tests
  - performance evaluation at 3KM or finer scales

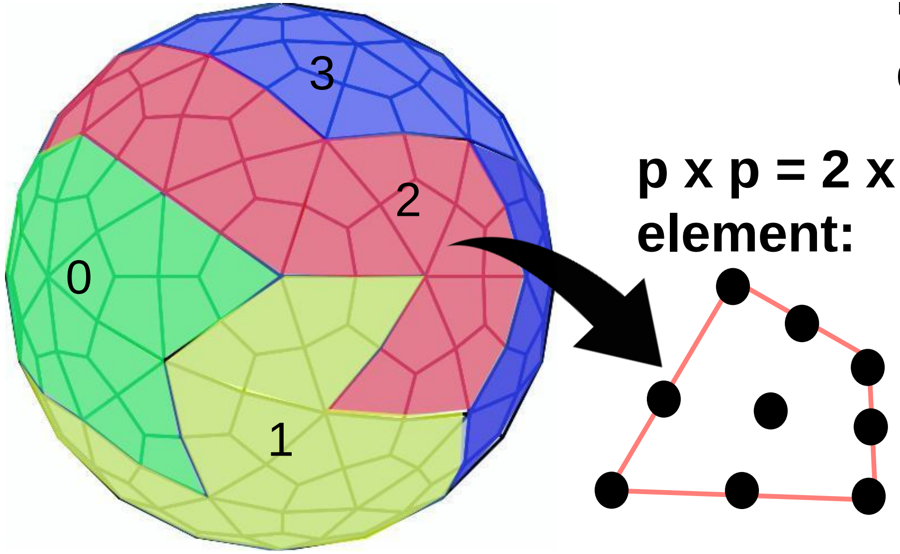




GeoFLOW Objects (C++)



Domain Decomposition and Data Layout



p x p = 2 x 2 element:

Domain decomposition done in horizontal in subdomains ‘owned’ by Task0, Task1, Task2, Task3,…

3D elements stacked radially/vertically.

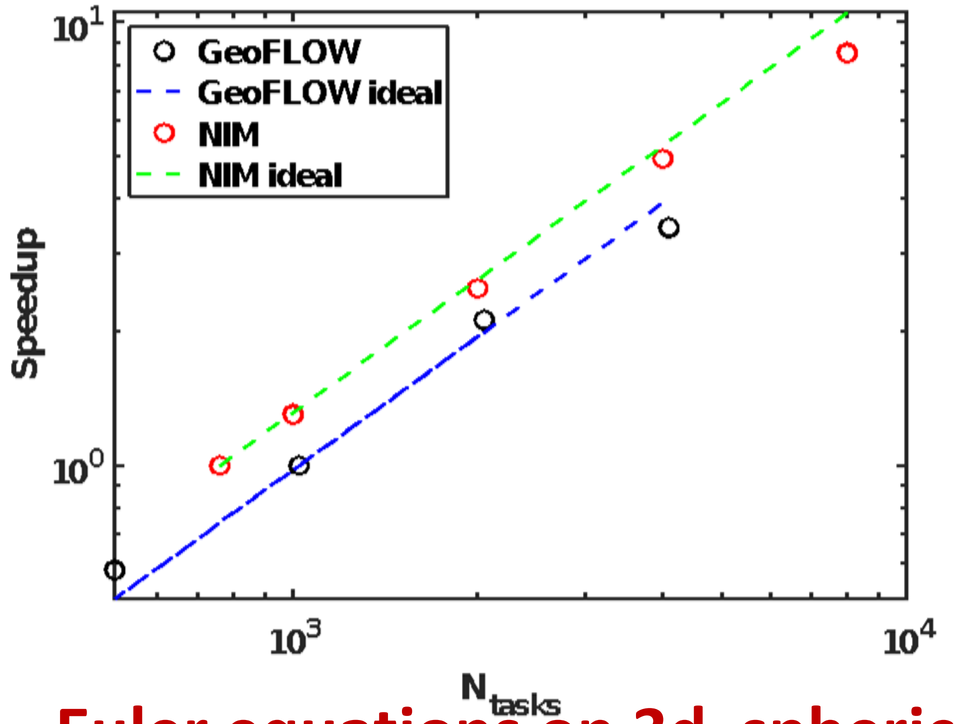
- State data, T, for each element , **i**s accumulated into a contiguous vector on each task for easy vectorization:

$$
T = \begin{matrix} [T_0\ T_1\ T_2\ \ldots\ T_{E0}] \\ [T_0\ T_1\ T_2\ \ldots\ T_{E1}] \\ [T_0\ T_1\ T_2\ \ldots\ T_{E1}] \\ [T_0\ T_1\ T_2\ \ldots\ T_{E2}] \end{matrix}
$$

[Subscript refers to entire 2d/3d high-order element’s worth of data]

- Spectral convergence requires communication between collocated node points only
- All differential operators are also vectorized
- Enables ‘easy’ application of directives…
- …Should map well to accelerators: being tested

Performance Results: Strong scaling - Preliminary



Computational Efficiency = $\langle T_i N_i / T_{i-1} N_{i-1} \rangle$

FiniteVolume NIM = 0.92

GeoFLOW = 0.95

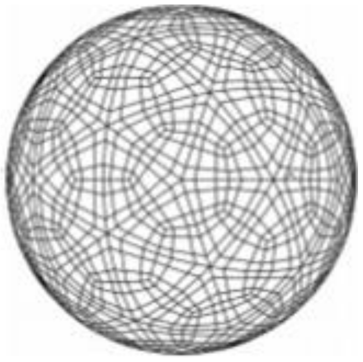
Euler equations on 3d spherical grids:

NIM: 2,621,440 horiz. polygons (Δx = 14 km), 96 vertical levels

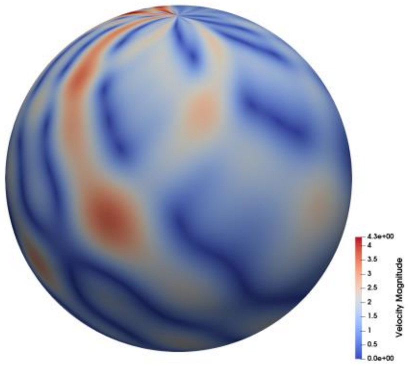
GeoFLOW: 61,440 horiz. elements, p x p = 8 x 8;
3,932,160 horiz. polygons (Δx = 11 km),
12 p=8 vertical elements, 96 vertical levels

Visualizing GeoFLOW Data

- What
  - Load GeoFLOW data into a robust 3D visualization environment and navigate through layers of complex data at interactive rates
  - Quickly observe small-scale features superimposed with large scaled features
- Why
  - Size of scientific data continues to increase
  - Need enough compute resources to dynamically visualize and navigate through complex 3D scientific data at interactive rates
- How
  - Use VAPOR software (developed at NCAR) to provide fast, flexible, progressive access to high resolution data



GeoFLOW on an icosahedral grid



Rendering of a GeoFLOW variable

Summary Points

- Convergence on earth system models
  - WMO, ICAMS, ESCAPE, etc
  - Urgency in communicating challenges to decision makers
- Major investment to improve models
  - Adaption, rewriting codes
  - Select algorithms with best capabilities
  - Choose modern languages
- Co-design, collaboration essential
  - Scientists, mathematicians
  - Software engineers
  - Computer scientists
  - Industry