The Influence of Exascale Architectures on Earth System Modeling in NASA

Can the compute capacity at NASA keep pace with science demands?

Bill Putman – NASA/GMAO

Dan Duffy – NASA/CISTO

Tsengdar Lee – NASA/HQ

Tom Clune : Ben Auer : Hamid Oloso : Matt Thompson : Atanas Trayanov – GMAO/SI-Team

GEOS: A Scale-Aware Modeling System

"GEOS is a comprehensive global model for simulation, assimilation, and prediction on weather and climate time-scales"

1. Weather Analysis and Prediction

- *near-realtime analyses, assimilation products, and forecasts*
  - In support of NASA's satellite missions and field experiments
  - Generating atmospheric products for a broad community of users.

2. Seasonal-Decadal Analysis and Prediction

- Coupled Earth-System models and analyses of subseasonal to seasonal variability
  - National Multi-Model Ensemble (NMME) project
  - Chemistry-Climate Model (CCM)
  - Coupled Model Intercomparison Project (CMIP)

3. Reanalysis for Climate

- Modern-Era Retrospective analysis for Research and Applications (MERRA-2)
  - Hi-Resolution global downscaling of reanalyses

4. Global Convection Allowing

- Global simulations at the forefront of model and computing capability
  - These form the basis for *Observing System Simulation Experiments*.

Supported Mission	GEOS Data Used	GMAO's Mission Page
AMSR Unified Rainfall Team (NASA - JAXA) Advanced Microwave Scanning Radiometer Unified Rainfall Team	FP-IT	
CALIPSO (NASA - CNES) Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation	FP-IT	
CATS (NASA) Cloud-Aerosol Transport System	FP-IT	
CERES (NASA, several platforms) Clouds and the Earth's Radiant Energy System ▪ CERES — Climate records of the radiation balance ▪ FLASHFlux — Rapid-releast of TOA radiative fluxes	G5-CERES	
MLS (NASA: EOS-Aura) Microwave Limb Sounder	FP-IT	
MODIS (NASA: EOS-Terra and EOS-Aqua) Moderate Resolution Imaging Spectroradiometer	FP-IT	
MOPITT Measurement of Pollution in the Troposphere	FP-IT	
MPLNET (NASA) Micro-Pulse Lidar Network	FP, FP-IT	
OCO-2 (NASA) Orbiting Carbon Observatory-2	FP-IT	
OMPS-LP (NASA-NOAA: Suomi-NPP) Ozone Mapping Profiler Suite	FP-IT	
SMAP (NASA) Soil Moisture Active Passive	FP	SMAP Level-4 Products
TES (NASA: EOS-Aura) Tropospheric Emission Spectrometer	FP-IT	

GEOS: A Scale-Aware Modeling System

"GEOS is a comprehensive global model for simulation, assimilation, and prediction on weather and climate time-scales"

1. Weather Analysis and Prediction

- *near-realtime analyses, assimilation products, and forecasts*
  - In support of NASA's satellite missions and field experiments
  - Generating atmospheric products for a broad community of users.

2. Seasonal-Decadal Analysis and Prediction

- Coupled Earth-System models and analyses of subseasonal to seasonal variability
  - National Multi-Model Ensemble (NMME) project
  - Chemistry-Climate Model (CCM)
  - Coupled Model Intercomparison Project (CMIP)

3. Reanalysis for Climate

- Modern-Era Retrospective analysis for Research and Applications (MERRA-2)
  - Hi-Resolution global downscaling of reanalyses

4. Global Convection Allowing

- Global simulations at the forefront of model and computing capability
  - These form the basis for *Observing System Simulation Experiments*.

GEOS: A Scale & Aerosol Aware Modeling System

"Our main tasks in atmospheric modeling will be to make the transition to a global non-hydrostatic model and to develop the physical parameterizations appropriate to high resolution, particularly those related to moist processes and to aerosol-cloud interactions at the microphysical level."

1. The Finite-Volume Cubed-Sphere (FV3) dynamics

- *Global to regional non-hydrostatic capability:*
  - Efficient NH dycore with improved effective resolution
  - Global stretched grids and nesting capability

2. From MERRA to MERRA-2

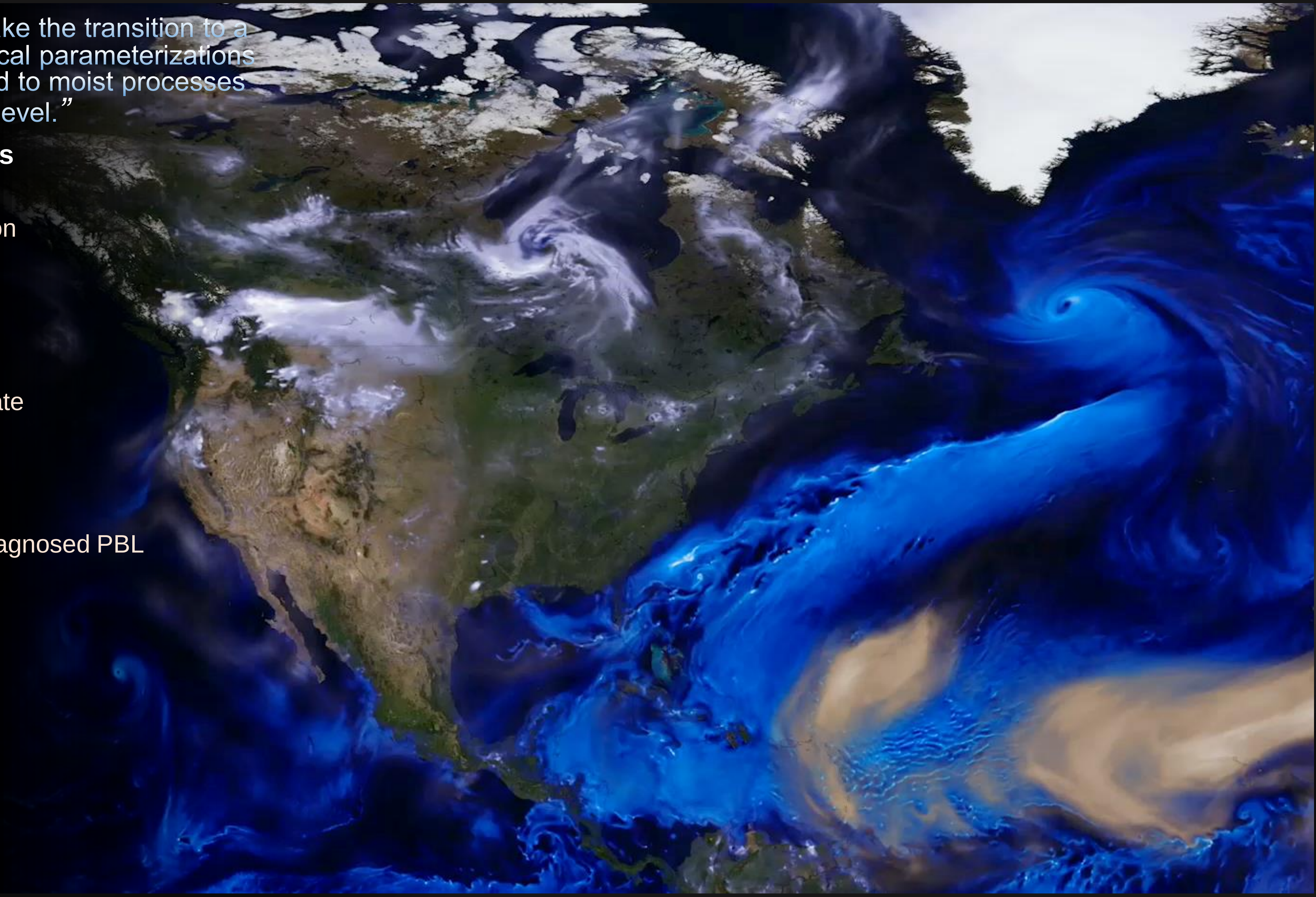
- *Resolution-aware model physics:*
  - Stochastic Tokioka Limiter in RAS Convection
  - Critical Relative Humidity for Large-Scale Condensate

3. Evolved Physics

- *Incremental updates from MERRA-2 physics:*
  - GMTED filtered topography & retuned GWD
  - Updated turbulent length scale in Louis based on diagnosed PBL
  - Enhanced shallow & mid-level convection in RAS
  - Convective vs Stratiform ice settling
  - Realistic liquid & ice cloud radii

4. Advanced Physics

- *Scale-aware & Aerosol-aware:*
  - Morrison Gettelman Barahona Cloud Microphysics
  - GFDL 6-phase microphysics
  - Grell-Freitas Convection + UW Shallow Convection
  - EDMF / SHOC Turbulence and Boundary Layer
  - RRTMG Radiation



6-km GEOS Aerosol Replay (2017 Hurricanes)

GEOS Developments toward an Earth System Model

"…developing an Earth System Model (ESM) and Integrated Earth System Analysis (IESA) capability, which links simultaneous analyses of Earth system components and captures the necessary interactions between them"

1. Ocean GCM, Sea-Ice and Biogeochemistry

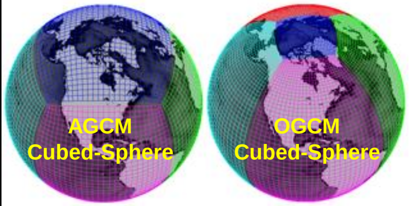
- *Evolution to the GEOS S2S v3 and IESA:*
  - MOM4 => MOM5
  - MITgcm Hi-Res Ocean-Atm coupling
  - Development of CICE
  - Inclusion of NOBM Biogeochemistry to complete the carbon cycle

2. Wave Model

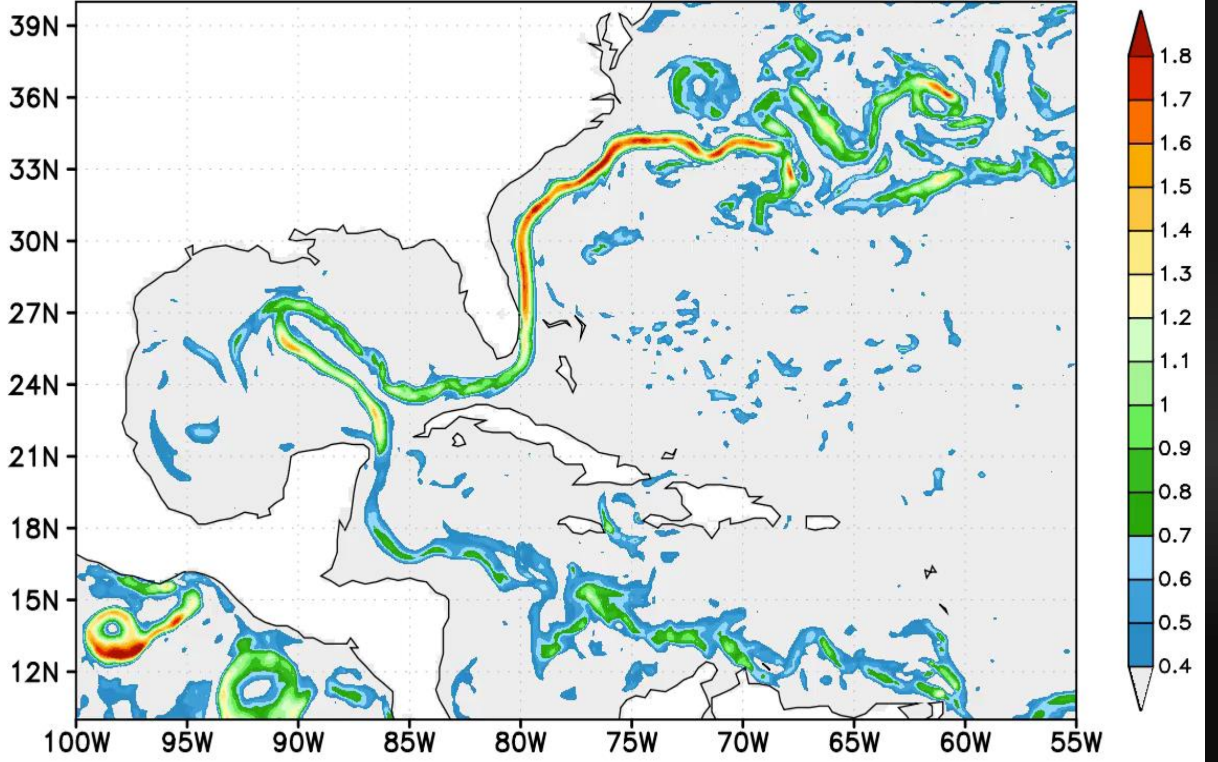
- *Improving air-sea fluxes:*
  - Implementing the Miami Wave Model *(MAP16 funded)*
  - Inclusion of sea-spray effects *(feeding into aerosol model)*
  - Strengthen ocean-atmosphere coupling in the IESA

3. Land Processes

- *Catchment Land Surface Model developments:*
  - Dynamic phenology and carbon storage
  - Improved river routing to predict catchment transport of surface water
  - Updated snow model



12-km GEOS Atmosphere
coupled with
9-km MITgcm Ocean



Coupling of Eddy-permitting Ocean with Convection-permitting Atmosphere

Andrea Molod (GMAO), Ehud Stobach (UMD) and Dimitris Menemenlis (JPL)

GEOS 1.5km Global Simulation

"Global simulations approaching 1km resolution begin to simulate our atmosphere at the resolution of our global observing system"

1. Form the basis for Observing System Simulation Experiments

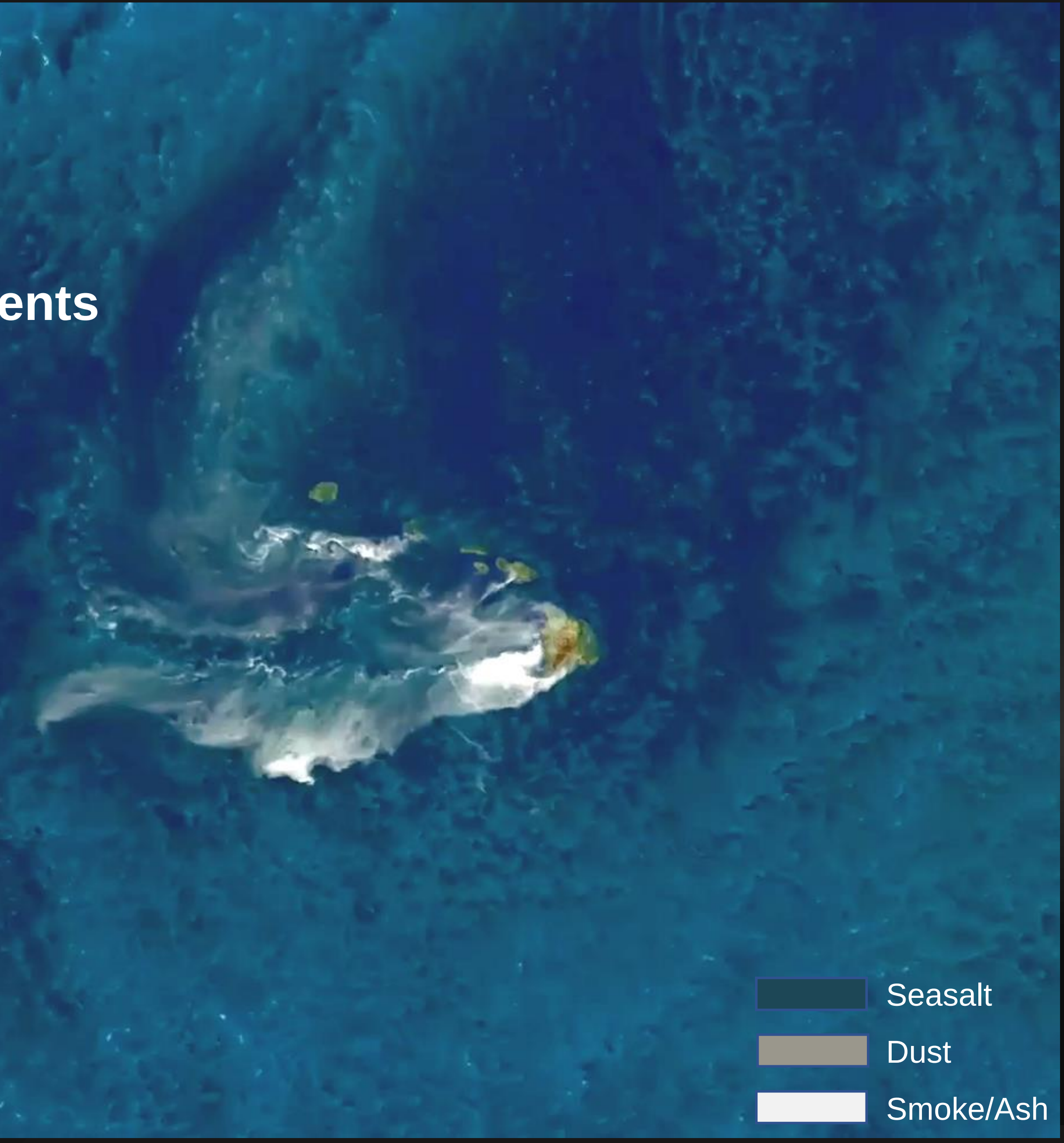
- *Known as 'Nature-Runs' used to simulate the observing system*
  - Satellite simulators estimate existing observations and errors
  - Instrument teams can use these for designing new observations

2. Evaluate new science and resolved scale processes

- While production systems remain highly parameterized
  - Shallow convection only
  - 2-Moment Morrison Gettelman cloud microphysics
  - Aerosol cloud interactions and gaseous transport

3. Stress our modeling infrastructure

- Optimize memory usage and I/O support
  - Resolve external memory issues when scaling (ESMF and MPI)
  - Offload output streams to asynchronous worker nodes



GEOS Simulated Aerosols Around Hawaii June 15-18, 2012

GEOS Project DYAMOND Experiments

Project DYAMOND: Simulating August 2016 (40-days)

Dynamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains

https://www.esiwace.eu/services/dyamond

1. GEOS Global Simulations (AMIP and "REPLAY")

- 12km (c768) 72-levels
- 6km (c1536) 72-levels
- 3km (c3072) 72-levels *[waiting for new compute nodes]*

2. Scale-aware parameterization

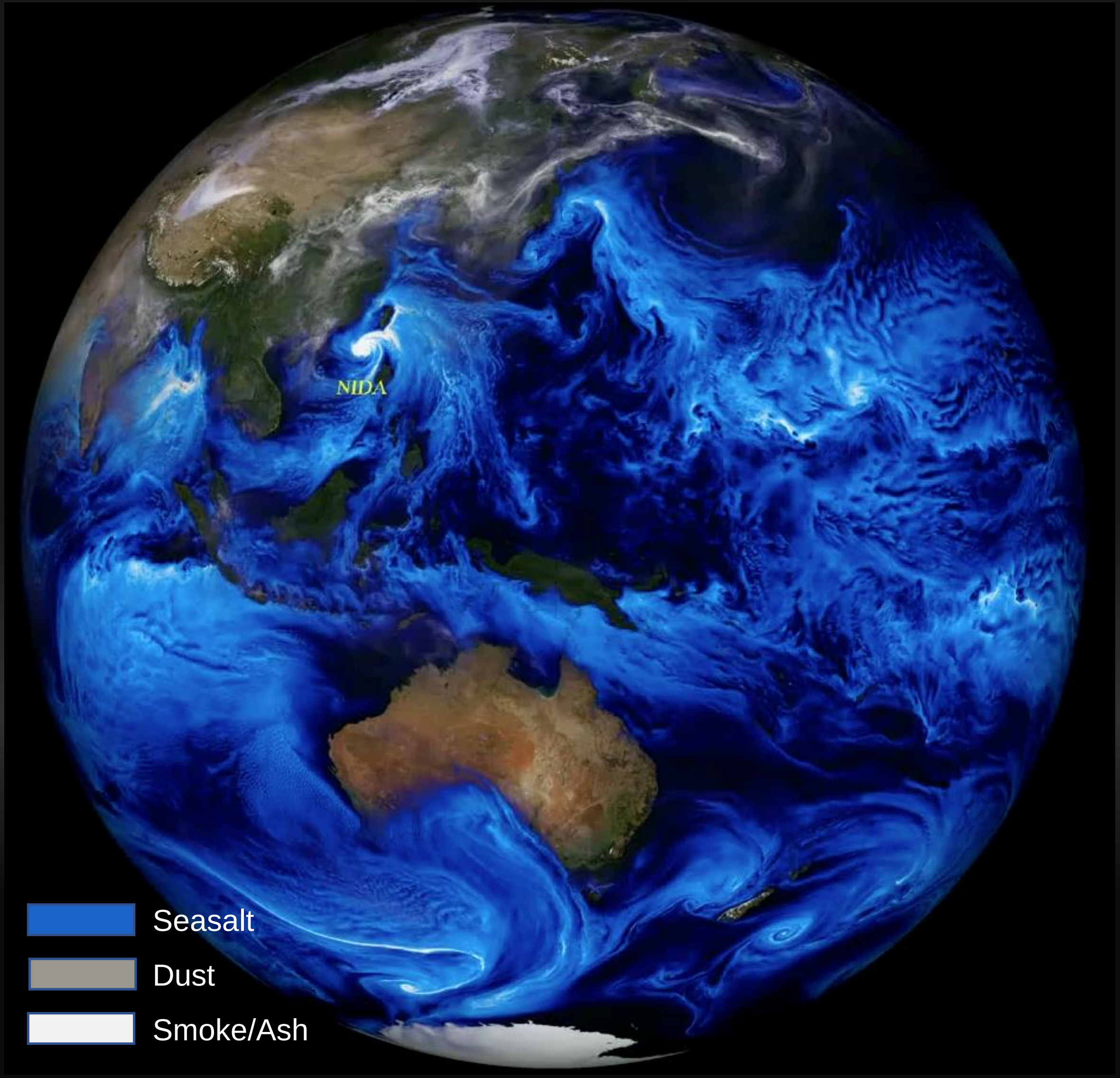
- Grell-Freitas
- Grell-Freitas (disabling deep convection)

3. Cloud microphysics

- Single-moment GEOS/GFDL microphysics
- Two-moment Morrison-Gettelman-Barahona

4. Aerosol & Gaseous species

- GOCART aerosols: Organic/Black Carbon, Sulfate, Dust, Seasalt
- Gaseous species: CO, CO_2, SO_2, O_3



6-km GEOS Replay : Aerosols : August 2016

Compute Expansion at NCCS since 2010

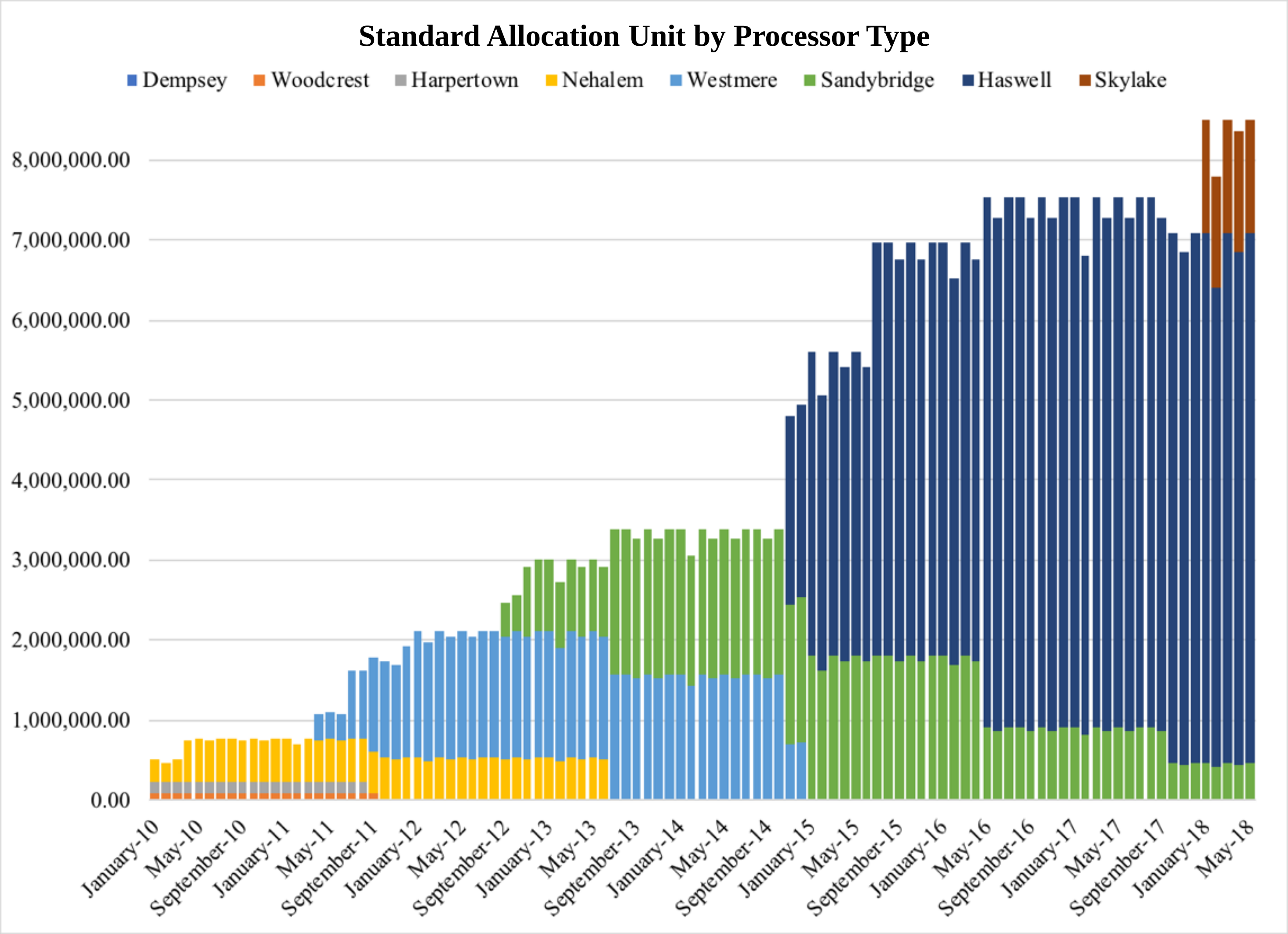
Growth rate: ~10x from 2010-2015 ~2x from 2015-2019

The addition of Scalable Unit 14 (SCU14)

Processors: The Supermicro FatTwin server nodes each have dual 20-core Intel Xeon Gold 6148 “Skylake” processors (2.4 GHz).

Instruction Set: Skylake uses a different instruction set, namely Advanced Vector Extensions 512 (AVX-512).

Interconnect: Connecting the Skylake nodes at 100 gigabits per second is Intel’s Omni-Path Architecture.



NCCS
NASA CENTER FOR CLIMATE SIMULATION
HIGH-PERFORMANCE SCIENCE

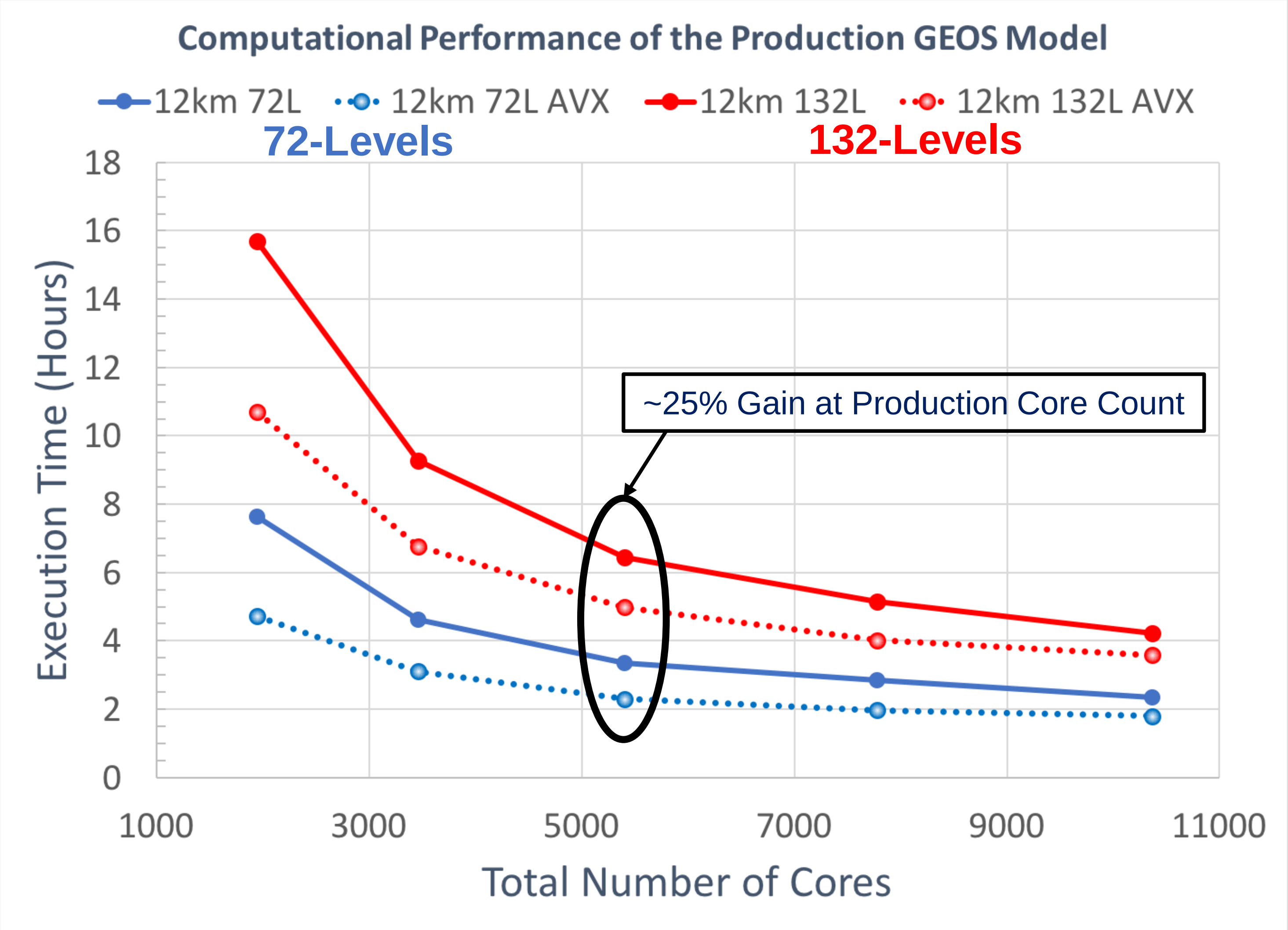
The addition of Scalable Unit 14 (SCU14)

Processors: The Supermicro FatTwin server nodes each have dual 20-core Intel Xeon Gold 6148 "Skylake" processors (2.4 GHz).

Instruction Set: Skylake uses a different instruction set, namely Advanced Vector Extensions 512 (AVX-512).

Interconnect: Connecting the Skylake nodes at 100 gigabits per second is Intel's Omni-Path Architecture.

Gain in GEOS using AVX-512 instructions



GEOS: Production 13-km GEOS Configuration

1. History *(All model & analysis output fields)*

- *Significant bottleneck without asynchronous output*

2. Aerosols *(Advection + GOCART)*

- *More than 25% of the total execution*
  - 57 advected species
  - Overhead from reading and regridding of emissions
  - Computation of transport and wet-deposition

3. FV3 *(only 22% of total execution)*

- *FV3 is very efficient*
  - Scales well even to 5400 cores at c720

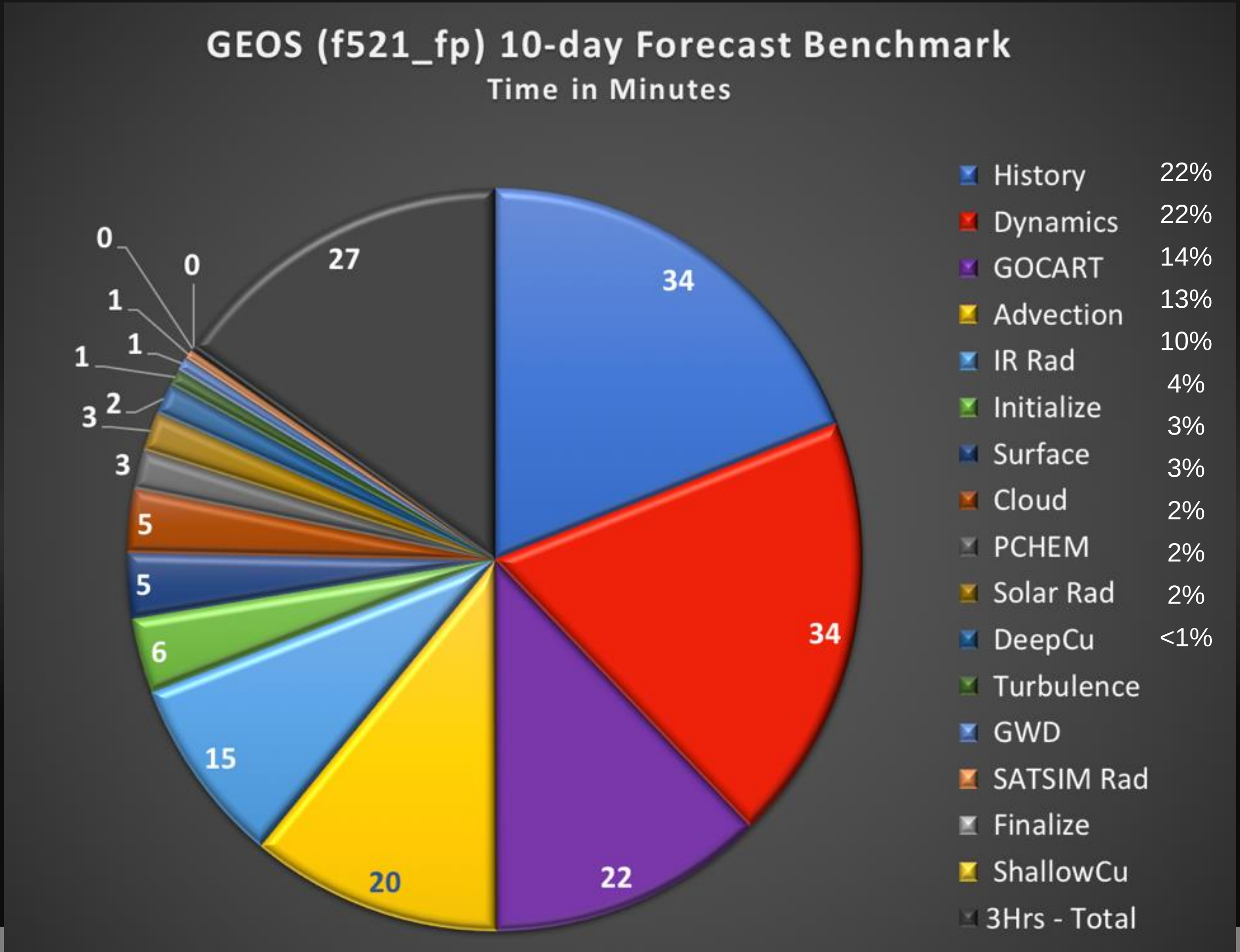
4. Physics & Radiation

- *Long time-scales (for now…)*
  - RRTMG LW (3600s update timestep)
  - Chou-Suarez SW is quite efficient
  - Cloud physics + GWD + Turbulence < 5%

GEOS NWP Model Profile

Component times in minutes

10-day forecast in 2.5 hours using 5400 MPI cores



Total Walltime: 2.5 Hours

FV3: Interagency Collaboration

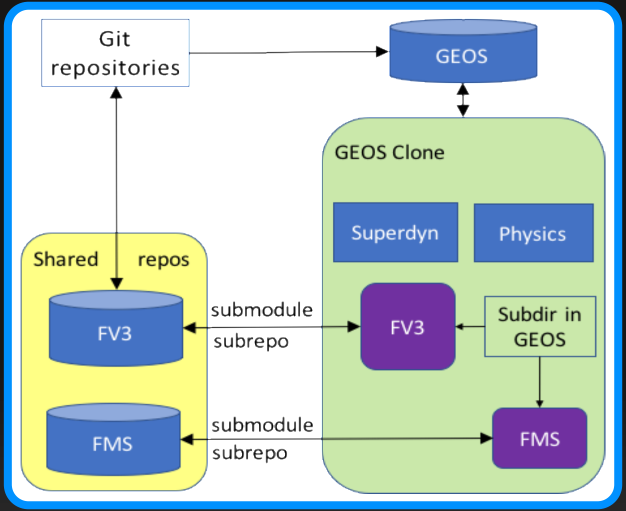
Shared FV3 component lives in separate Git repos

FV3 exists as a normal subdirectory in GCM
- Most users unaware
- Requires some minor refactoring

Two variant Git approaches
- Git **submodule**
  - Lightweight – just links are stored
  - Most users should treat subdir as static
- Git **subrepo**
  - Files stored in both repos
  - Unaware users can modify subdir contents

Hosting shared components
- Ideally via a public site (e.g., GitHub)
- But can use read-only clones at each end

Released versions to NOAA VLab



GMAO
GEOS
- Data Assimilation
- Weather Prediction
- S2S Prediction
- Reanalysis
- Global Mesoscale

GFDL
AM3 CM3 HiRAM
- Climate (IPCC)
- S2S Prediction
- Weather
- High-Resolution

EMC
GFS GEFS SFS
- Operational DA
- Operational NWP
- S2S Prediction
- Reanalysis
- *Future CAM and CAM Ensembles*
- *Hurricanes, Space Weather, WoF …*

GISS
modelE
- Climate (IPCC)
- Exo-planets

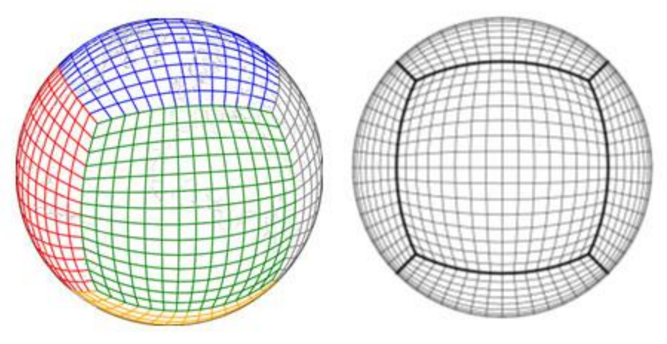
Harvard
GEOS-Chem
- Chemical Transport
- Composition

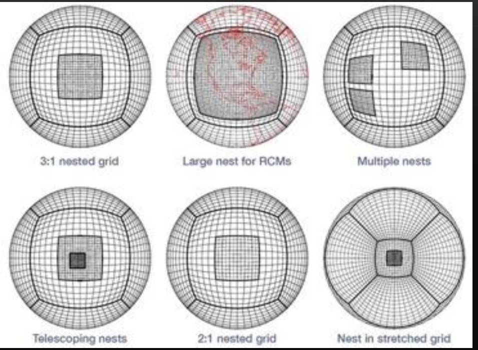
ESRL
NGGPS
- Real-time FV3-GFS
- Advanced Physics

NCAR
CESM
- Climate (IPCC)
- Data Assimilation

JCSDA Joint Center for Satellite Data Assimilation

JEDI
- Data Assimilation





FV3: Interagency Collaboration

GEOS is significantly more expensive than GFS

Most of this expense lies in aerosol, chemistry and cloud interactions

12-13km 10-day forecasts on 1536 Xeon Cores

c768 64L	fv3GFS
FV3	3004
Physics	497
Radiation	693

c768 72L	GEOS
FV3	3370
Physics	3300
Radiation	2665
Chemistry	4805
Extra Tracers	1850

Time spent in model components (seconds)
NOAA data courtesy of Lucas Harris (GFDL)

2018 NASA Campaigns Supported by the GEOS system at GMAO
A significant connection to clouds and aerosols

2018 Campaigns Supported *[Revisit this page for additions and updates]*

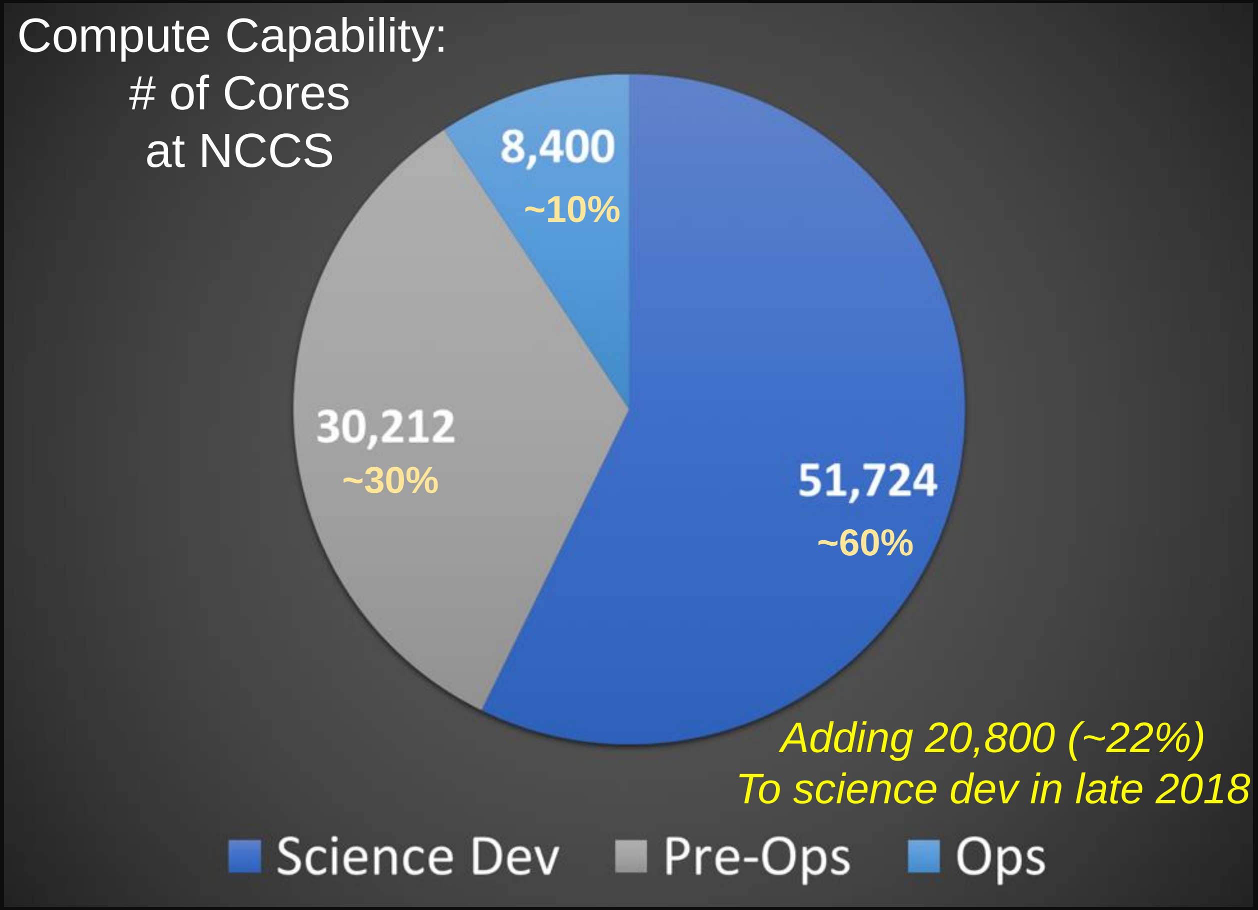
Campaign Name	Support Timeframe	GMAO Support page
SOCRATES Southern Ocean Cloud, Radiation, Aerosol Transport Experimental Study	Jan 10, 2018–Mar 10, 2018	SOCRATES Portal
ACE-ENA 2018 Aerosol and Cloud Experiments in Eastern North Atlantic	Jan 11, 2018–Feb 20, 2018	ACE-ENA 2018 Portal
ASPIRE Advanced Supersonic Parachute Inflation Research Experiments	Mar 12, 2018–Apr 07, 2018	
ATom-4 Atmospheric Tomography Mission	Apr 18, 2018–May 21, 2018	ATom-4 Portal
WE-CAN Western wildfire Experiment for Cloud chemistry, Aerosol absorption and Nitrogen	Jul 01, 2018–Sep 30, 2018	WE-CAN Portal
ABOVE Arctic - Boreal Vulnerability Experiment	Jul 15, 2018–Sep 15, 2018	ABOVE Portal
ORACLES ObseRvations of Aerosols above CLouds and their intEractionS	Sep 01, 2018–Nov 30, 2018	ORACLES Portal
CAMP2EX Cloud, Aerosol and Monsoon Processes Philippines Experiment (CAMP2Ex)	Sep 15, 2018–Nov 30, 2018	CAMP2EX Portal

Current GEOS Modeling Capabilities at NCCS

Modeling Component	2018 Capability	
	Resolution	*Cores*
Climate/Reanalysis	50 km	O(1000)
Sub-Seasonal to Seasonal	50 km Atm 50 km Ocn	O(1000)
Deterministic Medium-Range	12 km *72 Levels*	8,000
Pioneering Global Embedded-Regional OSSEs	1-3 km *72 Levels*	O(10,000)
Radiative-Convective-Equilibrium Large-Eddy	500 m	O(1,000)
Single Column	132 Levs	1

Compute capacity: 3.5 petaflops

Utilization of available resources at NCCS (ignoring NAS):



Projected GEOS Modeling Capabilities at NCCS

Modeling Component	2018 Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*Cores*	*Resolution*	*Cores*	*Resolution*	*Cores*
Climate/Reanalysis	50 km	O(1000)	25 km Atm 25 km Ocn	O(10,000)	13 km Atm 13 km Ocn	O(100,000)
Sub-Seasonal to Seasonal	50 km Atm 50 km Ocn	O(1000)	25 km Atm 25 km Ocn	O(10,000)	13 km Atm 13 km Ocn	O(100,000)
Deterministic Medium-Range	12 km *72 Levels*	8,000	6-9 km *132 Levels*	600,000	3-6 km *~200 Levels*	6 million
Pioneering Global Embedded-Regional **OSSEs**	1-3 km *72 Levels*	O(10,000)	500 m – 1 km *132 Levels*	O(10 million)	100-500 m *~200 Levels*	O(100 million)
Radiative-Convective-Equilibrium **Large-Eddy**	500 m	O(1,000)	10-100 m	O(100,000)	1-10 m	O(1 million)
Single Column	132 Levs	1	132 Levs	1	~200 Levs	1
Compute capacity:	3.5 petaflops		~10x Increase		~100x Increase	

What have we left out…

Ensembles

Coupling in NWP

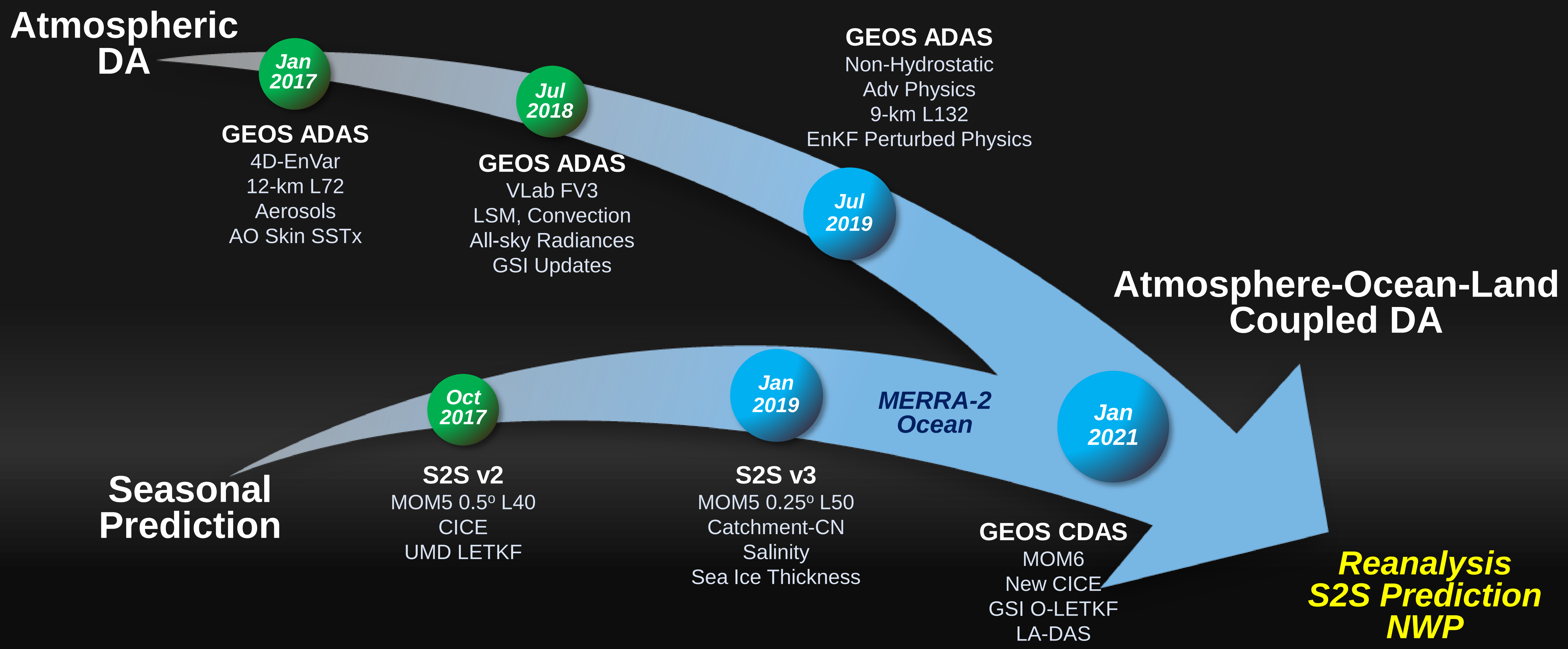
Physics Complexity

Chemistry

Easily a 1000x increase

Need for Excascale computing capability

GEOS: Toward a Coupled Integrated Earth System Analysis



JEDI: Scalable/Flexible Data Assimilation Framework

Joint Effort for Data assimilation Integration (JEDI)

What computational benefits will JEDI provide for data assimilation at NASA and beyond?

1. DA on the native grid --- *a pathway to 4DVar*

- *GFS & GEOS run on the FV3 cubed-sphere (CS) and GSI Lat-Lon (LL):*
  - GSI requires transforms from CS to LL and LL to CS
  - In a 4DEnVar approach this is only done once per DA cycle
  - To enable 4DVar using TLM and Adjoint a native grid approach saves O(100) horizontal transforms from CS to LL and LL to CS

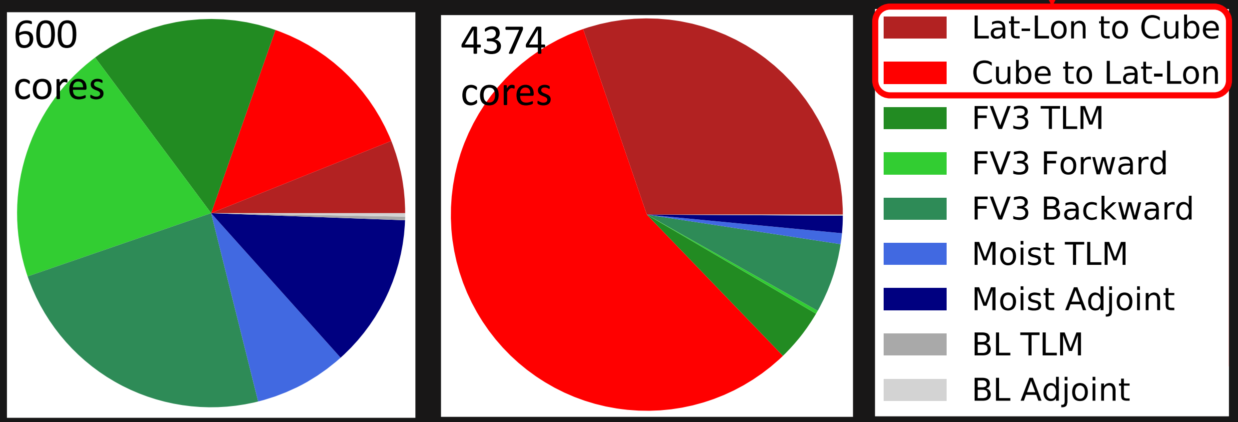
2. Alternative application of the static BKG ErrCov

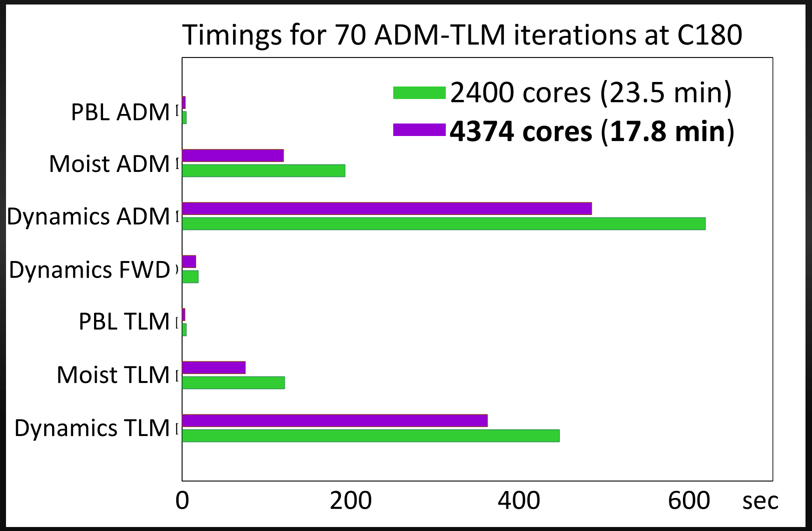
- *GSI spreads the background error covariance through recursive filters:*
  - *Presents limitations on scalability requiring full 2D horizontal slabs*
  - Limited to the product of the number of levels and variables O(400)
  - JEDI-OOPS will implement alternatives to spread this information

3. Unification, Coupled DA and beyond…

- *JEDI will unify DA amongst Earth system components:*
  - Easily leverage scientific developments across components
  - Provide a natural pathway to coupled DA

GSI CS to LL Bottleneck





FV3 ADM-TLM

70 iterations permits a viable 4DVar solution at respectable inner-loop resolution (C180)

JCSDA
Joint Center for Satellite Data Assimilation
A multi-agency research center created to improve the use of satellite data for analyzing and predicting the weather, the ocean, the climate and the environment

Hybrid MPI+OpenMP Scaling and Optimization in GEOS

We are enabling hybrid MPI+OpenMP capability throughout GEOS

1. FV3

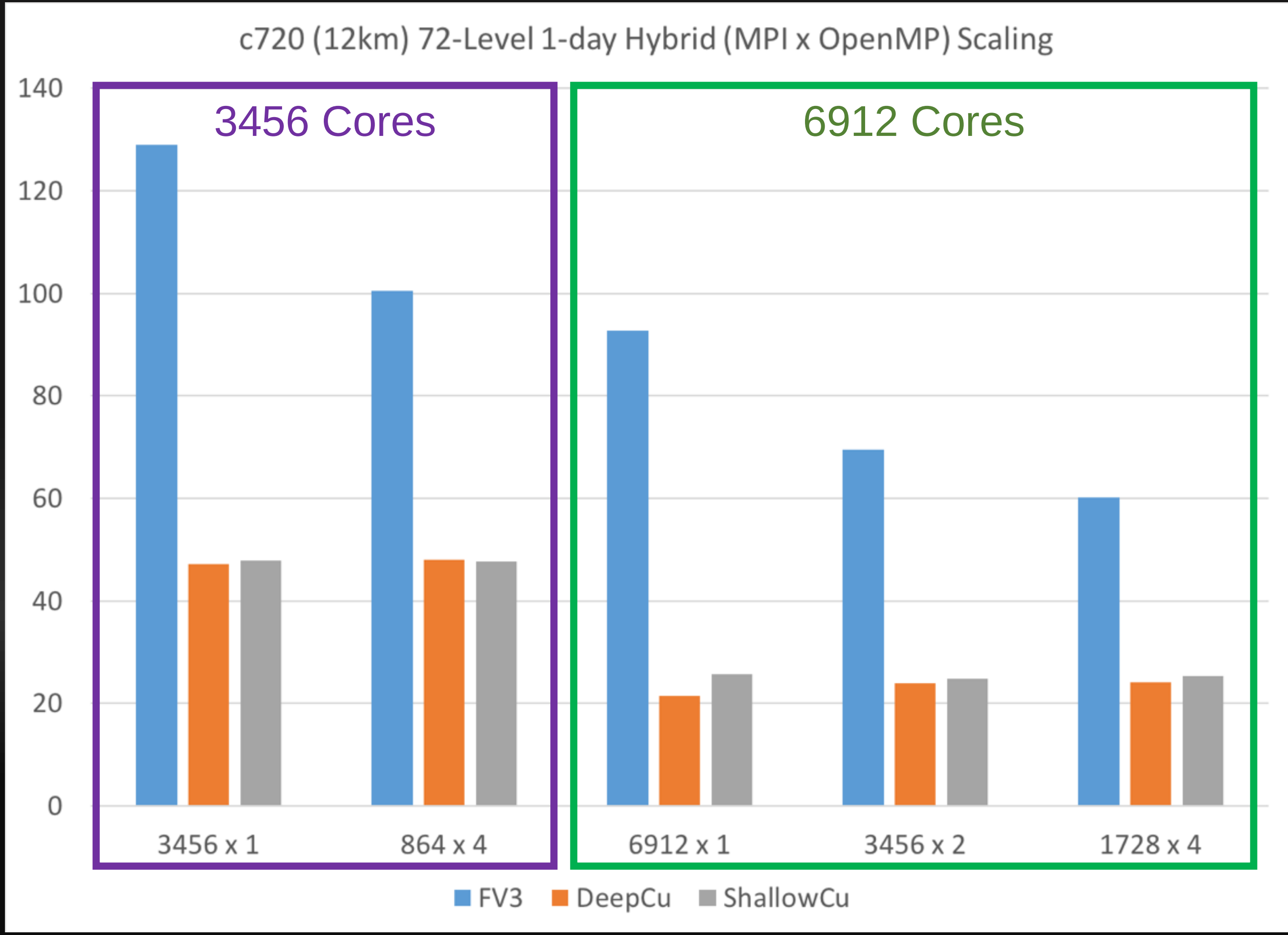
- MPI (comm:work ratio) limits scalability
  - *Halo updates*
- Hybrid threading is essential for FV3 scaling

2. Physics components

- Scale well in either case
- No communication (for now…)
- But.. Not currently written with OpenMP

3. Two hybrid approaches

- Hybrid FV3 + MPI Physics
  - *Requires MPI redistribution between FV3 & Physics*
- Add OpenMP parallelism throughout the physics
  - *A large volume of code to address…*



DeepCu is Grell-Freitas ShallowCu is UW Park Bretherton

Maintenance of GEOS GPU Enabled Code

Our production GEOS code maintains support for GPUs

1. **Some physics kernels do well**
   - Radiation (RRTMG in particular)
2. **FV3 takes significant re-writing**
   - An older simplified GPU version of FV3 exists
   - Advection kernels do well for big domains
   - Data exchange is prohibitive
3. **Infrastructure limitations**
   - ESMF connects all GEOS components
     - *Internal arrays are stored as ESMF pointers*
   - Data exchanges are expensive
     - *CPU-GPU data exchange on entry/exit to every GEOS model component*
   - Not a viable end-to-end solution at the moment

C360 (25-km) – 16 CPU Cores (Xeon) v 1 GPU (K40)

	72 Levels			137 Levels		
	CPU (seconds)	GPU (seconds)	Speedup	CPU (seconds)	GPU (seconds)	Speedup
GWD	7.0	13.5	0.52	39.3	72.9	0.54
TURB	7.2	14.4	0.50	51.5	84.7	0.61
CLOUD	15.3	55.2	0.28	88.8	317.3	0.28
LW Chou-Suarez	64.0	32.0	2.01	221.9	120.0	1.85
SW Chou-Suarez	12.3	8.9	1.38	24.2	16.1	1.50
LW RRTMG	102.4	29.1	3.52	207.8	47.9	4.34
SW RRTMG	30.9	10.0	3.07	56.1	16.0	3.50

	50-km	25-km	12-km	6-km	3-km
FV3 (transport)	1.1	1.8	3.4	5.4	7.2

Single node GPU speed over CPU with increasing problem size.
This data is quite old, no active development on GPUs with GEOS.

Machine Learning and Artificial Intelligence

NASA is undertaking a number of activities to increase our corporate knowledge and enhance our ML capabilities across disciplines.

1. Exploring a wide range of science and engineering applications

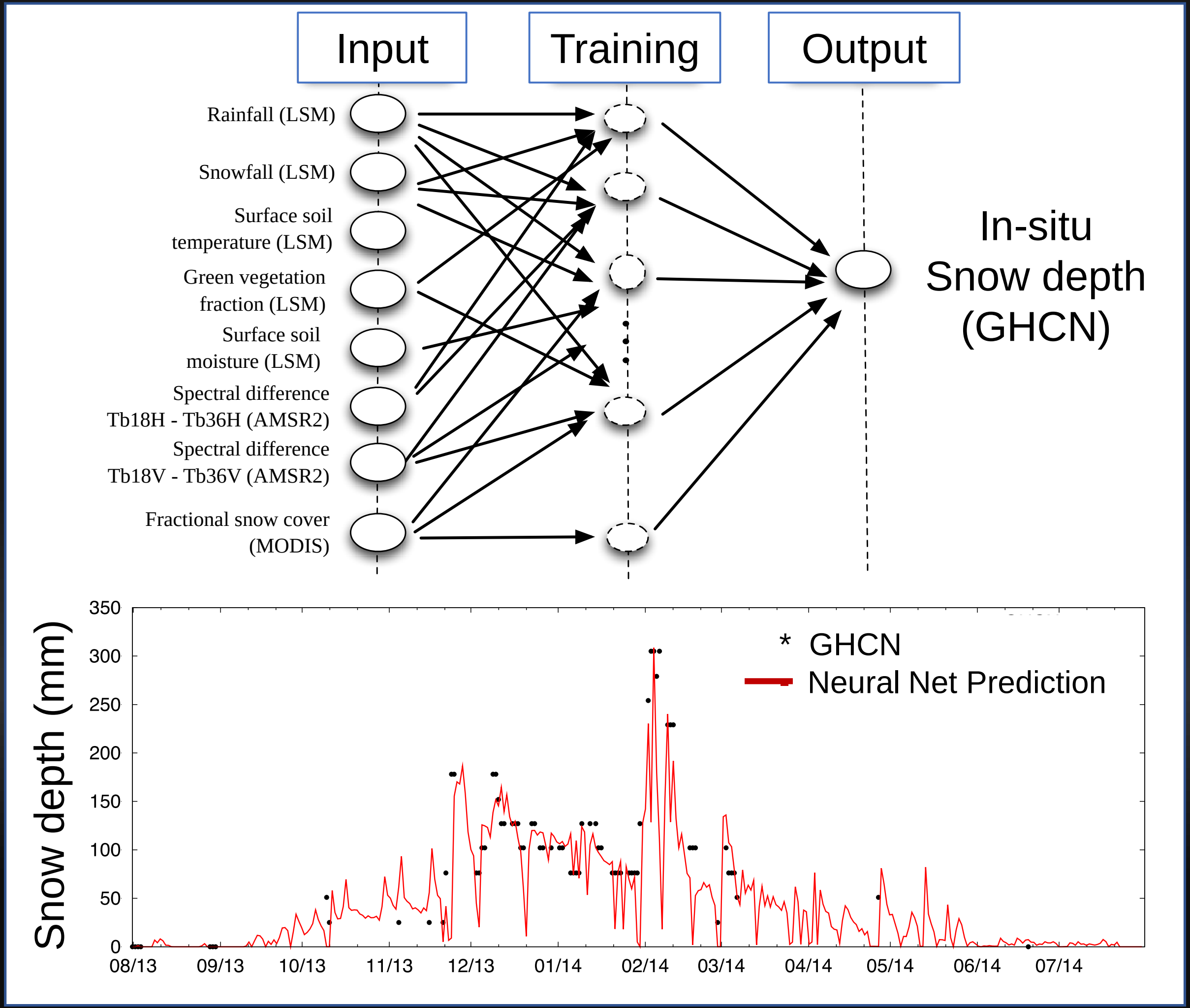
- Land Information Systems (LIS) data assimilation
- Improving the skill of remotely sensed observation retrievals using deep learning (e.g., snow water equivalent)
- Trained model components for Goddard Earth Observing System
- Advanced Hyperspectral Imaging through Integrated Compressive Sensing/Inpainting via Machine Learning
- Counting trees from space in high resolution imagery

2. Goddard Strategic AI Group

- *NASA Goddard Workshop on Artificial Intelligence*
  - *Dates: Nov. 27-29, 2018*

3. Increasing ML capabilities at Goddard

- Specialized computing resources designed for machine learning
- Currently have a small number of systems
- Expect significant demand for more capability in the near future



LIS data assimilation ML snow-depth example

Summary

Can the compute capacity at NASA keep pace with the science demands?

- The science demands of GEOS (in particular aerosols and clouds) will quickly exceed the compute capacity at NCCS
- Optimization efforts are ongoing
- Flexibility in modeling infrastructure means science capability and development often times exceed our optimization capacity to support these enhancements
- Joint development efforts with NOAA and the Unified Forecast System community are beneficial for model and DA developments
- Similar joint efforts are required to address the scalability and exascale capability of our systems