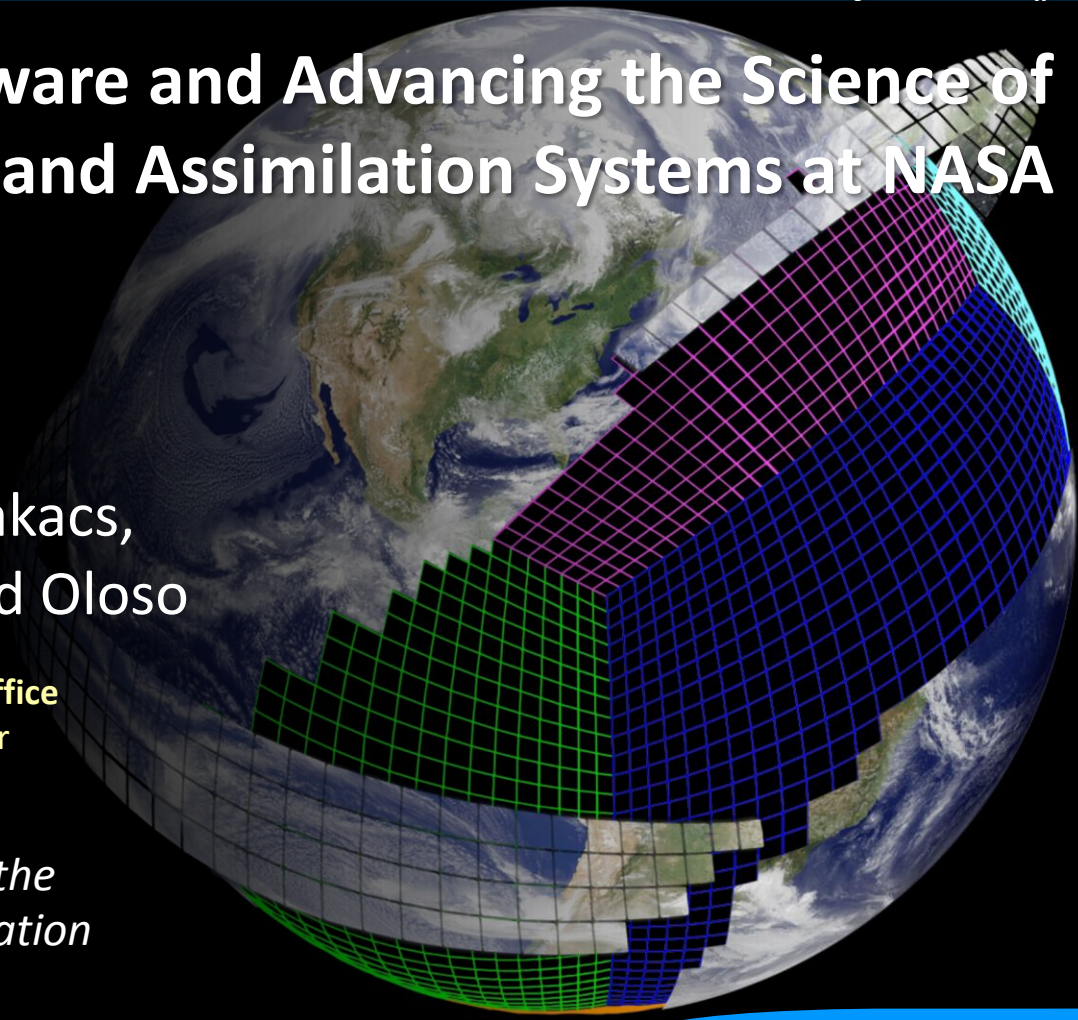
Scaling the Software and Advancing the Science of Global Modeling and Assimilation Systems at NASA

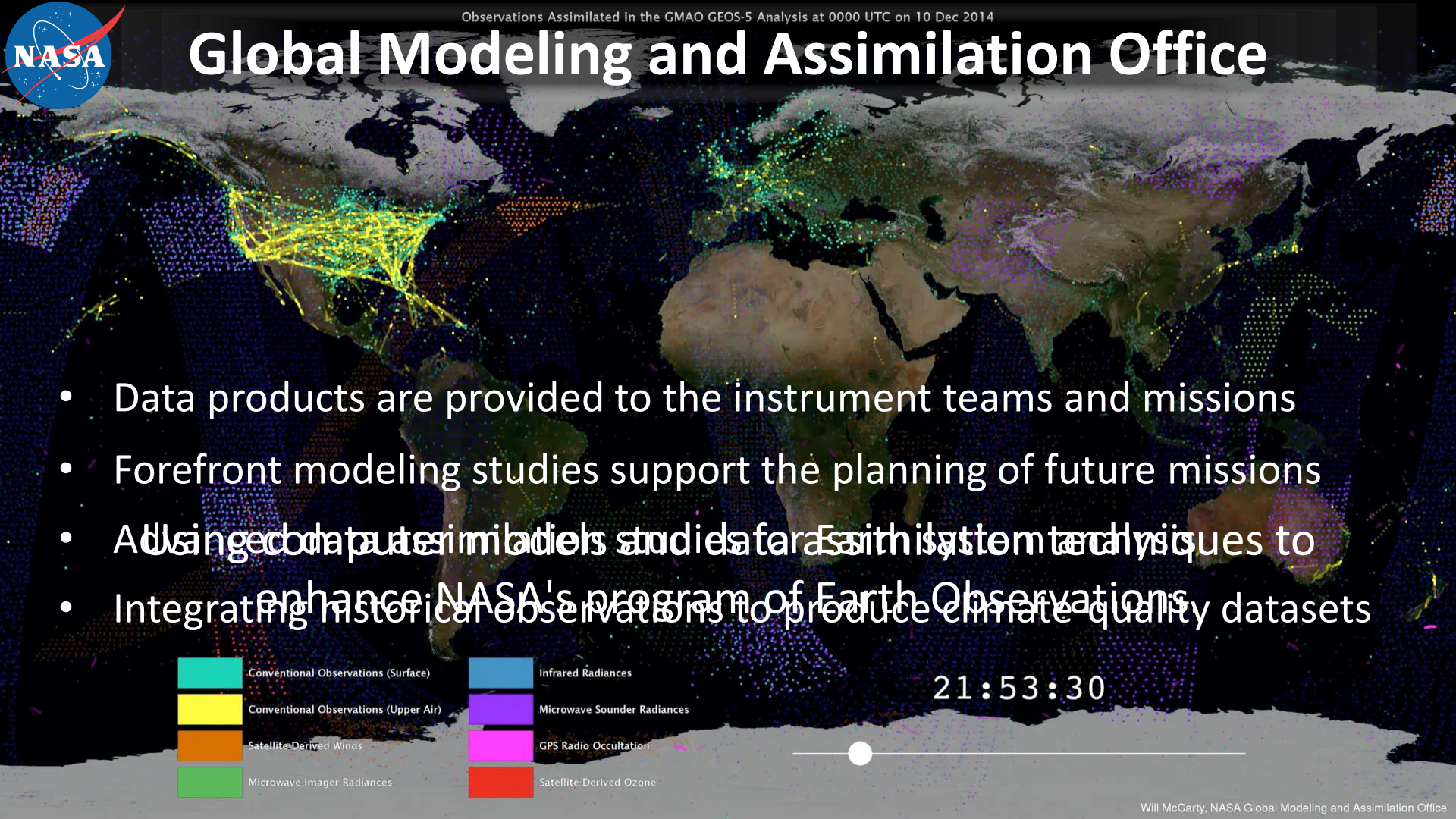


Bill Putman

Max Suarez, Lawrence Takacs,
Atanas Trayanov and Hamid Oloso

Global Modeling and Assimilation Office
NASA Goddard Space Flight Center

with extensive support from the NASA Center for Climate Simulation



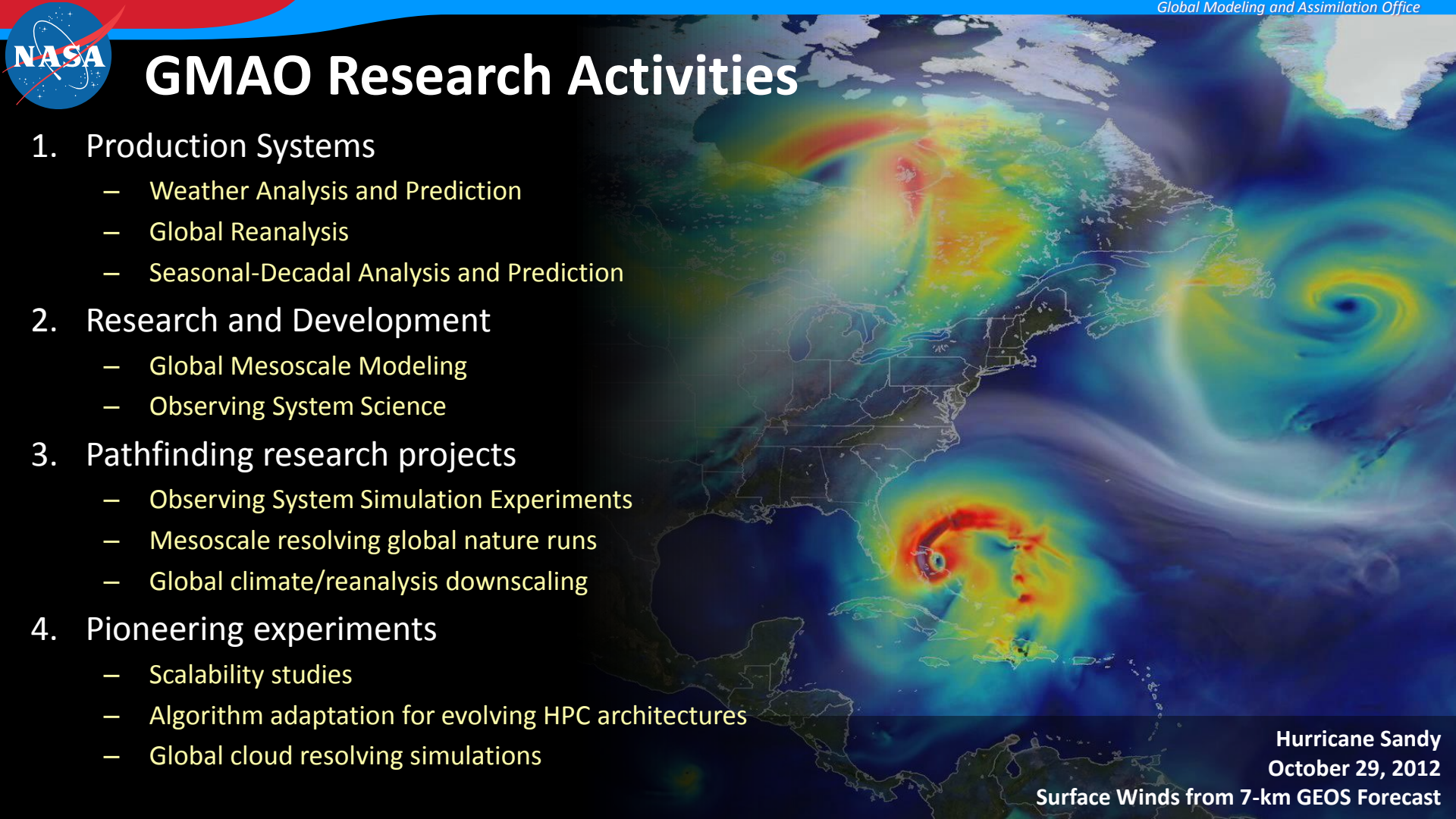
Global Modeling and Assimilation Office

- Data products are provided to the instrument teams and missions
- Forefront modeling studies support the planning of future missions
- Advanced data assimilation studies for Earth system analysis
- Integrating historical observations to produce climate-quality datasets

Using computer models and data assimilation techniques to enhance NASA's program of Earth Observations

GMAO Research Activities

1. Production Systems
   - Weather Analysis and Prediction
   - Global Reanalysis
   - Seasonal-Decadal Analysis and Prediction
2. Research and Development
   - Global Mesoscale Modeling
   - Observing System Science
3. Pathfinding research projects
   - Observing System Simulation Experiments
   - Mesoscale resolving global nature runs
   - Global climate/reanalysis downscaling
4. Pioneering experiments
   - Scalability studies
   - Algorithm adaptation for evolving HPC architectures
   - Global cloud resolving simulations



Hurricane Sandy
October 29, 2012
Surface Winds from 7-km GEOS Forecast

GEOS Atmospheric Model

Finite Volume Cubed-Sphere *(FV3 collaboration with NOAA GFDL)*

- Split explicit time-stepping *(C-D grid)*
- Monotone advection
- Vertically Lagrangian
- Hydrostatic or Non-Hydrostatic

Single-Moment or ***Two-Moment*** moist cloud microphysics options

Relaxed Arakawa Schubert Convection *(stochastic Tokioka entrainment limiter)*

- Grell Freitas and UW shallow convection options

Chou Suarez or ***RRTMG*** radiation options

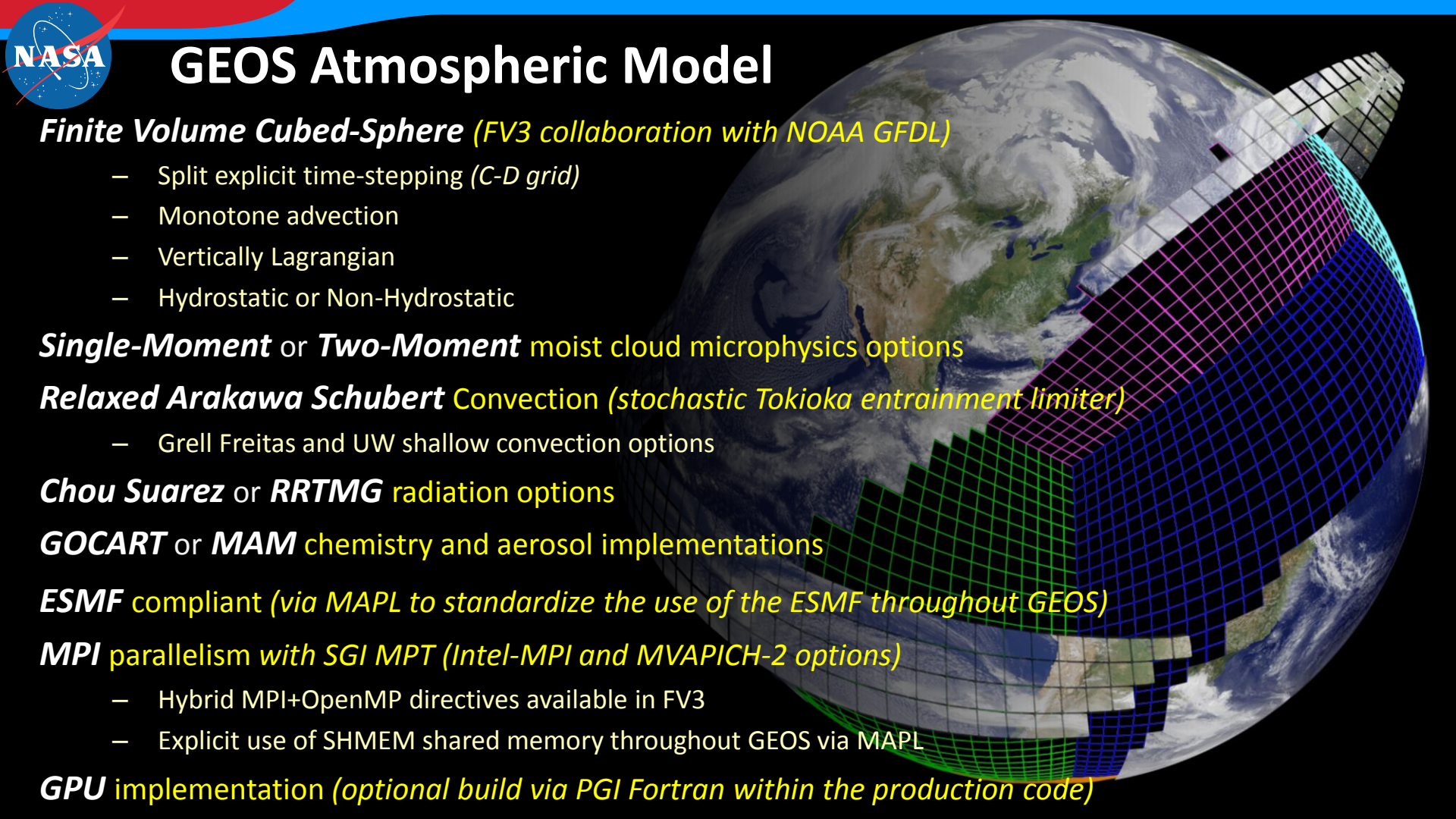
GOCART or ***MAM*** chemistry and aerosol implementations

ESMF compliant *(via MAPL to standardize the use of the ESMF throughout GEOS)*

MPI parallelism *with SGI MPT (Intel-MPI and MVAPICH-2 options)*

- Hybrid MPI+OpenMP directives available in FV3
- Explicit use of SHMEM shared memory throughout GEOS via MAPL

GPU implementation *(optional build via PGI Fortran within the production code)*



GMAO Core Production Suite with GEOS

Modeling Component	Current Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*X86 Cores*	*Resolution*	*X86 Cores*	*Resolution*	*X86 Cores*
Deterministic Medium-Range	10-15 km *72 Levels*	6,000	3-7 km *132 Levels*	600,000	1-3 km *200 Levels*	6 million

2016 *Goddard Earth Observing System (GEOS)*

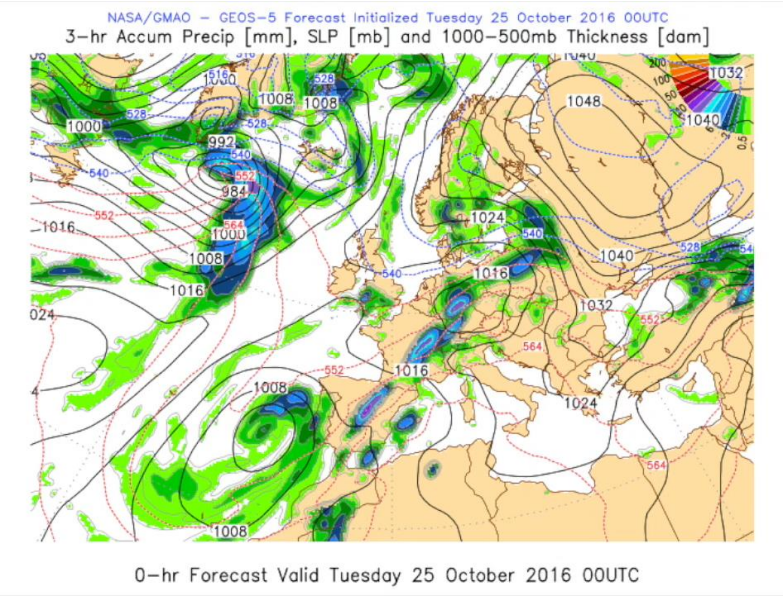
Model Production Deterministic Configuration

25-km with 72-levels on 840 Xeon Haswell Cores

3-D Hybrid Ensemble-Var Data Assimilation

Hydrostatic FV3 Dynamics

GOCART Aerosols (47 tracers)



GMAO Core Production Suite with GEOS

Modeling Component	Current Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*X86 Cores*	*Resolution*	*X86 Cores*	*Resolution*	*X86 Cores*
Deterministic Medium-Range	10-15 km *72 Levels*	6,000	3-7 km *132 Levels*	600,000	1-3 km *200 Levels*	6 million

2016 *Goddard Earth Observing System (GEOS)*

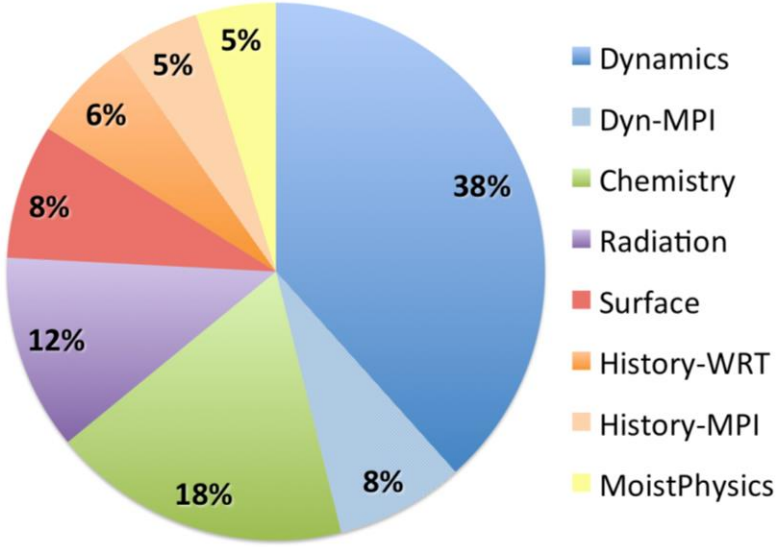
Model Parallel Deterministic Configuration

12.5-km with 72-levels on 5400 Xeon Haswell Cores

4-D Hybrid Ensemble-Var Data Assimilation

Hydrostatic FV3 Dynamics

GOCART Aerosols (47 tracers)



GMAO Global Reanalyses for Climate

Modeling Component	Current Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*X86 Cores*	*Resolution*	*X86 Cores*	*Resolution*	*X86 Cores*
Climate/Reanalysis	50 km	100	10-25 km	8,000	5-10 km	100,000

Modern-Era Retrospective Analysis for Research and Applications

[MERRA-2 1980-present]

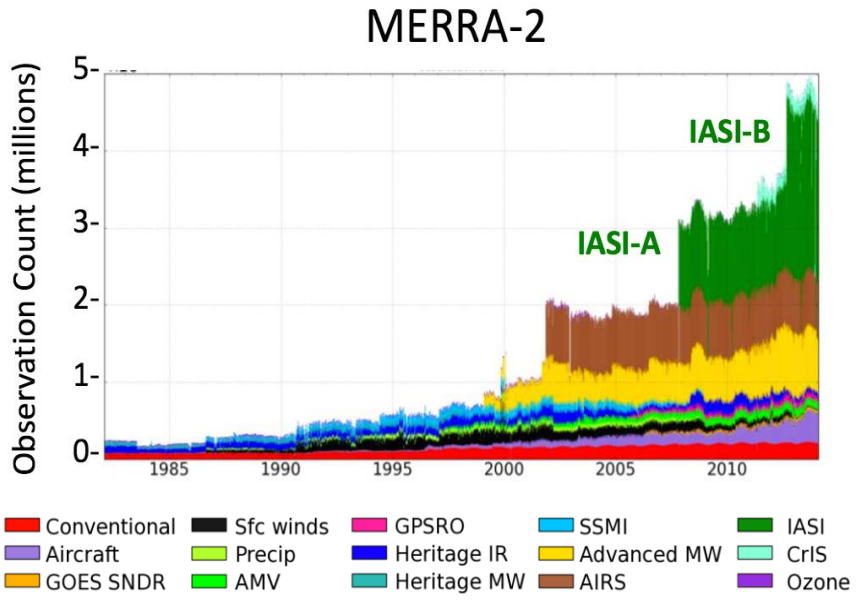
Production Configuration

50-km with 72-levels on 100 Xeon Haswell Cores

3-D Variational Data Assimilation

Hydrostatic FV3 Dynamics

GOCART Aerosols including CO SO_2 and O_3 (~50 tracers)



GMAO Seasonal Forecasting System

Modeling Component	Current Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*X86 Cores*	*Resolution*	*X86 Cores*	*Resolution*	*X86 Cores*
Seasonal	50 km Ocn 50 km Atm	100	25 km Ocn 25 km Atm	800	10 km Ocn 10 km Atm	16,000

MERRA-2 Driven Ensemble-OI Ocean Analysis

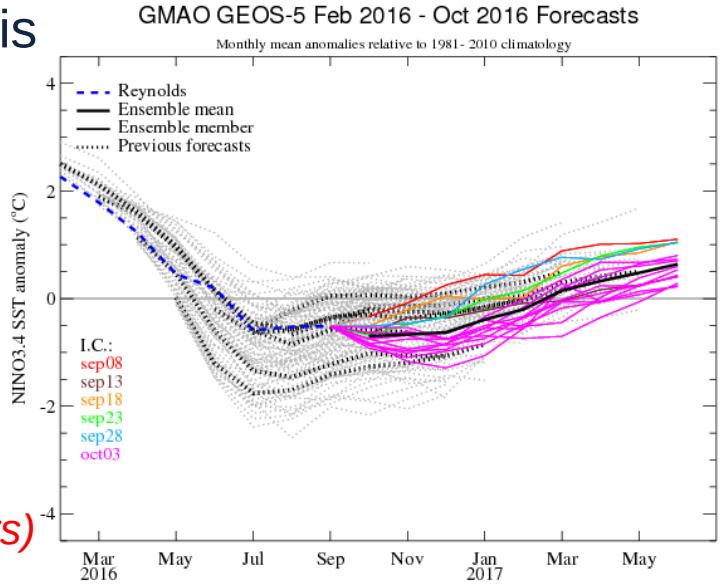
Production Configuration part of *NMME*

50-km 72-lev Atmosphere

50-km 40-Lev Ocean
on 100 Xeon Haswell Cores

31-ensemble members per month

GOCART Aerosols including CO SO_2 and O_3 *(~50 tracers)*



GMAO Seasonal Forecasting System

Modeling Component	Current Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*X86 Cores*	*Resolution*	*X86 Cores*	*Resolution*	*X86 Cores*
Seasonal	50 km Ocn 50 km Atm	100	25 km Ocn 25 km Atm	800	10 km Ocn 10 km Atm	16,000

MERRA-2 Driven Ensemble-OI Ocean Analysis

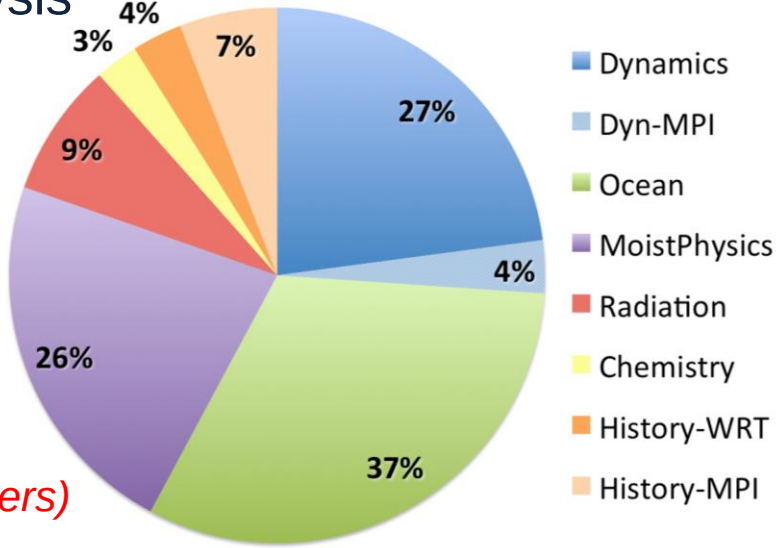
Production Configuration part of *NMME*

50-km 72-lev Atmosphere

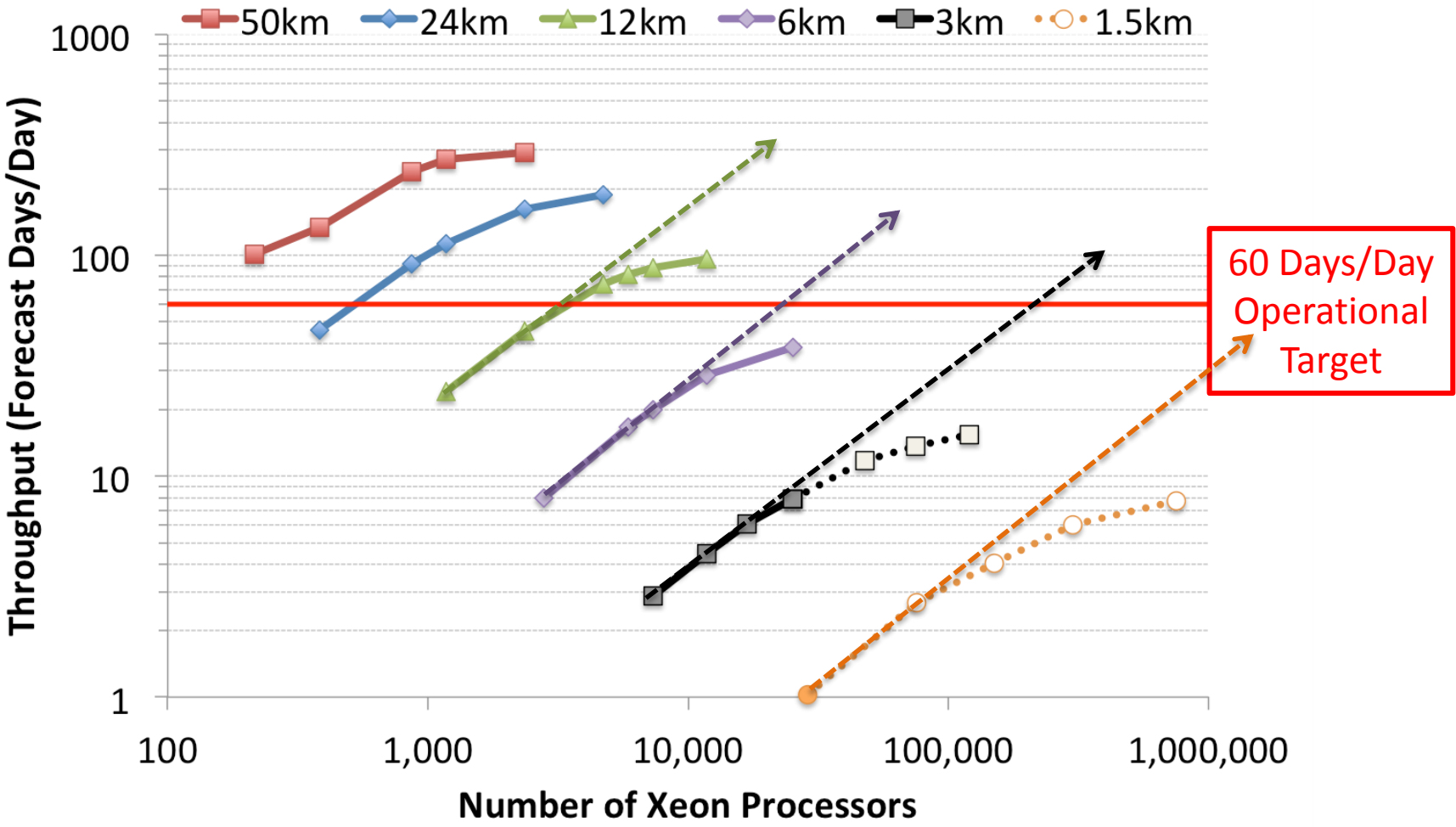
50-km 40-Lev Ocean
on 100 Xeon Haswell Cores

31-ensemble members per month

GOCART Aerosols including CO SO_2 and O_3 (~50 tracers)

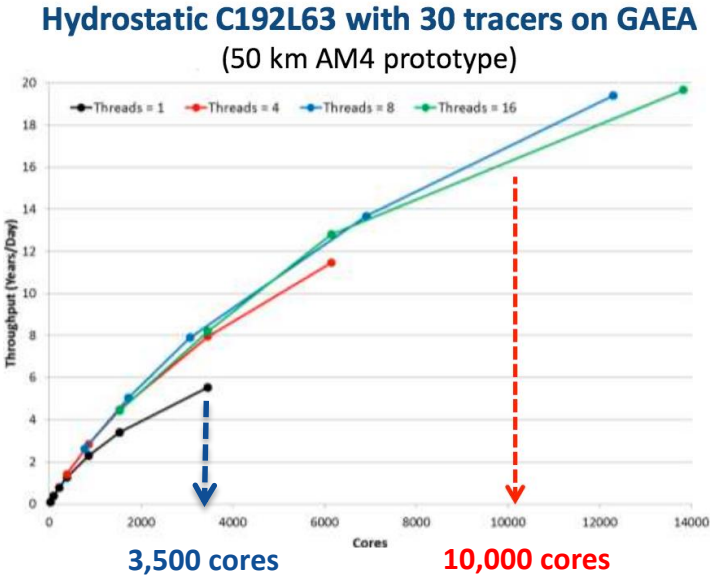


GEOS (72 Levels) MPI Scaling on Discover at NCCS



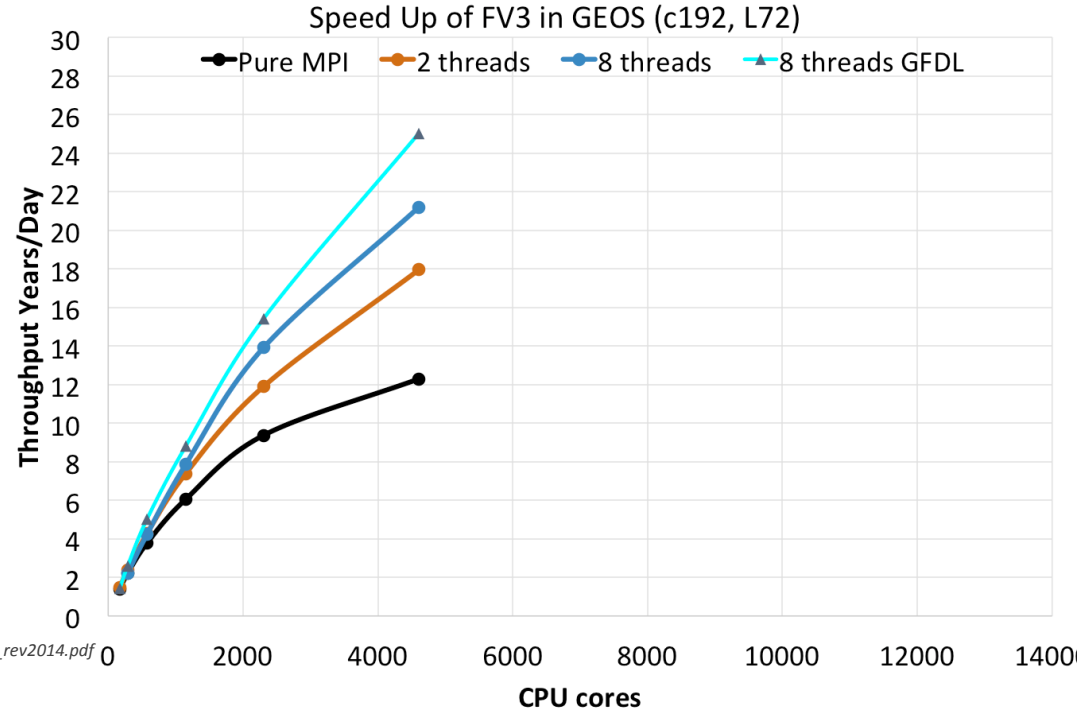
Hybrid MPI+OpenMP in GEOS

FV3 Scaling at NOAA-GFDL



http://www.gfdl.noaa.gov/cms-filesystem-action/administrative/2014review/presentations/02.lin_gfdl_rev2014.pdf

FV3 Scaling at NASA-NCCS



GPU Evaluation of GEOS Physics

C360 (25-km) – 16 CPU Cores (Xeon Sandybridge) v 1 GPU (K40)						
	72 Levels			132 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**

- Components with ***low computation-to-data-movement*** suffer (GWD, TURB, CLOUD)
- Data movement on/off the GPUs on entry/exit of each physics component hinders performance
- To truly exploit the efficiency of the GPUs, ***all of the physics would need to reside on GPU***

GEOS Parallel History Asynchronous I/O Server

GCM Communicator

400 nodes

with SHMEM

2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data
2d Slice of data

Horizontal Regridding
Bit Shaving

Horizontal Regridding
Bit Shaving

Horizontal Regridding
Bit Shaving

Horizontal Regridding
Bit Shaving

Horizontal Regridding
Bit Shaving

AllToMany 2D-Gather

Node-to-Node with SHMEM

ManyToFew

IO Communicator

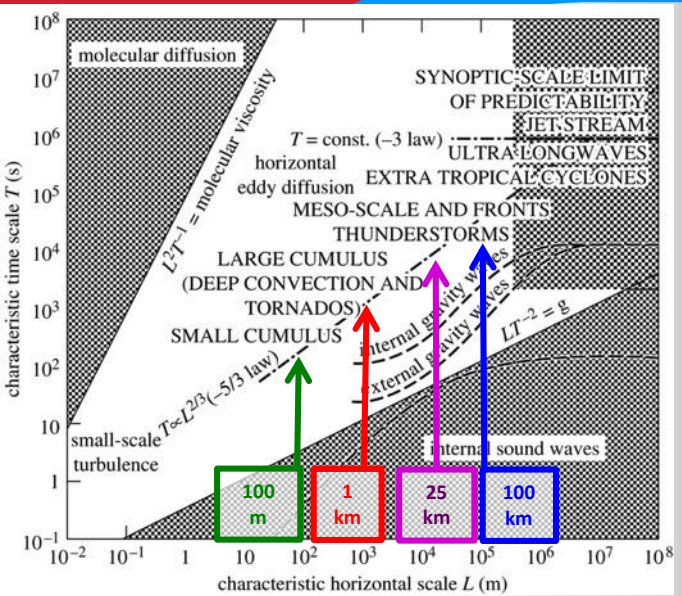
3 nodes

Worker Nodes
Stores Slices
Chunk and Compress
Write Collections

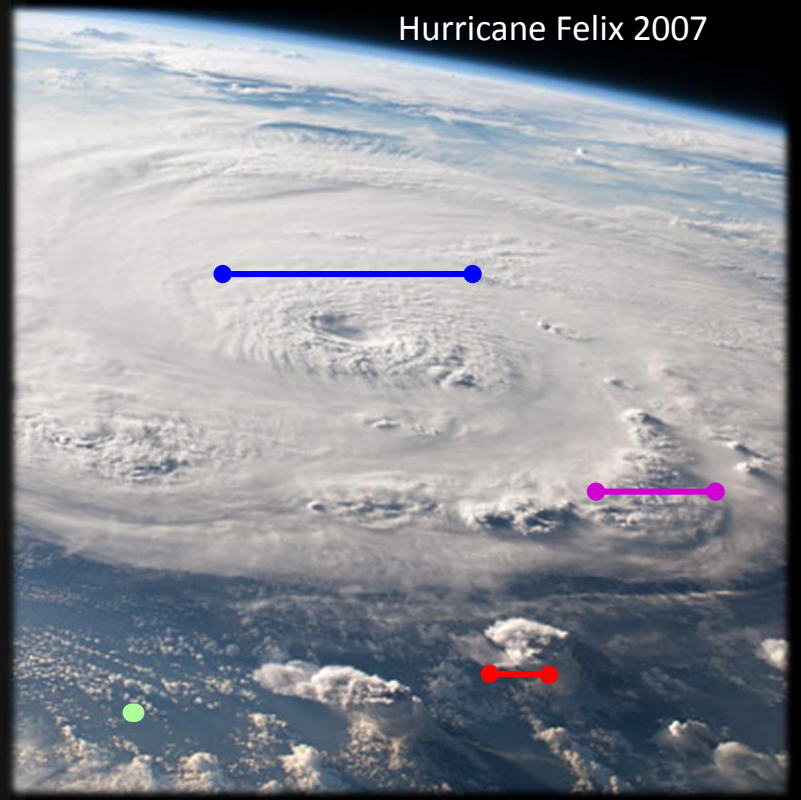
Worker Nodes
Stores Slices
Chunk and Compress
Write Collections

Multiple Writers

Science and Computing Required to Increase Resolvable Scales



Resolution (km)	Resolvable ~7x (km)	Computing (X86 Cores)
25.0	200	800
12.5	100	6,400
3.0	20	462,963
0.1	1	6,400,000,000
10 (m)	100(m)	21,600,000,000,000,000



GMAO Pathfinding/Pioneering Projects

Modeling Component	Current Capability		Year 2020-2025		Year 2025-2030	
	Resolution	*X86 Cores*	*Resolution*	*X86 Cores*	*Resolution*	*X86 Cores*
Embedded-Regional Pioneering Global OSSEs	1-10 km *72 Levels*	30,000	500 m – 1 km *132 Levels*	10 million	100-500 m *250 Levels*	1 billion

12-km Full Chemistry Global Nature runs for Climate OSSEs

7-km Global Nature Run (G5NR, 2-Years) for OSSEs

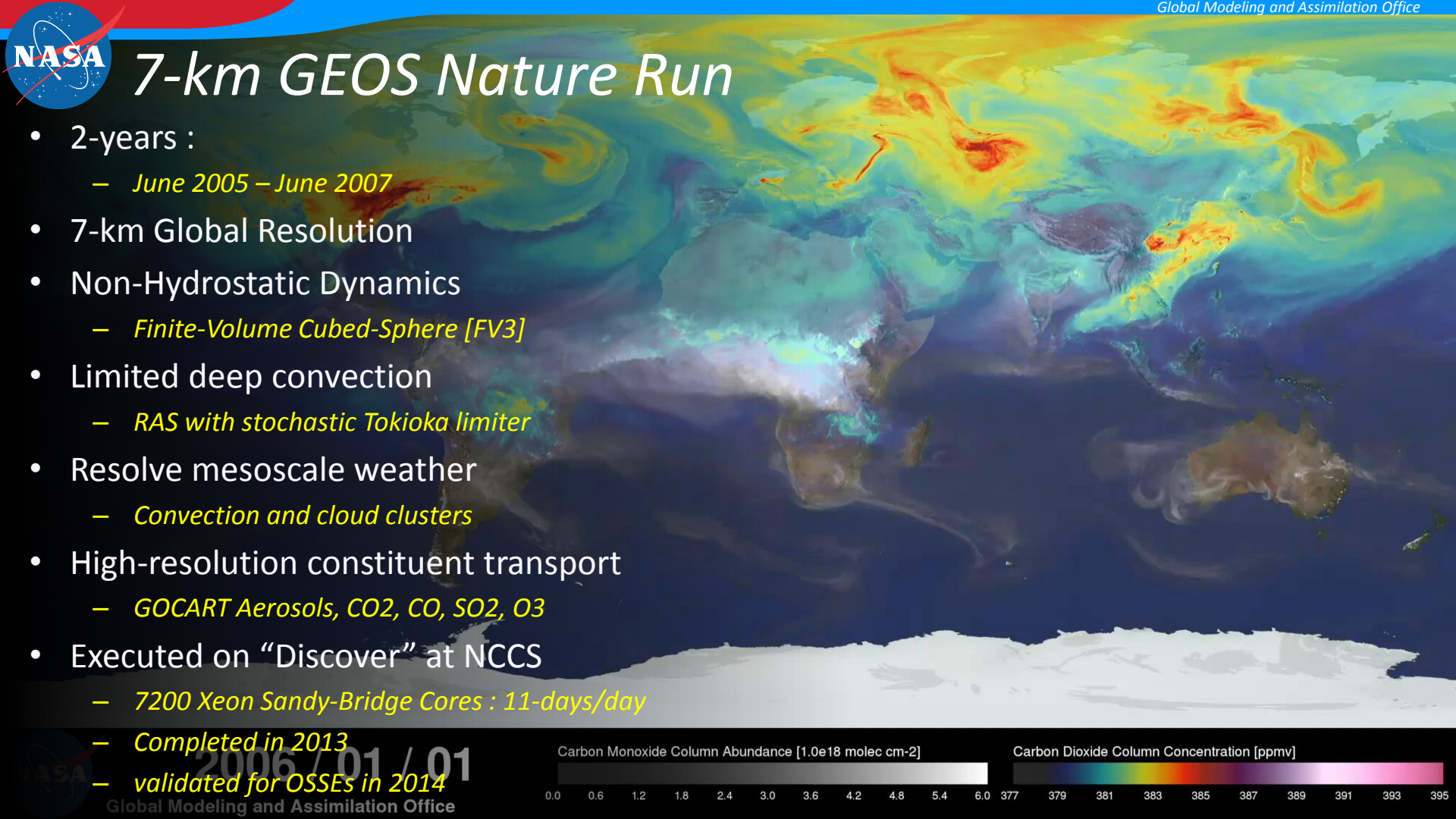
1.5-km Pioneering Global Cloud Permitting Simulations

'Downscaling' Replays of Reanalysis/Climate Simulations

With an emphasis on informing decisions for the NASA decadal survey to prepare for future observing systems and new instruments through OSSEs.

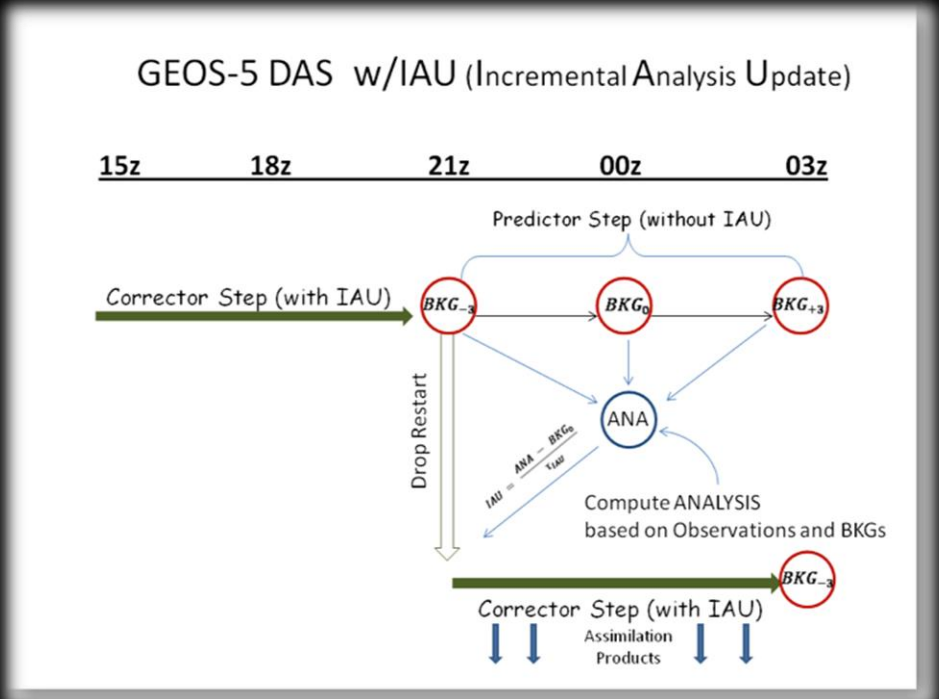
7-km GEOS Nature Run

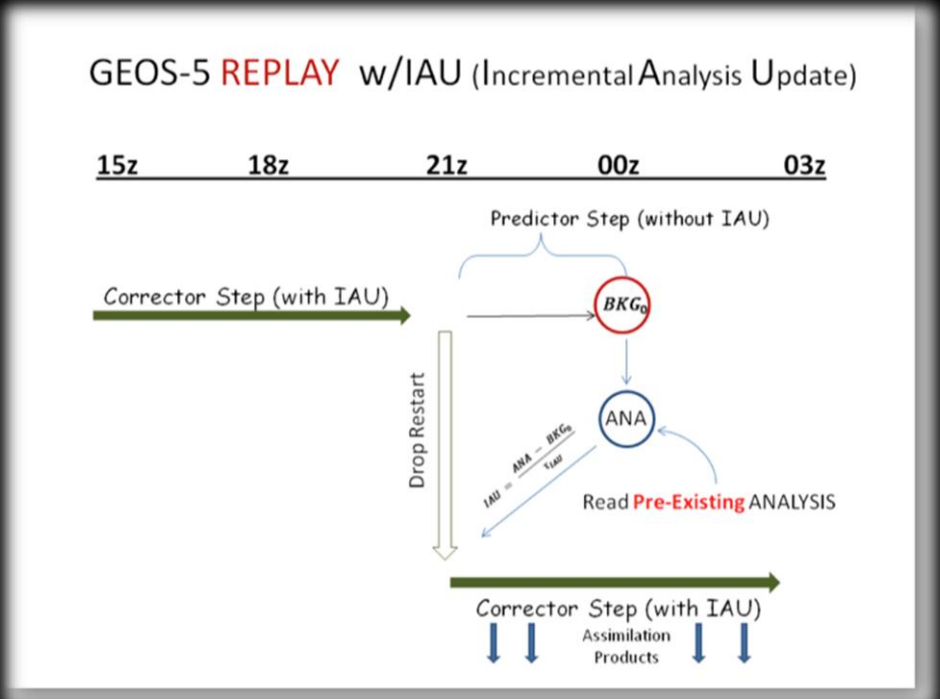
- 2-years :
  - *June 2005 – June 2007*
- 7-km Global Resolution
- Non-Hydrostatic Dynamics
  - *Finite-Volume Cubed-Sphere [FV3]*
- Limited deep convection
  - *RAS with stochastic Tokioka limiter*
- Resolve mesoscale weather
  - *Convection and cloud clusters*
- High-resolution constituent transport
  - *GOCART Aerosols, CO2, CO, SO2, O3*
- Executed on “Discover” at NCCS
  - *7200 Xeon Sandy-Bridge Cores : 11-days/day*
  - *Completed in 2013*
  - *validated for OSSEs in 2014*



GMAO Pathfinding Projects

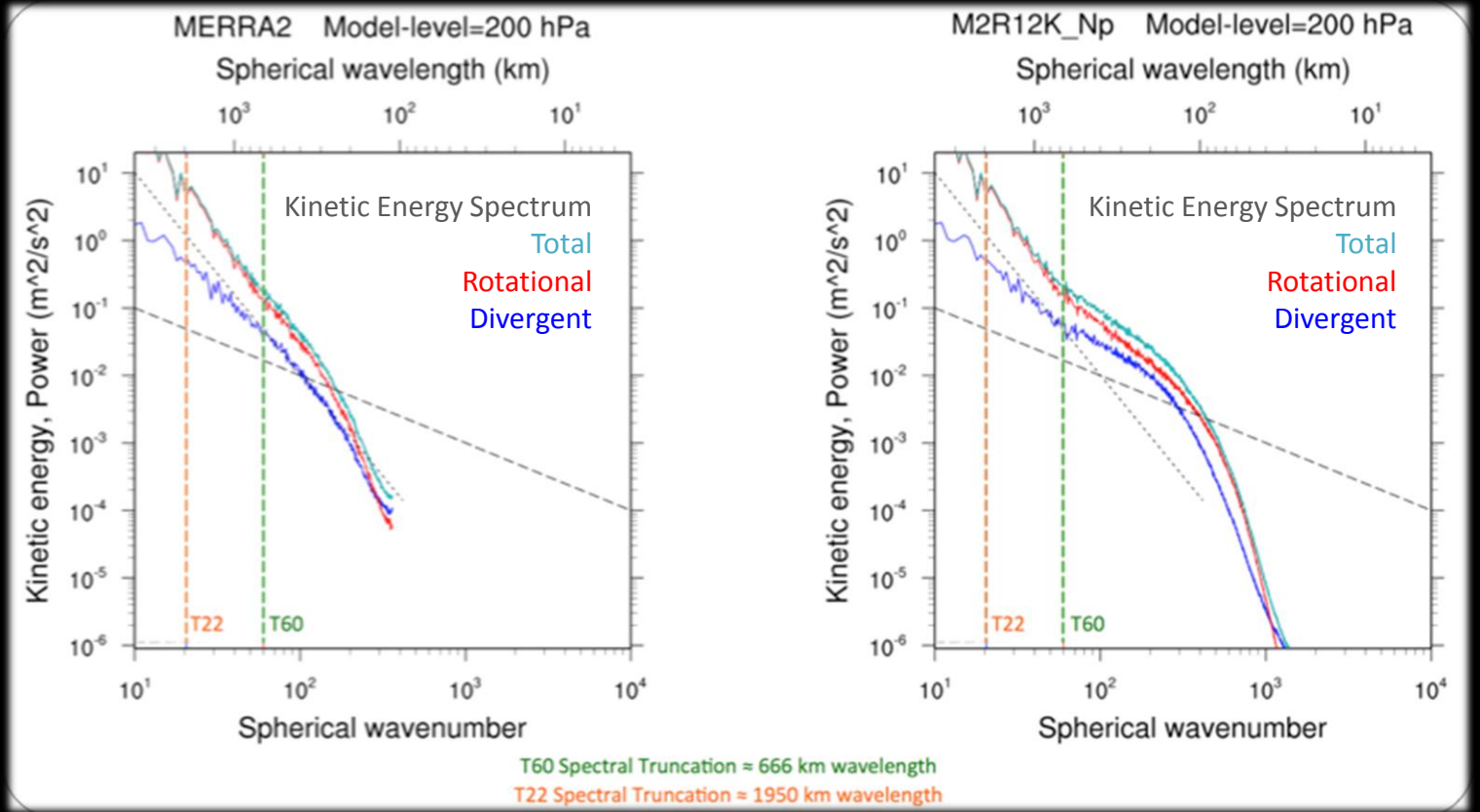
A '*Replay*' Capability for Global Reanalysis/Climate Downscaling





GMAO Pathfinding Projects

The *MERRA-2 Replay at 12.5-Km [M2R12K]*



The *MERRA-2 Replay at 12.5-Km [M2R12K]*

MERRA-2 'downscaled' to 12.5 km
16 Years: 2000-2015

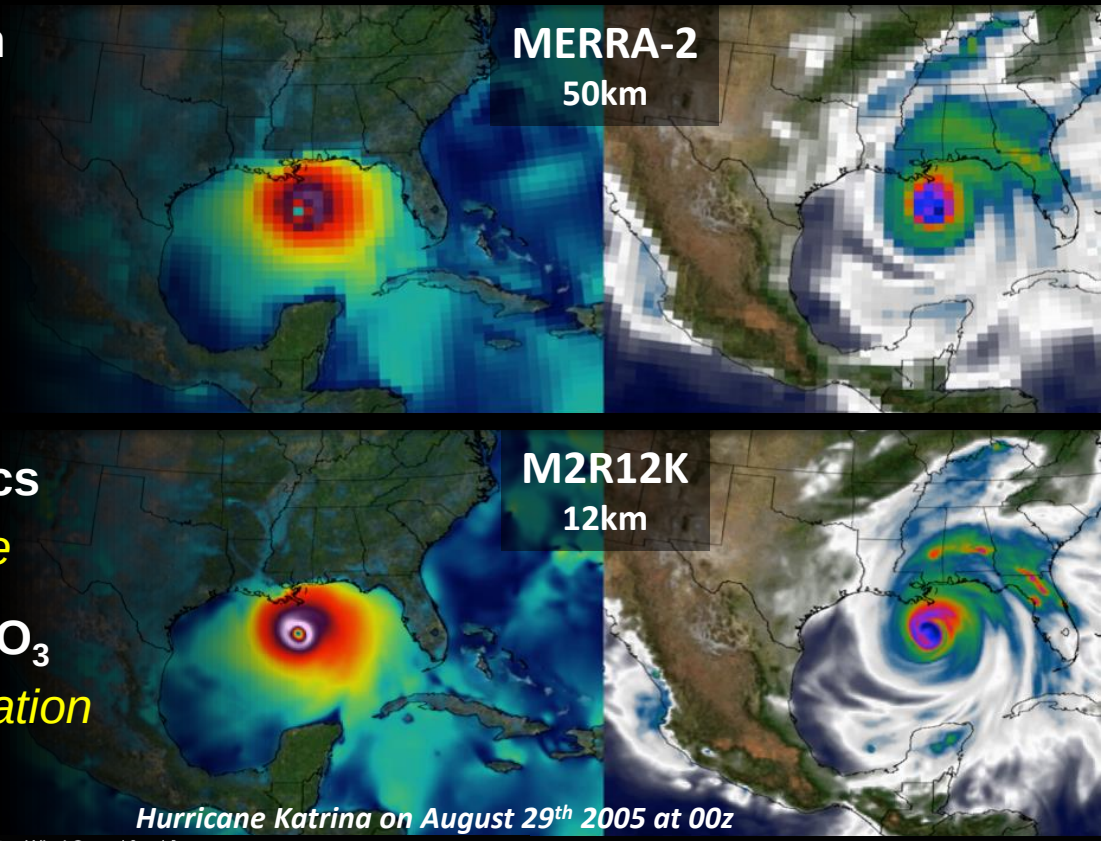
4600 Xeon Haswell Cores
3 weeks/day in two streams

Non-Hydrostatic FV3 Dynamics
4th order divergence damping

Single-Moment Cloud Microphysics
Repartitioning of resolved/convective

Includes Aerosols, CO_2, CO, SO_2, O_3
Running with GAAS Aerosol Assimilation

Produced more then 4 PB of data



The *MERRA-2 Replay at 12.5-Km [M2R12K]*

MERRA-2 'downscaled' to 12.5 km
16 Years: 2000-2015

4600 Xeon Haswell Cores
3 weeks/day in two streams

Non-Hydrostatic FV3 Dynamics
4th order divergence damping

Single-Moment Cloud Microphysics
Repartitioning of resolved/convective

Includes Aerosols, CO_2, CO, SO_2, O_3
Running with GAAS Aerosol Assimilation

Produced more then 4 PB of data

6-km GEOS Replay

Run on 10,800 Intel Xeon Haswell cores

1. Replaying our Production Analysis at 6km
2. Daily 10-day forecasts at 00z
3. Non-Hydrostatic FV3
4. Two-Moment Microphysics
5. Supported the ORACLES science team
   - ObseRvations of Aerosols above CLouds and their intEractionS
   - designed to study key processes that determine the climate impacts of African biomass burning aerosols

September 19, 2016 forecast
Aerosols including carbon, dust, sulfate

1.5-km Global GEOS

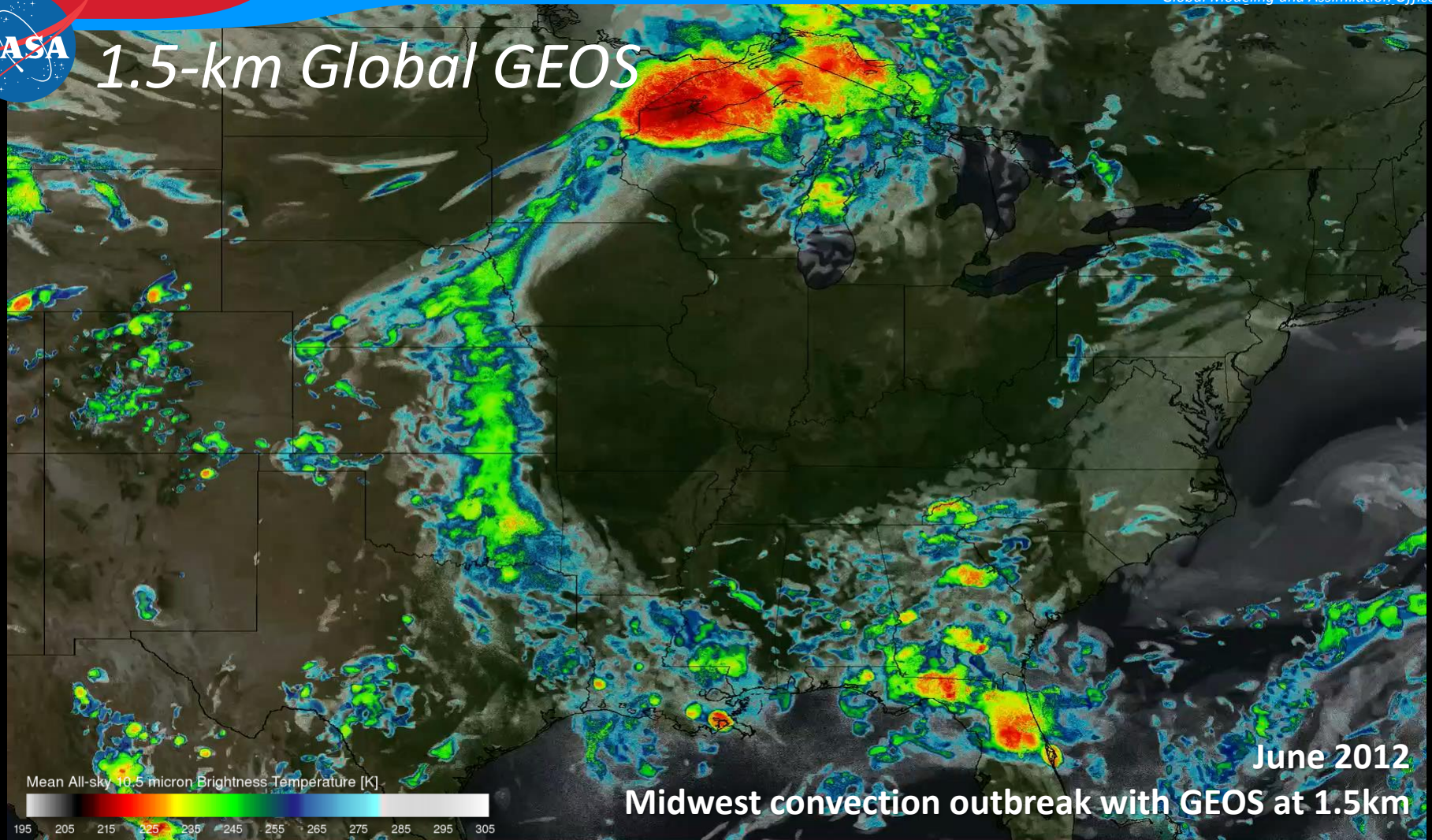
Run on 30,240 Intel Xeon Haswell cores

1. GEOS global state variables ~5 terabytes
2. 1.5-km GEOS uses ~110 of 128 GB per node
3. ESMF initialization scaled badly *(memory)*
   - 3 GB per node for each component
   - 22% overhead (40 GB used out of 128 GB avail)
   - Removed unneeded components *(ESMF resolved)*
4. Offload any I/O
   - GOES I/O server manages worker nodes
     - regriding, chunking and compression
     - Asynchronous writing *(model continues execution)*
   - 3-days produced >60 TB of output at 10-minute intervals

June 2012

sulfate/carbon/seasalt aerosols in GEOS at 1.5km

1.5-km Global GEOS

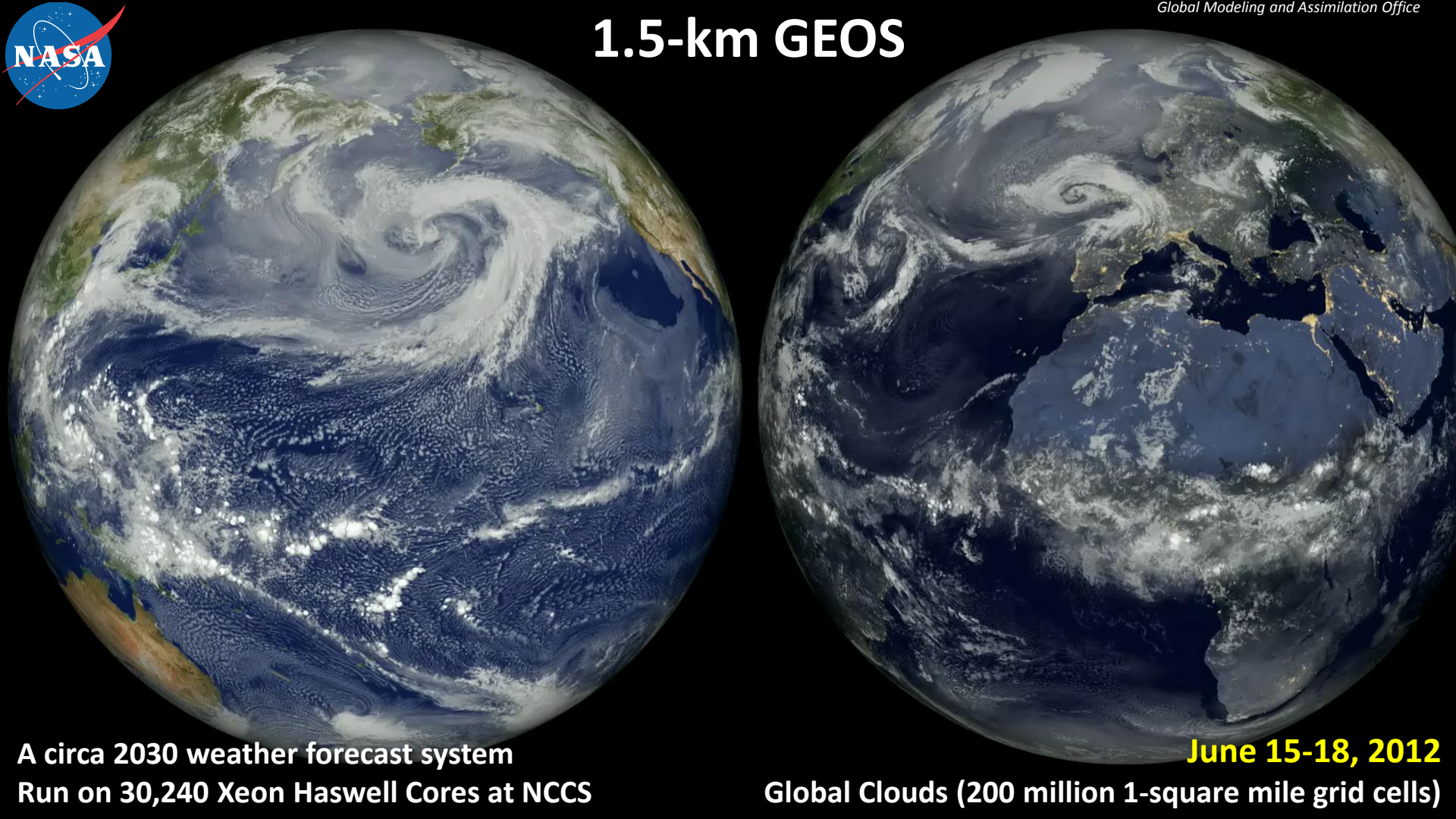


June 2012

Midwest convection outbreak with GEOS at 1.5km

Global Modeling and Assimilation Office

1.5-km GEOS



A circa 2030 weather forecast system
Run on 30,240 Xeon Haswell Cores at NCCS

June 15-18, 2012
Global Clouds (200 million 1-square mile grid cells)