National Aeronautics and
Space Administration

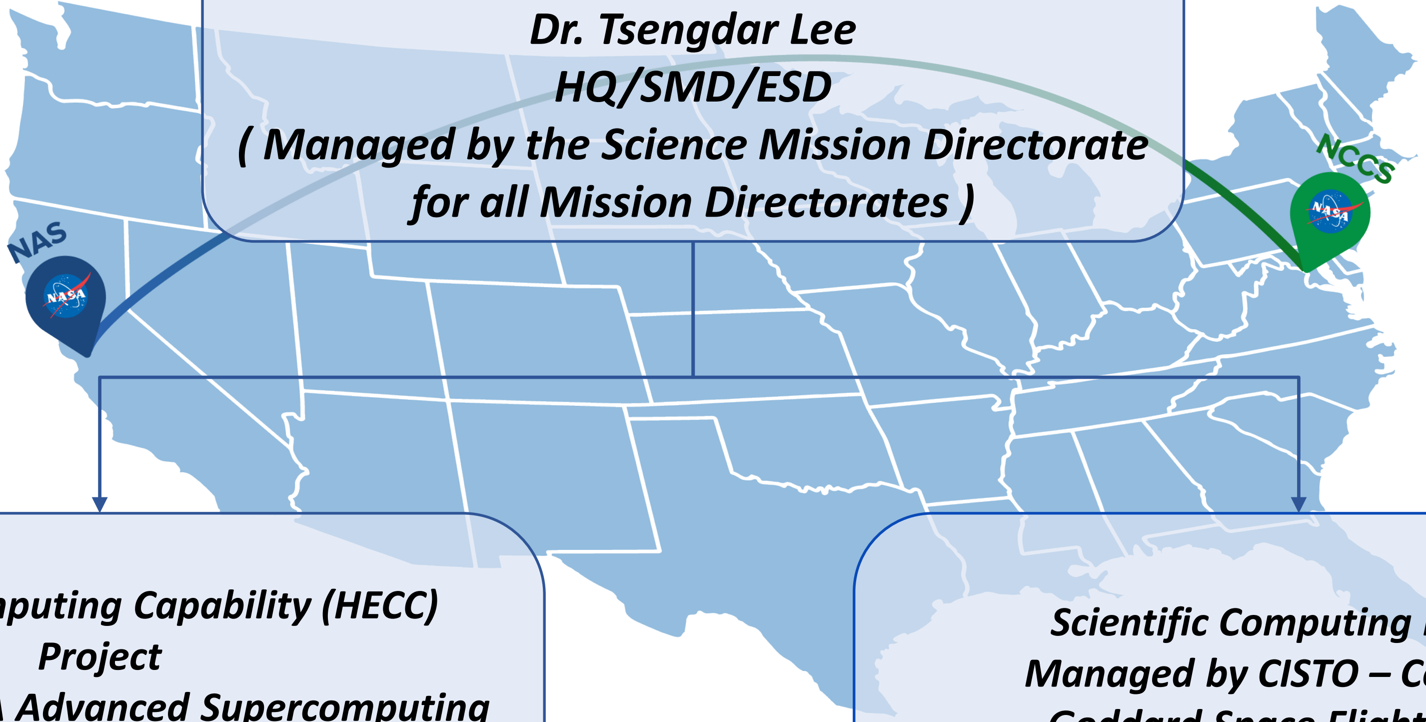
Three Decades of Supercomputing at NASA

Supercomputing for Space

Kevin Murphy, NASA Chief Science Data Officer
Tsengdar Lee, NASA HQ
Dr. Piyush Mehrotra, Ames Research Center
William Thigpen, Ames Research Center
Dr. Daniel Duffy, Goddard Space Flight Center
Laura Carriere, Goddard Space Flight Center

High End Computing (HEC) Portfolio Organization

NASA High End Computing (HEC) Portfolio
Dr. Tsengdar Lee
HQ/SMD/ESD
(Managed by the Science Mission Directorate for all Mission Directorates)



High End Computing Capability (HECC) Project
Hosted by: NASA Advanced Supercomputing (NAS) Division,
Ames Research Center

Dr. Piyush Mehrotra, William Thigpen

Scientific Computing Project
Managed by CISTO – Code 606,
Goddard Space Flight Center
Primarily Science Funded

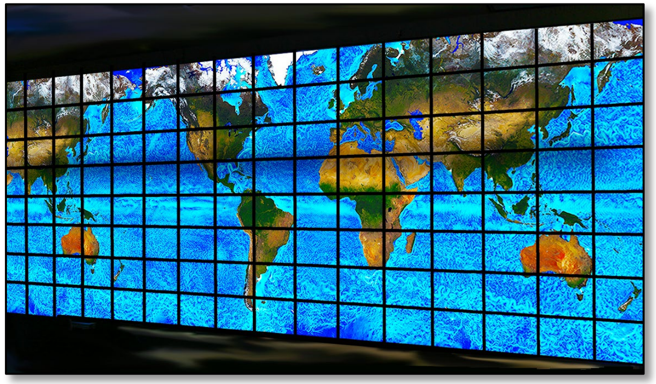
Dr. Daniel Duffy, Laura Carriere

High-End Computing Capability (HECC) Project

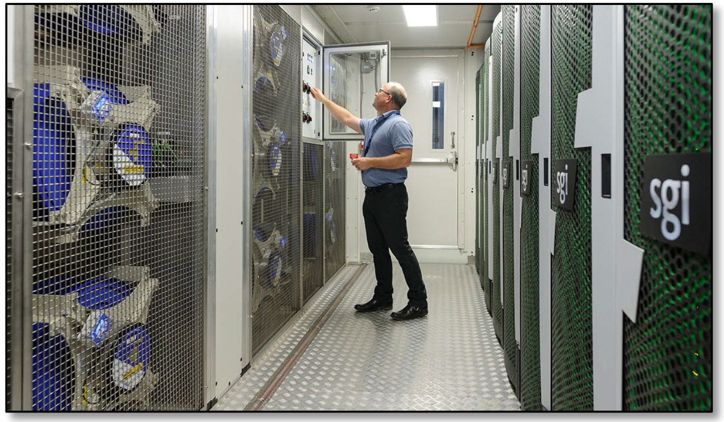
Pleiades: 9.2 PF;
11,141 CPU-nodes; 121 GPU-nodes

Integrated Support Services

Aitken: 8.4 PF; 2152 nodes



Hyperwall; 245M pixels
128 screens (23x10 ft.)



Electra: 8.32 PF; 3,456 nodes

Facilities

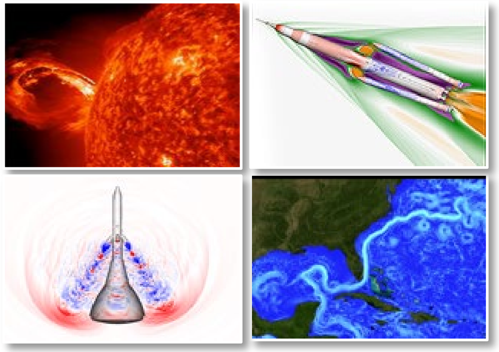
Capabilities at NASA Ames Supercomputing Center

HECC Hardware Assets

5 Compute Clusters

- Aitken 20 Racks / 3,200 nodes / 13.1 PF / 10,204 SBU/hr
- Pleiades 155 ½ Racks / 11,196 nodes / 7.12 PF / 8,020 SBU/hr
- Electra 24 Racks / 3,456 nodes / 8.32 PF / 4,815 SBU/hr
- Cabeus 10 Racks / 57 nodes / 2.08 PF / 1,595 SBU/hr
- Endeavour 2 Racks / 2 nodes / 155 TF / 81.2 SBU/hr

1 Visualization Cluster	245-million-pixel display / 128 nodes / 2.61 PF
10 Lustre File Systems	88.4 PB
3 NFS File Systems	3.3 PB
2 BeeGFS File Systems	1.5 PB
1 VAST File System	8.9 PB
Archive System	1,000 PB



NAS Facility Extension

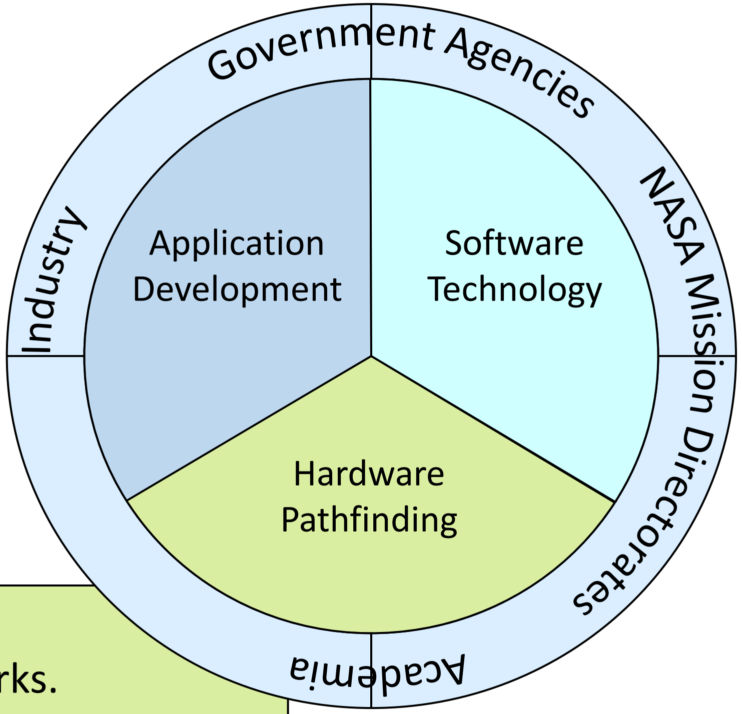
- One-acre 30 MW site to house HPC systems in modules.

HECC provides a suite of complimentary services to the user community to enhance the scientific and engineering results obtained from the hardware assets.

- Systems: Customized solutions including compute and storage solutions to meet specific project or mission requirements. Cloud access for immediate or non-standard computing.
- Cloud Resources: Customized cloud access for hybrid data pipelines, traditional simulations and exploratory research.
- Application Performance and Productivity: Software solutions provided to research/engineering teams to better exploit installed systems.
- Visualization and Data Analysis: Custom visualization during traditional post-processing or concurrent during simulation to understand complex interactions of data.
- Networks: End-to-end network performance enhancements for user communities throughout the world.
- Data Analytics: Exploitation of data sets through neural nets and emerging new techniques.
- Machine Learning: Custom environments to enable learning through advanced data techniques.
- Custom Data Gateways: Custom data portals to support diverse programs and projects.

R&D for the Future at HECC

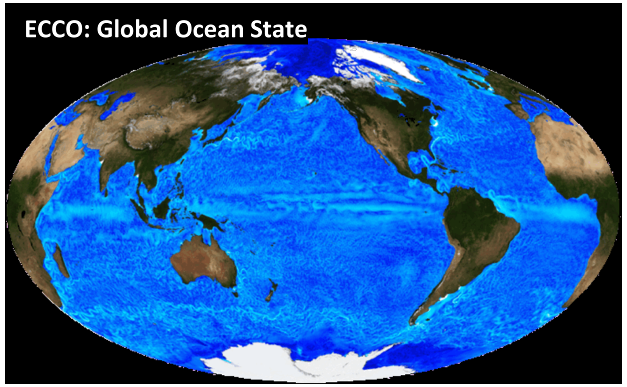
- New algorithms to exploit hardware realities.
- Performance optimization including code refactoring for new architectures.
- Mixed precision and machine learning/AI in computation.
- New methods for visualization and data analytics.
- Application enhancements to exploit I/O advancements.

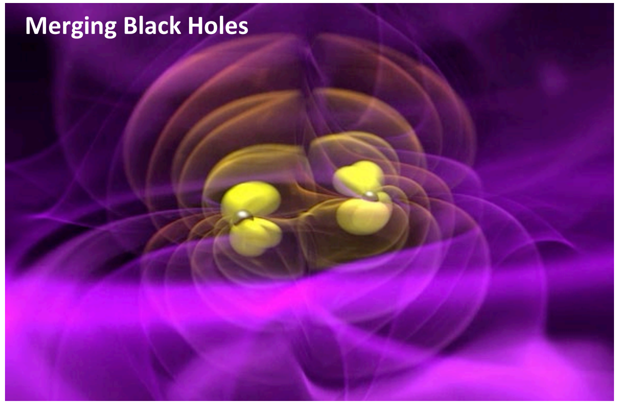


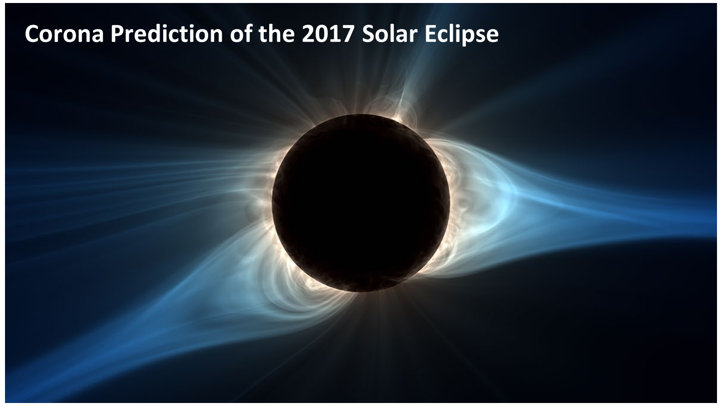
- Research and develop programming approaches, for accelerated and non-accelerated architectures.
- Optimize and port techniques/tools for visualization, data analytics, AI, and machine learning.
- Research and develop efficient math libraries.
- New file system software technologies.
- Research and develop machine learning techniques for problem solving and system error tracking/prediction.
- Evaluate software packaging technologies, such as containers.

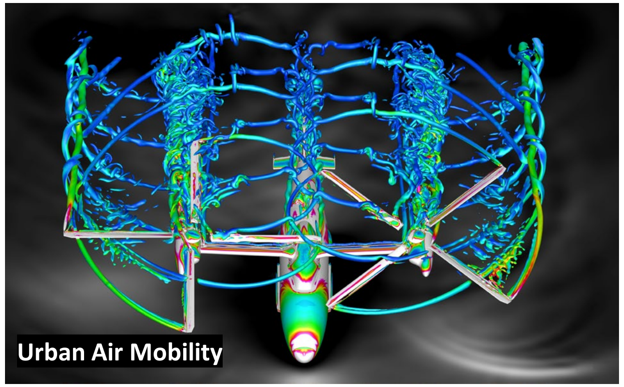
- Examine requirements.
- Develop necessary benchmarks.
- Assess I/O approaches.
- Develop models to predict performance.
- Make recommendations for pathfinding systems.

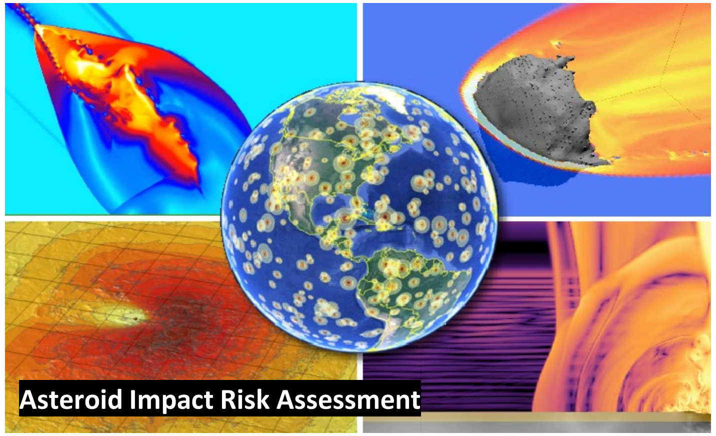
Representative applications across MDs @ NAS

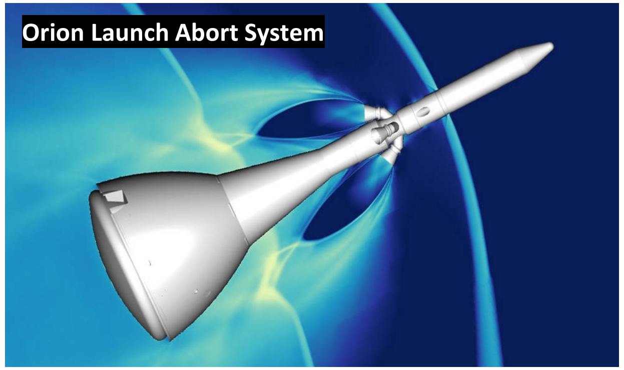


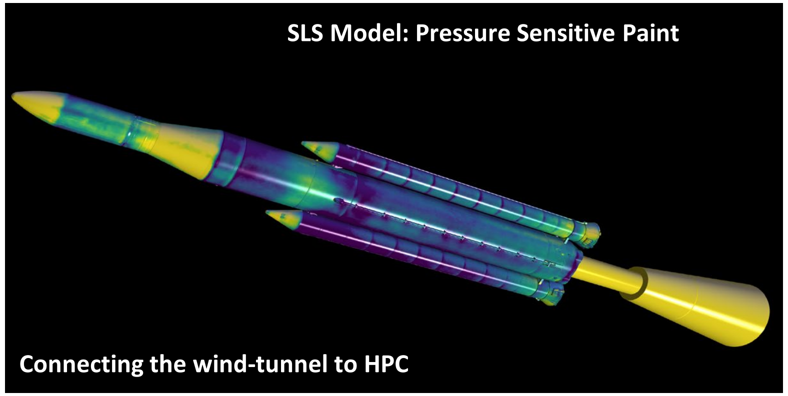








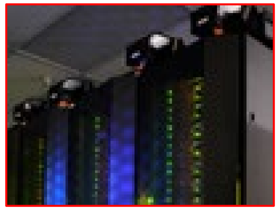




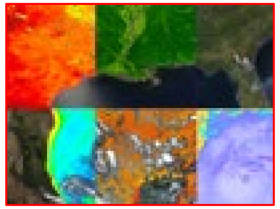
Scientific Computing

Scientific Networking

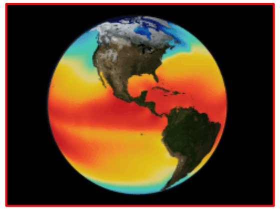
High Performance Computing (HPC) Cloud Computing



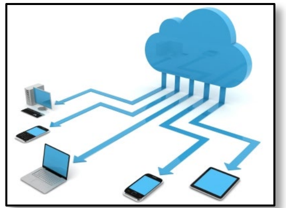
Data Science, AI/ML



Scientific Visualization Studio (SVS)



Science Managed Cloud Environment (SMCE)



NASA Center for Climate Simulation

- High Performance Computing (HPC, Discover Cluster)
- Private Science Cloud (Advanced Data Analytics Platform – ADAPT)
- Accelerator-based systems to enable AI/ML and Exascale Computing
- Large shared data storage for curated data
- Science Managed Cloud Environment (SMCE) using the Amazon Web Services (AWS)
- Advanced Software Technology Group (ASTG)

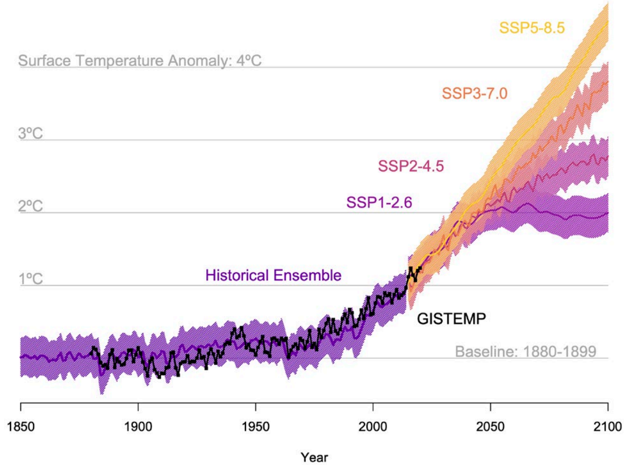
Target For Next 10 - 15 Years

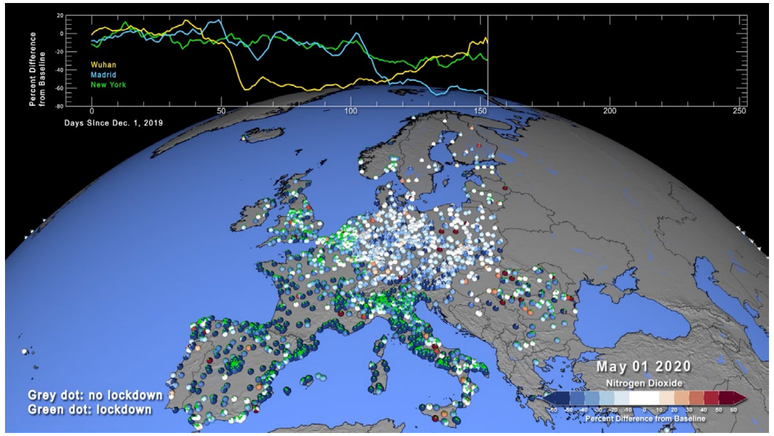
- Exascale Computing Requirements
- +1,000x growth in computational capacity
- >3,000x growth in storage capacity

Fun Fact: Discover (~127,000 compute cores) can execute 7,700,000,000,000,000 ($7.7x10^{15}$) floating point operations per second; that is the equivalent of every person on earth multiplying two numbers together ever second for over 10 days straight!

Representative Applications

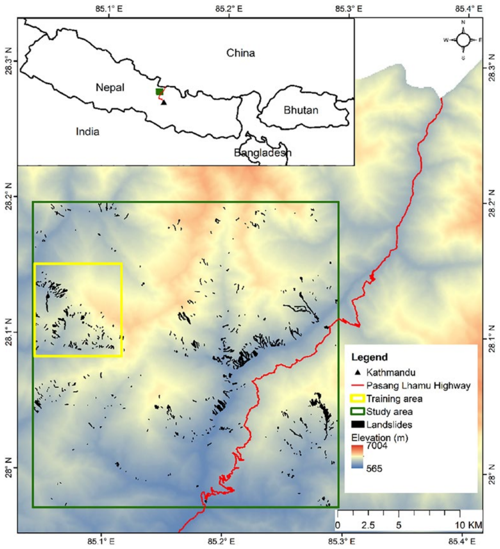
NASA GISS Global mean surface temperature anomalies.

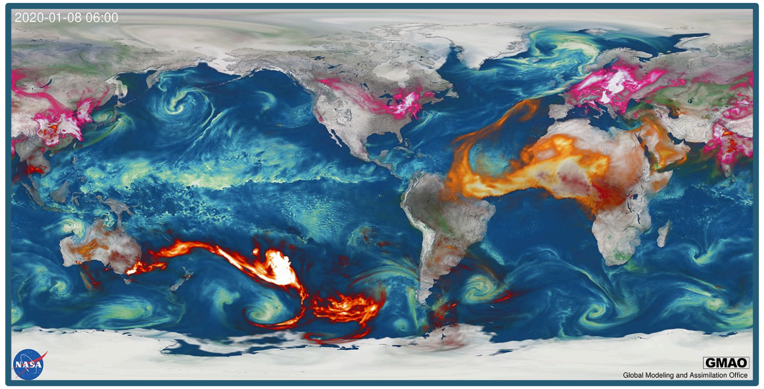




Decline of Air Pollution During COVID-19 Pandemic

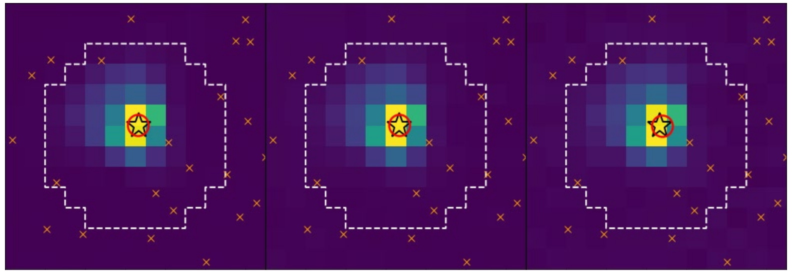
Landslide mapping

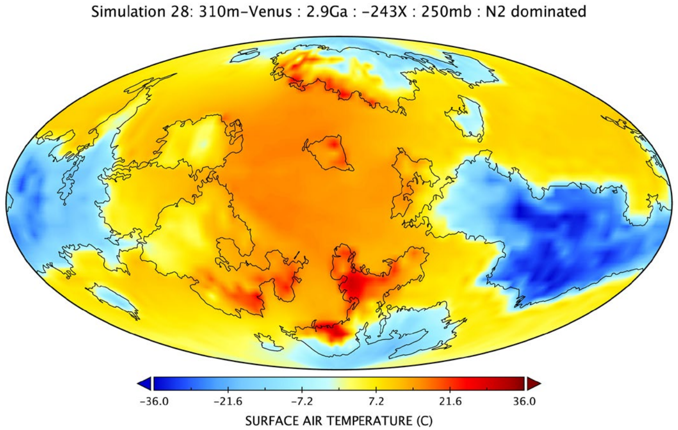




NASA GMAO High Resolution Global Atmospheric Modeling

Light Curve Analysis – sextuple star system





Climate Simulations of Venus – Exoplanet Climate

Backup

NASA/NOAA/DOE Partnership

NASA
- Have systems on which the development can take place; not many systems: order of 10 nodes
- Development of GOES GridTools DSL
- Possible development of DSL physics packages
- Sharing lessons learned; regular tag ups

NOAA
- Porting FV3UFS using GT4Py
- Coordinates on integration and development efforts
- Sharing lessons learned; regular tag ups

DOE
- Transfers code porting experiences (possibly through hackathon)
- Allows access to DOE systems within 6 months for scalability and performance testing
- Sharing lessons learned; regular tag ups

Code Modernization to Exascale Computing Architecture

GOES Domain Specific Language (DSL) Project: DOE/NASA/NOAA Partnership

Primary Goal:

Develop the next generation GEOS model using a Domain Specific Language (DSL) to enable an increase in resolution and scalability to meet future GMAO requirements.

Strategy:

- Recognize that accelerator-based systems provide a promising platform to meet future application requirements.
- Adoption of a DSL provides an opportunity for portability and scalability across multiple platforms, including more classic CPU systems and accelerator-based systems.
- Provides a higher-level language/interface for model development.
- Also provides for the ability to optimize implementations for specific architectures under the covers.

Requirement:

- Production support for coupled data assimilation (NWP, S2S and Reanlysis)
  - NWP: 6-km ATM coupled to 6-km OCN [4D-EnvVar, 4-cycles per day]
  - S2S: 25-km ATM coupled to 25-km OCN [31-Ensemble members 30- to 90-day predictions]
  - Reanalysis (MERRA-3): 12-km ATM coupled to 12-km OCN
- Ultra-high resolution global Nature Runs for OSSEs
  - Coupled 3-km ATM and 3-km OCN
  - Carbon and chemistry: 3- to 1-km ATM
  - Mesoscale global convection resolving: sub-1km ATM

HEC Portfolio Management Background

- HECC (High-End Computing Capability) project was established in 2006 to maintain a computational capability and capacity including computing, storage, networking, facility and human capital to support the agency needs
- OCFO, SMD, and ARC collectively designed the project and required the budget contributions from all the mission directorates and supporting offices based on the past utilization and the projection of future needs
- The budget and management responsibility were assigned to Earth Science Division in SMD by the Agency Strategic Management Council
- SMD/ESD is the steward of the HEC budget since 2006
- Initially HEC was part of the Shared Capability Asset Program and now part of the Agency Capability Portfolio governed by the Mission Support Council
- Agency passed the portfolio management policy NPD 8600.1 in 2018 and management authority for HEC portfolio was given to SMD

Collaborations

Current Collaborations

- Single contract across the two centers
  - Exchange of best practices
  - Joint solutions to similar issues, e.g., in operations
- HEC Needs Assessment

Potential Collaboration

- Exploring novel architectures/approaches for Exascale computing
- Transitioning applications to new architectures
- AI/ML infrastructure and algorithms
- Hybrid Cloud strategy
- Research in HPC topics
- Workforce Development

Need Center Leadership help in HPC Advocacy to HQ