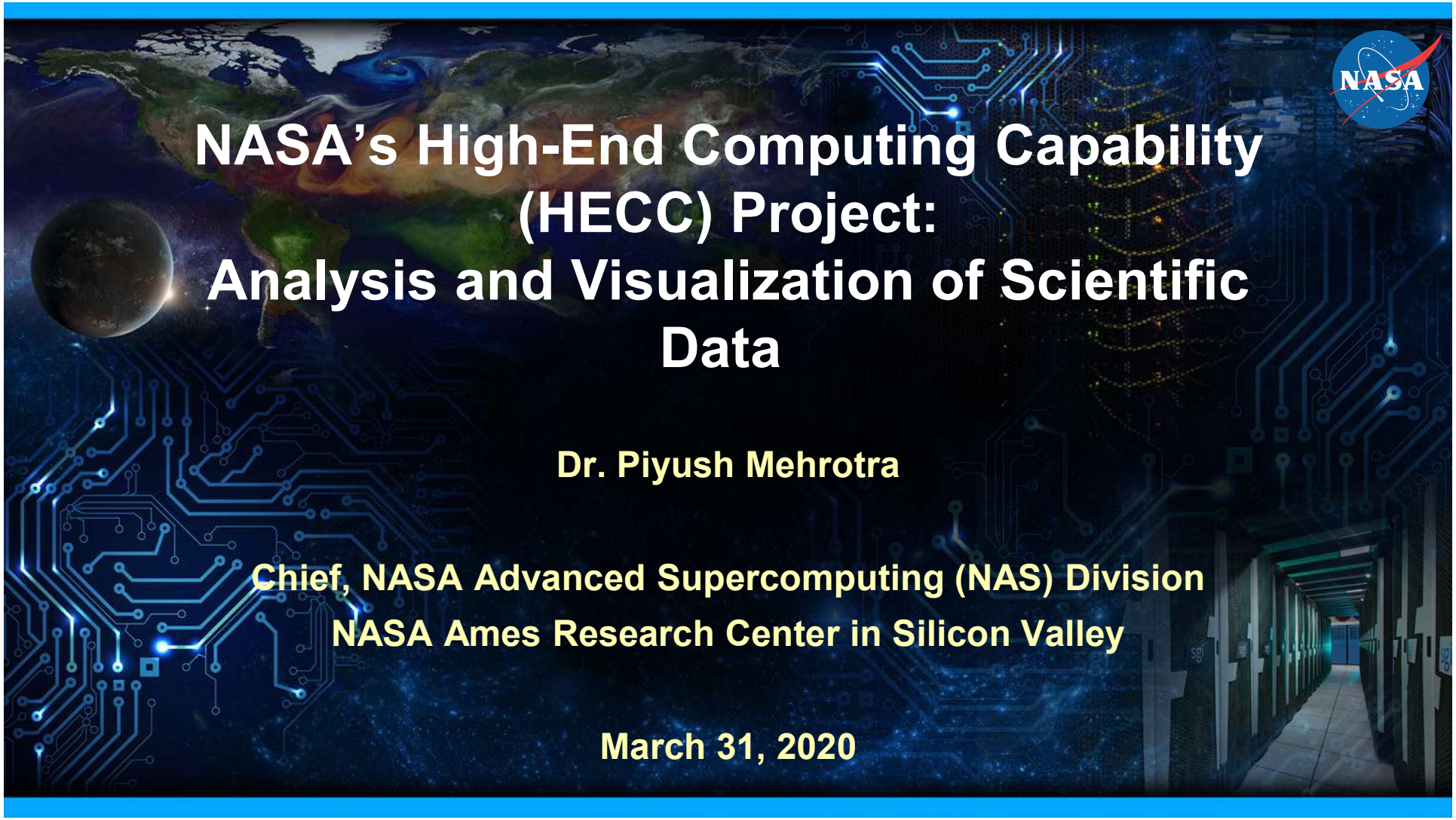


NASA's High-End Computing Capability (HECC) Project: Analysis and Visualization of Scientific Data

Dr. Piyush Mehrotra

Chief, NASA Advanced Supercomputing (NAS) Division

NASA Ames Research Center in Silicon Valley

March 31, 2020

High-End Computing Capability (HECC) Project

NASA's premier Supercomputer Center
(hosted by the NAS Division at the Ames Research Center)

Charter: to meet the supercomputing needs of all NASA's Mission Directorates

Over 600 science & engineering projects with more than 1,600 users across the nation & the world

Pleiades: 7.25 PF
245,000+ cores

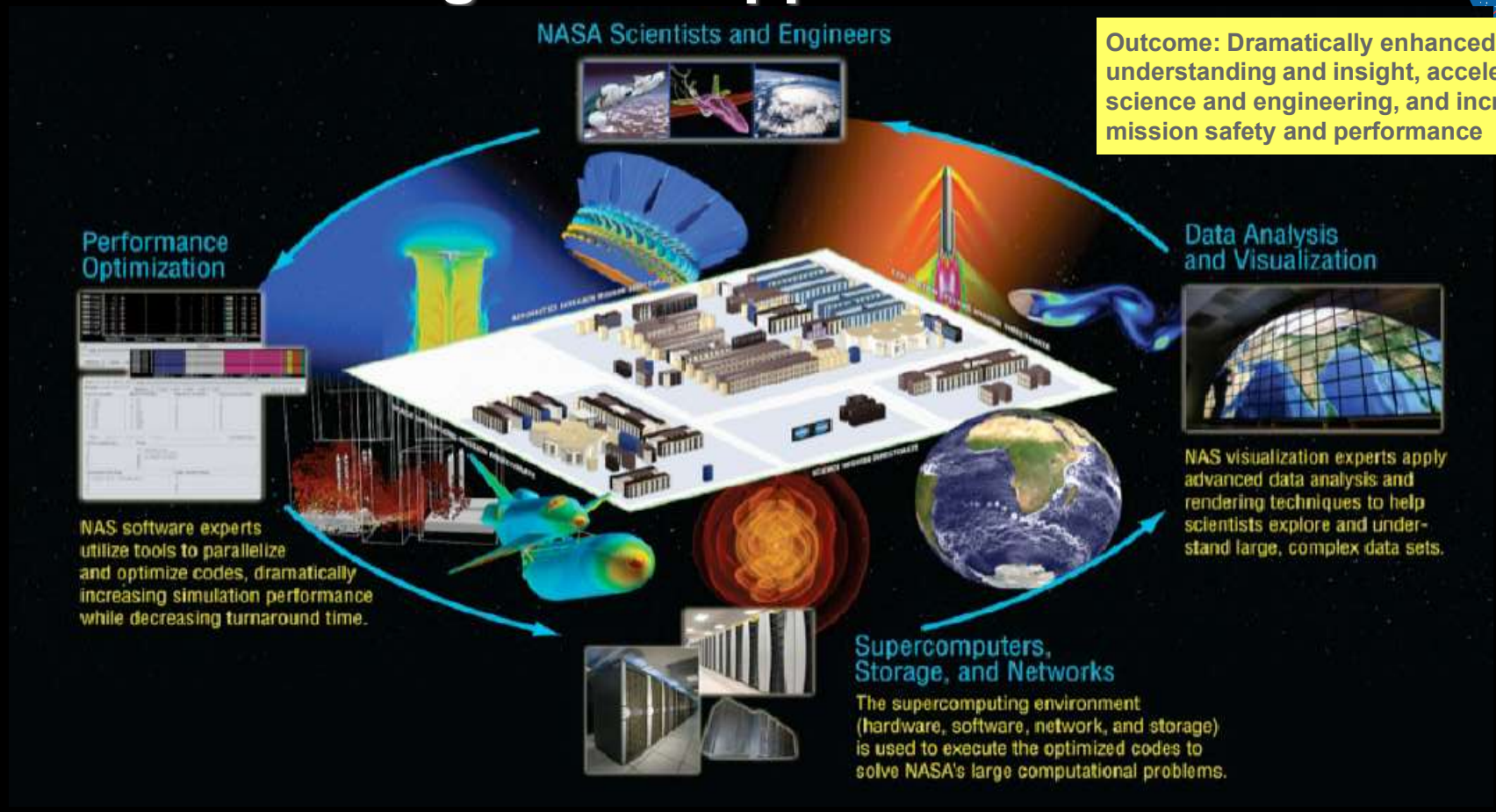
Electra: 8.32 PF
124,416 cores

Aitken: 3.69 PF
46,080 cores

Global Storage Capacity :
Online: ~50 PB
Archival: 1 EB

http://www.nas.nasa.gov/hecc

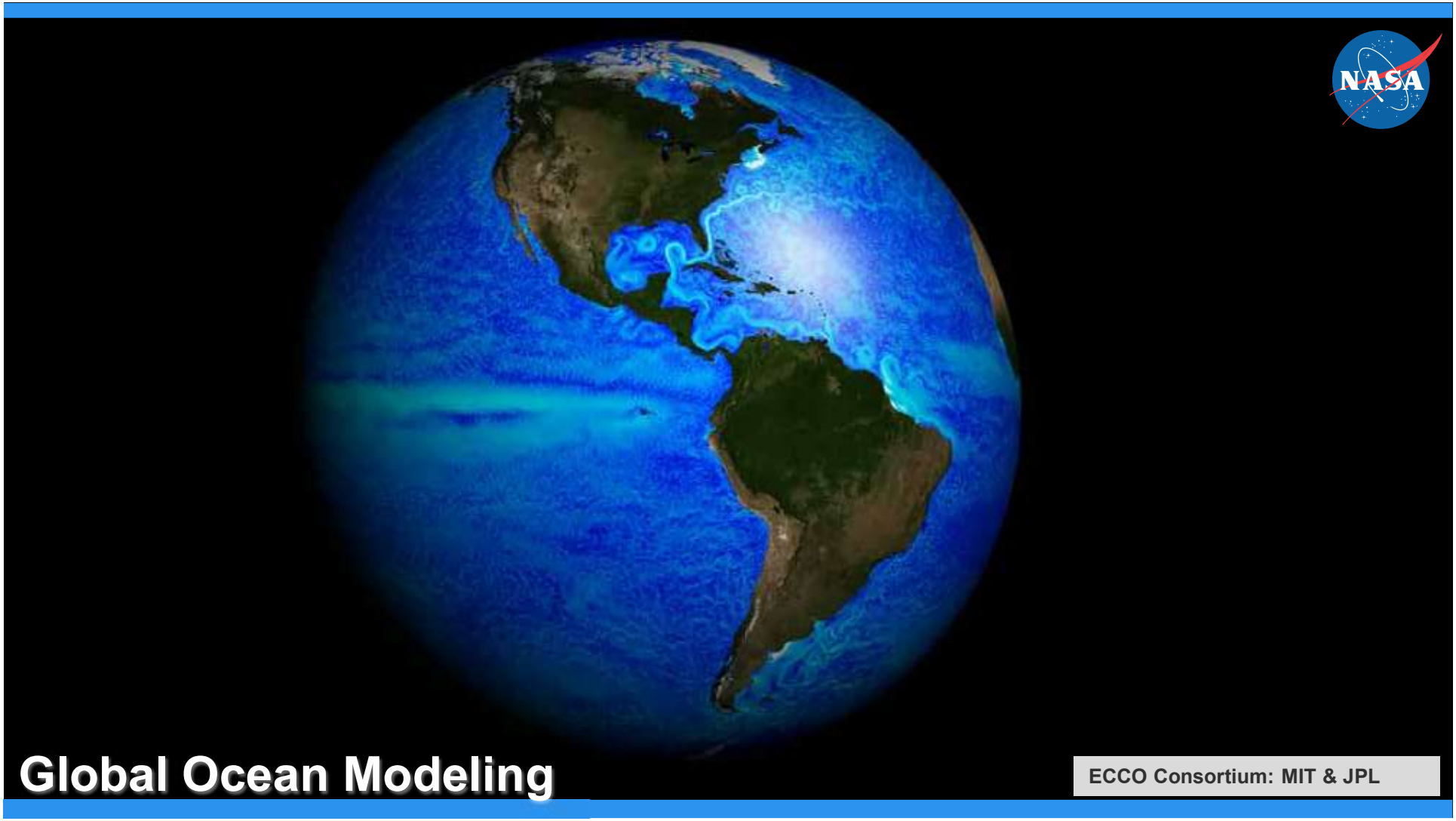
Integrated Support Services



Outcome: Dramatically enhanced understanding and insight, accelerated science and engineering, and increased mission safety and performance

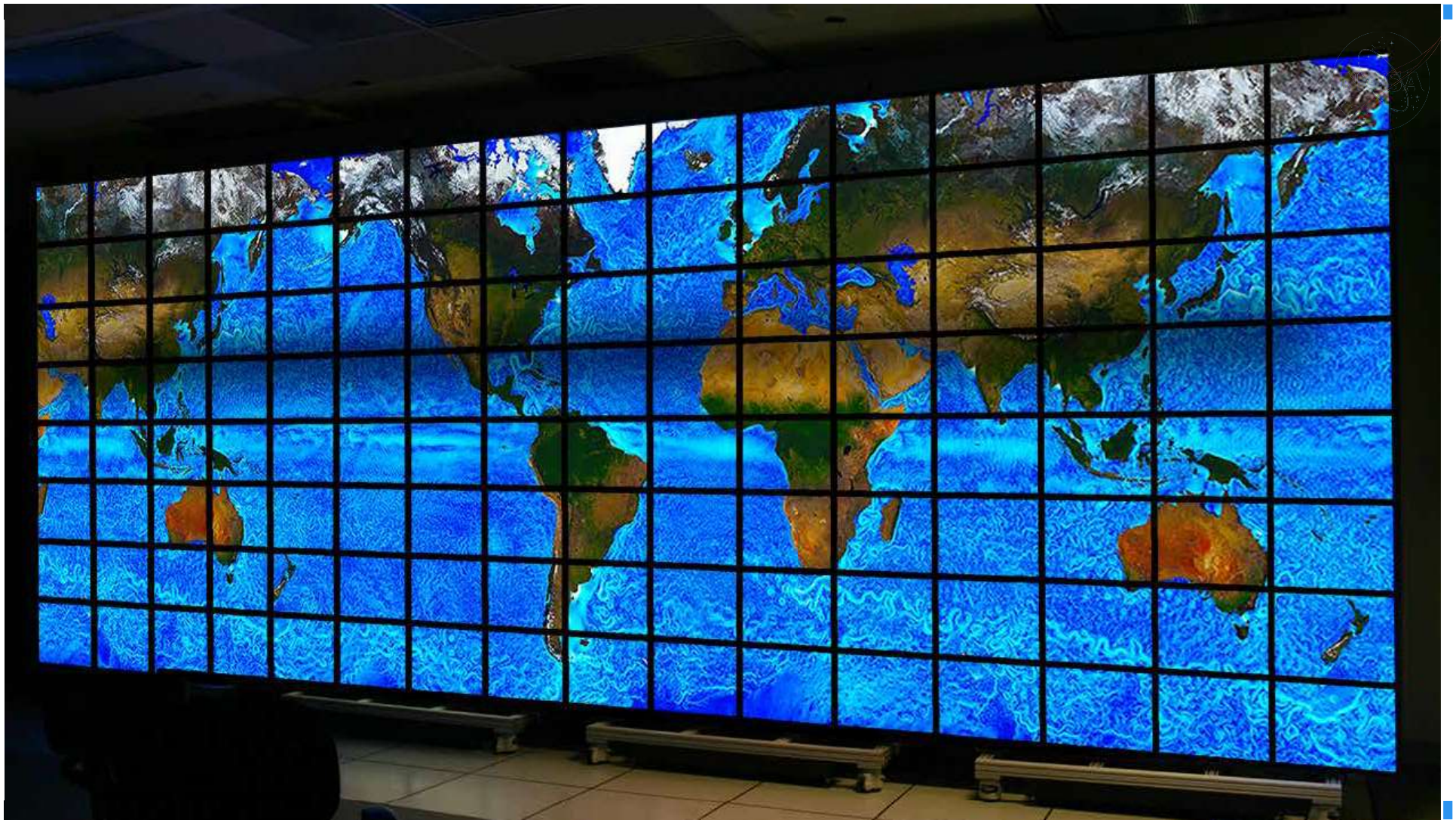
Advanced Visualization

- *Supercomputing-scale visualization system* to handle massive size of simulation results and increasing complexi of data analysis
  - 8x16 LCD display (23 feet x 10 feet)
  - 245 million pixels
  - Backed by 128-node GPU-based system
- *Two primary modes*
  - Single large high definition image
  - Sets of related image(s)/animation(s) (e.g. parameter study)
- *High-bandwidth to HPC resources*
  - *Traditional Post-Processing*: Direct read/write access to Pleiades file systems eliminates need for copying large datasets
  - *Concurrent Visualization*: Runtime data streaming allows visualization of every simulation time step - ultimate insight into simulation code without increase in traditional disk I/O



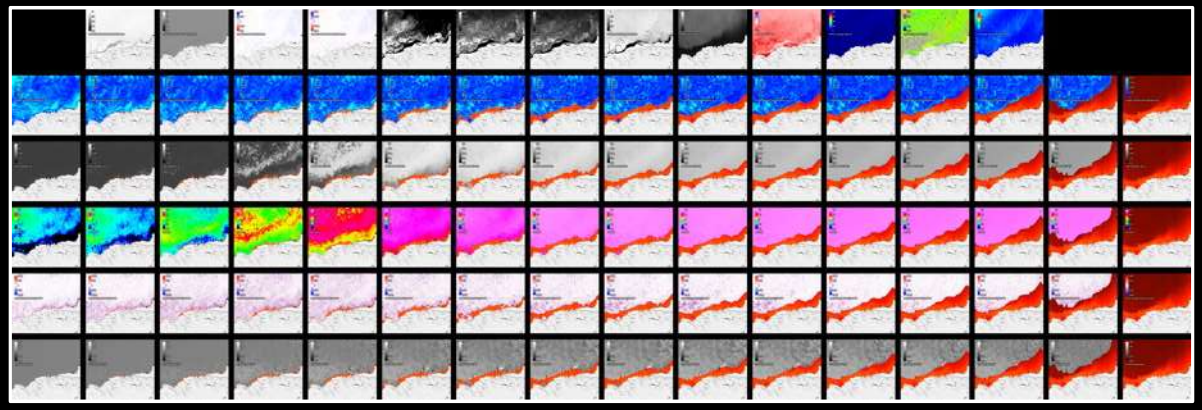
Global Ocean Modeling

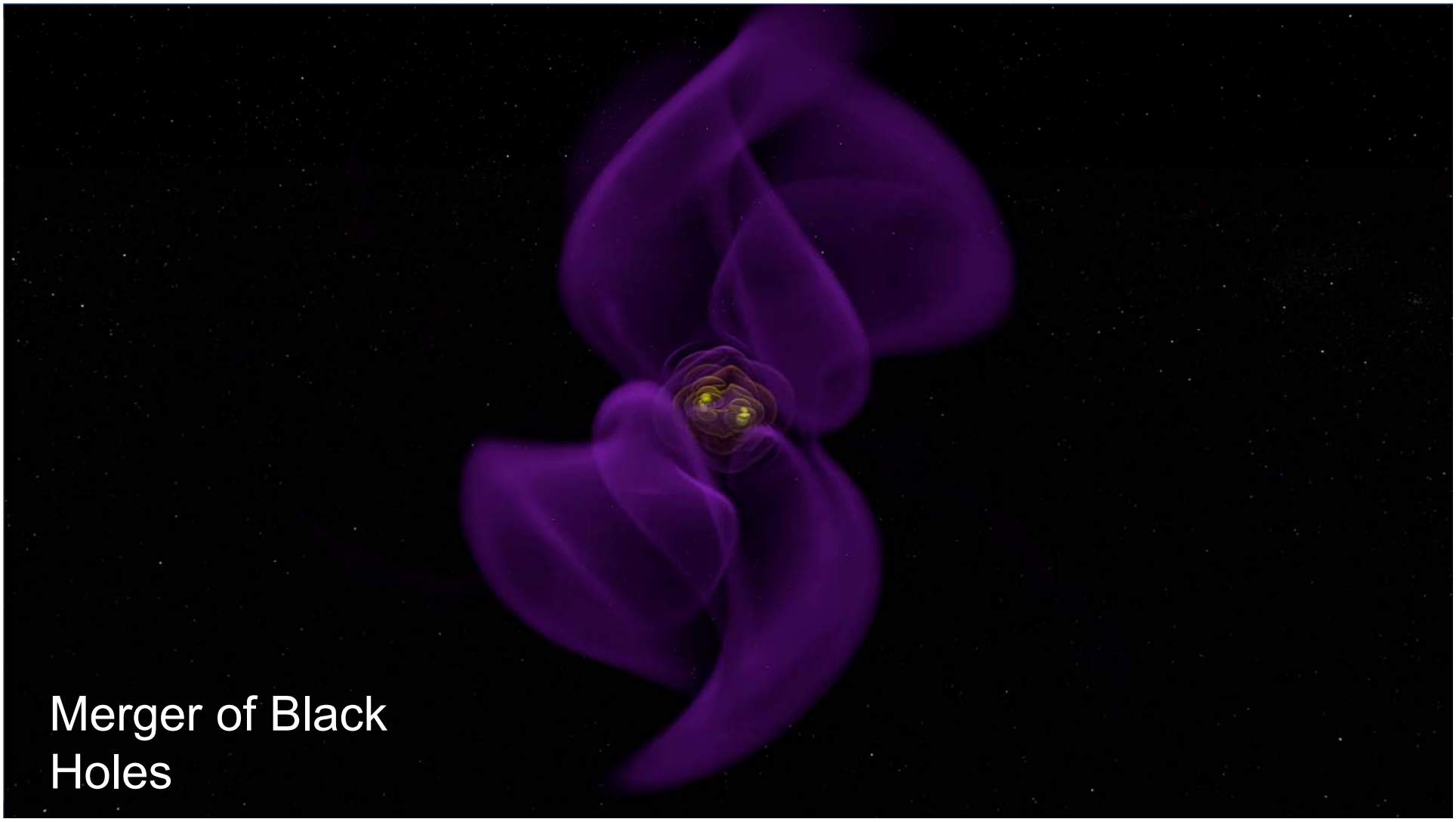
ECCO Consortium: MIT & JPL



Depth analysis using a view showing multiple depths of a single north-south section of the globe - horizontal velocity magnitude around the Americas, with depth increasing to the right. The orange is the ocean floor

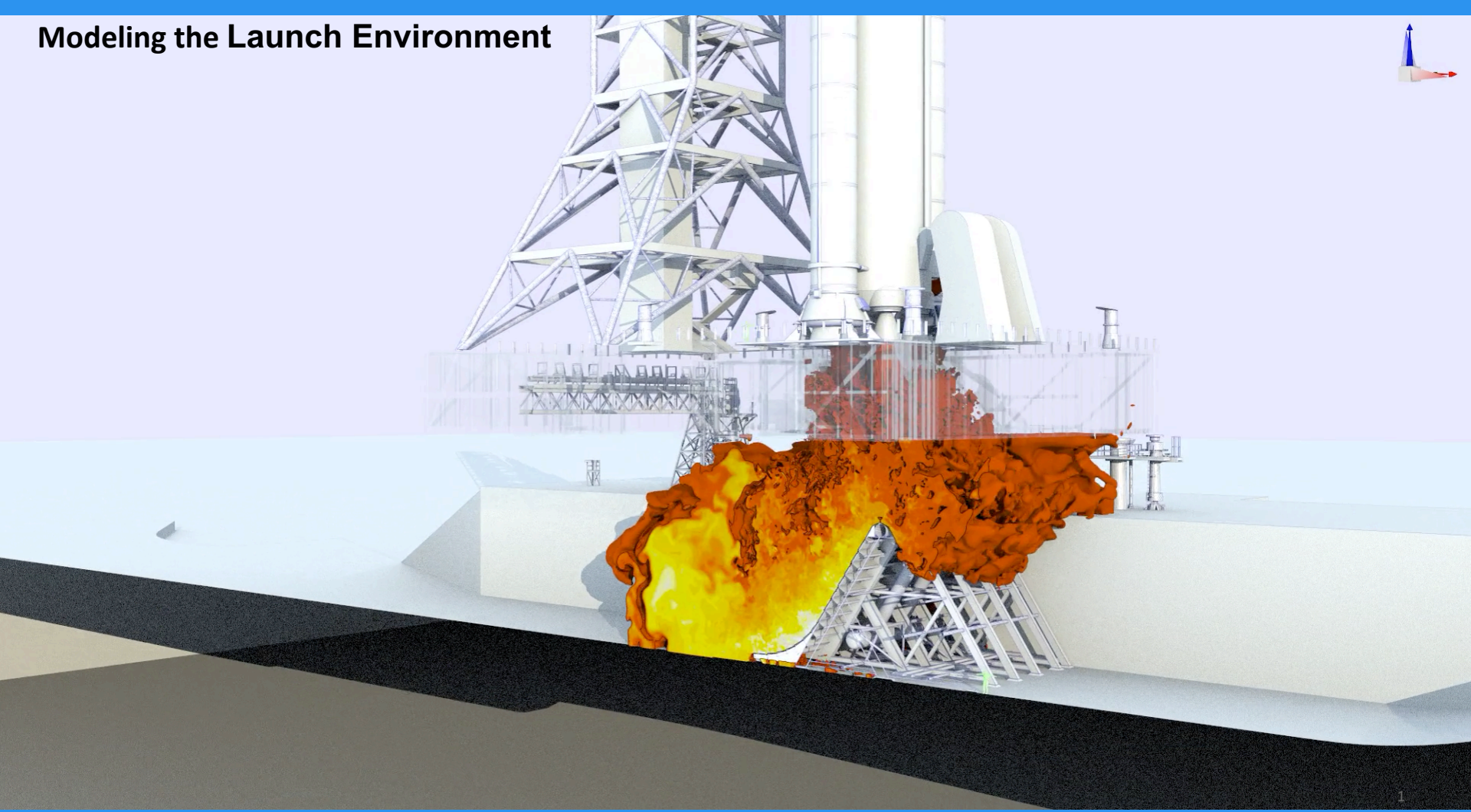
Finding correlations made easy by showing all data for a given region; depth increases left to right for the 3D fields



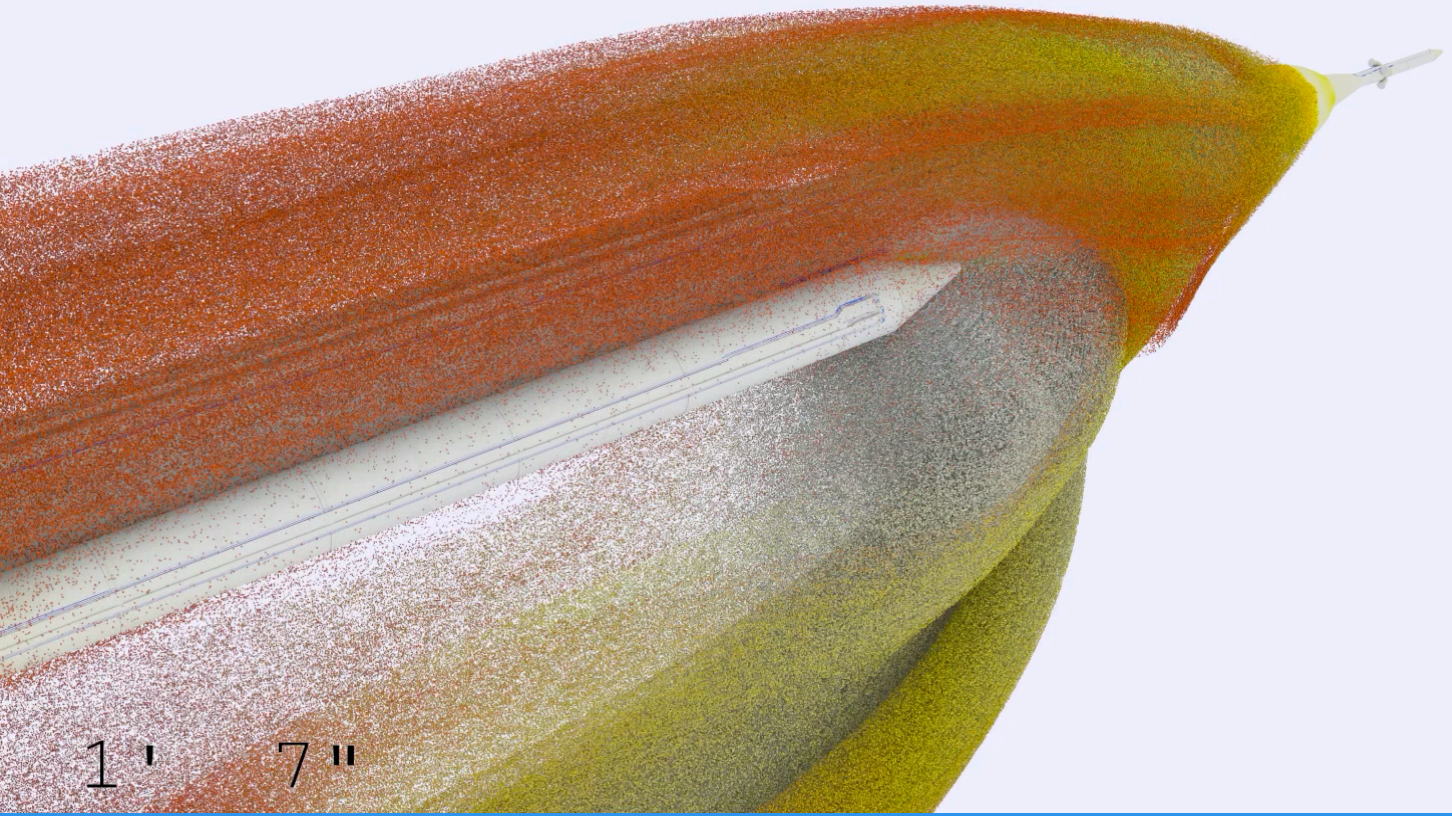


Astrophysicists interested in galaxy formation black holes origins of the universe
Scientists created a database of gravitational wave patterns different sizes and spins of black holes
LIGO observatory detected gravitational waves in 2016
Used database to determine parameters
And then modeled

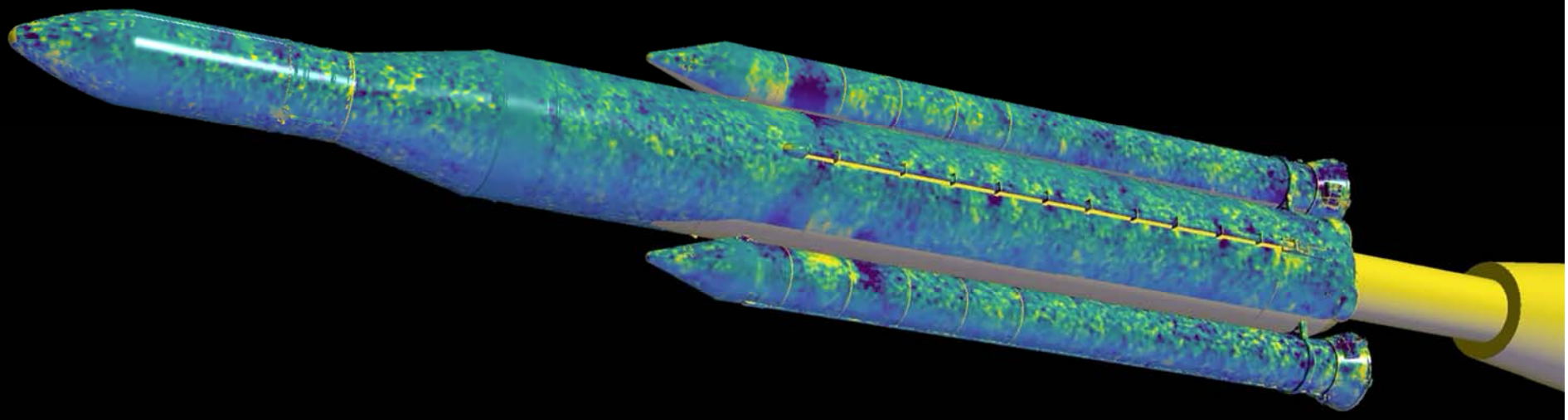
Modeling the Launch Environment



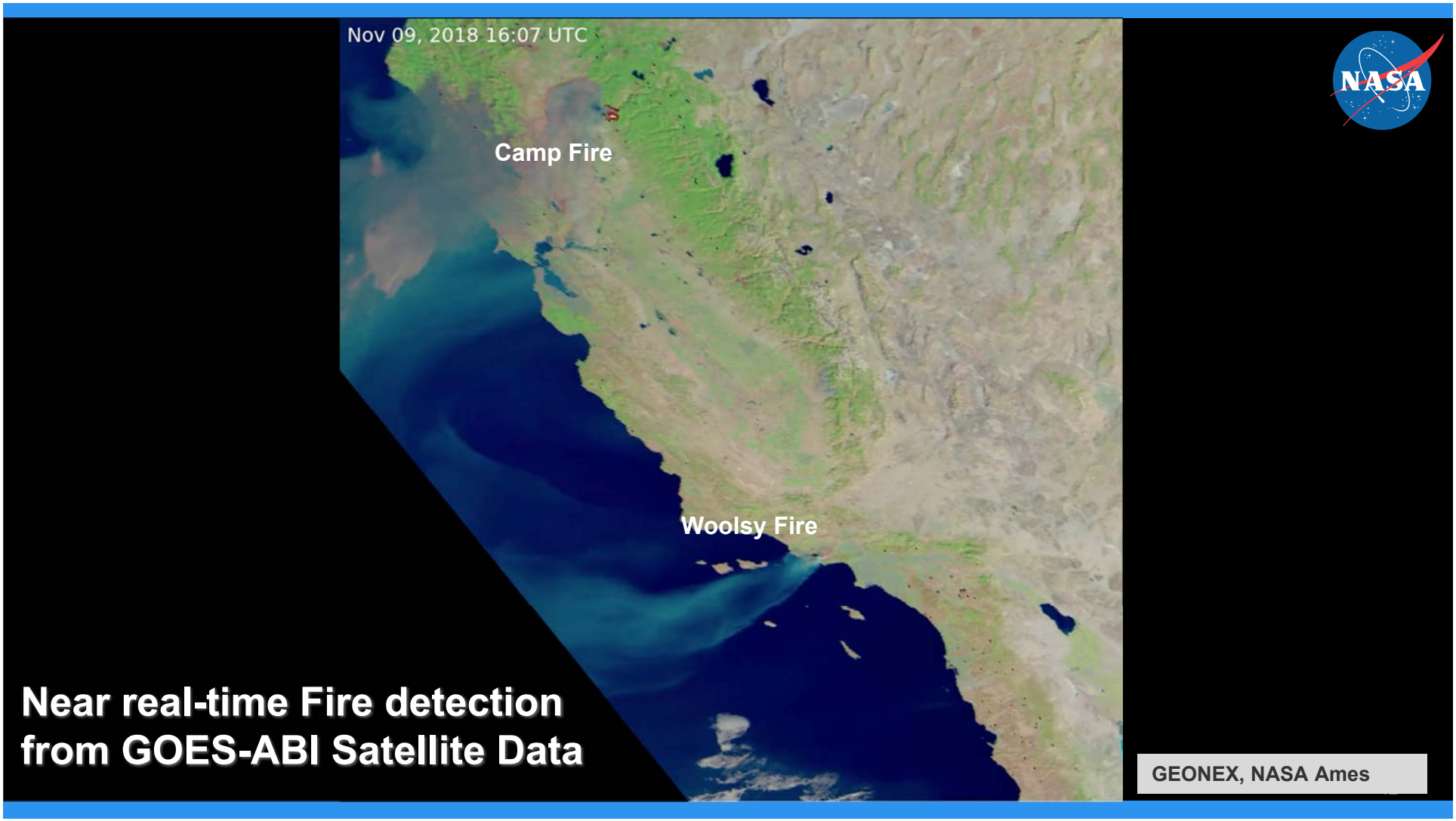
Space Launch System – Stage Separation



Wind tunnel runs of SLS model using pressure sensitive paint to estimate pressure/loads



Near real-time analysis of test data using Supercomputer resources
with potential for computer-guided data acquisition



geostationary satellites (22300 miles)
provide data every few minutes
land surface temperatures, snow dynamics, vegetation, drought impact etc.

Sample NASA AI/ML Projects

- **Feature detection**
  - Shock waves & vortices from flow data
  - Exoplanet identification from Kepler data
  - Artifact identification, e.g., trees, from satellite data
- **Prediction**
  - Solar flares/space weather from solar surface magnetic fields data
  - Asteroid properties from light curves
  - Solar cells current-voltage properties from IV curves
- **Anomaly detection**
  - Aviation safety issues from flight data
  - Systems behavior, e.g., ISS control operations
  - Ammonia and CO_2 concentration using carbon nanotube sensors on the ISS
- **Interactive assistants/robots that can learn for ISS crew**
- **Autonomous rovers**

- **Mission Support**
  - Email Classification/Records Management
  - Scientific Document Tagging
  - Network Traffic Anomaly Detection
  - Service Desk Ticket Analysis & Trending
  - Detect CUI content in documents

NASA's Data Analysis Challenges

NASA data is increasing in volume and complexity from multiple sources: observational, simulation and experimental

- **Discovery**
  - Multi-petabytes of data with little metadata to support discovery/sharing
- **Data Access**
  - Siloed data with non-intuitive access interfaces or APIs
  - Distributed storage not co-located with computational resources
- **Tools/Algorithms development at scale**
  - One-off algorithms/tools/workflows
- **Analysis infrastructure**
  - Heterogeneous hardware requirements
  - Optimized data management and archiving
- **Workforce & Culture issues**

- SMD's "Strategy for Data Management and Computing for Groundbreaking Science 2019-2024" (Dec 2019)
- Agency focus on Digital Transformation and Data Strategy including Governance

Conclusions

- Today's supercomputers are enabling ever larger simulations – using tens to hundreds of of thousands tcores running for weeks.
- The increased computing capability has allowed for a dramatic increase in the fidelity of the simulations and the ensuing results
- The enhanced quality and granularity of the data along wit the data and visualization analysis tools has supported the decision makers, increased our understanding of the Earth system, solar system and the universe while also having a direct impact on our daily lives.

Questions?

piyush.mehrotra@nasa.gov

http://www.nas.nasa.gov/