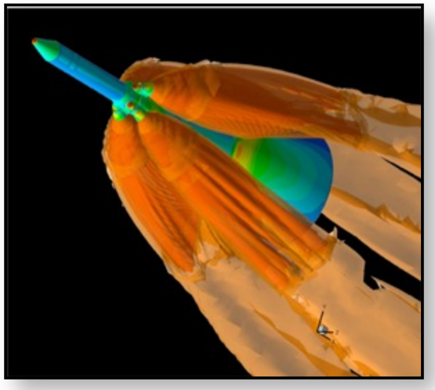
Environmentally-Friendly Supercomputing at NASA

William W Thigpen
Manager, High-End Computing Capability Component
Assistant Division Chief for High-End Computing
william.w.thigpen@nasa.gov

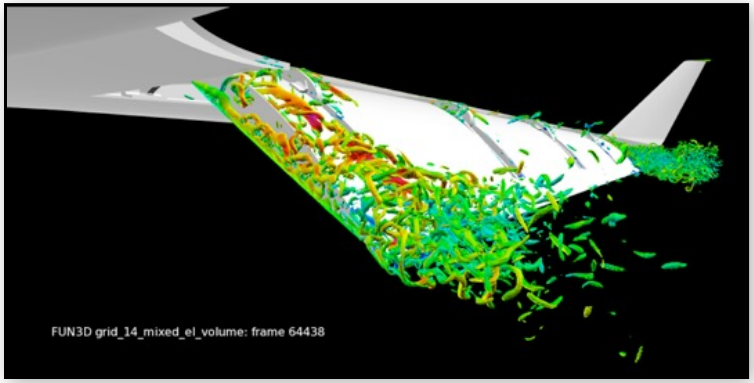
NASA's HEC Requirements: Capacity

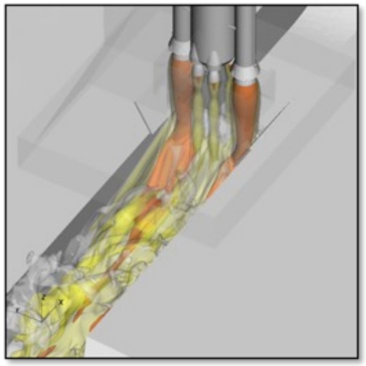
Engineering-related work requires HEC resources that can handle large numbers of relatively-low CPU-count jobs with quick turnaround times.



Over 1500 simulations utilized ~2 million processor hours to study launch abort systems on the next-generation crew transport vehicle.

The formation of vortex filaments and their roll-up into a single, prominent vortex at each tip on a Gulfstream aircraft.

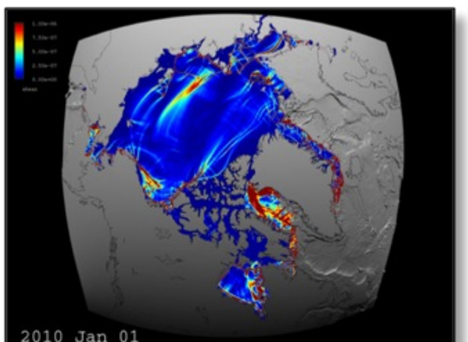




Over 4 million hours were used over a four-month project to evaluate future designed of the next generation launch complex at Kennedy Space Center.

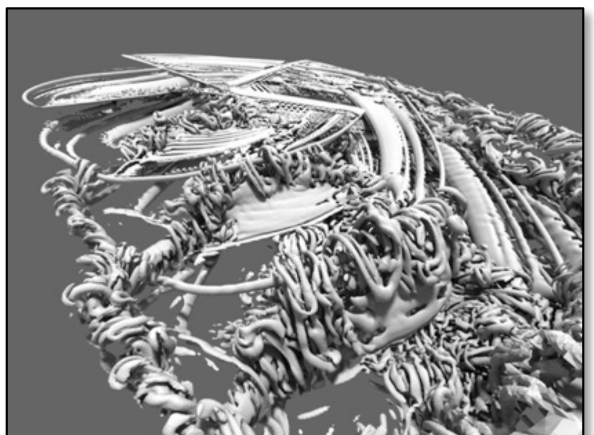
NASA's HEC Requirements: Capability

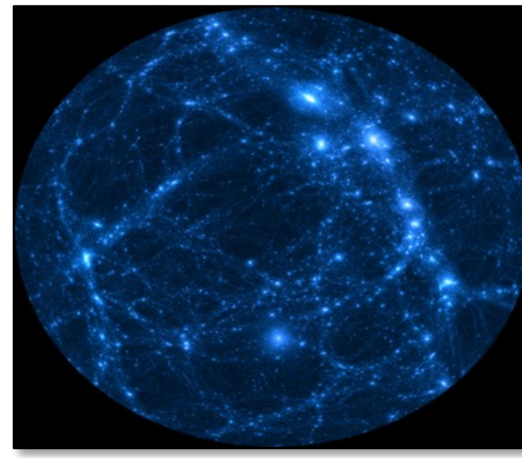
Aeronautics and science related work require HEC resources that can handle high fidelity relatively-large CPU-count jobs with minimal time-to-solution. Capability enables work that wasn't possible on previous architectures.



NASA is looking at the oceans, running 100's of jobs on Pleiades using up to 10,000 processors. Looking at the role of the oceans in the global carbon cycle is enabled by access to large processing and storage assets.

For the first time, the Figure-of-Merit has been predicted within experimental error for the V22 Osprey and Black Hawk helicopter rotors in hover, over a wide range of flow conditions.

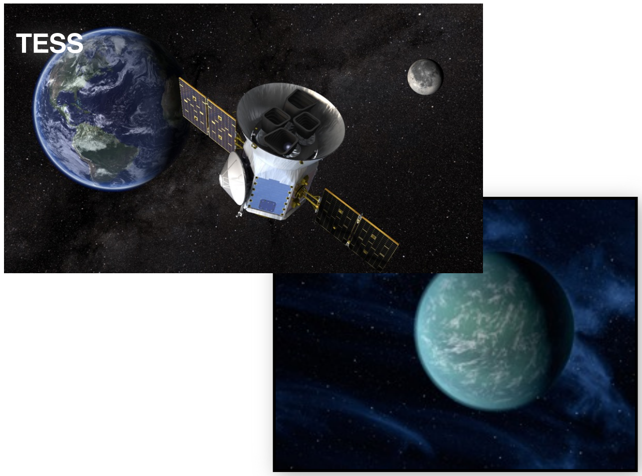




To complete the Bolshoi simulation, which traces how the largest galaxies and galaxy structures in the universe were formed billions of years ago, astrophysicists ran their code for 18 straight days, consuming millions of hours of computer time, and generating massive amounts of data

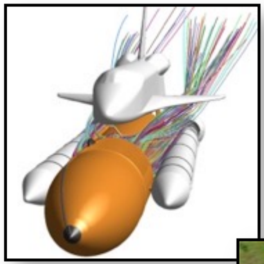
NASA's HEC Requirements: Time Critical

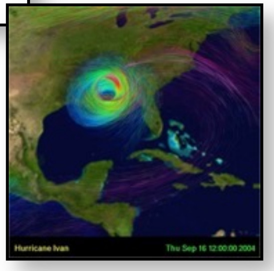
NASA needs HEC resources that can handle time-sensitive, mission-critical applications on demand (maintain readiness).



HECC worked with TESS to design a data pipeline that incorporated lessons learned from earlier missions enabling enormous planetary transit searches to be completed in less than a day with accuracy and effectiveness of the software pipeline.

REENTRY



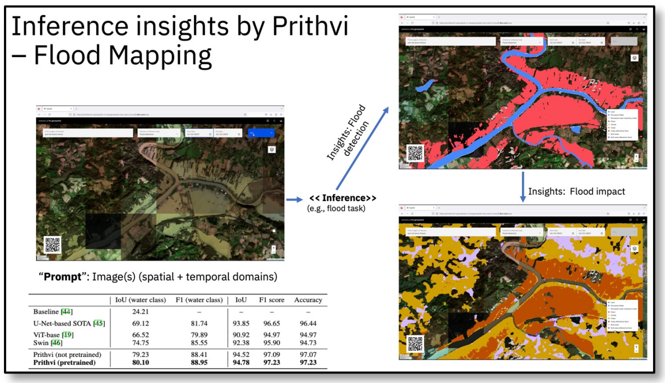


STORM PREDICTION

UAVSAR produces polarimetric (PolSAR) and interferometric (repeat-pass InSAR) data that highlight different features and show changes in the Earth over time.

NASA's HEC Requirements: Artificial Intelligence

NASA needs HEC resources that can enable discovery through the use of computers to explore massive data sets.



	IoU (water class)	F1 (water class)	IoU	F1 score	Accuracy
Baseline [44]	24.21	–	–	–	–
U-Net-based SOTA [45]	69.12	81.74	93.85	96.65	96.44
ViT-base [19]	66.52	79.89	90.92	94.97	94.97
Swin [46]	74.75	85.55	92.38	95.90	94.73
Prithvi (not pretrained)	79.23	88.41	94.52	97.09	97.07
Prithvi (pretrained)	**80.10**	**88.95**	**94.78**	**97.23**	**97.23**

HECC is working with the Earth Science Division in the Science Mission Directorate, the Department of Energy and industry to develop an Earth Science foundation model to do predictions and analysis.

Across NASA, teams are evaluating, testing and utilizing AI for design, scheduling, identifying areas of interest and predicting future states based on current. conditions.

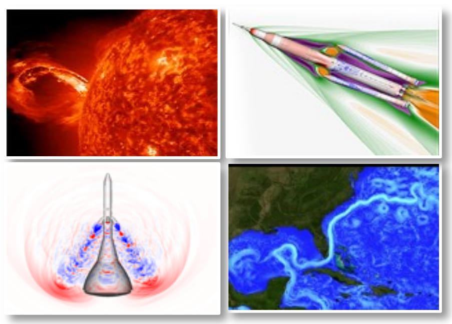
Meeting Today's Requirements

HECC Hardware Assets

5 Compute Clusters

- Aitken – 20 Racks / 3,200 nodes / 13.1 PF / 10,204 SBU/hr
- Pleiades – 143 Racks / 10,310 nodes / 6.83 PF / 7,600 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	1-billion-pixel display / 128 nodes / 2.61 PF
13 Lustre File Systems	118.1 PB
3 NFS File Systems	3.3 PB
2 BeeGFS File Systems	1.5 PB
1 VAST File System	2.5 PB
Archive System	1,572 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.

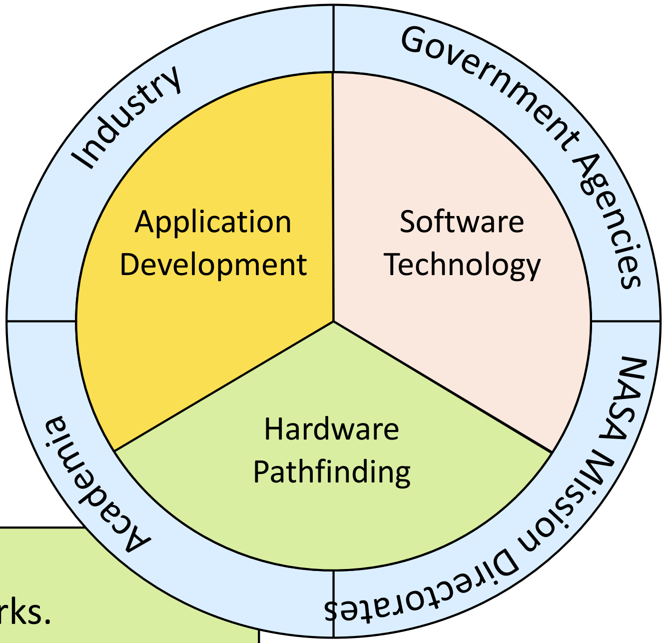
- Systems: Customized solutions including compute and storage solutions to meet specific project or mission requirements. Cloud access for immediate or non-standard computing.
- Cloud Resources: Customized cloud access for hybrid data pipelines, traditional simulations and exploratory research.
- Application 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.

Cabeus Expansion

- **Current**
  - 55 nodes w/4 NVIDIA V100 GPUs
  - 2 nodes w/8 NVIDIA V100 GPUs
  - 2.08 PF
- **Adding**
  - 128 nodes w/4 A100 GPUS (~5 PF)
- **Total System**
  - 185 nodes 748 GPUs 7.67 PF

Preparing for the Future

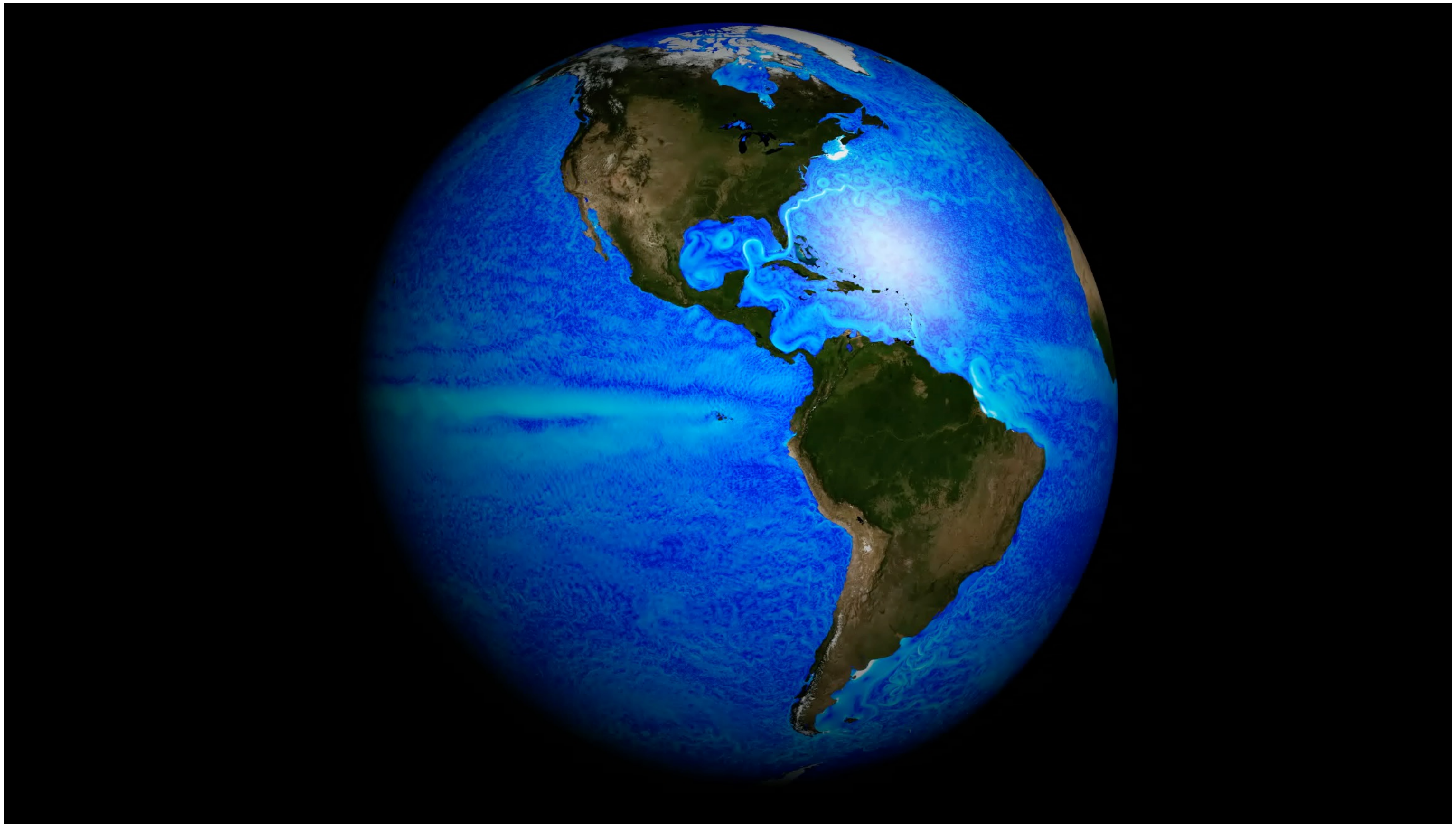
- New algorithms to match hardware realities.
- Mixed precision and machine learning/AI in computation.
- New methods in visualization and data analytics.
- Application enhancements to exploit I/O advancements.
- Fault resilience.
- Performance optimization.



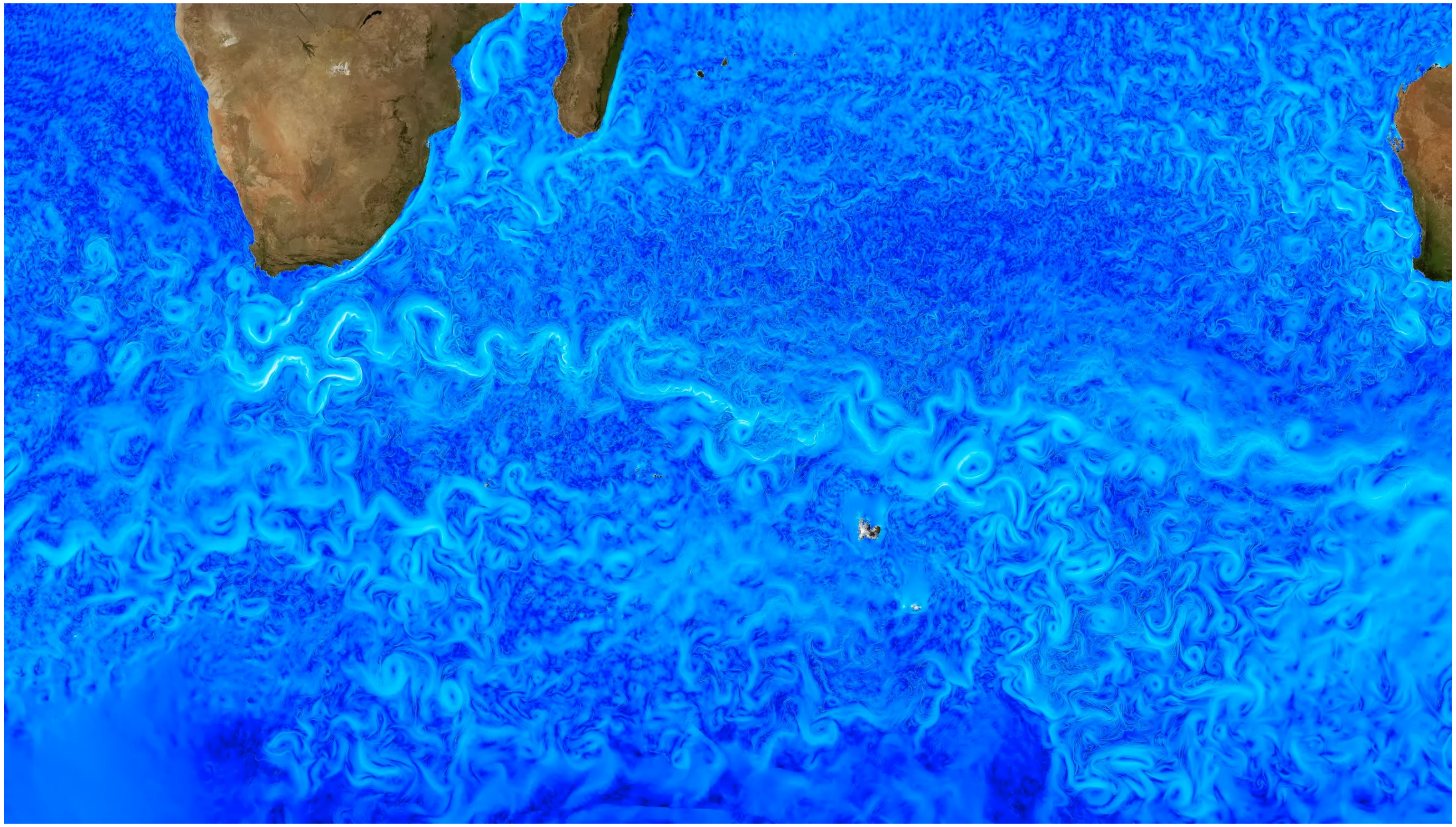
- Research and develop programming approaches, for accelerated and non-accelerated architectures.
- Optimize and port techniques/tools.
- Optimize and port techniques/tools for visualization, data analytics, AI, and machine learning.
- Research and develop efficient math libraries.
- New file system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

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

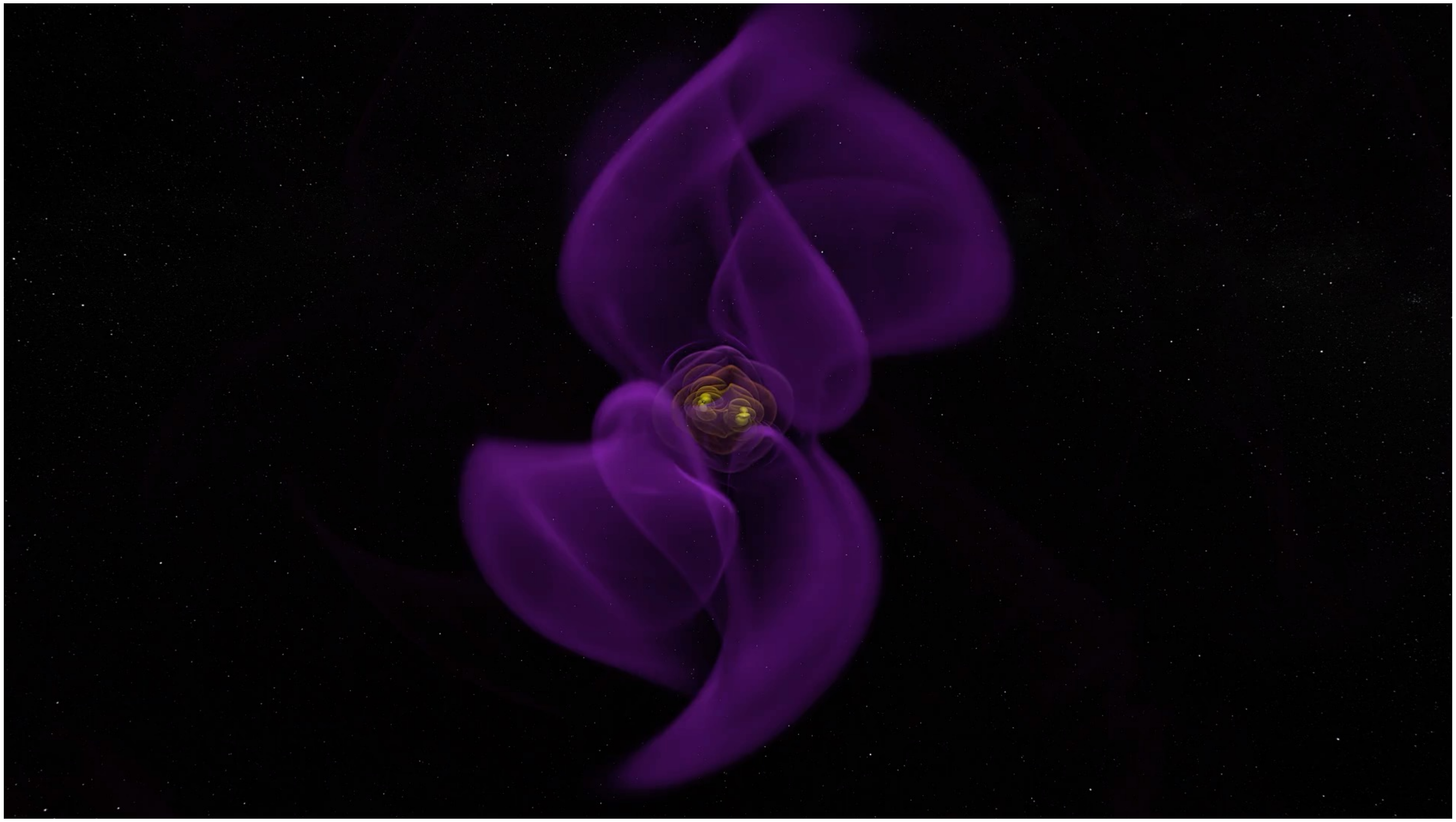
Earth Science



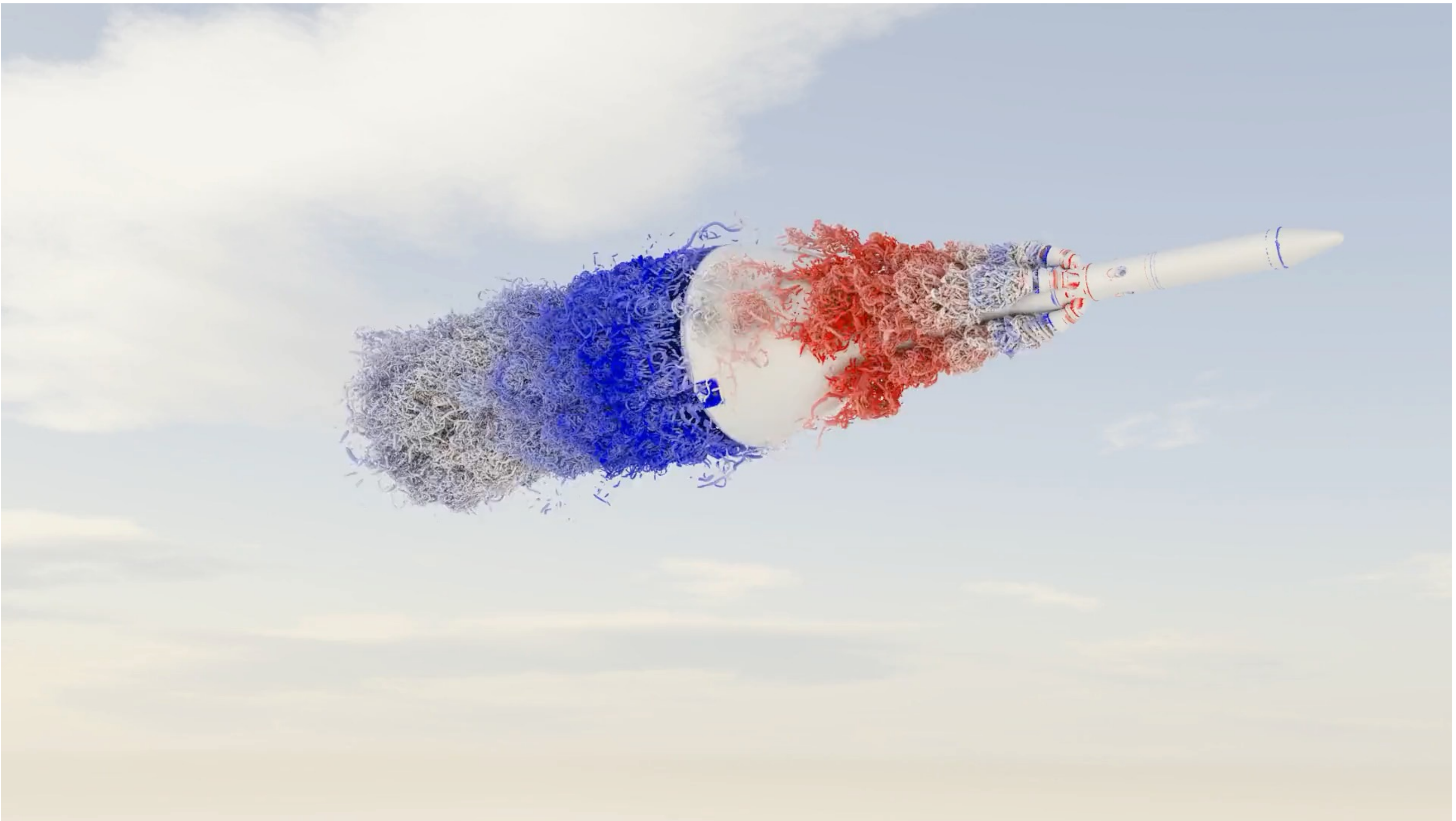
Earth Science Detail



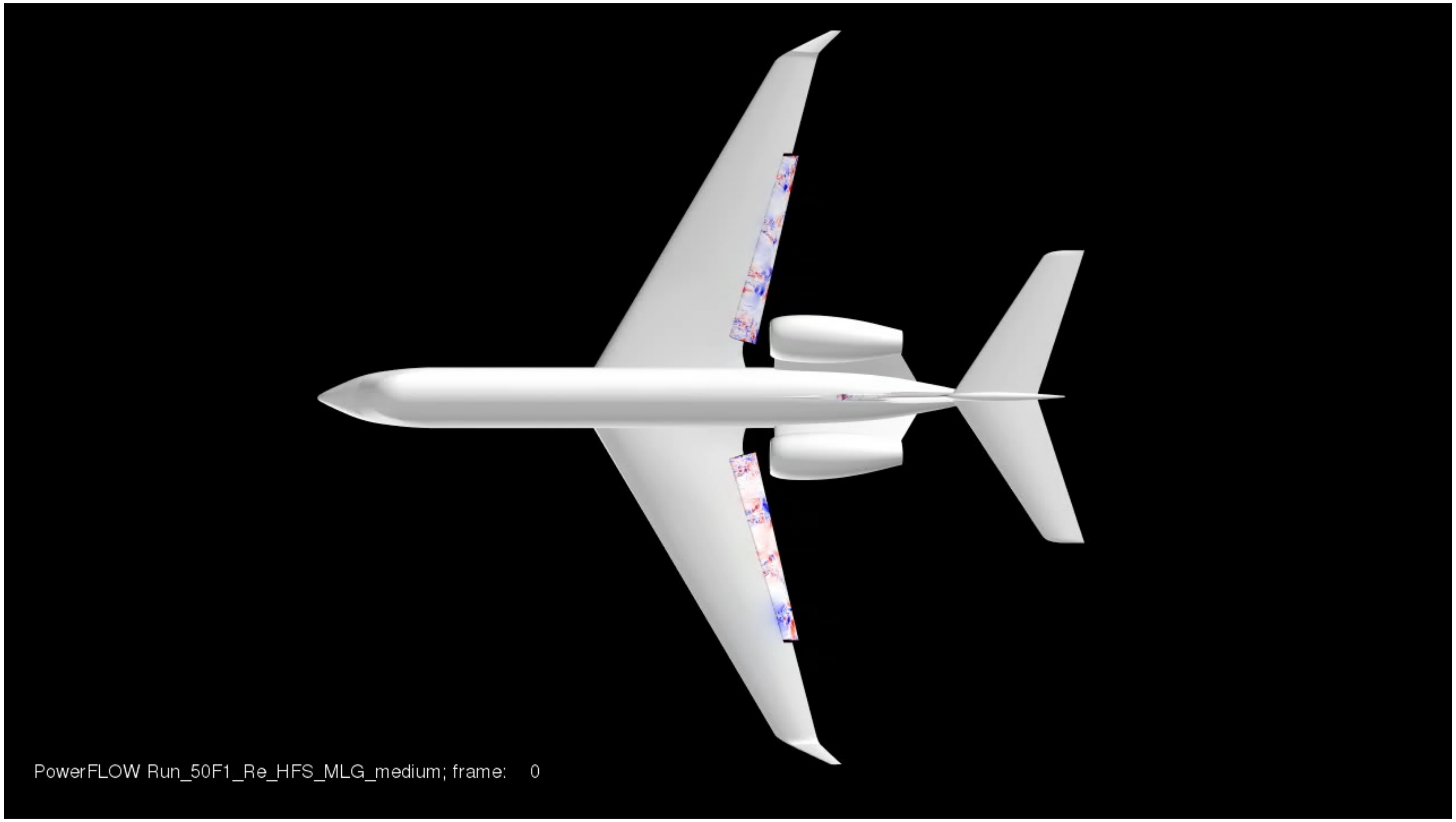
Colliding Black Holes



Orion Launch Abort



Flexible Control Surfaces



What Drives Increased Focus on Sustainability

- The National Environmental Policy Act of 1969 committed the United States to sustainability, declaring it a national policy “to create and maintain conditions under which humans and nature can exist in productive harmony, that permit fulfilling the social, economic and other requirements of present and future generations.
- The NASA Sustainability 101 website states: “NASA is an agency that leads by example and will continue to spur profound changes in our knowledge, culture, and expectations. NASA’s sustainability policy is to execute NASA’s mission without compromising our planet’s resources so that future generations can meet their needs. Sustainability involves taking action now to enable a future where the environment and living conditions are protected and enhanced. In implementing sustainability practices, NASA manages risks to mission, risks to the environment, and risks to our communities, all optimized within existing resources.”
- The HEC Capability Performance Capability Agreement (essentially the contract between the portfolio and NASA) states: “The third performance measurement area is focused on reducing the environmental impact of computation on our environment. The HEC Portfolio leads the Agency in reducing the amount of electricity and water needed to run computing systems. Moving forward, this will continue to be an area of focus. To measure this, HECC uses the Power Usage Effectiveness (PUE) metric— the most common method for calculating energy effectiveness in data centers. A facility’s PUE is the ratio of the total energy used by the facility to the energy used by the IT equipment.”

What in HEC Contributes to Sustainability

- Facility
  - New Facilities
  - Old Facilities
- System Operations
  - Procurement of new systems
  - Component Failures
- Code Performance
  - Effective utilization of node resources
  - Effective design of code

Facility Improvements

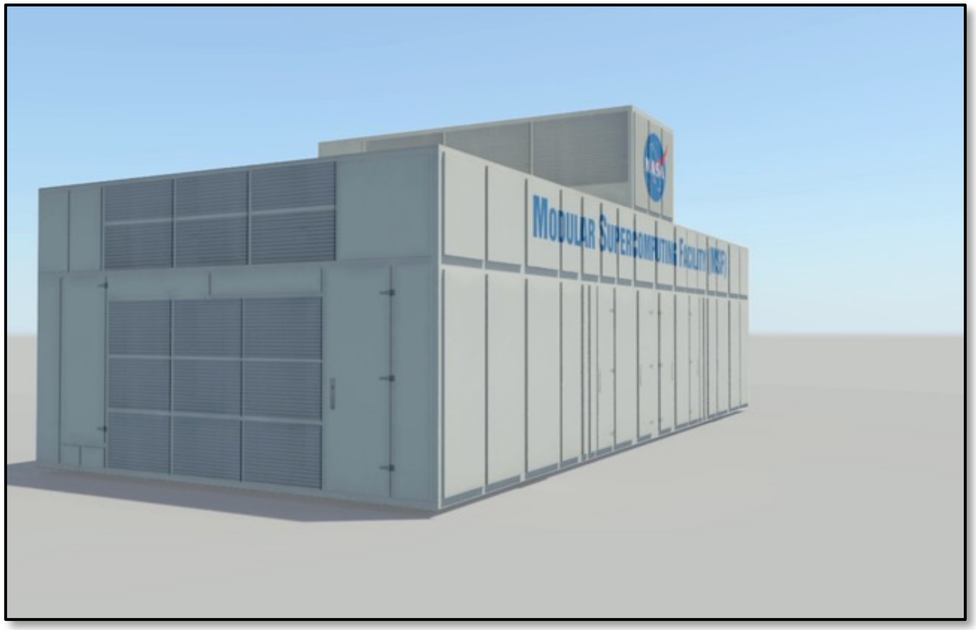
What Drove the Movement to Modular

- **The calculation had been very simple…**
  - When the cost of maintaining a group of nodes for three years exceeded the cost to replace those nodes with fewer nodes that did the same work, we replaced them.
- **In the early 2010's, not so much…**
  - We look at the total computing our users get and procure new nodes within our budget and remove enough nodes to power and cool the new nodes.
  - This means that we are not able to actually realize all of the expansion we are paying for.
  - Contracted a 3rd party survey in 2014-15, to determine the "best" approach for expanding the facility.

But That's Not All

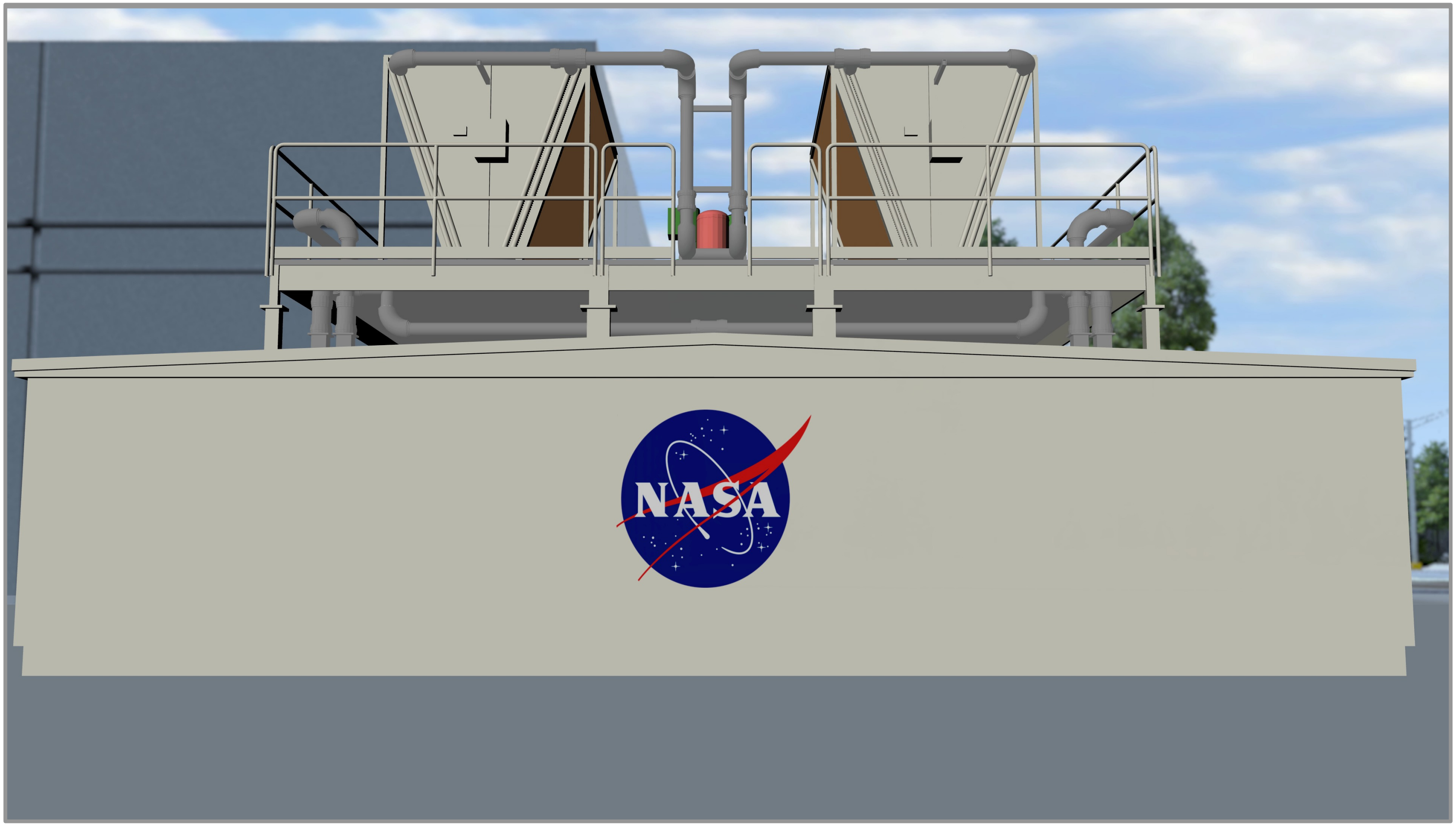
- **Our computer floor is limited by power and cooling.**
- **Our current cooling system:**
  - – Open-air cooling tower with four 50HP pumps.
  - – Four 450-ton Chillers.
  - – Seven pumps for outbound chilled water.
  - – Four pumps for inbound warm water.
- **Our Electrical System:**
  - – Nominally, the facility is limited to 6MW.
  - – 25% is used for cooling (1.34 PUE).
  - – 4MW – 5MW for computing.

First Prototype



Second Prototype Provided more Density

Module 2 Assembly



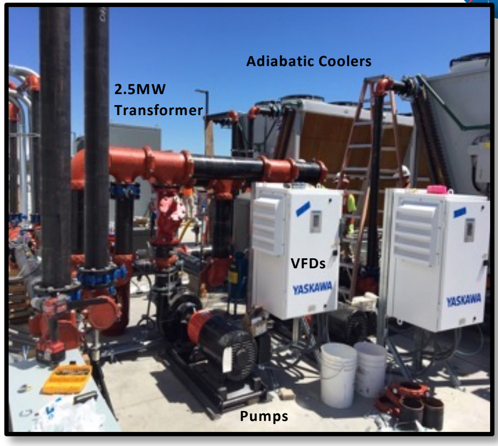
Annual Energy Impact from Prototypes

Electra System (2,304 Nodes)	N258 Facility	R&D 088 Facility	% Savings
Water Utilization Per Year	2,920,000 gal*	128,000 gal**	96%
Cooling Electricity per year	2,424,150 kwh°	221,026 kwh°°	90.9%

* Assumes 2,304 nodes represent 20% of N258 facility load
** Year 1 (Oct 2017-Sept 2018 actuals)
° 1.34 PUE (Oct 2017-Sept 2018 actuals)
° ° 1.031 PUE (Oct 2017-Sept 2018 actuals)

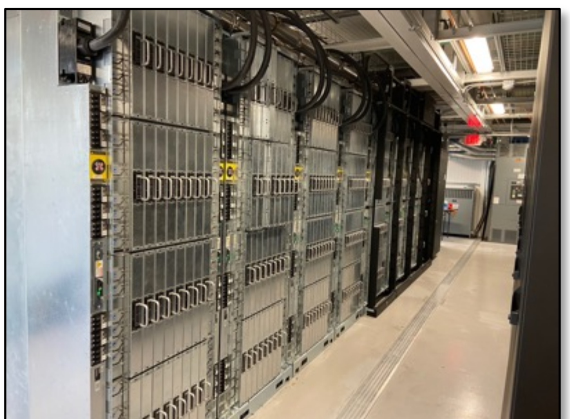
Facility Expansion 2017

- **In 2017, after successful completion of the prototype modules, HEC committed all future expansion to a modular data center facility.**
- **The 1-acre cement platform with a potential draw of 30 MWs has space for 12 compute modules and 3 data module.**
- **Deployed energy-efficient and cost-effective infrastructure capable of housing a 1,000,000-core.**
  - One-acre site can hold 12 compute modules and 3 data modules.
  - Site is 14 feet above mean sea level.
  - A 115kV breaker & 30MW 115kV–25kV transformer was installed in the N225B substation.
  - Conduit holds the electrical cables connecting N225B to R&D-099 and data cables connect N258 to R&D-099.
  - 25kV switchgear was placed on the site to distribute up to 30 MW to the modules.
  - A fire road and fencing were built.



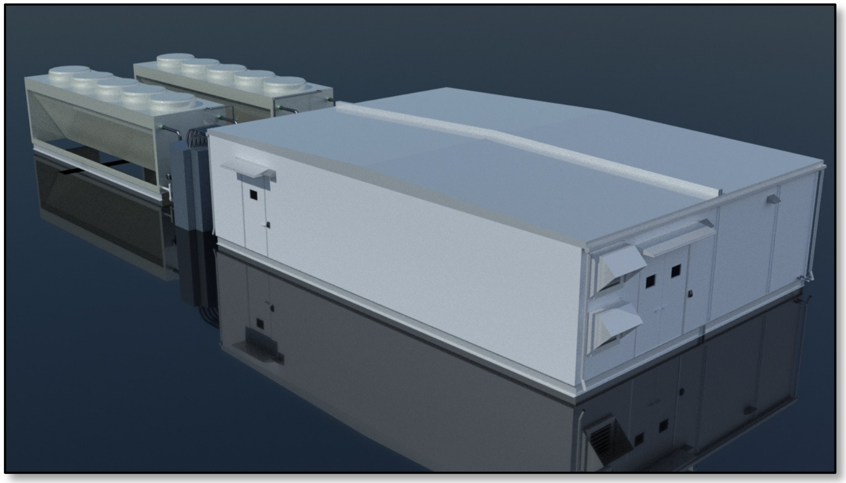
Module 1 Fully Deployed

- **The facility was commissioned on July 24, 2019.**
- **Since the initial deployment consisting of 1 row of racks, Aitken has expanded 3 times and is now fully populated with 3 rows.**
  - The system, Aitken, was named after the lunar crater, Aitken, part of the very large South Pole-Aitken basin. The crater was named after the astronomer Robert Aitken (12/31/1864 – 10/29/1951).
  - Aitken is connected to N258 through four 288-strand Single Mode fiber cables.
- **Aitken is now the largest (in terms of cores and memory) and fastest computer at NASA.**



Module 2 FY24

- A second module is scheduled to be delivered in March 2024.
- Will have 50% more power (3 MW) than module 1 and will house new nodes that will be added to the Aitken Supercomputer.
- The building management system will be separate but will be monitored in the control room in N258.
- Cables enabling the complex connectivity required for an HEC system will pass through openings planned to enable this.



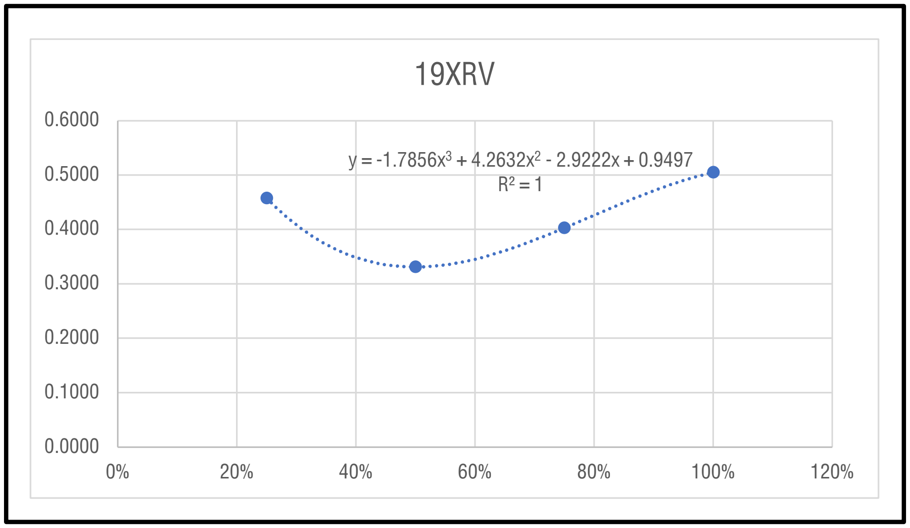
Modules Impact NASA and Drive Further Improvements

- **The impact of the modules on the PUE of the HECC component have had a positive impact on HECC and NASA.**
- **N233A will be phased out over the next few years as a data module is deployed on R&D099.**
- **The PUE in N258 will be reduced over the next 5 years after replacing the aging chillers.**

Facility	PUE - Oct 2022 thru Sept 2023
N258	1.37
N233A	3.14
R&D088 Mod 1	1.023
R&D088 Mod 2	1.045
R&D099	1.044
All Modules	1.042
All HECC	1.23

N258 Chiller Replacement Project

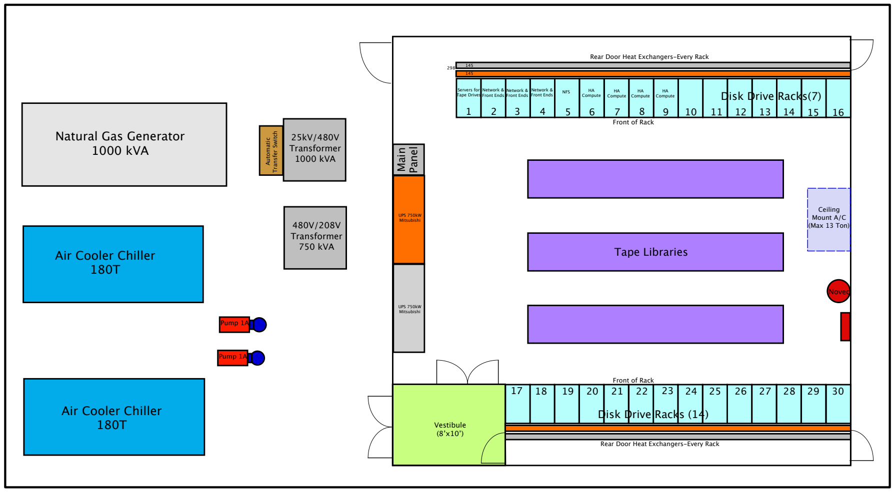
- **Project to replace our out of service chiller led to the redesign of our 80's era chilled water plant. The new design includes:**
  - Replacing all four 450 Ton chillers with four new 700 Ton Chillers.
  - Specifying the new chillers with Variable Frequency Drives
  - Operating the new chillers at 50% of design capacity to meet our 1100 Ton load
  - Installing new 480V transformer and electrical distribution to move away from the current 2400V; medium voltage VFDs cost as much as chillers!
- **The new chiller operating point will be 0.32 kW/Ton versus our current 0.55 kW/Ton (3 chillers at 350T each)**
  - Saving 2,200,000 kWh of energy annually.
  - Realizing $300K annual energy savings.
  - This reduces the N258 PUE from 1.37 to 1.30.



Performance Curve of new 700 Ton Chiller. Three chillers at 50% load meets our typical 1050-1100 Ton heat load while operating in the most energy efficient region of the curve.

Data Module FY25

- The Data Module will house the data storage (tape libraries, disk drives, solid state drives), distribute power to the data storage, cool the data storage, protect the data storage from fire and incorporate an entry vestibule for equipment loading.
- Total IT equipment power draw estimated at 584 kVA: 26 racks at 17kVA each, 4 racks at 34kVA each, and 3 tape libraries at 2kVA each .
- Chillers necessary to meet the environmental requirements for magnetic tapes; 61° to 77° F, 20% to 50% RH.
- UPS with generator backup for the entire module load, including infrastructure equipment.
- Target PUE for the module is 1.40.



Plan View of Data Module & Infrastructure

Operations Improvements

Lots of moving parts

- **The HECC component has 6 compute clusters and numerous file systems shared between all of the clusters.**
- **With ~18K Nodes and close to 4K switches, failures are common.**
- **If 10% of the nodes are down, we are losing close to $600K/year in power with no return.**
- **Actions**
  - Constant monitoring of compute node failures
  - Rigorous testing for return to service (a node that fails during a job affects all the nodes it is running with)
  - Fabric tests that check the stability of the networks
  - Annual process to determine if decommissioning is indicated.

Procurements

- Energy consumption is now considered when looking at the total cost of ownership
- Energy cost is based on energy/unit of work and takes into account where the IT will sit and what resources are needed to fully operate.

Software Considerations

Improving Code Performance

- The role of the Application Performance and Productivity Team.
- Migration to newer technologies that are higher performing while being more energy efficient.
- Monitoring job utilization of system resources.

Application Performance and Productivity Team

- **Help with application-related issues, for example:**
  - Porting your code to our systems from other systems.
  - Basic performance analysis and optimization.
  - Setup for parallel execution of numerous serial jobs.
- **In-house developed tools for applications and job scheduling:**
  - Info about running jobs
  - Node availability
  - License availability
  - Job startup
  - Memory monitoring
- **Specific results of these activities have resulted in runtime improvements in extreme cases of 70x**

Migration to New Technologies

- **Technologies like GPUs can potentially have game changing affects on time to solution facilitating better science.**
  - Work with industry and other government institutions to host and/or facilitate hackathons to rapidly indoctrinate our users with the approaches to run on the technologies.
  - Provide testbeds through on-premises or cloud test systems to evaluate the possible impact of the technology.
  - Provide users with the ability to try their codes on new technologies to see if it is a good fit for them.
  - Any migration of a portion of the user community to new more efficient technologies have a positive impact on all users.
- **Incorporation of faster networks and storage systems reduces time to solution and reduces energy requirements.**

Job monitoring

- **Looking at node availability is important but not all of the story.**
  - Nodes are comprised of many components and it is important that we get the most work for energy used.
  - Through monitoring we can identify jobs that are not using the processors on the node in a balanced manner leading to waste.
  - Working with the user we can determine the most sufficient means to run on the targeted architecture
    - Our Rome nodes have 128 cores on a node
    - It is sometimes more efficient to run on twice the number of nodes utilizing half the cores. This can result in a runtime that is less than half the original time.

Questions

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