NASA
High-End Computing Program
User Needs Assessment 2020

HEC Program Manager: Tsengdar Lee

December 11, 2020

Coordinator: Michael M. Little

Executive Committee

Tsengdar Lee, NASA Headquarters

Aaron Pina, NASA Headquarters

Piyush Mehrotra, NASA Ames Research Center

William Thigpen, NASA Ames Research Center

Daniel Duffy, NASA Goddard Space Flight Center

Robert Ferraro, Jet Propulsion Laboratory

Executive Summary

NASA's use of High-End Computing (HEC) resources, dating back to 1987, has been and remains a major and growing factor in the effectiveness and execution of NASA missions. Our understanding of the complex interaction of natural phenomena and constructed machinery is enhanced by sophisticated models, made possible with HEC. Data-driven models require the ability to process instrument output into usable data products. As observing strategies evolve to collect larger volumes of data, algorithms become more complex, resulting in the fusion of multiple sources across several interacting scientific and engineering disciplines into more sophisticated data products, all needing much larger computing capacities.

Supercomputing systems enable fast communications among large numbers of processors to enable accurate simulations of non-linear, non-equilibrium processes to produce high-fidelity model outputs. HEC systems aid in predicting the behavior of these processes under specific conditions, with accurate models run with sufficient resolution. Currently, the NASA HEC Program addresses eight classes of problems, which are often interrelated:

- Physics-based models.
- Data-driven models using machine learning.
- Science data processing from observations and instruments.
- Systems engineering.
- Mission planning and design.
- Analysis of data and model output for science and engineering.
- Uncertainty quantification and risk reduction.
- Control of scientific instruments in experiments and observation.

The HEC Program has the most direct impact on NASA missions in helping to solve large-scale technical problems. There are three primary ways the program supports these missions: discovery and use of new knowledge; support for managing complex and pathfinding programs; and creation of an inspirational work environment. Using HEC resources allows missions to reduce risk and cost, quantify uncertainty, improve mission planning and design, accelerate schedules, and improve the NASA workforce by attracting and retaining those most energized and inspired by working with HEC resources on meaningful problems at the very edges of science, engineering, and computing.

Learning the needs of the NASA workforce and researchers is foundational to developing strategic plans for the HEC Program. Every few years, the HEC Program conducts a user needs assessment to evaluate how researchers and users are consuming HEC resources, where they see gaps in service, and how they anticipate future HEC needs. The results are used along with other engineering studies to formulate a strategic plan.

From June 1–19, 2020, the HEC Program asked scientists, engineers, and researchers to discuss how they use NASA's HEC resources or how they expect to use them in the future. During this workshop, the users provided the HEC Program with context for their use cases and a greater understanding of how the presented projects fit into the larger NASA mission. Over 250 researchers and programs contributed use cases. These use cases have provided the foundation for the HEC Needs Assessment Report, which also provides an overall analysis of the major uses of the HEC Program and potential areas for growth.

Users of HEC Program resources stated that they generally consumed all the resources allocated to them, and that they were able to occasionally use extra resources. In discussions of future needs, the users described how they would use additional capacity, capabilities, and support. These users' needs largely fell into two categories: speed and volume. Several end users referenced long queue times and the high level of schedule pressure sometimes associated with runs. Complementary to this issue is the need to run larger-scale models and integrated models. Some users described restructuring their code and creating workflows to fit within the existing HEC processes.

Around the discussion of future workflows, users identified their need to understand how to use emerging architectures and computational capabilities. Users expressed a desire to have NASA HEC resources and services assist with migrating codes to future capabilities, leveraging portable code design, and redesigning applications past their initial generation. Among these emerging technologies include exascale computing. While other government agencies have larger scale resources available than NASA, users identified that proposals to use other agencies' resources are time consuming, which can take away from primary mission research.

The second section of this report provides more details of the use cases, and ways that the HEC Program supports each NASA Mission Directorate. These include the Aeronautics Research Mission Directorate, Human Exploration Allocation Group, and the Science Mission Directorate.

During the workshop, in addition to presenting their use cases and identifying perceived barriers they encounter while using HEC resources, HEC users described their major requirements from NASA's HEC Program. These include:

- Substantially more capacity than currently available in the current architecture.
- Improved processor access to the data.
- Expanded support by supercomputing experts to domain experts.
- Increased availability of new architecture.
- Stable funding streams for the ongoing evolution of frequently re-used codes.
- Management policies to enable widespread external collaboration, including computer security restrictions tailored to the applications.
- Management by metrics of user effectiveness instead of processor workload.

While the current HEC Program structure has the capacity to adequately address some of these problems, it may be inadequate to address expectations and demands associated with some of the Agency's anticipated future objectives. This is because of growing needs for multi-point analyses and uncertainty quantification to reduce mission risks, for planning missions that require atmospheric re-entry, and for fully realizing the reuse of facility codes without stable funding. This report outlines how HEC currently supports those mission objectives, and how the current program is constrained in supporting those objectives in light of these growing needs.

CONTENTS

FIGURES

APPENDIX A FIGURES

TABLES

APPENDIX A TABLES

Overview of NASA High-End Computing

NASA established its High-End Computing (HEC) Program in 1987 in response to the need for a stable investment stream to acquire, maintain, and refresh high-cost computing technologies important to NASA's programs—at the time, primarily Aeronautics research (Bailey & Kutler, 1988). Over the past 30 years, the various programs were consolidated to create a unified investment stream supporting a NASA-wide capability with broad-based policies and procedures responding to the needs of NASA's user community. Today, NASA's HEC Program delivers high-end computing systems and services to NASA's aeronautics, exploration, science, and space technology missions.

While HEC-supported organizations conduct daily interaction with their users, periodic reviews of users' needs are also conducted as part of programmatic strategic planning. Recent reviews occurred in 2008 and 2013. Today, four challenges warrant a new review and a new strategic plan:

(1) Growth of machine learning as a major workload on NASA's supercomputers.
(2) Conclusion of the era of supercomputing dominated by Moore's Law.
(3) Evolution of new computing architectures as a major, if not primary, source of computing capacity
(4) Growth of data volume and of model output.

NASA's HEC Program is an agency-wide resource managed by the Science Mission Directorate (SMD) with funding from all NASA Mission Directorates. The program uses an integrated management approach to computing and support for its two projects (Figure 1):

(1) The *High-End Computing Capability* (HECC) Project operated by the NASA Advanced Supercomputing (NAS) Facility at Ames Research Center (ARC).
(2) The *NASA Center for Climate Simulation* (NCCS) Project operated by the NCCS within the Computational and Information Sciences and Technology Office (CISTO) at Goddard Space Flight Center (GSFC).

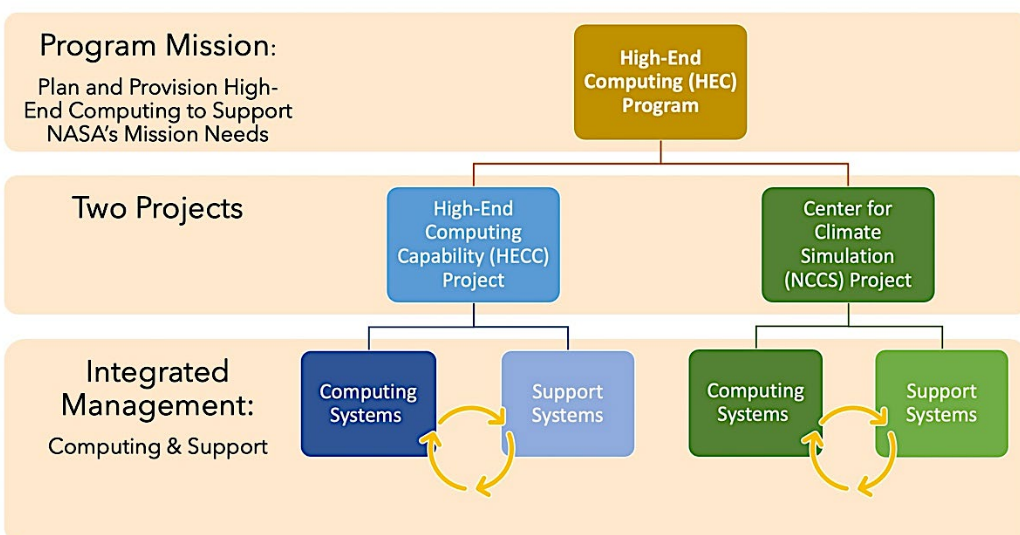

FIGURE 1. NASA HIGH-END COMPUTING PROGRAM HAS TWO PROJECTS WITH INTEGRATED MANAGEMENT OF COMPUTING SYSTEMS AND SUPPORT SYSTEMS WITHIN EACH PROJECT. (CREDIT: PETER WILLIAMS)

HECC serves the broad spectrum of NASA customers, while NCCS focuses on the Science community. Combined, these facilities provide more than 565,000 processor cores and over 26 quadrillion (thousand trillion) floating point-operations per second (petaflops) peak computing power to the NASA user community. The HEC Program manages these investments by providing high-level oversight and guidance. HECC's integrated resource and service offerings include high-speed networks, archival storage systems, system performance and application optimization, 24x7 user services operations, data analysis, and scientific visualization. In parallel with supercomputer growth, HECC continually develops advanced visualization techniques, including a cutting-edge visualization system, hyperwall which is connected directly to supercomputers, allowing scientists to run sophisticated concurrent visualizations (Table 0-1).

TABLE 0-1. NASA HEC COMPUTING SYSTEMS AND RELATED RESOURCES AT THE NASA ADVANCED SUPERCOMPUTING (NAS) FACILITY AND THE NASA CENTER FOR CLIMATE SIMULATION (NCCS). (SOURCE: HTTPS://HEC.NASA.GOV/ABOUT/OVERVIEW.HTML)

	NAS	NCCS
Systems	**Aitken** • Aitken Overview **Electra** • Electra Overview **Pleiades** • Pleiades Overview **Endeavour** • Endeavour Overview **Merope** • Merope Overview	**Discover** • Discover Overview ADAPT Private Science Cloud • ADAPT Overview SMCE Commercial Cloud • SMCE Overview
Storage	*Online:* 29 petabytes of RAID disk capacity *Archive Capacity:* 1,040 petabytes (1 exabyte) • Archival Storage Overview	*Online:* 75 petabytes of RAID disk capacity *Archive Capacity:* 150 petabytes • NCCS Storage Systems
Networking	SGI NUMAlink Voltaire InfiniBand • Networking Resources Overview	Mellanox Technologies InfiniBand
Visualization and Analysis	**Hyperwall-2** • Hyperwall-2 Overview	**Data Visualization Theater** **Hyperwall** **Hyperwall Cluster** **Control Station** • Data Visualization Theater Overview **ADAPT—Advanced Data Analytics Platform** • ADAPT Overview **DataPortal** • DataPortal Overview **Remote Visualization** • Remote Visualization Overview

The complete HEC ecosystem consists of both the elements supplied by the HEC Program as well as those supplied by the users. While the HEC Program supplies hardware, software management, and user support capabilities, the users supply their own expertise and knowledge as well as code that runs on the systems. The central focus revolves around using supercomputers, which have highly interconnected nodes enabling them to run large-scale, non-linear, non-equilibrium simulations on a linear platform. In this ecosystem, HEC users, their capabilities, and the code they acquire or develop, as well as the HEC Program-supplied components (Table 0-2), are all essential components for meeting NASA's mission objectives. Perhaps most importantly, the systems must be balanced for them to work efficiently for science applications. This is evidenced by the use cases contributed by the user community. Appendix A further describes the entire ecosystem.

TABLE 0-2. NASA HEC SUPPORT SERVICES OVERVIEW. (SOURCE: HTTPS://HEC.NASA.GOV/ABOUT/SERVICES.HTML)

User Services Support • NAS User Services • NCCS User Services Group **Advanced Visualization and Analysis** • NAS Visualization and Data Analysis Services • NCCS Data Visualization Theater • NCCS Remote Visualization • CISTO: Scientific Visualization Studio • NCCS Advanced Data Analysis Platform (ADAPT) • NCCS: NASA Climate Data Services	**Application Optimization** • NAS Application Optimization Services • NCCS User Services Group **High-Speed Networking** • NAS End-to-End Networking Services • CISTO: High-End Computer Networking • NASA Communications Services Office **High-End Computing Systems Development** • NAS System Services **System Performance Mass Storage** • NAS Data Storage Systems • NCCS Storage Systems

Three Mission Directorate groups are allocated a share of the SBU's available each fiscal year, based on a long standing agreement. User requests for accounts and allocations are submitted through the single online request tool and routed for review by the relevant Mission Directorate representative using their own criteria, as described on the HEC website. The Mission Directorate processes produce a list of allocations of SBUs which are used to notify the users of their available compute resources. As part of the HEC strategic planning, this process and the relative MD shares are expected to be re-examined.

As described in Appendix D1, the HEC User Needs Assessment is one of several studies and analyses feeding into HEC strategic planning. These include

- environmental impact,
- engineering estimate of computation and storage requirements,
- facilities capacity,
- Federal policies,
- workforce,
- technology assessment,
- business and financial analysis.

A comparative analysis with other supercomputing centers for best practices and resource sharing is also planned.

1. NASA HEC Applications

There are eight general classes of problems NASA scientists and engineers solve using HEC. The HEC Needs Assessment team organized the workshop sessions and grouped use cases into affinity groups based on these problem classes:

- Physics-based Models.
- Data-driven Models.
- Science Data Processing.
- Systems Engineering.
- Mission Planning and Design.
- Analysis of Data and Model Output.
- Uncertainty Quantification and Risk Reduction.
- Control of Instruments and Experiments.

FIGURE 2. EIGHT PROBLEM CLASSES ADDRESSED BY NASA HEC USERS. (CREDIT: PETER WILLIAMS)

1.1 Physics-Based Models

Physics-based models analyze equations representing large, complex, non-linear, and non-equilibrium systems with appropriate initial and boundary conditions. Typically, these are sets of partial differential equations constrained to operate in a known regime. While various numerical methods are used to solve these sets of equations, often the physical or transformed domain is digitized into a collection of discrete pieces to enable the solver. Solving these complex problems requires high computational demand.

For many of the complex problems NASA faces, physics-based models can solve those that cannot be solved with experimental systems and must be dealt with through simulation. For example, space-time coherence cannot be studied in a wind tunnel because air flow over a simple wing—described by the complex interactions of Navier-Stokes equations and thermodynamics—can only be solved with complex numeric techniques. Similarly, achieving accurate estimations for many flight regimes requires solving compounding problems. The addition of complex geometries, such as wing tips, roots, fuselages, and engines, makes computational jobs even more demanding and necessitates an HEC system.

Physics-based models allow NASA decisions to better reflect a robust understanding of risk. Unwarranted risks can be removed from consideration, low-probability/high-consequence risks can be better understood, and high-probability risks can be mitigated. This improves mission planning and design reduces costs while increasing confidence, and supports systems engineering.

Likewise, the characterization of weather in the atmosphere or water flowing through a river demands complex techniques with similarly high computational loads running on multiple nodes simultaneously, and requires each node to quickly communicate its results in each step to nearby nodes for each iteration. The reporting of states in each process creates additional loads in order to communicate and store data about each grid box for intermediate states as well as the final state—representing the characteristics of one time step—and then repeated for the duration of the computational experiment. Similarly, modeling the origin, structure, evolution, and destiny of the universe requires fully cosmological, high-resolution simulations of galaxy formation. Investigators run models almost continuously to examine different hypotheses and impose large demands for interconnected computing nodes, writing several gigabytes of output for each time step.

Whether modeling engineered devices in an operational environment or modeling natural phenomena in a robust natural system, physics-based models allow NASA to advance science and engineering beyond what is possible with experimental systems and do so while significantly reducing risk to human participants and experimental facilities such as wind tunnels. Full-physics models can also capture the understanding of phenomena by the scientific community, as it evolves. For example, the Land Information System (LIS) encapsulates physical Earth system models, data assimilation algorithms, optimization, and uncertainty estimation algorithms, all providing description of the current understanding of the physical processes.

1.2 Data-driven Models

The discipline of machine learning employs various approaches to teach computers to accomplish tasks where no fully satisfactory algorithm is available. The result of machine learning is a form of artificial intelligence about a system, whether engineered, natural, or interactive. Learning occurs through machines analyzing large volumes of observational data or model-generated output about a system to characterize and understand the system and its behavior better, more richly, and more robustly. Machine learning has four generally accepted functionalities:

- Classification.
- Prediction.
- Identification.
- Detection.

Scientists applying machine learning techniques can characterize physical systems when the full-physics models are not fully developed or are so computationally demanding that a run will not produce an answer in the available time. While the primary tools in use today are based on principles of supervised machine learning, unsupervised methods also make important contributions. Machine learning already has useful applications in many areas at NASA and is rapidly expanding into new NASA programs (Table 1-1).

TABLE 1-1. NASA EXAMPLES OF MACHINE LEARNING FUNCTIONALITIES AND CAPABILITIES.

Action	NASA Program Example
Segment and Classify	NASA Earth Exchange (NEX) analysis of Landsat data to characterize land surface types
Forecast	Potentially supplement full-physics models in aeroelastic simulation for certification by analysis
Forecast System Behavior	Laminar-turbulent transition across flight speed regimes from full-physics models
Super-Resolution Imaging	Deep learning applied to single images of hyper-spectral images, expanding to support time series
Spacecraft Health	Analysis of telemetry feeding into a digital twin
Classification	NEX analysis of GOES-16 satellite images to detect wildfire smoke plumes
Denoising Data	High-data-volume TESS all-sky data with 20,000 target stars to identify and discard bad data
Anomaly Detection	Bolide detection from GLM lightning data
Precursor Analysis	Determination of causes for aviation safety incidents

Supervised machine learning requires sufficient volume of data, whether observational, simulated, or experimental. Those data also must be properly labeled. Generally, the body of data available is divided into two subsets, one for machine training and one for learning validation, with some approaches applying a third subset for testing purposes. Each subset must be selected to reflect all the variations in data to create confidence in the accuracy of the model's ability to forecast future behavior of the system. Nevertheless, as in overfitting in statistics, machine learning can overtrain the model so closely that it can't be generalized to be used with new data. Often this occurs with observation sets that are too narrow or too similar. Training of machine learning models can run through large volumes of data and may become long-running jobs stretching into days or weeks in extreme cases and become computationally expensive. However, the payoff is a trained model that executes very quickly and inexpensively.

Users report successful reduction of training time when machine learning occurs in new hardware architectures – specifically, a Graphical Processing Unit (GPU) cluster designed for machine learning. This type of acceleration can make a substantial difference in meeting a tight schedules by reducing the computational time involved in training supervising learning models. It also can offload CPU-based supercomputers.

Similarly, users at both NASA HEC Centers report successful experiments in hardware acceleration of machine learning. Initial indications are that hardware acceleration is beneficial in autonomous observing, robotics, processing of models (approximation of partial differential equations), computer vision, analysis of large volumes of data, and text analytics. The use cases presented during the Needs Assessment indicate the usefulness of this class of computing and suggest it will continue to grow at NASA.

Machine learning can accelerate the discovery process, as illustrated by the Transiting Exoplanet Survey Satellite (TESS) mission. Traditional techniques are labor intensive, but two different machine learning techniques running on a HEC supercomputer accelerated and systematized the process of looking for multi-planet systems. The first, unsupervised deep learning methods, created high-dimensional embedding spaces to find light curves that exhibit similar features among the millions in the dataset. The second, supervised machine learning, performed the classification of light curves from stars observed. Both workflows received support from HEC resources that helped improve the automation and streamlining necessary to proceed without human intervention.

Unsupervised machine learning is less widely used at NASA, but even this limited use allowed discovery of unexpected characteristics. For example, a project under development at GSFC uses genetic programming algorithms to forecast ocean chlorophyll-a characteristics under varying climate conditions. Other applications led to the discovery of magma movement using Global Navigation Satellite System (GNSS) data and of potential spacecraft landing sites on other planets.

Like physics-based models, machine learning can reduce risk, allowing a more robust understanding of systems without the need for physical or virtual experiments—or, when experiments are not possible, as is often the case when planning missions to other planets, moons, or asteroids. Machine learning can produce new scientific discoveries and new engineering applications.

1.3 Science Data Processing

Science Data Processing (SDP) refers to workflows that translate instrument and sensor output into meaningful measures of system characteristics. Instrument output data is transferred from the point of collection to a storage system; processed to integrate appropriate instrument characteristics along with measurement characteristics, like time and location; and then processed through a series of algorithms to produce measurements, observations, or visualizations. Some important characteristics are not directly observable, requiring the fusion of data from multiple instruments or measurement techniques to create a numeric representation of the system state.

With SDP, the quantification of uncertainty in the representation of those indirectly observed characteristics is equally important. Each data point can reflect the results of complex calculations. As output volume from instruments and experimental facilities grow, HEC becomes more critical to the SDP stage of an experiment. Once meaningful measures are available, experiment success can be evaluated and assessed through a variety of techniques.

In 2019, Project Red Rover connected the core competencies of HEC to experimental facilities to run wind tunnel tests more efficiently. During the test campaign, over 600 data points and more than 150 terabytes of data were acquired and securely transmitted from the NASA Ames Research Center's Unitary Plan Wind Tunnel to the computing center at the NASA Advanced Supercomputing (NAS) facility. In preparation, NAS experts parallelized portions of the code used to process the unsteady pressure sensitive paint (uPSP) data, dramatically reducing processing time from one day per test point to one minute per test point. By connecting the power of this high spatio-temporal measurement to the power of the Pleiades supercomputer, along with the ability to quickly process and visualize the results on the NAS facility's hyperwall, near-real-time design decisions become possible—establishing a new path toward decreasing design cycle time and mitigating over-design of aerospace vehicles due to insufficient tools.

The Kepler Space Telescope, the Transiting Exoplanet Survey Satellite (TESS), the Stratospheric Observatory for Infrared Astronomy (SOFIA), and the Geostationary Carbon Observatory mission (set to launch in 2022) all rely on HEC resources to perform SDP. In NASA's Earth Science Division, purchases of commercial satellite data are being re-processed using HEC resources to improve calibration and geo-location, enabling their use in NASA research programs.

SDP, made feasible by NASA HEC resources, expedites mission design and vehicle design by making real-time decisions possible, reducing design cycle time, and reducing the risk of over-design of aerospace vehicles. SDP also supports fiscal and financial efficiency by allowing re-processing of commercial data to enable use in NASA programs.

A growing component of SDP is the creation of new data products through fusion of data from multiple instruments. To date, this is largely done by individual scientists or small teams for specific analyses, however, users indicate the need for large scale data fusion will grow.

1.4 System Engineering

Engineering studies as part of system design rely on HEC to run the complex models and analyses necessary to optimize system performance. Over time, the complexity of the optimization tools has grown to achieve the fidelity needed for these complex systems. These are essential to the design of new generations of launch vehicles and their components.

Modeling and simulation allows behavior prediction under conditions where physical experiments are not possible. The modeling tools, however, require many runs to ensure the confidence in the results necessary to minimize risk and optimize design. NASA researchers have developed, validated, and expanded multiple codes to a wide range of conditions and collaborative environments. For example, some codes can be broadly shared and, thus, commercial firms use these codes to support product development and deployment. Other codes in support of NASA's work are subject to International Trafficking in Arms Restrictions (ITAR) and Export Administration Regulations (EAR), limiting where the codes can be shared. These restricted codes can only be run on appropriately authorized NASA HEC assets. This is a source of additional computing load when commercial firms enter into a formal agreement with NASA to conduct their analysis.

Evolution of the *digital twin* approach to system design and operations—so innovative some call it a new paradigm—is expected to increase demand for computation, particularly in near-real-time applications (Glaessgen & Stargel, 2012). Conceptualized by NASA and the U.S. Air Force as a tool for certification, fleet management, and sustainment, the digital twin integrates ultra-high-fidelity simulation with the on-board integrated vehicle health management system, maintenance history, and all available historical and fleet data to mirror the life of its flying twin and enable unprecedented levels of safety and reliability. HEC is supporting development and integration of this technology, including the models and real-time data stream, with the need expected to grow rapidly for spacecraft, crewed and uncrewed aerospace vehicles, and related missions. Further information can be gained from Appendix D0.3, the Keynote speech at the Workshop by Dr. Karen Willcox.

Model-based systems engineering (MBSE) is a systems engineering methodology that focuses on creating and using computational models as the primary means of information exchange between subsystems (or their engineering owners) instead of document-based information exchange. Such subsystems may correspond to those within the flight and ground systems, e.g., attitude control, power, ground antenna; they may also correspond to stakeholder objectives and science traceability, e.g., measurement needs, instrument specifications. MBSE provides the means to quantify interdependence between requirements vertically and cross-functionally in the project design phases, and provides traceability of final specs during the verification and validation phases. In such a framework, HEC can be used to support tradespace analysis to optimize design across potentially conflicting objectives and constraints, verification across a multitude of unit tests, automating validation across varying scenarios and use cases to stress the system, fault trees and root cause analysis, etc. HEC resources for MBSE are especially useful in spacecraft robotics, or in applications with large uncertainties, to capture system response and performance across varying parameters, e.g., cyclone monitoring with constellations, tasking instruments for observing weather predictors.

A major element of systems engineering is the evaluation of operational characteristics and anomalies—in extreme cases, accident or mishap investigations. HEC assets are used to support these unexpected activities in response to some specific event and generally provided on a quick-reaction basis. The most extreme case is the Columbia Accident Investigation (Gehmann, Barry, Deal, Hallock, & Hess, 2003), in which HEC resources were redirected to run the engineering analyses needed to support the investigation. At a much smaller scale, engineering studies of component changes are important to reduce risk, such as the recently started P3 pylon engineering assessment, which will improve the operational availability..

1.5 Mission Planning and Design

Historically, mission planning and design efforts were manual optimization processes where simulations of orbital geometries, uplink/downlink schedules, and observation coverage were run offline with the results returned to the main discussion. Today, the critical need is to perform integrated, multipoint mission planning optimized with all the major simulations running together to characterize the mission performance more fully. Two primary drivers for this change are the emergence of interfering traffic in a denser low Earth orbit environment, and the increased demand on downlink capacity for large data volumes. Optimization cannot be performed using single-point solutions, but the entire range of solution space must be evaluated.

As an example of this new complexity, when the mission is a constellation of satellites, tools such as NASA's Trade-space Analysis Tool for designing Constellations (TAT-C) and the AGI's Systems Tool Kit (STK) require more and more computational capacity to generate alternative choices, as well as additional capacity to permit analysis of those options. As low Earth orbit becomes crowded, analysis of any maneuver requires simulations to ensure collision avoidance while maintaining the ability to accomplish the observational mission. Similarly, planning for missions to celestial bodies has grown in complexity to achieve a more precise landing. Future missions to the Moon and Mars, where proximity to previous equipment or settlements must be considered, will require computationally intense, multi-point analysis to obtain the optimum flight paths and to minimize the risk of the mission.

Some missions require vehicles to perform entry, descent, and landing (EDL) into an atmosphere. EDL scenarios were simulated for Mars lander missions for relatively small payloads and are currently being simulated for up to two metric tons, limited by the computing capacity available. Two EDL elements require full physics simulations. First, parachutes are used to slow and stabilize the vehicle; in some cases, they require multiple deployments. Second, supersonic retropropulsion is used to slow descent from hypersonic speeds and to create a low-impact landing.

Because contemporary mission simulations involve multiple interacting disciplines, such as flight mechanics, flight environment, radiative heating, impact structural dynamics, and response of the thermal protection systems, they must couple thermal, structures, and fluids codes. Current flight mechanics simulations only involve inputs from bank angle modulation for guidance. The primary code for use in flight mechanics is Program to Optimize Simulated Trajectories II (POST2), which has been used successfully to solve a wide variety of atmospheric ascent and entry problems, as well as exo-atmospheric orbital transfer problems. POST2 includes atmosphere, aerodynamics, and gravity. This code was developed at NASA's Langley Research Center (LaRC) over 40 years ago and is in wide use by other agencies, universities, and commercial space companies. Because of its importance, POST2 code development has been funded periodically to upgrade, enhance capabilities, and migrate into C/C++ from the original FORTRAN. An integration of FUN3D into POST2 is ongoing. A typical mission planning exercise involves running thousands of simulations to quantify uncertainty and to obtain statistics for the landing location. Development of an Onboard Autonomous Trajectory Planner (OATP) has begun as a necessary step to landing 20 metric tons on Mars to support crewed landings there.

Satellite mission design requires study of how the proposed design will impact the capabilities of the future observing system. An assessment of the cost-effectiveness of adding the proposed instruments to weather forecast modeling systems. Observing System Simulation Experiments (OSSEs) are conducted by simulating the key characteristics of the new instruments on a forecasted future environment, called a *nature run*. The Goddard Modeling and Assimilation Office (GMAO) collaborates with many Earth science mission design projects to provide simulations of a future atmospheric environment along with simulations of the instrument characteristics. NASA's Jet Propulsion Laboratory (JPL) also uses OSSEs in satellite instrument design tradeoffs, including swath coverage, location of ground stations, instrument duty cycles, sampling frequencies, etc. OSSEs create a significant demand for computation and data storage to enable optimization of the investment in new missions, but significantly improve the probability of a new mission's success in providing observational data that will improve future forecasts.

Application of HEC to mission planning and design allows mission schedules to have a more aggressive timeline with less risk to vehicles, payloads, and crews, whether those of the subject mission or those of other missions. There are direct, indirect, and cumulative benefits to NASA mission goals, such as advancing science, fiscal responsibility, and engineering excellence.

1.6 Analysis of Data and Model Output

In addition to the physics-based and machine-learning-based modeling and analysis described in previous sections, the HEC Program supplies computational capacity and other special capabilities to enable analysis of both model output and observational data as might be produced by Science Data Processing (SDP). Specifically, sophisticated statistical techniques for both machine learning and visualization require powerful computing capability. Similar techniques can be used for both areas, whether generated by the output of instruments measuring actual conditions or by the output of models. As the need to handle larger data volumes grows along with the need for fast-turnaround analyses, the research community and the HEC staff develop faster techniques for producing computed products.

Once data are available, HEC supports further processing to conduct analyses for science and engineering. Depending on the mission requirements and objectives, researchers use HEC to perform data cleanup, identify trends and anomalies, categorize imagery, collect states into bins, or further characterize the performance of a system. HEC supports data assimilation to fill in blank spots or to smooth the data to reflect continuity of physical systems.

Data visualization is an especially powerful and increasingly relevant analytical tool because it can compare theory (simulations) against measurement, or produce more understanding from long time series of measurements. Simple plots of one parameter against another for small volumes of data can be generated on a small computer. However, HEC's greater computational capacity and sophistication become necessary, as data volumes and geospatial resolution increase, or time series have finer steps. Still more HEC resources become required for interactive visual analysis of large volumes of data. NASA's HEC Program, with its integration of processors, storage, and communications in conjunction with sophisticated visualization tools and specialized staff, is essential to producing an accurate and meaningful data-driven visualization.

Historically, non-destructive evaluation (NDE) of aerospace structures was an experimental discipline, seeking to detect, quantify, and characterize damage inside materials and structures through mechanisms that leave the structure intact and unchanged. It is another area where computational capabilities have greatly extended the discipline beyond what can be done in an experimental facility. By providing simulation, modeling, and data analysis tools, computational NDE provides information during the design stage that leads to inspectable designs as well as designs that either remove or compensate for the consequences of possible failure modes, damage, and flaw cases. New techniques for analyzing large volumes of data and integrating them into models reduce inspection time and cost, illustrate non-obvious trends and relationships, and support decision making. Such techniques are also useful in reducing the subjectivity of human interpretation of NDE results.

Data analysis and visualization of large volumes of model output often go hand in hand. For example, researchers in the Estimating the Circulation and Climate of the Ocean (ECCO) consortium run large scale, high-resolution simulations on HEC assets to produce global maps of the Earth's ocean and sea-ice system at ~1-kilometer (km) resolution. The model output is then analytically compared with in situ sensors and NASA ocean satellite observations for validation. A key feature of ECCO efforts to reduce time to science is the proximity of the data to the computing platform and the harmonization of the analytic tools with the data structures, thereby enabling the science team to focus on analysis, not data wrangling. They can easily perform feature detection, classification and segmentation, and clustering of similar features, among other analyses. Researchers also use the NAS hyperwall to examine scalar and vector fields at various depths, enabling a better understanding of how circulation, chemistry ,and biology collectively interact with atmospheric carbon. This combination of analytics and visualization enables the investigation of specific events, such as how a pollutant plume or debris field might spread from its source.

The NASA Earth Exchange (NEX) analyzes the data from geostationary satellites, both in retrospective and in near real time. With HEC resources, NEX is able to detect and locate wildfire origins within minutes of the smoke becoming visible. Other NEX projects have analyzed time series data from a combination of satellites to evaluate cause and effect relationships, such as drought in the Amazon rain forest. NEX also was able to scale up machine learning and computer vision techniques on HEC computing resources to develop a forest disturbance map for the entire U.S. at 30-meter (m) resolution from Landsat data. This capability not only allows an estimate of the biomass, but also track causes of change and their impact on the carbon inventory.

Because of the availability of HEC resources, the Advanced Rapid Imaging and Analysis (ARIA) science data system has been able to analyze synthetic aperture radar (SAR) data at various stages of processing in order to triage failure modes and develop a methodology for automatically redirecting images into error-correcting workflows, if needed. This operational analytics capability enables much higher throughput for a lower cost. This project accelerated the development of efficient algorithms in preparation for the NASA-Indian Space Research Organization SAR (NiSAR) mission.

Heliophysics integrated modeling relies on the high-capacity, highly interconnected HEC resources to create model output for studies. The output from these models can then be used to detect and visualize the characteristics of the important solar processes, such as acoustic waves excitation, magnetic structures formation, solar corona structure and dynamics, jets and eruptions, and small-scale dynamos.

In the SMD's Astrophysics Division, the RomulusC cosmological simulation of a galaxy cluster has been demonstrated with the highest level of resolution to date and requires significant HEC resources. This simulation has allowed the study of ultra-diffuse galaxies, and analysis of the large volumes of model output on NASA resources has revealed much about their physical processes.

> *Over the past century, major advances in astronomy and astrophysics have been largely driven by improvements in instrumentation and data collection. With the amassing of high-quality data from new telescopes, and especially with the advent of deep and large astronomical surveys, it is becoming clear that future advances will also rely heavily on how those data are analyzed and interpreted. New methodologies derived from advances in statistics, computer science, and machine learning are beginning to be employed in sophisticated investigations that are not only bringing forth new discoveries but are placing them on a solid footing. (Siemiginowska et al., 2019, p. 3)*

In the Planetary Science Division, model output for studying planetary and exoplanetary atmospheres and the Juno spacecraft's datawave using Jovian dynamo simulations relied on computational resources and analytic tools available through the HEC Program.

1.7 Uncertainty Quantification and Risk Reduction

The goal of Uncertainty Quantification (UQ) is to produce reliable model predictions with associated confidence levels in the presence of uncertainty. High-quality UQ reduces risk and allows designers to decrease unneeded margins, thus reducing cost. UQ also gives decision makers more understanding of the nature of alternatives and confidence in decisions. Two different UQ approaches are being used by HEC users, one using multi-model Monte Carlo simulations, and the other using physics-informed deep learning. A third approach will efficiently fuse the two approaches through use of physics-informed generative adversarial networks (PI-GAN) (Yang et al., 2019). UQ techniques require many more simulation runs than alternative approaches. The advantage, however, is that those runs provide significant payoff in understanding the strengths and weaknesses of predictions and their limitations for use in decision making.

The aeronautics community has identified improved management of uncertainty—especially automated management—as a significant need, for example in the CFD 2030 study (Slotnick et al., 2014), because of the fundamental role it plays in increasing solution accuracy, understanding the sensitivity of results under varying conditions, and enabling future advances in simulation capability. This is especially true for reliably predicting turbulent-separated flows, currently recognized by the aeronautics community as an obstacle to maturing use of computational fluid dynamics (CFD) in aerospace design. Improved management of uncertainty is essential to producing experimental data needed to reduce uncertainty further, which is a prerequisite for creating digital twins—essential for certification by analysis.

In the science modeling community, UQ is considered to be a critical estimate of the usability of model predictions. Only with the availability of computing capacity can UQ be adequately evaluated. In Earth science, discussed previously in this report, UQ is a growing feature of both machine learning and full-physics models—using ensemble prediction approach with its high demand for HEC resources—in hydrology, atmospheric composition research, geophysics, and air quality and climate modeling. In Astrophysics, it figures prominently in planning for the next decade.

1.8 Control of Instruments and Support Equipment

Analyses of instrument output to influence future measurements and observations is an evolving HEC application. As instruments and support equipment become more sophisticated, complex, and expensive and seek to satisfy growing demand, more complex, low-latency interactions between the science data processing of the output and instrument control become necessary.

Today, HEC supplements experimental facilities to help guide selection of experiments and tune them to maximize the value of data collected. HEC is crucial for these purposes, because of requirements associated with high data volumes, complex algorithms, and fast turnaround. For instance, HEC helps experimental investigators configure follow-ons tests, shorten test campaigns, reduce costs, and make collected results more useful. The ultimate objective of the Red Rover project, for example, is computer-aided control of wind tunnel tests. This integration of physical experiments and HEC resources promises to unleash new knowledge while reducing risks to facilities and other investments.

Both NASA and commercial space organizations seek to optimize the effectiveness of their investments in certain remote sensing scenarios through the use of autonomy (NASA, 2018a). A future use of HEC will be to perform the low-latency, complex calculations needed to assess observational data and, in turn, direct instruments onto more valuable observations. Semi-autonomous or even autonomous operations will rely on low-latency collection and processing of instrument output to feed into the control subsystem. One application is enabling models to task satellites to collect data needed to improve the skill level of weather forecasting.

NASA HEC resources support learning needed to engineer mission instruments and support equipment, as well as to operate those instruments and equipment autonomously in distant and/or harsh operating environments. With HEC, instruments and equipment are developed faster with less risk and less cost, yet greater confidence in operational products. Once operational, those products produce better science and more relevant data, and it all happens more safely and more durably than would otherwise be possible.

2. HEC Impact on NASA Missions

HEC has the most direct impact on NASA missions in solving large-scale technical problems faced by the agency, whether scientific or engineering, as described in Section 1.1. More details about anticipated needs of specific use cases are described in this chapter.

HEC supports NASA missions in three primary ways:

- Discovery and Use of New Knowledge.
- Mission Management and Decisions.
- NASA Work Environment.

2.1 Discovery and Use of New Knowledge

> *NASA inspires the world with our exploration of new frontiers, our discovery of new knowledge, and our development of new technology. Our work benefits Americans and all humanity. Since NASA's inception in 1958 to present day, the Agency's history is written with each unique scientific and technological achievement. We have landed people on the Moon, visited every planet in the solar system, touched the Sun, and solved some of the core mysteries of our home planet (NASA, 2018b, p. 6)*

Very little of the inspiration NASA seeks to offer the world is possible today without the unique computational capability supplied by the HEC Program. Since 1987, NASA's HEC has evolved to become an ever-more effective tool for achieving the agency's objectives, and has become pervasive in its use throughout the agency. Because it has been readily accessible to the research community, the HEC Program enables discovery and exploration. This document describes specific mechanisms in NASA HEC Applications (Chapter 1, p. 5).

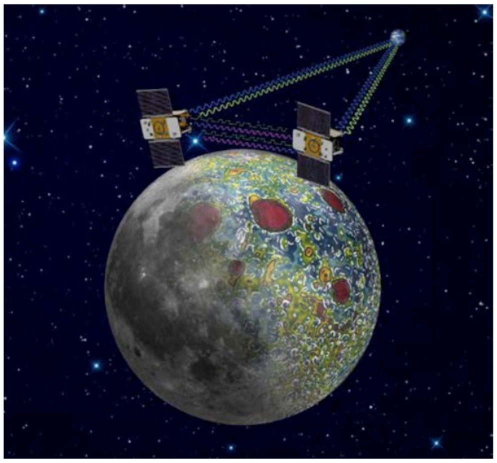

FIGURE 3. DISCOVERY AND USE OF NEW KNOWLEDGE INCLUDES GLOBAL GRAVITY MAPPING OF THE EARTH'S MOON IN 2012 THAT IS STILL PRODUCING NEW INSIGHTS. (CREDIT: NASA)

2.2 Mission Management and Decisions

HEC supports management decision making in NASA missions and programs throughout the agency by reducing risk and cost, quantifying uncertainty, improving mission planning and design, accelerating schedules, and improving the NASA workforce. This is especially valuable to complex and pathfinding programs.

Mission management decisions require a robust understanding of mission risks, and each mission decision is examined and informed by many others. Selecting a course of action requires information about the projected future outcome of many alternatives and about the confidence in those projections. Use of models to predict these outcomes and to quantify the uncertainty of those forecasts has become pervasive among NASA's mission teams. During the Space Shuttle era, for example, HEC resources performed quick turnaround analysis of tile loss during missions, essential for re-entry decisions (Figure 4). The team assessed the validity of these predictions based upon robust understanding of each alternative and of the analytical confidence appropriate for the results.

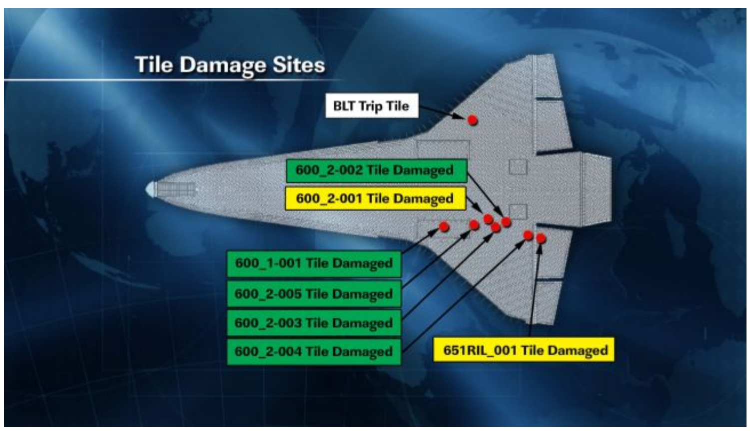

FIGURE 4. STS-134 POST-MISSION MANAGEMENT TEAM BRIEFING MATERIALS WERE DEVELOPED IN PART WITH HEC RESOURCES. (CREDIT: NASA)

Risk reduction is an extension of understanding risk. The HEC Program improves the chances of successfully understanding and reducing risks by providing a computational environment to experiment with situations that cannot be created physically. Some experiments examine failure modes—situations where a physical experiment would yield insight only at catastrophic and expensive scales. In scenarios that would occur on other planets, it is not possible to create experimental conditions, so applying physics-based models is the only way to gain insight for many mission planning decisions. In still other cases, the required experimental facilities are in high demand and scheduled for other tests during the period in which information is needed for an informed program decision.

Reducing risk while optimizing mission performance is exceptionally difficult when multi-point solutions are required, which is especially true when seeking solutions across multiple, interacting disciplines. For example, for many decisions an optimum solution is only possible with continuous refinement as well as evaluation of the viability of the results. Today, single-point solutions provide insight into the viability of a mission, but multi-point solutions are needed to optimize mission performance. All HEC problem classes include challenges related to multi-point solutions, especially for problems involving multiple disciplines.

Another way HEC supports mission management is by assisting NASA experimental facilities that are frequently overbooked or require long setup times. Models developed to reflect the historical results of previous tests are used to build a virtual instrument or test facility, which can reduce the extent of testing needed to verify model fidelity. Computational experiments are often used to obtain faster, cheaper insights and to provide more flexibility for running variations on test articles. Running computational experiments in parallel leads to insights in shorter times than otherwise possible in a physical facility while still retaining confidence in the fidelity and results of the model.

Lastly, HEC computations permit informed decisions based on forecasting system characteristics in non-observed states, undergirded by a validated understanding of the science used in the models. Informed by HEC computations, users can evaluate the level of confidence appropriate for a decision. With computational uncertainty quantification, mission planners can put numbers on the risk level, based on empirical data and valid simulations rather than conjecture.

2.3 NASA Work Environment

HEC impacts the work environment as much as it impacts the work itself. Providing exciting, interesting, and meaningful work allows HEC to contribute to NASA recruitment and retention (Figure 5). Historically, for example, new employees working in NASA's science and engineering programs were attracted by the unique and inspiring nature of the agency's work. A related benefit was the availability of unique and inspiring resources with which to tackle those projects, including colleagues, facilities, and computational resources. More recently, however, key hires have been lost to competing high-technology companies and other government agencies because the current NASA computing environment is seen as comparatively less attractive and, thus, less competitive, or because there are unanswered questions about the adequacy of resources to pursue the potential hires work.

FIGURE 5. NASA INTERNS FIND EXCITING WORK AS PART OF NASA MISSIONS SUPPORTED BY HEC RESOURCES. (CREDIT: NASA)

NASA workforce development and retention is another important area to which the HEC Program contributes. Specifically, it contributes significantly to the skill level of NASA's workforce with training and frequent opportunities for researchers to present their work and teach others how to practice more and more advanced techniques. Most science and engineering students graduate without the skills or knowledge needed to use modern high-performance computing systems, with only a passing knowledge of tools from resources such as GitHub and Jupyter Notebooks to code tuning tools. Training offered by the HEC centers gives them that introduction and helps to strengthen those skills to the point where they can be effective. Users further advance their skills in numerical methods and specific analytic and visualization techniques through training from the HEC centersAnother impact on the work environment is in attracting and growing collaborative team efforts.

HEC queueing strategies, often driven by mission directorate decisions regarding prioritization, affect the NASA work environment in important ways. Policies related to the availability of NASA HEC resources have contributed to democratization of NASA's research, which is often seen as a positive consequence. The computing resources are accessible to any NASA researcher, including grant recipients, based upon individual mission directorate allocations. For example, NASA's Science Mission takes the position that there is not a good way to predict the potential scientific value (outcome) from a particular computational modeling project. Science Mission's queueing strategy is fundamentally flat. Beside the "operational queue," which addresses the time sensitive mission development and engineering workload, all the other research and development workload are treated equally and prioritized the same.

These queueing strategies, however, have led to unintended inequities for those researchers addressing emerging science and engineering problems that require high-node-count jobs. For example, many NASA HEC users expressed frustration in using the agency's HEC resources because of the way current queueing strategies affect high-node-count jobs. Such jobs, which can only run on supercomputers, can be delayed—in some cases for weeks—while smaller jobs requiring lower node counts run with little or no delay. The HEC Program has created an informal process for exception handling, but this occurs infrequently, and the process is not suitable for regular campaigns involving several high-node-count jobs running regularly. As this involves tradeoffs among NASA polices, Mission Directorate management decisions and priorities it will be considered as part of an engineering study feeding into HEC strategic planning.

The HEC Program significantly contributes to the collaboration of teams formed across different organizations and locations. Because the resources are managed as an agency asset, teams that are formed because of shared interest and complementary skills work on the same problem, see the same results, and share ideas to advance the project's timeline. Many of NASA's projects involve international collaborations with other space agencies, universities, and companies. While funding cannot be sent overseas, in cases where sharing access to HEC facilities is possible, it has contributed to the teams' success. Many researchers indicated that this should be expanded in some way that removes some of the security restrictions for work where the availability of shared computing is more important than the confidentiality of the code and data.

3. Current NASA HEC User Needs

NASA HEC use cases presented and discussed during the assessment workshop identified current unmet needs. These user-identified needs for HEC generally fall into five groups:

- Increased Operational Compute Capacity.
- Understanding Emerging Architectures and Technologies.
- Improved Performance of Storage.
- Expanded User Support Services.
- Improvements to HEC Management.

3.1 Increased Operational Compute Capacity

Generally, HEC users can effectively use whatever additional capacity they are provided and then some. This HEC Needs Assessment, however, found that users were able to cogently and specifically state how they would use additional capacity, capabilities, and support. Figure 6 shows the current programmatic estimate of growth in demand for the HECC. A quantitative engineering estimate of future user demand will be generated as part of the engineering analyses feeding into the strategic plan.

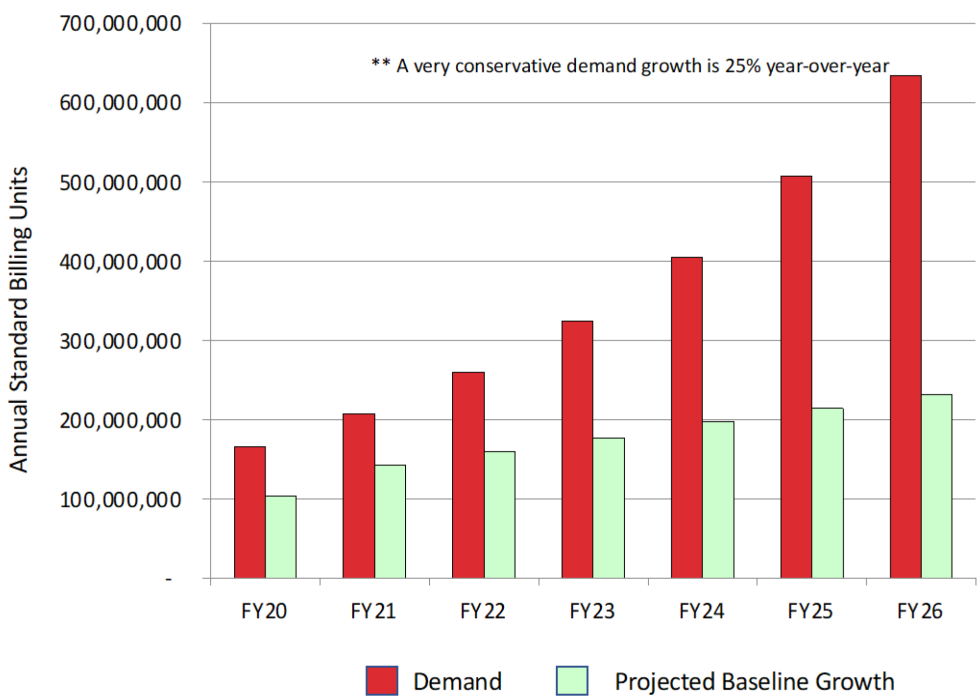

FIGURE 6: ESTIMATED GROWTH OF HECC DEMAND

The NASA research and engineering communities expressed a need for expanded operational capacity in conventional cluster supercomputing for running jobs of all varieties, sizes, and urgencies. Many users indicated that they restructured their codes and created workflows to fit within the HEC keyhole. Users indicated three compelling and related needs:

- Turnaround Time: Short wait times in large-job queues are needed to meet schedule deadlines, both for model development and model runs supporting studies.
- Throughput: Schedule pressure is particularly high when a project nears a design review, and many runs are needed to show the characteristics of alternative designs. The ability to run sufficient jobs in the clock time available to support these reviews increases confidence in the recommended design by management.
- Capability: Capacity that allows for larger-scale models and integrated model runs, rather than executing a set of components one after the other. The sub-optimized workflow adds significant clock time and researcher labor to accomplish the studies needed for many projects, particularly in optimization problems previously described.

NASA's HEC Program is the primary resource for addressing complex modeling problems to meet mission needs. NASA relies on HEC to make significant leaps in progress, yet growing mission needs for fidelity and resolution cause these models to grow more complex, and thus demand more compute capacity. It is now a struggle or, in some cases, nonviable to fulfill engineering requirements (defined as computational resources needed for performing certain engineering analysis or optimization to meet a design, development, testing, or deployment deadline). Alternatively, for research and analysis projects, requirements are often driven by the fierce international science and technology (S&T) competition (e.g., who gets to discover or develop a solution and publish a paper first.) NASA's scientists and engineers are out-competed when their peers at other research laboratories nationally and internationally may publish papers on the most advanced high fidelity and high-resolution computational modeling results. Some current investigations, especially those at the cutting edge of either engineering or research, are at or beyond NASA's HEC capacity while the demand continues to grow. To run in this environment, compromises in workflow or code execution strategies result in fragmentation into stand-alone components that can fit within the resources available, thereby increasing clock time and overhead needed to complete projects.

NASA researchers need access to advanced computing architectures leading into the exascale environment to understand better how to scale up their code to work as an integrated application in an efficient way. Since the agencies with platforms of the scale NASA users need are focused on the development of exascale applications, they insist on running the largest and most complex jobs they can find. The Department of Energy (DOE) does not provide early access to new architectures unless the NASA application is funded under their Exascale Computing Project. Many of NASA's study and analysis job requirements do not meet their minimum threshold as currently configured. The DOE's leading-edge machines run individual jobs scaling up from 1–10 million cores. Many NASA jobs, such as those using 4D eddy-resolving CFD codes, do not currently fit the "shape" of the desired job topology at DOE. Because of the lack of NASA resources at this scale, NASA jobs and workflows are designed to fit within the restricted space available and, therefore, have no path to expand to meet the threshold constraints. Current directorate-level procurement and management policies for measuring system utilization can force workflow and code architecture decisions into smaller jobs to run in a reasonable timeframe. The problem, as identified by participants in this needs assessment, is that users lack a mechanism to easily and regularly use the entire platform without waiting for weeks for the job scheduler to start execution. There is no intermediate scale computing resource to allow NASA users to ramp up to exascale-sized systems.

Use of computational resources at other government agencies is limited by funding and policies regarding NASA's relationship with the agencies, labor-intensive proposal processes, and security limitations. NASA's ITAR restrictions preclude running on open-access machines, such as DOE's Summit facility whose international users represent a challenge to these restrictions on the codes and the data. Most other agencies require formal proposals to be submitted, with the unpredictable outcome of selection processes applying agency agendas that are different than NASA's. These proposals to other agencies are time-consuming to formulate and require months of background preparation, all of which take time away from the primary mission research. To make progress in strategic investigations and studies, it is not possible to carry out a 5–10-year plan based on the use of another agency's resources when there is no certainty that these resources will be accessible. Users indicated a need for experimentation on architectures that are available at other government agencies, supported by formal inter-agency agreements that create dependability in the relationship, consistent with *The National Strategic Computing Initiative* (COUNCIL, 2019) and related plans.

Access to the nation's most capable computing system, such as the previous example of NASA engineers using the DOE's Summit machine to perform model simulation, is critical. However, it is not sustainable. NASA scientists and engineers need to be able to access sustainable computational resources steadily to make predictable progress. Because of the sporadic access to the nation's most capable machine for NASA researchers, they can only be used for limited purposes and special cases.

3.2 Understanding Emerging Architectures and Technologies

Over the past five years, all indications point to a rapid shift in the type of processors that will yield the highest capacity and most capability in supercomputing. GPU and ARM-based supercomputers hold the highest places in the TOP500 lists of the world's most powerful systems, and more recently AMD-based processor systems have also demonstrated potential. One plausible future is that new architectures will supplant the conventional cluster supercomputers in use at NASA today—similar to the way clusters replaced vector supercomputers—perhaps with an entirely new architecture or perhaps with changes to the conventional architecture, such as the number of cores or deeper memory hierarchies.

Several notable challenges discussed earlier in this assessment relate to emerging architectures and technologies. For example, uncertainty quantification in all physics-based models in Earth and space science, fundamental aeronautics, space vehicle design, and space technology requires significantly more model execution and model output. Execution time is long and can't be interrupted or pre-empted without improved storage to handle the increase in model output. Similarly, adaptive discretization is important in many incompressible flow problems since it is often necessary to resolve details on multiple levels. In fluid structure interaction problems, it is critical to resolve the turbulent flows in the wakes behind objects in order to accurately predict even large-scale behaviors (Ilie, 2019). Lastly, challenges associated with collaborative work within and beyond the HEC environment should be considered. Ideally, for example, many of the capabilities identified during this needs assessment also would be available outside the HEC Program's environments, to all other compute environments at the agency (and in cases, outside of the agency for our partners to use on NASA projects).

Large-scale GPU platforms are mentioned in several case studies as likely emergent architecture based on current experience with the technology. For example, based on limited experiments run on DOE GPU-based supercomputers and small NASA GPU clusters, some HEC users see a future need to access large-scale GPU platforms as operational capabilities, although mostly for machine learning applications rather than engineering applications, like those of physics-based modeling. The limited access to date has given researchers indications for how to exploit this type of platform. In the case of both full-physics models and machine learning workflows, researchers have begun to understand, for example, which problems are well-suited to this environment and the kind of savings in conventional processor load that can be achieved. More capacity is needed to experiment with and eventually exploit the scaling that can be achieved with these systems. Initial experiments by the LaRC High-Performance Computing Incubator Program running codes on the Summit machine at the Oak Ridge Leadership Computing Facility gave indications of the advantages of scaling up some codes to reduce clock time and to improve fidelity, resulting in improved answers to mission questions.

Additionally, some new workloads involving machine learning and other forms of artificial intelligence have demonstrated better cost performance on GPU or neuromorphic processors. The rapid adoption of these analytic tools will contribute substantially to the rapid growth in load on NASA's supercomputing centers over the next decade and materially drive the need to understand the value of these alternative architectures.

Users also indicated the emergence of workflows involving analyses running on multiple platform types. Currently, the interfaces and management of such workflows are handled mostly manually, but automation will soon be necessary to maintain progress against mission needs. These heterogeneous workflows will require coordination among different platforms for job submission, similar to the way job submissions coordinate processors within a single image today. The emerging need for these workflows is recognized, as is the likely complexity associated with managing the workflows given the increased complexity.

Users identified several types of computing environments warranting experimentation:

- Conventional cluster computing.
- ARM processor-based computing.
- Commercial cloud computing.
- Private cloud computing.
- Neuromorphic computing.
- Power9 processor-based systems.
- AMD-processor-based systems.
- GPU-based systems.

Case studies presented as part of this needs assessment identified at least four main areas where NASA researchers need to understand and use these emerging architectures and computational technologies:

- Migrating codes from existing environments to future ones.
- Leveraging portable code design that could be hosted on a variety of platforms.
- Redesigning applications beyond initial generation.
- Developing new applications and codes.

To these ends, researchers encourage the HEC Program to expand its own research into these environments to address the platform characteristics and how to tune code to yield optimum performance, as they have done with previous architectures. The success of such a partnership depends on the HEC staff understanding these new technologies, the operating systems, and auxiliary software, as well as how to integrate them into an operational capability.

Workshop participants also indicated that early knowledge of evolving HEC strategy is essential for identifying eventual legacy platforms promptly and, thus, reducing the amount of time over which multiple platforms are maintained. This strategy needs to be formulated based on experiments and research by HEC Program staff into the characteristics and behavior of these computing environments as applied to NASA research problems, with special attention to the challenges of maintaining applications and codes as systems evolve and mature. When a decision is made regarding HEC strategy, a related transition plan should explain timeframes for re-hosting with large capacity as soon as possible after the new architecture is available.

Computer security is a critical consideration when addressing these needs because every need involves different levels and considerations of computer security risk, whether at end-user locations, or on or off NASA centers. Because separate facilities are unaffordable, a mechanism for sharing a common facility with different security settings is needed to accommodate all the variations in environments, considering two primary factors:

- Breadth of collaborators, both people and organizations.
- Sensitivity of the applications and data.

NASA researchers also need professional assistance in high-performance computing to understand the capabilities and limitations of emerging architectures and technologies, including:

- Benchmarking through well-designed experiments and creation of representative kernel applications.
- Code development leading to an operational capability.
- Best practices for each HEC element, from code development to scaling, optimization, and execution.
- Cost considerations for the entire lifecycle including operation.
- Code migration implications.
- Technical characteristics of the systems and architecture that influence code implementation.
- Scaling.
- Alternative development models.
- Code compatibility across platforms.

Engineers need help in exploiting the emerging hardware. They need a team effort where the HEC experts may be working on several projects or may act as a tiger team to focus on a specific problem. If they are spread too thin and, thus, delay the schedule, this would be a failed approach.

Some specific needs for professional computer expertise that were widely identified include:

- Assistance in determining when to run projects on commercial cloud computing environments, including cost comparisons that include cost of migration.
- Performing research into improved techniques and algorithms.
- Establishing a technique for benchmarking NASA applications to compare across platforms without major code re-configuration.

Lastly, there are broader, overarching computing policies to address. For over five years, the Office of Science Technology Policy (OSTP) has pursued the National Strategic Computing Initiative (NSCI) to maximize the benefits of high-performance computing for the United States. NASA's role, as a deployment agency, is to develop HPC requirements to influence early stages of design of new HPC systems including viewpoints from the private sector and academia on target HPC requirements. Capable exascale computing will require scaling algorithms to billion-way parallelism, improving energy efficiency, and speeding memory access. NASA HEC users are studying the potential for advancing agency mission objectives using these tools and have identified several areas which could leverage this capability, identified in Appendix E, that are being used to influence the evolution of the next generation HPC. To continue this process, researchers need access to advanced computing architectures to understand how codes need to be redesigned, in order to be able to influence requirements.

3.3 Larger Storage Capacity for Data and Digital Output with Faster Access

The amounts of data provided by external instruments and generated by simulations for further exploitation has exploded and is forecast to continue its rapid expansion. As one of the needs assessment participants said, "We need a coherent strategy for model output retention and documentation." Practices are needed to ensure that observational data, reanalysis data, and model output are easily findable, accessible, interoperable, and reusable (FAIR) by the HEC assets. Relocating data to increase proximity to the processors is time consuming and labor intensive. From a practical point of view, in the case of HEC users, *accessible* means the storage devices should appear to the processors as directly accessible and available for processing without intermediate transfers, at least in read-only mode. An overall Data Management Plan for the HEC Program is needed to establish consistency and to achieve these objectives, recognizing that the nature of model output is different than observational data.

For some of the data that serve a wide user community, NASA NPD 2230.1 notes that data and model output must be described by consistent metadata adequate to enable re-use of the data. Preservation of model output for re-use involves creating metadata conforming to common standards to be re-usable, as well. Many datasets are not adequately described by metadata to ensure the characteristics of the model run with the associated constraints are re-usable. For Earth science, metadata standards have been developed by NASA's Earth Observing System Data and Information System (EOSDIS) Program for observational data and, for environmental modeling, by the American Geophysical Union (AGU) and the National Center for Atmospheric Research (NCAR). Data policies are being drafted at the agency and SMD levels that will complement the standards being developed in other parts of the agency.

3.4 Expanded User Support Services

HEC users vary in their experience with high-end computing systems, with NASA facilities, and with the expectations and demands of working with or for a federal agency. Through its support services, the HEC Program provides users with a wide variety of value-added services to help them quickly and efficiently accomplish their mission computing needs (Table 0-2, p. 3).

All seven of the current emphasis areas for support services respond to and meet user needs well. Nevertheless, in their case studies, users indicated, with widespread agreement, that extensive user support services are essential to their ability to advance the state of modeling on HEC assets. Case studies identified three areas of need that require attention as users anticipate eventually migrating their applications onto emerging computational capabilities. Those areas are (1) code development; (2) attracting and retaining human resources; and (3) collaboration. In addition, users noted that existing demand means many services are currently heavily used, and more capacity is needed. Lastly, discussion of these areas of need and of current trends recognized by HEC user support services suggest the traditional model of application support may need to change because of the growing expectation among HEC staff and users to co-design, often working side-by-side to migrate applications, develop code, and identify and meet new or emerging needs.

3.4.1 Code Development

The HEC Program has long helped the domain experts improve their code performance. HEC sites contribute new code development techniques in these partnerships. When the domain experts understand modern coding techniques, HEC staff provide sounding boards for discussion. In other situations, HEC staff help the domain experts understand techniques and aids available in HECC or NCCS facilities. The use of modern software development practices (e.g., Agile, CI/CD, containers, code reuse) is on the rise but is not well facilitated by current HEC capabilities. Promotion of commonly shared and available tools for all HEC users across the agency is vital. However, many instances of tools have grown from local need and current point solutions (separately at ARC, GSFC, and elsewhere) are not well scaled. Users identified more specific examples, including:

- GitLab at GSFC is not easily accessible by users at other centers.
- The NASA GitHub server is not widely used other than for local point solutions.
- Cloud computing resources for scientific and engineering computing are not easily available.

Some specific services that users indicated would be useful include:

- GitHub at various levels of confidentiality.
- A cost estimator for deciding where to run code, particularly for estimating commercial cloud computing costs.
- Automatic metadata generation for provenance.
- A microservices registry.
- A common, agency-wide container registry for code sharing.
- A common, shared artifact repository for both NASA and third-party commercial and open-source binary artifacts and libraries.
- Publication of a common hosting infrastructure baseline.

Users indicated that there is a widespread need to make commercial cloud computing available across the agency for scientific and engineering use at different security levels. To use the commercial cloud computing capabilities, users need to understand (1) their strength and weaknesses better, (2) the unique features they offer, and (3) the realistic cost of performing work on those platforms. The mechanisms for users to pay for these services needs to be smoother and easier to access.

HEC users need additional support for code performance analysis (tuning, enhancements), including:

- New approaches to software development that tune code continuously, particularly where the approaches are important to successful code migration.
- Testing and version control.
- Expanding the current partnerships between domain expertise and computer science expertise into wider communities, with appropriate funding for both.
- Performance monitoring at both HEC sites.
- Review of major codes for flaws or potential improvements.
- A policy that identifies inefficient codes to the responsible mission directorate.

These needs are motivated by the need for code modernization, migration, and portability. Some users identified the need for programming models that permit code transfer to any platform and help in understanding the tradeoff between portability and performance on a specific platform. Most projects do not have the budget to address code conversion or modernization, so when the code is originally designed, help is needed in considering how to do so is a way that permits evolution. More urgently, migration to other processors is inevitable. Support services are needed to coach and guide proper software development. In addition, domain experts will need help migrating their codes efficiently because people who have a combination of domain-specific science and computer science backgrounds are rare.

Code owners and sponsors can frequently find funding to support a new capability or function, but are infrequently funded to re-write code for performance, modernization, or a new platform. Across all mission directorates, HEC users identified the need for NASA to establish recognition for selected facility codes that widely impact agency research or flight missions and a way to support their maintenance and evolution. This involves a process for including critical tool development in the budgetary process with sufficient funds to develop, enhance, and maintain selected codes. Where such processes exist, they vary between organizations. These facility codes all need to include the recognition of their value and a commitment to their continuation, subject to reviews like the Non-Advocate Review.

Commonly used facility codes at NASA include:

- Cart3D,
- LIS/LDAS
- GEOS
- LAVA
- GISS Model E
- OVERFLOW
- FUN3D
- MAPSS

The extent of the demand will be better understood after a solicitation on Open Source software maintenance closes and the results are analyzed.

3.5 Provide Leadership in Attracting and Retaining Human Resources

HEC support services rely on a strong culture of customer service combined with state-of-the-knowledge skills. In today’s competitive marketplace, hiring and maintaining qualified staff for the HEC centers and development teams around the agency presents a challenging problem. Users indicated that a collaboration with HEC is needed to establish a strategy for upgrading the human resources necessary both for the projects and the HEC centers. This includes both the acquisition of new talent and strengthening the skills of the existing HEC user and staff community.

Users observed that this is an opportunity to provide equitable opportunities and promote diversity, equity, inclusion, and accessibility training to foster a more inclusive environment and leverage more opportunities for recruitment. This will be addressed in the HEC strategic planning.

Adoption of machine learning at NASA will need more than hardware to accelerate its application to achieve mission directorate objectives. The need for knowledge sharing about machine learning will be a key aspect that can be facilitated by the user group, among others. Other organizations within NASA are also addressing this area, and a coordinated strategy should include the Digital Transformation Initiative and broader sharing of local symposia and demonstrations. Aspects of knowledge sharing include access and understanding of existing NASA machine learning source code, related software products (such as containers), providing an ability to share solutions without running them locally (microservices), and infrastructure that can capture (and promote) metadata indicating quality of these solutions to downstream consumers.

3.6 Expand Ability to Collaborate Safely and Widely

Collaboration is a key element of meeting most NASA mission objectives. Such collaborations include a wide range of participants and vary by project:

- Across NASA centers.
- Among other government agencies.
- Aerospace industry and commercial space firms.
- Academia.
- Foreign researchers and government agencies.
- Through NASA's SBIR/STTR program.

Users need the HEC Program to continue their past support, but also to broaden the scope of collaboration among multiple parties, as dictated by the project and mission directorate prioritization decisions, to enable access the compute resources, data, or model output to perform analysis without transferring data to local facilities. The need for remote collaboration in science and engineering has been highlighted by the recent COVID-19 quarantine, which makes even intra-center collaboration a remote activity. Individual Principal Investigators and co-researchers on a given project need to be able to share the same processing environment and equal access to input data and digital output for analysis and visualization through the HEC Program, regardless of whether they have NASA credentials.

Tools beyond typical videoconferencing have demonstrated the joint exploration of ideas. Many of the users reported that the HEC Program enables encounters among other research teams through visualizations, file system access, and other real-time methods. Such tools permit teams of researchers to examine and discuss results together. Currently, tools are provided at an agency level with full capabilities to NASA-badged individuals only. Tools provided to non-NASA-badged individuals are partially constrained. The COVID-19 quarantine encouraged the emergence of new tools that are even more supportive of these partnerships. HEC Program support should also evolve to leverage the best ones available.

Some programs (e.g., Atmospheric Composition, Space Weather) have a significant number of international collaborators. By supplying a common work environment to the entire team, NASA benefits by advancing the state of the art without funding the labor. For example, improvements to the GEOS-Chem model are funded by several different programs but collaborators need a way to include their contributions into a common body of code instead of creating several different branches. A common and accessible GitHub repository and a common code development environment would save funding and reduce the cost of re-integrating the various enhancements. A collaborative environment is necessary where the confidentiality aspects of computer security are appropriate to not halt collaboration, but also protect integrity and availability.

As an example, in space weather modeling a collaboration was established among NASA, NOAA, AFOSR, AF/XOW, SMC, the Office of Naval Research, and NSF under the auspices of the Committee for Space Weather and the National Space Weather Program Council. Tasked to aid in the development of models for specifying and forecasting conditions in the space environment, the Community Coordinated Modeling Center (CCMC) relies on contributions of all the partners and grant recipients. Validation of models and tools to analyze them is a group effort. Once accepted, models are used by a wide range of space scientists to assess conditions between the Earth and the Sun. Most investigations involve workflows of multiple models, analytic tools, and visualization capabilities. The success of the CCMC depends on active collaboration and coordination among all space weather stakeholders. Interagency collaboration is essential to achieve the goal of bridging the gap between the research and operations communities. International collaboration will allow for better leveraging of resources with research, observations, and models provided by the worldwide space weather community. HEC enabled the development and sharing of tools, but with limited resources and the emergence of modern high-performance computing and commercial cloud computing, the opportunity is ripe to accelerate model and tool development to achieve the needed capabilities and to provide the research and space operations communities access without impediment.

3.7 Improvements to HEC Management

The most important improvement to management of HEC resources in in computer security. Implementing computer security on a shared system, particularly one shared across all mission directorates and centers, is a difficult compromise. The relationship between confidentiality, integrity, accessibility, and availability is different for each community involved and often for each program within that community. A single set of security control implementations is, at best, disappointing to some customers and a showstopper to others. NASA has chosen to apply the most restrictive requirements for confidentiality to its in-house computing assets, which forces the less restricted projects to compromise their availability to team members. Many HEC users have highly restrictive proprietary, ITAR, or EAR Center Contribution Agreement (CCA) restrictions and cannot run without these restrictions. Several NASA projects that should use HEC Program resources find it expedient to use non-NASA assets to accomplish their work so that the full range of collaborators can share the same computing resources. Several solutions should be examined during the next phase of the HEC modernization program, including acquisition of additional hardware or a flexible system security model with time-of-day dependent Authorization to Operate (ATO) at different levels that protect restricted work but are less constraining for non-SBU projects.

Users identified the need to review scheduling and prioritization of jobs, as well as the allocation of SBUs among the Mission Directorates. These issues also involve policy, standing NASA management decisions and financial management and are being addressed separately in the policy analysis to feed into the strategic planning effort.

4. HEC Support to Aeronautics Research Mission Directorate

NASA's ARMD and its predecessors first recognized the ability to characterize fluid dynamics using numeric simulations on a supercomputer and used this as a basis for establishing the Numeric Aerodynamic Simulation (NAS) Facility in 1984 at Ames Research Center (now called the NASA Advanced Supercomputing Facility). Taking the long view, not only did the early researchers capture their understanding of fluid dynamics in full-physics computer models, but successively refined those models until they have achieved adequate fidelity that the Boeing 777 could be designed without wind tunnel tests. This improved design process saved considerable investment and accelerated time to delivery of the first aircraft. These models have been transformed into a series of well-respected and widely used computer codes, including FUN3D, OVERFLOW, Launch Ascent and Vehicle Aerodynamics (LAVA), etc. In 2014, ARMD chartered development of a new strategy which was published as CFD Vision 2030 Study: A Path to Revolutionary Computational Aerosciences (Slotnick et al., 2014). In parallel, modeling of thermodynamics, computational structures and materials, interaction with the fluid flow field, and control systems have all evolved over the past forty years. ARMD researchers have demonstrated new capabilities to accelerate completion of model runs with new types of supercomputers, including DOE's Summit at Oak Ridge National Laboratory.

Similarly, many efforts have undertaken the challenge of improving the interaction of the models with the data resulting from experiments on large-scale experimental facilities, such as wind tunnels. It has been long recognized that faster turnaround is key to effective use of these facilities and the experiment staff—and the key to rapid turnaround is analysis and visualization of the data. Fast turnaround of an initial run's data also helps to assess the impact of tunnel artefacts, such as walls, stands, etc., to make subsequent runs more effective. Integration of experimental facilities with quick turnaround in processing has been successful to the point that data analysis and visualization of the resulting output have can help guide the reconfiguration of the experiment almost without delaying the next run. Demonstration creates conversation, which improves the experiment strategy.

The emergence of machine learning and other artificial intelligence techniques offers attractive opportunities to improve the simulation of aircraft behavior under various flight regimes with higher fidelity and reduced computational load and clock time. These machine learning techniques can benefit the analysis of experimental data and model output. Alternative hardware for processing these techniques have placed new demands on High-End Computing Capability (HECC) resources at the NAS facility.

NASA's aeronautics programs focus on six areas of research that develop solutions to the major challenges and opportunities for aviation: a growing demand for mobility, the sustainability of energy and the environment, and technology advances in information, communications, and automation. The six research areas are:

- Safe, efficient growth in global operations.
- Innovation in commercial supersonic aircraft.
- Ultra-efficient commercial vehicles.
- Transition to low-carbon propulsion.
- In-time systemwide safety assurance.
- Assured autonomy for aviation transformation.

These are addressed within four Programs, described in Table 4-1.

TABLE 4-1. MAJOR ARMD PROGRAMS. (SOURCE: HTTPS://WWW.NASA.GOV/AERORESEARCH/PROGRAMS)

ARMD Program	Description
Advanced Air Vehicles (AAVP)	AAVP studies, evaluates, and develops technologies and capabilities for new aircraft systems, and explores far-future concepts that hold promise for revolutionary air-travel improvements. Innovative AAVP design concepts for advanced vehicles integrate technologies that focus on fuel burn, noise, emissions, and intrinsic safety. The goal: to enable new aircraft to fly safer, faster, cleaner, quieter, and use fuel or alternatives to fuel far more efficiently. Partnering with industry, academia, and other government agencies, AAVP pursues mutually beneficial collaborations to leverage opportunities for effective technology transition.
Airspace Operations and Safety (AOSP)	AOSP works with the Federal Aviation Administration, industry, and academic partners to conceive and develop Next Generation Air Transportation System (NextGen) technologies to further improve the safety of current and future aircraft. As radar-based air traffic control transitions to a NextGen satellite-based system to enhance safety, capacity, and efficiency on runways and in flight, AOSP-developed NextGen methods and means will provide advanced automated support to air navigation service providers and aircraft operators to reduce air-travel times and delays, and to ensure greater safety in all weather conditions.
Integrated Aviation Systems (IASP)	IASP conducts flight-oriented, system-level research and technology development to effectively mature and transition advanced aeronautic technologies into future air vehicles and operational systems. IASP focuses on the rigorous execution of highly complex flight tests and related experiments to support all phases of NASA's aeronautics research. For technologies at low technology readiness levels, IASP flight research aims to accelerate development and determine feasibility. For more mature technologies, IASP intends to reduce potential risk and accelerate transition to industry.
Transformative Aeronautics Concepts (TACP)	TACP solicits and encourages revolutionary concepts, creates the environment for researchers to experiment with new ideas, performs ground and small-scale flight tests, allows failures, and learns from them, and drives rapid turnover into potential future concepts to enable aviation transformation. Research is organized to aggressively engage both the traditional aeronautics community and non-traditional partners. Although TACP focuses on sharply focused studies, the program provides flexibility for innovators to assess new-technology feasibility and provide the knowledge base for radical aeronautics advances.

Other historical programs are identifiable in the HECC usage data, based on tags collected when allocations were made; not all of these, however, align to the current programs. No effort was made to adjust these tags, even for FY20, so some program consumption data may be understated.

ARMD researchers have constructed a set of production codes that are well-developed, tested, and validated, and widely used by the aeronautics community to aid vehicle design. Examples of well-supported codes with extensive validation history, which are well understood and used by many researchers, include OVERFLOW, FUN3D, Cart3D and LAVA. Community involvement allows multiple contributions outside the basic development teams to further refine both the understanding of the phenomena and the codes themselves. For example, LAVA generates Reynolds Averaged Navier Stokes (RANS) databases for ARMD projects, including the X57 Maxwell experimental electric aircraft, the X59 Low-Boom Flight Demonstrator (LBFD), the Novel Propulsion Airframe Integration (NPAI), and the NASA Common Research Model (CRM). As modeling and simulation are used for one project after another, fidelity improves, the history of validation is extended, and codes are expanded to handle new elements of a computational wind tunnel.

The actual ARMD usage of HECC baseline computational capacity for the past five years is shown in Table 4-2, encompassing over 400 separate projects, many of which have remained active through the entire five-year period. The HEC Program uses Standard Billing Units (SBU), described in Appendix A, as the metric by which allocations are made and consumption is measured.

TABLE 4-2: ARMD HECC USAGE BY SBU (2015-2020).

	FY20	**FY19**	**FY18**	**FY17**	**FY16**	**FY15**
SBU	30,713,041	26,493,134	25,490,311	18,287,640	14,367,649	12,563,867

The growth of allocations and actual usage show only part of ARMD's need for HEC. ARMD programs are hampered by the lack of capacity on HECC resources to the point that ARMD has purchased dedicated HECC hardware to expand that capacity, which is reflected in these numbers.

ARMD scientists and engineers are developing a strategy to exploit exascale computing when it is available in order to support the following disciplines:

- Computational fluid dynamics (CFD).
- Multidisciplinary simulations across speed regimes and ranges.
- Adjoint-based algorithms for error estimation, design, and uncertainty quantification, including chaotic systems.

Some observations gained from this planning include:

- More and more missions are demanding higher fidelity in space, time, and physics modeling.
- Waiting months for an answer on a few thousand CPU cores is not acceptable to users.
- The project must scale up and leverage emerging architectures (e.g., GPU-based simulation at capability-wide scale would be a game-changing technology for multiple missions).

- HECC resources are useful for CPU-based code development and production up to large scales, but very large scale runs are being accomplished using inefficient or labor-intensive workarounds, including splitting up jobs and re-assembling output. Frequent, fast-paced access to large resources is required to prepare applications to run at scale.

ARMD researchers emphasized several key obstacles to advancing the state of the art over the next 5–10 years to exascale computing environments. Due to the lead time to convert and validate code onto new, more effective architectures, they need early access and assistance in understanding the re-engineering needed. To a limited degree, this is accomplished today using informal arrangements with external facilities, labs, vendors, and academia. Users also indicated several specific issues with current applications, including performance portability, validation and verification in the face of asynchronous execution, concurrency in temporal direction, latency hiding, multiphysics, load balancing, increasingly disparate spatial and temporal scales, task-based vs, bulk-synchronous execution, asynchronous and mixed-precision algorithms, integration of edge computing with high performance computing, etc. (Appendix D8.14, Malik)

Table 4-3 shows ARMD's top 10 users of HECC resources based on FY20 SBUs.

TABLE 4-3: ARMD TOP TEN HECC PROJECTS IN FY20.

PI	Organization	Project Title	FY20 SBUs Consumed
Debonis, James Raymond	NASA/Glenn	Large-Eddy Simulations for Propulsion Flows	2,346,894
Moder, Jeffrey P	NASA/Glenn	National Jet Fuel Combustion Program CFD	1,936,197
Malik, Dr. Mujeeb R.	NASA/Langley	LES of High RE Flows	1,531,545
Murman, Scott M.	NASA/ARC	ARC for RCA	1,513,314
Georgiadis, Nicholas J.	NASA/Glenn	Advanced Propulsion Turbulent Predictions Methods	1,442,904
Khorrami, Mehdi R.	NASA/Langley	Evaluation of noise reduction concepts for flaps	1,236,947
Kiris, Cetin	NASA/ARC	High Fidelity CFD support for X-57 Aerodatabase	959,886
Jansen, Kenneth	University of Colorado, Boulder	UCB-ScaleResolvingSimulations	951,787
Deere, Karen A.	NASA/Langley	Electric Propulsion Concept	898,668
Khorrami, Mehdi R.	NASA/Langley	Airframe noise simulations of commercial supersonic transport concepts	802,796

4.1 Advanced Air Vehicles (AAVP)

The HECC Project supports the AAVP models and analyses with high node-count runs and visualization of simulation output using the HECC capability at Ames. AAVP is one of the largest users of the ARMD allocation; the history of AAVP project usage is displayed in Table 4-4. As a matter of explanation of this table, the HEC Program uses Standard Billing Units (SBU), described in Appendix A, as the metric by which allocations are made and consumption is measured. Major AAVP projects using HECC resources are listed in Table 4-5.

TABLE 4-4: AAVP HECC USAGE FROM 2015 TO 2020.

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	8,573,259	14,089,739	14,366,645	9,496,938	7,085,593	5,327,209

TABLE 4-5: TOP TEN AAVP PROJECTS USING HECC RESOURCES.

PI	Organization	Project Title	FY20 SBU's Consumed
Khorrami, Mehdi R.	NASA/Langley	Airframe noise simulations of commercial supersonic transport concepts	802,796
Vatsa, Veer N.	NASA/Langley	Active Flow Control Optimization Studies for Lifting Surfaces	583,870
Chaderjian, Neal M.	NASA/ARC	RANS/DES Code Development and Applications for Rotorcraft Aeromechanics and Design	564,881
Ashpis, David E.	NASA/Glenn	Low Pressure Turbine (LPT) Flows	533,205
Li, Fei	NASA/Langley	Transition simulations for 3D hypersonic boundary layers	428,305
Hah, Dr. Chunill	NASA/Glenn	Large eddy simulation for multi-stage turbomachinery	414,981
Bartels, Robert E.	NASA/Langley	High Fidelity Computation of Adaptive Aeroelastic Performance of Advanced Configurations	376,321
Raman, Venkatramanan	University of Michigan Ann Arbor	Direct Numerical Simulation of Scramjet Inlets	322,163
Celestina, Mark L.	NASA/Glenn	AATT IBLI Develop Numeric Testbed for Fan Response to Boundary Layer Ingestion	298,110
Pandya, Shishir A.	NASA/ARC	High Fidelity Simulation Support for Double Bubble BLI Concept	283,875

Two examples of the large-scale use of HECC resources are both efforts to predict the acoustic characteristics of aircraft. The first example is the development of high-fidelity, system-level airframe noise simulations for civil transports. This project's objective is to produce accurate prediction of airframe noise to help mitigate its impact on communities near airports. HECC resources are used to perform the high-fidelity, system-level airframe noise simulations and to develop/evaluate noise reduction technologies prior to expensive flight tests. These analyses use proprietary data from commercial aircraft manufacturers and are sensitive but unclassified (SBU). The workflow feeds the simulation output into a visualization for analysis, and into an acoustic propagation code, which is then used to determine airframe noise signatures and certification metrics. The project today uses 4,000–10,000 cores per simulation, generates over 150 terabytes of data each run, and needs high-throughput input/output to avoid holding up processing. The limitations on HECC capacity and configuration for large node-count models result in 30-day waits for jobs to run, and extended execution times due to shortage of capabilities like high-memory nodes and reliable high-bandwidth, fast I/O.

The second example is a project that uses the LAVA code to predict the near and far-field acoustics for aircraft, including rotorcraft, with CFD, within a short enough turnaround time to help with noise reduction during vehicle design. The project team used HECC computing to perform scale-resolving simulations and stored the data on HECC storage resources, performing five LAVA Lattice-Boltzmann simulations of an SUI quadcopter (requiring grids of 100–600 million cells and using approximately 200,000 SBUs over three months). This work demonstrated the capability to predict high-frequency broadband noise for small rotorcraft. The Advanced Air Mobility (AAM) market is extremely interested in new methodologies like Lattice-Boltzmann to better address noise concerns with multi-rotor aircraft. However, the project was deemed unsustainable because of the need to obtain a reservation to start each job requiring more than 4,000 cores. Without this barrier, more scale-resolving simulations could be performed, leading to a better understanding of the means for reducing the noise of AAM designs. (Appendix D6-4.4, Khorrami; Appendices D6-8.3 and 9.6, Cadieux; Appendix D6-2.6, Street)

Another area of HECC support for AAVP is the development of a multi-disciplinary analysis and optimization (MDAO) tool as part of the Revolutionary Vertical Lift Technology (RVLT) Project. Development of a high-fidelity MDAO tool couples high-fidelity computational aerodynamics, rotorcraft comprehensive analysis with structural dynamics, and noise prediction models for rotorcraft aeromechanics. The project’s objective to capitalize and improve unique vertical capabilities requires high-fidelity analysis and sensitivity analysis of fluids, structures, dynamics, propulsion, acoustics, and complex interactions among various disciplines. Accurate resolving of all the physics involved is a technical challenge requiring orders of magnitude more HECC resources than what can be currently accessed. HECC currently enables MDAO tool development based on Reynolds Averaged Navier Stokes (RANS) simulations and low-frequency rotor noise predictions. Optimization is limited to grids containing less than 10 million grid points, considering resource availability and runtime limits. A recent aero/structure optimization used 122 wall-clock hours on 2,200 CPUs for a two-point (hover and forward flight) Blackhawk rotor-blade optimization. The demonstration was performed with isolated rotors on coarse grids (no aeroacoustics). It is expected that tools like the one demonstrated here, that enable coupled analysis/design runs, will be needed for many designs in the next five years. Much larger computing resources and advanced architectures are needed for full geometry and acoustic optimization. These applications need orders of magnitude more resolution to resolve turbulence with higher fidelities, blade/engine/airframe acoustics, and interactions. Coupled solutions are compared with experiments for model validation. (Appendix D8.10, Malik, Nielsen, Schuster, Bushnell; Appendix D6-8.4, Wang)

Another project with a high usage history is the direct numerical simulation of scramjet inlets as part of the Hypersonic Technology Project. Researchers at the University of Michigan run an end-to-end simulation platform to predict high-altitude relight of aircraft engines. Low ambient air pressures and temperatures at high altitude can lead to engine flame-out and hamper relight attempts. In high-altitude relight, an aircraft engine must ignite within a certain period after its initial flameout. Although aircraft fuels are tested to evaluate their ignition characteristics at different operating conditions, it is difficult to experimentally replicate all the physical parameters that affect ignition ,which gives rise to the need for detailed computational models to provide insight into the complex relight process. The researchers ran simulations to generate enough samples to represent statistical effects of parameters such as turbulent flow field and initial kernel energy on ignition outcome. Outputs were used to reconstruct ignition probability, using uncertainty quantification. Results showed that the modeling framework can efficiently generate abundant high-fidelity data of turbulent forced-ignition processes, which can be applied as input for techniques used to study patterns of altitude relight problems, benefitting future engine design. (Appendix D6-9.3, Choudhari; Appendix D9.4, Tang)

CFD simulations of ARMD's RVLT Urban Air Mobility (UAM) concept vehicles are run on HECC resources with a primary focus on developing and validating tools for conceptual design. High-fidelity modeling simulations of UAM vehicles can capture critical aerodynamic interactions between the vehicle fuselage, wing, rotors, and rotor wakes that are not validated in NASA's lower-fidelity design tools. Without experimental data, high-fidelity modeling plays a critical role in assessing and validating design tools. The workload is determined by timing of project goals and milestones, wind tunnel test planning, and conferences. Post model-run analysis and flow field visualization on HECC assets are essential elements of the project's contribution. While most of this work is open, some aspects are ITAR-restricted and can only be performed on appropriately cleared facilities, like HECC. Most simulations take multiple queue submissions with 2,000 cores needed for 250 million grid points; queue wait times of 1–3 days or more introduce delays and waste engineering labor. When computational capacity is not available to run all the variations needed, the result is a reduction in the scope of the investigation. Output is typically 200 gigabytes (GB) to 1 TB of data, with 100 GB being stored in HECC's long-term archives. The primary measure of effectiveness of HECC resources is the turnaround time (queue wait time plus execution). Without HECC resources, this work would revert to lower-fidelity simulations on smaller, slower, local assets, reducing the scope of applicability of the design tools. As the UAM market expands, increased use of HECC resources is expected, with more projects and additional researchers needing access. (Appendix D6-8.1, Allan)

Similarly, using HECC resources, the use of the Reynolds-Averaged Navier Stokes (RANS) database is being expanded to support other AAVP projects, including RVLT projects.. HECC storage resources are also used to retain the data generated. HECC provides ARMD projects with "numerical wind tunnel" capability and feeds results into a database of characteristics that can be used by the research community for both research and design without having to obtain time in a physical facility except to validate models. This capability is expected to be a key component of aircraft certification process in the future. (Appendix D6-8.2, Kenway)

AAVP's computational aeroelasticity research is supported by HECC resources to analyze deformation under flight loading and to predict dynamic instabilities that could lead to catastrophic failures. Almost all of this work is restricted by ITAR or EAR99 regulations and can only be run on computing platforms with appropriate Authorization to Operate, further limiting the use of alternative computing environments. Some work is used to help plan wind tunnel experiments by providing expected stability limits and characterization of the wind tunnel effects on the experiment. More fundamental research leads to better understanding of aeroelastic physics and improved methodologies and tools. Output can be as large as 1 TB, depending on vehicle complexity. The nonlinearity of the aeroelastic behavior requires many iterations to detect changes in vehicle stability. As described above, limitations in wind tunnel time or capabilities restrict experiments for every parameter space or every design iteration and so computational experiments are needed to produce a comprehensive characterization of the dynamic states of a vehicle. As a result, there are as many as 50 separate high node-count jobs waiting in queue at any given time to run for as long as 216 wall-clock hours. Currently, these runs need shorter times in queue and higher node counts. Compromises are being made that affect the quality of the research and increase the time spent by researchers on housekeeping. For example, aeroelastic flutter studies are taking months to perform, and intuition is being used to select the runs to be made, potentially missing critical points of nonlinear behavior. Jobs are being simplified to select queues with short lead times. Uncertainty quantification and limit-cycle oscillations are essential but too expensive to conduct on the limited resources available. These would significantly improve the confidence in computational results and are needed to achieve Certification by Analysis. To meet the program objectives, shorter queue times with many more cores are needed for each job. The evolution of aeroelastic simulation over the next decade is depicted in Figure 7. As it matures, computational aeroelasticity is expected to be used more widely in the following applications (Appendix D6-9.3, Jacobson; Appendix D6-4.2, Warner):

- Certification by Analysis.
- Multi-disciplinary design optimization.
- Model-based engineering.
- Transonic flutter and limit-cycle oscillations.
- Unsteady launch vehicle loads.

Alternative architectures show great promise for increasing the number of cores usable by the models. The use of alternative architectures would require early access to systems for development and evaluation as well as the availability of HECC experts on this platform to help identify code modifications needed.

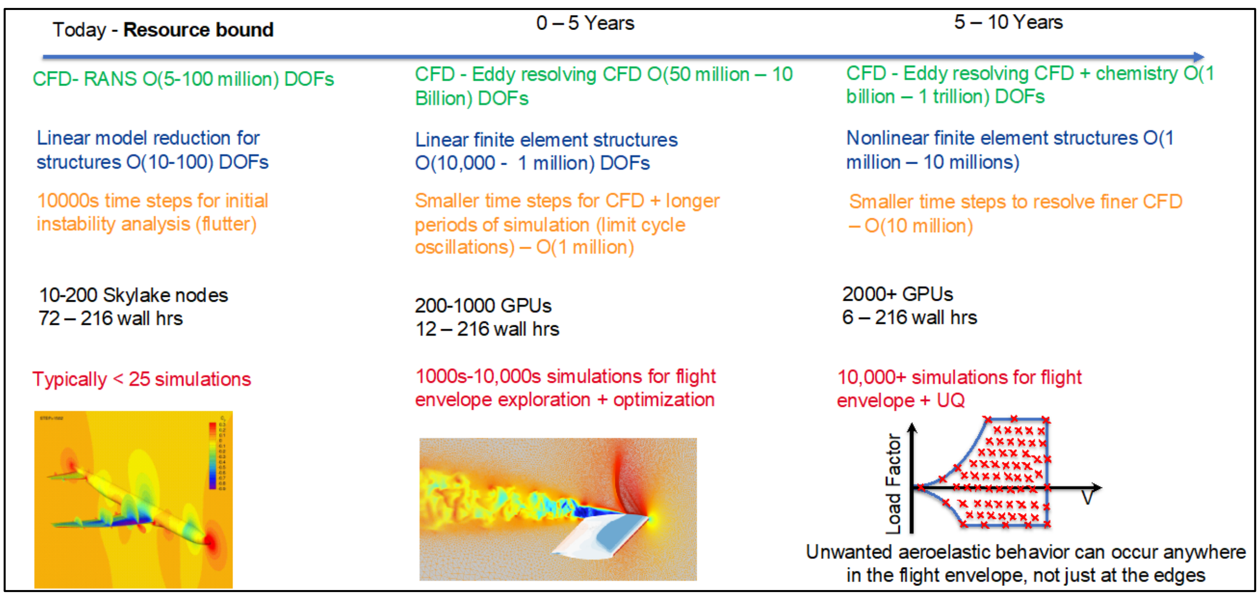

FIGURE 7: AEROELASTIC SIMULATION OVER THE NEXT DECADE. (CREDIT: K. JACOBSON, NASA LARC)

Another AAVP research program requiring HECC resources is multi-fidelity modeling of boundary-layer transition (BLT). Accurate prediction of the state of the flow (laminar vs. turbulent) is a cross-cutting requirement for modeling the aerothermodynamic environment of aerospace vehicles across the speed regime. Large-eddy simulation (LES) is also critical to this analysis. The myriad paths to laminar-turbulent transition and their sensitivity to external disturbances require a multi-fidelity prediction capability. Direct numerical simulation (DNS) must assume a critical role because of measurement (and facility) limitations, especially at high speeds. Overall HECC workflow includes a combination of high-order Navier-Stokes flow solvers, special purpose codes based on various levels of approximations to the Navier-Stokes equations, and deep learning codes that encapsulate the knowledge base to build regression models. The DNS simulations require substantial user-level monitoring because of the compromises on mesh size to allow timely and reasonable quality output on the typically congested and oversubscribed HECC system. A lot of effort is spent on balancing turnaround time and the need for user intervention. HECC resources are used continually for hierarchical simulations ranging from low-fidelity stability models, to phenomenological transport equation-based models, to DNS resolving of all spatiotemporal scales in the flow. Typically, multiple hierarchical computations are ongoing at the same time. Currently feasible mesh sizes (hundred million to a billion cells or somewhat larger) utilized on the HECC systems are barely adequate to serve the targeted goals. The highly competent and responsive user support staff cannot compensate for the limited hardware resources. Routine DNS of natural transition to enable the development of accurate, multi-fidelity reduced-order-models (ROMs) would require order(s) of magnitude larger resources than currently provided by HECC. In particular, the transition zone cannot be reliably modeled with lower order approximations such as LES and wall-modeled LES. Current codes easily scale up to 10,000–20,000 cores but are rarely exercised with more than 5,000 cores because of excessive wait times. (Appendix D6-9.4, Choudhary; Appendix D6-2.1, Korzun; Appendix D6-2.2, Kleb)

Note that the goal of physics-based transition prediction to support CFD Vision 2030 cannot be accomplished without the resources provided by the HECC project. The implications of not having this prediction capability are (1) substantially higher modeling uncertainty, (2) enhanced risk for human re-entry missions, and (3) inability to mature breakthrough concepts for reduced fuel burn. GPU resources are needed for deep learning analysis of the outputs to build regression models. The AAVP research team has evaluated substitutes for NASA's HECC resources, including DOE/DoD machines and reliance on academic partners with resources provided by National Science Foundation (NSF) sites. All of these options were considered inadequate for reasons of security, authorization and the application process. Another relevant issue is the increasing difficulty of working with external, academic partners. This research area is expected to grow significantly. Continued development of data-driven modeling will entail increased emphasis on simulation databases for new classes of vehicles. The grid sizes will continue to increase to accommodate more complex configurations or to enable higher spatial resolution. While high-fidelity DNS modeling of BLT is likely to remain a specialist's domain, the increased criticality for use with high-speed vehicles may lead to a modest increase in the number of HECC users of this use case. A significantly larger number of users can be expected at the low- to intermediate-fidelity modeling as that end of the capability matures and, also, as data-driven models for BLT are integrated into CFD codes. (Appendix D6-9.4, Choudhary; Appendix D6-2.1, Korzun)

4.2 Airspace Operations and Safety Program (AOSP)

AOSP uses a minimal amount of HECC resources primarily related to engineering studies and understanding the data related to safety. The HECC Project supports the AOSP models and analysis with high node-count runs and visualization of simulation output using the HECC capability at Ames. The history of AOSP project usage is displayed in Table 4-6. The HEC Program uses Standard Billing Units (SBU), described in Appendix A as the metric by which allocations are made and consumption is measured. The major AOSP project using HECC resources is listed in Table 4-7.

TABLE 4-6: AOSP HECC USAGE FROM 2015 TO 2020.

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	5,795	229,027	53,695	536,069	45,858	77,669

TABLE 4-7: AOSP PROJECTS ON HECC.

PI	Organization	Project Title	FY20 SBUs Consumed
Oza, Nikunj	NASA/ARC	Discovery of Precursors to Safety Incidents	5,795

A long-term hope of the U.S. aviation community is that the collection of sufficient data regarding safety mishaps can improve the safety record of the national airspace. NASA has been collecting and analyzing data on incidents and mishaps for over 30 years. Industry is currently using a rule-based approach to collecting the data, with rules typically involving 2–3 variables and corresponding to known safety issues—they cannot find previously unknown safety issues. Recent advances in artificial intelligence show a way to use data with more than 150 variables and thousands of flights per day acquired from aircraft operations (flight operations quality assurance, or FOQA), and trajectory data on takeoffs and landings at airports and metroplexes. There are very few operationally significant events to use for training. However, under the Aviation Safety Information Analysis and Sharing (ASIAS) partnership, new techniques for anomaly detection and precursor identification were developed. These include active learning, a machine learning technique in which the application, upon discovering an uncorrelatable input—using the existing labeled training data—queries domain experts to generate new labels, thereby reducing false alarms. Using this method, the project identified several safety-relevant anomalies not identified by the traditional method. The technique was demonstrated to improve one airline's own methods and is being developed as an operational capability within ASIAS for application to many airlines' data. HECC assets are used for testing for use with large datasets and operational applications. (Appendix D8-15, Oza)

4.3 Integrated Aviation Systems Program (IASP)

The HECC Project supports the IASP models and analysis with high node-count runs and visualization of simulation output using the HECC capability at Ames. The history of IASP project usage is displayed in Table 4-8. The HEC Program uses Standard Billing Units (SBU), described in Appendix A as the metric by which allocations are made and consumption is measured. Major IASP projects using HECC resources are listed in Table 4-9.

TABLE 4-8: IASP HECC USAGE FROM 2015 TO 2020.

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	3,900,784	5,611,229	4,474,068	3,783,550	3,220,486	4,173,518

TABLE 4-9: MAJOR IASP PROJECTS ON HECC IN FY20.

PI	Organization	Project Title	FY20 SBUs Consumed
Khorrami, Mehdi R.	NASA/Langley	Evaluation of noise reduction concepts for flaps	1,236,947
Kiris, Cetin	NASA/ARC	High Fidelity CFD support for X-57 Aerodatabase	959,886
Deere, Karen A.	NASA/Langley	Electric Propulsion Concept	898,668
Elmiligui, Alaa	NASA/Langley	Supersonic Low-Boom/Low-Drag Computations	372,520
Khorrami, Mehdi R.	NASA/Langley	Airframe noise simulations of a full scale aircraft	285,842
Dippold, Vance Fredrick	NASA/Glenn	Low Boom Flight Demonstrator Integrated Propulsion System Validation	111,404
Kiris, Cetin	NASA/ARC	CFD Support for Low Boom Flight Demonstrator	34,620
Melton, John E.	NASA/ARC	NAH X-Plane Analysis	898

HECC resources are used to generate large-scale databases of aircraft characteristics in a virtual wind tunnel capability. Database generation is often used to reduce the risk associated with a wind-tunnel or flight test. It can provide much needed guidance to the physical test, which reduces the time and costs associated with the experiment. It also helps with design decisions and NASA studies of system performance under varying conditions. HECC is essential for providing the required compute capacity for large scale database generation. Limited availability of HECC resources and delays in job execution are already limiting the scale of work that can be accomplished. (Appendix D6-8.2, Kenway; Appendix D9.8)

Some ARMD projects need to evaluate system behavior under extreme conditions. Three specific cases were described by users. First, system behavior predictions often need to forecast characteristics on the edge of the operational envelope, where physical experiments are too dangerous to run. Second, some knowledge is needed of system behavior under conditions that are too expensive to run multiple times, such as airframe drop tests. Third, some experimental facilities, such as wind tunnels and structures labs, are overbooked to the point that they are unavailable during the time the analysis is needed. Projects that need to assess behavior where experiments are impractical, infeasible, or impossible substitute computational experiments on HECC resources. As the CFD Vision 2030 Study puts it, “In many instances, CFD provides the only affordable or available source of engineering data to use in product design due to limitations either with model complexity and/or wind tunnel capability, or due to design requirements that cannot be addressed with ground-based testing of any kind.” (Slotnick et al., 2014, p. 6). Adding to the computational demand, these studies require multiple runs to create a probability distribution function rather than a single point solution. (Appendix D6-9.3, Jacobson)

HECC resources are used in evaluating the sonic-boom reduction achieved in the X-59 Quiet Supersonic Technology X-plane. Aerospace engineers at Ames ran high-fidelity computational fluid dynamics (CFD) simulations on Pleiades, Electra, and Endeavour supercomputers to help shape the design of the X-59, using codes including LAVA, Cart3D, FUN3D, USM3D, and PCBOOM. NASA's production-level Cart3D simulation package for CFD was used to determine the pressure field near the aircraft and evaluate the ground noise carpet of each major design evolution of the X-59, as shown in Figure 8. Cart3D was coupled with an atmospheric propagation solver to estimate the noise level on the ground. And with uncertainty quantification tools to provide uncertainty estimates in the pressure signatures due to variations in the aircraft's operating conditions and configuration. The simulations contributed to many design improvements such as reducing the noise generated by the nose of the aircraft, instrumentation probes, and secondary-air-systems inlets. Cart3D is also used to support supersonic wind tunnel tests. Necessary code tuning and improvements required for this code to be used in new ways was supported by the HECC Project's computer scientists and were deemed essential to the successful improvement of performance. (Appendix D6-1.4, Doebler)

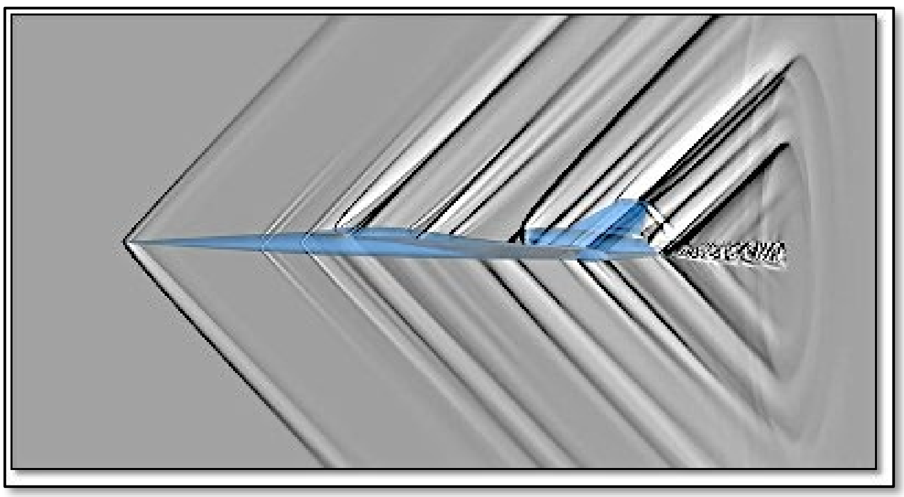

FIGURE 8: A FRAME FROM A VIDEO FROM A CART3D SIMULATION SHOWING THE COMPLEX SHOCK SYSTEM OF NASA'S X-59. DARK AND BRIGHT REGIONS REPRESENT SHOCKWAVES AND EXPANSIONS, RESPECTIVELY. WEAKER SHOCKS PROPAGATE FROM THE LOWER SURFACE OF THE AIRCRAFT, QUIETING SONIC BOOMS TO SONIC THUMPS ON THE GROUND. (CREDIT: M. NEMEC, M. AFTOSMIS, NASA AMES)

As with AAVP, the use of LAVA to populate a Reynolds-Averaged Navier Stokes (RANS) database is being expanded using HECC resources to cover IASP projects like the X57 Maxwell, the X59 LBFD, and Novel Propulsion Airframe Integration. The simulation output feeds into large-scale databases for a number of ARMD projects and HECC storage resources are used to retain the data generated. Over 100 LAVA simulations of the X57 Maxwell in a wide variety of flight configurations were performed. In preparation for the National Transonic Facility (NTF) wind tunnel testing campaign at LaRC, over 100 simulations of the NASA Boundary-Layer Ingestion (BLI) wind-tunnel model were run with the Common Research Model. HECC provides ARMD projects with "numerical wind tunnel" capability and feeds into a database of characteristics that can be used by the research community. (Appendix D6-8.2, Kenway)

4.4 Transformative Aeronautics Concepts Program (TACP)

TACP is the largest user of HECC resources in ARMD and supports a wide range of computational modeling and simulation efforts, including computational fluid dynamics, computational materials, and computational structures. The HECC Project supports the TACP models and analysis with high node-count runs and visualization of simulation output using the HECC capability at Ames. The Transformational Tools and Technologies (T3) project develops state of the art computational and experimental tools and technologies that are vital to prediction of aircraft behavior in flight. HECC is needed for critical ARMD milestones listed in Table 4-10.

The history of TACP project usage of HECC is displayed in Table 4-11. The HEC Program uses Standard Billing Units (SBU), described in Appendix A, as the metric by which allocations are made and consumption is measured. Major TACP projects using HECC during 2020 are listed in Table 4-12.

TABLE 4-10: T3 HECC USAGE SUPPORTS CRITICAL ARMD MILESTONES.

Technical Challenges (TCs) and Agency Annual Performance Indicators (APIs)
Develop and demonstrate computationally efficient, eddy-resolving modeling tools that predict maximum lift coefficient (CL*max*) for transport aircraft with the same accuracy as certification flight tests. (2025). Supporting U.S. industry to bring improved aircraft products to market faster and with greater confidence in performance.
Predict the sensitivity of lean blowout (and soot emissions) to changes in fuel composition occurring with the use of alternative fuels (or blends) where the relative difference in fuel sensitivity between simulations and experiments is less than 20%. (2021). Enables early assessment of promising candidate alternative fuels without costly and time-consuming experimental testing.
Complete detailed analysis of turbulent heat flux data obtained from NASA's Turbulent Heat Flux (THX) experiment to enable better computational tools for prediction and design of future air vehicle propulsion systems. (2020). Provide tools that enable accurate prediction of film cooling used for protection of surfaces from hot combusting gases in propulsion systems.

TABLE 4-11: TACP HECC USAGE FROM 2015 TO 2020.

	FY20	**FY19**	**FY18**	**FY17**	**FY16**	**FY15**
SBU	17,772,981	14,089,739	14,366,645	9,496,938	7,085,593	5,327,209

TABLE 4-12: TACP TOP TEN PROJECTS ON HECC.

PI	Organization	Project Title	FY20 SBU's Consumed
Debonis, James Raymond	NASA/Glenn	Large-Eddy Simulations for Propulsion Flows	2,346,894
Moder, Jeffrey P	NASA/Glenn	National Jet Fuel Combustion Program CFD	1,936,197
Malik, Dr. Mujeeb R.	NASA/Langley	LES of High RE Flows	1,531,545
Murman, Scott M.	NASA/ARC	ARC for RCA	1,513,314
Georgiadis, Nicholas J.	NASA/Glenn	Advanced Propulsion Turbulent Predictions Methods	1,442,904
Jansen, Kenneth	University of Colorado, Boulder	UCB-ScaleResolvingSimulations	951,787
Kiris, Cetin	NASA/ARC	TTT Workshop participations for V&V	791,920
Lawson, John	NASA/ARC	Aeronautical Sciences Computational Materials Science	777,515
Moin, Parviz	Stanford University	Validation of wall models for LES with application to the NASA Common Research Model	696,423
Vyas, Manan	NASA/Glenn	Large Eddy Simulation of Propulsion Flowpath	557,520

HECC supports TACP in the development of scale-resolving aeropropulsion CFD tools for ARMD, other mission directorates, and the broader aerospace community. This fundamental research into large-eddy simulation, numerical methods, and fluid dynamics uses massively parallel, computationally intensive codes, and cannot be done without the special capabilities of HEC, using large node-counts and reservations to be able to run the job. Use of a commercial supplier would be prohibitively expensive for the capacity required. HECC capacity and workload already restrict the widespread use of scale-resolving simulations due to cost and turnaround time. This also limits the size of the problems that can be analyzed. (Appendix D6-9.3, Cadieux)

Prediction of the maximum lift coefficient (CL_{max}) prediction is used to indicate the stall angle of an airfoil. This calculation is essential to enable aircraft certification by analysis, as shown in Figure 9, and is one of the major Technical Challenges (TC) of TACP (**Error! Reference source not found.**). HECC resources enable the solution of the nonlinear partial differential equations with billions of grid points needed for modern aircraft designs. The TC requires development/assessment of eddy resolving methods that require orders of magnitude more HECC resources than the current RANS-based methodology, which is unable to predict CL_{max} with the required accuracy. A significant number of simulations will be needed over the next five years employing varying fidelities (DNS, LES, and wall-modeled large eddy simulation, or WMLES), all of which require between 1 and 3 orders of magnitude more computational capacity compared to the use of RANS-based simulations, as shown in **Error! Reference source not found.**. The TC goal is to mature the WMLES technology for practical configurations, but DNS and LES are required for canonical configurations to assess lower fidelity (RANS and WMLES) tools and develop reduced cost physical models. The output data is compared with ongoing experiments to establish efficacy of computational tools, develop more efficient methods, and establish best practices for performing industrial engineering. Currently, either HECC CPU or GPU computing assets are used based on availability at the times the jobs are run. Timely simulations would require that the HECC Project expand its hardware capability by 100x in 5 years. (Appendix D6-8.6, Malik)

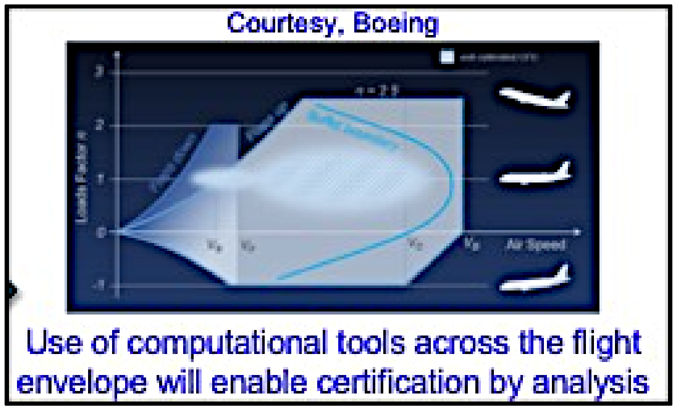

FIGURE 9: CLMAX MISSION CONCEPT. (CREDIT: BOEING CO.).

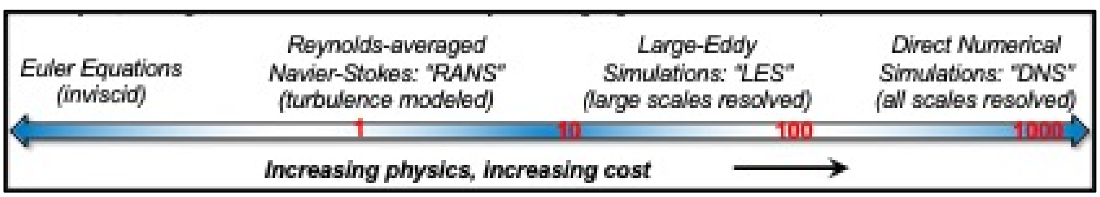

FIGURE 10: DEVELOPMENT OF ACCURATE COMPUTATIONAL TOOLS TO ENABLE AIRCRAFT CERTIFICATION BY ANALYSIS, SAVING HUNDREDS OF MILLIONS OF DOLLARS BY REDUCING FLIGHT CERTIFICATION TEST POINTS. (CREDIT: M. MALIK, NASA LARC)

HECC also supports several elements of the T3 Combustion research campaign by enabling the use of computational tools for advanced propulsion systems and alternative fuels assessment. One such customer is the multi-agency National Jet Fuels Combustion Program. The overall objective of this program is to develop combustion-related generic test and modeling capabilities that can improve the understanding of the impact of fuel chemical composition and physical properties on combustion, leading to accelerating the approval process of new alternative jet fuels, an example of which can be seen in Figure 11. Visualization of results is another HECC contribution that improves understanding of the results. This program is a major user of the ARMD allocation assigned to TACP and is currently limited in performing large-scale computations by the capacity of HECC resources available. Indications are that its demand for HECC resources is reasonably flat. (Appendix D6-8.8, Rogers)

HECC assets are used by TACP's research into computational materials, which encompasses development of tools to reduce time and cost required to design, develop, certify, and sustain materials. These capabilities are used by NESC and STMD, as well. One such development effort is to develop and validate atomistic simulations of materials which are used to predict properties of materials from physics-based first principles.

The Artificial Neural Network Molecular Dynamics Simulation (ALADYN) mini-application was developed under the HPC Incubator program with assistance from the HECC staff. ALADYN simulates aluminum crystal structure and behavior with quantum mechanics precision using machine learning. The optimized code is being implemented in a NASA-developed code for atomistic simulations, Parallel Grand Canonical Monte Carlo (ParaGrandMC), which is being used by other agencies and universities, as well. Computational materials research is critically dependent on HECC resources to support tool development; the tools are then used to accelerate development of high-performance, application-specific materials for a wide range of aerospace applications, e.g. metal alloys, composites, batteries, coatings, biosystems, nanosystems, etc. One example of this work is the development of shape memory alloys (SMAs), which have many benefits for morphing aircraft structures. HECC materials simulation can be used to design SMAs with application specific properties. The use case predicts phase transition between crystal structures and can be used to predict electronic structure using density functional theory (DFT) quantum mechanical modeling. Without HECC, the TACP would need to make an investment in a system or to try to use commercial cloud computing. (Appendix D6-8.9, Lawson)

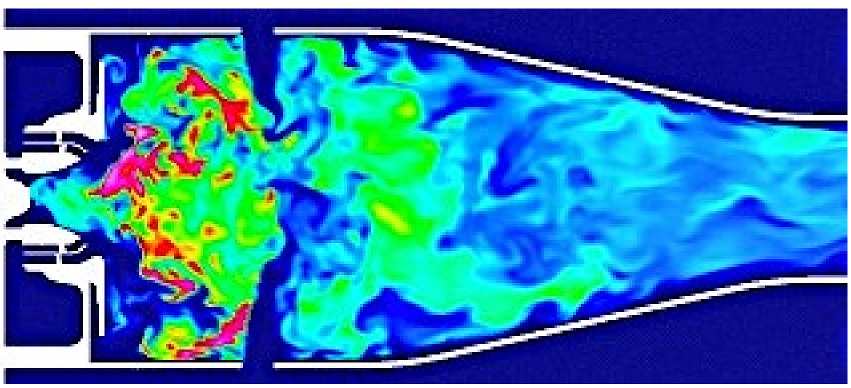

FIGURE 11: C1= GEVO ALCOHOL-TO-JET FUEL USED IN NJFCP EXPERIMENTS. (CREDIT: M. ROGERS, NASA ARC)

Similarly, durability, damage tolerance and reliability characteristics of materials are being simulated with finite element modeling and analysis. To optimize the most computationally intensive elements of the finite element analysis tools, the HPC Incubator funded a mini-app, FEMERA, to conduct code optimization to enable larger, more detailed, high-fidelity analyses using NASA resources more efficiently. This work was conducted in conjunction with the HECC Applications Performance and Productivity team and achieved a 1.85-fold speedup. The joint LaRC-HECC team also applied new features of OpenMP to coordinate computations between CPUs and GPUs and tested this on a GPU-based system. This project is an example of the partnering that HECC creates to support improvements that pay off for ARMD. (Appendix D6-4.1, Wagner)

5. HEC Support to Human Exploration Allocation Group

Three NASA Headquarters Organizations are managed in a single allocation group: Human Exploration and Operations Mission Directorate (HEOMD), the NASA Engineering Safety Center (NESC), and the Space Technology Mission Directorate (STMD).

5.1 Human Exploration and Operations Mission Directorate (HEOMD)

HEOMD manages the International Space Station (ISS) and develops the next generation of rockets, spacecraft, and other capabilities that extend human presence throughout the solar system.

Table 5-1 shows the actual usage for the major HEOMD programs. Note that the high usage in FY18 was due to the Launch Induced Environments studies, which required an additional 10 million SBUs. The HEC Program uses Standard Billing Units (SBU), described in Appendix A, p. 83, as the metric by which allocations are made and consumption is measured. HEOMD's usage is determined by the studies and engineering analyses that are needed for mission reviews and decisions. The top ten users of HECC resources in FY20 are listed in Table 5-2.

TABLE 5-1: HEOMD HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	14,184,753	19,036,950	24,162,201	12,727,952	4,950,455	5,390,831

TABLE 5-2: HEOMD TOP TEN HECC PROJECTS.

PI	Organization	Project Title	FY20 SBUs Consumed
Rogers, Stuart E.	NASA/ARC	CFD Support for Space Launch System (SLS) Ascent Aerodatabase	2,043,493
West, Jeff	NASA/MSFC	Launch Induced Environments	1,863,306
Krist, Steven E.	NASA/Langley	SLS Aerodynamics at LaRC	1,691,175
Kiris, Cetin	NASA/ARC	CFD support for the Launch Environment of the Next Generation Launch Vehicles	1,233,961
Kiris, Cetin	NASA/ARC	High Fidelity Simulations to Support LAS QM1 test	1,008,612
West, Jeff	NASA/MSFC	Fluid Dynamics Support for CATALYST Program	970,558
Canabal, Francisco	NASA/MSFC	MSFC SLS Thermal Environments	922,108
West, Jeff	NASA/MSFC	Tanks and MPS Simulations	782,166
Canabal, Francisco	NASA/MSFC	MSFC CCP Fluid Dynamics	696,331
West, Jeff	NASA/MSFC	Liquid Engine Simulations	461,889

Three application areas have indicated their need for exascale computing to achieve future objectives:

- Launch vehicle design.
- Computational fluid dynamics (CFD) virtual wind tunnel for scale-resolving models.
- Launch ground environment.

HECC primarily supports HEOMD with the capacity to run many cases through existing models for study, as well as mission design and evaluation in the full range of the mission’s portfolio, as seen in Figure 12.

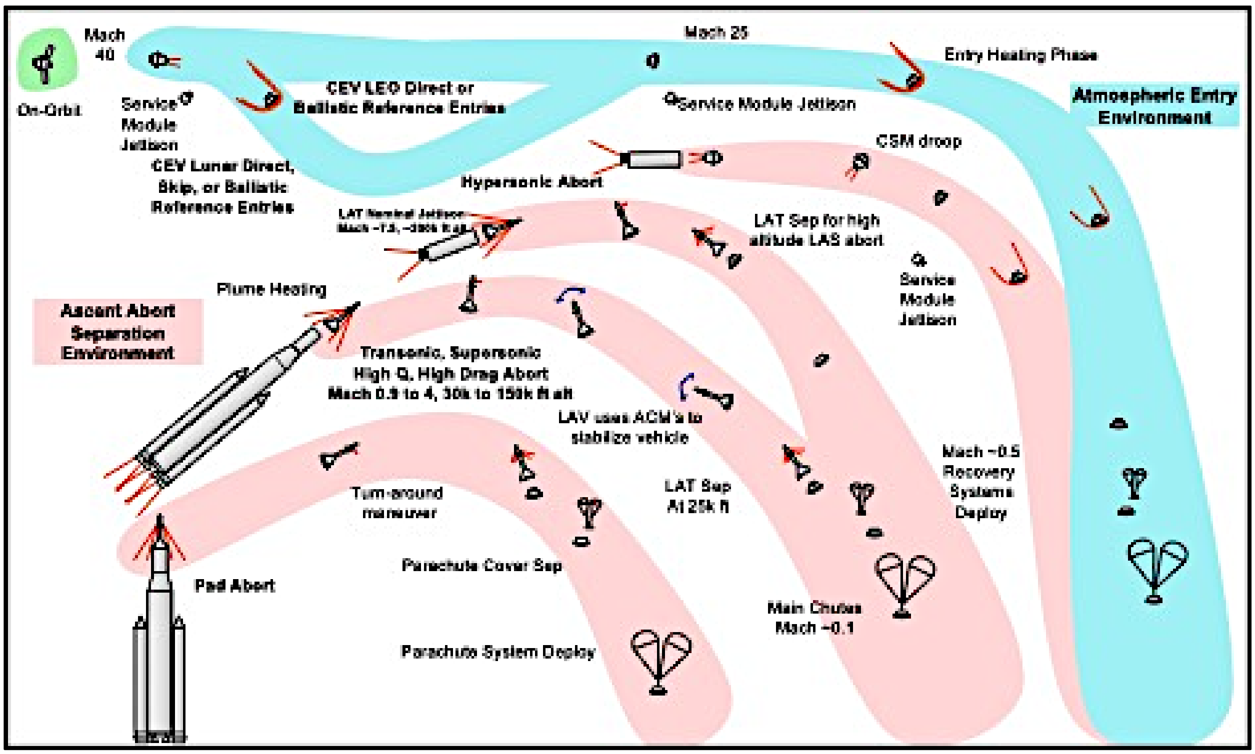

FIGURE 12: HECC SUPPORTS ENGINEERING FOR HEOMD FLIGHT. (SOURCE: APPENDIX D6-8.5, GOMEZ)

HECC supported the development of a booster separation aerodynamic database for the Artemis II mission. Aerospace engineers at NASA Ames ran CFD simulations on the Pleiades and Electra supercomputers to formulate the booster separation databases for NASA’s Artemis II flight configuration. Created using NASA’s FUN3D flow solver, the calculations include the aerodynamic effects of 22 different plumes during the booster separation event and 13 independent variables. Additionally, separate sets of computations were run to simulate possible effects caused by a core-stage engine failure. These databases are used by the Guidance, Navigation, and Control group at NASA Marshall to create dynamic simulations of the booster separation event for Artemis II to ensure the boosters can separate successfully without reconnecting to the core under all possible flight conditions. (Appendix D9.8, Appendix D8.10, Malik, Nielsen, Schuster, Bushnell)

The CFD simulations that construct these databases allow engineers to model flight conditions and complex interactions that are difficult to test in an experimental setting. (Appendix D6-8.3, Cadieux)

HECC resources are also used to run Launch Ascent and Vehicle Aerodynamics (LAVA) Cartesian adaptive mesh refinement (AMR) for HEOMD to predict (Appendix D6-8.3, Cadieux):

- Ignition over-pressure (IOP).
- Vibro-acoustic loads on launch vehicles.
- Main Flame Deflector (MFD) design and construction at NASA's Kennedy Space Center (KSC).

The value of HECC's contribution to this CFD work is calculated by evaluating experimental validation, ground and flight test data, and the extent to which design decisions and changes occur based upon CFD model output rather than test data. HECC resources are used to perform scale-resolving simulations and to preserve the output. For example, recent runs included 12 Orion Launch Abort simulations using the LAVA flow solver. Until recently, CFD was not used to determine the vibro-acoustic environment of those simulations. Now, in combination with test data, CFD is being used to drive down uncertainty margins. Projections of future work suggest it will be demand-driven rather than at a regular interval. However, the engineers can only run a handful of Orion Launch Abort scenarios, or KSC launch simulations in a year, primarily due to lack of concurrent computing resources within the allocation made to the HEOMD group. It is difficult to get one 16,000 core reservation for one month, even more so the 5–12 reservations needed in a year. The project's ultimate goal is to cover a database of scenarios in under one year. (Appendix D6-8.3, Cadieux)

HEC assets also play a key role in engineering design changes during construction. LAVA simulation results were directly used to make design decisions for the re-design of KSC Main Flame Detector (MFD) after KSC decided to stop construction. New LAVA simulations helped determine course corrections, which led to restarting construction. (Appendix D6-8.x, Cadieux)

The development of aerodynamic/aerothermal databases describing aerodynamic forces, moments, and heating distributions for Commercial Crew, Multi-Purpose Crew Vehicle (Orion spacecraft), and Space Launch System (SLS) spacecraft is another way in which HECC resources support HEOMD. Databases are developed to support design cycles, verification cycles, and flight-readiness analysis cycles. HECC resources have a critical role in the extension of databases to flight conditions and are needed on a daily basis. Without HECC, the vehicles would be designed with reduced performance to allow for greater margin for uncertainty and costs would increase. The development of these databases is impacted by the availability of compute resources and restrictions placed on the handling of ITAR, and CCP TPPI data and code. This work must be accomplished at an appropriately secured facility. (Appendix D6-8.5, Gomez)

Partnerships with commercial firms, such as Boeing and Dassault introduce commercially licensed code, such as PowerFLOW into the modeling environment, leveraging investments by other organizations. As part of the Commercial Crew Program, Sierra Nevada Corporation leveraged several of these codes to develop aerodynamic and aerothermal preflight databases for the reusable Dream Chaser® spacecraft, which were then used to predict behavior of the flight control system and the thermal protection systems (TPS). This involved over 20,000 unique high-fidelity CFD simulations to accurately characterize the aerodynamic forces on the airframe and control surfaces during atmospheric flight, thereby enabling the Dream Chaser® to provide payload capacity to and from the International Space Station (ISS). The results, shown in Figure 13, visualize an important flight control problem: as the thrusters fire at hypersonic velocities, they create a reaction by the control surfaces that significantly changes vehicle performance. (Appendix D9.1)

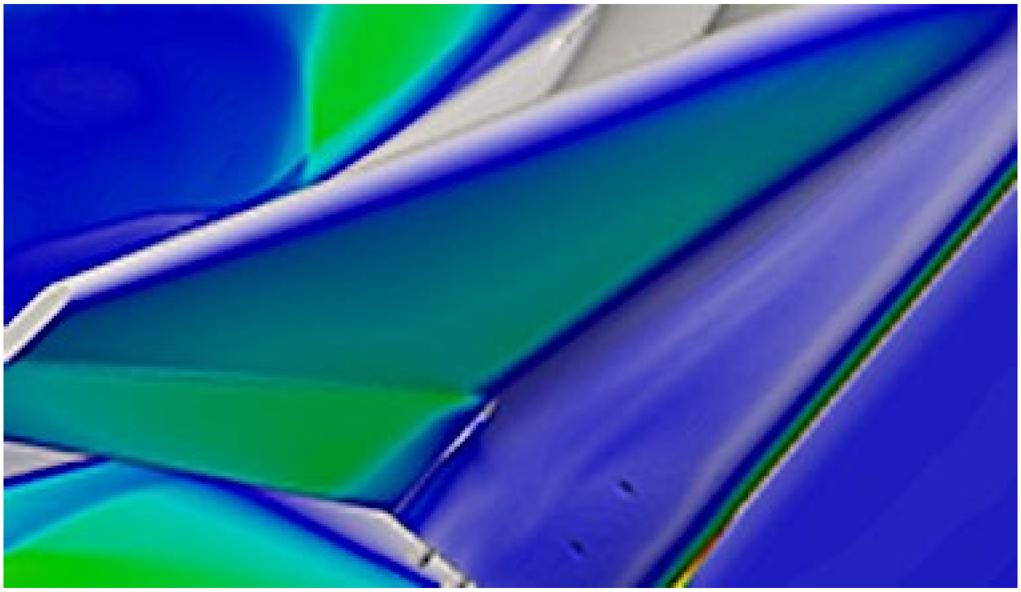

FIGURE 13: VISUALIZATION OF THE REACTION CONTROL SYSTEM THRUSTER FIRING AT HYPERSONIC VELOCITIES DURING ENTRY, DESCENT, AND LANDING OF THE DREAM CHASER SPACECRAFT. (CREDIT: MATT OPGENORTH, SIERRA NEVADA CORPORATION)

The availability of computing resources has an impact on critical path calculations for launch schedules. As the SLS is deployed, actual flight data will be used to refine simulations and models, and the SLS family of vehicles will require aerodynamic analysis for the next decade. Orion will move into a sustaining engineering stage. Rehosting of code is possible, although the performance improvements are not well understood, and expert assistance will be needed to migrate the code onto any new computing architecture. Today, the CFD analysis performed for the SLS is constrained by the current HECC capacity. Turnaround time is a large limiter for some of the desired moving-body and 6-degree-of-freedom cases. More processors and longer run times are the only current mitigations. The SLS Program often cannot afford to perform analyses due to the cost and load they would place on computing resources. However, more accurate, higher-fidelity analysis in the preliminary design cycles will reduce future design cycle time, or even eliminate entire design cycles. (Appendix D6-8.5, Gomez)

5.2 NASA Engineering and Safety Center (NESC)

At the core of the NESC is an established knowledge base of technical specialists pulled from the ten NASA centers and from a group of partner and organizations external to the agency. This ready group of engineering experts is organized into discipline areas called Technical Discipline Teams (TDT), whose members represent NASA organizations, industry, academia, and other government agencies. By drawing on the minds of leading engineers across the country, the NESC consistently optimizes its processes, deepens its knowledge base, strengthens its technical capabilities, and broadens its perspectives, thereby further executing its commitment to engineering excellence.

The NESC Technical Discipline Teams are:

- Aerosciences
- Avionics
- Cryogenics
- Electrical Power
- Environmental Control/Life Support
- Flight Mechanics
- Guidance, Navigation & Control
- Human Factors
- Loads and Dynamics
- Materials
- Mechanical Systems
- Nondestructive Evaluation
- Nuclear Power and Propulsion
- Passive Thermal
- Propulsion
- Robotic Spaceflight
- Sensors/Instrumentation
- Space Environments
- Software
- Structures
- Systems Engineering

NESC's technical evaluation and consultation products are delivered in the form of written reports that include solution-driven, preventative, and corrective recommendations. To further this goal, the NESC is currently leading NASA's efforts for independent data mining and trend analysis. Using engineering and safety data, the NESC established a Data Mining and Trending Working Group that includes representatives from all NASA centers as well as external experts. This group ensures that results are maximized and that the NESC comprehensively learns from previous efforts.

The actual usage of HECC capacity for the past five years for NESC is shown in **Error! Reference source not found.**. The HEC Program uses Standard Billing Units (SBU), described in Appendix A, as the metric by which allocations are made and consumption is measured.

TABLE 5-3: NESC HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	5,736,746	3,098,034	1,720,752	1,763,377	1,045,353	46,598

The NESC’s usage of HEC grew considerably, by a factor of three since 2018 due to individual, specific studies, primarily in the use and study of CFD. Table 5-4 shows the NESC’s HECC projects for FY20; all have been projects for at least the past two years.

TABLE 5-4: NESC HEC PROJECTS FOR FY20.

PI	Organization	Project Title	FY20 SBUs Consumed
Streett, Craig	NASA/Langley	Numerical Simulations for AeroSciences Research	2,845,845
West, Jeff	NASA/MSFC	Commercial Crew Program Propulsion Efforts under the NESC	2,635,847
Dalle, Derek Jordan	NASA/ARC	Aerodynamic Buffet Flight Test Support	163,262
Chwalowski, Pawel	NASA/Langley	Aeroelastic Prediction Workshop	83,053
Streett, Craig	NASA/Langley	AA-2 Drag Anomoly CFD	4,509

The objective of the SLS Buffet and Aeroacoustics Project, which includes several of the HECC projects listed in Table 5-4, is to produce an accurate prediction of unsteady loads on the launch vehicle which cannot be measured physically. The simulation of the turbulent flow field around the launch vehicle, leading to unsteady forces on structure, is important to risk reduction. Data provided to structure-dynamics analysts leads to validation of design specs, identification of design exceedances, suggesting alternative structural concepts, and optimal designs. This work is dependent upon NASA HECC resources both due to the security restrictions and need for predictability in the availability of the resources. Output data reflects several million points each at several million timesteps. Initiation of an SLS Program loads analysis cycle requires many disciplines to simultaneously begin a computational campaign that might need 5-20 cases, each requiring 10-25 million core-hours. Additional post-processing across several million timesteps with several million points at each time step produces input for structural dynamics codes will be needed. (Appendix D6-2.6, Street)

5.3 Space Technology Mission Directorate (STMD)

Technology drives exploration to the Moon, Mars and beyond. NASA's STMD develops transformative space technologies to enable future missions. The 2017 NASA Strategic Technology Investment Plan (STIP) identifies the ten most critical technology research and development investment areas as:

1) Propulsion and Launch Systems.
2) Human Health, Life Support, and Habitation.
3) Destination Systems.
4) Robotics and Autonomous Systems.
5) Scientific Instruments, Sensors, and Optical Communications.
6) Lightweight Space Structures and Materials.
7) Entry, Descent, and Landing.
8) Space Power Systems.
9) Advanced Information Systems.
10) Aeronautics.

Table 5-5 describes the actual usage, in SBUs, for STMD. Note that the HEC Program uses Standard Billing Units (SBU), described in Appendix A as the metric by which allocations are made and consumption is measured.

TABLE 5-5: STMD HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	5,788,746	3,496,219	2,374,810	1,915,225	887,543	1,129,199

STMD's usage has grown significantly in the past three years, doubling between 2017 and 2020. This is primarily a result of increased exploration requirements and design studies. The development of improved models to handle supersonic retropropulsion and drag devices for planetary atmospheric entry have also expanded.

Table 5-6 shows the ten largest HECC projects under the STMD allocation, based on projected FY20 consumption. Three types of projects receive the largest share of HECC allocation to STMD:

- Entry, descent, and landing (EDL), including both retropropulsion and drag devices.
- Computational materials.
- Computational structures.

These STMD projects are essential precursors to human or robotic exploration for several missions planned in the near future and represent high priorities. Several STMD projects have indicated their need for exascale computing, as described in Chapter 1, and are developing plans for how to evolve to leverage this potential capability when it becomes available:

- Mars powered descent simulations.
- Uncertainty quantification.
- Aerothermal analysis.
- Multi-physics simulations (CFD+thermal+finite element analysis).

TABLE 5-6: STMD TOP HECC PROJECTS IN FY20.

PI	Organization	Project Title	FY20 SBU's Consumed
Ihme, Matthias	Stanford University	Advanced Physical Models and Numerical Algorithms to Enable High-Fidelity Aerothermodynamic Simulations of Planetary Entry Vehicle	1,705,655
Barnhardt, Michael D.	NASA/ARC	Computational Support for the Hypersonic Entry Descent and Landing Project	970,922
Edquist, Karl T	NASA/Langley	Evaluation of CFD as Surrogate for High Supersonic Wind Tunnel Testing (Propulsive Descent)	687,964
Volkov, Alexey	University of Alabama, Tuscaloosa	Mesoscopic modeling of vertically aligned carbon nanotube forests	534,052
Canabal, Francisco	NASA/MSFC	Mars Lander Aero/Thermal Environments	460,196

PI	Organization	Project Title	FY20 SBU's Consumed
Yoon, Seokkwan	NASA/ARC	Arc heater modeling	254,674
Brehm, Christoph	University of Kentucky	Modeling Transitional and Turbulent Flows with Surface Ablation	242,386
Brehm, Christoph	University of Kentucky	A New Numerical Method for Fluid-Structure Interaction with Large Deformations for Parachute Simulations	129,922
Korzun, Ashley M.	NASA/Langley	Powered Descent Aerosciences for EDL Architecture Study	121,579
Yoon, Seokkwan	NASA/ARC	High-Fidelity Material Response Modeling	118,259

The Entry, Descent, and Landing (EDL) analysis supports planning for human-scale and high-mass Mars landers that require propulsive descent and landing with potentially significant plume-induced aerodynamics and environments. Figure 14 shows this mission concept. In addition, parachute dynamics are important for any atmospheric entry.

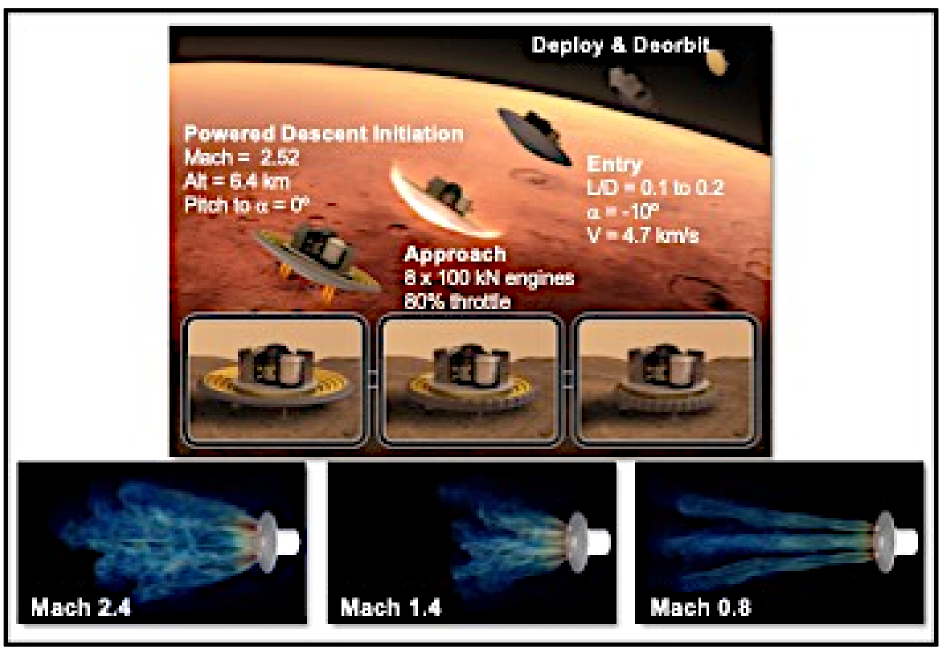

FIGURE 14: MISSION CONCEPT FOR MARS EDL AND RETROPROPULSION. (SOURCE: APPENDIX D6-2.1, KORZUN)

Limitations in ground testing require increased reliance on computational modeling and simulation for not only conceptual design, but also for flight implementation for mission infusion, vehicle performance evaluation, and risk characterization and reduction. Because of the sensitivity of the results during the maturation of these studies, flight implementation and realistic propulsion systems require increasingly controlled access, effectively eliminating external collaborators from industry and academia. EDL increasingly needs to use HECC for post-processing and scientific visualization, which is integral during and after solutions are completed. This data is used to quantify aeroscience impacts to vehicle performance and the definition of design requirements. (Appendix D6-2.1, Korzun)

Increasingly, EDL work requires computations larger than can be performed on NASA's HECC resources and requires the use of three different supercomputing facilities, including at the Department of Defense (DOD) and the Department of Energy (DOE). Output from intermediate steps are transferred from one facility to another, which delays the overall project schedule and adds labor to the process. In 2019, the DOE allocated $7.6M in computing to retropropulsion calculations. Today, only isolated, under-resolved, and low-fidelity simulations can be performed on HECC resources within the STMD allocation. Without adequate high-performance computing, exploration of increasing fidelities (space, time, and physical modeling) to keep up with mission planning requirements and eventual production database construction are entirely infeasible. Supersonic retropropulsion studies are computationally expensive, which, on NASA resources, experience multi-day delays waiting in queues. Without additional resources, mission directorates are unwilling to direct priority processing on a frequent basis. HECC resources have been limited to CPU-type processors, but experience on one EDL code indicates that high-capability GPU machines (such as Oak Ridge National Laboratory's Summit system) may accelerate the process and enable the high capacity required. (Appendix D6-2.1, Korzun, Appendix D6-2.3, Nielsen)

Uncertainty quantification (UQ) for expensive, high-fidelity models are essential and require additional runs to obtain the necessary variation in the data. UQ quantifies mission risk, decreases margins/costs, provides confidence to decision makers. Emerging paradigms like *digital twin* (DT) and *certification by analysis* (CBA) require rigorous UQ. UQ adds even more iterations to these already demanding computational model runs, as can be seen in Figure 15. Traditional Monte Carlo simulation-based approaches require tens of thousands of model evaluations to obtain sufficient sample population. Emerging data-driven (machine learning) approaches still require large volumes of data to train deep neural networks before they can be used. The most critical demand, from a security point of view, is availability; interrupted HECC availability results in unquantified or, at least, poorly quantified risks for missions. Lack of availability results in missed deadlines for research when analysis is particularly intensive, such as computing rare event probabilities and hyperparameter tuning for scientific machine learning. (Appendix D6-4.2, Warner)

A second aspect of EDL supported by HECC resources is the need to understand the extreme heating that occurs when a spacecraft enters an atmosphere. Aerothermodynamics is the estimation of this heat transfer in order to develop suitable thermal protection systems. The primary tool is computational fluid dynamics (CFD) simulation, specialized to address thermochemical non-equilibrium. Aerothermodynamic CFD is mission-enabling for human space exploration and many planetary space missions. Without adequate HECC resources, the various missions would be forced to acquire their own computational resources. These calculations are restricted as Sensitive, but Unclassified (EAR99, ITAR) and must be preserved as part of the mission design record. The current workload involves over 20 engineers running multiple cases each week on HECC resources. Data is significant for each of these studies—input is generally less than 1 GB, but outputs can be over 10 GB per simulation, and accumulating on the order of hundreds of simulations. Visualization is essential to understanding the resulting output. Faster turnaround times will improve engineering effectiveness and the potential for meeting mission objectives. (Appendix D6-4.3, Hill)

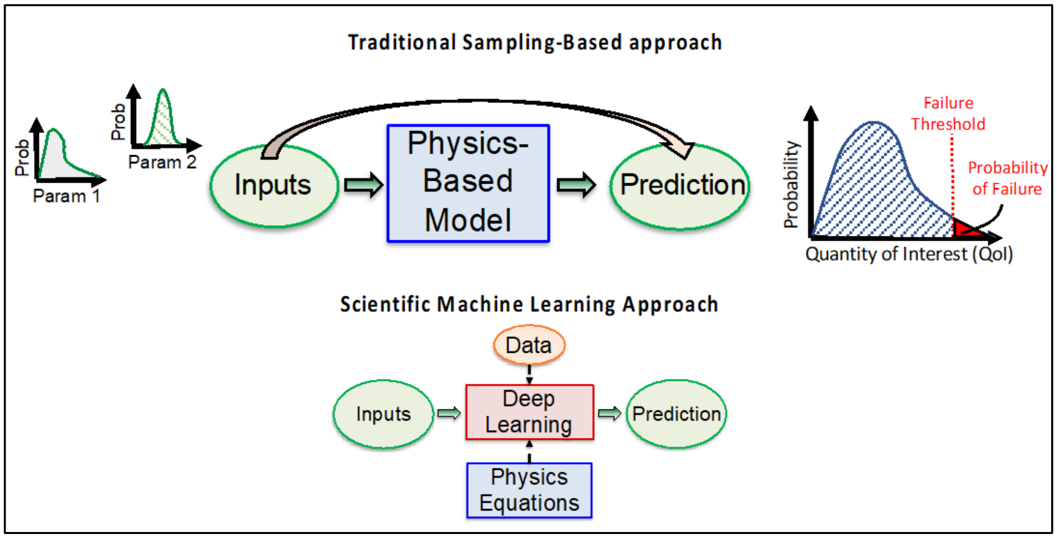

FIGURE 15: APPROACHES TO UNCERTAINTY QUANTIFICATION. (SOURCE: APPENDIX D6-4.2, WARNER)

A third aspect of EDL is the qualification of parachute entry systems. The current capability is restricted in fidelity because work is limited to one-off cases and adequate fidelity requires parametric analysis for parachute opening loads and failure analyses. Without HECC resources for modeling, extensive and expensive experimental campaigns are required, which carry greater residual risk. Current computational experiments run a few simulations per year, but a weekly cadence is expected in production, at which point current HECC resources are insufficient. The terabyte-scale results are used both for a risk assessment and also to inform design and manufacturing decisions. Analysis of the results relies on HECC visualization resources. These studies are expected to become more important to the broader commercial space community, which may need help in running the models for their own purposes. While GPU processing could enable more rapid and sophisticated analyses, technical assistance is necessary to migrate models into a new environment. The data volume to be analyzed will grow, as well. (Appendix D6-8.7, Barnhardt)

6. HEC Support to Science Mission Directorate (SMD)

The Science Mission Directorate (SMD) engages the nation's science community, sponsors scientific research, and develops and deploys satellites and probes in collaboration with NASA's partners around the world to answer fundamental questions requiring a view into and from space. SMD seeks to understand the origins, evolution, and destiny of the universe and to understand the nature of the phenomena that shape it. Organizationally, SMD is composed of five divisions as described in Table 6-1.

Note: The Biological and Physical Sciences Division's allocation has previously been carried in HEOMD. Due to a recent change, it is now carried in SMD, however this was not reconciled prior to the completion of this user needs assessment.

TABLE 6-1: SCIENCE MISSION DIRECTORATE DIVISIONS. (SOURCE: HTTPS://SCIENCE.NASA.GOV/ABOUT-US/SMD-VISION)

Division	Mission
Astrophysics	Research programs and missions necessary to discover how the universe works, explore how the universe began and developed into its present form, and search for Earth-like planets.
Earth Science	Technology development, applied science, research, flight mission implementation and operation to help us to understand our planet's interconnected systems, from a global scale down to minute processes.
Heliophysics	Studying key space phenomena and processes supports situational awareness to better protect astronauts, satellites, and robotic missions exploring the solar system and beyond.
Planetary Science	Understanding the history of our solar system and the distribution of life within it.
Biological and Physical Sciences	Pioneering scientific discovery and enabling human spaceflight exploration. (Source: https://science.nasa.gov/biological-physical)

SMD benefits primarily from the HECC resources, but the Earth Science Division (ESD) separately funds the NCCS, primarily for use in modeling climate. Limited allocations in the NCCS by other SMD divisions are authorized on a case by case basis through the allocation process.

Future growth of SMD's needs are difficult to predict quantitatively, but every major modeling user plans to increase resolution and time steps, and to add new physics to the models. All of these improvements are to provide more accurate and actionable forecasting. Some major increases include:

- The Geostationary Carbon Observatory (GeoCarb) mission, with computing scheduled to start in 2022, will add 3 million SBUs to each year's allocation after that, along with additional load from periodic re-processing.

- Development of new Goddard Earth Observing System (GEOS) components, which are expected to be completed in 2022 with higher vertical and surface resolution.
- Increased resolution of Goddard Institute for Space Studies (GISS) models in both vertical and surface dimensions.
- Increased resolution and time steps for Land Information System/Land Data Assimilation Systems (LIS/LDAS) will substantially increase their requirement for storage and processors, both in terms of job size and output volume and duration.
- Additionally, LIS/LDAS has matured to the point that several other users will start running models for specific scenarios, increasing the demand for processors and storage.

Experimentation with machine learning modeling will increase the demand for hardware acceleration using GPU platforms to perform training, while the conventional processors will be used to exercise the models. In order to generate the training data, the existing models will need to be run a number of times.

A number of SMD applications are preparing to be able to use exascale computing to improve model resolution, and performance:

- Uncertainty quantification.
- Full-physics atmospheric river transport.
- Subseasonal to seasonal coupled modeling and data assimilation.
- Physical oceanography modeling of circulation with tides.
- Modeling of black holes using 3D radiation magnetohydrodynamics (MHD).

SMD maintains and supports several facility codes and modeling frameworks that are migrating to new platforms, including GEOS, GISS, and LIS/LDAS. Codes such as these foster collaborations and are used by external users, contributing to NASA's role as a leader in Earth and space science.

Table 6-2 shows the actual usage for SMD programs of HECC resources for the past five years and Table 6-3 shows the NCCS resource usage. The HEC Program uses Standard Billing Units (SBU), described in Appendix A as the metric by which allocations are made and consumption is measured.

TABLE 6-2: SMD HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	52,914,596	37,298,819	29,879,752	28,584,274	16,930,359	14,530,800

TABLE 6-3: SMD NCCS USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	26,353,935	20,647,959	16,610,287	16,966,282	7,018,960	N/A

The top users of HEC resources in SMD are identified in Table 6-4 and for each division in their respective sections.

TABLE 6-4: TOP TEN SMD USERS OF HEC SBUS IN 2020.

PI	Organization	Title	FY20 SBUs Consumed
Oraiopoulos, Lazaros	NASA/GSFC	Improving clouds and radiation in GEOS-5 through algorithm development and evaluation against observations	7,970,424
Lucchesi, Robert A.	Science Systems and Applications, Inc.	Subseasonal to Decadal Climate Forecasts	6,598,209
Dong, Chuanfei	Princeton University	Integration of Extended MHD and Kinetic Effects in Global Magnetosphere Models	3,886,189
Gavin, Schmidt	NASA/GISS	GISS ModelE Development and Vision	3,560,453
Menemenlis, Dimitris	NASA/JPL	Requesting of Columbia High-End Computing Resources for ECCO2	3,028,190
Chan, Samuel	NASA/JPL	SWOT KaRIn Data Simulation	2,239,936
Dong, Chuanfei	Princeton University	Exoplanetary Space Weather, Climate and Habitability: Consequences of Atmospheric Loss	2,146,490
Gelaro, Ronald	NASA/GSFC	Atmospheric Data Assimilation Development	1,468,055
Putman, William	NASA/GSFC	GMAO - Systems and Data Synthesis	1,466,259
Lucchesi, Robert	SAIC	Global Data Assimilation Products	1,383,251

6.1 Earth Science Division (ESD)

In Earth Science, HEC enables science data processing as part of selected instrument elements of flight projects and ongoing missions. Multi-scale models require high levels of HEC resources, in each of the Research and Analysis (R&A) focus areas:

- Atmospheric Composition.
- Weather and Atmospheric Dynamics.
- Climate Variability and Change.
- Water and Energy Cycle.
- Carbon Cycle and Ecosystems.
- Earth Surface and Interior.

The Earth Science Technology Office and the Applied Sciences Program have limited needs for HEC resources.

Table 6-5 shows the actual usage for ESD programs for the past five years on HECC resources and Table 6-6 shows the usage on NCCS resources. The HEC Program uses Standard Billing Units (SBU), described in Appendix A as the metric by which allocations are made and consumption is measured. The top ten users of HEC resources in ESD are identified in Table 6-7.

TABLE 6-5: ESD HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	18,396,267	14,127,792	8,939,293	5,555,844	2,825,213	3,970,692

TABLE 6-6: ESD NCCS USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	23,791,526	20,260,330	16,112,960	16,130,966	6,679,776	N/A

TABLE 6-7: TOP TEN ESD USERS OF HEC SBUS IN 2020.

PI	Organization	Title	FY20 SBUs Consumed
Oraiopoulos, Lazaros	NASA/GSFC	Improving clouds and radiation in GEOS-5 through algorithm development and evaluation against observations	7,970,424
Lucchesi, Robert A.	Science Systems and Applications, Inc.	Subseasonal to Decadal Climate Forecasts	6,598,209
Gavin, Schmidt	NASA/GISS	GISS ModelE Development and Vision	3,560,453
Menemenlis, Dimitris	NASA/JPL	Requesting of Columbia High-End Computing Resources for ECCO2	3,028,190
Chan, Samuel	NASA/JPL	SWOT KaRIn Data Simulation	2,239,936
Gelaro, Ronald	NASA/GSFC	Atmospheric Data Assimilation Development	1,468,055
Putman, William	NASA/GSFC	GMAO - Systems and Data Synthesis	1,466,259
Lucchesi, Robert	SAIC	Global Data Assimilation Products	1,383,251
Cheng, Cecilia	NASA/JPL	Orbiting Carbon Observatory 2	1,091,088
Molod, Andrea M.	NASA/GSFC	Coupled predictions of ozone-climate: 1950-2100	1,006,381

In the Atmospheric Composition focus area, the Goddard Earth Observing System (GEOS) model has a lengthy history in the Goddard Modeling and Assimilation Office (GMAO). This model is community supported and receives enhancements and expansions through both limited dedicated funding and competitively selected projects. It is widely used for atmospheric data analysis, observing system modeling and mission design, retrospective reanalysis, climate and weather prediction, as well as for basic Earth science research. Recent advancements in GEOS has added significant complexity and increased resource requirements. As an example, adding active atmospheric chemistry has slowed its execution by up to a factor of eight when constrained to fit in the same computing capacity. To overcome these challenges, traditional optimization techniques are being pushed along with several efforts to apply machine learning to emulate the physics-based model components, such as chemistry, that will result in global air quality forecasts at half the computational cost. If successful, the extensive network of hundreds of research groups worldwide who use Figure 16 ozone modeling as analysis of more observational data improved understanding of the physical processes. The use of HEC is essential to meet three purposes:

- Analysis of the large volumes of data.
- Development of the model.
- Comparative analysis of the model output to the observational data to identify deficiencies.

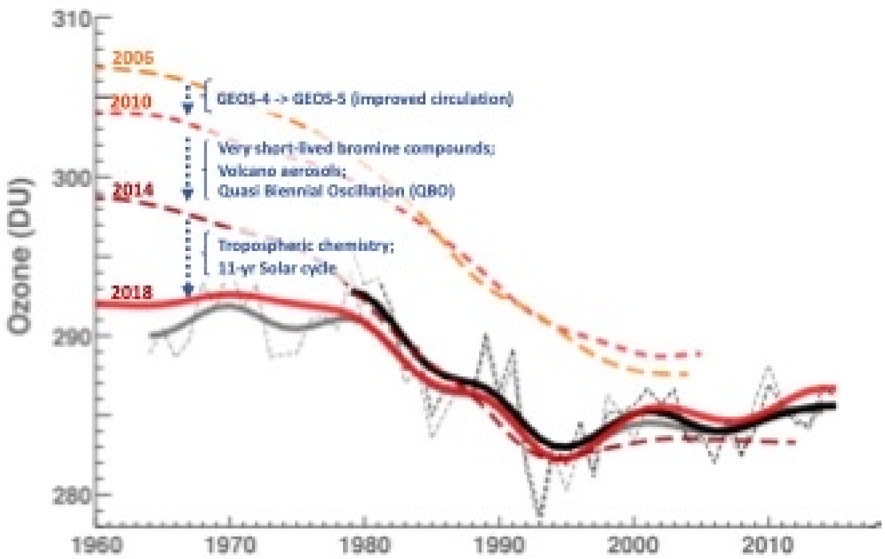

FIGURE 16: MORE INPUT OF NASA DATA BETWEEN 2006 AND 2018 IMPROVES MODEL PERFORMANCE. (SOURCE: APPENDIX D-7, COLARCO)

Atmospheric Composition relies on GMAO to develop Observing System Simulation Experiments (OSSE), which are detailed simulation of global observing systems (Figure 17). OSSEs are critical for pre-launch algorithm development and assessment as well as for conducting quantitative science trade studies during mission formulation. An OSSE is started with a high-resolution nature run (free running atmospheric model using only sea surface temperatures as input) of a comprehensive Earth system model, which is then sampled at the instrument footprint using an instrument simulator—all of which are computationally demanding. HEC resources are essential for realistic observing system simulations. Without these capabilities, one would need to resort to a limited number of cases and conditions, or to rely on low-resolution, low-fidelity simulations.

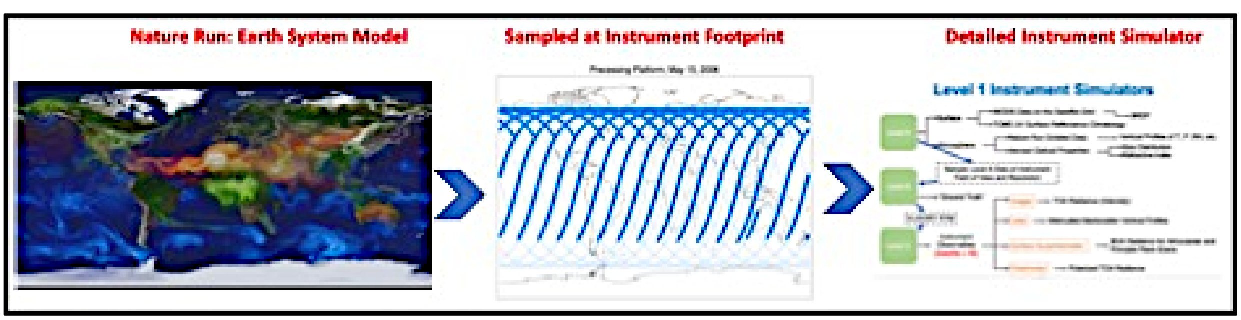

FIGURE 17: OSSE USE CASE CONCEPT. (SOURCE: APPENDIX D-7, DA SILVA)

HEC resources are also used to perform science data processing for the Aerosol Robotic Network (AERONET), performing inversions of ground-based remote sensing measurements to obtain atmospheric aerosol characteristics. AERONET has hundreds of sites worldwide, with measurements made hourly. Specifically, HEC is used for near real-time (NRT) processing of instrument output (1–3 times per day) and re-processing of data in retrospective (generally quarterly). Data transfer and scheduling are orchestrated between the AERONET server and the HEC resources. Availability of HEC resources is of concern in order to meet commitments for NRT data products, particularly during field campaigns. No significant growth in demand is expected for AERONET. If code migration is required, assistance and advice may be needed.

Another way in which HEC resources support Atmospheric Composition is by providing one stage of the science data processing for GeoCarb, which measures carbon constituents, such as carbon dioxide (CO_2) and methane (CH_4) HEC resources are allocated to provide the daily Level 2 (L2) Full Physics data products from 6% of the L1B products produced externally and used in instrument health. The L2 data are forwarded along to the next external step in the workflow. HEC resources are also used to process 100% of the L1B products with a one-week delay with output available within 60 days. HEC resources were specifically offered as an alternative to NASA funding the construction of a new data processing center at Colorado State University, but HEC received no additional funding for this allocation.

Biomass burning is a key contributor to issues in Atmospheric Composition in terms of air quality, and more broadly impacts respiratory and cardiovascular illness, the Earth's radiation budget, and nutrient availability. Important safety decisions are made about personnel and containment that would benefit from better quality near-realtime information. Machine learning techniques are being used to detect smoke plumes using geostationary satellite data (e.g., from GOES-R, a Geostationary Operational Environmental Satellite), providing high spatial and temporal resolution over large domains. Pixel-wise inference of full disc GOES-R is time consuming, but there is a potential to detect smoke pixels every ten minutes (or hourly detection) with GOES-R, and the method can be extended to other geostationary satellites (i.e., Himawari). Since this approach is highly parallelizable it can be run on HEC assets. GOES-R data is available freely on Amazon Web Services (AWS) (and may also be available for in-time processing at HEC). This approach can serve as a pipeline for other phenomena detections with timely visualization within NASA's Worldview tool.

Project DYAMOND (Atmospheric general circulation Modeled On Non-hydrostatic Domains) used NCCS resources to create a 3-km GEOS simulation of global convection for August 2016 using an ensemble of simulations in conjunction with global collaborations. A single instance of the final simulation required 40,000 cores at the NCCS for two weeks and produced 600 TB of data. However, working up to the final iteration required many runs to resolve performance issues, improve the capability to extract meaningful science from it, and assess the consequences of increasing complexity. The process was slowed by the availability of sufficient resources in balance with other GMAO needs.

Both HECC and NCCS support the evolution of a coupled atmosphere-ocean model, as described in **Error! Reference source not found.**, as part of the development of global simulations. The computing facilities provide large volumes of storage and post-processing visualization support and, through analysis, a catalog of events, such as tropical and extra-tropical storms, mesoscale convective systems, and other extreme phenomena. NCCS resources are needed to support this process, including the exploration of concepts like domain-specific languages and machine learning. A new workflow is being examined to reduce the processing load, which involves multiple stages with restarts and on-the-fly observation operators to reduce post-processing workloads. Additional HECC and NCCS consultation will be needed to evaluate the use of alternative architectures, such as GPUs.

TABLE 6-8: EVOLUTION OF A COUPLED ATMOSPHERE-OCEAN MODEL.

2020 Capability		by 2025		by 2030	
Resolution	***Cores/Data***	***Resolution***	***Cores/Data***	***Resolution***	***Cores/Data***
Atmosphere 6 km *137 Levels*	**O(10,000)** **500 TB** per simulation	**Atmosphere** 3 km *181 Levels*	**O(100,000)** **>5 PB** per simulation	**Atmosphere** 1.5 km *>200 Levels*	**O(1 million)** **>100 PB** per simulation
Ocean *4 km* *90 Levels*		***Ocean*** *2 km* *100 Levels*		***Ocean*** *1 km* *150 Levels*	

GMAO's GEOS composition forecasting (GEOS-CF) model produces global analyses and forecasts of atmospheric composition using NASA's space observations on constituent analysis and forecasting. Applications include global health and air quality assessments and support of NASA's in-situ field missions and basic research in transport, emissions, and composition. The production system is run on a daily cycle, generating four analyses and one forecast each day. Usage is almost continuous, with prior days available for R&D studies. Analysis of the data is facilitated by machine learning techniques using the GPU cluster with a rapidly growing application of this technology. Work relies heavily on HEC staff to assist in leveraging system capability and performance improvement. These large, complex models generate large node-count jobs. GEOS-CF is expected to be linked to the GEOS forward processing (GEOS-FP) model (Table 6-9, p. 66).

The Weather and Atmospheric Dynamics Focus Area also uses GMAO to perform OSSEs in support of decisions regarding mission selection for new instruments, as described above. A global high-resolution model with 4D ensemble variation data assimilation will be run for multiple 2–3-month periods for each proposed instrument. The frequency of experiments depends on demand from the instrument teams. High-fidelity output is created for process studies and to generate training data for machine learning tool development. For example, the GOES Program has funded the GMAO to use the OSSE system to help establish a trade space in quantifying the impact of future hyperspectral infrared (IR) sounders in geostationary orbit. A similar study was performed to understand the value of alternative approaches to measuring winds from space. The current workflow is linear and segmented into steps, as described in **Error! Reference source not found.** New nature runs would require 40,000 cores, with each perturbation of the configuration requiring 5,000 cores. This will improve the fidelity of the model reflecting complexities that are not handled today. It also will create the opportunity for a more elegant, single validation for each nature run.

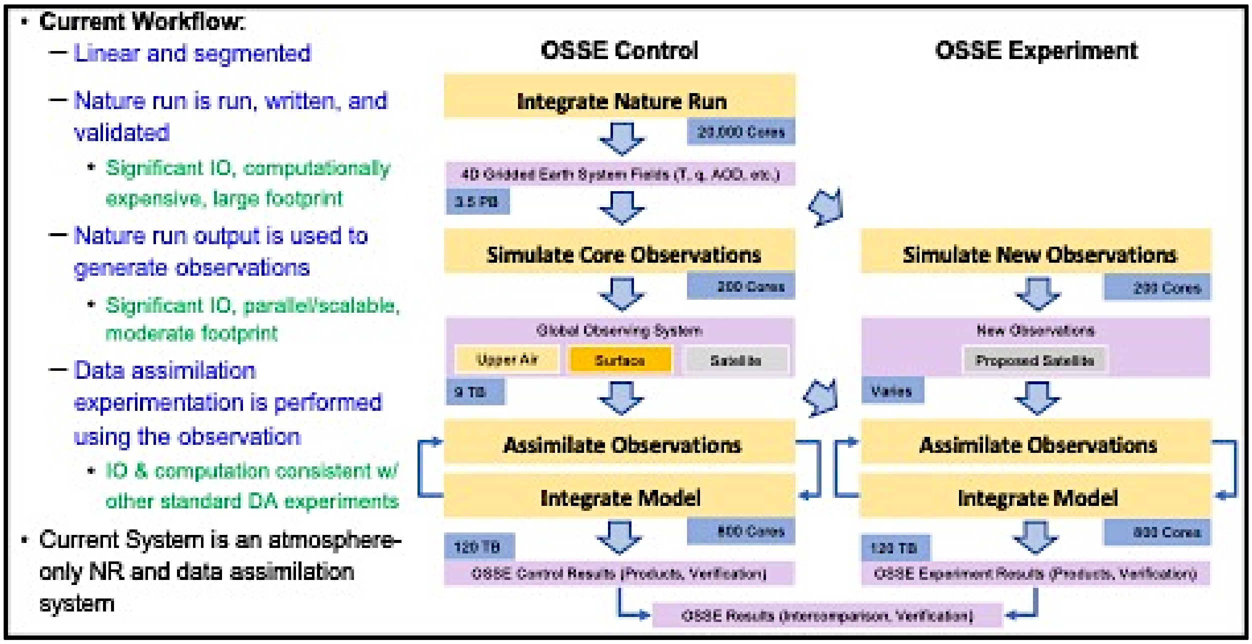

FIGURE 18: CURRENT OSSE WORKFLOW. (SOURCE: APPENDIX D-7, PRIVÉ)

The GEOS forward processing (GEOS-FP) product is a high-resolution, near-real-time, advanced data assimilation weather analysis & prediction system that supports NASA Instrument teams, field campaigns and scientists to plan, develop, and assess the contribution from new and under-utilized Earth Observing System missions for weather and climate. GEOS-FP supports a large number of both NASA and external customers, including the NASA-NOAA Joint Center for Simulation and Data Assimilation (JCSDA). It requires thousands of cores of HEC resources for each run, which are performed four times each day, each requiring 1–4 hours. While a copy of GEOS-FP is running, the next generation version, GEOS-FPP, is also run for testing. Availability of disk space has limited GEOS-FP from providing customers with full-resolution output (GEOS-FP runs at 12.5 km, but output is written at 24 km). Dedicated disk space is needed to avoid contention with allocations and performance. Within 18 months, the model is expected to double the number of vertical levels and increase the assimilated all-sky radiances with a substantial growth in the number of cores required. Future workloads are described in Table 6-9, but this does not address the growth in both the ensemble members and resolution for the data assimilation, nor the ensemble model prediction capability that is likely is needed for the future. In other words, the future is probably 32 to 64 runs of the last column in the table below. (Appendix D6-19.3, Todling)

TABLE 6-9: FUTURE HEC FOOTPRINT FOR GEOS-FP. (SOURCE: APPENDIX D6-3, TODLING)

	Present	Future
Inputs	368 GB	~ 2.4 TB
Outputs	559 GB	~ 1.4 TB
Products	102 GB	~ 1 TB
Cores	7,688	~ 45,000
Throughput	2 days/day	~ 2 days/day

In the Weather and Atmospheric Dynamics Focus Area, HEC resources are used to further develop and to forecast with the NASA Unified Weather Research Forecast (NU-WRF) model, which is used to feed many of the other models. HEC capabilities are used to analyze ceilometer output to develop accurate estimates of the Planetary Boundary Layer (PBL). The current limited experiments will be broadened in extent and expanded to include space-based lidar instruments to detect the upper extent of the PBL.

In the Climate Variability and Change Focus Area, NCCS resources are used to support the GISS Model E, both for its development and its use in analyzing climate change. One recent study was to understand the Sun's effect on climate over long timescales (centuries). To investigate complex, remaining questions about how radiation from the Sun affects Earth's climate, Duke University and NASA GISS scientists ran multiple century-long simulations under a variety of solar conditions at NCCS. After a 100-year control run to bring the ocean to equilibrium, the team ran two sets of simulations on the NCCS Discover supercomputer. One entailed three simulations covering 110 years (10 solar cycles) and the other ran four simulations covering 160 years. The computations revealed that long-term changes in the Sun's output cause clear effects on Earth's climate that can be seen from the surface to the upper stratosphere. The study also showed that the tropics and high latitudes respond to solar variability at different wavelengths, especially during winter. In contrast, the 11-year solar cycle has clear impacts in the stratosphere but relatively weak effects on surface climate that are similar in magnitude to natural variability. Also, in the Climate Variability and Change Focus Area, HEC supports the development and use of the Ice-Sheet and Sea-Level System Model (ISSM), which estimates the change in sea level resulting using ice sheets, glaciers, and ocean thermal contributions. The model is a finite element methods-based, fully unstructured mesh (e.g., no grids) consisting of computations using the sea-level solver Glacial Isostatic Adjustment (GIA) and an ice-sheet dynamics solver. It performs uncertainty quantification by running the solvers many times using the Parallel Sandia National Labs Dakota samplers. A typical experiment runs for five days on 1,000 CPUs, requires highly interconnected nodes, and generates at least 50 TB of data per run.

HEC resources support the development and operations of the MERRA-3 (Modern-Era Retrospective analysis for Research and Applications) analysis and reanalysis system, which is described in Figure 19. MERRA-3 has two objectives:

- Develop a coupled Earth system data assimilation capability to explore the combined value of NASA observations in air, land, ocean, and ice.
- Produce comprehensive Earth system reanalysis to place these observations in a climate context, enabling a broad range of research and applications.

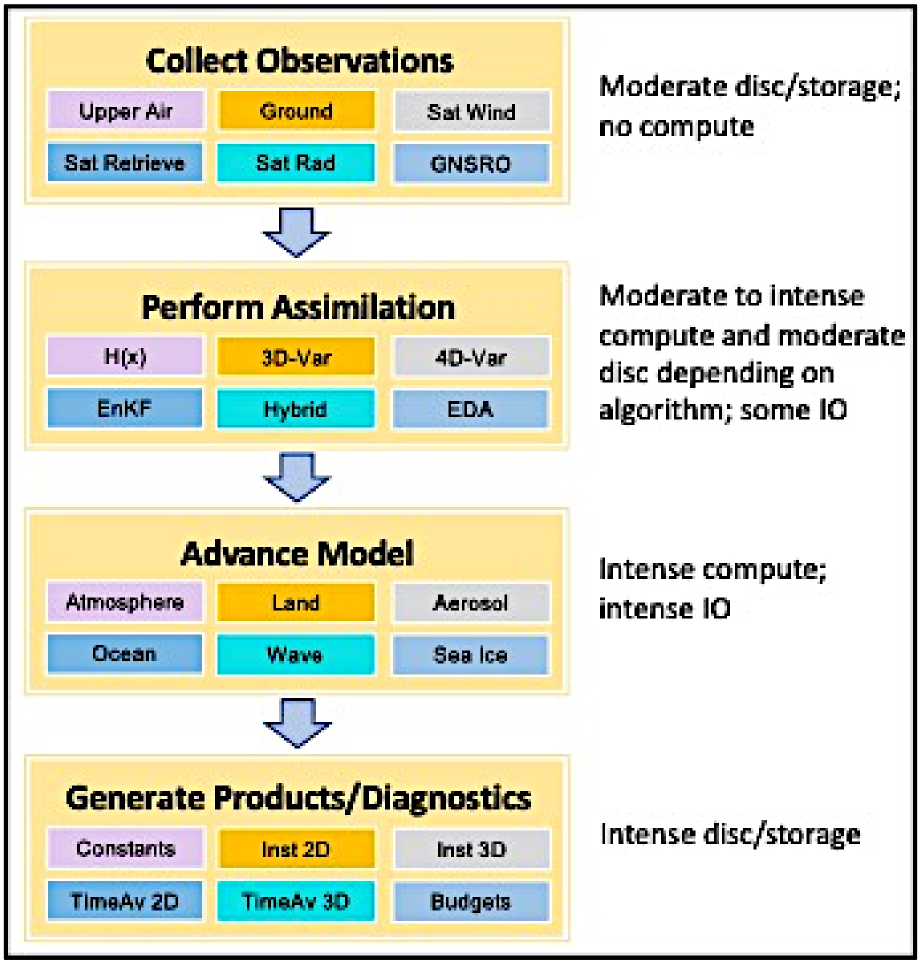

FIGURE 19: COUPLED ANALYSIS/REANALYSIS: MERRA-3 AND BEYOND. (SOURCE: APPENDIX D-7, GELARO)

HEC resources are essential for reanalysis production and prerequisite testing and are needed on a near-continuous basis. Without these capabilities it would be impossible to meet the GMAO deliverable of a realistic GEOS-based coupled reanalysis. Commercial cloud resources could be used as a supplement, but cost would likely be prohibitive for long-duration production. A reanalysis runs in round-the-clock, continuous data assimilation cycles, and typically in multiple streams simultaneously. Because these streams run behind real time, there is no need (and considerable detriment) to stop the cycle. The output represents a synthesis of NASA modeling and observations required to demonstrate and explore the value of these assets for understanding the Earth system and advancing national capabilities in Earth-system analysis and prediction. Future growth is expected to re-organize the workflow so that interim I/O is reduced but memory requirements are expanded, and over 10,000 cores will be required. Much larger output volumes are expected. Help from supercomputing experts at HEC facilities will be needed for any re-hosting imposed.

GMAO's development of subseasonal-to-seasonal-to-decadal prediction systems (S2S2D) is designed to demonstrate the beneficial impact on forecast predictability and prediction skill of adding new interactive Earth system components in the model, increased resolution and forecast ensemble size, and forecast initialization from a modern assimilation that incorporates new ocean data types. The new version of GMAO's seasonal forecast system includes higher ocean resolution and a larger ensemble size. The advantages of increased ocean resolution include more realistic bathymetry and surface currents. GMAO's new ensemble strategy takes maximum advantage of computing resources and provides more accurate predictability estimates and increased forecast skill. In order to provide this level of forecast and analysis, a complicated series of steps must be performed. Workflow begins with coupled assimilation, as all input data are available. Initial conditions for forecasts may need to be transferred between HECC and NCCS, ensemble perturbations are calculated, and forecasts are run. Then, input data for ocean analysis are downloaded, pre-processing of input data occurs, and assimilation is run. Ensemble perturbations are then calculated, and forecasts are run. Latency is up to a few days behind real time, as reanalysis depends on arrival of data and on multiple near-real-time procedures. Although scripts automate much of the process in both phases, there is frequently need for interaction by staff.

HEC resources will be used to conduct a retrospective suite of sub/seasonal forecasts at high resolution, spanning the period 1981–present, consisting of an ensemble of 40 sub/seasonal forecasts per month. The forecasts will also continue into near-real time. A global weakly-coupled reanalysis will also be run and used to initialize the forecasts. Without HEC resources, GMAO would need to resort to a limited number of ensemble members, rely on low-resolution forecasts, or lose near-real time capabilities. Each of these choices would substantially affect the success of S2S activities by making estimates of predictability difficult, degrading prediction skill, and delaying system upgrades. Output files need to be accessed by external collaborators and other users, such as participants in multi-model activities to which GMAO contributes. Because of the massive dataset size, there will be a need for partners to execute algorithms on the same nodes where data resides. The retrospective phase requires 180 wall clock hours on 122 nodes for each month of simulation; 8–10 forecasts are running at the same time. The near real time phase requires 122 nodes for each of 8–10 concurrent forecasts for 72 hours for each month, plus 100 nodes for 24 hours each month for post-processing. Queue structure can severely impede the progression of the experiment, as can file system stability and throughput. Increased horizontal and vertical resolution in the atmosphere and ocean is expected. In addition, the expansion of the model is expected to include more interactive components of the Earth system, such as interactive vegetation and ice sheet models. Increased use of new data types, such as sea ice thickness, during assimilation is also expected.

HECC supports the growing understanding of the ocean's role in the Climate Variability and Change through the Estimating the Circulation and Climate of the Ocean (ECCO) project, including computation, storage, and visualization resources. ECCO researchers run simulations on HEC resources to produce global "maps" of Earth's ocean and sea-ice system at an unprecedented resolution (~1 km horizontal grid). Simulations are produced with the Massachusetts Institute of Technology general circulation model (MITgcm) on up to 70,000 cores on the Pleiades supercomputer, and are compared to observational data from NASA satellites, ocean sensors, and ship-borne and mooring data. Researchers can use this modeling tool to:

- Investigate fundamental questions such as how the circulation, chemistry, and biology of the ocean collectively interact with atmospheric carbon.

- Determine how a pollutant plume or debris field might spread from a particular ocean location, or where and when heat is absorbed by or released from the ocean.

As HECC storage and analytic capabilities continue to increase, this work can potentially develop into a transformative strategy for understanding and predicting the impact of global ocean circulation on climate.

Another chemistry-climate modeling effort supported by HEC resources is the Chemistry-Climate Model (CCM). The project develops and integrates atmospheric chemistry and aerosol components within the GEOS model that:

- Inform international assessments (e.g., the Montreal Protocol).
- Support the utility of NASA satellite measurements (e.g., scientific analysis, *a priori*).
- Advance research on coupled Earth system response to atmospheric chemistry on seasonal-to-decadal timescales.

The CCM project recognizes a need to move toward higher resolution full stratospheric and tropospheric chemistry simulations to support recent satellite missions, such as NASA's Tropospheric Monitoring Instrument (TROPOMI) and the Geostationary Environment Monitoring Spectrometer (GEMS), and upcoming ones, such as NASA's Tropospheric Emissions: Monitoring of Pollution (TEMPO). HEC resources are essential for realistic chemistry simulations. Without these capabilities, the GMAO would only be able to perform a limited number of cases and conditions, or to run only at low resolution, degrading the effectiveness of the activity. The current model runs one simulation year in two wall clock days using over 1,000 cores. Future runs are expected to need 5,000 cores. Output is shared with external collaborators. Because of the massive size of datasets, there are benefits to allowing partners to perform some analysis on the same nodes where data resides. Options to share output by web services and OPeNDAP (Open-source Project for a Network Data Access Protocol) are necessary. Contributions to assessments require multiple ensemble members of ~150 simulated year duration; more HEC resources would allow sufficient throughput for higher spatial resolution simulations. Longer integrations (e.g., assessments) require saving and archiving monthly global gridded output; data volume can be high depending on number variables saved, prescribed by the assessment protocol. The main constraint on advancing this model is the scalability of compute environment; higher resolution and higher complexity simulations are needed in the future but require ever more compute nodes and storage volume. Inability to get timely access to the required number of nodes results in a sacrifice of model resolution and degrades the quality of the modeling products; lack of storage space for data archive (current and historical) weakens scientific utility of CCM output.

In the Water and Energy Cycle Focus Area, HEC resources are being used to construct models and to help validate them. HEC supports the Land Information System (LIS) model, which relies on the capabilities of supercomputing clusters to handle the non-linear, non-equilibrium dynamics of the water cycle. LIS is a software framework for high-performance terrestrial hydrology modeling and data assimilation, developed with the goal of integrating satellite- and ground-based observational data products and advanced modeling techniques to produce optimal fields of land surface states and fluxes.

LIS supports a breadth of current and planned NASA research interests and priorities, such as:

- Development of OSSEs in support of future missions such as Snow, MC, SBG, and NOS.
- High-resolution (1-km) land reanalysis over the Western U.S. (WWAO), HMA (HMA).

- Fully coupled modeling and data assimilation systems (MAP) integrating LIS, WRF, WRFhydro, and ParFlow.
- Multi-variate data assimilation for hydrologic science and applications (GRACE-FO, SWOT, NISAR).
- Advanced machine learning techniques for data fusion, information extraction.

The LIS team also runs analyses to evaluate the stability of the model and the quality of the output. HEC assets are essential to both the continued development of capabilities and the operational forecasting applications of the model for decision makers. Access to a large volume of input data is also required. All of these require special HEC capabilities such as high-performance file systems and networking fabrics, large storage capacities, and the availability of high node-count processing environments. LIS is gaining wider acceptance by both decision makers and scientists, with a growing demand for new users to perform their own forecasts. Data storage capacity, turnaround time, and external access are the biggest limits that LIS experiences. In the future, as models and data move towards higher spatial/temporal resolution, more computational power will be needed. In turn, greater storage space (either local or virtual) will be required for both remote sensing data and model output. For example, the significant large volumes expected from future sensors (i.e., NISAR) will necessitate virtual data storage solutions and modifications to the code to support non-local data access. These limitations affect LIS by negatively impacting schedules and imposing additional project management. Slow turnaround times and difficulty in getting external access makes it difficult to review results and to track mission success. Limited data storage capacity forces scientists to perform mundane tasks of shuffling data just to keep projects moving forward. Lack of computational power and storage space will limit the use of future, higher-resolution remote sensing data and novel developments in model physics.

In the Carbon Cycle Focus Area, NASA's Biodiversity Program, much of the research is through competitively selected grants, many of which are awarded to universities and non-NASA organizations. As a result, many elements of the program use non-NASA computing for modeling, as the security model for accessing the computing resources is too restrictive.

In another element of the Carbon Cycle Focus Area, the University of Virginia Forest Model (UVAFME) is used to estimate biomass starting from climate change scenarios and characterizes the changes in the forest composition down to the individual tree level, as depicted in Figure 20. The application of this third-party model enhances the understanding of the evolution of arctic boreal region under NASA's decade-long Arctic Boreal Vulnerability Experiment (ABoVE) on HEC resources and could not have been run at this scale without the use of HEC.

The Orbiting Carbon Observatory-2 (OCO-2) mission has also used HEC resources for their science data processing workflow. OCO-2 is a satellite with a three-channel imaging grating spectrometer that returns high-resolution spectra of reflected sunlight in molecular oxygen (O_2) A-band and carbon dioxide (CO_2) bands. It is being used to study carbon dioxide distribution in the atmosphere and to detect emission hotspots and volcanoes since its nominal operation in 2014. Orbiting Carbon Observatory-3 (OCO-3), the immediate successor to OCO-2, was installed on the International Space Station in May, 2019. It has been in nominal operations phase since August, 2019. The algorithms for extracting CO_2 values from the observations are computationally complex, requiring about five minutes to process a single sounding. This limits the number of processed soundings to about 6% of the one million each day that can be translated into L2 data products during the forward processing. The use of HEC resources enables the operational data systems to reprocess all the cloud-free scenes on a monthly basis with the most current calibration coefficients, yielding more extensive and higher quality L2 products on a production schedule. In addition, a comprehensive re-processing campaign for the entire lifetime (2014–2020) was enabled by the use of HEC assets. Without these resources, the cost of replacement assets (either cloud or on-premises) would preclude this monthly re-processing limiting analysis to lower quality initial processing inputs or limited special studies. (Appendix D8.7, Chu)

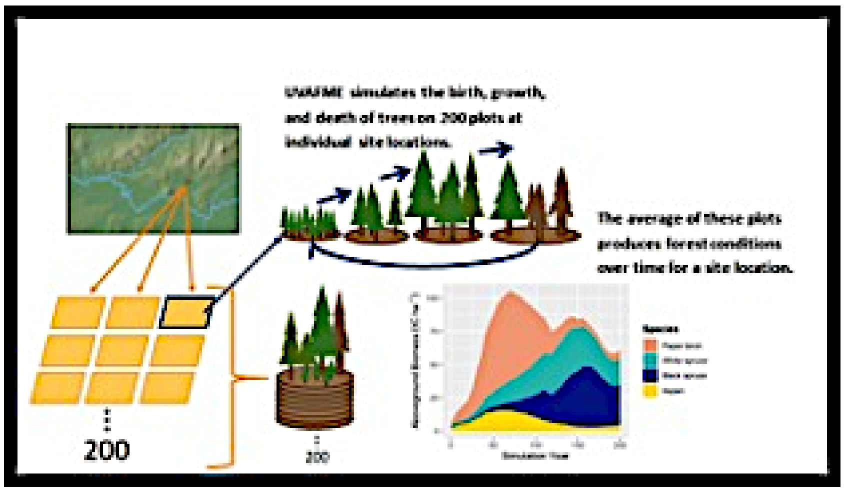

FIGURE 20: SIMULATIONS OF THE BIRTH, GROWTH, AND DEATH OF TREES PRODUCES FOREST CHANGES OVER TIME FOR A SITE LOCATION. (CREDIT: A. FOSTER, UNIVERSITY OF VIRGINIA) (SOURCE: APPENDIX D-9, HEC MSR INPUTS4)

Another example of the use of HEC resources in support of the Carbon Cycle Focus Area is the Harmonized Landsat Sentinel-2 Program, which is producing a harmonized surface reflectance product. Landsat and Sentinel-2 data represent the most widely accessible moderate-to-high spatial resolution multispectral satellite measurement. Following the launch of the two Sentinel-2 satellites in 2015 and 2017, the potential for synergistic use of Landsat and Sentinel-2 data creates unprecedented opportunities for timely and accurate observation of Earth status and dynamics. Harmonization of the Landsat and Sentinel-2 data is of paramount importance for the scientific community. This research project prototypes the harmonization for the entire North America and other globally distributed test sites. Processing includes both forward processing, as the observations are made available, and re-processing, which provides a uniformly calibrated data product based on the accumulation of results over a year of observations. (Appendix D6-9.5. Ramachandran)

Started as a land-use land-change contribution to the Carbon Cycle Focus Area, the NASA Earth Exchange (NEX) enables research and applications with regional, continental, and global Earth observations. NEX has introduced many advanced techniques, including early machine learning tools, into the Earth science community. NEX is an analytic center which retrieves data and re-organizes it into an analysis-ready format. Data is acquired from any relevant source (both NASA and non-NASA) for analysis, and the NEX platform provides the collaborating scientists with access to usable data, tools, support, and computational resources under an Authorization to Operate (ATO) to conduct the research. NEX supports between 20 and 30 research teams, and uses both HECC and commercial cloud computing resources, depending on the type of workflow and project objectives. Commercial cloud costs are difficult to forecast. Currently, CPUs are used by recruiting thousands of cores in an embarrassingly parallel fashion (process per granule or granule set). The special HECC capabilities that are critical for this specific use case are high-performance file systems with fast interconnects, networking (in and out of NAS), and storage management. NEX measures the impact of HECC resources using three metrics: 1) is it possible or impossible, 2) costs of accomplishing the work using HECC resources versus those using a commercial cloud, and 3) the traditional method of measuring impact—publications and citations). Storage is the main limit; user education and limited messaging services are also constraints. Without addressing these challenges, the number of NEX users is likely to remain constant. (Appendix D9.19, Appendix D6.3-3)

In support of the Earth's Surface and Interior (ESI) Focus Area, the study of earthquakes requires the use of complex models to reflect the understanding of ongoing crustal deformations, particularly in the Southern California area. QuakeSim, an application started in 2001, was developed to integrate observational data, finite element analysis, and analytic tools into an earthquake forecast system. It has evolved into QUAKES-A, whose goal is to reconcile Global Navigation Satellite System (GNSS) and synthetic aperture radar (SAR) data in a mathematical and geospatial framework to produce a gridded crustal deformation reference model. Together they harmonize the data, the analytic tools, and the computing environment to provide researchers and disaster response decision makers a mechanism for evaluating surface deformation. The QuakeSim system and QUAKES-A analytic center framework use Indiana University's GeoGateway, which relies on NSF's XSEDE Jetstream supercomputing environment. Figure 21 shows the operational concept for this multi-source, data intensive computing system. (Appendix D8.24, Donnellan)

Another example of HEC support to the ESI Focus Area is in the use of NISAR (launching in 2022) to measure plate boundary deformation using SAR output. HEC provides the computational platform for generating L2 deformation maps from Sentinel 1 SAR L1 data, and could be used to run analysis code to process the existing archive of Sentinel 1 data. Subsequently, processing would then be needed to keep up with new data. More specifically, HEC processing jobs are initiated from on-demand requests for processing over an area of interest (AOI) input from the Aria science data system in AWS. Aria's workflow processing dispatches across both AWS and HEC, but large SAR compute jobs are assigned to run on HEC's Pleiades supercomputer. (Appendix D6-9.7, Owen)

This workflow is ideal for bulk processing, but less ideal for lower latency real-time processing. The resulting L2 displacement maps, distributed by the NAS Data Portal, are used in time series analysis to understand plate boundary deformation, a key science objective of ESI. Ease of access by outside researchers is more important than the negligible confidentiality considerations. This type of high-throughput disk I/O would be improved, and system wear reduced, by attaching SSD on the nodes to avoid high demand on the Lustre file systems. It is expected that this type of analysis will grow after NISAR launches—and so will the number of users wanting to run on-demand analysis. (Appendix D6-9.7, Owen)

As model improvements become available and model resolutions increase, their applicability across science broadens and more researchers require forecast outputs and specialized model runs, all of which will significantly increase the demand on HEC resources. (Appendix D6-9.7, Owen)

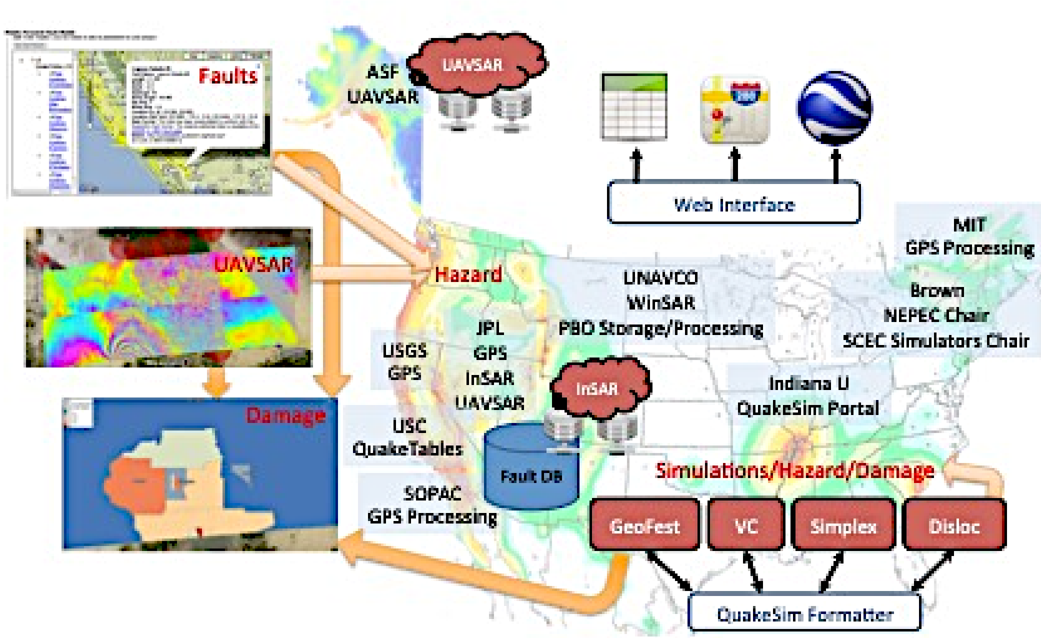

FIGURE 21: QUAKESIM INTEGRATES MULTIPLE DATA SOURCES INTO A CRUSTAL DYNAMICS ANALYTIC CENTER. (SOURCE: APPENDIX D8.24, DONNELLAN)

All of the focus area models can also be used to identify observations needed for forecast quality improvements. Such requirements can be used as input into the process of tasking Earth observing satellites, described conceptually as the New Observing Strategy (NOS). This is particularly useful in prioritizing spacecraft operations when demand is high; some observations are more valuable than others. A 2019 SMD autonomy workshop developed design reference missions using constellations of satellites operated autonomously and driven by the needs of various models for input to improve forecast quality. In addition, such a concept supports the tasking of satellites to support various science studies, prioritizing the collection of data in areas when and where the phenomenon of interest is occurring. In both cases, model execution and analysis of output place new demands on the HEC platforms to supply computational capacity. (Appendix D6-9.2, Le Moigne)

6.2 Astrophysics

Astrophysics uses HEC resources primarily for data analysis and modeling. Table 6-10 and Table 6-11 list the Astrophysics usage data for the 226 projects supported by both HECC and NCCS resources, respectively. The HEC Program uses Standard Billing Units (SBU), described in Appendix A as the metric by which allocations are made and consumption is measured. The top ten users of HEC resources in the Astrophysics Division are identified in Table 6-12.

Table 6-10: Astrophysics HECC Usage (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	11,557,732	10,080,198	9,627941	11,384,978	6,298,592	1,679,443

Table 6-11: Astrophysics NCCS Usage (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	688,125	58,615	96,398	408,804	40,870	N/A

Table 6-12: Top Ten Astrophysics Users of HEC SBUs in 2020.

PI	Organization	Project Title	FY20 SBUs Consumed
Tumlinson, Jason	Space Telescope Science Institute	Resolving the Milky Way and Nearby Galaxies with WFIRST	1,022,696
Jiang, Yanfei	Unknown company or sponsor	Structures and Winds of Massive Stars from Three Dimensional Global Radiation Magnetohydrodynamic Simulations	839,273
Peeples, Molly	Space Telescope Science Institute	Figuring Out Gas & Galaxies in Enzo: The Gas-Galaxy Connection at z>2	822,242
Jiang, Yanfei	Unknown company or sponsor	Radiation Magnetohydrodynamic Simulations of Accretion Disks	803,980
Cen, Renyue	Princeton University	BUFFALO	649,503
Toomre, Juri	University of Colorado, Boulder	Tiny Stars, Strong Fields: Exploring the Origin of Intense Magnetism in M Stars	576,424
Lazzati, Davide	Oregon State University	Gamma-Ray Bursts as a window into extreme physics	504,955
Hopkins, Philip Fajardo	California Institute of Technology	The Milky Way: A Billion Particles on FIRE	435,697
Brooks, Alyson Michelle	Rutgers University	Small Statistics No More: a suite of simulated dwarf galaxies to interpret observations	399,966

PI	Organization	Project Title	FY20 SBUs Consumed
Ruiz, Milton	University of Illinois, Urbana-Champaign	Studies in Theoretical Astrophysics and General Relativity	310,060

HEC resources support Astrophysics in performing the science data processing of the Transiting Exoplanet Survey Satellite (TESS) instrument output using Science Processing Operations Center (SPOC). The SPOC was originally developed for the Kepler science pipeline, which also ran on HEC resources. TESS downlinks about 1 TB of data every two weeks. The pipeline outputs calibrated pixels and light curves on the Pleiades supercomputer. The SPOC will also search for periodic transit events and generate validation products for the transit-like features in the light curves. All TESS SPOC data products are archived to the Mikulski Archive for Space Telescopes (MAST). The availability of HEC resources at no cost is essential to this mission, which is cost-capped. The extended TESS mission plans an additional data type and an increase in the number of images each month. (Append D6-3.4, Tenenbaum, Jenkins)

HEC resources are used also to support the study of nonintuitive high-dimensionality phenomena. Creating light curve data from TESS on NCCS CPUs, Astrophysics researchers then use deep learning methods on NCCS GPU resources to create representations of high-dimensionality embedding spaces to identify similar features among the millions in the dataset (Figure 22). The impact is the discovery of 50 planet candidates, over 200 heartbeat stars, more than 10 potential triple star systems, more than 20 potential quadruple star systems, and one possible sextuple star system. (Append D8.2, Powell)

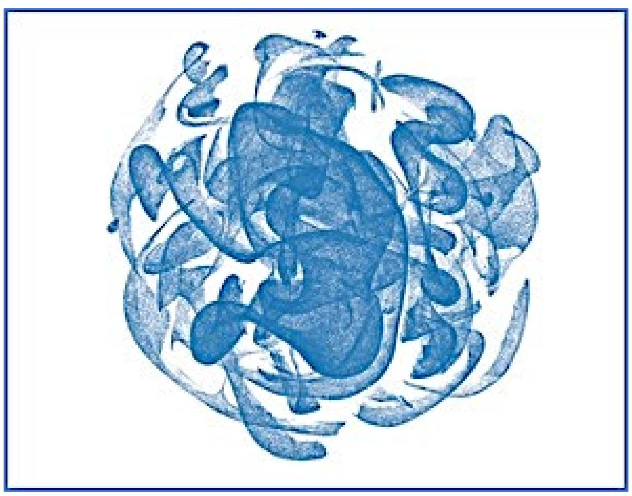

FIGURE 22. A TWO-DIMENSIONAL PROJECTION OF THE HIGH-DIMENSIONAL SPACE OF TESS LIGHT CURVE REPRESENTATIONS. (CREDIT: BRIAN P. POWELL, NASA GODDARD SPACE FLIGHT CENTER).

HEC resources are also used to support the study of super-Eddington accretion flows onto black holes using a global 3D radiation magnetohydrodynamics (MHD) simulation. The project solves the time-dependent radiative transfer equation for the specific intensities to accurately calculate the angular distribution of the emitted radiation. Turbulence generated by the magneto-rotational instability provides self-consistent angular momentum transfer. These simulations were able to better account for vertical advection and offer implications for the growth of supermassive black holes in the early universe. The simulation results also provided a basis for explaining the spectrum and population statistics of ultraluminous X-ray sources. (Appendix D8.3, Jiang)

HECC resources are also used in simulating the cosmic fog around galaxies. Researchers from the Space Telescope Science Institute and Johns Hopkins University are running cosmological simulations to model how galaxies and gas change through time. Using the Enzo cosmological hydrodynamic code, the "Figuring Out Gas & Galaxies In Enzo (FOGGIE)" project scientists model the co-evolution of galaxies and their gas with a focus on resolving the ultra-diffuse circumgalactic medium (CGM) with unprecedented fidelity. The simulations reveal a richly structured CGM full of churning turbulent gas, small clouds, and tenuous hot gas. Results are used to help interpret real observations made by NASA's Hubble Space Telescope and other observatories. Hubble observations show that low-ionization gas, which should be relatively cool and have higher density, often has kinematic structure very similar to more highly ionized gas, which is expected to generically be hotter and lower density. With more data and a range of investigations to run, the demand for capacity is expected to grow substantially. (Appendix D9.51, Peeples)

6.3 Heliophysics

As described on the Heliophysics Division website[1],

> *The Science Mission Directorate Heliophysics Division studies the nature of the Sun, and how it influences the very nature of space — and, in turn, the atmospheres of planets and the technology that exists there. Space is not, as is often believed, completely empty; instead, we live in the extended atmosphere of an active star. Our Sun sends out a steady outpouring of particles and energy – the solar wind – as well as a constantly writhing magnetic system. This extensive, dynamic solar atmosphere surrounds the Sun, Earth, the planets, and extends far out into the solar system.*
>
> *Studying this system not only helps us understand fundamental information about how the universe works, but also helps protect our technology and astronauts in space. NASA seeks knowledge of near-Earth space, because – when extreme – space weather can interfere with our communications, satellites and power grids. The study of the Sun and space can also teach us more about how stars contribute to the habitability of planets throughout the universe.*
>
> *Mapping out this interconnected system requires a holistic study of the Sun's influence on space, Earth and other planets. NASA has a fleet of spacecraft strategically placed throughout our heliosphere – from Parker Solar Probe at the Sun observing the very start of the solar wind, to satellites around Earth, to the farthest human-made object, Voyager, which is sending back observations on interstellar space. Each mission is positioned at a critical, well-thought out vantage point to observe and understand the flow of energy and particles throughout the solar system – all helping us untangle the effects of the star we live with.*

[1] Downloaded December 2020.

Researchers supporting the Heliophysics Division utilize HEC resources to analyze large volumes of data and further expand and operate forecast models Table 6-13 and Table 6-14 show the actual usage for Heliophysics programs for the past five years in on HECC resources and four on NCCS resources respectively. The HEC Program uses Standard Billing Units (SBU), described in Appendix A, p. 83, as the metric by which allocations are made and consumption is measured. The top ten users of HEC resources in Heliophysics Division are identified in Table 6-15.

TABLE 6-13: HELIOPHYSICS HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	14,479,093	6,825,219	7,697,471	5,944,176	3,268,446	4,870,080

TABLE 6-14: HELIOPHYSICS NCCS USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	207,927	300,961	315,636	320,234	272,101	N/A

TABLE 6-15: TOP TEN HELIOPHYSICS USERS OF HEC SBUS IN 2020.

PI	Organization	Project Title	FY20 SBUs Consumed
Dong, Chuanfei	Princeton University	Integration of Extended MHD and Kinetic Effects in Global Magnetosphere Models	4,533,436
Featherstone, Nicholas Andrew	University of Colorado, Boulder	The Solar Dynamo Revealed	1,214,859
Brown, Benjamin P.	University of Colorado, Boulder	Fundamental Stellar Dynamo Physics and Touchstone Stars	799,396
Chen, Li-Jen	NASA/GSFC	PIC simulations to support the Magnetospheric Multiscale mission	705,127
De Pontieu, Bart	Lockheed Martin	Interface Region Imaging Spectrograph (IRIS) Small Explorer	619,322
Zhao, Junwei	Stanford University	Frequency-Dependent Helioseismic Analysis on Solar Meridional Flow, Center-to-Limb Effect, and Sunspots	436,717
Berchem, Jean	University of California, Los Angeles	Plasma acceleration and energization at Earth’s magnetosphere	408,011

PI	Organization	Project Title	FY20 SBUs Consumed
Pogorelov, Nikolai	University of Alabama, Huntsville	Pickup Ions in the Outer Heliosphere and Beyond	392,819
Opher, Merav	Boston University	Heliosheath Flows with a Tilted Solar Magnetic Field	280,717
Kitiashvili, Irina Nikolaevna	NASA/ARC	Interaction of Quiet-Sun Magnetic Fields with the Chromosphere	265,300

Models capture the evolving understanding of the solar corona dynamics and the origin of space weather events. Advances in understanding are enabled by HEC resources, which are used to test and improve theories via high-fidelity simulations. In Figure 23, a model generates a three-dimensional numerical simulation of a solar corona jet, visualized with still frames and animation. Each experiment consists of many runs; the user monitors output for faults or errors so as to stop, make corrections and restart the simulation. NASA's NCCS enables NASA researchers to take a leadership role in this work; without this resource, the work and personnel would be transitioned to NSF projects and other government agencies with collaboration by NASA researchers. (Appendix D8.8, DeVore)

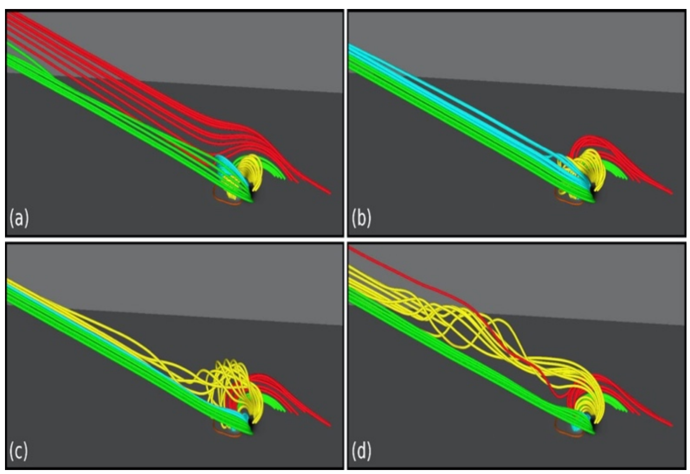

FIGURE 23: VISUALIZATION OF SOLAR CORONA JET FROM MODEL OUTPUT. (SOURCE: APPENDIX D8.8, DEVORE)

To further understand solar flares and storms, in 2018 researchers employed NCCS resources to model a gigantic superflare and coronal mass ejection from Kappa Ceti, a Sun-like star 29 light years from Earth. The simulated explosion released ~700,000 times the energy used by all humans on Earth over an entire year—an event as powerful as our Sun's famous 1859 storm known as the Carrington Event. A solar storm of this magnitude today would widely damage communications and electrical power infrastructure, with $40 billion in daily economic losses in the U.S. alone. Astoundingly, observations show superflares ten times more powerful erupting from Kappa Ceti. (Appendix D9.15, Lynch)

The computations used the Adaptively Refined Magnetohydrodynamics Solver (ARMS) The published ARMS simulation ran on 256 cores of the NCCS Discover supercomputer for ~200 hours. It produced 300 gigabytes of data, later moved to NCCS local storage for analysis and visualization, yielding another 300 gigabytes of data. Ben Lynch from Space Sciences Laboratory at the University of California, Berkeley, said, “From a theorist and modeler’s perspective, the availability of high-performance computing infrastructure like NCCS Discover is absolutely necessary for advancing our first-principles understanding of complex, multi-scale physical systems such as solar and stellar atmospheres. These resources dramatically increase our capacity for space weather forecasting through the ability to model the dynamic, time-dependent evolution of our own Sun-to-Earth system and to facilitate the application of these tools to more exotic astrophysical environments.” (Appendix D9.15, Lynch)

HECC resources are being used to derive high-fidelity 3D radiative models to reproduce the multiscale solar dynamics from the interior to the corona to enhance predictive capabilities of activity manifestations, along with impacts on the space weather conditions and Earth’s environment. These codes used in these models, such as StellarBox, require nonstop execution and are used for:

- Calibration and interpretation of observations from NASA’s space missions.
- Understanding the observed phenomena.
- Development capabilities to predict space weather conditions.
- Verification and validation new data analysis methods.
- Support new instrumentation development and design to identify requirements and specifications.

The output is analyzed using different approaches such as feature identification and tracking, statistical analysis, machine learning techniques, and 3D visualization**.** Large-scale, non-linear models and post processing of the data, application testing, and revision require substantial multi-core computational resources (4,000–10,000 CPUs) for a given campaign. In the current environment, this requirement carries with it extremely long queue wait times. Future growth is expected for the investigation of large-scale solar eruptions and extreme events, as well as the application of 3D radiative MHD modeling to the full spherical Sun, as a basis for modeling from first principles. (Appendix D6-1.3, Kitiashvili)

HECC supports the Magnetospheric Multiscale Mission (MMS) which consists of four identical spacecraft flying in a tetrahedral formation, launched in 2015. NASA missions need simulations to guide interpretations and analyses of data to achieve physical understanding, and to develop predictive capabilities such as space weather forecasts. The mission team performs simulations ranging from global simulations for the Earth, Moon, and asteroid Psyche, to support missions including the MMS (primary mission support), Radiation Belt Storm Probes (RBSP), THEMIS (nightside magnetosphere dynamics), ARTEMIS (Moon studies), and Psyche (discovery mission to the metallic asteroid 16 Psyche). Figure 24 shows visualizations of some analysis results. Lunar simulations support the Human Exploration and Operations Mission Directorate, in addition to SMD. Without HECC, the research would have to compete for time on DOE and NSF supercomputers, disrupting the regular flow of the missions. (Appendix D8.9, Chen)

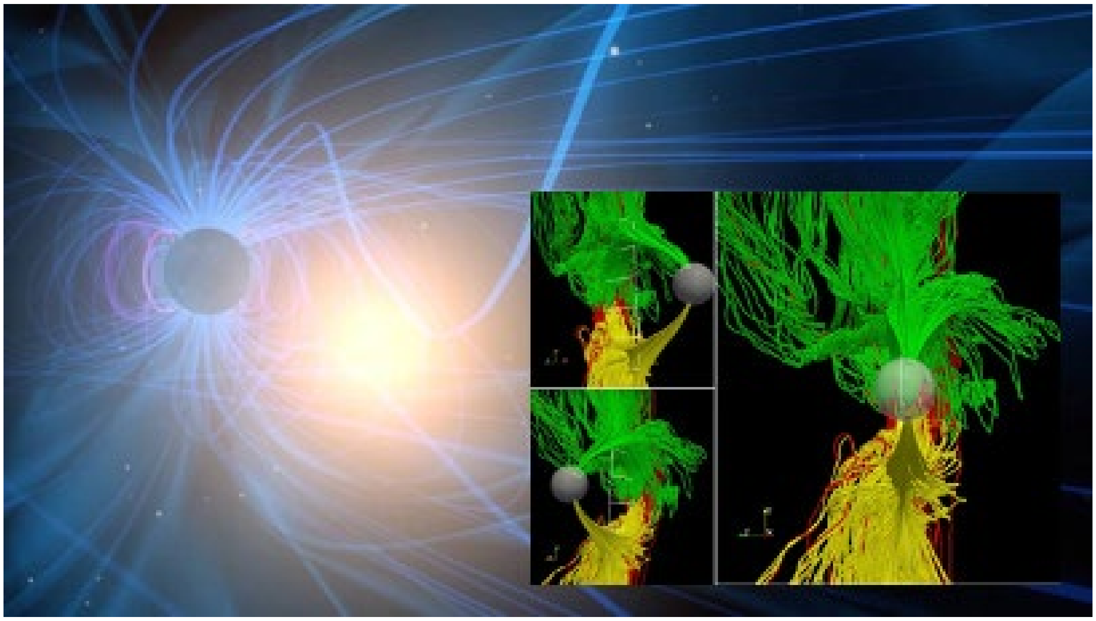

FIGURE 24: BACKGROUND SHOWS MAGNETIC RECONNECTION FOR THE MAGNETOSPHERIC MULTISCALE (MMS) MISSION WITH BRIGHT POINT ILLUSTRATING A MOMENT OF RECONNECTION. (CREDIT: NASA/GODDARD SPACE FLIGHT CENTER.) INSET SHOWS A 3D GLOBAL HYBRID SIMULATION OF EARTH'S MAGNETOSPHERE AS LARGE-SCALE MAGNETIC FLUX ROPES FORM. (CREDIT: NASA AFFILIATED RESEARCHERS HOMA KARIMABADI, UNIVERSITY OF CALIFORNIA, SAN DIEGO; BURLEN LORING, UNIVERSITY OF CALIFORNIA, BERKELEY)

6.4 Planetary Science Division (PSD)

HEC supports many aspects of Planetary Science Division's missions and research, primarily in modeling and data analysis. The entry, descent, and landing (EDL) modeling work is described under the STMD section. However, understanding physical processes that may be different than those observed on Earth are also supported by HEC.

PSD missions and research utilize HEC, enabling researchers to analyze large volumes of data and to further expand and operate forecast models. Table 6-16 and Table 6-17 show actual usage by PSD programs for the past five years of HECC and NCCS resources, respectively. The HEC Program uses Standard Billing Units (SBUs), described in Appendix A as the metric by which allocations are made and consumption is measured. The top users of HEC resources in the PSD Division are identified in Table 6-18.

TABLE 6-16: PSD HECC USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	9,416301	6,265,610	3,615,047	5,699,276	4,538,108	4,010,585

TABLE 6-17: PSD NCCS USAGE (2015-2020).

	FY20	FY19	FY18	FY17	FY16	FY15
SBU	40,933	28,053	85,102	106,274	25,718	N/A

TABLE 6-18: TOP PSD USERS OF HEC SBUS IN 2020.

PI	Organization	Project Title	FY20 SBUs Consumed
Dong, Chuanfei	Princeton University	Exoplanetary Space Weather, Climate and Habitability: Consequences of Atmospheric Loss	2,307,215
Dong, Chuanfei	Princeton University	A model database of unmagnetized planetary space environment with an interactive user interface	854,383
Dong, Chuanfei	Princeton University	Solar wind interaction with Venus: From the planetary interior to interplanetary space	680,454
Dong, Chuanfei	Princeton University	Mercury's Dynamic Magnetosphere Under Varying External Conditions	676,785
Dong, Chuanfei	Princeton University	MAVEN Extended Mission 3	613,643
Bellan,Josette	NASA/JPL	Rocket Plume Cratering of the Lunar Regolith	385,819
Huang, Xinchuan	SETI Institute	Highly accurate ro-vibrational line lists for HCN and HNC for use in studies of planetary and exoplanetary atmospheres	262,272
Yadav, Rakesh Kumar	Harvard University	Understanding the JUNO data using Jovian dynamo simulations	239,585

In studying the origin of the planetary system around our Sun, several potential theories have been put forward. One approach starts with the application of first principles and has translated into a model by researchers working at Harvard and the University of Utah. The model simulates planetary formation processes from a circumstellar disk of gas and dust, in scales from 1-micron particles to planet-scale objects. One interesting element of this study is the formation of the Pluto-Charon binary system, which was discovered just 40 years ago. Pluto's largest moon, Charon, was discovered and subsequently observed to have a virtual tidal lock with Pluto. The model is being used to study this binary planet-moon, along with the evolution of the Kuiper belt objects. The calculations required are processor intensive and can only be run on a supercomputer—typically requiring 1.6 million CPU hours for an investigation, several of which are run in the course of a year. (Appendix D8.1, Kenyon)

Another aspect of the HEC Program's support to PSD is the modeling of thermoelectric materials from first principles. These materials will lead to a new generation of Radioisotope Power Generators (RPG) essential to deep space missions. Since their introduction 50 years ago, RPGs, using conventional materials to convert heat to electricity, are only about 6% efficient. As power demand increases and mission lifetime extends, the need for more efficient materials also grows since the mass allocations to the power unit are fixed. Recent developments in materials synthesis, novel compounds with complex structures, the ability to engineer with increasing precision micro- and nanostructure features—coupled with improved scientific understanding of electrical and thermal transport in such engineered materials—have created the need for in-depth theoretical simulations with fast turnaround times.[51] Models simulate electronic band structures, and a large number of runs are used to compute density of states and thermoelectric transport coefficients. Output is analyzed at JPL, but must also be maintained for long-term retention for further review and potential re-use. Given the complexity of fabricating experimental samples, it is essential to use the simulation studies to guide the experimental process. This strategy has resulted in demonstrated performance improvements to 15% efficiency since its introduction, so far. In order to understand alternatives, the ability to explore the parameter space depends upon the availability of no-cost NASA resources. The project expects slight growth in load, 20–30%, over the next decade. (Appendix D6-3.5, von Allmen)

6.5 Biological and Physical Sciences Division (BPSD)

Due to the BPSD Division's recent move into the Science Mission Directorate, further information is not available. An examination of the HEOMD allocations does not indicate significant use of HEC resources.

Appendix A

Assessment Methodology Notes

A.1. HEC Ecosystem

In order to consider the full range of user needs, one must consider the entire HEC ecosystem. One breakout of the various elements is displayed in Figure A- 1. The relationships among the elements vary with Center, but all must be considered in supplying the users' needs. Most users recognize that every element of the ecosystem in Figure 1 affects their ability to use HEC resources to meet mission needs, and that a healthy and balanced ecosystem is necessary. Many of the issues raised by the users stretch beyond the actual hardware systems into how they are allocated and managed, how funding is channeled into the HEC Program, and how the program's performance is measured.

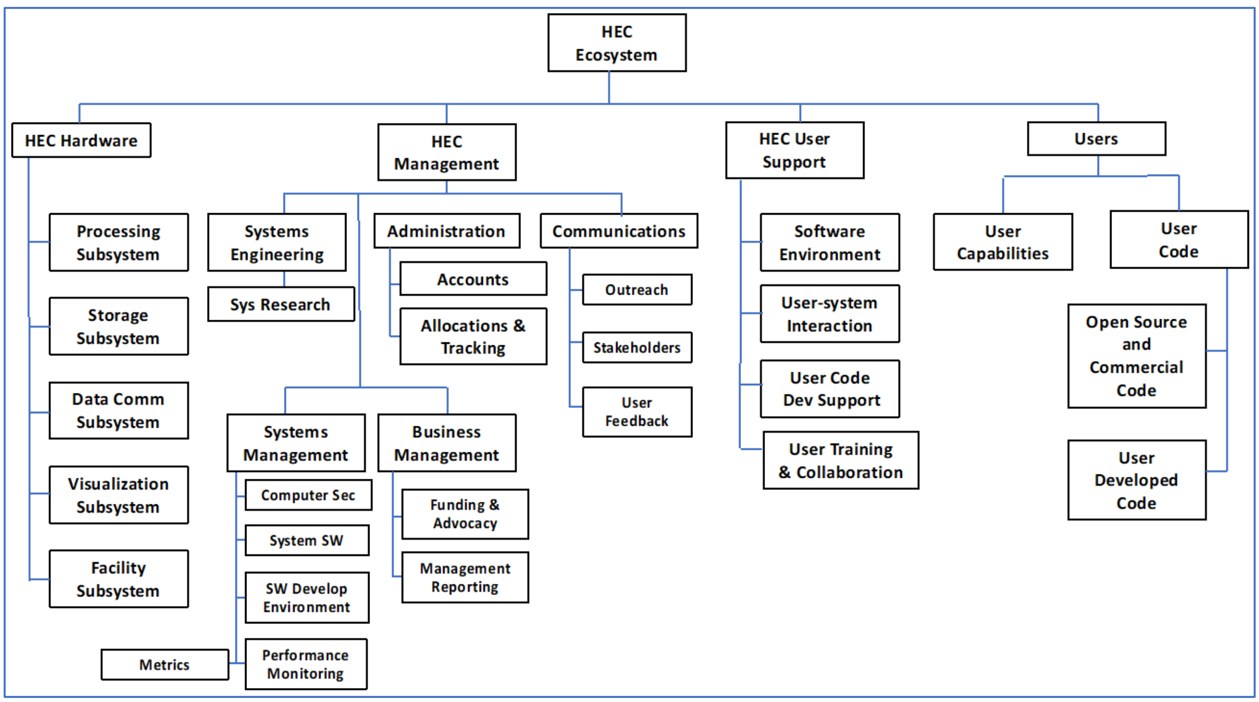

Figure A- 1. Elements Of A High-Performance Computing Ecosystem (Credit: Peter Williams)

Three general classes of software run on HEC assets:

- Custom-developed code for individual research projects but not made available for use outside the project.
- Facility-class NASA codes, such as FUN3D, GEOS, and NU-WRF.
- Commercial or open-source software developed elsewhere and licensed to HEC users. Some software packages are widely used (e.g., ArcGIS, compilers, or TensorFlow) and acquired and managed by the HEC Program. Other software is acquired by the individual project and only needed and licensed for their use.

A.2. Standard Billing Units (SBU) Allocation and Usage

The High-End Computing Capability (HECC) project and the NASA Center for Climate Simulation (NCCS) use a common Standard Billing Unit (SBU) for allocating and tracking computing time usage. The NASA SBU is a way of standardizing work across dissimilar architectures. Representative codes are run on each architecture and their run times are then compared to the baseline values, currently calculated for Pleiades Broadwell nodes. A conversion factor is then calculated to convert the individual system's accounting values to SBUs.

The goal is that users will be charged the same number of SBUs for a given workload regardless of the environment they choose to run in.

Each architecture at HECC and NCCS determines the Minimum Allocatable Unit (MAU). This is the smallest unit of the hardware resources that the scheduling software will allocate to a job. For Aitken, Pleiades, Electra, and Merope at HECC and Discover at NCCS, the MAU is a node; for Endeavor (HECC), the MAU is 8 cores.

The SBU conversion factor represents the difference in efficiency of a particular architecture in performing the baseline workload on its MAU compared to the Pleiades baseline MAU. The larger the conversion factor, the faster the architecture completes the work.

In October 2018, a new set of SBU conversion factors went into effect that uses runtime on a Pleiades Broadwell node as a baseline. SBUs identified in this report have all been converted to use the SBU2 factor.

HECC updates the cost per SBU each October—the beginning of the federal fiscal year (FY)—based on the total SBUs available at that time and the full cost of HECC investment for that fiscal year. The historical cost per SBU for HECC is listed in Table A- 1:

TABLE A- 1. HECC COST PER SBU.

	FY20	**FY19**	**FY18**	**FY17**	**FY16**	**FY15**	**FY14**	**FY13**	**FY12**
Cost per HECC SBU ($)	0.47	0.55	0.65	0.97	1.05	1.13	1.54	2.26	3.23

A.3. The Charging Formula

The formula for SBUs charged to a job is as follows:

SBUs charged = number of MAUs x number of wall clock hours x SBU conversion factor

Given that an SBU is a representation of an amount of useful work, a job should be charged similar amounts of SBUs whether it is running on a more efficient or less efficient system.

A.4. Use Case Collection Methodology

User needs were identified and collected in the form of use cases describing how scientists, engineers, and researchers use HEC or expect to use it in the future (Cockburn, 2001). This mechanism allowed for (1) the context in which statements as to users' needs could be placed and (2) an improved understanding of how their work fits into the larger NASA mission objectives. A workshop was scheduled in a series of sessions stretching from June 1, 2020, through June 19, 2020, to discuss a representative sample of use cases. Over 250 individual researchers and program managers were invited to contribute use cases. Invitations went to HEC users and to NASA researchers who are using computing resources from other government agencies. An evaluation of the responding organizations and projects identified gaps in creating a representative sample of the user community, and specific representatives were contacted and interviewed to ensure input was sufficiently comprehensive to provide insight into the complete scope of needs. Subsequently, a review of the plans published by other government agencies was conducted for any additional NASA discipline or possible user community that might have been overlooked. Additional input was collected from the NASA LaRC High-Performance Computing Initiative workshop, held July 9-10, 2020, and from HEC Monthly Status Reports to the Science Mission Directorate (SMD). In total, 75 use cases were collected, some of which represented a condensation from multiple invitees. While the primary scope of the assessment focused on the NASA users of HEC, input was collected from the full range of scientific and engineering research conducted by NASA employees, contractors, and grant recipients. In some cases, collaborating partners in other government agencies were considered. Users who are envisioning needing HEC resources for future NASA activities, even if not in the current model, were also included. This broad scope permitted consideration of needs that were pathfinding for NASA's current community, reflecting possible future projects. One such pathfinder in coordinating multiple supercomputing processing centers is the Event Horizon Telescope, which operates in conjunction with the Chandra X-ray Observatory and other NASA telescopes.

A.5. Invitees

The roles of invitees were predefined in the invitation. Several different approaches were taken to ensure invitations included a wide range of users to ensure diverse participation, including the user name list as well as discussions with each computing center's user liaison staff. Participants, identified in Table A- 2, were expected to share a use case; in some cases, groups of related participants shared a single use case. Observers identified in Table A- 3 are managers and HEC staff learning better the issues involved. The actual attendees at each session are identified in the notes section for each Affinity Group in Appendix D (separate document).

TABLE A- 2. INVITED PARTICIPANTS IN WORKSHOP.

Name	Site	Division
Henderson, John	Atmospheric and Environmental Research	ESD
Barnhardt, Michael	ARC	HEOMD
Ciotti, Robert	ARC	HEC
D'Souza, Sarah N	ARC	STMD
Dalle, Derek	ARC	HEOMD
Denison, Marie	ARC	STMD, ARMD TTT, ARMC AETC
Dungan, Jennifer Lee	ARC	Earth Science (E)
Garcia, Joseph	ARC	HEOMD
Gomez, Reynaldo	ARC	HEOMD
Henze, Chris	ARC	HEC
Hood, Robert	ARC	HEC
Jenkins, Jon	ARC	PSD
Jin, Henry	ARC	HEC
Jones, Daniel Ray	ARC	HEOMD
Kahre, Melinda	ARC	Planetary Science (P)
Kiris, Cetin	ARC	ARMD
Kitiashvili, Irina Nikolaevna	ARC	Heliophysics (H)
Lawson, John	ARC	ARMD
Lee, Michael	ARC	ARMD
Michaelis, Andrew	ARC	Earth
Murman, Scott M.	ARC	ARMD/HEOMD
Nemani, Rama	ARC	Earth
Ranjan, Shubha	ARC	HEC
Ricca, Alessandra	ARC	Astrophysics (A)
Rogers, Tamara Marie	ARC	Astrophysics (A)
Rogers, Stuart E	ARC	HEOMD
Roozeboom, Nettie	ARC	HEOMD SLS
Schuh, Michael	ARC	HEOMD
Ventura Diaz, Patricia	ARC	ARMD RVLT
Wright, Michael	ARC	STMD, PSD
Yoon, Seokkwan	ARC	ARMD TTT, STMD

Name	Site	Division
Aftosmis, Michael	ARC-TNA	ARMD, SMD, HEOMD
Barad, Michael	ARC-TNA	ARMD/HEOMD
Cadieux, Francois	ARC-TNA	ARMD/HEOMD
Kenway, Gaetan	ARC-TNA	ARMD
Madavan, Nateri	ARC-TNA	ARMD
Nemec, Marian	ARC-TNA	ARMD, SMD, HEOMD
Rogers, Michael	ARC-TNA	ARMD
Wong, Man Long	ARC-TNA	ARMD
Hill, Jeffrey	ARC-TSA	HEOMD, PSD
Vanderkam, Jeremy	ARC-TSS	HEOMD
Springer, Tony	ARMD	ARMD
Slotnick, Jeff	Boeing	ARMD
Hopkins, Philip Fajardo	CalTech	Astro
Brown, Benjamin	University of Colorado Boulder	Helio
Featherstone, Nick	University of Colorado Boulder	Helio
Toomre, Juri	University of Colorado Boulder	Astro
Cronk, Heather	Colorado State University	ESD
Partain, Phillip	Colorado State University	Earth
Van Den Heever, Susan Claire	Colorado State University	Earth Science (E)
Schulthess, Thomas	ETH Zurich	
Sobel, Adam	GISS/COL	Earth
Debonis, James	GRC	ARMD
Dudek, Julianne Conley	GRC	ARMD TTT
Johnson, Susan M	GRC	ARMD RVLT
Moder, Jeffrey P	GRC	ARMD
Wey, Changju Thomas	GRC	ARMD
Cruz, Carlos	GSFC	HEC
Dolan, Jim	GSFC	HEC
Knowland, Emma	GSFC	Earth
Ormseth, Reid	GSFC	HEC
Pfaff, Bruce	GSFC	HEC
Shute, Jim	GSFC	HEC
Srivastava, Sujay	GSFC	HEC
Thomas, Brian	GSFC	Helio X-ray
Pelissier, Craig	GSFC 606	HEC

Name	Site	Division
Kuo, Kwo-sen	GSFC 612	Earth
Yuan, Tianle	GSFC 613	Earth
Chin, Mian	GSFC 614	Earth
Giles, David	GSFC 618	Earth
Holben, Brent	GSFC 618	Earth
Osmanoglu, Batu	GSFC 618	Earth
Wooten, Maggie	GSFC 618	Earth
Akaike, Yosui	GSFC 661	Astro
Krizmanic, John	GSFC 661	Astro
Antiochos, Spiro	GSFC 670	Helio
Pulkkinen, Antti	GSFC 670	Helio
Attie, Raphael	GSFC 671	Helio
Ireland, Jack	GSFC 671	Helio
Kirk, Michael	GSFC 671	Helio
Narock, Ayris A.	GSFC 672	Helio
DeVore, Carl	GSFC 674	Helio
Lapenta, Giovanni	GSFC 674	Helio
Macneice, Peter	GSFC 674	Helio
Mazarico, Erwan	GSFC 698	PSD
DelGenio, Anthony	GSFC GISS	PSD
Kelley, Max	GSFC GISS	Earth
Ruedy, Reto	GSFC GISS	Earth
McCarty, William	GSFC GMAO	Earth
Fok, Mei-ching H	GSFC-6730	Helio
Collado-Vega, Yaireska M	GSFC-6740	Helio
Kuznetsova, Maria M.	GSFC-6740	Helio
Mays, M. Leila	GSFC-6740	Helio
Le Moigne, Jacqueline	GSFC/407	Earth
Carroll, Mark	GSFC/606	HEC
Peters-Lidard, Christa	GSFC/610	Earth
Matsui, Toshi	GSFC/612	Earth
Tao, Wei-Kuo	GSFC/612	Earth
Colarco, Peter	GSFC/614	Earth
Kumar, Sujay	GSFC/617	Earth
Nearing, Grey	GSFC/617/	Earth

Name	Site	Division
Griffith, Peter	GSFC/618	Earth
Hoy, Elizabeth	GSFC/618	Earth
Schmidt, Gavin	GSFC/GISS	Earth
Clune, Tom	GSFC/GMAO	Earth
da Silva, Arlindo	GSFC/GMAO	Earth
Gelaro, Ron	GSFC/GMAO	Earth
Lucchesi, Robert	GSFC/GMAO	Earth
Molod, Andrea	GSFC/GMAO	Earth
Ott, Lesley	GSFC/GMAO	Earth
Pawson, Steven	GSFC/GMAO	Earth
Putman, Bill	GSFC/GMAO	Earth
Kenyon, Scott	Harvard Center for Astrophysics	Astro
Rosen, Anna	Harvard Center for Astrophysics	Astrophysics (A)
Powell, Brian	Johns Hopkins University	Astro
Altinok, Alphan	JPL	Other
Bellan, Josette	JPL	ARMD TTT, PSD, STMD
Bue, Brian	JPL	Earth
Chan, Samuel	JPL	Earth
Cheng, Cecilia	JPL	Earth
Fukumori, Ichiro	JPL	Earth
Hua, Hook	JPL	Earth
Kiessling, Alina	JPL	Astro
Kurowski, Marcin	JPL	Earth
Larour, Eric Yves	JPL	Earth
Mccoy, Kelli	JPL	Planets
Menemenlis, Dimitris	JPL	Earth
Owen, Susan	JPL	Earth
Rabinovitch, Jason	JPL	Planets, STMD
Rebbapragada, Umaa	JPL	Asto
Su, Hui	JPL	Earth
Vance, Steven Douglas	JPL	Planets
Verkhoglyadova, Olga	JPL	Helio
Villarreal, Michaela Nicole	JPL	Planets
von Allmen, Paul	JPL	Other
Amar, Adam	JSC	HEOMD

Name	Site	Division
Greathouse, James S.	JSC-EG311	HEOMD
Brehm, Christoph	University of Kentucky	STMD
Daughton, William	Los Alamos National Lab	Helio
Carter, Melissa B.	LaRC	ARMD Comm SST
Edquist, Karl	LaRC	HEOMD
Elmiligui, Alaa	LaRC	ARMD
Glaesgen, Ed	LaRC	ARMD
Heeg, Jennifer	LaRC	ARMD Conv Aero Solns
Jelley, Ben	LaRC	Earth
Khorrami, Mehdi R	LaRC	ARMD
Korzun, Ashley M.	LaRC	STMD
Krist, Steven	LaRC	HEOMD
Lang, Chris	LaRC	Explore
Lee-Rasch, Elizabeth	LaRC	CAS
Loubeau, Alexandra	LaRC	ARMD Comm SST
Malik, Mujeeb R	LaRC	ARMD
Nielsen, Eric	LaRC	CAS
Rivers, Melissa B.	LaRC	ARMD AETC
Streett, Craig	LaRC	NESC
Wagner, David	LaRC	Explore
Warner, Jim	LaRC	Explore
Yamakov, Vesselin	LaRC	Explore
Yatheendradas, Soni	GSFC	Earth
Yeratapally, Sai	LaRC	Explore
Anderson, Kyle	LaRC D302	ARMD CAS
Buning, Peter	LaRC D302	ARMD CAS
Choudhari, Meelan	LaRC D302	ARMD CAS
Allen, Brian	LaRC D303	ARMD
Balakumar, Ponnanpalam	LaRC D303	ARMD
Kleb, Bill	LaRC D305	ARMD
Bauerie, Robert	LaRC D306	ARMD
Jacobson, Kevin	LaRC D308	ARMD
Boyd, Doug	LaRC D321	ARMD
Kazachenko, Maria	LASP	Heliophysics (H)
Diachin, Lori	Lawrence Livermore National Laboratory	External

Name	Site	Division
Riaz, Amir	University of Maryland	HEOMD
Gombosi, Tamas	University of Michigan	Helio
Toth, Gabor	University of Michigan	Helio
Hill, Chris	MIT	Earth
Canabal, Francisco	MSFC	HEOMD SLS/STMD
Ramachandran, Rahul	MSFC	Earth
West, Jeff	MSFC	HEOMD
Guhathakurta, Madhulika	GSFC	Helio
Thompson, Barbara	GSFC	Helio
Hansen, KC	NASA HQ DG	PSD
Koehn, Patrick	NASA HQ	Helio
Rempel, Matthias	NCAR	Earth
Schuster, David	NESC	NESC
Keesee, Amy	University of New Hampshire	Heliophysics (H)
Govett, Mark	NOAA Earth Systems Research Laboratories	Earth
Yungster, Shaye	Ohio Aerospace Institute	ARMD TTT
Kothe, Doug	ORNL	External
Bhattacharjee, Amitava	Princeton University	Helio
Dong, Chuanfei	Princeton University	Helio
Ostriker,Eve	Princeton University	Astrophysics (A)
Brooks, Alyson Michelle	Rutgers University	Astro
Opgenorth, Matthew	Sierra Nevada	HEOMD
Nakada, Kazumi	GSFC	Earth Science (E)
Ihme, Matthias	Stanford University	STMD
Peeples, Molly	Space Telescope Science Institute	Astro
Cattaneo, Fausto	University of Chicago	Helio
Foster, Ian	University of Chicago	Earth
Germaschewsk, Kai	University of New Hampshire	Helio
Mavriplis, Dimitri	University of Wyoming	ARMD
Miesch, Mark	University Corporation for Atmospheric Research	Earth
Jiang, Yanfei	University of California Santa Barbara	Astro
Halem, Milt	University of Maryland Baltimore County	Earth
Volkov, Alexey	University of Arizona	STMD

Name	Site	Division
Dumitrica, Traian	University Minnesota	STMD
Miranda Braganca, Vivian	University of Arizona	Astro
Stephani, Kelly	University of Illinois Urbana-Champaign	
Primack, Joel	University of California Santa Cruz	Astro
Smith, Ben	University of Washington Polar Science Center	Earth
Fedorov, Alexey	Yale University	Earth
Keller, Christoph	GSFC	Earth
Legrande, Allegra	GSFC	Astro
Thompson, Matt	GSFC	Earth
Way, Michael	GSFC	Astro

TABLE A- 3. OBSERVERS INVITED TO WORKSHOP.

Name	Site	Role	Division
Biswas, Rupak	ARC	Observer	Exploration
Hartman, Blaise	ARC	Observer	HEC
Manning, Ted	ARC	Observer	HEC
Mathias, Donovan	ARC	Observer	HEC
Mehrotra, Piyush	ARC	Observer	HEC
Thigpen, William	ARC	Observer	HEC
Dunbar, Jill	ARC	Scribe	HEC
Moyer, Michelle	ARC	Scribe	HEC
Ognoskie, Lindsay	ARC	Scribe	HEC
Pitta, Katie	ARC	Scribe	HEC
Hill, Jo	GSFC	Observer	GSFC Management
Thomas, Brian	GSFC	Facilitator	Helio X-ray
Clampin, Mark	GSFC 6000	Observer	GSFC Management
Cohen, Jarrett	GSFC 606	Scribe	Earth
Keefe, Sean	GSFC 606	Scribe	HEC
Irons, Jim	GSFC 6100	Observer	GSFC Management
Cole, Marge	GSFC/407 KBR	Facilitator	Facilitator
Carriere, Laura	GSFC/606	Observer	HEC
Duffy, Daniel	GSFC/606	Observer	HEC

Name	Site	Role	Division
Little, Mike	GSFC/606	Coordinator	HEC
Peirce, Bob	GSFC/606	Observer	HEC
Salmon, Ellen	GSFC/606	Observer	HEC
Williams, Peter	GSFC/606	Facilitator	Facilitator
Hedin, Daniel	HEOMD	Observer	HEOMD
Considine, David	HQ	Observer	Earth
Hook, Elizabeth	HQ	Facilitator	HEC
Lee, Tsengdar	HQ	Observer	HEC
Murphy, Kevin	HQ	Observer	EOS-DIS
Pina, Aaron	HQ	Observer	HEC
Mason, Angela	HQ ARMD	Observer	ARMD
Ferraro, Robert	JPL	Observer	General
Petraska, Karen	HQ OCIO	Observer	HQ
de la Beaujardiere, Jeff	NCAR	Observer	ESD
Willcox, Karen	University of Texas Austin	Speaker	ARMD

Appendix B

Acronym List

AAVP	Advanced Air Vehicles Program
ABoVE	Arctic Boreal Vulnerability Experiment
ADAPT	Advanced Data Analytics PlaTform
AETC	Aeronautics Evaluation and Test Capabilities
AFOSR	Air Force Office of Scientific Research
AF/XOW	Air Force Directorate of Weather
AI	Artificial Intelligence
AMD	Advanced Micro Devices, Inc., a commercial processor firm
AOSP	Airspace Operations & Safety Program
ARC	NASA Ames Research Center (NASA/ARC)
ARM	A commercial processor formerly known as Advanced RISC Machine
ARMD	Aeronautics Research Mission Directorate
ASIAS	Aviation Safety Information Analysis and Sharing
Cart3D	Cartesian 3D software package
CCA	Commerce Control Act which created the Commerce Control List governing export licenses
CFD	Computational Fluid Dynamics
CI/CD	Continuous Integration/Continuous Delivery, a software engineering practice
CNDE	Computational Non-destructive Evaluation
DOE	Department of Energy
DNS	Direct Numerical Simulation
DPLR	Data Parallel Line Relaxation
EAR	Export Administration Regulations
ESD	Earth Science Division
FAIR	Findable, accessible, interoperable, and reusable policy
FOQA	Flight Operations Quality Assurance
FUN3D	Fully Unstructured Navier Stokes in 3D flow solver
GOES	Goddard Earth Observing System model, currently version 5
GFC	Goddard Private Cloud
Github	Software development platform for code management

GLM	Geostationary Lightning Mapper instrument on GOES-16
GMAO	Goddard Modeling and Assimilation Office
GOES	Geostationary Operational Environmental Satellite (now GOES-16)
GPGPU	General Purpose Graphics Processor Unit
GPU	Graphics Processor Unit
GRC	NASA Glenn Research Center (NASA/Glenn)
GSDOP	Ground Support Development & Operations Program
GSFC	NASA Goddard Space Flight Center
HEC	High-end computing
HECC	High-End Computing Capability (at NASA/ARC)
HEOMD	Human Exploration and Operations Mission Directorate
HPC	High-performance computing
IASPO	Integrated Aviation Systems Program Office
ISS	International Space Station
ISSM	Ice-sheet and Sea-level System Model
ITAR	International Trafficking in Arms Regulations
LARC	NASA Langley Research Center
LAVA	Launch, Ascent and Vehicle Aerodynamics computational framework
LES	Large Eddy Simulation
LDAS	Land Data Assimilation System
LIS	Land Information System
LPT	Low pressure turbine
ML	Machine Learning
MPCV	Multi-Purpose Crew Vehicle
MSFC	NASA Marshall Space Flight Center (NASA/MSFC)
NAS	NASA Advanced Supercomputing facility (located at NASA/ARC)
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCCS	NASA Center for Climate Simulation (located at NASA/GSFC)
NDE	Non-destructive Evaluation
NESC	NASA Engineering and Safety Center
NEX	NASA Earth Exchange (located at NASA/ARC)
NLCS	National Leadership Computing System Initiative
NOAA	National Oceanic and Atmospheric Administration
NSF	National Science Foundation

NU-WRF	NASA Unified Weather Research Forecast model
OCIO	NASA Office of the Chief Information Officer
ORNL	Oak Ridge National Laboratory
OSTP	US Government Office of Science and Technology Policy
OVERFLOW	A 3D time marching implicit Navier-Stokes code
PATRAN	Commercial finite element analysis modeling software
POST2	Program to Optimize Simulated Trajectories II
PPBE	US Government Planning, Programming, Budgeting and Execution
PSD	Planetary Science Division
SDP	Science data processing
SLS	Space Launch Systems
SMD	Science Mission Directorate
STK	Satellite Toolkit
STMD	Space Technology Mission Directorate
TACP	Transformative Aero Concepts Program
TESS	Transiting Exoplanet Survey Satellite
TPS	Thermal Protection System
TTT	Transformative Tools and Technologies Program
UQ	Uncertainty Quantification
UVAFME	University of Virginia Forest Model Enhanced for individual trees

Appendix C

Bibliography

(ASCAC), A. S. C. A. C. (2015). *Exascale Computing Initiative*. Washington, DC: US DOE

Aguilo, M. (2020). *PLATO: Optimization-Based Design*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Aldenbrook, K. (2020). Ames Unitary Plan Windtunnel. Retrieved from https://www.nasa.gov/centers/ames/orgs/aeronautics/windtunnels/index.html

Anonymous. (2018). Welcome to the HySDS Wiki (Hybrid Science Data System). Retrieved from https://hysds-core.atlassian.net/wiki/x/AoAC

Bailey, F. R., & Kutler, P. (1988). *NAS—The First Year*. Paper presented at the 16th Congress of the International Council of Aeronautical Sciences, Jerusalem, Israel. https://www.icas.org/ICAS_ARCHIVE/ICAS1988/ICAS-88-4.8.1.pdf

Biswas, R. (2019). *Capability Portfolio Status Report: High End Computing Capability* Unpublished internal report. NASA. Exploratory Technology, Ames Research Center.

Bonadonna, M., Lanzerotti, L., & Stailey, J. (2017). The National Space Weather Program: Two decades of interagency partnership and accomplishments. *Space Weather, 15*(1), 14-25. doi:10.1002/2016sw001523

Budalakoti, S., Srivastava, A. N., Akella, R., & Turkov, E. (2006). *Anomaly Detection in Large Sets of High-Dimensional Symbol Sequences*. Paper presented at the 2006 SLAM International Conference on Data Mining https://ntrs.nasa.gov/api/citations/20060046368/downloads/20060046368.pdf

Caers, J. K. (2018). Quantifying uncertainty about Earth's resources. *Eos, 99*. Retrieved from https://doi.org/10.1029/2018EO097471

Carriere, L. (2020). *HPC at the NASA Center for Climate Simulation – Current Work and Emerging Trends.* . Paper presented at the HPCI Technical Interchange Meeting, NASA Langley Research Center.

Casper, J., Lockard, D., Khorrami, M., & Streett, C. (2004). Investigation of Volumetric Sources in Airframe Noise Simulations. In *10th AIAA/CEAS Aeroacoustics Conference*.

Chan, W., Barmore, B., Kibler, J., Lee, P., O'Connor, N., Palopo, K., . . . Zelinski, S. (2018). *Overview of NASA's Air Traffic Management - Exploration (ATM-X) Project*. Paper presented at the AIAA Aviation Technnology, Integration, and Operations Conference, Atlanta, GA.

Chang, Y.-T., Hood, R. T., Jin, H., Heistand, S. W., Cheung, S. H., Djomehri, M. J., . . . Kokron, D. S. (2018). *Evaluating the Suitability of Commercial Clouds for NASA's High Performance Computing Applications: A Trade Study*. Retrieved from https://ntrs.nasa.gov/citations/20190002611

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., . . . Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment, 219*, 145-161. Retrieved from http://www.sciencedirect.com/science/article/pii/S0034425718304139

Cockburn, A. (2001). *Writing Effective Use Cases*: Dorling Kindersley Pvt Limited.

Colket, M., Heyne, J., Rumizen, M., Gupta, M., Edwards, T., Roquemore, W. M., . . . Sankaran, V. (2017). Overview of the National Jet Fuels Combustion Program. *AIAA Journal, 55*(4), 1087-1104. doi:10.2514/1.J055361

Consortium, E., Fukumori, I., Wang, O., Fenty, I., Forget, G., Heimbach, P., & Ponte, R. M. (2020). *Synopsis of the ECCO Central Production Global Ocean and Sea-Ice State Estimate (Version 4 Release 4)*. Retrieved from http://doi.org/10.5281/zenodo.3765929

Consortium, I. (2002). *Concept of Operations for the Community Coordinated Modelling Center*. Retrieved from https://ccmc.gsfc.nasa.gov/CONOPS_final.pdf

COUNCIL, N. S. T. (2019). *NATIONAL STRATEGIC COMPUTING INITIATIVE UPDATE: PIONEERING THE FUTURE OF COMPUTING*. Washington, DC: Office of Science and Technology Policy

D'Souza, S. (2018). *Pterodactyl: Integrated Control Design for Precision Targeting of Deployable Entry Vehicles*. Paper presented at the 15th International Planetary Probe Workshop

Das, S., Matthews, B. L., & Lawrence, R. (2011, 20-23 June 2011). *Fleet level anomaly detection of aviation safety data.* Paper presented at the 2011 IEEE Conference on Prognostics and Health Management.

Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Dev., 8*(10), 3071-3104. doi:10.5194/gmd-8-3071-2015

Garner, R. (2020). Transiting Exoplanet Survey Satellite (TESS) Mission. Retrieved from https://www.nasa.gov/tess-transiting-exoplanet-survey-satellite

Gehmann, H., Barry, J., Deal, D., Hallock, J., & Hess, K. (2003). Columbia Accident Investigation Board. Volume One. Retrieved from https://www.nasa.gov/columbia/home/CAIB_Vol1.html

Glaessgen, E., & Stargel, D. (2012). *The digital twin paradigm for future NASA and U.S. air force vehicles*. Paper presented at the Conference: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Special Session on the Digital Twin, Honolulu, HI.

Hemminger, J. C., Sarrao, J., Crabtree, G., Flemming, G., & Ratner, M. (2015). *Challenges at the Frontiers of Matter and Energy: Transformative Opportunities for Discovery Science*. Retrieved from https://www.osti.gov/servlets/purl/1283188

Holdren, J., & Donovan, S. (2016). *National Strategic Computing Initiative Strategic Plan*.

Ilie, M. (2019). Fluid-structure interaction in turbulent flows; a CFD based aeroelastic algorithm using LES. *Applied Mathematics and Computation, 342*, 309-321. doi:https://doi.org/10.1016/j.amc.2017.10.059

Jones, G. (2020). *Quantum Chemistry with Quantum Computers*. Paper presented at the HPCI Technical Interchange Meeting NASA Langley Research Center.

Kappes, B. (2020). *Efficiently Connecting Exascale HPC and Physical Experiment Through Collaborative Data Management*. Paper presented at the HPCI Technical Interchange Meeting, NASA Langley Research Center.

Kitiashvili, I., & Wray, A. (2019). *Structure and Dynamics of the Overshoot Layer in Rotating Main-Sequence Stars with Shallow Convection Zone*. Paper presented at the AGU Annual Meeting.

Kleb, W. L., Park, M. A., Wood, W. A., Bibb, K. L., Thompson, K. B., Gomez, R. J., & Tesch, S. H. (2019). *Sketch-To-Solution: An Exploration of Viscous CFD with Automatic Grids*. Paper presented at the AIAA Aviation and Aeronautics Forum and Exposition (AIAA AVIATION 2019). https://ntrs.nasa.gov/api/citations/20200002620/downloads/20200002620.pdf

Kogge (Ed.), P., Bergman, K., Borkar, S., Campbell, D., Carlson, W., Dally, W., . . . Yelick, K. (2008). *ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems*. Retrieved from Washington, DC: http://www.cse.nd.edu/Reports/2008/TR-2008-13.pdf

Kokron, D. (2020). *Performance Tuning Experience from Seven Langley Applications*. Paper presented at the HPCI Technical Interchange Meeting, NASA Langley Research Center.

Lee, T. (2020, October 22, 2020). High-End Computing Program: Standard Billing Units Retrieved from https://hec.nasa.gov/user/policies/sbus.html

Lynch, J. (2020). Sticking the landing on Mars: High-powered computing aims to reduce guesswork (2020, July 30) Retrieved from https://phys.org/news/2020-07-mars-high-powered-aims-guesswork.html

Mansour, N. N., & Kitiashvili, I. (2018). *Heliophysics Integrated Modeling: From Sun-Earth to Star-Exoplanet*. Paper presented at the Cambridge Workshops of Cool Stars, Stellar Systems and the Sun (Cool Stars 20)

NASA. (2018). *NASA Strategic Plan 2018*. Washington, DC

NASA. (2018, October 10-11, 2018). *2018 Workshop on Autonomy for Future NASA Science Missions: DRM Team Summary Reports.* Paper presented at the 2018 Workshop on Autonomy for Future NASA Science Mission, Carnegie Mellon University.

NASA. (2019). Femera (Version LAR-19531-1): NASA Technology Transfer Program. Retrieved from https://software.nasa.gov/software/LAR-19531-1

National Academies of Sciences, E., and Medicine. (2016). *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts* Washington, DC: The National Academies Press.

Nielsen, E. (2020). *Petascale Mars Powered Descent Investigations Using FUN3D on the GPU-Based Summit Supercomputer*. Paper presented at the HPCI Technical Interchange Meeting, NASA Langley Research Center.

Peeples, M. (2018). Figuring Out Gas & Galaxies in ENZO. Retrieved from https://foggie.science/

Pesnell, D. (2020). SDO: Solar Dynamics Observatory. Retrieved from https://sdo.gsfc.nasa.gov/

Schmidt, G., Miller, R., Bauer, S., Russell, G., Ackerman, A., Canuto, V., & Way, M. (2019). *GISS Model Development FY20-FY24*. Unpublished internal report submitted in response to a request from Jack Kaye. NASA. Earth Science Division

Siemiginowska, A., Eadie, G., Czekala, I., Feigelson, E., Ford, E., Kashyap, V., . . . Young, C. (2019). *Astro2020 Science White Paper: The Next Decade of Astroinformatics and Astrostatistics*.

Siemiginowska, A., Eadie, G., Czekala, I., Feigelson, E., Ford, E. B., Kashyap, V., . . . Young, C. A. (2019). The Next Decade of Astroinformatics and Astrostatistics. *Bulletin of the American Astronomical Society, 51*, 355. Retrieved from https://ui.adsabs.harvard.edu/abs/2019BAAS...51c.355S

Sinyuk, A., Holben, B. N., Eck, T. F., Giles, D. M., Slutsker, I., Korkin, S., . . . Lyapustin, A. (2020). The AERONET Version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to Version 2. *Atmos. Meas. Tech., 13*(6), 3375-3411. doi:10.5194/amt-13-3375-2020

Slotnick, J., Khodadoust, A., Alonso, J., Darmofal, D., Gropp, W., Lurie, E., & Mavriplis, D. (2014). *CFD Vision 2030 Study: A Path to Revolutionary Computational Aerosciences*. (NASA/CR–2014-218178). Hampton, Virginia: Langley Research Center Retrieved from https://ntrs.nasa.gov/citations/20140003093

Sorensen, B., Norton, A., Joseph, E., & Conway, S. (2019). *Special Report for NASA. Exploring Options for a Bespoke Supercomputer Targeted for. Weather and Climate Workloads*. Retrieved from NASA HEC Program: https://hec.nasa.gov/news/reports/NASA-Hyperion-Research-Bespoke-Weather-HPC-Exec-Summary.pdf

US DOE, O. o. S. (2015). *DOE Advanced Scientific Computing Advisory Committee (ASCAC) Report: Exascale Computing Initiative Review*: United States. Department of Energy. Office of Science.

Vetter, J. S., & Yoo, A. (2002, 16-22 Nov. 2002). *An Empirical Performance Evaluation of Scalable Scientific Applications.* Paper presented at the SC '02: Proceedings of the 2002 ACM/IEEE Conference on Supercomputing.

Wagner, D. (2020). *Computational advances in materials modeling at atomistic and continuum levels*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Walsh, T. (2020). *High-Performance Computing for Solving Inverse Problems in Multiphysics*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Warburton, T. (2020). *Optimizing high-order finite element methods for heterogeneous architectures*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Warner, J. (2020). *Efficient Uncertainty Quantification Using High Performance Computing and Scientific Machine Learning*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Wilkie, T. (2015). Exascale: expect poor performance. Retrieved from https://www.scientific-computing.com/analysis-opinion/exascale-expect-poor-performance

Williams, A., & Green, J. (2020). *Current State of HPC for Entry, Descent, and Landing*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Wood, S. (2020). *From Mini-Apps to Stabilized Finite Elements in FUN3D*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.

Yang, L., Treichler, S., Kurth, T., Fischer, K., Barajas-Solano, D., Romero, J., . . . Karniadakis, G. (2019, 17-17 Nov. 2019). *Highly-scalable, Physics-Informed GANs for Learning Solutions of Stochastic PDEs.* Paper presented at the 2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS).

Yip, W. (2020). NASA Earth Exchange (NEX). Retrieved from https://www.nasa.gov/nex

Zakrajsek, J., Woerner, D. F., & Fleurial, J.-P. (2017). *NASA Special Session: Next-Generation Radioisotope Thermoelectric Generator (RTG) Discussion*. Jet Propulsion Laboratory, California Institute of Technology: NASA

Zubair, M. (2020). *Optimization of Solver for Computational Materials and Structures Problems on Emerging High Performance Architectures*. Paper presented at the HPCI Technical Interchange Meeting at NASA Langley Research Center.