26th International Conference on Parallel Computational Fluid Dynamics
Parallel CFD2014
T. Kvamsdal (Ed)

NASA ADVANCED COMPUTING ENVIRONMENT FOR SCIENCE AND ENGINEERING

RUPAK BISWAS AND PIYUSH MEHROTRA

NASA Ames Research Center
Moffett Field, CA 94035, USA
Email: {rupak.biswas, piyush.mehrotra}@nasa.gov

Key words: Computational Fluid Dynamics, Supercomputing, Big Data, Collaborative Environments, Quantum Computing.

Summary. *High-fidelity modeling, simulation, and analysis, enabled by supercomputing, are becoming increasingly important to NASA's broad spectrum of missions. This paper describes NASA's advanced supercomputing environment at Ames Research Center that is geared toward solving the space agency's most challenging science and engineering problems.*

1 INTRODUCTION

High-fidelity modeling, simulation, and analysis, enabled by supercomputing, are becoming increasingly important to NASA's mission to pioneer the future in space exploration, scientific discovery, and aeronautics research. While scientific and engineering advancements used to rely primarily on theoretical studies and physical experiments, today computational science contributes equally in such achievements. In addition, computational modeling and simulation serves as a predictive tool that is not otherwise available. As a result, high performance computing (HPC) is now integral to the space agency's work in all mission areas. Anchored by the Pleiades supercomputer at the NASA Advanced Supercomputing (NAS) Division, the High End Computing Capability (HECC) Project provides a fully integrated environment to satisfy NASA's diverse modeling, simulation, and analysis needs.

In addition to providing production supercomputing cycles and services to scientists and engineers, HECC serves as the agency's expert source for evaluating emerging HPC technologies and maturing the most appropriate ones into the production environment. The NAS Division investigates advanced IT technologies such as GPU and Xeon Phi accelerators, cloud computing, collaborative environments, integrated hardware/software solutions for Big Data analytics, and adiabatic quantum computing for hard discrete optimization problems. The overall goal is to provide a consolidated leading-edge environment to support NASA's computational and analysis requirements for science and engineering applications.

2 SUPERCOMPUTING

NASA's HECC Project provides world-class computing, storage, and visualization resources to users across the nation to accelerate the development of innovative technologies, ensure new scientific discoveries, efficiently develop complex engineering systems, and reduce risks of NASA missions. The HECC integrated environment includes premier supercomputing systems, a vast data management and archive capability, high-speed networks, and an array of support services. The environment is continuously upgraded and maintained for maximum availability and reliability to help users meet NASA's computational challenges.

2.1 Computing environment

At the heart of this environment is Pleiades [1], a petaflops-scale supercomputer used by over 1,200 scientists throughout the U.S. It is one of the world’s most powerful systems and represents NASA’s state-of-the-art technology for meeting its computational needs, enabling scientists and engineers to conduct key large-scale, high-resolution modeling and simulation. Initially installed in October 2008, Pleiades is a distributed-memory SGI ICE cluster that has been upgraded several times over the years. As of April 2014, the system consists of 184,800 cores of multiple generations of Intel Xeon processors, and has a peak performance of 3.6 PF. The 11,176 nodes comprising the system are interconnected in a hypercube topology using a dual-plane InfiniBand network.

The computing environment provides users with 20 PB of online disk storage along with 115 PB of tape archive storage capacity attached to the supercomputing and data analysis systems, allowing users to archive and retrieve important results quickly, reliably, and securely. In addition, an in-house developed visualization system—one of the largest resolution and most powerful in the world—provides a supercomputer-scale environment to visualize and explore the very large, high-dimensional datasets produced by NASA supercomputers and instruments. The hyperwall, consisting of 128 AMD Opteron based nodes driving 128 monitors, helps researchers display, analyze, and study both observational and simulation datasets in meaningful ways, allowing the use of different tools, viewpoints, and parameters. NAS has also developed a sophisticated multistream concurrent visualization methodology to process and display data in real time while applications run on Pleiades.

2.2 Integrated support environment

To help achieve exceptional performance and groundbreaking results for NASA’s scientific and engineering users, HECC combines cutting-edge HPC technology and techniques with an emphasis on providing reliable, real-world production capability. Key to this integrated approach is the collaboration among its interdisciplinary product teams with both broad and deep knowledge of advanced systems and technologies.

The NAS Division’s expertise in developing and delivering HPC technologies is at the center of HECC’s integrated environment. The systems team evaluates new architectures; acquires, installs, and manages the supercomputing systems; and develops custom software tools and advanced IT security methods. Network engineers provide end-to-end networking services that provide users with access to resources that are critical to the success of their computational projects. The application performance and productivity team is dedicated to helping users optimize application performance, improve code scalability, and effectively port their software to the supercomputers. In addition, the team evaluates advanced hardware and software technologies to identify and leverage those best suited to meet evolving high-end computational challenges for NASA. Finally, visualization experts develop and implement advanced software tools and data analysis technologies customized to help scientists and engineers make new discoveries and design novel air and space vehicles and launch facilities. The team's extensive repertoire includes a sophisticated concurrent visualization framework, which together with the hyperwall makes it feasible to render and store animations showing every simulation time step, letting users see rapid processes in their models—often for the first time. Concurrent visualization has proven crucial to, for example, understanding the complex nature of rotor wakes (Figure 1).

2.3 Sample applications

Rotorcraft: Many of the complex physical phenomena associated with both aircraft and rotorcraft flight are still not well understood and are difficult to predict accurately. Rotorcraft challenges include blade deformations, noise, vortical wakes, and blade-wake interactions. Over the last decade, CFD researchers have supported industry by developing improved physics-based computational tools that will lead to reduced rotorcraft noise pollution and better heavy-lift capabilities in situations where airports are not practical or available, such as in rescue mission scenarios. Using advanced simulation methods, NASA's Rotary Wing Project has achieved significant improvements in rotorcraft modeling capabilities [2]. Typically, these multidisciplinary flow simulations requires 5–24 hours to compute each rotor revolution and 1–2 weeks to run each case using 1,500–4,600 cores on Pleiades. For the simulation shown in Figure 1, grid sizes varied from 60–750 million grid points, and up to 100 TB of disk space were needed to render movies of the unsteady flow simulations.

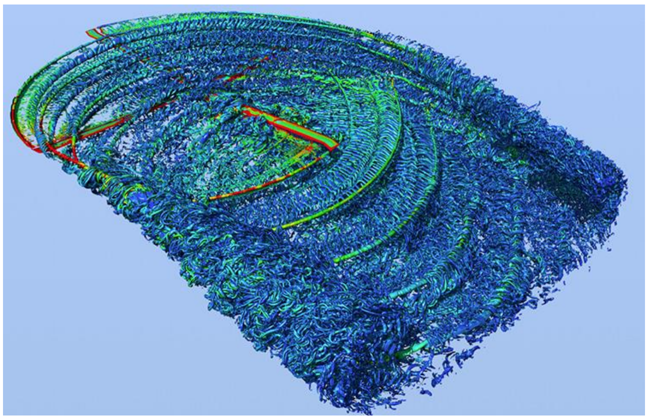

Figure 2. Image from a simulation of a Black Hawk helicopter rotor in forward flight.

Space Launch System: CFD simulations are being used to support the design of NASA's next-generation space launch vehicles [3]. This includes characterizing aerodynamic performance, providing distributed line loads and pressure signatures on the vehicle for structural analysis, performing stage separation analysis, and predicting

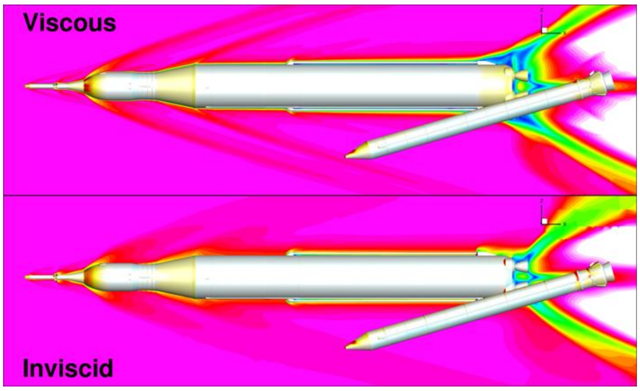

Figure 1. Images from simulations of SLS solid rocket booster separation showing flow around vehicle and boosters.

the launch environment during ignition and takeoff. Simulation results have been used during multiple design analysis cycles of the Space Launch System (SLS) (Figure 2). They are an efficient source of critical design data due to the quick turnaround times and minimal cost to produce results for a large number aerodynamic performance databases and pad configurations. HECC resources at NAS enable fast and efficient turnaround times for CFD simulations of space vehicles and launch environments. Pleiades allows viscous simulation databases with hundreds of cases to be completed in under a week using 200–300 cores per simulation.

Launch Environments: The NAS Division's supercomputing capability is also used to help prepare NASA's 21st century launch complex and determine whether the existing launch pad flame trench and deflector system used for the Space Shuttle can withstand flow created by next-generation launch vehicles (Figure 3). By combining engineering methods with high-fidelity aerodynamics simulations run on Pleiades, NASA teams are using the OVERFLOW and CART3D CFD codes to run simulations that will help streamline the development of these vehicles. Work includes analyzing potential design shapes for exploration vehicles, generating aerodynamic databases for each major design iteration, and conducting extensive probabilistic risk assessment models for vehicle ascent. Each simulation requires about a week of computing time using up to 500 Pleiades cores.

Figure 3. Image from a SLS launch configuration simulation showing an exhaust plume structure.

Climate Modeling: The consortium for Estimating the Circulation and Climate of the Ocean (ECCO), a joint venture led by Massachusetts Institute of Technology (MIT) and NASA's Jet Propulsion Laboratory, develops advanced computational tools to study ocean currents and their interactions with Earth's atmosphere, sea ice, and marine terminating glaciers [4]. The objective is to help monitor and understand the ocean's role in climate variability and change, and to improve the representation of ocean-climate interactions in Earth system models. These computations involve billions of observations and control parameters, and trillions of predictive model variables—numbers that will only rise as the resolution of ocean models increases and new data is collected—making the role of supercomputers even more vital to the process

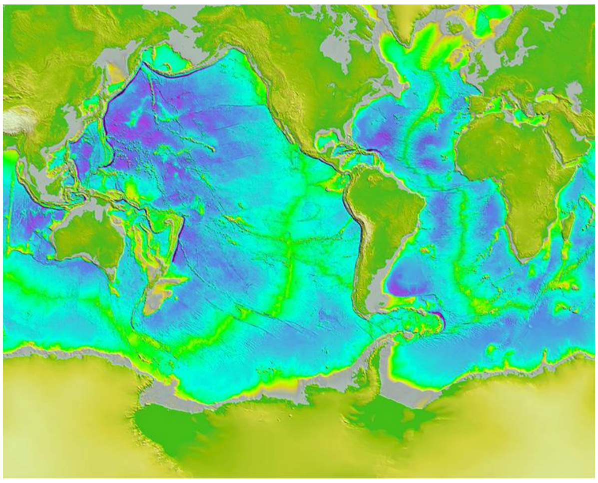

Figure 4. Image from a simulation showing ocean depth (bathymetry).

of creating, storing, and analyzing global ocean circulation. NAS works closely with scientists to optimize the MIT General Circulation Model (MITgcm) code and the ECCO infrastructure to work more efficiently on Pleiades. Typical simulations (Figure 4) use about 252,000 processor hours on Pleiades per month.

3 DATA INTENSIVE SUPERCOMPUTING ENVIRONMENT

NASA is the curator and host for large quantities of observation data, estimated to be multiple petabytes annually. Similarly, large-scale simulations on supercomputers are produce terabytes and sometimes petabytes of data from each run. HECC supports a user base that is at the forefront of data intensive and data driven science. Our users' codes utilize and generate very large datasets and analyzing these datasets to extract knowledge is a fundamental part of their workflows. With NASA data, discovery is challenging because in addition to sheer volume, there is a wide variety of scientifically complex data residing in many distributed archive centers. As a result, users often have to move and manage very large amounts of data using sophisticated tools. Our goal is to provide them with easy access to exascale processing platforms that meet both their computational and analytic needs.

Recently NAS surveyed several HECC users to better understand the challenges faced in their analysis efforts. Guided by the results of the survey [5] as well as those from other studies addressing the convergence of Big Data and supercomputing, NAS is developing a Data Intensive Supercomputing Environment (DISE). The goal is to address the problems of discovery of data and tools, access to and movement of data, storage and management of data, scaling of the analysis and analytics algorithms/tools, workflow execution and management, and dissemination of the results. DISE will also provide the heterogeneous HPC resources and adequate data storage to support all the workflow components of a typical NASA scientist. The main focus is to co-locate the big data with big compute systems to reduce transfer time and support analytics across multiple, large data sources. In addition, the environment will help reduce scientists' time and effort in implementing and executing data-intensive applications, and will reduce software development costs by eliminating software silos.

4 COLLABORATIVE ENVIRONMENTS

The exponential increase in the volume of Earth science observational data poses a serious challenge to scientists who spend an ever-increasing percentage of their time just to compile the long-term, global datasets needed to understand ongoing changes in the Earth system. The NAS Division hosts the NASA Earth Exchange (NEX), a collaborative scientific research platform that enables enhanced and more efficient use of Earth observations in order to understand how our planet is changing [6]. NEX combines Earth system modeling, remote sensing data, and state-of-the-art supercomputing in a collaboration environment for conducting research in Earth science. Using this platform, users can explore and analyze large Earth science data sets, run modeling codes, collaborate on new or existing projects, and share results within or among communities. NEX provides access to approximately 1.2 PB of Earth science data via a three-tiered infrastructure consisting of a collaboration portal, a sandbox environment, and high-performance computing resources. For the 400 registered users at this time, NEX reduces the need for duplicated efforts in downloading data, developing pre-processing software tools, and expanding local compute infrastructures while accelerating fundamental research, developing new applications, and reducing project costs.

5 QUANTUM COMPUTING

The core of many real-world problems consists of extremely difficult combinatorial optimization challenges that cannot be solved accurately by classical supercomputers in a realistic time frame. Such problems frequently occur in the areas of image classification, complex decision systems, machine autonomy, verification and validation, planning and scheduling, and others. A key reason for this difficulty is that classical computing systems as they are architected today are not designed to handle the exponential growth in computational complexity as the number of variables increases. In contrast, quantum computers use a physical evolution process that harnesses quantum mechanical effects.

NASA recently established the Quantum Artificial Intelligence Laboratory (QuAIL) to explore the potential for quantum computers to tackle optimization problems that are difficult or impossible for traditional supercomputers to handle [7]. Through an innovative partnership with Google and Universities Space Research Association (USRA), QuAIL houses a 512-qubit D-Wave Two system at the NAS facility. The research objectives are to determine whether this system can solve certain problems of interest to NASA using quantum annealing. We have conducted extensive benchmarking of the system and some early results from various application areas such as mission planning and scheduling, system diagnostics and anomaly detection, and knowledge extraction from data fusion are expected later this year.

REFERENCES

[1] R. Biswas, et al., “Pleiades: NASA’s First Petascale Supercomputer,” *Contemporary High Performance Computing: From Petascale toward Exascale*, edited by J. Vetter, Chapman and Hall/CRC, 2013, pp. 309–338.
[2] N. Chaderjian, “Advances in Rotor Performance and Turbulent Wake Simulation Using DES and Adaptive Mesh Refinement,” 7th International Conference on CFD, July 2012.
[3] J. Blevins, J. Campbell Jr., D. Bennett, R. Rausch, R. Gomez, C. Kiris, “An Overview of the Characterization of the Space Launch System Aerodynamic Environments,” 52nd AIAA Aerospace Sciences Meeting, AIAA 2014-1253, Jan 2014.
[4] Estimating the Circulation & Climate of the Ocean (ECCO), http://www.ecco-group.org/
[5] P. Mehrotra, L. H. Pryor, “Supporting ‘Big Data’ Analysis and Analytics at the NASA Advanced Supercomputing (NAS) Facility, NAS Technical Report 2014-02, Jan 2014.
[6] NASA Earth Exchange (NEX), https://nex.nasa.gov/nex/
[7] Quantum Artificial Intelligence Laboratory (QuAIL), http://www.nas.nasa.gov/quantum/