Impact of the Columbia Supercomputer on NASA Space and Exploration Missions

Rupak Biswas†, Dochan Kwak†, Cetin Kiris†, Scott Lawrence‡

†NASA Advanced Supercomputing Division
‡Systems Analysis and Integration Branch
NASA Ames Research Center, Moffett Field, CA 94035
{Rupak.Biswas, Dochan.Kwak, Cetin.C.Kiris, Scott.L.Lawrence@nasa.gov}

Abstract

NASA’s 10,240-processor Columbia supercomputer gained worldwide recognition in 2004 for increasing the space agency’s computing capability ten-fold, and enabling U.S. scientists and engineers to perform significant, breakthrough simulations. Columbia has amply demonstrated its capability to accelerate NASA’s key missions, including space operations, exploration systems, science, and aeronautics. Columbia is part of an integrated high-end computing (HEC) environment comprised of massive storage and archive systems, high-speed networking, high-fidelity modeling and simulation tools, application performance optimization, and advanced data analysis and visualization. In this paper, we illustrate the impact Columbia is having on NASA’s numerous space and exploration applications, such as the development of the Crew Exploration and Launch Vehicles (CEV/CLV), effects of long-duration human presence in space, and damage assessment and repair recommendations for remaining shuttle flights. We conclude by discussing HEC challenges that must be overcome to solve space-related science problems in the future.

1. Introduction

NASA’s Columbia supercomputer is a 10,240-processor Linux-based SGI Altix cluster that has increased the space agency’s computing capability tenfold and revitalized its high-end computing (HEC) effort. Constructed in just four months at the NASA Advanced Supercomputing (NAS) facility at Ames Research Center, Columbia has enabled U.S. scientists and engineers to perform significant, breakthrough simulations since it became fully operational in October 2004. Performing at a peak of 62 teraflops (Tflop/s), Columbia has amply demonstrated its capability to accelerate NASA’s Vision for Space Exploration, which seeks to establish a sustained human presence in the solar system, and to make new discoveries and develop revolutionary technologies and capabilities to benefit humanity.

Columbia is having a significant impact on all four of NASA’s Mission Directorates, particularly in Space Operations and Exploration Systems. Among the many projects being computed on Columbia, scientists have conducted high-fidelity computations to analyze the Space Shuttle Main Engine (SSME) liquid hydrogen flowliner. Results are being used to identify the root cause of the flowliner crack problem and to obtain more accurate flight rationale for the shuttle’s return to flight.

Researchers are also running high-fidelity computational fluid dynamics (CFD) codes on Columbia for future space vehicle designs, including the Crew Exploration and Launch Vehicles (CEV/CLV), and are building realistic models to simulate flight risks for these new spacecraft. Risks and performance issues during both the ascent and entry/descent/landing phases are being carefully analyzed.

NASA scientists are also employing the Columbia system for developing comprehensive, integrated models of human body functions during and post-flight recovery, and to develop countermeasures to protect astronauts. These Digital Astronaut models are particularly important to understand human physiological behavior for long-duration space missions.

The integrated HEC environment in which Columbia operates contributes significantly to the successful results obtained on this 20-node supercluster, currently ranked the fourth fastest in the world [1]. This environment is comprised of massive storage and archive systems, high-speed networking, high-fidelity modeling and simulation tools, application performance op-
timization, and advanced data analysis and visualization.

While the Columbia supercomputer represents a remarkable leap in NASA's HEC resources, many space-related science challenges remain that cannot be solved without petascale-size systems. In this paper, we discuss three application areas that will benefit greatly from next-generation technology and methods.

2. Columbia System Description

NASA's Columbia supercomputer is a 10,240-processor system composed of twenty 512-processor nodes, twelve of which are SGI Altix 3700s, and the remaining eight are double-density SGI Altix 3700 Bx2's. Each node is a shared memory, single-system-image (SSI) system, running the Linux operating system. Four of the Bx2 nodes are tightly linked to form a 2048-processor shared memory environment.

Each processor in the 2048-CPU subsystem is an Intel Itanium 2, running at 1.6 gigahertz (GHz), with 9 megabytes (MB) of level 3 cache (the "Madison 9M" processor), and a peak performance of 6.4 gigaflops (Gflop/s). There is a total of 4 terabytes (TB) of shared memory, or 2 gigabytes (GB) per processor. One other Bx2 node is equipped with these same processors. The remaining 15 nodes have Itanium 2 processors running at 1.5 GHz, with 6 MB of level 3 cache, and a peak performance of 6.0 Gflop/s. All these nodes also have 2 GB of shared memory per processor.

Within each node of Columbia, the processors are interconnected via SGI's proprietary NUMAlink fabric. The 3700's utilize NUMAlink3 with a peak bandwidth of 3.2 GB/s. The Bx2's have NUMAlink4 where the bandwidth is doubled to 6.4 GB/s. The 20 nodes are connected to each other by Voltaire InfiniBand fabric, as well as via 10- and 1-gigabit Ethernet connections. The four Bx2 nodes in the 2048-CPU subsystem use NUMAlink4 among themselves as well as the other fabrics. Columbia is connected to 440 TB of online RAID storage through a Fibre Channel switch. This capacity is being upgraded to almost 1 petabyte (PB). The archive (tape) storage capacity is 10 PB.

On each 512-processor node, the primary features and benefits are as follows:

- Low latency to memory (less than 1 microsecond), which significantly reduces the communication overhead;
- High memory bisection bandwidth, Columbia being the first system (in November 2004) to exceed 1 TB/s on the STREAM benchmark [2];
- Global shared memory and cache-coherency, which enables simpler and more efficient programming paradigms than message passing;

- Large shared memory (1 TB), which allows bigger problems to remain resident on the system.

These features make Columbia particularly well suited for large-scale compute- and data-intensive applications. Typical problems are physics-based simulations involving a discretized grid of the physical domain that is partitioned across multiple processors. In addition, applications requiring dynamic load balancing and/or adaptive gridding are much easier to control on Columbia, leveraging shared memory programming models such as OpenMP and MLP [3].

The development and operating environment on Columbia features a 64-processor SGI Altix front-end, a Linux-based operating system, Altair PBS Professional job scheduler, Intel Fortran/C/C++ compiler, and SGI ProPack software.

3. Integrated HEC Environment

The NAS facility's integrated HEC environment provides resources, capabilities, and services that enable NASA scientists and engineers to make highly effective use of the Columbia system. This supercomputing resource has accelerated technology developments in several scientific areas and is used to conduct rapid engineering analysis to reduce design cycle times and cost. Columbia enables the NASA user community to model and analyze data up to 10 times faster than before, and to view their results at much higher fidelity. Dramatic improvements in turnaround time and accuracy mean that answers to problems that were until recently considered unsolvable are now feasible within reach [4].

But true high performance computing is not enabled only by supercomputing hardware resource such as the Columbia system. The NAS facility therefore provides a fully integrated HEC environment that is comprised of massive storage and archive systems, high-speed networking, high-fidelity modeling and simulation tools, application performance optimization, and advanced data analysis and visualization. The NAS staff has world-class expertise in these wide-ranging areas to provide an environment that enables NASA to accelerate vehicle design cycle time, conduct extensive parameter studies of multiple mission scenarios, facilitate scientific discovery, and increase safety during the entire life cycle of space exploration missions.

3.1. System Technologies

The NAS HEC team of engineers and computer scientists conduct supercomputer system analyses based on several factors such as NASA mission needs, overall system performance, code compatibility and
interoperability, critical user applications, and memory and data storage capacity for system balance. Results are used to determine new system architectures and to specify the requirements for facility infrastructure improvements. To protect the integrity of users’ data and applications, NAS security experts analyze customer requirements, then design and implement custom, state-of-the-art security models that meets those requirements.

3.2. Network Design

NAS network engineers design and implement high-speed local and wide area networks that allow geographically distributed NASA users to efficiently access Columbia and rapidly transfer data to and from their local sites. Work in this area includes high-level and detailed network architecture design and engineering; hardware selection, testing, installation and configuration; and security testing. These engineers are experts in a wide range of protocols and applications. They also employ sophisticated tools to diagnose and tune end-to-end user application performance.

3.3. Modeling and Simulation Tools

Coupled with the HEC resources at NAS, we develop simulation and modeling tools to help NASA achieve scientific goals not previously attainable. For example, our scientists have created molecular-level models, using Columbia and parallel software, to develop a sophisticated biochemical model of DNA damage caused by space radiation. Scientists also run various six-degree-of-freedom simulations to complete key foam debris analyses for shuttle launches. In the future, we will be able to make these simulations so fast that a complete flight can be simulated in real time.

3.4. Code Parallelization and Optimization

The NAS team ports complex modeling codes and performs sophisticated parallel performance analysis and optimization to enable significantly increased performance on the Columbia architecture. Researchers are now getting results in hours and days instead of months. For example, the performance of a magnetohydrodynamics code used to study aspects of planet formation and accretion of gas onto its central star was improved dramatically using the NAS-developed CAPO parallelization and optimizing tool [5]. These code improvements affect not only computational performance but also the fidelity of the models themselves.

3.5. Data Analysis and Visualization

The NAS facility’s visualization staff designs and implements methods and environments that provide scientists advanced capabilities in viewing and steering their computational data and simulations, such as space shuttle debris analyses and nuclear combustion in supernovae. The hyperwall graphics hardware helps researchers display, analyze, and study ultra-large high-dimensional datasets in meaningful ways, allowing the use of different tools, viewpoints, and parameters. We research and apply advances in IT to enhance the understanding of computational simulation and experimental data. The staff also consults with and assists researchers in creating animations, videos, and acoustical data to help analyze and present scientific results.

4. Space Mission Applications

The integrated Columbia supercomputing environment provides key support for all four of NASA’s Mission Directorates: Aeronautics Research, Science, Space Operations, and Exploration Systems. As a National Leadership Computing System (NLCS), Columbia’s highly advanced architecture is also available to a broader science and engineering community to solve research problems of national interest.

Within the Space Operations and Exploration Systems Mission Directorates (SOMD/ESMD), we illustrate in this paper the impact of Columbia with a sampling of three recent applications areas: the SSME flowliner analysis for the Space Shuttle’s Return-to-Flight, CEV/CLV simulation assisted risk assessment, and blood circulation in the human brain under altered gravity.

4.1. Space Shuttle Main Engine Flowliner

NASA’s SOMD focuses on providing critical capabilities that enable human and robotic exploration missions in and beyond low-Earth orbit, including the International Space Station and the Space Shuttle. SOMD is also responsible for leadership and management of the Agency’s space operations related to launch services, space transportation, and space communications. In this section, we discuss the computational analysis of cracks in the Space Shuttle Main Engine (SSME) turbopump.

In May 2002, numerous cracks were found in the SSME#1 flowliner; specifically, at the gimbal joint in the liquid hydrogen (LH2) feedline flowliner. While repairs were made to existing cracks on all orbiters, scientific investigations have continued because the actual cause of the original cracks was not conclusively
established, and remaining shuttle flights require long-term investigations into the root cause of the problem.

Recently, high-fidelity computations were conducted on Columbia to analyze the SSME LH₂ feedline flowliner [6]. Various computational models were used to characterize the unsteady flow features in the turbopump, including the Low-Pressure-Fuel-Turbopump (LPFTP) inducer, the orbiter manifold, and an experimental hot fire test article used to represent the manifold. Unsteady flow originating from the LPFTP inducer is one of the major contributors to high frequency cyclic loading that results in fatigue damage to the flowliners.

The flow fields for the orbiter manifold and the hot fire test article were computed and analyzed on Columbia for similarities and differences using an incompressible Navier-Stokes flow solver, INS3D, to compute the flow of liquid hydrogen in each test article for two computational models [7].

The first computational model included the LPFTP inducer; by studying the inducer model alone, scientists were able to compare unsteady pressure values against existing data. To resolve the complex geometry in relative motion, an overset grid approach [8] was employed, which contained 57 overlapping zones with 26.1 million grid points.

The second computational grid system added the flowliner geometry. This grid system, which was very similar to the ground test article, consisted of 264 overset grids with 65.9 million grid points, and is shown in Figure 1. The flowliner component alone contained 212 grids and 41 million points.

To speed up and automate the grid generation procedure, scientists at the NAS facility developed scripts to perform the various steps in the grid generation process prior to use of the INS3D flow solver. They also developed special procedures to automatically create grids for each component type.

Two parallel programming paradigms were implemented into the INS3D code: the Multi-Level Parallelism (MLP) [3] and the hybrid MPI+OpenMP models. Multiple-node computations showed that point-to-point implementation of the MPI+OpenMP code performs more efficiently than the master-worker version of the MPI code [9]. This turbopump application currently exhibits some of the best scalability on the Columbia system.

Results of the CFD calculations confirmed the presence of backflow, caused by the LPFTP inducer. This is illustrated in Figure 2. The region of reverse flow extended far enough upstream to interfere with both flowliners in the gimbal joint. Computed results for the test article were verified by correlation with pressure measurements, and confirmed a strong unsteady-interaction between the backflow caused by the LPFTP inducer and secondary flow in the bellows cavity through the flowliner slots. It was observed that a swirl on the duct side of the downstream flowliner is stronger than on the same side of the upstream flowliner—causing more significantly stronger unsteady interactions through the downstream slots than those observed in the upstream slots. Despite the progress made in extremely complex simulations such as these, it crucial to further extend these models for flight rationale.

4.2. Crew Exploration Vehicle Risk Assessment

NASA's ESMD is developing a collection of new capabilities and supporting technologies, and conducting foundational research to enable sustained and affordable human and robotic exploration. As part of this work, Columbia is being used for the high-fidelity modeling and simulation of next-generation spacecraft and launch vehicles design. NASA is particularly focused on building the Crew Exploration Vehicle (CEV), the first new U.S. human spacecraft in 30 years. The CEV will transport a maximum of six crew members to and from the International Space Station and up to four astronauts to and from the moon, with the first CEV mission planned around 2010, when the
shuttle will be retired. The CEV will also support future Mars missions.

The CEV design includes Launch Abort System (LAS) for crew escape, similar to that used in other spacecraft, including the Apollo capsule. Several computational modeling and simulation tools suitable for analyzing abort scenarios have recently been developed and enhanced for use on Columbia. Staff at NASA’s Ames and Glenn Research Centers have collaborated on this work under the Simulation Assisted Risk Assessment (SARA) project. The SARA team developed a Probabilistic Risk Assessment (PRA) approach and demonstrated how risk analysis can be applied to launch abort using the Apollo configuration [10]. A PRA identifies the best level of fidelity for modeling critical failure modes associated with launch abort. Columbia was then used to conduct higher-fidelity modeling on specific failure modes. Two failure modes considered during the past year included booster explosion and those caused by re-contact with the booster during separation. Each of these modes required the application of high-fidelity aerodynamic simulation.

One of the more prominent failure modes analyzed (using Apollo data) was a possible catastrophic failure of the booster, leading to detonation of the propellant, which create blast wave overpressures that could fatally damage the LAS. This is illustrated in Figure 3.

As the risk model was being developed, it became clear that the exact nature and magnitude of the explosion is a contributor to abort risk. In other words, the type of booster and the nature of the failure it was likely to encounter determined the environments under which the crew escape system must operate to ensure a successful abort. The process for characterizing this interaction has to be carefully modeled and simulated.

One of the weaknesses found in an engineering-level model was the effect of headwind as the CEV ascends. In order to account for these effects in the risk analysis, high-fidelity blast wave models were built and simulated on Columbia using the Overflow Navier-Stokes code [11]. Results indicated that headwinds significantly affect the nature and magnitude of the shock wave as it impacts an escaping CEV. Therefore, the warning time required to initiate the abort sequence is also affected. Additional work in high-fidelity simulations remains to help engineers generate requirements for the LAS.

Another failure mode dependent on high-fidelity simulation involves the ability of the LAS to achieve “clean” separation of the CEV from the booster stack in the event of impending catastrophic failure; in other words, the CEV must not scrape or re-contact the booster stack. This failure mode was especially demanding because it involved complex proximity aerodynamics—modeling transonic flow and the complex
flow at the small gap (or cavity) between the CEV and booster stack at separation. Both Navier-Stokes simulations, using Overflow, and Euler simulations, using FlowCart [12] were applied, and their results validated against transonic wind tunnel and abort flight test data from the Apollo era [13].

All these cases are computationally expensive to simulate. A single steady-state simulation required approximately 3,500 processor-hours on Columbia. The complexity of the geometry and the flow-field required about 30 million grid points, which enabled good scalable performance up to at least 250 CPUs. In all, approximately 20 cases were computed using Overflow at various ascent trajectories and separation thrust levels. Including the computation of the initial static solution, each case required approximately 20,000 processor-hours on Columbia. All failure modes benefited greatly from the HEC resources at the NAS facility. These tools and processes, utilizing the Columbia resources, will likely be applied to analyze the actual LAS design, and to further understand the CEV failure modes and their impact on the vehicle’s survivability.

4.3. Human Presence in Space

NASA’s ESMD is also responsible for conducting biological and physical research necessary to ensure the health and safety of crew during long-duration space flight. Astronauts in space are exposed to hostile environments such as radiation and altered gravity. Scientists at the NAS facility are therefore using the Columbia system to develop biomedical simulation tools to study the impacts of altered gravity on astronaut health and performance, and the long-term impact on humans in space.

For astronauts, blood circulation and body fluid distribution undergo significant changes due to the stressful environment of microgravity, both during and after space flights. Many studies on physiological changes under weightlessness have been performed (especially in conjunction with the Space Shuttle program), including diverse physiological functions affected by the nervous system such as heart rate, blood pressure, hormone release, and respiration. In particular, altered blood supply to the brain and consequent delivery of oxygen to certain parts have non-trivial impact on the health and safety of astronauts, making it essential to understand what happens to arterial wall mechanics and resulting blood flow patterns under microgravity.

Researchers at NASA have initiated development of a computational model of the human circulatory system as a medium for connecting various biomedical functions of astronauts in space. Various arterial network models were developed; the high-fidelity CFD was then coupled to study local regions of interest, such as branches of the carotid artery and in the brain arteries under altered gravity. Physical models required for CFD simulation were introduced, including: a model for arterial wall motion due to fluid-wall interactions; a shear thinning fluid model of blood; and a model for the auto-regulation mechanism.

The high-fidelity local analysis is illustrated in Figure 4 using blood flow through the Circle of Willis (COW), the circle of arteries at the base of brain that supplies blood to its various parts. This computational approach was applied to localized blood flow through a realistic carotid bifurcation and then connected to the COW using a geometry obtained from an anatomical data set. A three-dimensional COW configuration was reconstructed from human-specific magnetic resonance images using an image segmentation method.

Through the numerical simulation of blood flow in the model problems, it was observed that altered gravity has considerable effects on arterial contraction and/or dilatation and consequent changes in flow conditions. The resulting numerical procedure was validated using laboratory test cases. Based on the consistency and good quantitative results produced by this procedure, this computational model was then extended to study blood circulation under altered gravity.

Figure 4. Human-specific geometry of the cerebral arterial tree reconstructed from magnetic resonance images (left) is used in conjunction with supercomputing technology to establish large-scale continuum fluid simulations (right).

The computed results [14] have provided evidence that altered gravity has considerable effects on arterial contraction and/or dilatation and resulting flow patterns. In future work, studies will be performed on the Columbia supercomputer to better understand the quantitative influence of altered gravity on the entire human circulatory system. Although the simulation procedure was validated using simple test models, scientists are very interested in being able to compare...
high-fidelity simulation results based on clinical geometry to clinical data in the future, eventually leading to the development of a whole-body algorithm for the Digital Astronaut program.

5. Emerging IT Challenges

Many of Columbia’s scientific and engineering users have stated that the system has allowed them to successfully complete investigations they never allowed themselves to dream of previously. Now, these users are envisioning what they can accomplish when even more powerful computing systems are available. NASA has already begun planning its next-generation supercomputer to meet the ever-increasing demand for computational resources required for breakthrough space science discoveries.

Obstacles to reaching these goals are wide ranging, and include not just hardware limitations, but network communication between nodes, scalable algorithms, and computer science tools, among others. With the Digital Astronaut program, for example, increased computational capabilities will enable scientists to extend simulations from just the arterial system to the entire body, and couple it with other systems such as the respiratory system. It will be feasible to bridge between the macroscopic and microscopic (molecular) scales, thereby extending studies from the capillary to the cell level, all ultimately leading to accurate prediction of astronauts’ performance during long-duration space flights.

5.1. Computing Hardware and Storage

Well-known scalability bottlenecks exist in high-performance computer hardware; in a recent report [9], it was observed that performance scalability is affected on the Columbia system when communication occurs between multiple boxes. With access to a system having 10 times the number Columbia’s processors, we could, for example, increase the fidelity of the current propulsion subsystem analysis to full-scale, multi-component, multidisciplinary propulsion applications, including modeling systems for new and existing launch vehicles to reach full flight rationale. But sophisticated domain decomposition, optimized sorting, and latency-tolerant algorithms, as well as faster interconnects are necessary to make these applications scalable to very large processor counts.

As the amount of data computed on HEC systems increases, the issue of managing, archiving, and retrieving this data will become a serious challenge. The NAS facility now stores over 1 TB of data per day. We are examining not just the day-to-day storage requirements, but also the overall strategy to plan for and ensure enough data storage space is available to meet our users’ requirements. Also, as vendors move to multicore processors, the issue of getting the data to memory becomes an even bigger challenge.

5.2. Modeling and Simulation

While the current NAS HEC environment includes high-fidelity modeling and simulation, taking it to an integrated system level will require the combination of existing simulation software frameworks and the development of new advanced physics-based predictive technologies. Multidisciplinary computations are critical for expanding the current operational envelope, for developing revolutionary new concepts for vehicles and propulsion systems, and for planning an entire exploration mission. The need for modeling, computing, and analyzing various scenarios is rapidly increasing the demand for large parallel computing resources to achieve efficient solution turnaround time.

The physical complexity of modeling complete multi-component systems with stationary and moving sections is not the only limiting factor; computational complexity also exists, including factors such as: the scalability of highly parallel algorithms that do not sacrifice the robustness or convergence of the underlying numerical scheme; the ability to efficiently handle multiple physical length and time scales through adaptive remeshing; and producing stable and accurate numerical boundary condition treatments which are robust for many physical problems.

5.3. Performance Enhancement Tools

At present, user tools such as code parallelizers, debuggers, and performance analyzers to make the best use of supercomputing resources are acutely lacking in the HEC market. The NAS staff is evaluating vendor products and helping to improve them, but industry needs to accelerate their efforts to create more useful tools. Static and dynamic dependence analysis, relative debugging, and the ability to gather and analyze trace data are some examples of near-term IT challenges. Development of appropriate programming paradigms and techniques for current and future NASA supercomputing systems, and efficiently porting some of the Agency’s most important mission-critical applications to demonstrate high levels of sustained performance is vital to the sustained success and impact of HEC. Tools that can enhance user productivity and code maintainability are also critical in reducing the total life cycle cost of NASA applications and increasing their return on investment.
6. Summary and Conclusions

The 10,240-processor Columbia supercomputer has dramatically improved NASA’s high-end computing (HEC) capability and capacity. Enabled by the NASA Advanced Supercomputing (NAS) facility’s multidisciplinary integrated environment, U.S. researchers are obtaining results heretofore deemed impossible. The Columbia HEC environment is a shared-capacity asset for the agency, and supports all four NASA Mission Directorates: Aeronautics Research, Science, Space Operations, and Exploration Systems.

In support of the agency’s Space Operations and Exploration Systems Mission Directorates, NAS researchers have utilized Columbia to identify the root cause of the SSME flowliner crack problem to obtain a flight rationale for future shuttle missions. They have developed modeling and simulation tools to analyze risks for crew escape on future space flight missions. Simulations tools are also being developed to perform high-fidelity analysis of the human circulatory system, so as to determine the short- and long-term impacts of altered gravity on astronauts in space. These three applications, highlighted in this paper, represent a small portion of the space mission-related work being conducted on Columbia.

- The Columbia supercomputer was critical to NASA’s Return to Flight effort, where scientists performed six-degree-of freedom foam debris transport analyses and visualization to forecast space shuttle damage, and for damage assessment and repairs during the successful Space Shuttle Discovery flight in July 2005. In addition, many Science areas, such as global ocean modeling, hurricane track prediction (including Hurricane Katrina), cosmology, and astrophysics, as well as fundamental Aeronautics Research in the subsonic, supersonic, and hypersonic regimes depend heavily on the integrated Columbia HEC environment.

Significantly more can be accomplished when even more powerful computing systems become available. However, obstacles to reaching these goals are wide ranging, and include many IT challenges that must be overcome. Barriers such as those mentioned here must be removed before efficient use of greater computational resources can be achieved and HEC can have an even bigger impact on NASA space science missions.

7. References


