Fall 2014

DOD HPC INSIGHTS

A publication of the Department of Defense High Performance Computing Modernization Program

Getting the Latest from Space To Win on the Ground

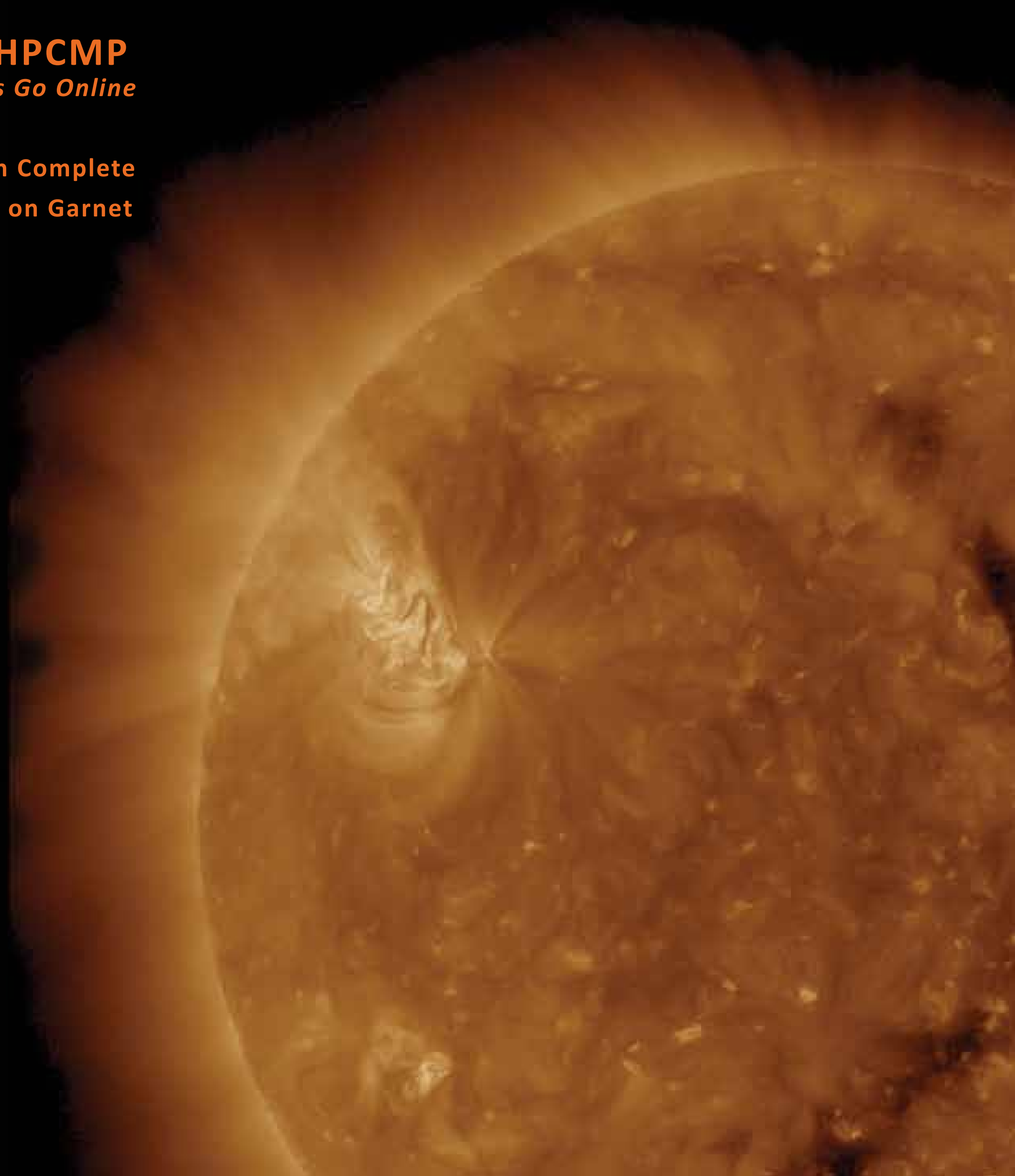
Advanced Solar Weather Modeling

New to the HPCMP

Cray XC30 Systems Go Online

DREN III Transition Complete

New Queue Policy on Garnet



Contents

First word

Cover Photo:
Observed solar emission recorded by the Solar Dynamics Observatory on July 11, 2010. Solar emissions can disrupt terrestrial and near-earth communications and interfere with DOD assets on the ground and in space.

First word

John E. West
HPCMP Director

In the months following the transition of the High Performance Computing Modernization Program from the Office of the Secretary of Defense to the Army, the HPCMP team—that's all of us, including users and our leadership in the five-sided building—spent time reflecting our successes over the first two decades of this program and studying our current structure to ensure we are positioned for another two decades of success. The result of this exercise was a set of five execution imperatives that have guided the first three years of Army leadership of this program:

1. *Maintain leadership in the provision of HPC software, hardware, and expertise to the RDT&E community of the Department of Defense*
2. *Increase our impact on the mission of the department by serving as the bridge for application of HPC to emerging challenges in decision support, acquisition engineering, test and evaluation, and new science and technology efforts*
3. *Increase the productivity of the DOD workforce as they use our systems, services, and expertise*
4. *Lead the way in the application of promising new technologies to DOD HPC challenges*
5. *Develop and mature a DOD workforce prepared to take full advantage of HPC technologies*

Despite the intense organizational, financial, and operational challenges we have faced over the past three years, we have managed to make significant progress in these areas.

Our position as the leading provider of HPC in DOD is strong. Our centers are flourishing today, deploying a mix of both very large-scale and intermediate-sized purpose-built supercomputers that provide a computational capability specifically suited to the DOD's application needs, and unavailable anywhere else in the private or public sectors. By the end of 2015 we will have increased our aggregate computational capability to well over 20 petaFLOPS. With the successful transition to DREN III this year, we have quadrupled the capabilities of the network that connects us all together. And at the end of FY14 the centers are launching a new initiative to refine and focus their capabilities and deployed machines to match your requirements even more closely, and to serve the current and emerging needs of the department even better. Look for more on this initiative in the coming months.

Our impact on the DOD is growing in exciting ways, led today by the CREATE program. A large part of the HPCMP's software investment, CREATE is developing complex, multiphysics applications that bring the fidelity of HPC simulations to the design and development of next-generation warfighting platforms. CREATE today is being applied to over 70 air, sea, and land weapon systems and platforms. Along with our deep partnership with the department's Engineered Resilient Systems initiative, CREATE and our HPC capabilities are helping to bring powerful new tools to the maintenance of military systems today and to the future of warfighting tomorrow.

Increasing the productivity of the DOD's users of HPC remains a focus of this program, and it is baked into everything we do. Part of the centers initiative I mentioned above will include an examination of the ways in which we deploy new hardware—and how that hardware is acquired, configured, and deployed—to ensure that when you are ready to work our machines are up, running, and ready for you. Our focus on productive use of HPC by nonspecialists using CREATE tools has also led to a multiyear investment in Web-enabled access to supercomputers and applications, and broad deployment of Web portal capabilities into the HPCMP centers began earlier this year.

(continued on page 28 "First Word")

Advanced Solar Telescope Weather Modeling

by Skip Williams; AFRL/RDSM Chief Engineer
R. Bruce Duncan; Boeing Company
Jon Linker, Cooper Downs, Roberto Lionello, and Janvier Wijaya; Predictive Science Inc.
C. Nick Arge and Carl Henney; AFRL/RVBXS

A team led by the AFRL/RD Directed Energy Directorate is performing software applications engineering development on advanced solar and corona modeling for improved space weather awareness and forecasting. This development has utilized state-of-the-art solar modeling and previous software applications and data. Selected previous legacy codes have been ported to the MHPCC DSRC Riptide platform, and code modification development, testing, and analysis of modeling results with improved performance have been realized.

Visualization of data provided by NASA's Solar Dynamics Observatory.
The National Solar Observatory (NSO) data used for this work are produced cooperatively by National Science Foundation (NSF) and the NSO. The NSO is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under cooperative agreement with the NSF.

"These transient [solar] phenomena can interfere, disrupt, and harm DOD space and ground-based assets."

High resolution coronal simulation, this work.

Background

Solar flares are localized events that produce intense electromagnetic emissions, including items such as high-energy Extreme Ultraviolet (EUV) and X-rays and energetic particles from the Earth's sun. Coronal mass ejections (CMEs) are large eruptions of magnetized coronal plasma that propagate out into space and accelerate energetic particles along the way. These transient phenomena can interfere, disrupt, and harm DOD space and ground-based assets. Flares and CMEs, for instance, heat Earth's atmosphere causing it to expand and change spacecraft orbits. They also disrupt radio communications between the ground and orbiting satellites. Energetic particles produced by flares, and especially CMEs, can damage vital instruments on satellites, rendering them useless. Since the DOD and US Air Force rely heavily on satellites, a knowledge of (or the potential for) disturbances to the geospace environment resulting from solar activity is critical for Space Situation Awareness (SSA).

While the overarching physical process behind CME and flare generation is thought to be the release of energy stored in the magnetic field, the actual detailed physics behind their eruptions is still poorly understood. State-of-the-art, physics-based numerical magnetohydrodynamic (MHD) models of the corona and solar wind are presently the best available tools to explore the physics behind CME and flare generation. Therefore, this project performed further development, modification, and improvements of physics-based numerical CORona-HELiosphere (CORHEL) MHD models of the corona and solar wind to better explore the physics behind CME and solar flare generation. This project utilized successful CORHEL and the Air Force Data Assimilative Photospheric Flux Transport (ADAPT) software application models and adjusted them for increased modeling performance. The ultimate goal of this R&D thrust is to achieve a routine simulation capability of realistic events using such models within a decade. CORHEL software improvements have been made to provide improved 3-D coronal and solar wind solutions including quantities and emissions of 3-D plasma and magnetic fields. The enhanced ADAPT-CORHEL space weather modeling and solar photosphere map capabilities are striving to enable significantly increased solar event forecasting capabilities for SSA. Inputs, processing, outputs, concept of operations, and objectives for this are depicted in Figure 1.

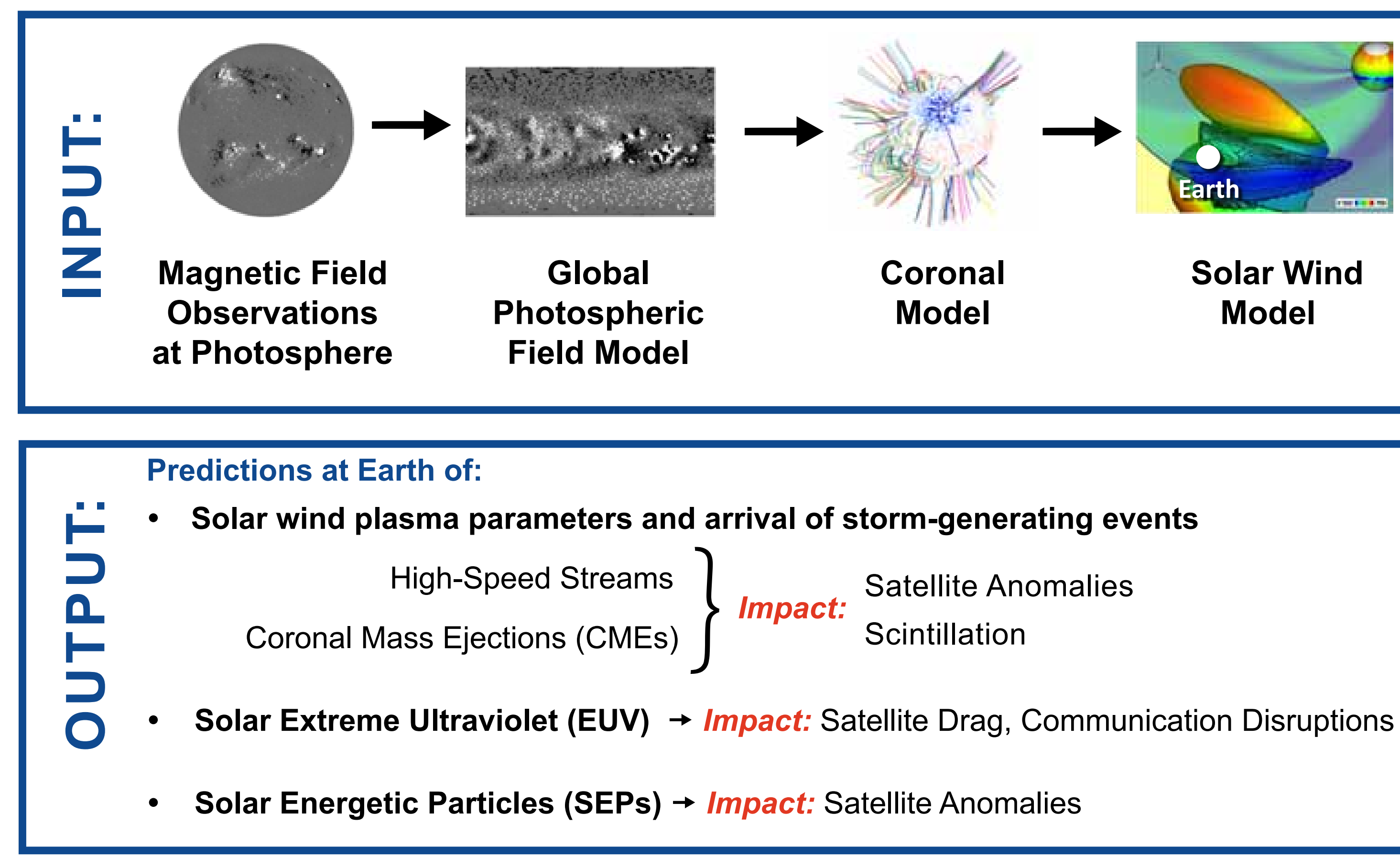


Figure 1. Advanced modeling of the Sun-Earth system for improved Space Situational Awareness

The ADAPT model is used to yield a global solar photospheric magnetic field distribution map and serves as a primary input to coronal and solar wind models, resulting in improved high-quality "snapshots" of the sun's magnetic field. ADAPT evolves the solar magnetic flux using well-understood transport processes where measurements are not available. It updates modeled flux with new observations using data assimilation methods, which rigorously take into account model and observational uncertainties. ADAPT provides more realistic estimates of the instantaneous global photospheric magnetic field distribution than those provided by traditional synoptic maps. CORHEL model software is then used, with principal inputs of global photospheric magnetic field maps, such as those provided by ADAPT. CORHEL is an advanced 3-D magnetohydrodynamic coupled set of models and tools of the ambient solar corona and solar wind. CORHEL outputs 3-D plasma and magnetic field quantities and emission data such as Extreme Ultraviolet (EUV) rays.

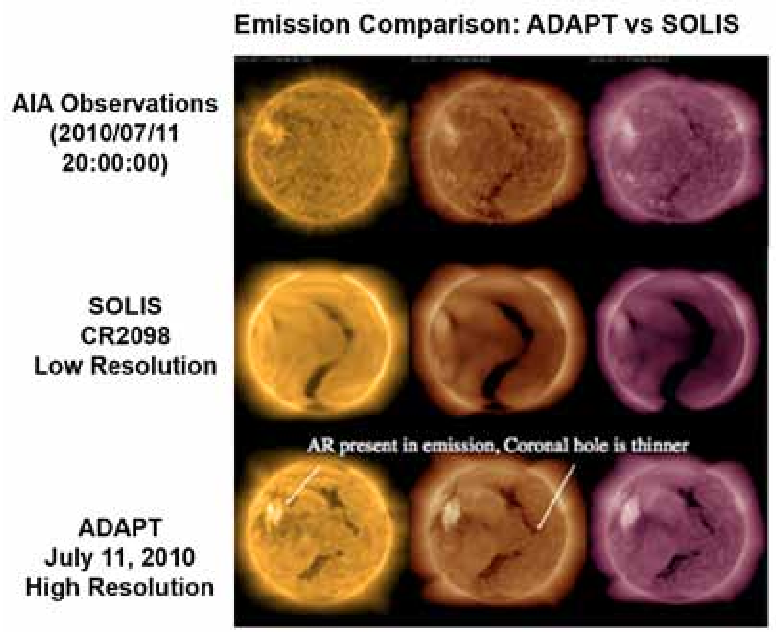


Figure 2. Comparison of observed and simulated emission in the 171Å, 193Å, and 211Å lines of the AIA instrument aboard SDO.

Status and Achievements

Figure 2 shows a summary example of the new capabilities developed as a result of this project. The top row shows observed emission from the Atmospheric Imaging Assembly (AIA) instrument on board the Solar Dynamics Observatory (SDO) recorded on July 11, 2010 at 20:00. Note the active region near the east limb (left-hand side of the solar disk) and the extended coronal hole near disk center. The second row shows the simulated emission from a CORHEL MHD model at moderate resolution using a Synoptic Optical Long-term Investigations of the Sun (SOLIS - magnetograph located at Kitt Peak Observatory, Arizona) synoptic map for Carrington Rotation (CR) CR2098. This is a standard run available with CORHEL at the start of this project. The third row shows a high-resolution CORHEL MHD model computed using a high-resolution ADAPT map for July 11, 2010. These capabilities were added to CORHEL as a result of this project. When the ADAPT map is used for the boundary condition of the simulation, an active region that emerged is visible in emission near the east limb, as is also seen in the observations. The extended coronal hole region near the disk center in the high-resolution model is thin and discontinuous, more closely resembling the observed coronal hole.

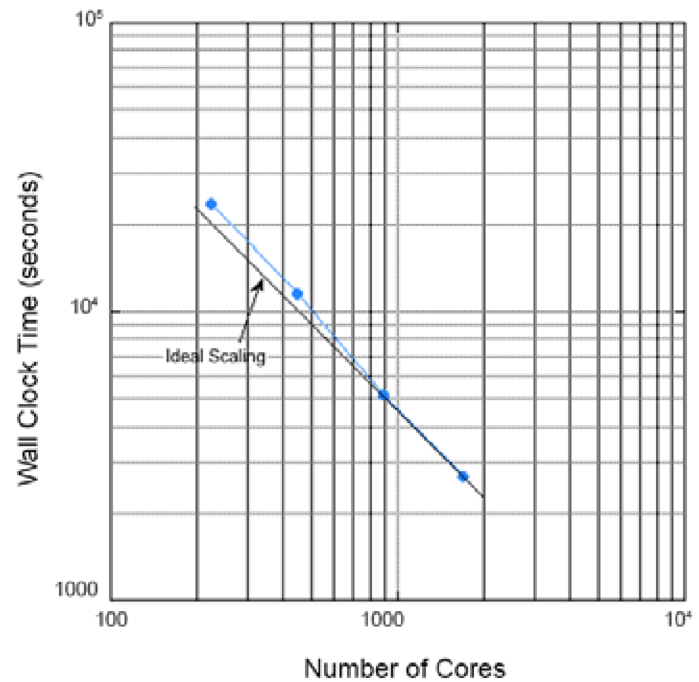


Figure 3. Scaling of CORHEL MAS Code on Riptide

The research and development performed to date has advanced further understanding of this key scientific domain. This further understanding has included utilizing models and simulations with comparisons to other actual solar instruments observations, while realizing enhanced high-resolution vital software applications capabilities. Key parallel code components have provided essentially ideal scaling when applying increasing amounts of supercomputing compute cores to reduce wall clock time as shown in Figure 3, while aiding the delivery of increased resolution modeling results.

Project Achievements

Several significant achievements and capabilities have been realized to date in the development and demonstration of this project. A summary of these include:

- CORHEL has been ported to the MHPCC DSRC Riptide supercomputer.
- CORHEL has been modified to work with ADAPT photospheric maps.
- CORHEL results with ADAPT versus standard observatory maps have been compared, showing the advantages of using ADAPT maps.
- CORHEL has been compared when using low- and high-resolution ADAPT maps, and with high-resolution simulations. High-resolution CORHEL runs with high-resolution ADAPT maps have shown great promise for revealing increased details of coronal structure.
- Different coronal heating models have been compared.
- ADAPT maps have been processed for time-dependent runs, and demonstrated in a prototype Magnetohydrodynamic Algorithm outside a Sphere (MAS) run. MAS is a semi-implicit time-dependent 3-D MHD model of the solar corona and the inner heliosphere.
- CORHEL has been shown to perform distributed parallel processing and run efficiently on MHPCC Riptide, and to scale ideally well at high quantities of processors.
- ADAPT has been modified to incorporate far-side solar data. Initial simulations, with and without far-side data, have been performed and comparisons are ongoing. The team plans to attempt a much more ambitious calculation and additional computational runs, with an unprecedented number of grid points using the high-resolution ADAPT map capabilities.
- The team continues to advance further understanding of this key scientific domain, utilizing modeling and simulations with comparisons to other actual solar instruments observations, while developing and realizing improved high-resolution CORHEL and ADAPT capabilities.

Achievements and Summary

As summarized in this article, the AFRL/RD, the MHPCC DSRC, the AFRL/RV solar project, and the National Solar Observatory (NSO) collaborated to provide an improved ADAPT photospheric mapping model and an improved 3-D MHD CORHEL model. The collaboration focused on the capability of these models to utilize rapid cadence, as well as work with the impending National Solar Observatory (NSO) Daniel K. Inouye Solar Telescope (DKIST) on Maui to receive daily high-resolution vector measurements of active region magnetic fields.

ADAPT and CORHEL have been improved and optimized. ADAPT has been improved to provide high-quality "snapshots" of the Sun's global magnetic fields as a key input into CORHEL, which has high-resolution and time-dependent capabilities. ADAPT also includes helioseismic far-side active regions data (far-side active regions detected using helioseismology) to further advance solar wind predictions.

CORHEL has been improved to provide 3-D coronal and solar wind solutions that include 3-D plasma and magnetic field quantities. The coronal plasma parameters can be computed permitting realistic simulation of Extreme Ultraviolet (EUV) and X-ray emission. The ADAPT-CORHEL integrated flow output can provide predictions at Earth of solar wind plasma parameters and the arrival of storm-generating events, such as high-speed streams. It can also provide background solutions for the simulation of Coronal Mass Ejections (CMEs). These solar space weather events can significantly impact satellites and possibly produce scintillation. Large solar flares increase solar EUV emission, which increases orbital drag. Radio noise or bursts associated with large solar flares can disrupt or degrade communications. Solar Energetic Particles (SEPs) generated in CME shocks can cause satellite anomalies. Accordingly, the new and improved ADAPT-CORHEL capabilities will provide advanced solar modeling and forecasting capabilities for improved DOD space situational awareness (SSA) to help protect the nation's space assets, national defense, and national security. This work is possible only through the continued joint efforts of the above entities.

New to the DOD HPCMP

High Performance Computing Portfolio:

3 Systems, 2.7 PetaFLOPS, and In-Situ Visualization

Three Cray XC30 High Performance Computing (HPC) systems have been installed at the Air Force Research Laboratory (AFRL) and Navy DOD Supercomputing Resource Centers (DSRCs)

"In the month or so that I've had access to Lightning, [Armstrong, and Shepard] I've generated results that would've otherwise taken over a year...I am very pleased with the performance, stability, and usability of the new... XC30 systems."

Jerry Boatz (AFRL/RQ, Aerospace Systems Directorate, Rocket Propulsion Division at Edwards Air Force Base)
HCPCMP user modeling chemical reactions on the surface of pristine aluminum nanoparticles, which are ingredients for new rocket fuels.

"Using the latest Intel compiler [on Lightning] I have been able to run a 2.5 billion element very high-resolution Large-Eddy Simulation (LES), of Mach 2.25 turbulent flow over a 24-deg corner in the presence of both sidewalls. All in all, [Lightning] looks to be a great workhorse for future computational studies."

Nick Bisek (AFRL/RQHF, Aerospace Systems Directorate, Hypersonics Sciences Branch)
HPCMP user modeling hypersonic calculations as part of an HPCMP Frontier Project

Lightning (AFRL DSRC)

Named for the Air Force F-35 aircraft, has 57,200 compute processors and 32 NVIDIA Tesla K40 nodes.

Armstrong & Shepard (Navy DSRC)

Named for NASA astronauts and naval aviators Neil Armstrong and Alan Shepard, are comprised of 30,592 and 30,144 compute processors, respectively; both systems have 124 Xeon Phi accelerator nodes each. Shepard is also outfitted with 32 NVIDIA Tesla K40 nodes.

"Two of my team members...are reporting that their codes are running 3-4x faster on Armstrong than the same codes do on [the Cray XE6] Garnet...or [the IBM iDataPlex] Kilrain."

Dr. Peter Mardahl (AFRL/RD, Directed Energy Directorate, High Power Microwave Division at Kirtland Air Force Base),
HPCMP user involved in the design of a high power microwave sources for electronic warfare and nonlethal weaponry

"On Armstrong we have...measured the highest I/O speeds and highest teraFLOPS-speeds ever seen on any HPC hardware running STRIPE (that includes [existing HPCMP] systems...). For the solution of...a few hundred million degrees of freedom, STRIPE's in-core solver could successfully be tested, for the first time, on Armstrong."

Börje Andersson (U.S. Air Force Academy),
Armstrong pioneer user studying the structural integrity of various DOD aircraft

In-Situ Visualization

As our users' datasets grow larger, moving files off the HPC systems to perform scientific visualization on smaller clusters has become onerous.

We have added 32 NVIDIA Tesla K40s each to Lightning and Shepard to enable our users to investigate their data on the HPC system itself—which leads to a faster turnaround in optimizations of not only the scientific application itself, but often the scientific assumptions and design choices fed into the application.

For more information on these systems, visit centers.hpc.mil

the Systems

The High Performance Computing Modernization Program (HPCMP) provides over 429,000 cores and 7.5 petaFLOPS to its users in the pursuit of improved scientific research, weapons design, force protection, and software development for the Department of Defense

Maui High Performance Computing Center DSRC
Kihei, Maui, Hawaii

RIPTIDE
IBM iDataPlex

Processor Cores:	12,096
Memory (GB):	24,192
Disk (TB)	2,150
teraFLOPS:	253

Service Date:
5/17/2013

Army Research Laboratory DSRC
Aberdeen Proving Ground, Maryland

PERSHING
IBM iDataPlex

Processor Cores	20,160
Memory (GB)	45,696
Disk (TB)	2,370
teraFLOPS	419

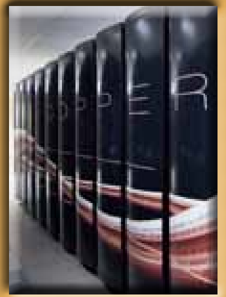
Service Date
1/7/2013

Engineer Research and Development Center DSRC
Vicksburg, Mississippi

COPPER
Cray XE6m

Processor Cores	14,720
Memory (GB)	29,440
Disk (TB)	443
teraFLOPS	138

Service Date
11/19/2012



GARNET
Cray XE6

Processor Cores	150,912
Memory (GB)	301,824
Disk (TB)	3,125
teraFLOPS	1,509

Service Date
7/18/2013

Air Force Research Laboratory DSRC
Wright-Patterson Air Force Base, Ohio

LIGHTNING
Cray XC30

Processor Cores	57,200
Memory (GB)	152,704
Disk (TB)	3,520
GPGPUs	32
teraFLOPS	1,281

Service Date
9/1/2014

PREDATOR
SGI UV

Processor Cores	1,004
Memory (GB)	4
Disk (TB)	88
teraFLOPS	22

Service Date
11/21/2013

SPIRIT
SGI ICE X

Processor Cores	73,440
Memory (GB)	146,880
Disk (TB)	2,458
teraFLOPS	1,528

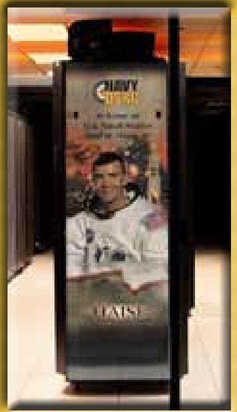
Service Date
3/25/2013

Navy DSRC
Stennis Space Center, Mississippi

HAISE
IBM iDataPlex

Processor Cores	19,812
Memory (GB)	40,832
Disk (TB)	2,473
Coprocessors	24
teraFLOPS	407

Service Date
1/7/2013



SHEPARD
Cray XC30

Service Date
9/1/2014

Processor Cores	30,592
Memory (GB)	82,752
Disk (TB)	1,560
Coprocessors	124
GPGPUs	32
teraFLOPS	822

KILRAIN
IBM iDataPlex

Processor Cores	19,812
Memory (GB)	40,832
Disk (TB)	2,473
Coprocessors	24
teraFLOPS	407

Service Date
1/7/2013

ARMSTRONG
Cray XC30

Service Date
9/1/2014

Processor Cores	30,144
Memory (GB)	81,728
Disk (TB)	1,560
Coprocessors	124
teraFLOPS	786

Early Experiences with a Heterogeneous Core HPC System at the ARL DSRC

by Thomas M. Kendall, Technical Director ARL DSRC

Following the Technology Insertion 2012 (TI-12) system acquisition cycle, the ARL DSRC integrated a heterogeneous processing system to allow ARL and the High Performance Computing Modernization Program (HPCMP) to gain insight into the applicability and performance of heterogeneous systems for relevant Department of Defense (DOD) HPC workloads. Heterogeneous systems and homogenous systems based on high-core-count processors are emerging as the leading candidates to enable supercomputing systems of the future to continue to deliver exponentially increasing performance over time.

FOB System		
	Quantity	Description
Login Nodes	3	2 X 2.6 GHz, 8 core, Intel E5-2670 CPUs 64 GB Memory 2 X 400 GB Solid State Disk FDR-10 Infiniband Interface Dual port 10 (gigE) Interface
X86 Coprocessor Nodes	48	2 X 2.6 GHz, 8 core, Intel E5-2670 CPUs 64 GB Memory FDR-10 Infiniband Interface Intel Phi 5110P, 60 Core Coprocessor with 8 GB Memory
GPGPU Nodes	16	2 X 2.6 GHz, 8 core, Intel E5-2670 CPUs 64 GB Memory FDR-10 Infiniband Interface Nvidia K20 GPGPU with 5 GB memory
Storage Nodes	2	2 X 2.6 GHz, 8 core, Intel E5-2670 CPUs 64 GB Memory FDR-10 Infiniband Interface 2 X LSI quad port miniSASx8 Interfaces

ARL's FOB system, short for Forward Operating Base, enables the HPCMP to evaluate the suitability of both heterogeneous processing and high-core-count processors. FOB is an IBM system with the same processor and interconnects as the TI-12 production systems at the ARL and Navy DSRCs. This similarity allows direct comparison of results to those systems: Pershing, Hercules, Kilrain, and Haise. The differences between the FOB system and the TI-12 production systems are limited to FOB's inclusion of 48 Intel Phi x86 coprocessors and 16 NVIDIA K20 general-purpose graphical processing units (GPGPUs). Both the Phi and the K20 each have a peak floating-point performance of approximately one teraFLOPS (TF) for double-precision operations, enabling two independent approaches to reach one TF per node to be evaluated.

Intel Phi 5110p processing cards are installed in 48 nodes of FOB. The Intel Phi coprocessor has the potential to enable improved performance with modest changes to existing application programs. Each Phi coprocessor utilized in FOB contains 60 cores supporting four threads each, 1.01 TF peak double-precision performance, and eight gigabytes (GB) of memory.

Several programming models are applicable to the Phi. The first programming model leverages the primary advantage of the Phi over other accelerators. Because the Phi is based on x86 technology, serial, OpenMP, and Message Passing Interface (MPI) applications written in C, C++, and FORTRAN can be built with the Intel compiler to run natively on the Phi without modification. Another supported programming model offloads individual, computationally-intense loops using compiler pragmas to identify regions of code that can be executed on the Phi. This is referred to as offload mode. There are several Intel Math Kernel Library functions that can implicitly utilize offload mode if a Phi coprocessor is detected and an environment variable is set. A third model is MPI with multiple program, multiple data (MPMD), where MPI ranks are scheduled to both the Phi and the Intel Sandy Bridge cores. Due to the performance differences between the 16 Sandy Bridge cores and the 60 Phi cores on each node, careful load balancing is required to make efficient use of the combined processing capabilities. A final programming model utilizes Open Computing Language (OpenCL), a framework for developing programs for heterogeneous systems.

The NVIDIA K20 GPGPUs are installed in 16 nodes of FOB. Unlike earlier GPGPUs, the K20 is optimized for high performance on double-precision workloads and has full support for error-correcting code (ECC) memory. The peak floating-point performance is 1.17 TF. Each K20 has 2,496 cores and five gigabytes of memory. To leverage the processing capability, applications can be written in NVIDIA Compute Unified Device Architecture (CUDA), OpenCL, or via library calls for common mathematical functions.

The system's software consists of Red Hat Enterprise Linux 6.5, Intel Cluster Toolkit, IBM Platform LSF, and NVIDIA CUDA. The Phi can be directly accessed from LSF or via an interactive session from a compute node with an attached Phi. Unlike the GPGPU, the Phi can be accessed via secure shell (ssh), resulting in a bash session. FOB's General Parallel File System (GPFS)-based home directories are available directly from the Phi coprocessors. This enables applications to be built on the hosting compute node or a login node and executed without the need to transfer files to the Phi's limited onboard storage.

In partnership with IBM, Intel, and Mellanox, FOB is running a software stack that enables native MPI applications to run across Phi coprocessors distributed across multiple compute nodes. MPI applications have been successfully executed using Sandy Bridge only, Phi only, and mixed Phi and Sandy Bridge configurations. MPI latencies across distributed Phi coprocessors are higher than typical of modern homogenous HPC systems, which will limit the range of applications that can achieve a high degree of performance using traditional MPI approaches. As a result of these higher latencies, applications with high computational intensity will have the best potential to perform using a traditional MPI model. The table below summarizes the MPI ping-pong latency as measured with the Intel MPI Benchmark for communications among various combinations of local and remote Sandy Bridge and Phi pairings. The worst-case latency of 10 microseconds (μs) is nearly five times that of the latency between two Sandy Bridge processors, reiterating that scalable performance will only be possible if the ratio of computation to communication is quite high.

Source	Destination	Latency (μs)
Sandy Bridge	*Sandy Bridge (on node)*	*2.2*
Sandy Bridge	*Sandy Bridge (on another node)*	*2.1*
Sandy Bridge	*Phi (on same node)*	*5.2*
Sandy Bridge	*Phi (on another node)*	*6.6*
Phi	*Phi (on same node)*	*3.7*
Phi	*Phi (on another node)*	*10*

ARL DSRC partnered with Sandia National Laboratories to evaluate the performance of a pre-released version of CTH on the Phi in both Phi-only and heterogeneous modes. CTH is a shock physics application that is heavily used across the HPCMP, and particularly at the ARL DSRC.

HPCMP customers with an interest in accessing FOB to evaluate distributed heterogeneous processing should inquire to the Consolidated Customer Assistance Center at help@ccac.hpc.mil.

FOB at the Army Research Laboratory

A Computational Framework for Hierarchical Multiscale Modeling

by Carrie E. Spear, Ken W. Leiter, Richard C. Becker, Oleg Borodin, David A. Powell, Mark J. Motsko, and Jaroslaw Knap, ARL DSRC

Over the last few decades, multiscale materials modeling (MMM) has become a dominant paradigm in materials science and engineering. MMM strives to replace material models entangled in empiricism and coarse phenomenology with a new class of high-fidelity models. This new class of models is derived by means of systematic analysis of the fundamental mechanisms associated with the entire range of material behaviors. Usually, this range is divided into a hierarchy of length and time scale and the relevant material behavior is extracted and modeled individually at each scale. Subsequently, these at-scale models are combined into a multiscale hierarchy of models, yielding a multiscale material model. However, the process of linking individual at-scale models is commonly quite cumbersome as there are currently few methodologies and tools for accelerating the process. The primary goal of our effort is to simplify building of multiscale material models via development of numerical methods, algorithms, and tools for expedient at-scale models assembly. To this end, we have developed a Python-based computational framework for scale-linking in MMM modeling. Since, in general, variability among at-scale models is remarkable, it is important that the framework be flexible and generic. For example, it must allow for multiple evaluations of an at-scale model at once, it must allow asynchronous coupling between models, and it must have the ability to leverage a variety of computational resources.

In Figure 1, we present an algorithm governing computational framework for a two-scale model hierarchy. Here, F(f,u) represents the macroscopic model, f(U) the microscopic model, u=G(u) represents the input filter, and g[f(u)] the output filter. The warehouse collects requests for microscopic model evaluations while the scheduler evaluates requests for microscopic model evaluation given available computational resources. The monitor is used to detect microscopic model failure and enables fault tolerance for the simulation.

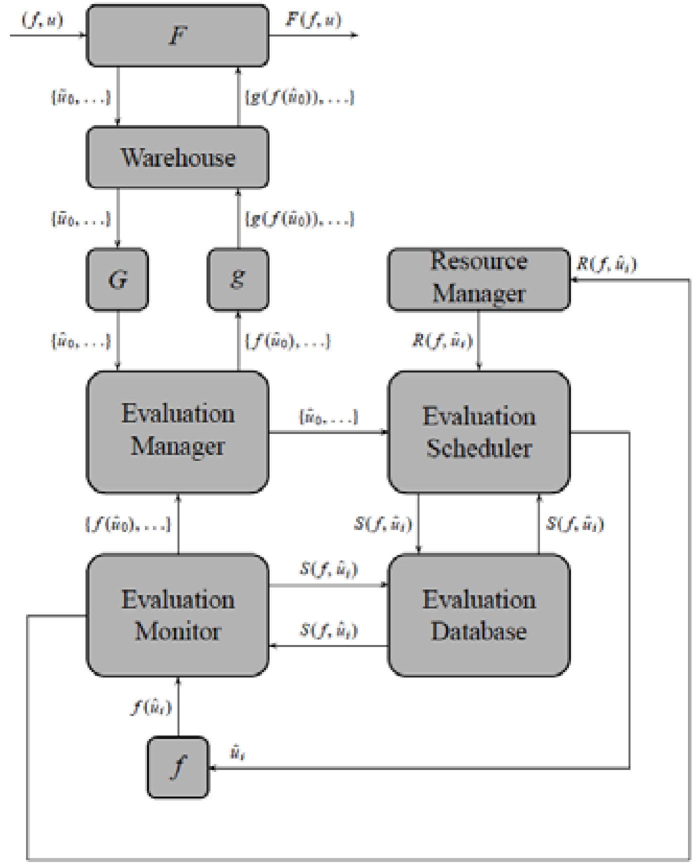


Figure 1: Hierarchical Multiscale (HMS) Framework Module

To assess the above framework, a proof-of-concept application comprised of a subgrid model of a dynamic deformation of a composite material was employed. We leveraged a commonly used finite element code to simulate a Taylor impact experiment involving a composite cylinder impacting a rigid anvil. In this two-scale example, our macroscopic model was a parallel explicit finite element simulation incorporating 1,036,800 elements, while our microscopic model was a finite element model incorporating 729 elements, with 81 elements occupied by the fiber.

In addition, we were able to leverage the HMS framework to perform high-throughput screening of electrolytes. In this case, our macroscopic model was an in-house-developed, high-throughput algorithm that was created to analyze the results of the microscopic model, a commonly used electronic structures code. The algorithm stored results of its operation into a database for subsequent analyses and visualization via the Scalable Multitiered Architecture for Research Tools (SMART).

We learned some valuable lessons while running, debugging, and profiling our framework. For example, Python's subprocess command turned out to be extremely computationally expensive. Our original HMS implementation launched a new subprocess command each time a microscopic model was evaluated, resulting in a substantial (negative) impact on performance. While working to improve the performance and extend the functionality of our framework, we encountered Python's global interpreter lock (GIL). The GIL prevented us from multithreading some of the modules in the HMS framework to achieve additional concurrency. The above lessons led us to redevelop our HMS framework in C++.

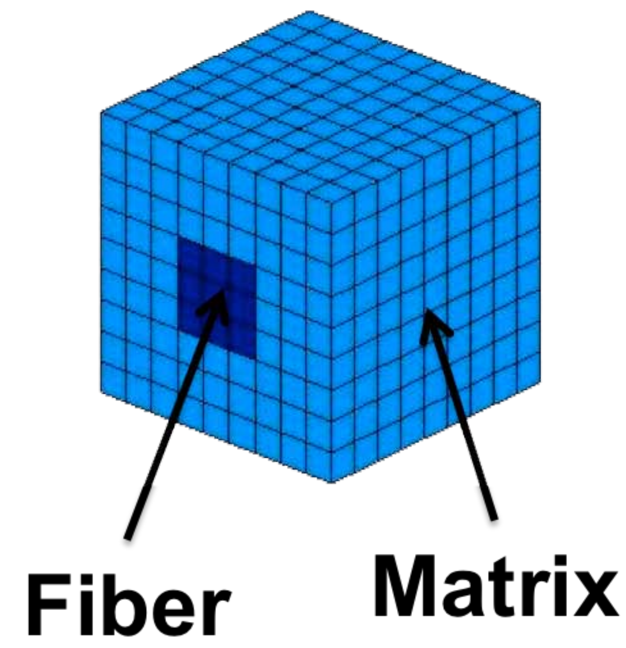


Figure 2: Microscopic model, finite element model of an elastic composite material

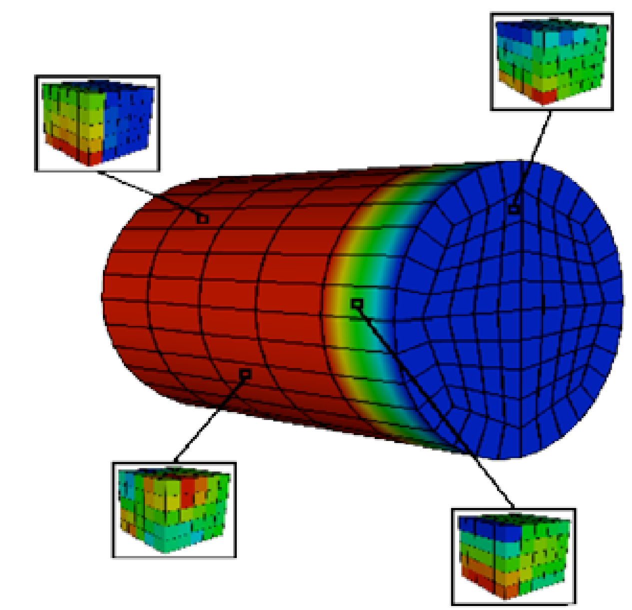
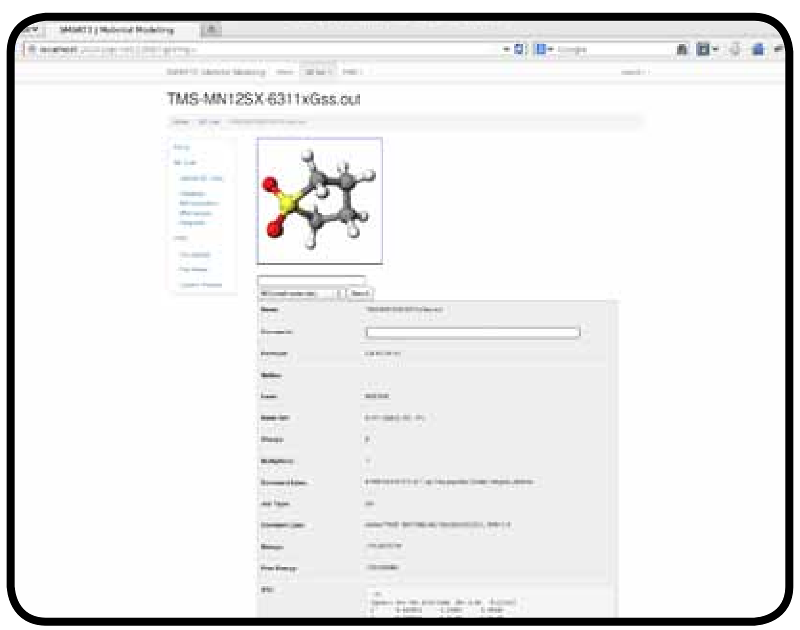


Figure 3: Macroscopic model, a finite element model of a solid body

Our team has successfully developed and tested a generic and flexible computational framework for scale-linking in MMM. We have also established that our methodology is applicable beyond MMM. As mentioned earlier, we were able to successfully leverage our framework to perform high-throughput screening of battery electrolytes, as well as deformation of a composite material. The HMS framework was also leveraged to perform a distributed analysis of graph data.

We would like to thank the staff at the AFRL DSRC for assisting us with acquiring 40,000 processors to run a large capability example on Spirit. With their assistance we were able to execute a simulation involving a composite cylinder impacting a rigid anvil at 150 m/s, resulting in an effective calculation of over 2.2 billion elements.

Figure 4: SMART Quantum Chemistry Interface



New Queue Policies Implemented on Garnet

by Christopher Borchert, ERDC DSRC

During the fall of 2013, the ERDC DSRC upgraded its Cray XE6, Garnet, to 150,912 cores. Before entering production, new queue policies were implemented to allow for faster throughput of jobs requesting a large number of cores. On May 1, 2014, the ERDC DSRC enhanced the queue policy to allow more smaller-core jobs to run, as well as to provide long-walltime jobs access to more cores.

When a user submits to the standard queue, for example, PBS routes the job to one of three execution queues: large (standard_lg), small (standard_sm), or long-walltime (standard_lw). Queues ending in "_lg" contain jobs that request between 5,120-102,400 cores and have walltimes of at most 24 hours. Queues ending in "_sm" contain jobs that request 5,120 cores or less and have walltimes of at most 24 hours. Queues ending in "_lw" contain jobs that request 2,048 cores or less and have walltimes between 24 and 168 hours.

The queue policy modification increased the number of running small ("_sm") jobs to 40 per user. It also reduced the number of running long-walltime ("_lw") jobs to four per user, while the maximum size increased to 2,048 cores. Likewise, it removed the previous limit of total cores allowed for long-walltime work for all users. Finally, it is no longer necessary for users to specify the "standard_lw" queue; they can simply request the "standard" queue with a walltime of more than 24 hours.

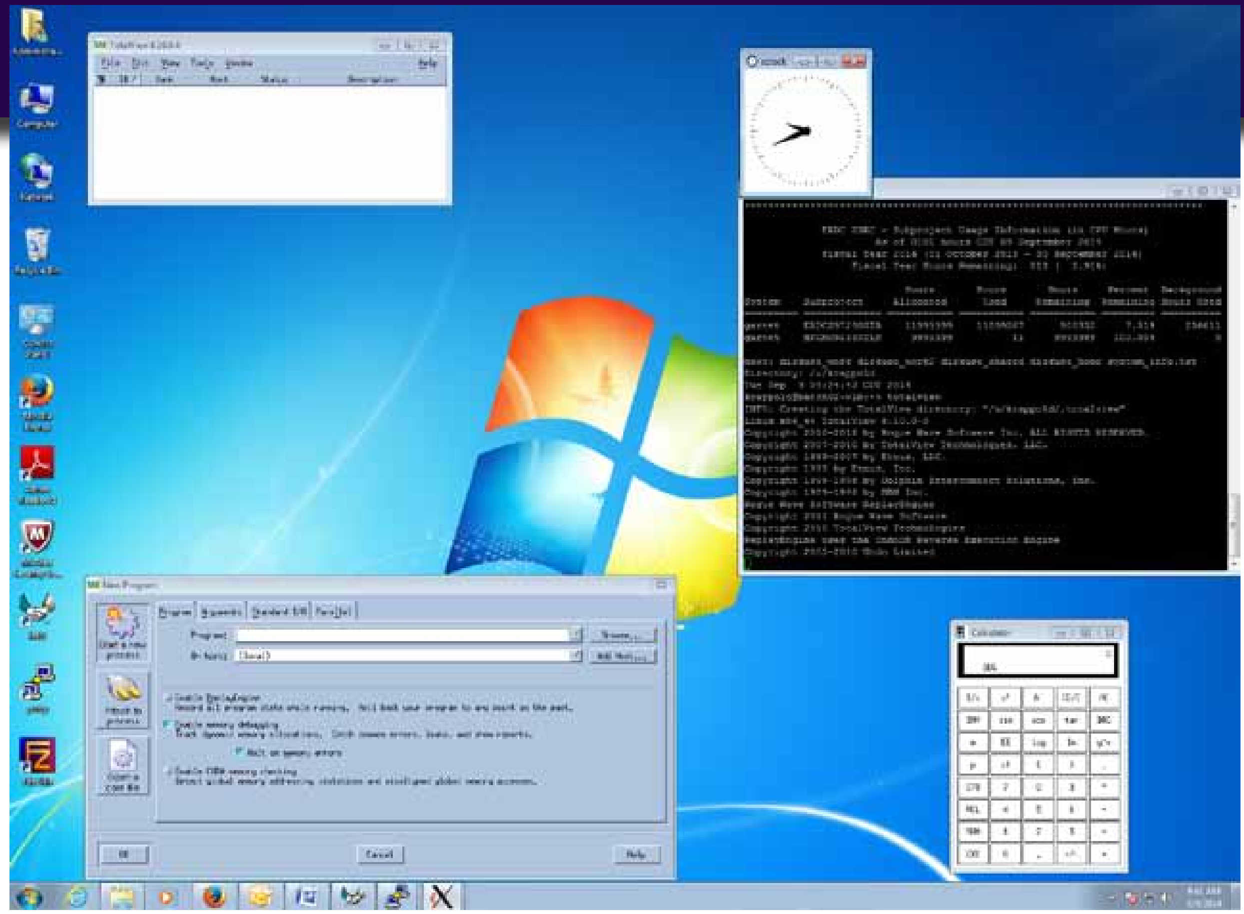
At the time of job submission, a formula is used to set initial priority. This gives the highest priority to jobs requesting large core counts. The formula is the queue's priority + (number of cores * 0.005). The formula for calculating initial job priority is:

$$\textit{Initial Job Priority} = \textit{Queue Priority} + (\textit{Number of cores x } 0.005)$$

With the new queue policy changes on Garnet, the ERDC DSRC enables better efficiency and job throughput for capability-class workloads. If you need additional information, please contact the Consolidated Customer Assistance Center (CCAC) at 1-877-222-2039 or help@ccac.hpc.mil.

The various parameters and limits associated with queues on Garnet

Route	Queue	Priority	Job Max	Core Min	Core Max	Walltime
	urgent	900	---	32	5k	24:00:00
	test	600	---	32	---	24:00:00
	debug	49	2	32	8k	1:00:00
high	high_lg	45	2	5k	100k	24:00:00
high	high_sm	45	40	32	5k	24:00:00
high	high_lw	45	4	32	2k	168:00:00
challenge	challenge-lg	40	2	5k	100k	24:00:00
challenge	challenge-sm	40	40	32	5k	24:00:00
challenge	challenge-lw	40	4	32	2k	168:00:00
standard	standard-lg	20	2	5k	100k	24:00:00
standard	standard-sm	20	40	32	5k	24:00:00
standard	standard-lw	20	4	32	2k	168:00:00
	transfer	20	3	1	1	12:00:00
	background	-1000	6	32	5k	4:00:00



Kerberos and Xserver Install without Admin Privileges

ERDC DSRC

Putting the right tools in place to access high performance computers can oftentimes conflict with your organization's security policies, policies which cover the desktops and laptops you use for your everyday work. This can be especially true if you are looking for something more than just a command-line prompt on your terminal.

The HPCMP provides several options for overcoming these challenges. The HPCMP Portal and Secure Remote Desktop (SRD) are resource access tools that facilitate ease of access to HPCMP resources with minimal or no install. You can learn more about these tools at
http://tinyurl.com/hpcmp-srd and ***http://tinyurl.com/hpcmp-portal***

The ERDC DSRC documented one more approach at
http://tinyurl.com/hpcmp-xwin

This tutorial will guide you through the installation of Kerberos and an X11 server without having administrator privileges on your desktop or laptop. As long as you can write to a directory on your local system, this approach should work for you. Try it and let us know how it works for you.

If you have any questions or comments, contact CCAC at 1-877-222-2039 or help@ccac.hpc.mil.

DREN III Transition

Completes Multiyear Journey

by Ralph McEldowney,
Associate Director for Networking, HPCMP

Transition to the Defense Research and Engineering Network (DREN) III, the network that connects the High Performance Computing Modernization Program's (HPCMP) computational scientists and engineers across the United States with its high performance computing systems, was completed on June 13, 2014, bringing a multiyear project to a successful conclusion. The transition finished five days ahead of schedule with no loss of service to users of the previous network, DREN II. The DREN III service provider, CenturyLink, and the HPCMP Networking and Security staff worked incredibly long hours for nine months to transition 149 DREN II sites and more than 440 enclaves to the new network. However, this was just the final phase in a multiyear journey.

The HPCMP and DREN Technical Advisory Panel (TAP) began drafting the DREN III Performance Work Statement (PWS) in 2009. The PWS was based on lessons learned from DREN II, with the addition of next-generation technologies and bandwidth capacities. The PWS also incorporated input from industry as a result of an Industry Days event held in September 2009. Activities significantly accelerated with the release of the DREN III PWS in January 2011. After a comprehensive source selection evaluation in a full and open competitive environment, the DREN III contract was awarded to CenturyLink in December 2012. Each of these activities was required before the sites could even consider transitioning to the new network.

DREN III is a 10-year, firm-fixed-price, indefinite-delivery/indefinite-quantity wide-area network services contract for more than 260 sites worldwide. It is valued up to $750M. DREN III includes Internet Protocol (IP) services, as well as next-generation transport technologies such as Ethernet and Optical Services. DREN III includes bandwidth capacities starting at 50 megabits per second (Mbps), and increases by four orders of magnitude, up to 100 gigabits per second (Gbps). For comparison, the maximum bandwidth at DREN II sites was 2.4 Gbps.

The transition to DREN III finally began in October 2013, following a successful nine-month Initial Performance Capability (IPC) demonstration of the first ten sites. Through IPC, CenturyLink demonstrated the DREN III network met all PWS requirements. During and after IPC, the DREN transition team contacted each site's point-of-contact (POC) and worked with them to transition each enclave from DREN II to DREN III. Sites with a small number of enclaves typically transitioned easily. Sites with a large number of enclaves, or a complex DREN II configuration, typically required significant time and effort to complete transition. However, with great cooperation from each site POC across nine months of transition activities, the DREN III transition finished five days ahead of schedule, on June 13, 2014, with no loss of service to DREN II users. The DREN II network remained active until the end-of-day on June 18, 2014, when that contract expired and the DREN II network was finally shut down.

During this journey—which started with Industry Days in September 2009, included a comprehensive source selection evaluation and contract award in 2012, and culminated in the final site transition in June 2014—the HPCMP team experienced three DREN program managers, three PWS/transition managers, two contracting officers, and two financial managers. Fortunately, there was great stability on the transition team itself, with experienced engineering, operations, and security team members. Without their extreme dedication and passion for the DREN mission, the multiyear transition from DREN II to DREN III would not have been successful. As the transition comes to a close, the HPCMP team now looks forward to its next journey in the DREN program: providing flawless high-speed networking to HPCMP users.

Using Multipath Secure Copy (MPSCP)

To Improve HPC File Transfer Performance

by John Skinner, Navy DSRC

Multipath Secure Copy (MPSCP) is a high performance utility for transferring large files between High Performance Computing Modernization Program (HPCMP) Department of Defense (DOD) Supercomputing Resource Center (DSRC) systems. The MPSCP utility achieves high transfer rates by spreading data over multiple streams using Transmission Control Protocol/Internet Protocol (TCP/IP) sockets, and it calls the Secure Shell (SSH) command internally to initiate a connection that has been authenticated by HPCMP Kerberos. MPSCP has a basic command line syntax similar to the nonkerberized remote copy (rcp) command and HPCMP-kerberized secure copy (scp) command.

The MPSCP command is installed on all HPCMP computational and archival storage systems. Configuration settings and environment variables used by MPSCP are set automatically in a default user login session. Documentation is available via the 'man mpscp' command.

Note that the maximum number of data streams per transfer for MPSCP may be limited to prevent negative system impact and overloading input/output (I/O) bandwidth. On Navy DSRC systems, the limit is four streams per transfer. You must have a valid, unexpired HPCMP Kerberos ticket on the system where you start the MPSCP transfer. Although MPSCP can be used to transfer small files, it is strongly recommended that you employ the 'tar' command to assemble related files such as source file trees and related datasets from large computational runs into individual larger tar files, and then compress the tar files with the 'gzip' command, which supports very large individual file sizes.

Best performance gains using MPSCP are obtained when you transfer individual large files rather than many small files, especially when using MPSCP to archive data to the long-term archival storage servers. The utility also has a "-v" option that can be used for debugging and a "-p" that can be used to preserve modification times.

More information on these options can be found in the online man page or by contacting the Consolidated Customer Assistance Center (CCAC) at 1-877-222-2039, or by e-mail to help@ccac.hpc.mil.

EXAMPLE MPSCP USAGE

Example 1 HPC server transfer to archival server

```
haise% mpscp -w4 -v twofiles.tar newton:\
/u/b/skinman/haise-twofiles.tar
twofiles.tar
100% 2610MB  290MB/s   00:09
2737223680 bytes received in 8.3 seconds
(3.2e+05 Kbytes/s)
```

Example 2 HPC server transfer from archival server

```
haise% mpscp -w4 -v newton:/u/b/skinman\
/haise-twofiles.tar haise-twofiles.tar
haise-twofiles.tar
100% 2610MB  326MB/s   00:08
2737223680 bytes received in 7.8 seconds
(3.4e+05 Kbytes/s)
```

Example 3 HPC server transfer to $WORKDIR on another HPC server

```
haise% mpscp  -w4 -v -m /site/MPSCP\
/MPSCP-1.3a/bin/MPSCP twofiles.tar \
kilrain:$WORKDIR/kilrain-twofiles.tar
twofiles.tar
100% 2610MB  653MB/s   00:04
2737223680 bytes received in 4.5 seconds
(6e+05 Kbytes/s)
```

Example 4 HPC server transfer from $WORKDIR on another HPC server

```
haise% mpscp  -w4 -v -m /site/MPSCP\
/MPSCP-1.3a/bin/MPSCP kilrain:$WORKDIR\
/kilrain-twofiles.tar kilrain-twofiles.tar
kilrain-twofiles.tar
100% 2610MB  522MB/s   00:05
2737223680 bytes received in 4.8 seconds
(5.6e+05 Kbytes/s)
```

Example 5 HPC server transfer to a remote DSRC server over DREN

```
haise% mpscp  -w4 -v twofiles.tar \
garnet02.erdc.hpc.mil:/lustre/work2\
/skinman/haise-twofiles.tar
twofiles.tar
100% 2610MB  131MB/s   00:20
2737223680 bytes received in 20 seconds
(1.4e+05 Kbytes/s)
```

Example 6 Transfer started on remote DSRC server back to Navy over DREN

```
garnet% mpscp  -w4 -v haise-twofiles.tar \
haise.navo.hpc.mil:\scr/skinman\
/haise-twofiles.tar
haise-twofiles.tar
100% 2610MB  145MB/s   00:18
2737223680 bytes received in 17 seconds
(1.5e+05 Kbytes/s)
```

Here is a transfer of same 2.55 gigabyte tar file from Haise to Newton using the HPCMP-kerberized 'scp' command. Note the MPSCP test in Example 1 above was 19 times faster than the traditional SCP transfer shown below.

Example 7 HPC server transfer to archival server using SCP

```
haise% /usr/bin/scp -v twofiles.tar newton:/u/b/skinman/twofiles.tar
twofiles.tar
100% 2610MB  16.8MB/s  02:35
Transferred: sent 2740033352 in 157.8 seconds
```

User Productivity Enhancement, Technology Transfer, & Training PETTT

By Dr. Casey Church, HPTi (a wholly-owned subsidiary of Engility Corp.), PETTT Program Director

The Department of Defense (DOD) High Performance Computing Modernization Program (HPCMP) is well known for the state-of-the-art HPC hardware provided to DOD users at the five DOD Supercomputing Resource Centers (DSRCs). In an effort to maximize the extent to which users can leverage this hardware investment, HPCMP's PETTT program provides user productivity enhancement, technology transfer, and training via its 32 scientific HPC subject matter experts (30 of whom hold a doctorate degree). Likewise, the PETTT program's training cadre and its oversight of Pre-Planned and Special Projects serve to bring the latest technology from academia and industry to the DOD.

Productivity Enhancement

PETTT provides user productivity enhancement in response to requests for expert advice and assistance in developing and using HPC applications. Users can submit requests directly to a local on-site or via the centralized PETTTrequests@hpc.mil. Requests may span the gamut from "tier-3 support" to "advanced HPC consulting services." Although the 32 "on-sites" are largely collocated with the DSRCs (Table 1), the enterprise service model means tasks can be allocated based on availability and specialized expertise, independent of location.

Technology Transfer

A core element of the PETTT program is bringing new tools and technology to DOD users through PETTT projects. Competitively awarded by the HPCMP, Pre-Planned Projects have a one-year duration, are executed by academia or industry, and are focused on tool and software development. As a means of transferring technology into DOD's HPC workforce, the HPCMP issues a call for 2-page concept proposals each January, with about 40 percent solicited for full proposals. Each year, typically 20-25 projects are awarded, with about 40 percent executed by academic principal investigators (PIs) and about 60 percent by industry PIs.

Special Projects connect users who have their own funding to HPC experts for more in-depth or sustained efforts. Advantages include ease of access (the sponsor only needs to provide a 1–2 page statement of work), low overhead cost through leveraging the PETTT Program Management Office structure, ability to engage experts from academia and industry, and access to the 32 PETTT on-site subject matter experts. In addition, knowledge is synergized across 15-20 special and 20-25 Pre-Planned Projects.

Recent Examples of PETTT User Support

- Ported Boron Carbide Equations of State model into application agnostic code.
- Profiled Matlab code to identify speedups and research alternative implementations of the underlying optimization routines.
- Gathered a full sweep of benchmark data on Phi and K20 accelerators for multiple workloads and provided a draft manuscript describing the benchmark code and its capabilities.
- Performed a strong scaling study using a 512x512 horizontal resolution on 16, 32, 64, 128, 256, and 512 cores in dual stream and single stream mode and collected performance data for varying vertical resolutions on 2, 4, and 8 cores.

Recent Examples of PETTT Training

- Performance Evaluation and Optimization with TAU
- Adios I/O
- Introduction to HyperMesh and HyperView
- Object-oriented, Parallel and Functional Programming in Modern Fortran
- Visualization with Paraview
- Electromagnetic Simulation With Xpatch
- Scientific and Parallel Computing Using Python

Training

The HPCMP provides HPC training through the PETTT Training program, which facilitates between 30-35 events annually at sites across the country. Users may submit requests to PETTTtraining@hpc.mil, which the HPCMP will then evaluate. In some instances, members of the PETTT team provide the instruction, but in the majority of cases, PETTT contracts with software developers and vendors to tap into the most current and detailed information available. The PETTT Training team identifies a venue, provides training laptops configured for the training, webcasts and video archives the training (whenever feasible), and collects post-course survey feedback. A list of upcoming courses and prior-course archives can be found on the Online Knowledge Center (OKC) at https://okc.erdc.hpc.mil/xwiki/bin/view/Training/CourseSchedule (authentication required).

For more information on PETTT Projects, email: PETTTprojects@hpc.mil.

Facilitating collaboration and the opportunity to maximize the DOD's return on its HPC investment is another key element of PETTT Training. An example of this is the recent Density Functional Theory Technical Interchange Meeting. This meeting brought together DOD and Department of Energy (DOE) scientists, experts from academia and industry, and the PETTT Computational Biology, Chemistry, and Materials Science cadre to discuss algorithms for computing energies and gradients of large and complex systems using density functional theory or other quantum mechanical methods. As a secondary focus, researchers are finding that as more efficient algorithms are being adapted to modern high-performance computing architectures, the question of how to best take advantage of such powerful techniques frequently arises. Dialogue between participants in this forum will accelerate the research of the DOD scientists participating in person, as well as affect a broader community through a forthcoming related book by Springer Scientific Publishing.

PETTT's Subject Matter Expert On-Sites

Affiliation	ACE	COM	CEA	CFD	CSM	CWO	ENS	EQM	IMT/FMS	SAS	SIP
AFRL	1.5	2	1	1.5	1						1
ARL	2	2			1		2		1		1
ERDC	1			2	1			1	1		
Navy	4	1	1	1		1.5				0.5	

Table 1

Connecting with PETTT

As an enterprise-wide support organization, PETTT services can be obtained by either contacting local on-sites or through centralized contact points.

PETTT Site Leads

AFRL	Steve Wong	Steven.Wong@engilitycorp.com
ARL	Mike Lasinski	Michael.Lasinski@engilitycorp.com
ERDC	Nathan Prewitt	Nathan.Prewitt@engilitycorp.com
Navy	Sean Ziegeler	Sean.Ziegeler@engilitycorp.com

Centralized Contacts

Reactionary Assistance Support	PETTTrequests@hpc.mil
Training Requests	PETTTtraining@hpc.mil
Projects	PETTTprojects@hpc.mil
Technical Lead	Hugh.Thornburg@engilitycorp.com
Adv. Comp. Environments Lead	Sean.Ziegeler@engilitycorp.com
Program Oversight	Casey.Church@engilitycorp.com
PETTT COR	Brad.Comes@hpc.mil

PETTT Support for DSRCs

When three small Cray XE6 systems were combined into Garnet's current configuration of 150,912 compute cores, the result was the largest system in the DOD HPCMP. But in order to take advantage of such a system, one must consider I/O performance and analysis of the resulting output data.

With the purpose of investigating various I/O and visualization coprocessing libraries, PETTT's "100K+Core Challenge" project employed previously developed code that had been tested and proven to scale well on more than 270,000 cores on DOE systems.

The project's approach was to test various methods independently and benchmark each one. The first method, Smart POSIX, was originally written by PETTT staff for fast I/O on thousands of cores. The next four methods were all provided by the ADIOS library, and ranged from a similar POSIX method to various forms of MPI-IO. The final method used coprocessing (the analysis or visualization of data in-core as the model executes). This project used ParaView Catalyst and a library named Qiso, which PETTT staff wrote. All of the I/O methods were benchmarked using six different job sizes. The coprocessing methods were benchmarked using the first five of the six. Job sizes were increased by using more compute cores and by adding resolution to the grid with performance of each I/O and co-processing methods shown in Figures 1 and 2.

PETTT by the numbers....	
User Productivity Enhancement	
Scientific HPC Subject Matter Experts on staff (30 with doctorate degrees)	32
User Assistance actions per year (primarily in optimization, porting, profiling, visualization, meshing, and accelerator utilization)	600+
Technology Transfer & Training	
Training events per year	30+
HPCMP-funded Pre-Planned Projects done through PETTT each year	20 - 25
User-funded Special Projects per year	20+
Academic and Industry Partners under subcontract each year	50+
Enhancing the productivity of the HPCMP's user community through HPC-related science and technology support, training, collaboration, and technology transfer	

Execution of this project highlighted the criticality of I/O optimization at large computational scales and the potential benefits of moving post-processing analyses and visualization inside the execution realm whenever possible in order to minimize I/O requirements. The final technical report can be obtained from the OKC at https://okc.erdc.hpc.mil/okc/ContentAdmin/projectDetail.jsp?project=CCM-KY04-013

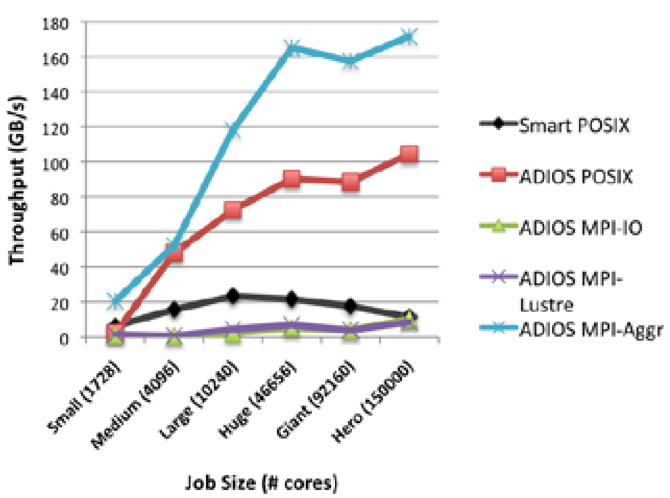


Figure 1. Minimum throughput results for the 5 I/O methods and 6 job sizes

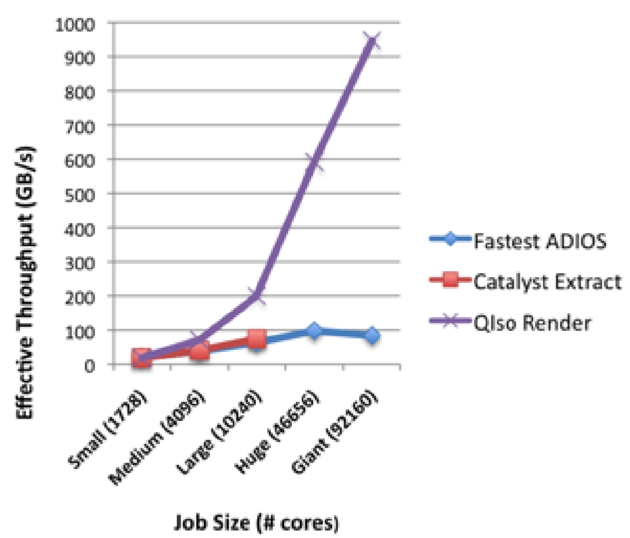


Figure 2. Effective throughput for the two coprocessing methods as compared to fastest I/O methods.

From the Director's Desk AFRL DSRC

by Jeff Graham, Director AFRL DSRC

It is with great excitement that I write about some of the activities underway at the Air Force Research Laboratory DOD Supercomputing Resource Center (AFRL DSRC) this year. I look at 2014 as a year of celebration and opportunity. With a new High Performance Computing (HPC) system arriving as part of the HPCMP's acquisition process, two new Dedicated HPC Project Investment (DHPI) systems being planned at AFRL, the largest Frontier project ramping up, and a significant outreach initiative taking shape under the wings of our new Technical Director—not to mention a new HPCMP Integrated Technical Services contract on the horizon—we are clearly living in exciting times!

First things first, we are delighted to have a new Cray XC30 system, Lightning, named in honor of the F-35 Joint Strike Fighter, passing into the production phase for an eager customer base. The new system sports 57,200 processor cores that propel this supercomputer into the astounding petaFLOPS realm (a petaFLOP is one quadrillion [1×10^{15}] floating point calculations per second). Cray, Incorporated has been in the supercomputer business for as long as there has been a supercomputer business, and we look forward to focusing this incredible resource on the key mission-critical projects that will increase the effectiveness of our warfighters across the service branches. Nearly doubling our capability at the AFRL DSRC, this system should provide much needed elbow room for customers who have been feeling a bit cramped on Spirit lately.

(continued on page 28 "Director of the AFRL DSRC")

From the Director's Desk ERDC DSRC

by Robert M. Hunter, Acting Director, ERDC DSRC

Ninety thousand one hundred and sixty two…that's the core count size of jobs that are no longer shocking to the staff of the ERDC DSRC when jobs of this size are queued on Garnet, the ERDC DSRC's Cray XE6. Trust me; we were shocked when we saw such a feat being accomplished. At 90,162 cores, a job of that size consumes 60 percent of the 150,912 available compute cores on Garnet. That's large! In addition, what makes the running of these large core count jobs a real success is that they go from job submission to job execution in fairly short order. Jobs in the 60K to 90K core count range are experiencing queue wait times of 24 hours or less. That is outstanding for a system that routinely stays at 98 percent or greater utilization.

So, how have we done it? I believe it all starts with identifying target workloads for a given high performance computing (HPC) system, then developing queuing policies with those workloads as the basis. From the outset, Garnet was targeted to support large jobs. How large? No one knew for certain, and "large" isn't defined as a singular quantity. The term encompasses a spectrum of core count sizes. But, historically, the right-hand end of the "large" spectrum was quantified by approximately 25 percent of a system. Garnet has clearly exceeded that benchmark and defined a new one at just over 50 percent of a system. So, back to my original question of, "How have we done it?" It took a lot of data analysis, HPC scheduling expertise, new queue policies, and a little bit of good old-fashioned luck.

You can learn more about the analysis and scheduling expertise that contributed to the queue policies in place on Garnet today by reading Chris Borchert's article entitled "New Queue Policy on Garnet," in this edition of HPC Insights. As for the luck…well, everyone can use a little bit when you are working an intractable problem.

Paul W. Kervin

2014 AFRL Fellow

By Betty Duncan, MHPCC DSRC

The Air Force Research Laboratory (AFRL) celebrated its best and brightest in an October ceremony highlighting the brilliant work and promising future of AFRL Fellows. The AFRL Detachment 15, located on Maui, HI, announced with great pride that Mr. Paul W. Kervin had received the Air Force Fellow Award in recognition of his outstanding work and achievements at the AFRL.

A Fellows designation is the highest AFRL honor possible, with just 164 (0.6 percent) of the laboratory's scientists and engineers in research and development and technical program management having been chosen to receive the title.

Mr. Kervin was drafted into the Army while he was attending graduate school at Oklahoma State University, OK and was sent to Vietnam. After his Army service, he was hired by the U.S. Air Force. He was then sent to training for a year at the Optical Sciences Center in Tucson, AZ. While working at the Airborne Laser Laboratory (ALL), he became the Chief of the ALL Optical Diagnostic and Analysis Group. Mr. Kervin served as Chief Scientist of the Strategic Defense Initiative Organization Relay Mirror Experiment (RME) from design through on-orbit operations and final analyses. The RME team went on to win the International Society for Optical Engineers (SPIE) 1991 Technical Achievement Award of the Year. Mr. Kervin became the Chief Scientist for the Air Force Maui Optical and Supercomputing Site, the research and development (R&D) component of the Maui Space Surveillance System (MSSS). He has been the principal investigator of multiple projects utilizing High Performance Computing Modernization Program resources, and he currently leads the development of non-imaging space object identification on the Maui High Performance Computing Center (MHPCC) IBM iDataPlex system, Riptide.

Mr. Kervin collaborated with the Jet Propulsion Laboratory to use their expertise coupled with MSSS assets in support of the Planetary Defense Mission. He developed the Raven small telescope program, which was certified as a contributing sensor by the Space Surveillance Network in 2001. For this work, Mr. Kervin garnered the 2001 Air Force Chief Scientist's Engineering Achievement Award. He partnered with NASA's Orbital Debris Program Office in the 1990s, which led to the installation of a Raven R&D telescope on the Kwajalein Atoll and his selection as a U.S. representative to the Inter-Agency Space Debris Coordination Committee, an international governmental forum for the worldwide coordination of activities related to debris in space. In 2009, Mr. Kervin also negotiated a memorandum of understanding between the U.S. Air Force and the Australian Government's Department of Defence to allow the installation of Raven telescopes in Australia.

Mr. Kervin is the founder and technical chairman of three venues for technical interchange: the annual AMOS Conference, which has received high praise from the Commander of the Air Force Space Command; the annual classified Space Situational Awareness Conference; and the semi-annual Non-Imaging Space Object Identification Workshop. During his career he has published more than 700 journal articles and conference proceedings. His career with the government extends for more than 40 years.

Additional major AFRL Awards that Mr. Paul Kervin has received include the AFRL Directed Energy Directorate Commanders Cup in 2006 and the Sparks Award in 2014.

.

"Lightning" Strikes the AFRL DSRC

Virginia Watson Ross, Ph.D., AFRL DSRC

The DOD High Performance Computing Modernization Program (HPCMP) recently acquired a $20.8M Cray XC30 Supercomputer called Lightning, which is located at the Air Force Research Laboratory's DOD Supercomputing Resource Center (AFRL DSRC) at Wright-Patterson Air Force Base. Part of the Technology Insertion 2013 (TI-13) procurement, Lightning is named in honor of the F-35 Joint Strike Fighter. Lightning has a theoretical peak performance of 1.2 petaFLOPS, which nearly doubles the computing power of the AFRL DSRC and thereby enables the DSRC to stay on top of the continuously expanding computational demands of the DOD research community.

Employing proven commercial off-the-shelf technology, Lightning includes 57,200 Intel Xeon E5 Ivy Bridge processor cores, 32 NVIDIA Tesla K40 graphical processing units (GPUs), low-latency Cray Aries interconnect with Dragonfly network topology, 150 terabytes of total memory, and total disk space of 4.5 petabytes.

Lightning was installed under a very aggressive schedule. The contract was awarded on February 5, 2014, and the system was delivered three months later on May 5, 2014. The speediness of this transaction made it possible for the DSRC to get the newest, yet proven, technology into the hands of DOD researchers as quickly as possible. Having just a short time to install Lightning is one thing, but preparation for such a move was much more challenging. The new system placed significant demands on electrical and cooling infrastructure, and design drawings needed to be firmed up to locate the cabinets, coolant pipes, power lines, and cable trays—not to mention actually installing them. This process involved coordination among the various facility contractors, Cray engineers, and local staff.

Acceptance testing on Lightning was completed on June 30, 2014. The AFRL DSRC planned several phases of pioneer user testing. During the initial pioneer phase, highly-skilled programmers executed large jobs across major portions of the system, sometimes having the entire computer running just one job. These users were the code builders and application owners. A second wave of testing involved comparison testing by pioneer users with accounts on both Lightning and the Navy DSRC's new Cray XC30 systems, Armstrong and Shepard, which are very similar to Lightning. This collaboration compared system performance to help ensure compatible user environments on the different systems, enabling users to switch computers more easily when executing jobs. A third wave of pioneer user testing taxed Lightning with large, memory-demanding applications.

Lightning is a powerful addition to the AFRL DSRC's computing lineup, having the potential to be a key technology enabler for DOD researchers, and it brings the promise of more computing power to bear in the service of our Nation and its warfighters.

Navy DSRC Architecture

Supports High-priority, Time-critical Requirements

By Bryan Comstock, Chief Technologist and Dave Cole, Acting Director, Navy DSRC

Beginning with the Department of Defense (DOD) High Performance Computing Modernization Program (HPCMP) Technology Insertion 2011/2012 (TI-11/12) supercomputer procurement, the Navy DOD Supercomputing Resource Center (DSRC) implemented a strategy to acquire two identical unclassified HPC systems for each TI cycle in order to provide the ability to support DOD HPCMP high-priority, time-critical requirements. Based on this architecture, the Navy DSRC is able to address the needs of high-priority projects characterized by time-critical job activity that occurs on a regular or recurring basis. By deploying twin IBM iDataplex systems, the Navy DSRC has provided support to the U.S. Army Test and Evaluation Command (ATEC) to test the feasibility of meeting the group's operational weather forecasting requirements.

Weather support for developmental testing of military systems is a force multiplier. It improves effectiveness by identifying when weather suitable for testing will occur, helps avoid lost days, and ensures the safety of personnel and materiel on the Army test ranges. During the mid-1990s, the ATEC and the Research Applications Laboratory (RAL) at the National Center for Atmospheric Research (NCAR) began a collaboration to provide ATEC's weather forecasters with a high-resolution, customized numerical weather prediction system that uses the most advanced and effective data-assimilation techniques to harness the abundant and detailed observations from the ranges. The culmination of that effort is the Four-Dimensional Weather System, or 4DWX, which is now implemented at Dugway Proving Ground, Utah to support eight Army test ranges.

In early May, the HPCMP granted ATEC a Dedicated Support Partition (DSP) for fiscal year 2014 (FY14) to evaluate if a DSP could meet ATEC's ensemble forecasting requirements. The DSP project was suggested by the HPCMP Office as an alternative to ATEC's initial request for the acquisition of a very large Dedicated HPC Project Investment (DHPI) system at Dugway for the purpose of executing 4DWX ensembles for all the ATEC test centers. ATEC's requirement to generate accurate ensemble forecasts for eight Army test ranges four times a day lends itself well to the DSP model—a dedicated subset of nodes on an HPC system. However, the challenge is to provide network, HPC systems, and related IT support infrastructure availability sufficient to support ATEC's time-critical requirements.

The Navy DSRC implemented the initial DSP, established for code porting, with 32 dedicated nodes (512 cores) on its iDataPlex system Kilrain. Provided the pilot DSP is successful, the project will expand to 2048 cores (128 nodes) on Kilrain in FY15. This amounts to a potential 20M hours of computational time committed to operational Army weather forecasting.

Current work focuses on automating management of input data with a goal that includes executing an ensemble of 4DWX instances. Once multiple ensembles are working routinely with data flowing in and out, ATEC will have a good indication as to whether or not the DSP approach will work operationally. Eventually, when all of the test centers are supported and the ensembles are operating stably, ATEC may consider using the DSP approach for all its 4DWX deterministic computational requirements, as well as ensemble forecasting requirements.

Continuing with the Navy DSRC strategy to acquire two identical HPC systems each TI cycle, the Center received two identical Cray XC30 HPC clusters (Armstrong and Shepard) as a result of the DOD HPCMP TI-13 procurement. These HPC systems are also available to support DOD HPCMP high-priority, time-critical requirements. In FY15, a project will be initiated to test the feasibility of providing a capability to submit jobs via a common point of entry for both systems and to access home directory data from either system. If successful, this will enable the establishment of collective allocations on the Cray XC30 HPC clusters, which would allow users to execute allocated jobs on either system.

NEW LEADERSHIP

at Maui High Performance Computing Center DSRC

By Betty Duncan, MHPCC DSRC

Major Micah O'Neal,
Acting Director
Maui High Performance Computing Center (MHPCC)
DOD Supercomputing Resource Center (DSRC)

In June 2014, Major Micah O'Neal took the reins as Acting Director of the Maui High Performance Computing Center (MHPCC) Department of Defense (DOD) Supercomputing Resource Center (DSRC) upon the departure of former MHPCC DSRC director David Morton. Major O'Neal has enthusiastically embraced the challenges and opportunities this leadership role provides and is developing a vision for the MHPCC DSRC that includes the continuing support and development of specialized services and software such as Dedicated System Partitions (DSPs) and the HPC Portal.

Major O'Neal provides war-winning capability and support for several DOD missions, including technical tasks for research and development, supercomputing resources, and collaborations with multiple government and civilian agencies.

Major O'Neal began his career with the Air Force in 1993 as a basic training trainee. His career spans various assignments and locations serving as a services apprentice, education and training manager, program manager, and deployed contracting officer at product centers, major commands and major command headquarters, and AFRL. Major O'Neal's experience covers program management and contracting.

In 2002, Major O'Neal received a Bachelor of Science degree in Business Management from Saint Leo University in Florida. In 2006 he received a Master of Business Administration from Strayer University in Washington, DC. To date, he has earned many outstanding honor awards in his Air Force career.

MAJOR AWARDS AND DECORATIONS

- Defense Meritorious Service Medal
- Air Force Commendation Medal with Oak Leaf Cluster
- Joint Service Commendation Medal
- Air Force Organizational Unit Award
- Air Force Outstanding Unit Award with 2 Oak Leaf Clusters
- Iraq Campaign Medal
- Global War on Terrorism Expeditionary Medal
- Global War on Terrorism Service Medal
- Armed Forces Expeditionary Service Ribbon w/Gold Border

First Word *(continued from page 1)*

New technology is in our DNA; the HPCMP is the largest S&T program in the Army, and has remained one of the largest in DOD. Our emphasis is always on deploying technologies that are ready for you to use effectively, and this is reflected in the increasing deployment of accelerators and associated software tools, training, and expertise through the centers and PETTT to ensure you can make effective use of these technologies. We are also increasing our investments in technology futures; over the past several years we, along with DOE and other federal partners, have made significant investments in new parallel runtime environments and programming models. While it will be a few more years before these investments yield tools that are useful to the general HPC community, they hold much promise for more effective expression of parallel work—a key aspect of using the systems we deploy today more effectively and preparing for the exascale systems of the next decade. Finally, we are also modestly increasing our investments and partnerships in novel computational hardware. Although other federal and academic partners will continue to lead here, we have a meaningful contribution to make in this area as well.

An HPC-capable federal workforce is the key to successful application of the HPCMP's expertise to the DOD's challenges. In recent years we have increased our focus on workforce development, and in 2013 I established a top leadership position in our office to manage this investment. During this time we have grown our efforts to connect early with high school and undergraduate students to create a computational pipeline for the department, and we have increased our efforts to develop effective faculty mentors for these students year-round. We are also starting an exciting effort with the Defense Acquisition University to explore ways in which we can introduce formal "HPC 101" training into the on-demand training provided by DAU to the DOD workforce. We have a long way to go on this front, but a successful partnership with DAU will provide a general introduction to the ways in which computation can help the department make better decisions and increase operational effectiveness. Of course our in-depth training for HPC practitioners through PETTT continues and is an important aspect of the services we offer.

The DOD is still experiencing substantial change as new demands in when and how the department supports the Nation translate into new acquisition, financial, security, and legal requirements. The structural changes I mentioned in this space over a year ago are still underway today. These changes have sometimes meant that important capabilities we would like to deliver to you are being delayed. Although these forces are beyond our control, we share your passion for your job, and we feel the frustration you are experiencing.

Through it all, however, we must remember that the work we are doing together—and the work we make possible for the DOD—is part of a much bigger picture. Together we are improving the safety and effectiveness of our Nation's fighting forces. This is a job worth doing, even when our circumstances make it more difficult than we would otherwise wish.

From the Director's Desk AFRL DSRC *(continued from page 21)*

Next, the larger of the first two Frontier projects established by the HPCMP is off and running on our other supercomputer, the 72,000-core SGI ICE X system named Spirit. Dr. Ryan Gosse and Dr. Nicholas Bisek, along with a diverse research team, have been experimenting with variations of their code that can currently scale well running on as many as 46,000 cores. Dr. Gosse and Dr. Bisek are addressing an urgent need within the Air Force to deliver long-range precision strike capability. Their goal is to design effective hypersonic transport vehicles using sophisticated computational techniques that can harness Spirit's maximum potential. To do that, their team plans to use up to 890 million core-hours per year by 2017. We are taking every precautionary measure imaginable to identify and remove any potential roadblocks that could hinder their progress over the next three years.

Along with the new HPCMP large-scale systems, we are working with two separate groups from the AFRL Sensors Directorate to design, implement, and maintain awarded DHPI systems. These systems offer an alternative to our customers who are not well-served by the shared nature of systems like Spirit and Lightning. These research groups receive funding for smaller-scale, dedicated supercomputers to advance their research objectives. Gotcha team members Linda Moore, Mike Minardi, and Steven Scarborough are working with Kevin Schoen on our staff to develop the "John Carter" system, named in honor of our late colleague. Another team, led by Robert DeSonia and Mike Butcher, is modeling countermeasures for infrared heat-seeking missiles using classified computational capability purchased by the HPCMP. This research will lead to improved missile avoidance systems for U.S. airborne platforms, and it will reduce costs by evaluating the readiness for live-fire test and evaluation (T&E) efforts. Also, we are developing more systems for placement in the Specialized High-performance Advanced Research Computing (SHARC) Lab that occupied the older computer room floor in Building 676. We are currently working with Mark Hagenmaier to design a $700K system to meet his requirements. In addition, we are exploring the potential of a redeployed system at the Sensors Directorate using parts of Harold and the old Gotcha DHPI system to aid Bob Simpson's team in their research objectives. At this rate, we may fill up that floor with SHARC systems!

And perhaps most importantly, Mr. Kelly Dalton, our new Technical Director, began his tenure with a new and comprehensive outreach project focused on Air Force needs for supercomputing. This effort has been endorsed by the Executive Director and Chief Scientist of AFRL and has several key objectives. First, we want to increase awareness within the Air Force of our powerful computational tools and our other resources.

Second, we want to reach out to our Air Force and DOD customers who are underserved or unfamiliar with how to make HPC systems work for them. Our third objective is to streamline access to HPCs so that new users can experiment with proof-of-concept projects first before committing to a formal project. Finally, we aim to work closely with each Technical Directorate within AFRL and other Air Force organizations to gain a better understanding of how they are using HPC systems, and to help them "right-size" their allocation to their mission objectives.

We are putting substantial focus on reaching communities outside the science and technology customer base. We particularly want to increase HPC usage within the DOD acquisition community. We can accomplish this goal by reaching out to the modeling and simulation specialists who conduct system effectiveness studies and provide input into program milestone decisions, and by talking to groups who perform test and evaluation of systems for the warfighter.

Also noteworthy is the transition to the HPCMP Integrated Technical Services (HITS) contract, which will replace the existing technical services contract with Lockheed Martin that provides the workforce required to operate four of the five DSRCs in the program. We worked hard to award this contract, and now we are anxious to maximize the impact of the new agreement. Through all of our activities this year, it is easy to see that the future of the AFRL DSRC shines bright with opportunities to increase the impact of high performance computing for the DOD.

DOD HPC INSIGHTS

is a semiannual publication of the High Performance Computing Modernization Program.

Editorial Staff

Managing Editor
Rachelle Hintson

DREN
Stephanie Parker

CREATE
Kevin Newmeyer

DAAC
Dr. Mike Stephens

UAG
Soumya Patnaik

PETTT
Casey Church
Lisa Diletti

WD
Pam Gilliam

HPCMPO
Cathy McDonald

HPCMPO Editor
Denise O'Donnell

Design / Layout
Ron Compton

Associate Director for HPC Centers
Christine E. Cuicchi

DSRCs

AFRL
Gregg Anderson

MHPCC
Betty Duncan

ARL
Brian Simmonds

Navy
Bryan Comstock

ERDC
Scotty Swillie
Nikkita Mays

Questions or Comments
feedback@centers.hpc.mil

The High Performance Computing Modernization Program (HPCMP) provides the Department of Defense supercomputing capabilities, high-speed network communications, and computational science expertise that enable DOD scientists and engineers to conduct a wide range of focused research, development, and test activities. This partnership puts advanced technology in the hands of U.S. forces more quickly, less expensively, and with greater certainty of success.

Today, the HPCMP provides a complete advanced computing environment for the DOD that includes unique expertise in software development and system design, powerful high performance computing systems, and a premier wide-area research network. The HPCMP is managed on behalf of the Department of Defense by the U.S. Army Engineer Research and Development Center.

To learn more about the HPCMP, please visit our website at: hpc.mil.

To learn more about our High Performance Computing (HPC) centers, please visit: centers.hpc.mil.

HPCMP User Customer Service Center
1.877.222.2039
centers.hpc.mil

POWERFUL RESOURCES

AFRL DSRC

US Air Force Research Laboratory
Wright-Patterson AFB, Ohio
www.afrl.hpc.mil

ARL DSRC

US Army Research Laboratory
Aberdeen Proving Ground, Maryland
www.arl.hpc.mil

ERDC DSRC

US Army Engineer Research & Development Center
Vicksburg, Mississippi
www.erdc.hpc.mil

NAVY DSRC

US Navy DOD Supercomputing Resource Center
Stennis Space Center, Mississippi
www.navydsrc.hpc.mil

MHPCC DSRC

Maui High Performance Computing Center
Kihei, Maui, Hawaii
www.mhpcc.hpc.mil