

# Navigator

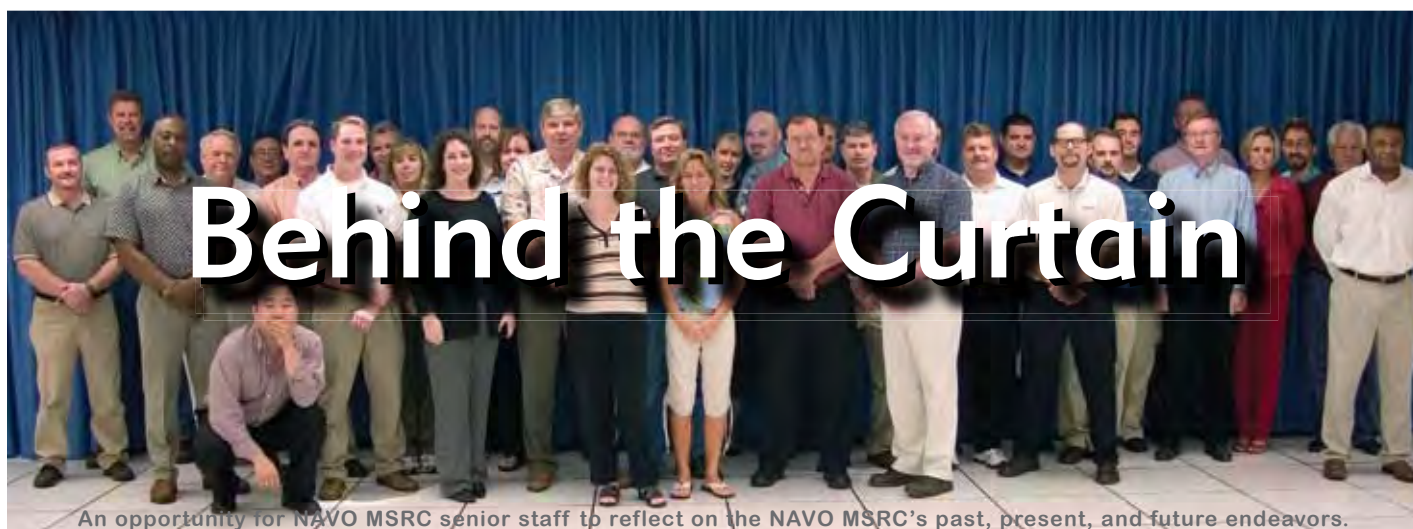
NAVO MSRC

FALL 2008



***Providing HPC Resources  
to DoD Communities and Beyond...***

| Report Documentation Page                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                    |                                     |                                                           | Form Approved<br>OMB No. 0704-0188                  |                                 |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|-------------------------------------|-----------------------------------------------------------|-----------------------------------------------------|---------------------------------|
| Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. |                                    |                                     |                                                           |                                                     |                                 |
| 1. REPORT DATE<br><b>2008</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                    | 2. REPORT TYPE                      |                                                           | 3. DATES COVERED<br><b>00-00-2008 to 00-00-2008</b> |                                 |
| 4. TITLE AND SUBTITLE<br><b>NAVO MSRC Navigator. Fall 2008</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                    |                                     |                                                           | 5a. CONTRACT NUMBER                                 |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5b. GRANT NUMBER                                    |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5c. PROGRAM ELEMENT NUMBER                          |                                 |
| 6. AUTHOR(S)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                    |                                     |                                                           | 5d. PROJECT NUMBER                                  |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5e. TASK NUMBER                                     |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 5f. WORK UNIT NUMBER                                |                                 |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br><b>Naval Oceanographic Office (NAVO),Major Shared Resource Center (MSRC),1002 Balch Boulevard,Stennis Space Center,MS,39522</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                    |                                     |                                                           | 8. PERFORMING ORGANIZATION REPORT NUMBER            |                                 |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                    |                                     |                                                           | 10. SPONSOR/MONITOR'S ACRONYM(S)                    |                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     |                                                           | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)              |                                 |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT<br><b>Approved for public release; distribution unlimited</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                    |                                     |                                                           |                                                     |                                 |
| 13. SUPPLEMENTARY NOTES                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                    |                                     |                                                           |                                                     |                                 |
| 14. ABSTRACT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                    |                                     |                                                           |                                                     |                                 |
| 15. SUBJECT TERMS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                    |                                     |                                                           |                                                     |                                 |
| 16. SECURITY CLASSIFICATION OF:                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                    |                                     | 17. LIMITATION OF ABSTRACT<br><b>Same as Report (SAR)</b> | 18. NUMBER OF PAGES<br><b>20</b>                    | 19a. NAME OF RESPONSIBLE PERSON |
| a. REPORT<br><b>unclassified</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | b. ABSTRACT<br><b>unclassified</b> | c. THIS PAGE<br><b>unclassified</b> |                                                           |                                                     |                                 |



An opportunity for NAVO MSRC senior staff to reflect on the NAVO MSRC's past, present, and future endeavors.

***"When fortune comes, seize her firmly by the forelock, for, I tell you, she is bald at the back."***

*- Leonardo Da Vinci*

When the NAVO MSRC found itself with a two-thirds empty machine room floor after the IBM P4 system MARCELLUS was removed, Director Tom Dunn seized the opportunity to kick off a number of facility upgrades in anticipation of the demands that new High Performance Computing (HPC) systems would place on the Center's infrastructure. Dave Cole, Acting Associate Director (Plans and Programs), details these upgrades on page 5—along with a long-anticipated update on the dogs that accompanied him and his wife, Leona, as they evacuated from Hurricane Katrina in 2005. The facility upgrades, overseen by Jennifer Rabert and later, new government staff member Rob Thornhill of the NAVO MSRC, were begun and completed in 2008 in time for the arrival of our two new HPC systems, DAVINCI (IBM P6) and EINSTEIN (Cray XT5), and our new mass storage server, NEWTON (Sun M5000).

***"The most incomprehensible thing about the world is that it is at all comprehensible."***

*- Albert Einstein*

The computing power that these new systems bring will make the NAVO MSRC the most powerful in the Department of Defense High Performance Computing Modernization Program—for the time being. It is a function of the industry that those on top are never on top for long; we continue to lean on our IBM P5+ workhorses BABBAGE and PASCAL while diversifying and vastly increasing the scientific computing capabilities we provide to our users with our new arrivals. Four Capability Application Projects (CAP) will run on both DAVINCI and EINSTEIN, providing researchers with the

capability of running jobs of up to 4,256 and 12,736 cores in size, respectively. We continue to take pride in our role in nurturing terascale scientific research and engineering in support of the warfighter.

***"If I have seen further, it is by standing on the shoulders of giants."***

*- Isaac Newton*

Onsite expertise and data storage: these two elements provide the critical foundation for the computational

## In Other Words...

**Christine Cuicchi**  
Computational Science and Applications Lead,  
NAVO MSRC

science accomplished on our systems. Although it has received decidedly less press than our incoming HPC systems, our new Sun M5000 archive server NEWTON will provide even more stability and faster data archive access to support the data storage demands, which are expected to double as EINSTEIN and DAVINCI are brought online. We have also strengthened the backbone of our Disaster Recovery infrastructure, as detailed on page 9. Over the past six months we have increased our original government staff of seven to eleven, strengthening our technical and facilities expertise; please take an opportunity to meet our newest team members starting on page 14.

It has been, and continues to be, our pleasure to adapt, expand, and advance to serve the needs of our user community.

**The Naval Oceanographic Office (NAVO)  
Major Shared Resource Center (MSRC):  
Delivering Science to the Warfighter**

The NAVO MSRC provides Department of Defense (DoD) scientists and engineers with high performance computing (HPC) resources, including leading edge computational systems, large-scale data storage and archiving, scientific visualization resources and training, and expertise in specific computational technology areas (CTAs). These CTAs include Computational Fluid Dynamics (CFD), Climate/Weather/Ocean Modeling and Simulation (CWO), Environmental Quality Modeling and Simulation (EQM), Computational Electromagnetic and Acoustics (CEA), and Signal/Image Processing (SIP).

NAVO MSRC  
Code 0TC  
1002 Balch Boulevard  
Stennis Space Center, MS 39522  
1-800-993-7677 or  
msrchelp@navo.hpc.mil

**NAVO MSRC Navigator**  
[www.navo.hpc.mil/Navigator](http://www.navo.hpc.mil/Navigator)

NAVO MSRC Navigator is a biannual technical publication designed to inform users of the news, events, people, accomplishments, and activities of the Center. For a free subscription or to make address changes, contact NAVO MSRC at the above address.

EDITOR:  
Gioia Furness Petro, [petrogio@navo.hpc.mil](mailto:petrogio@navo.hpc.mil)

DESIGNERS:  
Kerry Townson, [ktownson@navo.hpc.mil](mailto:ktownson@navo.hpc.mil)  
Lynn Yott, [lynn@navo.hpc.mil](mailto:lynn@navo.hpc.mil)

Any opinions, conclusions, or recommendations in this publication are those of the author(s) and do not necessarily reflect those of the Navy or NAVO MSRC. All brand names and product names are trademarks or registered trademarks of their respective holders. These names are for information purposes only and do not imply endorsement by the Navy or NAVO MSRC.

Approved for Public Release  
Distribution Unlimited

# Contents

## NAVO MSRC Update

- 2 In Other Words...

## Feature Articles

- 5 Five Dogs, a Hurricane, and Two IBM Power5+s - a Third Year Update
- 7 Advance Reservations: Back on Babbage and Better than Ever
- 9 Quarterly Tests Reveal Exceptional Performance Stability for IBM Platforms at NAVO MSRC
- 10 DAVINCI and EINSTEIN Bring Their Computing Power to NAVO MSRC

## The Porthole

- 12 Visitors to the NAVO MSRC
- 14 Welcome Aboard!

## Navigator Tools and Tips

- 16 Using PBS at the NAVO MSRC

## Upcoming Events

- 19 Coming Events



# Five Dogs, a Hurricane, and Two IBM Power5+s – a Third Year Update

Dave Cole, NAVO MSRC

## NAVO MSRC EXPANDS AGAIN

Located at Stennis Space Center (SSC) in Mississippi, the NAVO MSRC maintains and provides premier High Performance Computing (HPC) capability with primary emphases on support of the largest, most computationally intensive HPC applications and delivery of time-critical HPC services to directly support Department of Defense operations worldwide. The large scale, power hungry HPC systems used to satisfy the computational requirements place extraordinary demands on Center infrastructure.

After installation of the Technical Insertion 2004 (TI-04) systems, the building hosting HPC systems for the NAVO MSRC lacked sufficient space, power, and cooling capabilities to support installation of additional supercomputers.

In anticipation of future support requirements, our former Director Steve Adamec developed an innovative plan to renovate a steel reinforced concrete building that had been in mothball status for more than a decade at the Army Ammunition Plant complex located at the SSC.

Implementation of the plan had just begun when Hurricane Katrina struck the Mississippi Gulf Coast on 29 August 2005. The facility easily withstood the battering hurricane force winds including the newly installed metal roof rated for sustained wind speeds up to 150 miles per hour.

Three years ago, the Spring 2006 edition of the Navigator included an article that gave an account of my evacuation from the Mississippi Gulf Coast prior to landfall of Hurricane Katrina, a brief description of the journey home, and my eventual return to work at the NAVO MSRC.

The intent of the article was to provide a personalized Katrina account to provide insight into the many challenges faced by members of the NAVO MSRC community. It also told the story of the rescue of two beagles from a local kennel for a friend who was out of the country and their addition to our "pack" of terriers – an addition that brought the pet evacuation count to five small frisky dogs.

This article serves as a sequel to let the members of the High Performance Computing Modernization Program (HPCMP) community know about our recovery and to briefly describe

the results of significant facilities infrastructure modifications — modifications that were accomplished at a time when many NAVO MSRC members were rebuilding their homes and their lives. Oh yes, it also includes an update of how the five small frisky dogs have fared.

Renovation work at the new facility (See box, this page) resumed three weeks after the Gulf Coast and the NAVO MSRC began recovery efforts and was completed in time to support installation of the IBM P5+ systems BABBAGE and PASCAL.

The new facility provided approximately 11,000 square feet of 30-inch raised floor space with approximately 800 tons of cooling capacity, and a 1000 Kilowatt (KW) Uninterruptible Power System (UPS) for conditioned power. With the

**Continued Next Page...**



**NAVO MSRC space containing BABBAGE and DAVINCI.**

addition of a second 1000 KW UPS to support the installation of the Technical Insertion for Fiscal Year 2008 (TI-08) IBM P6 DAVINCI, 2000 KW of conditioned and backup generator power are now available for High Performance Computing (HPC) systems support.

With the decommissioning and removal of the IBM P4+ MARCELLUS system in January 2007 from the original NAVO facility, MSRC Director Tom Dunn recognized this as a unique opportunity to renovate our aging facilities infrastructure.

Renovation requirements were presented to the High Performance Computing Modernization Office (HPCMPO) in December 2007 and were subsequently approved for implementation. Upgrades to the original MSRC facility include:

- Removal of unused power whips and cables under the raised floor and replacement of floor water detector tapes.
- Replacement of the above floor fire suppression system

with overhead FM200 and removal of an existing water based system.

- Replacement of the raised floor that includes tiles rated at 2000 pounds per square inch.
- Installation of four 60-ton computer room units and modification of the false ceiling to provide sufficient cooling and air flow to accommodate large scale HPC systems.

Completed in time for the delivery of the Cray XT5 Einstein in September, the renovated space offers approximately 7500 square feet of 18-inch raised floor space with 800 tons of cooling capacity, 1000 KW of conditioned power, and 1750 KW of backup generator power for HPC systems support.

As described in the closing paragraph of the Spring 2006 edition of the Navigator, Hurricane Katrina severely impacted the lives of the members of the MSRC. Thirty percent of the MSRC team suffered catastrophic

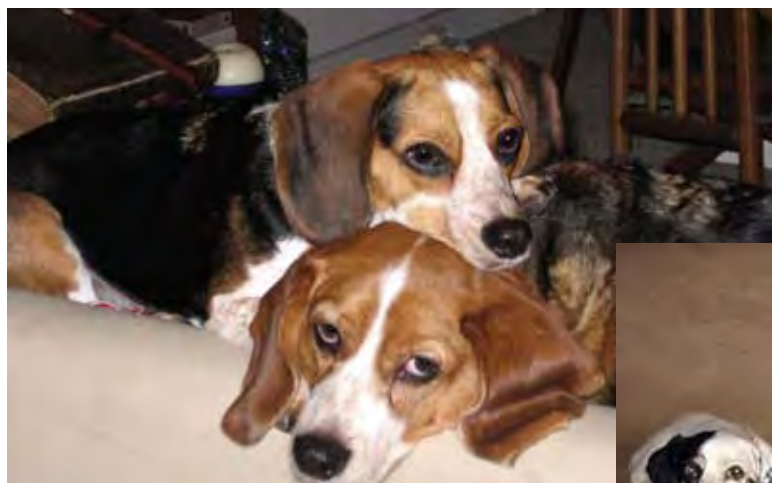
damage to their homes, rendering the home uninhabitable or totally devastated.

The transformation of a building in mothball status at the Army Ammunition Complex into a world-class HPC support facility at a time when many team members were rebuilding their homes and their lives demonstrates their commitment to the HPCMP community.

The recent renovation of the original MSRC facility further demonstrates the NAVO MSRC support of the HPCMP mission to accelerate the development of advanced defense technologies for the warfighter via supercomputing.

And last, but not least, the long-awaited update on the five small dogs: during the evacuation they experienced a sense of displacement and anxiety that was further deepened by their brief stay at a "two star" dog kennel in my hometown of Minden, LA.

Reunited with their owner, the two beagles have fully recovered and start each morning by waking up the neighborhood. The three "terrible" terriers are thrilled with their new home, which came with a big backyard enclosed by a privacy fence. In short, the five small dogs are as frisky as ever.



**Five frisky dogs - two singing beagles and three "terrible terriers" - taking it easy.**

# Advance Reservations: Back on Babbage and Better than Ever

Christine Cuicchi, Computational Science and Applications Lead

Want it all, and want it now? We can't give you the whole system on demand, but we can give you up to 256 processors with an advance reservation on the P5+ system, BABBAGE. The Advance Reservation System (ARS), which was available on KRAKEN, has been moved to BABBAGE for Fiscal Year 2009 and is available to any user with BABBAGE allocation.

As with the previous capability on KRAKEN, users may request a specific available run time to run interactive jobs, real-time simulations, or batch jobs. Users will authenticate to the front end of the ARS portal (link available at <http://www.navo.hpc.mil>) via Kerberos (Figure 1). First-

time users will be directed to request an account on the online reservation portal of the ARS. This account is not related to allocated accounts on the High Performance Computing (HPC) systems.

After establishing an online reservation portal account and logging in to the portal screen as shown in Figure 2, a user will be able to see how many processors are available for reservation on any given day in the calendar view. By simply clicking on the day being targeted for reservation, a user will see an hour-by-hour list of available processor counts (Figure 3).

Users may request--via drop-down menus--a number of processors (up

to 256, divisible by 16 processors per node) for a user-specified amount of time of up to 48 hours, delineated in one-hour increments. Users will also be able to specify, via a drop-down menu, the project to which the reservation hours will be charged. In the example shown in Figures 4 and 5, a request is being made for 48 processors (4 nodes) for 4 hours on 10 October 2008, starting at 1700 Greenwich Mean Time (GMT).

Checking the available hours for 10 October via the calendar now shows that there are 48 fewer processors available during the time period for

**Continued Next Page...**



The screenshot shows the 'Kerberos Login' page. It has a header with a logo and the title 'Kerberos Login'. Below the title, there are instructions: 'You need to enter your username in the Principal field.' and 'You need to enter your Kerberos password in the Password field.' There are three input fields: 'Principal' (with a dropdown menu showing 'NAVO HPC.MIL'), 'Password', and 'Token Code'. There are 'Login' and 'Clear' buttons. At the bottom, there is a note: 'You must first log in with your Kerberos principal and hardware token. Upon successful Kerberos login, you will be redirected to the reservation system login screen, where you will then login with your reservation username and password. If you experience multiple Kerberos login failures, please exit the browser window and open a new one and attempt your login again. If you would like to use your Fido/CAC instead of Kerberos, please read [www.navo.hpc.mil](http://www.navo.hpc.mil)'.

**Figure 1. ARS Kerberos login.**



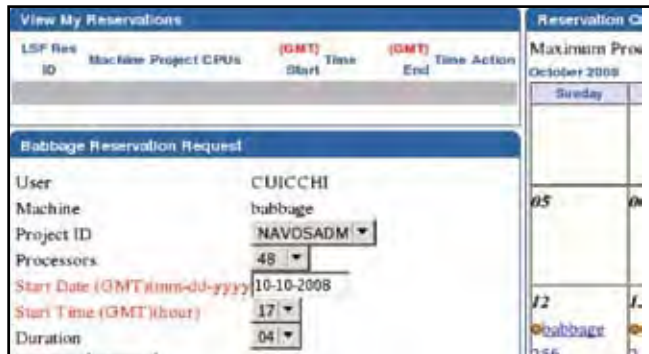
The screenshot shows the 'Sign In' page. It has a header with the title 'Sign In'. Below the title, there are instructions: 'Enter your Single Sign-On user name and password to sign in'. There are two input fields: 'User Name' and 'Password'. There are 'Login' and 'Cancel' buttons. At the bottom, there is a note: 'Unauthorized use of this site is prohibited and may subject you to civil and criminal prosecution.'

**Figure 2. ARS Portal login screen.**



| Day          |       | Processors Available |                        |
|--------------|-------|----------------------|------------------------|
| Oct 10, 2008 | 00:00 | 0                    | Oct 10, 2008 12:00 192 |
| Oct 10, 2008 | 01:00 | 0                    | Oct 10, 2008 13:00 192 |
| Oct 10, 2008 | 02:00 | 0                    | Oct 10, 2008 14:00 192 |
| Oct 10, 2008 | 03:00 | 0                    | Oct 10, 2008 15:00 192 |
| Oct 10, 2008 | 04:00 | 0                    | Oct 10, 2008 16:00 192 |
| Oct 10, 2008 | 05:00 | 0                    | Oct 10, 2008 17:00 192 |
| Oct 10, 2008 | 06:00 | 0                    | Oct 10, 2008 18:00 192 |
| Oct 10, 2008 | 07:00 | 192                  | Oct 10, 2008 19:00 192 |
| Oct 10, 2008 | 08:00 | 192                  | Oct 10, 2008 20:00 192 |
| Oct 10, 2008 | 09:00 | 192                  | Oct 10, 2008 21:00 192 |
| Oct 10, 2008 | 10:00 | 192                  | Oct 10, 2008 22:00 192 |
| Oct 10, 2008 | 11:00 | 192                  | Oct 10, 2008 23:00 192 |

**Figure 3. Hour-by-hour list of available processors.**



The screenshot shows the 'Babbage Reservation Request' form. It has a header with the title 'Babbage Reservation Request'. Below the title, there are several fields: 'User' (CUICCHI), 'Machine' (babbage), 'Project ID' (NAVOADM), 'Processors' (48), 'Start Date (GMT)(mm-dd-yyyy)' (10-10-2008), 'Start Time (GMT)(hour)' (17), and 'Duration' (04). There are 'Login' and 'Cancel' buttons. On the right side, there is a 'Reservation Calendar' showing the month of October 2008. The calendar has a grid with days of the week and dates. The date 10 is highlighted, and the number 48 is shown in the cell for 10/10. The calendar also shows the maximum processors available for each day.

**Figure 4. Selecting reservation options.**

which the reservation was requested (See Figure 6).

Once the reservation request is made, the user will receive an email containing a list of the specific nodes reserved as well as the Load Sharing Facility (LSF) reservation identification number (reservation\_id) to be used in submitting jobs to the reserved nodes. This reservation identification number must be used to submit batch jobs in the following manner:

**bsub -U reservation\_id ... < script**

Users who wish to run interactive jobs should submit them at the beginning of the reservation. Running the showq command on BABBAGE will allow users to see a list of reservations as well, as illustrated in Code 1.

A user can also check the status of a reservation at the ARS portal as shown in Figure 7.

All reservations will begin at the start time the user has requested. After receiving the confirmation email, users

will have the ability to cancel their reservations up to 30 minutes prior to the reservation start time (See Figure 7).

It is important to note that if the reservation is not cancelled, system utilization will be charged for these nodes regardless of how, or if, the nodes are used.

In its previous incarnation, the ARS ran on dedicated nodes, which ensured that the nodes were available to users for advance reservation at all times. In the spring of 2008, the NAVO MSRC added a backfill capability to ARS so that the nodes earmarked for advance reservation

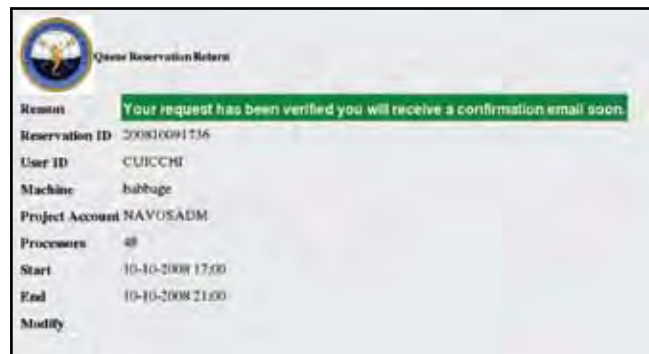
would still be available to regular batch queues when demand for advance reservations waned.

Users planning advance reservation requests should take into account that some, or all, of these nodes may be backfilled for up to 48 hours from the present time. Again, node availability by the hour can be found by clicking on the calendar day in which the user is interested in placing a reservation.

The NAVO MSRC will be a leading site in reproducing this advance reservation capability under the Portable Batch System (PBS) Pro workload management system. Future plans call for the possible

development of a single ARS portal that would handle advance reservation requests on a number of HPC systems across the Department of Defense HPC Modernization Program.

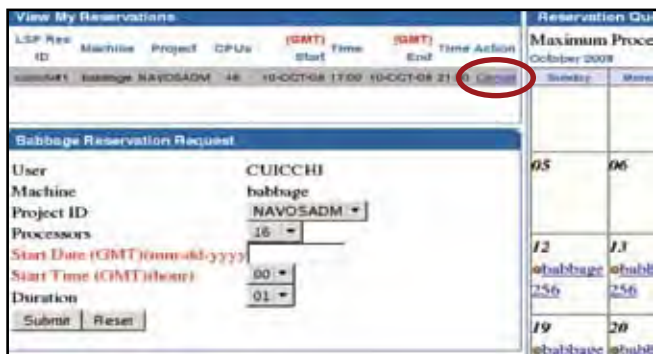
For more information about the ARS and access to the ARS portal, please visit: <http://www.navo.hpc.mil>.



**Figure 5. Notification of successful reservation request.**

| Day                | Processors Available   |
|--------------------|------------------------|
| Oct 10, 2008 00:00 | Oct 10, 2008 12:00 192 |
| Oct 10, 2008 01:00 | Oct 10, 2008 13:00 192 |
| Oct 10, 2008 02:00 | Oct 10, 2008 14:00 192 |
| Oct 10, 2008 03:00 | Oct 10, 2008 15:00 192 |
| Oct 10, 2008 04:00 | Oct 10, 2008 16:00 192 |
| Oct 10, 2008 05:00 | Oct 10, 2008 17:00 144 |
| Oct 10, 2008 06:00 | Oct 10, 2008 18:00 144 |
| Oct 10, 2008 07:00 | Oct 10, 2008 19:00 144 |
| Oct 10, 2008 08:00 | Oct 10, 2008 20:00 144 |
| Oct 10, 2008 09:00 | Oct 10, 2008 21:00 192 |
| Oct 10, 2008 10:00 | Oct 10, 2008 22:00 192 |
| Oct 10, 2008 11:00 | Oct 10, 2008 23:00 192 |

**Figure 6. Reduction of available cores due to requested reservation.**



**Figure 7. Reservation status and cancel option.**

|                                   |                         |                                                            |
|-----------------------------------|-------------------------|------------------------------------------------------------|
| <b>b5n1...cuicchi&gt; showq</b>   |                         |                                                            |
| <b>ADVANCED RESERVATIONS-----</b> |                         |                                                            |
| <b>RESV ID</b>                    | <b>PROC</b>             | <b>RESERVATION WINDOW</b>                                  |
| <b>res# 1 : time1=1223658000</b>  | <b>time2=1223672400</b> |                                                            |
| <b>cuicchi#1</b>                  | <b>48</b>               | <b>Fri Oct 10 17:00:00 2008 - Fri Oct 10 21:00:00 2008</b> |

**Code 1. Advance reservations as they appear in showq command output.**

# Quarterly Tests Reveal Exceptional Performance Stability for IBM Platforms at NAVO MSRC

Christine Cuicchi, Computational Science and Applications Lead, NAVO MSRC

Dr. Paul Bennett, Computational Science and Engineering (CS&E) Group, Engineer Research Development Center (ERDC) MSRC

Sustained system performance—it's something most users expect of the well-developed high performance computing (HPC) systems—and the IBM P4+ and P5+ platforms at NAVO MSRC have demonstrated excellent sustained performance over the past several years.

Every three months, Dr. Paul Bennett of the ERDC MSRC CS&E group runs a Sustained System Performance (SSP) test on Department of Defense (DoD) High Performance Computing Modernization Program (HPCMP) systems to evaluate and compare the systems' performance over their time in service to the Program.

The test, modeled after the National Energy Research Scientific Computing (NERSC) Center's SSP test<sup>1</sup>, consists of a number of application codes and input data sets that represent a composite of the HPCMP's systems workload. Each application is run on both "standard" and "large" size data sets, and the full suite as characterized by Dr. Bennett will, "stress the processing elements, main memory, and file I/O systems to determine the existence of compiler

optimization issues, issues with communication libraries, problems with I/O subsystems, problems with libraries that have been recompiled in a different way, software or hardware issues specific to the interconnect, changes to the runtime environment, or the application of security patches that adversely affect performance."

Figures 1 and 2 show the sustained system performance of KRAKEN (IBM P4+) and BABBAGE (IBM P5+) per application code over their lifetimes. System performance for both systems is presented as relative to the system's original Technology Insertion (TI) benchmarked application performance in TI-04 and TI-06, respectively. The SSP tests reveal that KRAKEN's relative performance increased over its time in service while remaining fairly stable over its last year of life and that BABBAGE's relative performance has been extremely stable with a slight increase in performance for a number of application codes. These SSP tests will also be run on DAVINCI and EINSTEIN once the two new systems enter their allocated phases at NAVO MSRC.

Figure 1

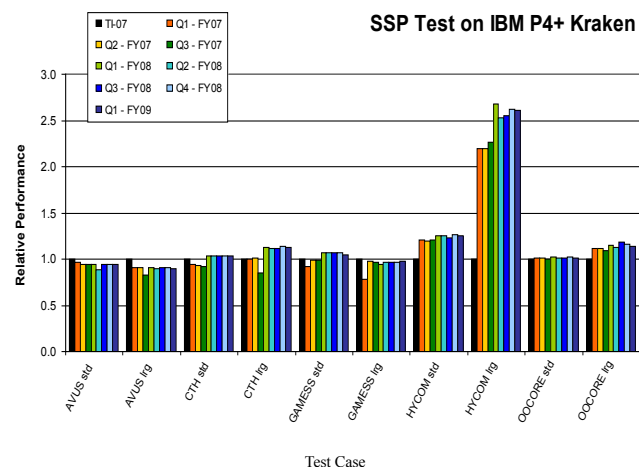
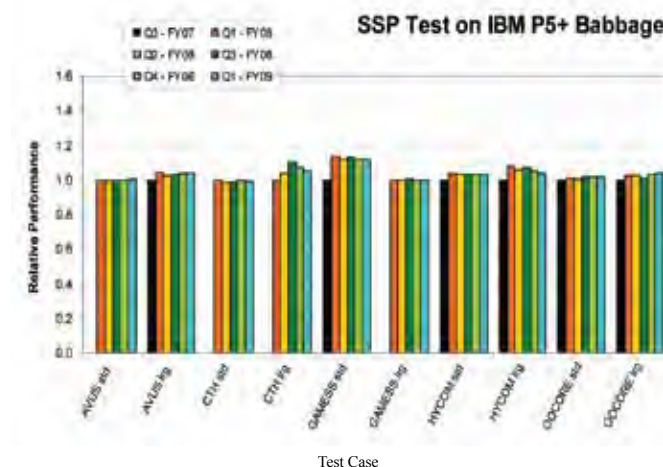


Figure 2



1. W. Kramer, J. Shalf, E. Strohmaier, "The NERSC Sustained System Performance (SSP) Metric," Paper LBNL-58868, Lawrence Berkeley National Laboratory, 2005.

# DAVINCI AND EINSTEIN BRING THEIR COMPUTING POWER TO NAVO MSRC

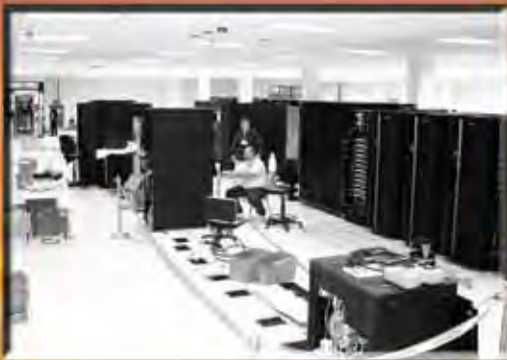
**T**he summer of 2008 has been an eventful one at the NAVO MSRC. Two major happenings were the installations of both the Cray XT5 EINSTEIN and the IBM P6 DAVINCI. With the completion of these installations and the decommissioning of our venerable IBM P4s (KRAKEN and ROMULUS), the total computing capacity at NAVO has quadrupled to 233 Teraflops (TFLOP).

To kick off this summer's festivities, the NAVO MSRC received DAVINCI, an 80 TFLOP IBM Power6 system packing 4,256 4.7 Gigahertz (GHz) compute cores. It also contains 10 Terabytes (TB) of main memory and 437 TB of disk space.

Another amazing feature of DAVINCI is that it is completely water-cooled – what IBM terms a “hydro-cluster.” Water is piped into the system and runs across the processors via copper plating to dissipate their heat. DAVINCI is also different in that each of its 14 cabinets has a door that passively removes heat (known as a heat-exchanger door). These doors also utilize water to reduce the amount of heat allowed to enter the computer room facility.

According to IBM, a hydro-cluster system similar in size to DAVINCI needs roughly 80 percent less air conditioning than a traditional air cooled system. Total energy consumption

## DAVINCI



With the installation of DAVINCI and EINSTEIN, the NAVO MSRC continues its mission to provide the best high performance computing technology available, with the best support possible, to the military and academic research communities.

should also be reduced by around 40 percent. These savings allow the NAVO MSRC to move towards becoming a more "green friendly" facility. Though the IBM Power series systems have been a mainstay at the NAVO MSRC for years, the DAVINCI system will prove to be a worthy successor.

The installation of the Cray XT5 EINSTEIN will forever be remembered as "the one that finally happened." As production of the system completed, Cray prepared to ship the system, and anticipation at the NAVO MSRC was on the rise. Little did we know that Mother Nature had other plans.

Initially, the system was to be delivered the week after Labor Day, but along came Hurricane Gustav to nix those plans. The shipment was delayed. As time drew near for the next attempt, Hurricane Ike decided to make his way into the Gulf of Mexico, bringing yet another delay in the shipment date. The system finally made its way to the MSRC in mid-September, but only after breaking the record for most hurricanes (two) to delay an installation. The previous record of one delay was held by the SAPPHIRE system at the U.S. Army Engineer Research and Development Center (ERDC) MSRC for its bout with Hurricane Katrina in 2005.

While already a record holder, EINSTEIN is also the largest in the Department of Defense in peak performance (117 TFLOP), number of cores (12,856), total available main memory (25 TB), and disk space (518 TB). It recently recorded a LINPACK number of 93 TFLOP. Another interesting EINSTEIN factoid is that it is one of a select number of Cray XT systems to have a vinyl skin applied to its exterior. The skinning of the new XT systems has become so popular that Cray now has a laminating press to apply it. The previous XT3 generation had a textured door that would not allow for an overlay. In response to the ERDC MSRC desire for its new XT4 system JADE to have camouflaged doors, Cray developed smoother cabinet doors to allow the skin to be applied. So, EINSTEIN, already an amazing system, looks good too. - Bryan Comstock, NAVO MSRC, Computer Scientist

## EINSTEIN



... AND THE BUILD  
GOES ON ...



# The Porthole



DoD HPCMPO, IBM, and Government Oversight representatives participate in the TI-08 Capabilities Test for the new IBM Power6 HPC system DAVINCI.

**CRAY XT5 EINSTEIN  
installation team.**



**IBM P6 DAVINCI  
installation team.**

**Teachers with the Mississippi State University-sponsored Industry-Education Partnerships Workshop view a briefing on the field of high performance computing and its role at the NAVO MSRC.**



**Casey Bretti (AFRL) and Sheila Carbonette (NAVO) participate in the Enhanced USER Experience team meeting at AFRL.**

**Lockheed Martin Database Administration Process Improvement Group gathered at NAVO September 23-25.**



# Welcome Aboard! NAVO MSRC Welcomes

## Tom Brown

Tom Brown arrived in September to be the Associate Director for Operations at the NAVO MSRC and the National Center for Information Processing and Storage (NCCIPS) data center at the Stennis Space Center, Mississippi. Tom comes to the MSRC and NCCIPS from Arnold Engineering Development Center, Tennessee, where he served in numerous leadership positions. Prior to his move to Stennis, he served as part of the Office of Secretary for Defense High Performance Computing Modernization Program Office (HPCMPO) senior staff as Deputy Centers Program Manager with responsibilities for managing and delivering services to all the DoD HPC MSRCs. "Over the years I've been fortunate to work with many of the great folks at NAVO as a colleague in the larger HPC community; now I'm proud to now be part of the Stennis team," said Tom.

Starting with the Department of the Air Force as a student aide, Tom has served in numerous Information Technology (IT) positions during his 30+ years of government service. His experience includes assignments at three Air Force bases with three different Major Commands (MAJCOMS) on Center and Headquarters MAJCOM staffs, and as senior member of the HPCMPO staff.

As a seasoned IT professional, Tom brings the NAVO MSRC and NCCIPS extensive knowledge of corporate leadership of Information Resource Management (IRM) programs and staff, communications, and computer systems planning and management.

He is also using his experience in IT acquisition, business/engineering computer and communications architectures, modeling and simulation systems design, software engineering, IT integration and program management, computer security, and telecommunications systems management to improve the MSRC and NCCIPS. He holds two masters degrees, and his executive training and experience includes the Defense Leadership and Management Program (December 2004 graduate) and Chief Information Officer Certification from the National Defense University.



## Rob Thornhill

Rob Thornhill is a native Mississippian who is the new NAVO MSRC Facilities Engineer. As such, he oversees all facilities planning, maintenance, and upgrades. With the recent renovation of the original MSRC space and the new facility and the installation of EINSTEIN and DAVINCI, Rob has been very, very busy. (See Five Dogs, a Hurricane, and Two IBM Power5+s – a Third Year Update on page 5 and DAVINCI and EINSTEIN Bring Their Computing Power to NAVO MSRC, page 10).

Though Rob is not new to the Naval Oceanographic Office (NAVOCEANO) (he started as a contractor in 2003 and joined the Government ranks in 2006), it wasn't until this spring that he joined the NAVO MSRC team.

Along the way from the University of Southern Mississippi to the NAVO MSRC, Rob took a detour through the aggregate mining industry ("That's a fancy way of saying 'I worked in a gravel pit.'"), where he learned the advantages of indoor work in the summer and winter, as well as maintenance and planning skills. These are

skills that he refined and applied to the information technology arena through his prior work with Navy Marine Corps Intranet (NMCI) implementation at NAVOCEANO.

Of his move to the NAVO MSRC from NAVOCEANO, Rob says, "While working with NAVOCEANO, I was presented an opportunity to become part of the MSRC team. I have always been somewhat in awe of the 'the big, fancy computers' and could not pass up the chance to actually become a part of it."



# New User Support Team Members

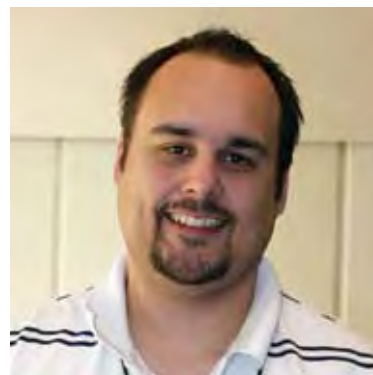
## Bryan Comstock

Bryan Comstock is a proud Gulf Coaster who recently joined the NAVO MSRC as the primary Government Point of Contact (POC) for the Consolidated Customer Assistance Center (CCAC), the Consolidated Software Initiative (CSI), the Customer Satisfaction Survey (CUSS) and the Enhanced User Environment (EUE). Taken down to its essence, this means: "I will have an active role in forming, monitoring, and modifying the policies, procedures, and activities that affect our users on a daily basis. As the backup Outreach and Challenge project POC, I assist the primary POC, Christine Cuicchi, in supporting the NAVO operational users and Challenge users. Being a new member of the government staff, I hope to continue to provide a high level of service in these areas."

Bryan started his road to the MSRC at the University of Southern Mississippi, which led to a position at the Mississippi State University's Engineering Research Center. Later, he joined the MSRC as a contractor providing user support analysis, and, after a brief foray as a Software Engineer for Planning Systems, Inc., he returned to the NAVO MSRC as part of its Government management team. One of the best parts of being part of the NAVO MSRC team, he says, is the opportunity to observe and

support the various scientific applications being run on the High Performance Computing (HPC) systems.

While he finds them all interesting, the ones relating to weather and mechanics are especially attractive: "Growing up on the Gulf Coast and seeing hurricanes firsthand, I've always had a fascination with weather; the Climate/Weather/Ocean (CWO) Modeling and Simulation Computational Technology Area (CTA) would have to be my favorite. Before going into computer science, I was considering becoming a mechanical engineer, so the Computation Fluid Dynamics (CFD) or Computational Structural Mechanics (CSM) CTAs also interest me."



## Morgan Harrison

Morgan Harrison, the newest NAVO MSRC computer scientist, comes to the NAVO MSRC after five years of system administration and user support that the Mississippi State University High Performance Computing Collaboratory (HPC2). Though part of the NAVO MSRC team for only three months (he jokes "I joined so I could say 'I'm from the Government and I'm here to help you.'"), Morgan is quickly integrating into the MSRC user support community and providing valuable technical expertise in systems management and overall MSRC system architectures.

Morgan's new position is, he states, the next logical progression in working with bigger, faster, more interesting High Performance Computing systems. Though he hasn't managed to do this yet at the NAVO MSRC, he says, "My favorite thing to do whenever I get a new computer is to pull the case off and see what the insides look like. When Operations and the System Administrators turn their backs, I might go at DAVINCI or EINSTEIN."

Working at the MSRC is exciting because, Morgan says, "Even though I haven't been part of the NAVO MSRC team long, I feel like I'm already contributing behind the scenes to improve our users' experience."



## Navigator Tools and Tips

# Using PBS at the NAVO MSRC

Sheila Carbonette, User Support, NAVO MSRC

Morgan Harrison, Computer Scientist, NAVO MSRC

Bryan Comstock, Computer Scientist, NAVO MSRC

The Portable Batch System (PBS) is available on the IBM P6 and CRAY XT systems at the NAVO MSRC.

This guide will serve as an overview of PBS and offer a comparison of PBS, Load Sharing Facility (LSF) and LoadLeveler queuing system commands, environment variables, and batch scripts.

In order to use PBS, users do not need to add anything to their startup files. PBS-related environment variables and paths have been added to the IBM and CRAY system files to set up PBS as part of the user login session.

PBS, LSF, and LoadLeveler are alike in that they all are products that schedule user jobs in a batch environment. All three schedulers have commands that allow users to submit jobs, check job and queue status, and hold/cancel jobs. The following tables provide a brief comparison of the commands and utilities of the schedulers as well as list the more common commands.

### PBS SCRIPT TIPS

Writing a PBS script is straightforward. However, a few tips regarding writing PBS scripts are always helpful:

- Do not use PBS environment variables in the #PBS directives.

- Standard output and error can be combined into one file using the “-j oe <filename>” option.

- Omitting the -o and/or -e option(s) sends std output and/or error to a filename in the format of:

<job name>.[oe]<job id>.

- Omitting a job name results in output and error filenames in the format of:

<PBS script name>.[oe]<job id>.

### SAMPLE PBS SCRIPTS

A PBS script to run a serial job will look something like this:

```
#!/bin/csh
```

```
#PBS -N serialjob                # Name of the job.
```

```
#PBS -o serialjob.out            # Appends std output to  
                                #file serialjob.out.
```

```
#PBS -e serialjob.err            # Appends std error to  
                                #file serialjob.err.
```

```
#PBS -A NAVOSLMA                 # Charging Project ID.
```

```
#PBS -l walltime=01:30:00        # Wall clock time of
```

### Queuing System Command Comparison

| PBS         | LSF           | LoadLeveler    | Description                                          |
|-------------|---------------|----------------|------------------------------------------------------|
| qsub script | bsub < script | lsubmit script | Submit a job script for execution.                   |
| qstat       | bjobs         | llq            | Show status of running and pending jobs.             |
| tracejob    | bhist         |                | Displays historical information about your jobs.     |
| qdel        | bkill         | llcancel       | Kill a job.                                          |
| qhold       | bstop         | llhold         | Hold a job.                                          |
| qstat -Q    | bqueues       | llclass        | Show configuration of queues.                        |
| qstat -Qf   |               | showqlimits    |                                                      |
|             | busers        |                | Displays information about users and groups.         |
|             | bpeek         |                | Displays the stderr and stdout of an unfinished job. |
|             | bacct         |                | Displays accounting information for finished jobs.   |
|             | bhosts        | llstatus       | Summarize load on each host.                         |

NOTE: With PBS “qsub” is used to submit a job while “qstat” can be used to check the status. The job remains in a pending state (PEND) until all resources are available.

Once the resources are available, the job is started by PBS and is now in a running state (RUN). “qstat -Qf” can be used to show the configuration of the queues.

```

#PBS -q standard          # Queue name.
#PBS -l select=1:ncpus=1  # Number of CPUs.
#
# Compile Fortran code
xlf90 -o serial.exe serial.f
# Run serial executable on 1 cpu of one node
./serial.exe
#End of Sample PBS Script

```

And a PBS scrip to run a parallel Open MP job will look like this:

```

#!/bin/csh
#PBS -N mpijob            # Name of the job.
#PBS -o mpijob.out        # Appends std output to file
                           #mpijob.out.
#PBS -e mpijob.err        # Appends std error to
                           #file mpijob.err.
#PBS -A NAVOSLMA          # Charging Project ID.
#PBS -l walltime=02:00:00 # Wall clock time of 2 hours.

```

While a PBS script to run a Parallel Message Passing Interface (MPI) job will resemble this:

```

#!/bin/csh
#PBS -N mpijob            # Name of the job.
#PBS -o mpijob.out        # Appends std output to file
                           #mpijob.out.
#PBS -e mpijob.err        # Appends std error to
                           #file mpijob.err.
#PBS -A NAVOSLMA          # Charging Project ID.
#PBS -l walltime=02:00:00 # Wall clock time of 2 hours.

```

Continued Next Page...

| Job Script Frequently Used Options      |                                                                     |                                                          |                                                                                         |
|-----------------------------------------|---------------------------------------------------------------------|----------------------------------------------------------|-----------------------------------------------------------------------------------------|
| PBS                                     | LSF                                                                 | Loadleveler                                              | Option                                                                                  |
| #PBS -N jobname                         | #BSUB -J jobname                                                    | ## job_name = jobname                                    | assigns name to job                                                                     |
| #PBS -M email_address<br>#PBS -m b      | #BSUB -B                                                            | ## notify_user = login_name<br>## notification = start   | sends email when job begins execution                                                   |
| #PBS -m e                               | #BSUB -N                                                            | ## notification = complete                               | emails finished job report                                                              |
| #PBS -e errfile                         | #BSUB -e errfile                                                    | ## error = errfile                                       | redirects stderr to specified file                                                      |
| #PBS -o out_file                        | #BSUB -o out_file                                                   | ## output = out_file                                     | redirects stdout                                                                        |
|                                         | #BSUB -a application                                                |                                                          | esub parameter                                                                          |
| #PBS -A project_name                    | #BSUB -P project_name                                               | ## account_no = project_name                             | assigns job to specified project                                                        |
| #PBS -l walltime=runtime                | #BSUB -W runtime                                                    | ## wall_clock_limit = runtime                            | sets the run limit of the job                                                           |
| #PBS -q queue_name                      | #BSUB -q queue_name                                                 | ## class = queue_name                                    | submit the job to the specified queue                                                   |
| #PBS -l select=[chunk<br>specification] | #BSUB -n num_procs<br>#BSUB -R "span[ptile=num_<br>procs_per_node]" | ## node = num_nodes<br>## tasks_per_node = num_<br>procs | specifies number of processors to use<br>Specifies resource requirements an<br>MPI job. |

| DAVINCI Queues Overview |                     |                     |                           |
|-------------------------|---------------------|---------------------|---------------------------|
| Queue Name              | Max Nodes/CPUs      | Max Wall-Clock Time | Comments                  |
| transfer                | 1 node/ 1 CPUs      | 6 hrs.              | Transfer jobs             |
| standard                | 32 nodes/256 CPUs   | 3 days              | Non-challenge jobs        |
| smp                     | 1 node/32 CPUs      | 12 hrs.             | Shared memory jobs        |
| share                   | 1 node/1 CPU        | 12 hrs.             | Serial jobs.              |
| challenge               | 128 nodes/1024 CPUs | 7 days              | Challenge & Priority Jobs |
| high                    | 1 node/ 32 CPUs     |                     |                           |
| debug                   | 32 nodes/256 CPUs   | 30 mins.            | Debug Jobs                |
| background              | 32 nodes/256 CPUs   | 4 hrs.              | Negative Allocation jobs  |

```

#PBS -A NAVOSLMA          # Charging Project ID.
#PBS -l walltime=04:00:00  # Wall clock time of 4 hours.
#PBS -q bigmem             # Queue name.
#PBS -l select=1:ncpus=8   # Request one
                           #8-processor "chunk"
#PBS -l place=excl         # Allocate node exclusively
#
# Run the OpenMP job with the IBM "poe" parallel job
starter.

```

## CONCLUSION

Using PBS on the IBM P6 and CRAY XT systems should increase the efficiency of the systems and, therefore, the productivity of the NAVO MSRC users.

The NAVO MSRC User Support team is available to assist you with porting your LSF and/or LoadLeveler script to PBS Pro. Users are invited to direct requests for assistance to the Consolidated Customer Assistance Center (CCAC) at 1-877-222-2039 or by email at [help@ccac.hpc.mil](mailto:help@ccac.hpc.mil).

| Environment Variable Comparison |              |                    |                                                                      |
|---------------------------------|--------------|--------------------|----------------------------------------------------------------------|
| PBS                             | LSF          | LoadLeveler        | Variable Description                                                 |
| PBS_JOBID                       | LSB_JOBID    | LOADL_JOB_NAME     | Unique job number.                                                   |
| PBS_ARRAY_INDEX                 | LSB_JOBINDEX | LOADL_STEP_ID      | Job index for array jobs.                                            |
| PBS_JOBNAME                     | LSB_JOBNAME  | LOADL_STEP_COMMAND | Name of the job.                                                     |
| PBS_TASKNUM                     | LS_JOBPID    | LOADL_PID          | Process ID of the job.                                               |
| PBS_ARRAY_ID                    |              |                    | Identifier for job arrays. Consists of sequence number.              |
| PBS_JOBDIR                      |              |                    | Pathname of job-specific staging and execution directory.            |
| PBS_JOBCOOKIE                   |              |                    | Unique identifier for inter-MOM job-based communication.             |
| PBS_QUEUE                       |              |                    | The name of the queue from which the job is executed.                |
| PBS_NODEFILE                    |              |                    | The filename containing a list of vnodes assigned to the job.        |
| PBS_NODENUM                     |              |                    | Logical vnode number of this vnode allocated to the job.             |
| PBS_ENVIRONMENT                 |              |                    | Indicated job type: PBS_BATCH or PBS_INTERACTIVE                     |
| PBS_MOMPORT                     |              |                    | Port number on which this job's MOMs will communicate.               |
| PBS_O_WORKDIR                   |              |                    | The absolute path of the directory from which the job was submitted. |
| PBS_O_HOME                      |              |                    | Value of HOME from submission environment.                           |
| PBS_O_LOGNAME                   |              |                    | Value of LOGNAME from submission environment.                        |
| PBS_O_LANG                      |              |                    | Value of LANG from submission environment.                           |
| PBS_O_PATH                      |              |                    | Value of PATH from submission environment.                           |
| PBS_O_MAIL                      |              |                    | Value of MAIL from submission environment.                           |
| PBS_O_QUEUE                     |              |                    | The original queue name to which the job was submitted.              |
| PBS_O_SHELL                     |              |                    | Value of SHELL from submission environment.                          |
| PBS_O_HOST                      |              |                    | The host name where qsub was executed.                               |
| PBS_O_SYSTEM                    |              |                    | The operating system where qsub was executed.                        |
| PBS_O_TZ                        |              |                    | Value of TZ from submission environment.                             |
| NCPUS                           |              |                    | Number of threads, defaulting to number of CPUS, on the vnode.       |
| OMP_NUM_THREADS                 |              |                    | Same as NCPUS.                                                       |
| TMPDIR                          |              |                    | The job-specific temporary directory for this job.                   |

## ***SNPD 09***

*Daegu, Korea  
May 27 – 29, 2009*

*10th IEEE/ACIS International  
Conference on Software Engineering,  
Artificial Intelligence, Networking,  
and Parallel/Distributed Computing*

## ***IST 09***

*Shenzhen, China  
May 9-10, 2009*

*IEEE International Workshop on Imaging  
Systems and Techniques*

## ***UGC 09***

*San Diego  
June 15 - 18*

*[http://www.hpcmo.hpc.mil/  
Htdocs/UGC/index.html](http://www.hpcmo.hpc.mil/Htdocs/UGC/index.html)*

## ***MSC 09***

*Saint Petersburg, RUSSIA  
July 8-10, 2009*

*IEEE Multi-conference on  
Systems and Control*

## ***VIS 09***

*Atlantic City  
October 11-16, 2009*

*Forum for visualization advances in  
science and engineering*

## ***ICCD 09***

*TBD*

*International Conference on  
Computer Design*

# **Coming Events**



**NAVO Major Shared Resource Center**  
**1002 Balch Boulevard, Stennis Space Center**  
**Mississippi 39522**