Summer 2000

The Resource

U.S. ARMY ENGINEER RESEARCH AND DEVELOPMENT CENTER NEWSLETTER

Program Review 2000

ERDC MSRC

Major Shared Resource Center

From the Director's chair

The recent High Performance Computing Modernization Program's (HPCMP's) Program Review 2000 (PR2000) was held during the last week of March. It was a milestone event for the HPCMP. We started this program in the early nineties but really did not formalize it with a definitive set of participants until 1996. At that point, we began to establish our business models, modes of intercommunication, working relationships, and friendships. PR2000 was the first time we stopped and looked back at what we had accomplished. The original goals of the program were in general to deploy high-end computers, foster the maturity of software systems, expand and train the Department of Defense (DoD) high performance computing (HPC) user base, contribute to the national HPC infrastructure, and hook it all together with a very high-speed network. The Program Review not only demonstrated that we have done a great job at achieving those goals but also that we have assembled a tremendous team of experts in HPC.

I look at PR2000 as a stake in the ground. It is a time when we look back and say, "Did we make it?" Sure there are areas for improvement, but overall I think the answer is unequivocally yes. The comments and ratings from the review board confirm this. It is also a time when we need to look forward and ask ourselves, "Where do we go from here?" A few goals might include applications that harness a larger pool of computational resources through distributed computing technologies, the development of applications that implement a coupled approach to multidisciplinary problems, and the increased use of interactive computing. Through the continued coordinated efforts of the Common High Performance Computing Software Support Initiative, Programming Environment and Training (PET), Shared Resource Centers, and Defense Research and Engineering Network, this program is poised to take on the next wave of computational challenges in the DoD.

Bradley M. Comes
Director, ERDC MSRC

About the Cover:
Users Group Conference 2000. Welcome to Albuquerque! The release of the Summer 2000 edition of *The Resource* was planned to coincide with the Users Group Conference 2000 so that we can personally present the newsletter to many of our readers.

Program Review 2000. The HPCMP Program Review 2000, held March 27-31, 2000, was a rare opportunity for the entire program to highlight the critical role the DoD HPCMP has contributed to the advancement of military and battlefield technologies. Pictured on the lower left is Dr. Delores Etter, Deputy Under Secretary of Defense for Science and Technology (PR2000 pictures courtesy of Mike Moore, Aeronautical Systems Center (ASC)) (see story, page 8).

Features

Departments

announcements

Dr. Art Cullati Steps Up as Systems and Services Director

Dr. Art Cullati has recently joined the U.S. Army Engineer Research and Development Center (ERDC) Major Shared Resource Center (MSRC) team as the Systems and Services Director. In this key position, he will oversee User Services, the HPC systems, and the network support teams.

Dr. Cullati has more than 30 years of experience in the computing field and 20 years of experience managing HPC centers. He joins us from his position as the Director of the National Environmental Supercomputing Center of the Environmental Protection Agency.

Dr. Cullati's experience includes developing software for the Minuteman Missile and managing the supercomputing installations at the Numerical Aerodynamic Simulation, National Aeronautics and Space Administration/Ames Research Center. He began his career in systems and applications work including designing, developing, implementing, and supporting his own original special-purpose operating system and writing a language translator that compiled structured English and generated object code in another vendor's computer. He was also the project leader of the benchmark effort for Cray Research, Inc., in its first commercial sale to Boeing Computer Services.

We look forward to Dr. Cullati's knowledge and experience contributing significantly to initiatives such as a common user interface, metacomputing, shared file systems, and remote disaster recovery.

tool tips

New Commands

Jay Cliburn

Recently, the ERDC MSRC added several useful commands for manipulating files on the Mass Storage Facility (MSF). These commands are available on all HPC hosts, as are manpages for each command. A short synopsis of each command follows.

`msfchmod`	change file permissions on the MSF
`msfcopy`	copy a file from the old MSF to the new MSF
`msfget`	copy a file from the MSF
`msfls`	list the contents of a directory on the MSF
`msfmcopy`	copy multiple files from the old MSF to the new MSF
`msfmget`	copy multiple files from the MSF
`msfmkdir`	create directories on the MSF
`msfmput`	copy multiple files to the MSF
`msfmv`	rename a file on the MSF
`msfput`	copy a file to the MSF
`msfrm`	delete a file on the MSF
`msfrmdir`	delete directories on the MSF
`msfstat`	check the status and availability of the MSF

Security Update

Mitch Baker

The ERDC MSRC has been working with several developers to provide better Windows tools to access the MSRC systems. Our first success in this endeavor is the integration of KRB5 with SecurID support into Kermit-95. Many may remember Kermit from the days of serial/dial-up access into the system. It is still around and greatly enhanced.

Listed below are the major features of Kermit-95 and contact information to obtain Kermit-95. Users will still need the Win32 KRB5 kit from Naval Research Laboratories to use the new KRB5 functions in Kermit. (It is available free of charge; please see *www.hpcmo.hpc.mil/Htdocs/SECURITY/kerberos_q.html* on how to get the software.) We would like to thank Jeffrey Altman, Columbia University, for all of his work to get the KRB5 with SecurID functioning.

Kermit-95 is the Windows 95/98/NT/2000 and OS/2 edition of Columbia University's well-known Kermit communications software. It handles both modem and TCP/IP connections with equal ease and offers 38 different terminal emulations, Kermit and XYZMODEM file transfer protocols, international character-set conversion (now including Unicode), numeric and alphanumeric paging, host mode for incoming connections (serial or network), and a scripting language to allow automation of any communications task that can be done by hand. The new Kermit-95+ package also includes C-Kermit 7.0, its file-transfer and client/server partner for UNIX, VMS, VOS, and other operating systems on CD-ROM, including source code and prebuilt binaries for hundreds of platforms. Also included are online manuals for Kermit-95 and C-Kermit, full source code for C-Kermit, and dozens of sample scripts.

Compliance: Year 2000, Euro, Win32, Winsock, TAPI (Telephony Application Programming Interface).

Communication methods: Direct serial, modem, TCP/IP Telnet and Rlogin, SuperLAT, DECnet. (SuperLAT and DECnet require the corresponding external network protocol stacks.)

Security: Kerberos IV, Kerberos V, SSL/TLS, SRP (subject to USA export restrictions).

Terminal emulations: VT320/220/102/100/52; VT320PC/VT220PC; Linux Console; ANSI, SCOANSI, AT386, Avatar; Wyse 30/50/60/160/370; Televideo 910/925/950; DG200/210/217; IBM HFT/AIXTERM/3151; HP 2621/HPTERM; QNX; Microsoft NTVT; Siemens Nixdorf BA80 and 97801; Heath-19; Hazeltine 1500; Volker Craig 404.

Special terminal features: Colors, selectable screen dimensions, international character-set translation, complete built-in default keymaps for each terminal type, flexible key and mouse-event redefinition, keystroke macros, URL hot spots, up to 2,000,000 lines of scrollback with bookmark and search capability, and special keyboard modes: PCTERM, EMACS, Word Perfect, Russian, Hebrew.

File transfer protocols: Kermit, XYZMODEM, ASCII. Kermit protocol includes client/server mode and international character-set translation.

Scripting: All operations can be programmed for automatic unattended execution in a consistent way, regardless of platform or connection method, using the portable Kermit scripting language, which includes file and communications input/output, block structure, looping, variables, arrays, functions, and structured programming features.

Languages: Kermit 95's user interface is in English, but in its communication functions it supports translation of all the character sets used for Western European languages (English, Italian, Portuguese, Norwegian, Icelandic, German, Spanish, French, etc.); Eastern European Languages (Czech, Polish, Hungarian, Romanian, etc.); Greek; Hebrew and Yiddish; languages written with the Cyrillic alphabet (Russian, Ukrainian, Belorussian, Bulgarian, etc.); languages written with the Hebrew alphabet (Hebrew, Yiddish, Ladino, etc.); and (in file transfer only) Japanese. New to version 1.1.20: Unicode, the Universal Character Set, in both UCS-2 and UTF-8 form, with full translation to/from all the other character sets.

Cost: Single copies are $64.00 U.S. dollars (USD), although quantity discounts are available. Low-cost bulk-right-to-copy licenses for 100 seats or more are available for prices beginning at $10 USD per seat. Academic site licenses are $2000 USD.

More information

www.columbia.edu/kermit/k95.html

off-campus

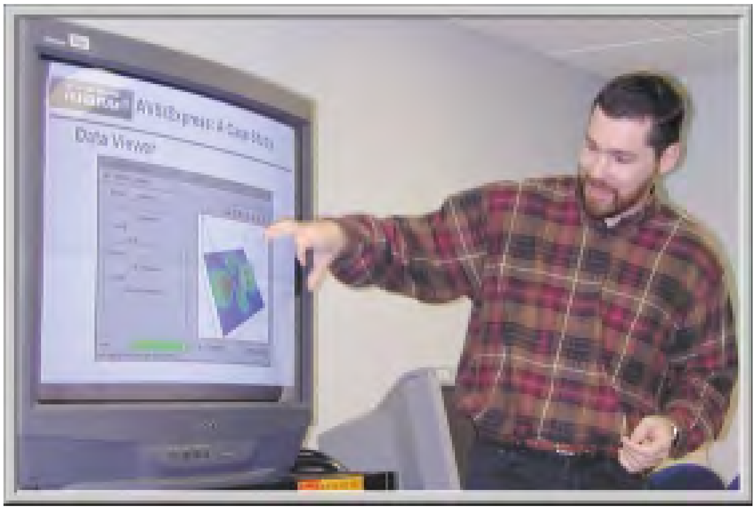
ERDC MSRC Presents Two Seminars

Scientific Visualization—Philosophy, Software, Video Editing

Dr. Kent Eschenberg

On February 8, 2000, three members of the ERDC MSRC Scientific Visualization Laboratory presented a half-day seminar to the faculty and students at Jackson State University (JSU), an academic partner to our PET Program. The seminar was part of a 2-day workshop on scientific visualization.

Dr. Kent Eschenberg, Acting Lead for the Scientific Visualization Laboratory, discussed the philosophy of the way visualization can support science and provided an introduction to visualization programming. Mr. Glen Browning reviewed software packages that can essentially eliminate programming and offered details on two packages regularly used at the ERDC MSRC Scientific Visualization Laboratory. Mr. David Longmire discussed basic principles and procedures for video editing. A list of resources in books, journals, and Web sites was also provided.



Glen Browning, ERDC MSRC visualization specialist, presents one of two case studies (using AVS/Express and Maya) as part of the ERDC MSRC visualization team presentation. AVS/Express and Maya are two tools commonly used for scientific visualization. The team presented approximately 4 hr of lecture and video.

Chuck Patrick, the scientific visualization specialist from JSU, works with JSU students during a demonstration of AVS (Application Visual Systems, Inc.) on the second day of the workshop. AVS is one of several visualization packages currently available at the ERDC MSRC.

at Jackson State University

Scientific Visualization Fundamentals

John E. West

Dr. Julie Baca, JSU Computer Science Department, assists one of her students at the seminar.

In a later visit in early spring, Mr. John E. West, ERDC MSRC, gave a seminar on scientific visualization at JSU. The seminar was designed to introduce the fundamental concepts of data visualization to students with no prior background in visualization or computer graphics. Dr. Julie Baca, JSU Computer Science Department, hosted the seminar, which was attended by graduate and undergraduate students in the Department of Computer Science. Dr. Baca had two goals for the seminar. First, she wanted it to serve as a supplement to enrich the experience her students receive in regular classes by introducing a new technology area in which computer science plays a critical role. Second, some of the students attending the seminar are beginning scientific visualization research on externally funded projects, and Dr. Baca saw the seminar as a way to jumpstart those students' preparation for the efforts to come.

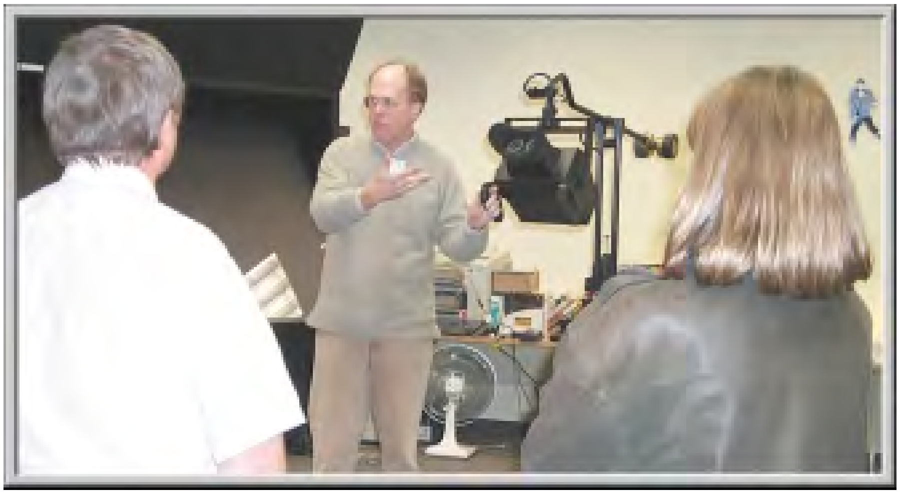
In all, the seminar was a big success, with follow-on sessions being planned to cover more advanced topics. Interactions like these are a welcome reminder that training the next generation of scientists and engineers for the HPC community is one of the most important ways in which we can have a lasting impact on the long-term future of our communities and our Nation.

John E. West, ERDC MSRC, conducts a scientific visualization seminar at JSU.

Corps of Engineers Personnel Attend ITL Computer Applications Class

During the week of January 24-28, 2000, 17 students from various Corps of Engineers' District offices attended the Computer Applications for Engineers and Engineering Managers (CAEEM) course held at the ERDC Information Technology Laboratory (ITL). The course, presented annually by the ITL Computer Aided Engineering Division, includes the use of computer-aided engineering design techniques and applications available within the Corps of Engineers with a "hands-on" emphasis.

David Stinson and Mike McCraney discuss the DoD HPCMP with the CAEEM class.



Dr. Kent Eschenberg presents scientific visualization techniques to the CAEEM class.

ERDC MSRC team members Mr. David Stinson, Dr. Henry Gabb, Mr. Mike McCraney, Mr. Mike Nassour, and Dr. Kent Eschenberg presented an overview of supercomputing at the ERDC MSRC. Topics of discussion included HPC applications, mass storage, networks, and scientific visualization.

U.S. Marine Corps Representatives

U.S. Marine Corps representatives from installations in North Carolina, Arizona, California, Virginia, and Hawaii in the United States and from two installations in Japan toured the ERDC MSRC HPC facility in February 2000. These representatives were at the ERDC ITL to attend the Marine Corps Geospatial Information Systems Users Group meeting at the CADD/GIS Technology Center for Facilities, Infrastructure, and Environment. This group convenes as required to exchange information, promote and sustain the use of GIS for integrated installation management, and assists in the development of policy for dissemination by Installations and Logistics, Headquarters, U.S. Marine Corps.

U.S. Marine Corps representatives visit the ERDC MSRC HPC facility.

Council of Colonels

The U.S. Army Training and Doctrine Command (TRADOC) Modeling and Simulation Advisory Council attended the Council of Colonels meeting held at the ERDC Headquarters, Vicksburg, MS, on March 15-16. This group meets quarterly to report on taskings and perform other business of the Advisory Council. Members visit various sites to learn more about Army organizations and their areas of expertise within the DoD.

Dr. Louis Turcotte, Assistant for Technology in the ERDC ITL, conducted a tour of the ERDC MSRC HPC facility and the Scientific Visualization Center for the 54 meeting attendees representing many agencies and commands.

Dr. Turcotte conducts tour of the computer facilities for the TRADOC Modeling and Simulation Advisory Council.

Mississippi State University President and Staff Members

Tim Ables discusses high performance computing with Dr. Portera.

The president of Mississippi State University (MSU), Dr. Malcolm Portera, along with some of the university's newest faculty members and a few senior administrators, visited the ERDC MSRC on March 30. Mr. Tim Ables, Acting Director, ITL, served as the tour guide, assisted by Mr. Dennis Gilman, ERDC MSRC.

The university is conducting a number of bus tours around Mississippi for new faculty to show them firsthand the impact that teaching, research, and service programs have in the State of Mississippi. This effort also helps acquaint faculty with the strengths and aspirations of Mississippi communities.

MSU is one of the ERDC MSRC's PET partners and the host university for the National Science Foundation (NSF) Engineering Research Center for Computational Field Simulation.

Program Review 2000

John E. West

The HPCMP had its first program-wide review the week of March 27-31, 2000. The review covered all aspects of the HPCMP in a series of tracks, including the network and software initiatives as well as the shared resource centers. The review was held in Fairfax, Virginia, and included an introduction by Dr. Delores Etter, Deputy Under Secretary of Defense for Science and Technology. Meeting attendees included service agency officials, members of the HPC Advisory Panel, and external reviewers as well as members from all facets of the HPCMP.

The four MSRCs had a significant role in the week's activities. The centerpiece of the MSRCs' participation in the review was the MSRC Post-Implementation Review, during which each Center Director highlighted the goals, characteristics, and processes that have contributed to the enormous success of this program. Each Center Director spent some time highlighting the way in which the shared resource centers cooperate across mission and service agency boundaries to deliver the latest in mission-critical computing technologies and service to the entire DoD research, development, test and evaluation community.

In addition to presenting MSRC and PET briefings during the Post-Implementation Review, the ERDC MSRC participated in a wide range of tracks during the review. Its staff members were part of sessions of the Mass Storage and Archival Systems Working Group and Networking Technical Advisory Panel sessions, the Service/Agency Approval Authority briefing and dialog sessions, the Requirements and Resource Allocation Review, and the Challenge Lessons Learned meeting.

Program Review 2000 was a rare opportunity for the entire program to highlight the critical role the DoD HPCMP has contributed to the advancement of military and battlefield technologies. It was also an opportunity for the program to look back as well as into the future.

Brad Comes, ERDC MSRC Director, making a presentation at Program Review 2000.

Program Review 2000

(Pictures courtesy of Mike Moore, ASC MSRC)

Massively Parallel Disk

Dr. Joseph Werne, Mr. Paul Adams, and Mr. David Sanders

Introduction

Current-generation, massively parallel computing platforms such as the 512-processor Cray T3E at the ERDC MSRC allow opportunities for scientific inquiry that were impossible just a few years ago. Complex nonlinear problems that offer only weakly nonlinear solutions through traditional analytic methods are being directly attacked daily through high performance computational technology at the Nation's Supercomputing Centers. The opportunities for scientists are growing, but so too are logistical difficulties. One of these difficulties is the subject of this article: massively parallel disk input/output (I/O), i.e., performance tuning associated with parallel data transfers to the massively parallel processing (MPP) disks, which are then migrated off-line to a long-term data storage facility. In particular, we describe a new technique called "Hierarchical Data Structuring" (HDS).[1] HDS can speed up massively parallel I/O-limited calculations. We present an example in which the speedup can range from 1.4 to 2.1 or more, depending on the number of data files involved.

As a vehicle for demonstrating the proper approach to massively parallel data handling on the ERDC MSRC Cray T3E, we describe high-resolution turbulence simulations in support of the Air Force AirBorne Laser (ABL) Program. The work is part of an effort of the Space Vehicles Directorate of the Air Force Research Laboratory. The scientists conducting the simulations are Drs. Joe Werne and Dave Fritts of Colorado Research Associates (CoRA).[2] Their computations at the ERDC MSRC involve the highest resolution pseudospectral calculations of atmospheric turbulence processes (such as wind shear and gravity-wave breaking) currently being conducted. (Please see the Spring 2000 issue of the ERDC MSRC's *The Resource* newsletter or the Spring 1999 *CEWES MSRC Journal* at *www.wes.hpc.mil* (under Publications) for further information on the ABL project.) We expect that the performance enhancement exhibited by HDS will be universally applicable to other architectures.

Project Description

The ABL turbulence solutions require massive computational resources and long-term disk storage. Tens of terabytes (TB) of numerical data will be generated by the project by the end of this year. Runs currently being conducted employ 500 processing elements (PEs) and generate 410 GB of data during one 20-hr job. Because of the large sizes of the data sets and limited high-speed storage space on the Cray T3E, ABL data must be shuttled off-line in real-time to the ERDC MSF. The MSF consists of an SGI Origin 2000 front end to a Cray J16/4512, which migrates data to three StorageTek robotic silos. Data are transferred between the T3E and the MSF across a High-Performance Parallel Interface (HiPPI) network, capable of 290-megabits-per-second (Mbps) sustained data transfers with bursts as high as 800 Mbps. Actual transfer rates vary according to the channel load, as the HiPPI channel is shared; however, the bandwidth is sufficient for this problem such that throughput is typically limited by read/write speeds on the Cray T3E and MSF. To facilitate the data movement, the ERDC MSRC has developed a suite of utilities to facilitate the storage and retrieval of large files on the MSF. The utilities are based on the SGI File Transfer Agent software package and provide fetch (`msfget/msfmget`), store (`msfput/msfmput`), status (`msfstat`), and file manipulation (`msfchmod/msfls/msfrm/etc.`) capabilities.

For optimal performance of the application program, I/O between Cray T3E memory and the

---

[1] Werne, J., Adams, P., and Sanders, D. "Massively parallel disk (I/O) on the Cray T3E." *In submission.*

[2] CoRA is a division of NorthWest Research Associates (NWRA).

T3E data disks must be considered carefully. For the ABL simulations, a poor I/O strategy can ruin an otherwise scalable algorithm. Two approaches to massively parallel disk I/O were considered:

1. Have each processor perform I/O to its own set of files.
2. Stripe one large datafile across multiple data disks, and instruct participating PEs to write to their portion of the global file.

Physically, the two solutions are the same, with separate processors writing data to separate disks. The main difference involves specifics of the implementations and the manner in which file names are handled. Though the user conceives a single file when employing (2), in reality, portions of the global file reside on all of the disks spanning the set of processors used. In order to stripe files across multiple disks, the `-p` and `-q` options of the Cray assign command must be used. The `-p` option indicates which partitions to use in the filesystem. The `-q` option indicates how much data will be placed on a single disk before moving on to the next disk. Users must be careful to preserve the intended file layout and use the Cray fdcp command when copying files from another machine. It is clear that parallel I/O via file striping immediately involves machine-specific commands, and portability can become limited.

Solution (1) presents a simpler alternative. Assurance that a processor's output files reside on its disk is accomplished trivially by instructing each processor to create its own files. Though restart data from a previous run will not be locally resident if file prefetching is done by another processor, presumably restart data are read only once, and poor I/O performance for the initial read is tolerable. To ensure that subsequent writes to the restart files are not hindered by the files' initial prefetched locations, the user should ensure that a new set of restart files is created for each run and that each processor creates its own restart file.

Dr. Werne implemented solution (1) with direct-access, fixed-record-length binary files for the ABL turbulence simulations. Output files are opened just before writing and then closed again after writing is complete. This approach minimizes the number of simultaneously open files and avoids problems with unflushed file buffers when a run is unexpectedly stopped.

The Cray T3E code scales linearly with the number of processors when I/O is turned off; however, the code's linear scaling is adversely impacted when run in production mode as a result of the massive amounts of I/O demanded by the project, even when solution (1) is properly implemented. We used two improvements to restore the code's scaling with the number of processors during production runs: (a) write-behind data buffering[3] and (b) HDS. The first improvement involves buffering output-data writes so that computation and I/O can be overlapped for as long as the array being written is not needed for subsequent computation. On the T3E this is accomplished with an option to the Cray assign command: -F bufa:num1:num2, where num2 is the number of buffers used, and num1 is the size in 4,096-byte blocks of each buffer layer. Dr. Werne uses `-F bufa:2000:4`. See on-line man "intro_ffio" for more information on this and other assign options. The second performance improvement, HDS, is a technique we recently discovered. It involves dividing and subdividing the output data directory into multiple subdirectories and sub-subdirectories, ending with tails that contain no more than NCPU files (NCPU is the number of processors used for a given calculation). For example, for three-dimensional volume files written at time levels 00, 01, etc., instead of a directory named "data" with the following layout:

```
data/:      file1.00.0000, file1.00.0001, ... file1.00.0499,
            file2.00.0000, file2.00.0001, ... file2.00.0499, ...
            file1.01.0000, file1.01.0001, ... file1.01.0499,
            file2.01.0000, file2.01.0001, ... file2.01.0499, ...
```

[3] Dr. Werne learned about this technique from David Cole and Mike Patterson at the Naval Oceanographic Office (NAVO) MSRC. Werne's specific implementation grew out of valuable correspondence with another DoD Challenge user, Alan Wallcraft, who conducts global and basin-scale ocean modeling on the NAVO T3E and the new ERDC MSRC IBM Power3 SMP.

HDS uses the following layout:

```
data/00/file1/:   file1.00.0000, file1.00.0001, ... file1.00.0499
data/00/file2/:   file2.00.0000, file2.00.0001, ... file2.00.0499
data/01/file1/:   file1.01.0000, file1.01.0001, ... file1.01.0499
data/01/file2/:   file2.01.0000, file2.01.0001, ... file2.01.0499, ...
```

The finer resolution via a directory tree ensures that the application program does not spend time wading through an overwhelming number of file names. For example, when using NCPU = 500, ABL simulations generate in excess of 72,000 individual files during the course of a 20-hr run. When real-time data migration to the MSF is accomplished according to plan, all 72,000 files do not reside on the T3E /tmp disks at the same time. They are collected into tar files and moved off-line immediately after they are closed. Nevertheless, a minimum of 9,000 files simultaneously reside on the T3E /tmp space during the run, and with the old data layout, this is sufficient to impact performance. For instance, when running with the old directory structure, ABL turbulence simulations required 82 min to complete 50 time-steps and do the associated disk I/O. When data from previous runs also occupied the directory, as much as 2 hr or more could be required. In fact, when performing tests to diagnose the source of the variable execution time, we intentionally allowed the data directory to fill with files that were not transferred when previous tests failed. Eventually, the directory became so full with 100,000 files that 50 time-steps were not completed in as many as 6 hr. Often an error "_sma_deadlock_wait" was encountered on the T3E. In contrast, now that we have implemented the new HDS strategy, execution time no longer varies. The 50-time-step tests finish in 59 min + 1 min and 20-hr dedicated runs demonstrate the same predictable performance regardless of the number of files in the directory tree.

Conclusion

We expect that the performance enhancement exhibited by HDS will be universally applicable to other architectures. Given its simplicity, we believe it could be implemented for all MPP applications that generate large amounts of data and large numbers of individual files. An additional benefit of HDS is that the user can easily navigate the directory tree and inspect data as they are generated by the application program. This is not feasible with a non-HDS directory containing 10,000 or more files.

Dr. Joseph Werne is a research scientist at the CoRA division of NWRA. He has 11 years of experience in high performance computing on shared-memory vector and distributed-memory parallel platforms. He has written highly optimized Fortran programs, including the ABL turbulence code, for both stably and unstably stratified turbulence simulations as part of NSF Grand Challenge and DoD Challenge applications since 1993. Dr. Werne receives support for ABL simulations from AFRL F19628-98-C-0030 and AFOSR F49620-98-C-0029. Related supercomputer work is sponsored by ARO DAAD191-99-C-0037.

Mr. Paul Adams recently moved from Applications Support Analyst to User Services Manager at the ERDC MSRC. He has 12 years of experience in the computer industry, providing programming and user support on Cray, SGI, and IBM platforms.

Mr. David Sanders is an Applications Support Analyst at the ERDC MSRC and a Systems Support Engineer with Cray, Inc. He has 16 years of programming and user-support experience on shared-memory, distributed-memory parallel, and distributed-shared-memory platforms. He wrote the suite of MSF data-handling utilities described above.

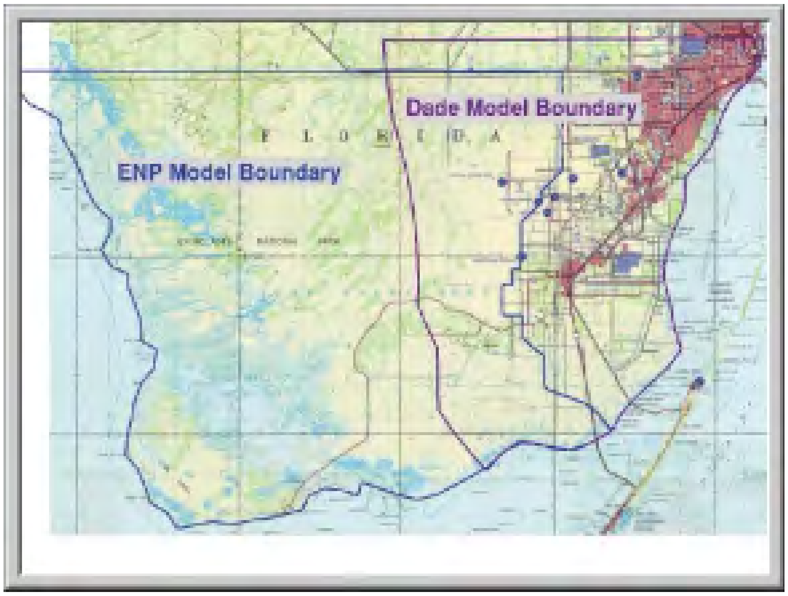
Interview with users...

Dr. Fred Tracy and Mr. Dave Richards

Dr. Fred Tracy, ERDC ITL, and Mr. Dave Richards, ERDC Coastal and Hydraulics Laboratory, spoke with us about using the ERDC MSRC Cray T3E for the remarkably difficult computational challenge of modeling the Everglades National Park (ENP) and the Dade County South Florida area using the finite element program FEMWATER-123. The ENP model involves both two-dimensional (2-D) overland and 3-D groundwater flow. The more sophisticated Dade model also includes 1-D canal flow. The goal is to do a 1-year simulation using 1- to 2-hr 3-D time-steps on a parallel high performance computing platform in a 24-hr time period. Since the canal, overland, and groundwater flow equations are all separately nonlinear, the required coupled solution seemed impossible on any computer. However, after much effort, the needed solution ran in 16.5 hr using 40 PEs on the Cray T3E. This application is being highlighted because of its computational complexity and because FEMWATER-123 is applicable to military applications where it is currently being successfully used for environmental remediation of military sites.

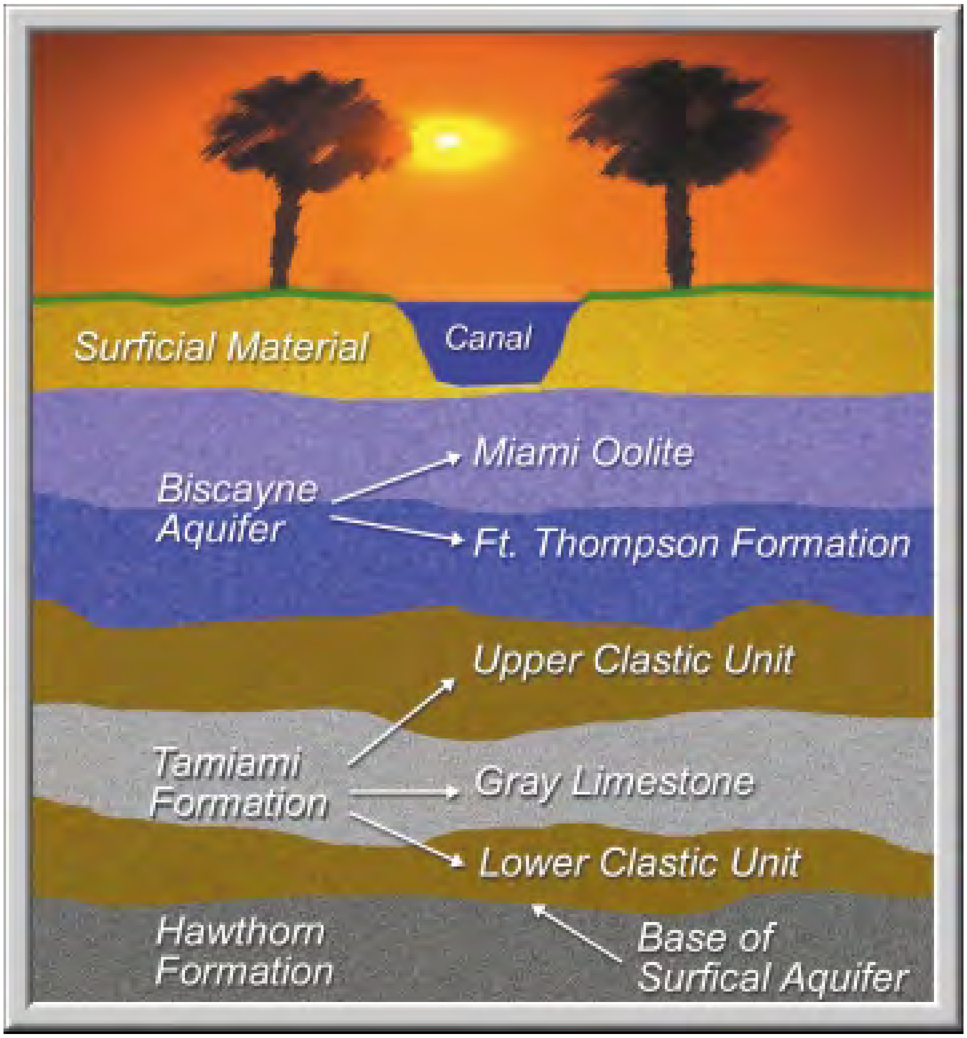
What are you modeling? The Jacksonville District of the U.S. Army Corps of Engineers and the South Florida Water Management District have responsibilities for managing water resources and, as a result, the natural environment of southern Florida. These include operating the extensive flood protection works and canals that significantly affect the surface and groundwater hydrologic processes of the region. While these flood protection works have served their purpose over the years, there has been an effort recently to revisit their design with the hope of returning the environment back to a more natural state. This is a difficult task since this may entail raising water levels, thereby affecting local flood risk. In addition, there are competing interests for the use of the water. Everglades National Park desires redistribution of water for ecological reasons; agricultural interests desire certain water levels to develop their businesses; and municipalities have increasing populations that want a clean water supply, recreational facilities, and good flood protection at the same time. Therefore, the flow of water in the ground, on the surface, and in the canals is being modeled using FEMWATER-123 to aid management in making these difficult decisions.

What are the computational challenges? Groundwater flow is characterized by many different heterogeneous materials as illustrated by the 3-D finite element mesh of the Dade model. Also, unsaturated 3-D groundwater flow by itself is highly nonlinear and often converges very slowly if at all. Add to that the challenges of creating an efficient scalable parallelized FEMWATER code, and you are already in a tough spot. It was first thought that FEMWATER would not converge at all for large problems, and if it were possible to do a parallel version, it would not scale very well. However, a message-passage interface version of the 3-D code was developed, along with improvements to the nonlinear solver. A parallel speedup of 1.8 when going from 64 to 128 PEs and a scaled speedup of 90 percent when going from 8 to 64 PEs on the Cray T3E were achieved, which is excellent.



South Florida project boundaries for the ENP and Dade models.

interview with users

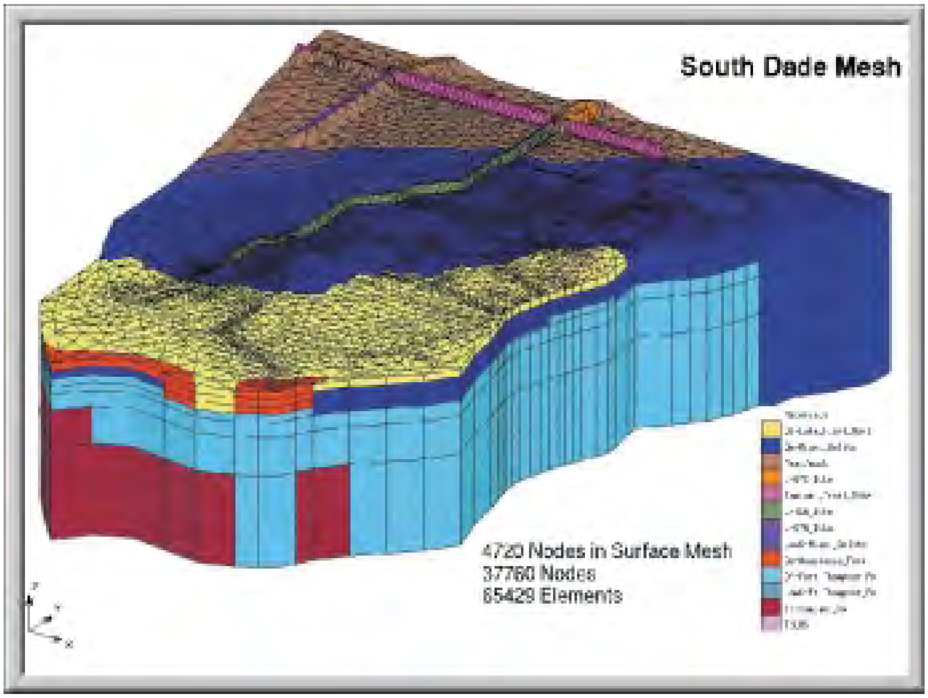


South Florida schematic geologic section.

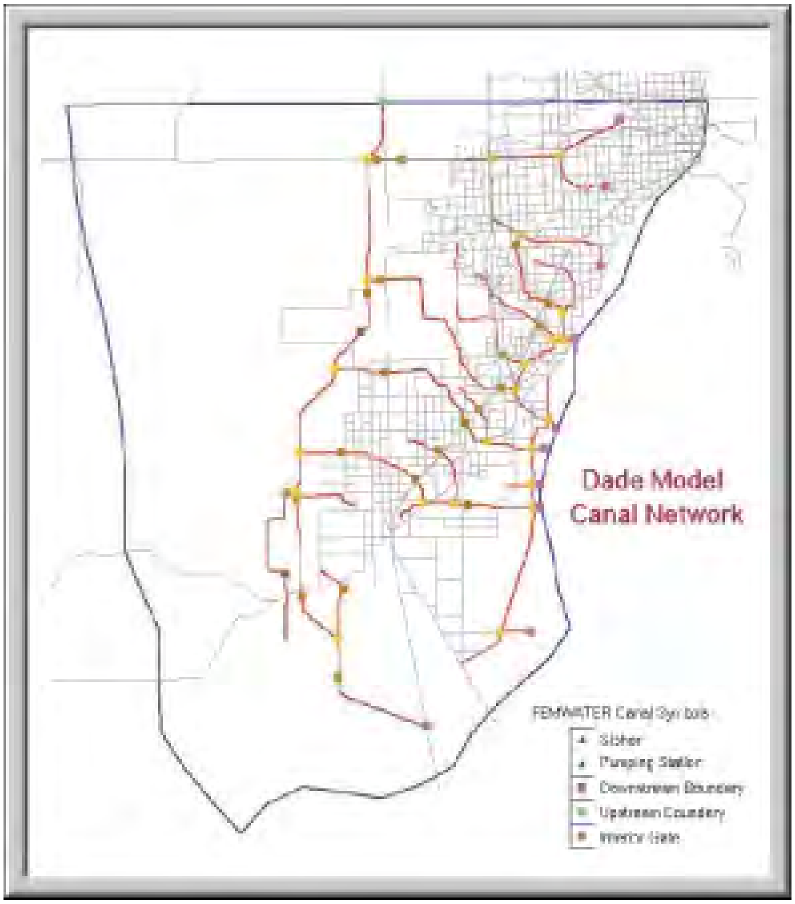
Added to the 3-D challenge are the 2-D overland flow computations on the 2-D mesh on the top of the 3-D mesh and the 1-D canal flow computations in the canal network. These are both nonlinear and also interact with the 3-D computation to form a highly nonlinear coupled system. At first, the 1-D solution used an explicit formulation, and it became unstable for time-steps greater than about 0.025 min. This resulted in the 1-D computations alone needing over 6 months of compute time using one T3E PE to accomplish the 1-year simulation. This was clearly unworkable. However, switching to an implicit formulation brought the computation time down to an acceptable level. The 2-D solution also took too long. However, by adjusting solver parameters, the 2-D solution then converged much faster.

Developing a parallel version of the FEMWATER-123 code presented other difficulties, such as how to partition the three separate meshes and parallelize the three separate program pieces efficiently. METIS was used to partition the 3-D mesh; the 2-D mesh was assigned a PE according to how the 3-D mesh was partitioned; and loop parallelization was used in the 1-D piece of code. Although more efficiency can be obtained, this was acceptable for the first version.

Finally, the interaction between the 2-D and 3-D solutions would not converge at first, as the solution oscillated and became very unstable in time. A new paradigm for boundary conditions was then developed, and the solution became very fast and stable. This allowed the successful completion of the 1-year simulation using 40 PEs



3-D finite element mesh for the Dade model with materials.



1-D Canal network for canal flow calculations for the Dade model.

in 16.5 hr on the Cray T3E, which was well under the needed time of 24 hr.

How large have your problems grown over the past 5 years? The complexity has grown by a factor of 10, and the size has grown by a factor of 2 or 3.

Do you anticipate them growing? Yes. The parallel version of FEMWATER (3-D only) and FEMWATER-123 will allow us to do even larger problems in the future with finer detail provided by larger meshes. The goal of doing a million-node 3-D mesh is now attainable.

What is the impact on the DoD? The parallel version of FEMWATER is currently being used to model the environmental remediation of military sites with specific application to pump-and-treat. Helping the DoD more efficiently clean up these military sites will result in saving millions of dollars for other use by the warfighter. Parallel FEMWATER-123 represents a significant advance in the state of the art and will be used in the future to model important DoD projects.

acronyms

Below is a list of acronyms commonly used among the DoD HPC community. You will find these acronyms throughout the articles in this newsletter.

ABL	AirBorne Laser
ASC	Aeronautical Systems Center
AVS	Application Visual Systems, Inc.
CAEEM	Computer Applications for Engineers and Engineering Managers
CoRA	Colorado Research Associates
DoD	Department of Defense
ENP	Everglades National Park
ERDC	Engineer Research and Development Center
HDS	Hierarchical Data Structuring
HiPPI	High-Performance Parallel Interface
HPC	High Performance Computing
HPCMP	High Performance Computing Modernization Program
I/O	Input/Output
ITL	Information Technology Laboratory
JSU	Jackson State University
MPP	Massively Parallel Processing
MSF	Mass Storage Facility
MSRC	Major Shared Resource Center
MSU	Mississippi State University
NAVO	Naval Oceanographic Office
NSF	National Science Foundation
NWRA	NorthWest Research Associates
PE	Processing Element
PET	Programming Environment and Training
PR2000	Program Review 2000
TRADOC	Training and Doctrine Command

The contents of this newsletter are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval or the use of such commercial products.

The ERDC MSRC welcomes comments and suggestions regarding *The Resource* and invites article submissions. Please send submissions to the following e-mail address:

info-hpc@wes.hpc.mil

training

Training Schedule*

June 27-28
Programming Tools

July 12-13
Workshop on Handheld Supercomputing
(to be held at Florida State University)

August 22-24
IBM Power3 SMP Applications Programming

*Additional courses may be offered. Please check the ERDC MSRC Web page at ***www.wes.hpc.mil***

ERDC MSRC Newsletter
Editoral Staff

Interdisciplinary Coordinator
Mary L. Hampton

Chief Editors
Rose Dykes
Mary Gabb

Visual Information Specialist
Betty Watson

ERDC MSRC Customer Assistance Center

Web site: *www.wes.hpc.mil*
E-mail: *info-hpc@wes.hpc.mil*
Telephone: 1-800-500-4722

Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the DoD. Your comments, ideas, and contributions are welcome. Design and layout provided by the Visual Production Center, Information Technology Laboratory, U.S. Army Engineer Research and Development Center.

Approved for public release; distribution is unlimited.

ERDC **Major Shared Resource Center**

U.S. Army Engineer Research and Development Center
ATTN: CEERD-IH (Rose Dykes)
3909 Halls Ferry Road
Vicksburg, MS 39180-6199