NPACI

PROCEEDINGS

of the
National Partnership for
Advanced Computational
Infrastructure

All-hands Meeting 2000
(AHM2000)

Held at the San Diego Supercomputer Center
February 8-12, 2000

NPACI All-hands Meeting 2000 Proceedings

A Message from NPACI Director Sid Karin

Hello —

We held our first All-hands Meeting in January 1998, just barely three months after the official start of NPACI in October 1997. Still green at the task of organizing such a big event, we commandeered an auditorium and a single additional room on campus and held meetings in every available space at SDSC. I think you'll agree that we've gotten better at managing the flow of participants and at tracking sessions so that attendees can customize their AHM experience.

With AHM2000, we also continued progress toward our original vision for these meetings. We went beyond presentations and demonstrations, ensuring even more of a working meeting. Sessions were engaging and interactive, issues were raised, discussions occurred, and collaborations were launched. We themed this year's meeting "crossing bridges" in acknowledgement of the complementary relationships within and between NPACI and the user community. The infrastructure is maturing, and the products of our efforts are beginning to find their place within academic research. Users we heard from at the meeting have been able to take advantage of NPACI tools and resources to further their investigations and leverage their funding. Next year, among other things, I expect that we will hear a lot about the impact of Blue Horizon, our new teraflops-capacity supercomputer that was formally dedicated during AHM2000.

I hope these proceedings reinforce the information you gathered at the NPACI All-hands Meeting. Please note that copies of slides presented at the meeting are available on the Web at **www.npaci.edu/ahm2000/presentations**.

Until next year....

"A spectacularly successful meeting."

–Sid Karin, NPACI Director

"Having a day just for EOT was very useful—I feel like I know the partners better now, and I could call any of them for a future collaboration or the replication of another program.... You did a good job of sprinkling EOT-related workshops throughout the other meeting days."

–Susan Ragan, Maryland Virtual High School
Winner of the Sun workstation at the AHM raffle

"The atmosphere of interchange was great. I liked interacting with the other attendees. I like the Wednesday-Friday format."

–Chuck Severance, University of Michigan

"The value of the AHM and the partnership for me is in the details—training and introductions to specific technologies and small group or individual interactions with NPACI staff at SDSC."

–Jim Beach, University of Kansas

"I'm glad that we did our tutorial on ADR/DataCutter—it helped us get our documentation and distribution put together.... The ability to run informal sessions (we had a PTE session, an alpha project session) was a key strength of the All-hands Meeting; there was also a high density of informal small group meetings that just came together. The AHM was one of the most productive meetings I've been to in years."

–Joel Saltz, University of Maryland
& programming tools and environments thrust area leader

"I don't go to SCXY anymore; I come to the AHM instead."

–Russ Altman, Stanford University
& Molecular Science thrust area leader

"It gave us the opportunity to learn what other research is being carried out... and, most importantly, we learned how to use the facility more efficiently."

–Chi-Yu Hu, California State University at Long Beach

"I enjoyed the poster session. This was an opportunity for one-on-one interaction with the folks who are doing the research, with all the details, and as such it was a good balance to the technical sessions, which provided more of an overview of their subjects."

–Tim Walsh, University of Texas

"You might be able to match this meeting next time, but there's no way you can top it."

–Tinsley Oden, University of Texas
& Engineering thrust area leader

NPACI All-hands Meeting at a Glance

Where these proceedings include more detailed information on a session, a page number is provided following the session title.

For PowerPoint slides, where available, see:
www.npaci.edu/ahm2000/presentations

Tuesday, February 8, 2000

Education, Outreach, and Training *(Education Session, pg. 10)*
Advanced Networking-MSI Grant: Status and How to Get Involved *(Education Session, pg. 10)*
Portals: What They Are and How They Can Support EOT Projects *(Education Session, pg. 10)*
EdGrid: Status and How to Get Involved *(Education Session, pg. 10)*
Guidelines on Self-evaluations for EOT Projects *(Education Session, pg. 10)*
Tracking Student Participants in PACI Programs *(Education Session, pg. 11)*
SDSC K-12 Science Enrichment Programs: Ideals, Realities, and Plans *(Education Session, pg. 11)*
Getting Involved with SC2000-2001 Education Program *(Education Session, pg. 11)*

The EOT-PACI Web Site *("Birds of a Feather," pg. 12)*

K-12/EdGrid Team *(Breakout Session)*
Undergrad Team *(Breakout Session)*
Access and Inclusion Team *(Breakout Session)*

Wednesday, February 9, 2000

Very Large Data Set Access and Manipulation: Active Data Repository, DataCutter, and the SDSC Storage Resource Broker *(Tutorial, pg. 3)*
CONTACTS: Joel Saltz **<saltz@cs.umd.edu>**, Alan Sussman **<als@cs.umd.edu>**, Tahsin Kurc **<kurc@cs.umd.edu>**

Accessing Deployed Data Collections *(Tutorial, pg. 3)*
CONTACT: Reagan Moore **<moore@sdsc.edu>**

Legion: How It Works and Hands-on Lab *(Tutorial, pg. 3)*
CONTACT: John Karpovich **<jfk3w@cs.virginia.edu>**

The Globus Toolkit *(Tutorial, pg. 3)*
CONTACTS: Ann Chervenak **<annc@isi.edu>**, Karl Czajkowski **<karlcz@isi.edu>**, Steven Fitzgerald **<sfitzger@isi.edu>**, and Carl Kesselman **<carl@isi.edu>**

Developing and Optimizing Applications for the IBM Teraflops and Sun HPC 10000 *(Tutorial, pg. 3)*
CONTACTS: Donald Frederick **<frederik@sdsc.edu>**, Tim Kaiser **<tkaiser@sdsc.edu>**, and Amit Majumdar **<majumdar@sdsc.edu>**

PTE Tools: KeLP & Titanium *(Tutorial, pg. 4)*
CONTACTS: Scott Baden **<baden@cs.ucsd.edu>** and Katherine Yelick **<yelick@cs.berkeley.edu>**

Hands-on Tools for Scientific Visualization *(Tutorial, pg. 4)*
CONTACT: Bernard Pailthorpe **<bap@sdsc.edu>**

Information Integration Using XML *(Tutorial, pg. 4)*
CONTACT: Chaitan Baru **<baru@sdsc.edu>**

Biology Workbench *(Tutorial, pg. 4)*
CONTACTS: Brian Saunders **<saunders@sdsc.edu>** and Shankar Subramaniam **<shankar@sdsc.edu>**

éTEACH: Instructional Technology Tool *(Tutorial, pg. 4)*
CONTACT: Michael Litzkow **<mike@cs.wisc.edu>**

Dedication Ceremony for the NPACI Teraflops IBM SP Supercomputer at SDSC *(Special Presentation, pg. 21)*

Programming Tools and Environments Thrust Area *(Breakout Session)*
Data-intensive Computing Thrust Area *(Breakout Session)*
Interaction Environments Thrust Area *(Breakout Session)*
Molecular Science Thrust Area *(Breakout Session)*
Earth Systems Science Thrust Area *(Breakout Session)*
Engineering Thrust Area *(Breakout Session)*
EOT Leadership Team Meeting *(Breakout Session)*

Thursday, February 10, 2000

Welcome to the All-hands Meeting 2000 *(Plenary, pg. 5)*
Bridges to the User Community: Maximizing the Impact of NPACI Infrastructure *(Plenary, pg. 5)*
Trends in High-end Scientific Computing *(Plenary, pg. 5)*

High-end Compute Resources *(Technical Session, pg. 6)*
Programming Tools and Environments Tools *(Technical Session, pg. 6)*
Instructional Technology and Asynchronous Learning *(Technical Session, pg. 6)*
Network Research and Tools *(Technical Session, pg. 6)*
Building Tools for Scalable Volume Visualization from Brains to Bays and Beyond *(Technical Session, pg. 7)*
What Users Want from NPACI *(Technical Session, pg. 7)*

Student and Researcher Poster Session *(Poster Session, pp. 14-20)*

Q&A on High-end Compute Resources *("Birds of a Feather," pg. 12)*
HPC Software Exchange *("Birds of a Feather," pg. 12)*
Scaling Successful EOT Programs *("Birds of a Feather," pg. 12)*

Friday, February 11, 2000

Future Directions and Challenges: A Panel Discussion *(Plenary, pg. 5)*

Developing a Cluster Strategy for NPACI *(Technical Session, pg. 7)*
Building Data Infrastructure and Data Collections Being Assembled *(Technical Session, pg. 7)*
Moving Computational Science into the Undergraduate Curriculum *(Technical Session, pg. 8)*
Strategic Applications Collaboration Projects (*Technical Session, pg. 8*)
Using Data Infrastructure *(Technical Session, pg. 8)*
Connecting with Minority-Serving Institutions *(Technical Session, pg. 8)*
Web Portals and Applications *(Technical Session, pg. 9)*
Using the Grid: Application Experiences within NPACI *(Technical Session, pg. 9)*
Computational Biology: How NPACI Resources Are Used by Biologists *(Technical Session, pg. 9)*
Mentoring Women and Minorities *(Technical Session, pg. 9)*

NPACI's Exhibit at SC99: The Good, Bad, and Ideas for SC2000 *("Birds of a Feather," pg. 12)*
Using Local Campus Infrastructure vs. PACI Resources: Conflict or Synergy? *("Birds of a Feather," pg. 13)*
EOT Participation in the Alpha Projects *("Birds of a Feather," pg. 13)*
Emerging Visualization Technologies: Java3D and VRML *("Birds of a Feather," pg. 13)*
Helping Partners Tune Their Networking Infrastructure *("Birds of a Feather," pg. 13)*

Very Large Data Set Access and Manipulation: Active Data Repository, DataCutter, and the SDSC Storage Resource Broker

Taught by Joel Saltz, Alan Sussman, & Tahsin Kurc, University of Maryland

The design and implementation of systems for managing and manipulating very large data sets, both on disk and in archival storage, were covered in this tutorial. Two systems—the Active Data Repository (ADR) and DataCutter—were described in detail. ADR is designed to optimize storage and processing of large disk-based data sets on a parallel machine or network of workstations, while DataCutter is designed to provide subsetting and filtering operations on data sets stored in archival (tertiary) storage systems in a grid environment.

Interfaces for customizing ADR for data-intensive applications were explained, and an example application was presented to illustrate the customization. Customization includes storing and indexing data sets into ADR and providing user-defined processing functions for the end application.

The design and current implementation of DataCutter were described, including the design and implementation of filter-based applications. Finally, integration of DataCutter with the SDSC Storage Resource Broker (SRB), to provide subsetting and filtering operations in that framework, was discussed.

Accessing Deployed Data Collections

Taught by Reagan Moore, SDSC

This session covered access to NPACI-supported data collections, including data collections that NPACI is helping establish and collections that have been developed independently of NPACI but are stored or available on NPACI resources. Presentations included information on data available within the collection, query mechanisms for locating data sets, and access mechanisms for reading data sets. The various data collections covered included the Storage Resource Broker access description (George Kremenek, SDSC); Protein Data Bank (Phil Bourne, SDSC); 2MASS, 2-Micron All Sky Survey digital library (Wen Piao Lee, University of Massachusetts); UC Berkeley Digital Library (Ginger Ogle, UC Berkeley); REINAS (Ted Haining, UC Santa Cruz); Neuroscience data (Ken Wong, Washington University); University of Maryland NASA archive (Joseph JaJa, University of Maryland); Alexandria Digital Library (Greg Hajic, UC Santa Barbara); and Digital Puglia Library (Massimo Cafaro, University of Lecce).

Legion: How It Works and Hands-on Lab

Taught by John Karpovich, University of Virginia

This full-day tutorial covered basic user information related to Legion. It provided background on the Legion system and instruction for running existing parallel codes within the Legion environment. The target audience was supercomputer users already familiar with parallel-processing tools, such as MPI and PVM.

The tutorial consisted of an introduction to the Legion system, philosophy, architecture, and object model, and was followed by an in-depth presentation of the user's view of Legion. Issues such as logging into the system, compiling and registering binaries, and using MPI were addressed.

A hands-on lab running applications on Legion demonstrated various kinds of applications that have been compiled and run using the program. These included legacy applications, parameter space programs, wide-area parallel applications, MPI, and coupled applications. Participants were encouraged to bring their own applications to try out.

The Globus Toolkit

Taught by Ann Chervenak, Karl Czajkowski, Steve Fitzgerald, & Carl Kesselman, University of Southern California

This tutorial was a practical introduction to programming for high-performance distributed computing systems, or "computational grids," and covered the capabilities of the Globus grid toolkit.

Emerging high-performance networks promise to enable a range of new applications such as remote computing, distributed supercomputing, tele-immersion, smart instruments, and data mining. However, development and use of such applications in practice is very difficult and time-consuming because of the need to deal with complex and heterogeneous systems.

The Globus grid programming toolkit is designed to help applications developers and tool builders overcome these obstacles so that they can construct "grid-enabled" scientific and engineering applications. It does this by providing a set of standard services for authentication, resource location, resource allocation, configuration, communication, file access, fault detection, and executable management. These services can be incorporated into applications and/or programming tools in a "mix-and-match" fashion to provide access to needed capabilities.

This tutorial's goal was to introduce the capabilities of the Globus toolkit and show attendees how Globus services can be applied in specific applications. A mixture of grid programming principles and detailed Globus toolkit examples was presented from the perspective of tool users.

Developing and Optimizing Applications for the IBM Teraflops and Sun HPC 10000

Taught by Donald Frederick, Tim Kaiser, & Amit Majumdar, SDSC

Topics covered in this class included an overview of parallel computing, the architecture of the new NPACI parallel machines (teraflops IBM SP and the Sun HPC 10000), development of parallel applications codes using MPI and OpenMP, compiling, and running jobs on these new machines. A hands-on lab provided attendees with the opportunity to compile, run, and parallelize example codes.

PTE Tools: KeLP and Titanium

Taught by Scott Baden, UC San Diego, & Katherine Yelick, UC Berkeley

KeLP is a run-time library for implementing applications with an underlying block structure, including irregular problems such as structured adaptive mesh refinement and multiblock methods. It supports run-time metadata and has the ability to manipulate geometric descriptions of data decompositions and data dependencies as first class language objects. The library complements Titanium by enabling the user to customize communication and data motion directly. This tutorial taught the basics of the KeLP programming model including fundamental KeLP data types and some example programs. An irregular application was included in the tutorial. Future KeLP developments were outlined, including enhanced interoperability (Abstract KeLP) and multi-tier programming for multi-tier platforms—such as the NPACI teraflops machine—that employ symmetric multiprocessor nodes.

Titanium is a Java-based language for large-scale scientific computing. It gives users access to modern program structuring through object-oriented technology and lets users write explicitly parallel code to exploit their understanding of the computation. The language is based on Java but adds SPMD parallelism constructs, primitives for layout and access to shared data structures, efficient multi-dimensional arrays, immutable classes, and region-based memory management. The Titanium language and optimizing compiler provide users of adaptive mesh refinement—and other techniques based on rectangular grids—with a convenient language to express these computations. It complements KeLP by providing modern compiler analyses and optimizations for this class of applications. This tutorial described the Titanium language and system as well as some of the advanced solvers that have been developed in Titanium.

Hands-on Tools for Scientific Visualization

Taught by Bernard Pailthorpe & Nicole Bordes, SDSC, & Ariel Shamir, University of Texas, Austin

This tutorial introduced the notion of scalable interrogative visualization. Several tools developed at the Center for Computational Visualization (University of Texas, Austin) were demonstrated, including Visual Eyes and Interrogative Visualization Tools for Scalable Visualization. Demonstrations used examples from scientific simulations and imaging data. These tools allow the user not only to see the data (volumes, iso-surfaces), but also to navigate and query for increased understanding (measurements and spectrum). Furthermore, using multi-resolution and compression techniques, such tools can be used on a variety of platforms ranging from laptop PCs to high-end visualization engines.

Information Integration Using XML

Taught by Chaitan Baru, Bertram Ludaescher, Richard Marciano, & Ilya Zaslavsky, SDSC; Yannis Papakonstantinou, UC San Diego; & the FedCon/Digital Government Group

Issues related to information integration, including evolving XML-based standards and technologies and related NPACI and SDSC projects were covered in this tutorial. XML, XSL, and the XML schema were presented and the SDSC MIX project—including the XMAS, DOM-VXD, and BBQ subprojects—was discussed. Participants were introduced to XML data servers and RDBMS support for XML, and were given an overview of AXL and the use of VML and SVG in the context of GIS and XML. The second half of the tutorial focused on information integration projects in NPACI as well as in various federal government agencies.

Biology Workbench

Taught by Brian Saunders & Shankar Subramaniam, SDSC

The Biology Workbench is a revolutionary Web-based tool that allows biologists to search many popular protein and nucleic acid sequence databases. Searching is integrated with access to a wide variety of analysis and modeling tools, all within a point-and-click interface that eliminates file-format compatibility problems.

Biology Workbench 3.2 (**workbench.sdsc.edu**) was introduced in this tutorial and preliminary plans for version 4.0 were discussed.

eTEACH: Instructional Technology Tool

Taught by Michael Litzkow, University of Wisconsin

eTEACH is an authoring tool that produces a Web-based presentation that can be viewed with a standard Web browser (Internet Explorer or Netscape). The authoring tool has a point-and-click interface to allow the author to choreograph a presentation, including up to two video streams, slides such as PowerPoint slides, and external URL references. This is all available through a dynamic table of contents that allows the viewer to access the video stream randomly. The authoring tool hides the complexities of file formats, XML DTDs, etc., and allows the non-technical author to produce multimedia presentations. eTEACH presentations can be designed for packaging on-campus and distance-learning lectures, research presentations, tutorials, online documentation, informal education modules, etc.

Welcome to the NPACI All-hands Meeting 2000

Presented by Bill Martin, Chair, AHM2000 Program Committee; Sid Karin, Director, NPACI; & Bob Dynes, Chancellor, UC San Diego

Bill Martin welcomed the participants to the third NPACI All-hands Meeting and noted that the high attendance was "a very clear instantiation of our partnership." Martin warned that the meeting turnout was higher than expected and that all sessions were likely to be crowded. The theme "From Building to Crossing Bridges" guided all planning decisions for AHM2000, including sessions focusing on the needs of the user community. Martin introduced the other Program Committee members and outlined the meeting schedule for February 10 and 11. Sid Karin and Bob Dynes each shared the positive impact that NPACI and SDSC have made at UC San Diego.

Following Chancellor Dynes' remarks, Bernard Pailthorpe, SDSC associate director for Scientific Visualization, introduced a two-minute video of a starship flight through the Orion nebula. The video was produced by SDSC for the Hayden Planetarium at the American Museum of Natural History's Rose Center for Earth and Space in New York City. Created using the Galactic MPIRE visualization package running on the Tera MTA at SDSC, the video was made by David Nadeau, Jon Genetti, and Greg Johnson for the planetarium's Digital Galaxy project.

Bridges to the User Community: Maximizing the Impact of NPACI Infrastructure

Presented by Peter Taylor, Chief Applications Scientist, NPACI, and Deputy Director for Computational Science, SDSC

This plenary talk reviewed the first two years of NPACI activities in the context of the external user community. Developments in capability computing and discovery environments have enabled new science to be done in a variety of disciplines. The challenge for the partnership now is to extend the power of this infrastructure to the larger user community. In addition to the five ongoing Alpha projects within the NPACI thrust areas (**www.npaci.edu/Alphas**), extension of the infrastructure is occurring through Strategic Applications Collaboration (SAC) projects, which represent a new strategy for cooperation and coordination with specific user groups. The goal for the SACs is for NPACI computer scientists to work closely with researchers to port and improve their codes on NPACI resources.

The first three SAC collaborations—with astrophysics, biochemical, and biomedical engineering users—were quite successful; for example, execution of ParTree code on an SP processor doubled in speed. Seven new SAC projects are in progress.

Plenary: Trends in High-end Scientific Computing

Presented by Paul Messina, Chief Architect, NPACI, & Associate Deputy Assistant Secretary for Advanced Simulation and Computing, DoE (on leave from Caltech)

This talk discussed the state of the art and current trends in advanced computing. Discussing the short term, Messina presented an overview of the DoE's Accelerated Strategic Computing Initiative (ASCI), a key element of the department's program to replace underground nuclear weapons testing with computer simulation and assessment. ASCI has enormous computing requirements, including speeds of up to 100 TeraOPS (the preferred DOE term for teraflops) by the year 2004. Other short-term activities were discussed, as were longer-term system projects such as HTMT and Blue Gene.

Messina also discussed trends in the software environment of advanced computing—including system software, tools, middleware, and gridware—as well as networks, databases and archival storage, and visualization.

The audience asked about requirements to do calculations on machines comparable to ASCI's Blue Mountain or Blue Pacific. Messina responded that though large calculations still take days to run, improved I/O speed on newer machines—such as SDSC's Blue Horizon—necessitate less user monitoring. Restarting the machine, if necessary, is also easier.

Future Directions and Challenges: A Panel Discussion

Chaired by Susan Graham, Chief Computer Scientist, NPACI
Panelists: Aron Kupperman, Caltech; Gwen Jacobs, Montana State University; Andrew Chien, UC San Diego; & Bill Carlson, Institute for Defense Analysis, Center for Computing Sciences

This plenary featured a panel discussion on the future directions and challenges of high-performance computing in fields such as chemistry and biology. The approximately 200 attendees at this session heard presentations from chemistry (Kuppermann), computer science (Chien), neuroscience (Jacobs) and high-end computing (Carlson). The impact of robust large-scale clusters on the high-performance computing landscape was discussed, with panelists agreeing that researchers are looking for more sophisticated problem-solving environments that will allow them to conduct experiments and integrate results across other research efforts. NPACI's new teraflops IBM SP system was presented as a significant resource to enable breakthroughs in such areas as climate modeling and modeling the human brain. Panelists felt that researchers would require more advanced data acquisition and database tools to work with these types of complex simulations. These tools should have seamless interfaces and provide access to federated data. Researchers emphasized that they do not want to "reinvent the wheel."

High-end Compute Resources

Chaired by Wayne Pfeiffer, SDSC

Pfeiffer gave an overview of NPACI's computing resources and the complementary tools of the six compute resources sites within the partnership. Details regarding individual machines were given by other speakers: Blue Horizon, the IBM teraflops SP (Jay Boisseau, SDSC); HP V2500 (Sharon Brunett, Caltech); Cray SV1 (Kent Milfeld, University of Texas, Austin); Sun HPC 10000 (Amit Majumdar, SDSC); and Tera MTA (Allan Snavely, SDSC). There were some questions regarding Blue Horizon's network performance, which Boisseau assured was stable, and the possibility of SDSC coordinating with Lawrence Livermore Laboratory, which also recently acquired a teraflops machine. According to Boisseau, SDSC and Livermore staff have already attended some IBM SP meetings. Regarding the Sun HPC 10000, an attendee asked about the differences between the various computational platforms. Brunett said the need for a particular platform depends on the program being run. Another attendee asked about compilers on the Tera MTA, and Snavely explained that the MTA compiler divides a program into threads then assigns the threads to the processors, which can each execute up to 128 threads concurrently. Snavely added that he and some of the other speakers would be available at the BOF "Q & A on High-end Compute Resources" to answer further questions. Questions regarding NPACI resources should be directed to NPACI Consulting at **www.npaci.edu/Consult/**.

Programming Tools and Environments Tools

Chaired by Joel Saltz, University of Maryland & The Johns Hopkins University

Presentations were made describing the status of projects within the PTE thrust area: KeLP (Scott Baden); NetSolve, SinRG, and RIB (Jack Dongarra); linear algebra libraries (Osni Marques); the NERSC-developed ACTS Toolkit (Marques); and other PTE activities including Titanium, Out-of-core Compiler, Planguages, ADR, and Meta-Chaos (Joel Saltz). The summary message from Joel Saltz was that NPACI PTE tools are gradually converging on a small set of application programming interfaces (APIs) for developing scientific applications. Some discussion, shortened by time constraints, covered the Advanced Computational Testing and Simulation (ACTS) Toolkit (**acts.nersc.gov**). The ACTS Toolkit is a set of DoE-developed software tools that make it easier for programmers to write high-performance scientific applications for parallel computers. There was discussion about the availability of ACTS or an ACTS-like environment to NPACI partners.

Instructional Technology and Asynchronous Learning

Chaired by Charles Severance, University of Michigan

Attendees of this session were greeted with a Web-run introduction prepared moments before by session chair Charles Severance using ClipBoard-2000, an asynchronous learning tool. The intro demonstrated the feasibility of capturing live video and quickly processing it to serve over the Web, and complemented the session's purpose: to share experiences using technology in education and training. The attendees also heard from Kris Stewart (San Diego State University) who discussed the Faculty Fellows program instituted at the Education Center for Computational Science and Engineering at SDSU. Stewart emphasized the importance of finding "reform-ready" faculty to work with on curriculum development using HPC and encouraged session attendees to contribute demonstrations, modules, or full curriculum that they have developed to her curriculum cache at the Ed Center.

Greg Moses (University of Wisconsin, Madison) presented éTeach, a tool he has developed that allows an instructor to pair audio and video with slide presentations. éTeach will be used by Moses at University of Wisconsin, Madison to reverse the traditional lecture-homework-discussion paradigm of undergraduate education. Students will access éTeach lectures at home over the Web, and class time will be devoted to interactive lessons to help students complete homework requirements. This teaching model will be evaluated.

éTeach has also been used by Max Zain and Ed Borbely of the Center for Professional Development at the University of Michigan; they are involved in the NPACI Resources group. They have been able to prepare packaged training modules on NPACI resources that can be used to train new users between "live" training sessions. The modules are fed from a database system and include slides and video.

Network Research and Tools

Chaired by Hans-Werner Braun & David Moore, SDSC

The National Laboratory for Applied Network Research (NLANR) was created to support the NSF's High Performance Connections program. The Measurement and Analysis group (**moat.nlanr.net/**), centered at SDSC, has as its goals to create an infrastructure to support measurements and network analysis, further high-performance networking environments, and analyze network header traces, performance, statistics, and routing. It has deployed active and passive monitoring devices (Intel PCs with OC_3 and OC_{12} interfaces) at key locations on the vBNS and other networks to support these activities. NLANR also has developed tools such as Cichlid for 3-D data visualization and TCPtune for tuning TCP stacks in Windows.

The Cooperative Association for Internet Data Analysis (CAIDA) was created October 1997 with seed funds from NSF to promote cooperation in engineering and maintaining a robust, scalable Internet infrastructure. Based at SDSC, CAIDA is now funded by NSF, DARPA, and its member institutions and companies. CAIDA measures, analyzes, and visualizes network topology, workload, performance, and routing. Its goal is to monitor, depict, and predict traffic behavior on current and advanced networks. CAIDA is developing and deploying tools to better engineer and operate networks and identify traffic anomalies in real time. These include skitter (active measurements of performance and topology), coral and cflowd (passive measurements), and manta and otter (visualization of multicast infrastructure). CAIDA also operates the Internet Engineering Curriculum Repository (IECR). See **www.caida.org/**.

For PowerPoint slides, where available, see:
www.npaci.edu/ahm2000/presentations

Building Tools for Scalable Volume Visualization from Brains to Bays and Beyond

Chaired by Art Olson, The Scripps Research Institute, & Bernard Pailthorpe, SDSC

This talk provided an overview of NPACI volume visualization toolkits being designed to enable efficient access to multiple multimodal, time-varying, tera-scale volume data sets. The toolkits support parallel architectures, access data sets on disk, where appropriate, by using ADR, and manage data movement among processors by using KeLP. When volume rendering is needed, the toolkits interface with VISTA (Volume Imaging Scalable Toolkit Acronym) as well as interactive hardware systems, such as the VolumePro rendering card and standard 3-D graphics pipelines. The toolkits are being designed to integrate easily with existing visualization software, such as Vtk, AVS, DX, and other visualization applications.

David Nadeau and Jon Genetti (SDSC) gave presentations on applications under development at SDSC (expression trees, volume tools, and VISTA) and commented that these programs are still most appropriate for use by seasoned programmers, though they are considering how to make them more user friendly. New partner Joerg Mayer (UC Davis) presented volume visualization of large-scale biomedical data, and Chandrajit Bajaj (University of Texas, Austin) talked about interrogative visualizations, the Visual Eyes program, and the new visualization "theater" under construction at the University of Texas, Austin.

What Users Want from NPACI

Chaired by Aron Kuppermann, Caltech

This session addressed several topics, including how users can submit their input (Donald Frederick, SDSC); NPACI Consulting, the Remedy database, and on-line documentation (Laura Nett-Carrington, SDSC); NPACI applications, tools, and how to work effectively with computational scientists and computer scientists (Abhijit Bose, University of Michigan); NPACI allocations (Bob Sinkovits, SDSC); and the NPACI User Survey (Nancy Wilkins-Diehr, SDSC). Frederick outlined ways to reach NPACI and said that e-mail (**consult@npaci.edu**) or the Web (**www.npaci.edu/Consult**) was most efficient. Questioned as to whether users can see responses to past surveys, Frederick and Wilkins-Diehr said they would take this idea into consideration, particularly since they have survey results from the past several years. Nett-Carrington referred users to the Frequently Asked Questions page (**www.npaci.edu/Documentation**) to see questions that come up repeatedly. Some were concerned that this on-line documentation is not aimed at new users and suggested that a New Users page with tips and tricks be created. Nett-Carrington said she would bring this up to the Scientific Computing Group.

Several suggestions were made regarding NPACI allocations, including that users be able to give a first choice, second choice, etc., for which machine they wish to use. Sinkovits replied that this is already possible as long as users specify which type of machine they are interested in. Other questions addressed how to make sure users get their awarded hours, whether over-allocation can become a problem (no, as hardware is always being added), and how users can find out which platform they should use (consult **www.npaci.edu/Allocations** and **www.npaci.edu/Resources/Systems/compute.html**). Responding to a question about the possibility of multi-year allocations, Sinkovits said they are not feasible, due to changing hardware resources and the Allocations Committee's desire to keep track of who does or does not use their allocated time.

Developing a Cluster Strategy for NPACI

Chaired by Andrew Chien, UC San Diego

In this session, the panel of speakers and attendees discussed the growing popularity of clusters and how NPACI could become a participant or a leader in the field. The speakers presented the pros, including a low cost-to-performance ratio, and cons, such as software and operating system issues, for large-scale clusters. The ensuing discussion identified that the rise of clusters, while not diminishing the need for centers like SDSC to support the highest-performance clusters, could force NPACI to change its operational model. Support of clusters may entail addressing challenges in software development and system administration.

To stay abreast of developments, it was suggested that NPACI install a cluster with 128-256 nodes (FY00), 256-512 nodes (FY01), and 512-1024 nodes (FY02). To become a leader in the field, according to the presenters, NPACI must install a cluster with 256 nodes (FY00), more than 1024 nodes (FY01), and 2048-4096 nodes (FY02). A paper by Chien, based on this discussion, is available at **www-csag.ucsd.edu/individual/achien/npaci00/cluster-strategy.html**.

Building Data Infrastructure and Data Collections Being Assembled

Chaired by Chaitan Baru, SDSC

This session provided an overview of NPACI's data infrastructure, including the High Performance Storage System (HPSS), the SDSC Storage Resource Broker (SRB), and XML-based technologies. Presentations focused on collections and contributions to the data infrastructure. Research efforts have focused on increasing distribution, centralizing access, enhancing curation, improving query, and providing rich metadata capability. Presentations included:

- CEED: 'Caveat Emptor' Ecological Data Repository (**ceed.sdsc.edu**). The first attempt to create a data repository for Earth Systems Science-related collections. Contact Don Sutton, SDSC.
- Protein Data Bank (PDB, **www.rcsb.org/pdb/**). The single international repository for processing and distribution of 3-D macromolecular structure data primarily determined experimentally by X-ray crystallography and nuclear magnetic resonance (NMR). Contact Peng Yang, SDSC.
- Building a Neuro-Anatomy Data Collection: What Do You Want to Do With the Data? Neuroscience collections (**www.npaci.edu/DICE/Neuro**) at UC San Diego. Contact Amarnath Gupta, SDSC.
- Community of Science, Inc. (COS, **www.cos.com**), a network of scientists and research organizations on the Web. Contact Muhammad Rabi, COS.

- Digital Embryo Project. SDSC, in association with the National Library of Medicine, is helping George Mason University create a digitized embryo repository of 700 specimens from the Carnegie collection housed at the Armed Forces Institute of Pathology. Contact Reagan Moore, SDSC.

Moving Computational Science into the Undergraduate Curriculum

Chaired by Kris Stewart, San Diego State University, & Rubin Landau, Oregon State University

Panelists Stewart, Landau, J.P. Bayard (CSU Sacramento), Roscoe Giles (Boston University), and Greg Moses (University of Wisconsin, Madison) discussed how they have used computational science in undergraduate curriculum. The attendees heard how Landau has introduced a B.S. degree in computational science at Oregon State University. Coursework includes introductory courses in computing supplemented by the application of computing skills to problems in physics. Landau anticipates that codes developed within NPACI, including MCell and others, may be used to demonstrate concepts in physics. Bayard discussed his work with the National Institute for Science Education (NISE) and research that has identified successful teaching methods that make use of technology. Stewart presented the Faculty Fellows program at the Education Center for Computational Science and Engineering at SDSU and how it has helped "reform-ready" faculty begin developing curriculum using high-performance computing technology. The Sociology Workbench was discussed as an NPACI tool that is useful in undergraduate instruction. Giles encouraged attendees to attend EOT education workshops (including the upcoming Shodor Computational Science Institute) as a way to get involved with outreach. He also talked about partnerships between EOT and others with similar objectives (BioQuest Curriculum Consortium) and presented the EdGrid, AN-MSI, and AIHEC projects of EOT as examples of how EOT-PACI is pursuing systemic educational change. Finally, Moses talked about using éTeach in engineering courses and how the Web-based lectures allow him to reverse the traditional teaching paradigm. Lectures are viewed over the Web, and class time is devoted to hands-on activities that help students complete homework assignments.

Strategic Applications Collaboration Projects

Chaired by Jay Boisseau & Bob Sinkovits, SDSC

Projects in the Strategic Applications Collaborations program (**www.npaci.edu/SAC/**) expedite the efforts of researchers working on Grand Challenge-class computational problems. NPACI staff work with researchers on short-term efforts to improve the performance of codes, ideally by using generic techniques that can be applied elsewhere. The projects also develop stronger working relationships between users and staff. Three initial problems, each with a "good potential for success" within a 3-12 month period, were a test of the SAC program: galaxy collision astrophysics (L. Hernquist, Harvard), biochemistry (P. Kollman, UCSF), and biomedical engineering (C. Peskin, NYU). These were successful, and seven additional SAC projects have begun: climate modeling (T. Barnett, SIO; W. Washington, NCAR), particle physics (T. Kinoshita, Cornell), cochlea modeling (E. Givelberg, University of Michigan), quantum chromodynamics (R. Sugar, UC Santa Barbara), magnetic recording (N. Bertram, UC San Diego), neuron modeling (J. Bower, Caltech), stellar atmospheres (P. Hauschildt, University of Georgia; E. Baron, Oklahoma University). Code speedups typically have been x2 to x12, but even small speedups are considered "successful" if hundreds of hours of processing are involved. Audience suggestions included performing code critiques for users, involving tools developers in the process, and giving presentations at application-specific conferences.

Using Data Infrastructure

Chaired by Heather Drury, Washington University

This session gave an overview of an emerging family of databases spanning a wide range of species (including human, monkey, mouse, and cricket), diverse types of experimental data (structural, functional, and developmental), and spatial scales ranging from microscopic (cellular, molecular) to macroscopic (systems). The databases also span a range of maturity from fully implemented, internationally used databases to ones that are still in design stage. Presentations included the development of a multi-resolution 3-D neuron database (Maryann Martone, UC San Diego); the Center for Computational Biology database and connecting remote data to distributed tools (Sandy Pittendrigh, Montana State University); a cortical surface management system, SuMS (James Dickson, Washington University); knowledge mechanics, a low-level approach to generating a federated knowledge management system (Gully Burns, University of Southern California); and molecular biology-related collections (Phil Bourne, SDSC). There is a strong need for a computational infrastructure that permits the federation of databases containing complementary information, such as those with information about the brain. This session also included a discussion of the first, practical steps in designing a "federation engine" to permit queries of multiple databases, as well as the prospects for having a small set of identifiable visualization tools applicable across multiple databases.

Connecting with Minority-serving Institutions

Chaired by Allison Clark, NCSA

PACI partners are encouraged to couple their existing programs and projects with Minority-Serving Institutions (MSIs) through EOT-PACI's new Advanced Networking Project with Minority-Serving Institutions (AN-MSI) program. Allison Clark, Program Director of Access and Inclusion Initiatives at the National Center for Supercomputing Applications (NCSA), chaired a panel discussion on why PACI partners should establish collaborations with MSIs. This panel consisted of Richard Alo (University of Houston-Downtown), John Hurley (Clark Atlanta University), and Mark Trebian (Ojibwa Community College).

Each panelist discussed his projects and programs, emphasizing that collaboration is critical to students' success. These MSIs want to reach the rest of the academic and research community, making them aware of their excellent faculty, diverse student body, and the strong impact these groups have had on their own research efforts.

For PowerPoint slides, where available, see:
www.npaci.edu/ahm2000/presentations

Under NSF's AN-MSI program, colleges and universities that traditionally serve African-American, Hispanic, and Tribal communities will develop the infrastructure and skills needed to take advantage of advanced computational tools and resources, such as the technology grid being prototyped by the Alliance and NPACI. Roscoe Giles and Allison Clark administer the EOT-PACI portion of the grant.

Web Portals and Applications

Chaired by Jay Boisseau, SDSC, & John Towns, NCSA

NPACI has developed the HotPage, a resource portal, to meet the needs of scientists who are not computational experts. The Alliance has developed MyGrid for the same purpose. Both portals provide informational and interactive services with the intent of streamlining the amount of specific knowledge required to access and use resources. For example, an interactive version of HotPage—anticipated to be available in Beta version at the end of March 2000—will provide users with a single HotPage account that will be their Web-based passport to all other services, including the ability to load and launch programs through any Web browser. At this session, attendees learned about plans to leverage HotPage and MyGrid efforts, creating a PACI portal that will have a common security and authentication scheme (using Grid Security Infrastructure from Globus). Attendees were asked to provide input about "community" features that should be included in the portal and were encouraged to participate in the Grid Forum meetings. Technology transfer and brokering relationships with others who have developed useful portals were also discussed as ways of making the portal development process more efficient. Finally, the possibility of creating a PACI "grid Web toolkit" was presented. The toolkit will aid in developing and deploying new portals and applications more rapidly.

Using the Grid: Application Experiences within NPACI

Chaired by Andrew Grimshaw, University of Virginia

A computational grid is a collection of heterogeneous and possibly geographically distributed resources that have been integrated to some degree into a single resource. Grid applications span a wide spectrum, including traditional HPC applications, high-throughput applications, and applications that use and integrate specialized resources at different sites. The NPACI grid effort (Metasystems thrust area) focuses on four projects: Legion, Globus, AppLeS, and Network Weather Service. The session began with a brief overview of grids, followed by a brief presentation on each project, focusing on application experiences and how applications can get the most out of the capabilities provided by the systems. For more information on each project,
see the following:

http://www.globus.org

http://legion.virginia.edu

http://apples.ucsd.edu

http://nws.npaci.edu/NWS/

Computational Biology: How NPACI Resources are Used by Biologists

Chaired by Russ Altman, Stanford University

Computational biology is emerging as a discipline of national interest and is particularly well supported by—and suited to the mission of—the NIH and various foundations. Within NPACI, tools have been developed for structural simulation, sequence and structural analysis, and to integrate biological resources. At this meeting NPACI scientists presented their current activities. The role of high-end computing in computational biology was discussed, with particular attention paid to the Bioinformatics Infrastructure for Large-scale Analysis Alpha project. Genomics was discussed at length, including various projects (MIA, MEME, MAST, PROFILES, 2HAPI, PDB) that involve molecular pattern recognition as a means of understanding the genome. Since NPACI bioinformatics tools are widely used, there was discussion of capitalizing on portal efforts to create a portal for bioinformatics, building on work in the Biology Workbench.

Mentoring Women and Minorities

Chaired by Anne Condon, University of British Columbia, & Ann Redelfs, SDSC

Students who are mentored feel more confident in their skills, bringing stability and direction to their career development. Mentoring can occur in a variety of forms, including one-time seminars, ongoing workshops, or one-on-one long-term relationships.

In this session, five panelists offered their advice and experiences mentoring people from underrepresented groups. Panelists included Andrew Bernat (University of Texas, El Paso), Sheila Castaneda (Clarke College), Anne Condon (University of British Columbia), J. Tilak Ratnanather (The Johns Hopkins University), and Richard Tapia (Rice University). Though mentoring is essential to student success, there is little recognition for excellence in mentoring, and few universities require that faculty who desire promotion or tenure mentor students. Consequently, many professors selectively mentor only the "best and the brightest" because they contribute more to research efforts. Those who assist students with the same potential but who need extra encouragement do so for purely personal rewards. Though all at this session were encouraged to become mentors (EOT has mentoring programs that can involve researchers), they were encouraged to enter into the arrangement thoughtfully: Mentoring involves exercise of power and responsibility, and poor mentoring can do real damage. Successful mentoring involves a "you can do it" attitude, offering respect, good listening skills, and the ability to emphasize accomplishments as well as strategies for overcoming problems.

Introduction to Education, Outreach, and Training

Presented by Roscoe Giles, Boston University; Greg Moses, University of Wisconsin, Madison; & Richard Tapia, Rice University

The sessions held on Tuesday, February 8, 2000, were devoted to issues in education and outreach and were run by the Education, Outreach, and Training Partnership for Advanced Computational Infrastructure (EOT-PACI) (**www.eot.org**). EOT-PACI is the joint organization formed of the education and outreach activities of NPACI and the National Computational Science Alliance, and is chaired by representatives from both partnerships (Giles, Moses, and Tapia). As an introduction to the day's events, the three chairs presented EOT-PACI as a success story of cross-partnership cooperation and collaboration, and encouraged the participants present to attend the sessions during the remainder of the All-hands Meeting. Researchers from both NPACI and the Alliance were present.

Advanced Networking-MSI Grant: Status and How to Get Involved

Presented by Allison Clark, NCSA, & Roscoe Giles, Boston University

EOT-PACI has received a subcontract for $1 million over four years from EDUCAUSE to participate in its Advanced Networking Project with Minority-Serving Institutions (AN-MSI) grant, funded by NSF. The grant helps minority-serving institutions of higher learning prepare for the next generation of information technology and computer networks. EOT-PACI envisions incorporating these institutions into the grid by connecting them to people, supercomputers, virtual environments, scientific instruments, educational tools, and large data sets via advanced high-speed networks. Participants at this session were encouraged to share curriculum, projects, and research with MSIs through the AN-MSI project. Ultimately, it is hoped that researchers and students at these schools will become involved in PACI science and technology activities.

Allison Clark has been visiting MSIs around the country, introducing them to the PACI program and learning their needs. Other grant activities will include providing Atlanta University with an Access Grid Node and training faculty and staff there to maintain and use it in distance learning. Training will serve as the basis for an envisioned Access Grid "cookbook" on how to use and access grid services. This publication will be helpful to other MSIs, especially once grid nodes become desktop-accessible. Atlanta will participate in the August 2000 Alliance Chatauqua remotely via its grid node.

Portals: What They Are and How They Can Support EOT Projects

Presented by Geoffrey Fox, Syracuse University; Reagan Moore, SDSC; & Roscoe Giles, Boston University

The EOT activities and products of NPACI and the Alliance lend themselves well to portal organization, and there are already EOT Web sites modeled on this concept (MICE, eTeach, Biology Workbench, RiverWeb, Digital Library projects, Sociology Workbench, Digital Insight, AMICO, etc.). The participants learned about portal architecture, including the mechanisms that allow portals to be customized and capture new knowledge as it is generated. Knowledge—data, visualizations, models, etc.—is represented as an "event stream" in the portal. A potential EOT proposal to the NSF Information Technology Research program will build an event-based model of learning and investigate how knowledge should be presented to students to inspire learning. Portal projects are an opportunity for interaction between EOT and other NPACI thrust areas, including Data-intensive Computing (DICE). Education portals may make information from the PACI projects more readily accessible for educational use and instruction. The long-term goal is to allow users to specify what type of information they want, how they want it analyzed, what they want it analyzed for, and then have an automated system find the information, process it, and deliver a final "knowledge product" to the user through the portal.

EdGrid: Status and How to Get Involved

Presented by Lisa Bievenue, NCSA

EOT-PACI, in collaboration with several colleges of education and the Illinois State Board of Education, has received a Preparing Tomorrow's Teachers to Use Technology (PT3) catalyst grant to host workshops, adapt technologies to support teacher education, and support collaborative group development of courses integrating modeling and visualization. The resulting project is called EdGrid and is intended to help pre-service teachers learn computing and technology-based methods for teaching math and science. Over the next two to three years—the grant is for at least two and up to three years—a testbed will be created to help EOT-PACI learn the needs of pre-service teachers with regards to computational science and technology. An evaluation will be conducted by the LEAD center to discern best practices for using technology in the classroom. Thematic workshops will be held, supported by the Access Grid nodes.

Two teams have been created within EdGrid. The first team will focus on certification processes at the state level and will involve Maryland Virtual High School, the University of Maryland, Mt. St. Mary's, Montgomery College, the University of Alabama at Birmingham, the ASPIRE program, and Lesley College. The second team will focus on working with administrators and faculty and will include the University of Illinois, Beloit College, and other universities. Because of emergency certification in California, the Cal State University system may become a partner in this effort. The participants at this session were asked to contact Lisa Bievenue if they would like to become involved with a workshop or on one of the two teams.

Guidelines on Self-evaluations for EOT Projects

Presented by Baine Alexander & Julie Foertsch, University of Wisconsin, Madison

Baine Alexander and Julie Foertsch of the Learning through Evaluation, Adaptation, and Dissemination (LEAD) Center conducted an interactive session on using self-evaluations to assess projects' progress and efficiently guide projects towards meeting goals. Evaluation of both strengths and weaknesses gives project

For PowerPoint slides, where available, see:
www.npaci.edu/ahm2000/presentations

leaders an opportunity to reflect on goals and the strategies being used to pursue them (formative evaluation), and facilitates the successful adaptation of programs and projects to other sites (based on summative evaluation). Evaluation results for the Rice University Spend a Summer with a Scientist program will help guide the replication of that program at the University of Wisconsin, Madison.

The eight steps of the self-evaluation process are: (1) establish the purpose of your evaluation; (2) articulate your project's overarching goals; (3) define the strategies that address the overarching goals; (4) define the metrics to measure progress towards goals (outcome metrics) and implementation of strategies (implementation measures); (5) decide what types of evaluation instruments will be used, with whom, and when (evaluation design); (6) construct evaluation instruments that use your metrics, (7) collect and analyze data; and (8) put the data to use (reporting findings summatively and using them formatively). Project leaders present in the audience offered example projects and were led through a mock self-evaluation.

Tracking Student Participants in PACI Programs

Presented by Jean Girves, University of Illinois, & Baine Alexander & Julie Foertsch, University of Wisconsin, Madison

Tracking student participation in education programs is important for facilitating cooperation from administrators, other students, faculty, and funding agencies. As the audience learned, each of these groups critiques programs based on specific outcome criteria that may be difficult to know in advance. Subsequently, a tracking program that accounts for many different variables—information about the students when they started, when they finished the program, and how they fared over a period of time after finishing—is desirable, and EOT-PACI will soon begin tracking student participants.

There are several questions that tracking should be able to answer: : Does the program work? What is the cost-benefit ratio? How much time and effort are required? How will the program benefit me? How do we compare with other schools? Will it make us look good? Will it benefit society?

Jean Girvin tracks more than 7,000 student participants of the Summer Research Opportunities Program (SROP) and has developed an Access database system for storing information, which she recommends. Based on her experience, she advises that a Web form be created for students to fill out, linked to the database. She also recommends that some sort of "hook" be attached to filling out the form, such as that the student can't participate in a related conference, etc., until data has been entered.

Student participants should be contacted once a year for several years following their participation to secure a long-range look at any benefit they received from program participation. They may become difficult to find after a time as known e-mail addresses and phone numbers expire. To mediate this, Girves recommends securing as much contact information as possible from participants at the beginning of the program. There are also programs that re-locate the students for a nominal fee.

SDSC K-12 Science Enrichment Programs: Ideals, Realities, and Plans

Presented by Rozeanne Steckler, Mike Bailey, & Cheryl Converse-Rath, SDSC

This session profiled some of SDSC's outreach efforts for girls and minorities. The activities presented were in addition to SDSC's "backbone" outreach efforts—tours, teacher mentoring, and partnerships. The original Science Scholars program, started by Steckler, donated desktop computers to girls for use in their homes, but the program was expensive and had a limited impact—only about 15 girls. New outreach efforts grew out of Science Scholars, however, including a partnership between SDSC, the San Diego City Schools District, and the Girl Scouts, which was identified for partnership because of their wide base in the San Diego schools. Thousands of youths from all ethnicities now participate in SDSC-led programs throughout San Diego.

Girls are GREAT (started in 1997) reaches 5,200 girls and has 15 paid staff participants. The staff work with girls on math and science at the girls' school sites, which are in inner city and border areas. SDSC is involved with curriculum development, equipment acquisition, Family Science Nights to get parental involvement, and Summer Day Camps designed to expose girls to the university campus environment.

Other girls in 7-12th grade at 10 schools throughout the county participate in the Girl Scout Science Interest Group. Teachers for this program receive training from SDSC volunteers on their own time, and the project is funded by SDSC and through additional fundraising. New SDSC EOT efforts include enVision for Kids—computational science curriculum for middle schools based on NPACI projects.

Getting Involved with the SC2000-2001 Education Program

Presented by Jeffrey Huskamp, East Carolina University

A proposal is being prepared for NSF, asking for funding to expand the education program at SC2000 and SC2001. The program committee hopes to launch a "Computational Science Leadership Program for High School Teachers," with NSF funding and participation of ACM, IEEE, the Alliance, NPACI, and others. The program would create two 100-teacher cadres, each organized into 25 teams. The program would be 18 months long and include a 5-day workshop at the SC conference followed by monthly seminars presented via VTC, a two-week summer institute, and an assessment of the impact that the program makes on the teachers' classroom practices and teaching. The program would be organized around computational science topics, including computational biochemistry, computational chemistry, computational physics, and environmental physics. The EOT-PACI participants attending this session were invited to get involved in organizing and running the program, as well as to contribute content for the computational science topics.

The EOT-PACI Web Site

Led by Raquell Holmes & Ed Boyce, Boston University

This session addressed redesigning the EOT-PACI Web site (**www.eot.org**). Redesign is critical, both for meeting the resource needs of educators and other EOT audiences, and for the impending NSF review of EOT-PACI. The current site is graphically intense, features stories and general information about EOT-PACI, has a search function, and provides lists of partner sites.

Participants at this BOF want the site to evolve into a portal, emphasizing comprehensive, user-friendly, and universally accessible information. Jennifer Matthews, Web and graphic designer at SDSC, presented a draft of the new site featuring an image jukebox of children and teachers in classroom settings, a consolidated list of links, and an area emphasizing classroom aides and lesson plans. Matthews will work with Ed Boyce and Jennie File of NCSA, who participated by phone, and the site will be ready for review by the EOT-PACI Leadership Team by April 2000.

Q&A on High-end Compute Resources

Led by Nancy Wilkins-Diehr, SDSC

This session, which was an extension of the technical session "High-end Compute Resources" that took place earlier that day (see page 6 for summary), provided an opportunity for questions and answers regarding NPACI architectures. Experts on NPACI allocable resources, such as Sharon Brunett of Caltech and Allan Snavely of SDSC, were available to address detailed questions from participants. The session was an informal lunch-time gathering, with participants split up into five groups, one for each of the five machines discussed at the "High-end Compute Resources" session: Blue Horizon (IBM teraflops), HP V2500, Cray SV1, Sun HPC 10000, and Tera MTA. Individuals could choose which group to join, and all discussions occurred concurrently.

HPC Software Exchange

Led by Jay Boisseau & Joakim Persson, SDSC

Throughout the high-performance computing community the production of software has not kept pace with advances in computing architectures. At the same time, as the computers themselves have become more complex, creating software for them has become a much more difficult task. Subsequently, there is a limited amount of software available. NPACI is now considering the state of software and deciding what actions are appropriate. The partnership's objective in this area is to become aware of software that is available, evaluate the software against the needs of NPACI users, and promote the use of the most suitable software. In part, this is accomplished by linking to repositories already in existence (for example, the National High Performance Software Exchange). There is little interest in creating an NPACI repository. The attendees at this meeting thought that NPACI should continue to work with existing repositories (such as NHPSE) to ensure that their links are accurate and active. Another suggestion from the audience was for NPACI to evaluate repositories and publish its evaluations, thereby providing users with some guidance about the software available through various repositories.

An attendee brought up cluster computing and said that it could be useful to have NPACI create "cluster in a box" software that could be used on clusters and migrated onto higher-capacity machines as the project matured. This idea will be taken into consideration.

Jay Boisseau pointed out that NPACI is also developing software through the thrust areas and Alpha projects. Typically, this software is evaluated just as vendors' products are, and "good" software is promoted and made available. He also said that NPACI would continue to use its influence with vendors to guide the creation of high-performance computing software within the commercial sector.

Scaling Successful EOT Programs

Led by Baine Alexander & Julie Foertsch, University of Wisconsin, & Cynthia Lanius, Rice University

Scaling successful programs, such as Spend a Summer with a Scientist at Rice University, will help EOT-PACI become a strong national program. However, scaling requires more than just adopting programs to a new location; it involves replicating and assimilating projects into current programs. Scaling faces other obstacles as well, including a predilection within academia for starting new programs rather than adopting successful ones already in existence.

Key factors for successful scaling include establishing the relevance and compatibility of the program to the needs of the adapters; establishing the quality and credibility of the program to be scaled through evaluation; creating a sense of community among the program adapters; providing opportunities for hands-on activities, group discussions, and workshops, where appropriate, within the new program; and providing interaction between the program developer and the adapter.

It is also important to consider such factors as demographics and location: Local programs often fail when they are scaled because they fail to adapt to the new context. To help ensure success, it is often beneficial to scale in cooperation with a larger, established program. For example, the Girl Scouts Science Interest Group at SDSC took advantage of its relationship with the Girls and Boys Scouts of America to grow into a county-wide program and will count on that relationship again if the program becomes national.

NPACI's Exhibit at SC99: The Good, Bad, and Ideas for SC2000

Led by Mike Gannis, SDSC

NPACI has had an exhibit at the annual SCXY conference—with posters and models, talks and demonstrations, and literature—since Sc97. This activity has been strongly encouraged by NSF but requires a significant amount of time and resources to plan and execute. All in attendance agreed that posters, one-on-one demos, and group presentations are valuable, and that organized HPC Challenge demos must be supported as high-profile, high-payoff events. Attendees also agreed that NPACI needs to continue to look for ways to more actively engage the partnership in SCXY planning and implementation. The participants advocated greater use of large rear-screen projection and/or flat-panel plasma displays. The expense of an exhibition booth and structure seems unavoidable. The SC2000

For PowerPoint slides, where available, see:
www.npaci.edu/ahm2000/presentations

booth may be downsized (to 20' x 30', or 30' x 30'), and should be coordinated with the Alliance. The exhibit in 2001 should be "memorable" to support the five-year NPACI grant renewal effort.

Using Local Campus Infrastructure vs. PACI Resources: Conflict or Synergy?

Led by Bob Sinkovits & Donald Frederick, SDSC

This session was offered to researchers as aid for transferring applications from desktop computers and campus servers to the high-end computing resources found within PACI. The focus eventually shifted, however, to developing a single (or multiple) campus mini-grid as a model for understanding the high-end computing grid. The question was raised whether PACI supercomputer resources would be needed in the future with the advent of cost-effective and easily accessible computer clusters. The discussion leaned toward using a well-developed campus mini-grid as an effective tool for sharing resources and as a preliminary step to moving to the PACI resources.

EOT Participation in the Alpha Projects

Led by Ann Redelfs, SDSC, & Greg Moses, University of Wisconsin, Madison

This session discussed the best methods for getting EOT involved in the Alpha projects. It was theorized that the biggest challenge facing this task was the definition of EOT within the Alpha projects. For example, some projects may consider working with graduate students to be education outreach. Reagan Moore, representing the Data-intensive Computing (DICE) thrust, agreed that this might be the case and listed several opportunities for EOT interaction with DICE projects. Martin Hadida, representing the Neuroscience thrust area and the Telescience for Advanced Tomography Applications Alpha project, agreed and said that there may be ways that EOT can work with that Alpha project. All in attendance agreed that one-on-one interactions between scientists and EOT members were necessary to accurately represent the goals of EOT and foster productive and meaningful collaborations. Meetings of this sort have begun and will continue.

Emerging Visualization Technologies: Java3D and VRML

Led by Dave Nadeau & John Moreland, SDSC

In this BOF, attendees learned about three emerging visualization technologies: VRML, X3D, and Java 3D. Dave Nadeau and John Moreland provided an overview of these technologies and their feature sets. They also covered the history and state of these technologies, including release schedules, quality of Beta releases, and vendor support.

Helping Partners Tune Their Networking Infrastructure

Led by Max Okumoto and Phil Papadopoulos, SDSC, & Matt Mathis, Pittsburgh Supercomputing Center

This session focused on techniques and methods for tuning network performance at NPACI partner sites and proactive plans to assist partners in their network-tuning efforts. Additional goals were to gather input on problems at remote sites and make technical contacts to address network difficulties in the future. Participants discussed building better tools for system administrators and how to put automatic controls in the end-system for users. System administrators should be "on-board mechanics," a philosophy drawn from mechanics that work on race cars while they're being driven. A collaboration was established between Tom Hacker of the University of Michigan and Max Okumoto. More information is available at NLANR On-Site (**www.ncne.nlanr.net/training/on-site/**), which describes a new training program for high-performance networking.

Poster Session

Posters and Demos

There were 27 posters (nine of them by students) and 16 demos (one of which included a student) that covered a wide range of topic areas; 15 of the 16 demos were computer-based.

Cluster Network I/O *[BEST STUDENT POSTER AWARD]*

Presented by Philip Buonadonna, UC Berkeley **[STUDENT]**

<philipb@cs.berkeley.edu>

Representing: Computer Systems; Resources; Metasystems

This poster presents our research efforts in investigating the various issues concerning cluster System Area Networks (SANs) and network-based I/O architectures. One of the challenges facing commodity-node distributed systems is high-performance, scalable communication. This study includes an examination of the low-level mechanisms and various high-level communication abstractions for distributed systems. One aspect of the study has focused on the Virtual Interface (VI) Architecture. VI is a proposed transport standard for SANs based on memory-to-network DMA controlled by user-level access to protected network resources. We have built a prototype VI Architecture system using the Myrinet programmable network interface and are using it to explore the distributed I/O space. Our first result is a distributed sorting application, Millennium Sort, that utilizes VI in conjunction with the River distributed I/O model. Millennium Sort achieved a Datamation sort benchmark world's record of 1.18 seconds on a cluster of 16 Intel Pentium-II PCs running Windows NT 4.0.

We have also built a port of Active Messages over VI and used it to examine the issues surrounding communication abstractions over this transport.

Our present work is centered around network-attached storage devices. Here, we are investigating various network-attached disk architectures and mechanisms to communicate with these devices. Future work will center on operating system abstractions for network-attached devices and the use of IP as the protocol for SANs.

Evaluation of the Tera MTA for Scientific Computing *[2ND PLACE STUDENT POSTER AWARD]*

Presented by Allan Snavely, SDSC **[STUDENT]**

<allans@sdsc.edu>

Representing: Computer Engineering (Hardware Multithreading); Resources

This poster provides an update on the ongoing evaluation of the Tera MTA (MultiThreaded Architecture) for scientific computing in the form of performance results from benchmarks and applications, along with some results from other supercomputers for reference.

NetSolve and IPARS

Presented by Dorian Arnold, University of Tennessee

<darnold@cs.utk.edu>

Representing: Metasystems; Programming Tools and Environments; Earth Systems Science

This demo uses NetSolve to interface a sub-surface fluid reservoir simulator, a repeat of our demo at SC99. Our demonstration emphasizes remote access and ease of computing and simulation that the NetSolve system provides. We have implemented an interface to IPARS via NetSolve that allows users to run complex simulations without having the software or visualization tools necessary to view the output. We have built a Web interface on top of this NetSolve interface to enable remote access. The graphical output is also rendered back to the user via the browser.

The KeLP Data Mover: A Platform-independent Communication Abstraction

Presented by Scott Baden, UC San Diego

<baden@cs.ucsd.edu>

Representing: Programming Tools and Environments

This poster presents the KeLP Data Mover, an abstraction for expressing machine-independent customized communication algorithms in a variety of applications where the data are organized into collections of structured blocks of data. The Mover supports multi-block and multi-level adaptive methods, and enables its user to express data motion using intuitive geometric operations that compose collective patterns restricted to movement of rectangular array sections. The Mover encapsulates the low-level details of the underlying communication substrate, providing a machine-independent interface. A prototype of the next-generation Mover employs a proxy to support latency-tolerant algorithm design and will be implemented on the NPACI teraflops platform. The changes to the Mover will be internal only, requiring no change in the user's interface.

Our Vision for the New Protein Data Bank

Presented by John Badger, SDSC

<badger@sdsc.edu>

Representing: Macromolecular Structure; Molecular Science

This poster describes the activities of the Research Collaboratory for Structural Bioinformatics (RCSB) in the operation of the PDB (**http://www.rcsb.org/pdb/**). The RCSB's member groups are SDSC, Rutgers University, and NIST. Data deposition, distribution, and query are highlighted on this poster.

Center for Visualization Prototypes

Presented by Mike Bailey, SDSC

<mjb@sdsc.edu>

Representing: Scientific Visualization; Interaction Environments

The SDSC Center for Visualization Prototypes (**cvp.sdsc.edu/**) has been researching a unique approach to visualization hardcopy—the fabrication of solid 3-D models. The project runs two manufacturing machines, a Helisys LOM 1015 and a Z Corporation Z402. Users can gain access to these machines by submitting geometry files over the Web. This project seeks more collaboration with NPACI researchers. If you have a need for this type of complex data visualization, please contact the CVP.

The Institute on Learning Technology, a Resource for SMET Faculty

Presented by Jean-Pierre Bayard, California State University, Sacramento, & National Institute for Science and Engineering (NISE)

<bayardj@csus.edu>

Representing: Education, Outreach, and Training

The Institute on Learning Technology provides a resource for science, mathematics, engineering, and technology (SMET) instructors seeking to engage in activities fostering inquiry-based learning through the use of computer-based technology. The institute's Web site features exemplary practices implemented at both the classroom and institutional levels, promoting higher levels of cognitive development (based on Bloom's taxonomy) and located at diverse types of educational institutions.

Computational Anatomy: Computing the Geodesic between Two Images

Presented by M. Faisal Beg, The Johns Hopkins University **[STUDENT]**

<mfbeg@cis.jhu.edu>

Representing: Biomedical Engineering; Engineering

We have derived and implemented an algorithm for deformable template matching of two images. The template matching is done by constructing high-dimensional diffeomorphisms (invertible maps) as flows. These flows result while minimizing a running cost on the velocity of the flow field as well as that associated with the difference between images being analyzed in a variational setting. The construction of such diffeomorphisms allows metrics to be calculated in comparing large volumetric image data. A gradient method is used to solve the nonlinear variational problem. Clinical applications are targeted at diagnosing neuropsychiatric disorders such as Alzheimer's disease, schizophrenia, and epilepsy via images of the hippocampus of the brain, which is responsible for memory function.

Pbody: A Parallel Adaptive N-Body Library

Presented by David Blackston, UC Berkeley **[STUDENT]**

<davidb@cs.berkeley.edu>

Representing: Computer Science; Parallel Algorithms; Programming Tools and Environments

The N-Body problem is a fundamental problem from computational physics with applications in many areas. There are many algorithms that solve this problem to any desired accuracy in a theoretically optimal fashion. Many have similar structures, and the Pbody library presented here is an attempt to unify these algorithms into a single computational structure. This library contains three of the better known N-Body algorithms (Barnes-Hut, Fast Multipole, and Anderson) and exhibits good parallel performance on a variety of platforms.

The Internet Engineering Curriculum Repository

Presented by Theresa Boisseau, SDSC & CAIDA

<theresa@sdsc.edu>

Representing: Education, Outreach, and Training; Network Engineering

This poster highlights initiatives by CAIDA (Cooperative Association for Internet Data Analysis) to make up-to-date Internet engineering curricula accessible to university faculty. Our demonstration involves a computer with a Web browser to allow people to see the Web site (**www.caida.org/iec/**) first hand and search the CD-ROM.

Calculations, Visualization, and Education: A Computational Science Infrastructure

Presented by Laura Brovold, SDSC & UC San Diego **[STUDENT]**

<brovold@sdsc.edu>

Representing: Computational Science; Interaction Environments; Education, Outreach, and Training

In this poster, we demonstrate our computational codes, how they integrate with our visualization codes, and how easy they are to use as an educational tool. Project examples are included.

Cichlid: 3-D Visualization of Remote Data Sets

Presented by Jeff Brown, SDSC & NLANR **[STUDENT]**

<jabrown@nlanr.net>

Representing: Networking; Interaction Environments

Cichlid is a software toolkit from the National Laboratory for Applied Network Research (NLANR) Measurement and Analysis team at SDSC that facilitates real-time visualization of remote data sets. Users can view, explore, and interact with high-quality 3-D data sets as the information changes with time. Cichlid produces several types of data visualizations, including graphs constructed of vertex and edge sets, 3-D bar charts with multiple "stacks" of data on each bar, and interpolated NURBS surfaces.

Cichlid features real-time data display, point-and-click user feedback, and dynamic data coloring and labeling. Sequences of displayed frames can be captured for later encoding as an animation, and single frames can be captured or re-rendered at arbitrary resolutions that exceed the available memory of the client computer. The source code to Cichlid is freely available and can be used on standard PC hardware and high-end workstations. It is written in C and based on free versions of the OpenGL and GLUT graphics libraries. Cichlid is being used on FreeBSD, Linux, Irix, and Win32 platforms.

To simplify development, Cichlid provides an API on the server side that represents the data set as an object and gives the developer the means to modify it, so the analysis code does not need to handle socket communication or graphics libraries. The Cichlid libraries take care of all data transport concerns, mirroring the objects to any clients that request them. In fact, Cichlid servers compile and run on systems without any other graphics libraries, such as X Windows systems.

Large Hydrogen (Antihydrogen) Formation Cross Sections

Presented by David Caballero, California State University, Long Beach **[STUDENT]**

<Astrokat@Access1.net>

Representing: Physics

Large hydrogen(antihydrogen) formation cross sections are found in the entry channel ps(n=2) + p in the multichannel, three-body collision system involving electron, positron, and proton(antiproton). Here we present the 3-D graphical representation of the 3-body interior channel wave functions, which points to the source of these large formation cross sections.

Interactive Web Database for 3-D Visualization of Gene Expression

Presented by James Carson, Baylor College of Medicine **[STUDENT]**

<jc691824@bcm.tmc.edu>

Representing: Computational Biology

The aim of this project is to develop procedures that allow rapid determination of gene expression patterns in the mouse brain by *in situ* hybridization (ISH) and build a model Web database with this information. This project concentrates on the brain, a complex organ whose detailed functions are still poorly understood. Our long-term goal is to determine the spatial expression patterns in the mouse brain of several thousand genes encoding signaling molecules, their receptors, and enzymes involved in signal transmission processes. The technology and computational infrastructure developed under this project eventually will allow mapping of the expression pattern of thousands of mouse genes within a time frame of a few years. Preliminary analysis using this approach begins with ISH of *mper1*, a gene involved in the control of circadian rhythm, performed on 7mm thick slices of brain from a three-day postnatal mouse. This series of 2-D data were used to create 3-D images of the brain anatomy and gene expression.

MCell/AppLeS/NetSolve

Presented by Henri Casanova, UC San Diego
<casanova@cs.ucsd.edu>

Representing: Metasystems; Neuroscience

MCell is a Neuroscience application that performs Monte Carlo simulations at the molecular level inside living cells. This project aims to deploy MCell in grid environments, with an emphasis on scheduling issues with respect to computation and storage requirements.

Early Results on the New SDSC IBM SP

Presented by Giridhar Chukkapalli, SDSC
<giri@sdsc.edu>

Representing: Scientific Computing; Resources

We present early results from running benchmark codes and full applications on the new IBM SP "teraflops" system at SDSC. These results include discussion of performance issues (including comparisons to other systems), programming models (including advantages and disadvantages of each), and portability of applications. We emphasize that the system is new and the results available by the All-hands Meeting are difficult to predict. However, we—and some of our users—have already run large-scale simulations to conduct leading-edge science.

Market-based Proportional Resource Sharing for Clusters

Presented by Brent Chun, UC Berkeley **[STUDENT]**
<bnc@cs.berkeley.edu>

Representing: Computer Science; Resources; Metasystems

Enabling technologies in high-speed communication and global process scheduling have pushed clusters of computers into the mainstream as general-purpose, high-performance computing systems. More generality, however, implies more sharing, and this raises questions in the area of cluster resource management. In particular, in systems where aggregate demand for computing resources can exceed the aggregate supply, how to allocate resources among competing applications is an important problem. Traditional solutions to this problem have focused mainly on global optimization with respect to system-centric performance metrics that ignore higher-level user intent.

We propose an alternative market-based approach based on the notion of a computational economy that optimizes for user value. This poster presents an abstract architecture for market-based cluster resource management systems and highlights our implementation of a computational economy for CPU time on a 32-node cluster of two-way SMPs at UC Berkeley.

RiverWeb℠: Building the Digital River Basin

Presented by David Curtis, NCSA
<dcurtis@ncsa.uiuc.edu>

Representing: Education, Outreach, and Training; Earth Systems Science

We demonstrate up to three recent prototypical applications developed as part of the NCSA-initiated RiverWeb Program. Bridging informal and formal education, RiverWeb harnesses advanced information technologies with potential to empower greater public engagement in science-based watershed management and policy. The applications we show are the following:

- The River Pilot Simulator: A museum-based, river tow barge navigation experience (NSF-funded Mississippi RiverWeb Museum Consortium).
- The Archetypal River Basin: A Web-based learning environment for high school science inquiry into water quality (EOT-PACI project with Maryland Virtual High School).
- The RiverWeb American Bottom Landing Site: A digital college and high school resource for teaching and learning in the humanities (NEH-funded collaboration between the Illinois State Museum, University of Illinois History Department, and NCSA).

Application-level Scheduling on the Computational Grid

Presented by Holly Dail and Otto Sievert, UC San Diego **[STUDENTS]**
<hdail@cs.ucsd.edu>
<otto@cs.ucsd.edu>

Representing: Grid Computing; Metasystems

Application-level Scheduling agents (AppLeS) provide a mechanism for adaptively and dynamically scheduling individual applications on distributed heterogeneous systems. This poster describes AppLeS methodology and gives an overview of AppLeS software.

Turbulence Studies in Electron Magnetohydrodynamics

Presented by Amita Das, UC San Diego
<amita@physics.ucsd.edu>

Representing: Physics

Electron Magnetohydrodynamics (EMHD) is the theory of magnetized electron fluid in the presence of self-consistent and external electric and magnetic fields. This fluid model is applicable to a variety of plasma phenomena occurring at fast electron time scales (e.g., lying between electron and ion gyrofrequencies). The EMHD model differs from other models (neutral fluid, MHD, etc.) on which turbulence studies have been carried out as it supports dispersive normal modes known as "Whistlers." Furthermore, in EMHD, there is an intrinsic length scale, namely the electron skin depth, above and below which the character of the equation changes.

We present results of our numerical studies wherein an understanding of how these features influence the turbulent EMHD state in terms of (i) diffusion of magnetic field lines in the presence of background turbulence and (ii) quantitative assessment of the wave vs. eddy nature of turbulence and its influence on spectral cascade.

Parallel *hp*-BE Modeling of the Human Auditory System

Presented by Leszek Demkowicz et al., University of Texas, Austin
<leszek@ticam.utexas.edu>

Representing: Engineering; Programming Tools and Environments

Our objectives are the following:

1. Develop a generic, parallel *hp*-BE code for acoustical simulations.
2. Apply the methodology to the simulation of various features of the human auditory system, thus providing an alternative to laborious experimental techniques.

For acoustical scattering problems on unbounded domains, the governing partial differential equation (Helmholtz) and boundary conditions are converted into the equivalent Burton-Miller boundary integral equation. This reduces the need for a volume grid (3-D case) as well as problems associated with unbounded domains. Discretization of this integral equation leads to dense, complex-valued linear systems.

The method has been implemented on distributed-memory machines using the PLAPACK (Parallel Linear Algebra Package) software. The parallel formation of the linear system is accomplished using the Application

Interface (API) mode, and the parallel solve is accomplished using one of the many parallel dense solvers provided by PLAPACK.

hp-Adaptive Finite-element Methods for Electromagnetics

Presented by Leszek Demkowicz et al., University of Texas, Austin

<leszek@ticam.utexas.edu>

Representing: Engineering; Programming Tools and Environments

Our objectives are the following:

1. Develop generic, parallel, 2- and 3-D *hp*-adaptive finite-element simulators for time-harmonic Maxwell's equations.
2. Apply the methodology to simulations of electromagnetic waves in the human head.

In *hp* finite-element methods, element size *h* and order of approximation *p* are varied locally to produce optimal meshes that deliver high accuracy at a minimum cost. *Geometrically graded meshes,* with *h* and *p* decreasing towards an incoming corner (2-D and 3-D) or edge (3-D) are used to capture singularities in the solution. Large, high *p* elements are used to minimize the *pollution error* in simulations of wave propagation problems, whereas small, low *p* elements are necessary to capture geometrical details. Only *hp* methods combine both in the same mesh.

Mathematically speaking, the *hp* method is the only approach that delivers *exponential convergence rates* (error in terms of number of degrees-of-freedom) for regular and irregular solutions.

While *hp* methods for elliptic problems have been known for more than two decades, we were the first to apply them to electromagnetics (*hp* edge elements) in 1997. Since then, through NPACI, four *hp* codes have been written in Fortran 90 and placed on the Web:

- 2Dhp90: a 2-D code (triangles, quads) for elliptic problems
- 2Dhp90_EM: a customized version of 2Dhp90 for Maxwell's equations
- 3Dhp90: a 3-D code (hexahedra) for elliptic problems
- 3Dhp90_EM: a customized version of 3Dhp90 for Maxwell's equations

Flying through the Orion Nebula: A 3-D Scientific Visualization Sequence for the Hayden Planetarium

Presented by Jon Genetti, SDSC

<genetti@sdsc.edu>

Representing: Scientific Visualization; Interaction Environments

The American Museum of Natural History's Rose Center for Earth and Space, a $210-million, seven-story exhibition and research facility opening February 19, 2000, includes the spectacular, all-new Hayden Planetarium. The 87-foot sphere will hold the world's largest virtual reality simulator—audiences will seem to fly through interstellar space surrounded by the stars and nebulas of the Milky Way galaxy, thanks to a scientifically accurate, high-resolution animation displayed by a seven-projector, computer-controlled image system. Visualizations made possible for the first time by NPACI's MPIRE volume-rendering software and the teraflops IBM SP and Tera MTA supercomputers at SDSC have given these images unprecedented realism. Our demonstration features a light box containing a 36' x 24' 3-D image of the heart of the nebula.

Telescience for Advanced Tomography Applications

Presented by Martin Hadida & Mark Ellisman, SDSC & NCMIR, UC San Diego

<marty@sdsc.edu>

<mhellisman@ucsd.edu>

Representing: Neuroscience

The Telescience Alpha project is developing an end-to-end, richly integrated system to enable telemicroscopy on biological specimens. This system will integrate use of remote-imaging instrumentation, distributed heterogenous parallel computing, federated and distributed databases and image archives, and component-based remote visualization tools. Telescience applications represent a new "closed-loop" paradigm for imaging systems wherein the process of data acquisition can be improved, in more or less real time, by feedback from the analysis. Part of our demo involves showcasing our "trans-Pacific telemicroscopy over IPv6" work at the National Center for Microscopy and Imaging Research (NCMIR) led by Mark Ellisman.

MPIRE Galaxy Renderer: Operation and Performance

Presented by Greg Johnson, SDSC

<johnson@sdsc.edu>

Representing: Scientific Visualization; Interaction Environments

SDSC and the American Museum of Natural History's Hayden Planetarium are working together to generate scientifically accurate, high-resolution imagery of stellar phenomena. SDSC researchers have developed a system called MPIRE (Massively Parallel Interactive Rendering Environment) that uses the CPU and memory capacity of high-performance computers to dynamically render images from multi-gigabyte 3-D data sets. A recent addition to MPIRE, called the Galaxy Renderer, employs a perspective viewing model to emphasize the size and relative position of features in gaseous nebulas and other extended objects. The MPIRE Galaxy Renderer has been ported to several shared-memory HPC systems, including the IBM Power3 SP, Sun HPC 10000, and Tera MTA. This poster describes the MPIRE Galaxy Renderer.

Computational Physics Educational Web Demonstrations

Presented by Rubin Landau, Oregon State University

<rubin@physics.orst.edu>

Representing: Education, Outreach, and Training; Computational Science; Computational Physics

We have new demonstrations that show how high-performance computing can be used to enhance a class and teach both physical concepts and computing. The demonstrations span various levels of physics and are designed to provide a glimpse into a physical world that is not otherwise visible.

Development of a 3-D Cell-centered Database

Presented by Maryann Martone, UC San Diego

<mmartone@ucsd.edu>

Representing: Neuroscience

We present a prototype for a Web-based database for 3-D data on neurons, their processes, subcellular organelles, and protein constituents. Data are derived using both light and electron microscopy at the National Center for Microscopy and Imaging Research. During the course of analysis, many types of image files are produced, including 2-D raw data, volume reconstructions, segmented data, quantitative data on neuronal morphology, and qualitative data on the localization of key proteins. To reduce the amount of manual data entry, the information within data files is entered automatically into the database via a utility, "ispider" (Image Search/Probe/Index Database Entry Robot), which automatically (overnight from a UNIX "cron" job or at

direct user request) traverses specified file systems and catalogs images found on them. A simple configuration language allows the maintainer or individual user to add new image types without rebuilding the program or include more detailed information than can be extracted directly from the image. The database is built on two main technologies, Web/Internet and Hughes Technologies' non-commercial MiniSQL database package.

Evaluating Titanium on the Tera MTA

Presented by Chang-Sunn Lin, Jr., UC Berkeley **[STUDENT]**
<cjlin@cs.berkeley.edu>

Representing: Programming Tools and Environments; Resources

We present the results of implementing the Titanium SPMD language on the Tera Multi-Threaded Architecture (MTA).

Repository in a Box (RIB)

Presented by Terry Moore, University of Tennessee, Knoxville
<tmoore@cs.utk.edu >

Representing: Programming Tools and Environments; Metasystems

Repository in a Box (RIB) is a toolkit to set up and maintain a repository of software and tools, in particular software and tools for high-end computing. RIB has become widely used by NASA, NSF, and DoD high-end computing centers for sharing and reusing high-end software, including use by the the Alliance enabling technologies parallel and distributed computing teams. A collaboration between the RIB and Globus teams is inputting the software deployment information maintained by RIB into the Globus Metacomputing Directory Service so that Globus users can make decisions about what machines to use based on installed software. This demo illustrates how RIB is being used to provide information to Alliance and Globus users about software deployed on Alliance platforms.

eTEACH: A Tool for Experimentation with Streaming Video in the Classroom

Presented by Gregory Moses, University of Wisconsin, Madison
<moses@engr.wisc.edu>

Representing: Education, Outreach, and Training

eTEACH and its associated authoring tool are demonstrated, and projects are discussed. The underlying design philosophy of eTEACH is presented.

Materials' Properties Calculations in the Engineering Thrust Area

Presented by Richard Muller, Caltech
<rpm@wag.caltech.edu>

Representing: Engineering; Chemistry

We detail the methods and the scaling of computational chemistry and materials science computations performed to compute materials' properties of interest in the Engineering thrust.

Mathematical Software for High-performance Computing and Some Applications

Presented by Almerico Murli, CCPS-CNR **[INTERNATIONAL AFFILIATE]**
<murli@matna2.dma.unina.it>

Representing: High-performance Mathematical Software; Programming Tools and Environments

We present some of the results of research activities carried out at the Center for Research on Parallel Computation and Supercomputers (CPS-CNR), Naples, Italy, whose aim is the development of mathematical software for high-end computing and its applications in relevant scientific areas. Among basic research activities at CPS—consisting of the development of numerical methods, algorithms, and software to be used as building blocks in the solution of application problems—we focus on those concerning *eigenproblems*, multidimensional quadrature, quadratic programming, fast transforms, and PDEs. In particular, basic parallel algorithm design issues and some performance results are presented.

CCB NeuroSys Database Project

Presented by Sandy Pittendrigh, Montana State University
<sandy@nervana.montana.edu>

Representing: Neuroscience

The poster describes the Center for Computational Biology's Database Software System, which pipes database query results to plug-in software tools on local and/or remote machines.

Computational Characterization of Heterogeneous Materials

Presented by Gregory Rodin, University of Texas, Austin
<gjr@ticam.utexas.edu>

Representing: Engineering

The poster describes our progress in integrated computational characterization of heterogeneous materials. Our approach combines CT data acquisition, fast O(N) numerical algorithms implemented in the code FLEMS, and parallel computing techniques based on the SDDA and Globus environments. Our objective is to bring these tools to mechanical engineering, material science, earth science, and other communities.

Parallel Tools at NCSA

Presented by Faisal Saied, NCSA
<fsaied@ncsa.uiuc.edu>

Representing: Scientific Computing; Programming Tools and Environments

This poster highlights parallel tools at NCSA available to the user community to enhance their productivity. These tools come from a variety of sources, such as NCSA Alliance partners, vendors, and the academic research community. This is an area where there is potential for fruitful NPACI/Alliance collaboration.

Scalable Interrogative Visualization

Presented by Ariel Shamir, University of Texas, Austin
<arik@ticam.utexas.edu>

Representing: Scientific Visualization; Interaction Environments

Numerical simulations running on supercomputing platforms and measurement techniques such as electromagnetic imaging can generate extremely large amounts of data. These data must be interpreted effectively to explain the underlying physical phenomena.

This project develops a suite of compressed multi-resolution representations and data-streaming techniques that adapt, in an error-controlled manner, to available computational resources. Used in conjunction with fully threaded visualization servers and clients for best parallel performance and I/O optimizations for out-of-core computations at both server and client ends, we attempt to provide seamless scalable performance. This project is concerned with defining and building client- and server-end data structures that interact in a query-based environment to support multi-resolution compression and out-of-core algorithms. This enables the effective use and visualization of large amounts of data on platforms ranging from high-end graphic displays to simple desktop or laptop PCs.

Estimating the Global Ocean Circulation

Presented by Detlef Stammer, Scripps Institution of Oceanography, UC San Diego

<dstammer@ucsd.edu>

Representing: Physical Oceanography; Earth Systems Sciences

Given the turbulent nature of the ocean flow field and the disparate techniques used to observe it, the only feasible approach to obtain a dynamically consistent picture of the changing flow field and related transport processes (among other things) is to combine the global data with a general circulation model (GCM) to produce the required estimates.

To bring ocean state estimation from its ongoing experiments phase into a semi-operational form suitable to support anticipated global experiments such as CLIVAR and GODAE, a NOPP node on "Estimation of the Circulation and Climate of the Ocean" (ECCO) was formed recently as a consortium between JPL, MIT, and SIO. The goals of ECCO are summarized here, based on preliminary results of ocean state estimations that are ongoing on the global scale and over the six years' time frame (1992–1997). The estimate is obtained by constraining the MIT global ocean general circulation model by the absolute and time-varying TOPEX/POSEIDON and ERS-1 and ERS-2 altimeter data relative to the EGM96 geoid model, monthly sea surface temperature fields, time-varying NCEP surface fluxes, and the Levitus monthly mean hydrography. The model is forced to consistency with those fields by using the model's adjoint to modify the initial temperature and salinity conditions over the full water column and adjust the time-varying meteorological forcing fields over the full estimation period.

Using the Biology Workbench as a Portal for Undergraduate Molecular Investigations

Presented by Ethel D. Stanley, Sam Donovan, BioQUEST Curriculum Consortium, Beloit College; Mark Trebian, Lac Courte Oreilles Ojibwa Community College; Kris Engelsen, Bertram C. Bruce, Umesh Thakkar, Scott Lathrop, Eric Jakobsson, NCSA Computational Biology Group, UIUC

<stanleye@beloit.edu>

<donovans@beloit.edu>

<mtrebian@lco-college.edu>

<engelsen@students.uiuc.edu>

<chip@uiuc.edu>

<uthakkar@ncsa.uiuc.edu>

<scott@ncsa.uiuc.edu>

<jake@ncsa.uiuc.edu>

Representing: Education, Outreach, and Training; High-end Computing; Biology; Molecular Science

A workshop featuring case-based investigations of molecular sequence data to detect genetic disorders, understand genetic fingerprinting, and recognize the power and limitations of molecular genetics will be held in early April at the American Indian Higher Education Consortium 19th annual conference in Albuquerque. This workshop is the result of a successful collaboration among three EOT-PACI (NPACI) partners, BioQUEST, AIHEC, and the NCSA Computational Biology group; it is the first of six such outreach workshops with other user communities scheduled this year.

Bioinformatics (sequence comparison and alignment) is changing not only how biologists do their research, but also the questions they ask. It is now possible to explore relationships among different life forms in a new ways, i.e., by building biological meaning from molecular sequence data. Students should use the tools of bioinformatics to investigate a wide range of biological concepts, but because these approaches to biological problem solving are so new, both students and teachers need access to rich data sources, powerful analysis tools, and concrete biological questions that will drive their investigations. One of the most integrated resources for teaching and learning bioinformatics is the Biology Workbench, a Web-based suite of tools for analyzing and visualizing molecular data.

Computational Science Imperative in Undergraduate Curricula: Working with the Faculty, within the System, by Example, and Supporting Change

Presented by Kris Stewart, San Diego State University

<stewart@sdsu.edu>

Representing: Education, Outreach, and Training

The NPACI Education Center on Computational Science & Engineering at SDSU has focused on promoting advanced computational tools in undergraduate curricula. However, NPACI developers are largely unaware of our clientele, and software tools are often developed as a by-product of narrowly focused research, while undergraduate faculty seldom have time to explore tools that are not going to be ubiquitously useful. We found that both groups could be brought together and their needs and expectations matched through a series of strategies involving building faculty interest groups, developing "light" or more universally appealing versions of software tools, focusing on regular assessment of learning outcomes, etc.

The Ed Center's recent activities—including the Faculty Fellows program, development of the Sociology Workbench, faculty workshops, etc.—were other steps to build an infrastructure for computational science in undergraduate curricula. This poster describes the Faculty Fellows program and demonstrates educational accomplishments of the two generations of fellows. We also discuss the results of the Ed Center's formative evaluation by the LEAD Center and instruments that can be used in the assessment of learning outcomes. One of these tools, the Sociology Workbench developed at the Ed Center by Ilya Zaslavsky (now of SDSC), is demonstrated during the session. SWB, an on-line survey data analysis system, has been used to analyze student and faculty surveys, and solicit feedback from previous attendees of the NPACI All-hands Meeting.

Visualization of Astronomy Data with ADR and MPIRE

Presented by Alan Sussman, University of Maryland

<als@cs.umd.edu>

Representing: Parallel Tools and Environments

We demonstrate volume rendering of a large 3-D astronomy data set using a parallel algorithm based on the one in SDSC's MPIRE scientific visualization system, implemented using the Maryland/Hopkins Active Data Repository (ADR) infrastructure to extend the algorithm to out-of-core data sets. The client for viewing the data and generating requests will run on a local workstation. The ADR data server runs on a Linux cluster at the University of Maryland, sending results over the Internet to the client.

Data Mining and Knowledge Discovery Tools: Rough Sets Methods

Presented by Roman Swiniarski, San Diego State University

<rswiniar@sciences.sdsu.edu>

Representing: Data Mining; Knowledge Discovery; Intelligent Systems; Data-intensive Computing; Resources

This poster includes a description of theory, techniques, and corresponding software systems for data mining and knowledge discovery of uncertain data. The presented technique is based on rough sets theory. Presented applications include image recognition.

NPACI HotPage: A Framework for Scientific Computing Portals

Presented by Mary Thomas, SDSC
<mthomas@sdsc.edu>

Representing: Scientific Computing; Resources

The NPACI HotPage provides authenticated access to NPACI resources via a Web interface. We present recent development efforts to create computing portals for projects, such as the pharmacokinetic modeling project, and what is required to port a computational problem to the Web. We also show users how to use the system and help them get accounts.

High-performance Computing in Indian Country

Presented by Mark Trebian, Lac Courte Oreilles Ojibwa Community College
<mtrebian@lco-college.edu>

Representing: Education, Outreach, and Training

The American Indian Higher Education Consortium (AIHEC) has been an Alliance partner since 1998. Since that time, member institutions have begun work on various high-performance computing projects and development of Information Technology curricula. A High-performance Computing Initiative is being developed that seeks to bring the technological advancements of Alliance work related to the grid to Indian country.

Processing Distributed Large Earth Science Data Using NPACI Infrastructure

Presented by Zengyan Zhang, University of Maryland
<zengyan@umiacs.umd.edu>

Representing: Data-intensive Computing; Earth Systems Science

We demonstrate how to extract land cover information from large amounts of remotely sensed data that reside at the University of Maryland and SDSC. Users submit requests through a simple Web interface to a server, which then coordinates with the SDSC SRB to locate the data and retrieve it. Any necessary processing (such as compositing, gridding with the selected projection, and atmospheric correction) is then performed as specified by the user.

The UMD Global Land Cover Facility (GLCF): An NPACI Digital Library Resource

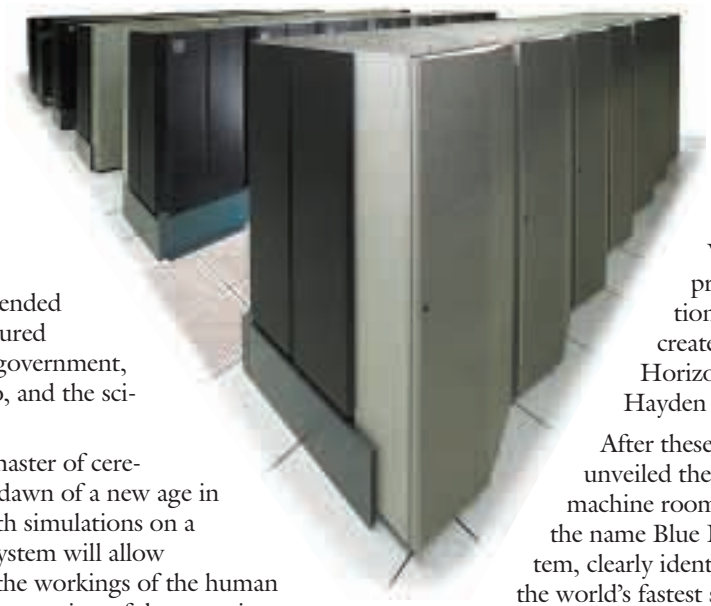
Presented by Zengyan Zhang, University of Maryland
<zengyan@umiacs.umd.edu>

Representing: Data-intensive Computing; Earth Systems Science

We demonstrate search, browsing, and retrieval tools on the University of Maryland Earth Systems Science digital library (GLCF) that holds more than 1TB of multi-sensor data, including Landsat (TM, MSS), AVHRR, land cover maps (1-km resolution), health of the coastal marshes of the eastern U.S., etc. Most of the GLCF data are not available anywhere else. Our demonstration features workspace architecture, multi-resolution browsing, processing and subsetting of large raster images, and online custom tailored product generation from large, remotely-sensed data.

Dedication Ceremony for the NPACI Teraflops IBM SP Supercomputer at SDSC

NPACI officially named its new 1,152-processor IBM RS/6000 SP system "Blue Horizon" at a dedication ceremony in the SDSC auditorium on Wednesday, February 9. Blue Horizon has a peak speed of 1.02 teraflops generated by 1,152 Power3 processors running at 222 MHz. The machine's total system memory is 576 GB, and its associated disk can store 5.1 terabytes of data. Test results show that the new machine will provide the capability to solve problems in hours or days that typically require weeks, months, or years on smaller machines. It is the most powerful machine in North America available for academic (non-classified) research and is ranked tenth in the world on the list of Top 500 Supercomputer Sites maintained by the University of Tennessee and the University of Mannheim.



The dedication ceremony was attended by more than 150 people and featured remarks by representatives from government, IBM, NPACI and UC San Diego, and the scientific research community.

NPACI Director Sid Karin was master of ceremonies. "The name suggests the dawn of a new age in scientific discovery," he said. "With simulations on a scale never before possible, this system will allow researchers to better understand the workings of the human nervous system, design the next generation of drugs against HIV and other diseases, and tackle complex issues of climate and the environment."

UCSD Chancellor Bob Dynes congratulated Karin and remarked upon the importance of data storage as well as computational power.

Jim Ritchie, an aide to Congressional Member Randy "Duke" Cunningham, read a congratulatory letter that endorsed the importance of computational science and computer technology in solving the nation's problems. UC San Diego is in Cunningham's district (CA-51).

Bob Borchers and Rich Hirsh, Director and Deputy Director respectively of the NSF's Advanced Computational Infrastructure and Research (ACIR) Division in the Directorate for Computer and Information Science and Engineering (CISE) that sponsors NPACI, spoke via teleconference from Reston, Virginia. They noted a significant increase in the CISE budget for the coming year and mentioned two upcoming NSF program opportunities for terascale systems.

IBM Vice President for Strategic Alliances and Pervasive Computing Lou Bifano said, "Our partnership with NPACI provides further evidence of our commitment to provide solutions for problems of global scale. We are in this for the long haul."

Susan Graham, NPACI's chief computer scientist and professor of Computer Science at UC Berkeley, noted that she was pleased for three reasons that the partnership had acquired Blue Horizon—because it fulfills the recommendations of the President's Information Technology Advisory Committee (on which she serves), because it will be of great benefit to NPACI researchers, and because she will get to use it. "Working on the problem is as much fun as finding the answers," she said.

J. Andrew McCammon, Joseph E. Mayer Professor of Theoretical Chemistry at UCSD, noted his 20 years of history in the use of supercomputers for computational biochemistry and structure-based drug discovery.

Bernard Pailthorpe, director of the Advanced Scientific Visualization Laboratory at SDSC, presented "a wonderful collaboration"—a three-minute video sequence created on 800 processors of Blue Horizon by SDSC staff for the new Hayden Planetarium in New York.

After these remarks, Karin and Bifano unveiled the system's name in the SDSC machine room. A large, colorful banner featuring the name Blue Horizon will hang over the SP system, clearly identifying for visitors that this is one of the world's fastest supercomputers.

The multiple towers of Blue Horizon occupies nearly 25 percent of the available floor space in the SDSC machine room.

Bob Borchers and Rich Hirsh, director and deputy director of the NSF Division of Advanced Computational Infrastructure and Research, participated in the dedication ceremony via VTC from the NSF offices.

SAN DIEGO SUPERCOMPUTER CENTER

A National Laboratory for Computational Science & Engineering

9500 Gilman Drive, La Jolla, CA 92093-0505 • 858-534-5000 • www.sdsc.edu • www.npaci.edu