SOFTWARE

THE GODDESS OF WISDOM WOULD HAVE APPRECIATED

ENHANCING RESEARCH WITH PORTALS

RETOOLING FIELD RESEARCH — WIRELESSLY
FROM THE DIRECTOR

1 Building a Community Grid

PORTALS

2 Powerful Data Management Tool Comes of Age

SCIENTIFIC INNOVATION

11 Computational Anatomy: An Emerging Discipline

EDUCATION

14 Kids Thrive on Science: SDSC and Houston Girl Scouts Extend Successful In-School Program

16 Wireless Networking Extends High-Speed Internet to Remote Academic Sites in Southern California

NEWS

20 Blue Titan and SDSC Developing Web Software

20 Plant Genomics Website at SDSC Attracts Attention

20 SDSC and Entropia Harnessing Desktop PCs

20 Lifemapper Produces Species Distribution Maps and Models

21 NSF Awards $3.28 Million to NLANR

21 SDSC’s Rozanne Steckler Receives ‘Spirit of Preuss’ Award

BACK COVER

Understanding Protein-Ligand Interactions

Photography by Gail W. Bamber

Any opinions, conclusions, or recommendations in this publication are those of the author(s) and do not necessarily reflect the views of NSF, other funding organizations, SDSC, UC San Diego, or the NPACI partner institutions. All brand names and product names are trademarks or registered trademarks of their respective holders.

NPACI

Building the Computational Infrastructure for Tomorrow’s Scientific Discovery

The National Partnership for Advanced Computational Infrastructure (NPACI) joins 41 partner institutions in 17 states, Australia, Italy, Spain, and Sweden, in creating the foundation for a ubiquitous, continuous, and pervasive computational environment to support research by the world’s scientists. NPACI is led by UC San Diego, funded primarily through the NSF’s Partnerships for Advanced Computational Infrastructure program, and has its focus of activities at the San Diego Supercomputer Center (SDSC).

NPACI INFORMATION
Fran Berman, Director
David Hart, Communications Director
dhart@sdsc.edu
858-534-8314

SDSC
SDSC is a campus research unit of UC San Diego and the focus of activities for NPACI. Since 1985, SDSC has served the U.S. scientific community as a national laboratory for computational science and engineering.

GENERAL INFORMATION
Phone: 858-534-5000
info@npaci.edu
www.npaci.edu
www.sdsc.edu
Subscribe to ENVision and Online at:
www.npaci.edu/Press/subscriptions.html

ONLINE
www.npaci.edu/Online
Online is a bimonthly, Web newsletter of the latest research and developments from NPACI and SDSC.

ENVision
www.npaci.edu/enVision
ENVision, ISSN 1521-5334, is published quarterly by NPACI and SDSC Communications. For a free subscription or to make address changes, visit the website or contact the editor.

EDITOR: Rex Graham, rgraham@sdsc.edu, 858-822-5408

DESIGNER: Gail Bamber, bamberg@sdsc.edu, 858-534-5150

WRITERS: Kimberly Mann Bruch (KMB), Cassie Ferguson (CF), Mike Gannis (MG), Rex Graham (RG), David Hart (DH), Linda McIntosh (LM), Merry Maisel (MM), Paul Tooby (PT)

Any opinions, conclusions, or recommendations in this publication are those of the author(s) and do not necessarily reflect the views of NSF, other funding organizations, SDSC, UC San Diego, or the NPACI partner institutions. All brand names and product names are trademarks or registered trademarks of their respective holders.
Building a Community Grid

For many of us who have been working in grid computing for the last decade, the current environment is exhilarating. At present, there is strong interest in the Grid from researchers, developers, funding agencies, the commercial sector, and a growing segment of resource-limited applications communities. If any time is the time for grid computing, this is it.

At the same time, there is a growing concern that the Grid has been oversold. In particular, the concern is that the Grid will not be able to deliver its promise of immense resources, high performance, and ubiquity to applications developers and users. This concern is legitimate, as the Grid will ultimately only be as successful as its user base.

The issue is not that the promise of the Grid has been oversold as much as that the difficulty of developing the requisite Grid infrastructure has been underestimated. Grid infrastructure is far from solely an integration project. Serious, long-term research must be conducted to support the use of the Grid as a virtual computing platform where an application can “plug in” anywhere and achieve reasonable performance. An impressive list of problems must be solved for the Grid to achieve its promise, including scheduling of Grid applications for performance as well as throughput, the development of policies that can promote stability of the Grid while crossing political, social, and technological boundaries, the development of programming environments which promote the development and deployment of adaptive, performance-efficient programs for the Grid (such as those being developed on the GrADS project), security, fault-tolerance, performance monitors, debuggers, and other Grid tools.

If we are thoughtful, serious, and organized about our approach to building the infrastructure, we should have a usable, useful, and performance-efficient Grid by the end of the decade that evolves with new hardware, new software, and new applications. This is a lot of work and much must be done in the meantime from infrastructure integration and application development to basic research.

How will we build this vision of the Grid? Building the Grid and developing a robust and usable software infrastructure requires unprecedented coordination and cooperation. Given that grid research started with a relatively small community that remains closely knit (some say too much so), there has been an opportunity for agreement on a basic service architecture approach and basic interfaces that have allowed the community to work together effectively. In particular, a community model has evolved for the Grid with a four-layer design.

The resource layer will always be the available resources. On the TeraGrid project, we have the opportunity to focus on the middle two layers with a larger group because each site will deploy the same processor architecture. With the Extensible Terascale Facility just proposed, the Grid will be heterogeneous. Heterogeneity is a fact of life on the Grid and provides both added capabilities and capacities to users while at the same time making the process of developing the software infrastructure considerably more complex. Facing the heterogeneity problem early on will be required to help us gain the critical experience needed to develop the Grid we have promised.

The next layer, common infrastructure, is especially important. No matter what hardware underlies the Grid and no matter which applications are deployed on it, we must have a layer of software that the community agrees on. Global Grid Forum standards, the Open Grid Services Architecture, Globus, and NSF’s Middleware Initiative (NMI) are all part of the current picture, and the community must agree on what services will be available and how they will be interfaced.

The next layer is comprised of user-focused and community-focused middleware, pre-NMI software, and other packages. While the common infrastructure layer is evolving largely as open source, the middleware layer will likely have proprietary software and vendor products, too. The National Partnership for Advanced Computational Infrastructure (NPACI) is developing an interoperable collection of software called the “NPACKage” that will be deployed at all NPACI resource sites on top of NMI and NPACI resources (constituting an NPACI grid) as well as at other sites in the emerging global Grid.

Finally, the applications layer provides the raison d’être for the Grid. Without applications that truly find the Grid a useful platform, the Grid can’t be successful. Note that if the Grid is developed properly, it will provide access for applications to any Grid resources, whether or not the application is distributed. It is critical for our community to work closely with the applications communities to develop useful application models for the Grid. The Grid user community of the next decade must be a broad community encompassing existing Grid application paradigms (embarrassingly parallel applications, parameter sweeps, staged applications, etc.) as well as applications with dependencies (such as those found in computational biology), data-oriented “killer apps” over the next decade, and applications with polynomials.

The development of this Community Grid Model, as well as serious research into the problems of developing the Grid that are not “low hanging fruit” is critical to fulfill the promise of the Grid. The rapid development of today’s Grids demonstrates that an enormous amount can be done in a relatively short period of time. With serious focus, serious resources, and serious cooperation, we should be able to achieve the full vision of the Grid and help usher in a new era of science, software, and technology.
Initially released four years ago, the SRB middleware has proven increasingly popular, with more than 200 registered users at over 50 sites. According to Arcot Rajasekar, director of the Data Grids Lab at SDSC, SRB installations administered at SDSC alone currently manage more than 6.4 million files and 40 terabytes of data, with additional data managed in many SRB installations administered at other institutions. The SRB is also an integral part of modern scientific portals such as NPACI GridPort (see story, page 8), the Biomedical Informatics Research Network (BIRN) portal (see story, page 6), and the Visible Embryo digital library.

AGENCY-WIDE INTEREST

“The SRB is coming of age, which can be seen from the number of different projects and communities using this software to increase their productivity in managing scientific data,” said Reagan Moore, co-director of the Data and Knowledge Systems (DAKS) program at SDSC. And beyond individual projects, agency-wide interest in the SRB is emerging. For example, NASA gathers vast amounts of data and manages collections that contain terabytes to petabytes of data, among the largest on the planet. NASA operates nine Field Centers and many of the nation’s scientific satellites, the contractor-operated Jet Propulsion Laboratory, and the Wallops Flight Facility. NASA is now using the SRB in several projects, including an Earth Science Information Partner (ESIP) collection for land habitat data, and the NASA Information Power Grid (IPG).

“Our goal for the IPG is to provide a revolutionary approach for NASA’s science and engineering activities through services that integrate widely distributed computing, data, and instrumentation resources into the user’s work environment,” said William E. Johnston, IPG Project Manager. “The SDSC SRB is an important component of the IPG, allowing us to manage—and transparently connect to—data resources that are very large, heterogeneous, widespread, and multi-institutional.”

A parallel data-mining application (Figure 1) developed by Tom Hinke, of NASA’s IPG Deployment Project, has demonstrated persistent and uniform access to large data sets in heterogeneous, multi-organizational archival storage systems using the SRB.

In another NASA-related project, the University of Maryland is leading an effort to set up a data grid that uses the SRB to link several major sites of the NASA-supported ESIP Federation. The ESIP Federation has been very successful in addressing the major needs of the NASA Earth Science Enterprise (ESE), including the introduction of novel products and information services to a very broad user community, ranging from environmental and climate researchers to land use planners and emergency planners.

Current ESIP holdings (Figure 2) total more than eight terabytes. “The SRB will enable the next evolutionary phase of the ESIP Federation by providing...”

---

Figure 1: High Speed Distributed Data Access

The NASA Information Power Grid uses the Storage Resource Broker to support the simultaneous mining of multiple large data sets located at four sites: SDSC, Caltech, NASA Glenn Research Center, and Washington University. This NASA image shows mesoscale convective systems, or storms, ranging from red (intense) to blue (storm-free areas). The black diamonds are areas between satellite swaths that lack data.
an integrated data management system over multiple administrative domains, distributed heterogeneous storage systems, and environments at the various ESIP sites,” said Joseph Jaja, director of the University of Maryland’s Institute for Advanced Computer Studies (UMIACS) and professor of electrical and computer engineering.

Other NASA researchers are investigating further uses for the SRB, such as providing a front end for large NASA archival storage in a High Performance Storage System (HPSS).

VIRTUAL ASTRONOMY

In addition to large-scale agency applications such as the NASA projects, the SRB is finding uses in dozens of projects ranging from astronomy and neuroscience to environmental science and data grids.

In astronomy, researchers from 17 institutions are collaborating to establish standards that will support the National Virtual Observatory (NVO), a discipline-wide, expandable database of astronomical images, catalogs, measurements, and scientific publications that will unite more than 100 terabytes of data collected from 50 ground- and space-based telescopes and instruments.

By linking all of this data, along with analysis and visualization tools, in the form of easily-accessible Web services, the project will make a “virtual observatory” available to professional researchers, amateur astronomers, and students alike, greatly broadening and speeding astronomy research and education.

The SRB is used as a data grid within the NVO, and is already managing two important collections, the 2-Micron All Sky Survey (2MASS) and the Digital Palomar Observatory Sky Survey (DPOSS). Moore is a member of the NVO project’s senior staff. Together, these are the largest data collections under SRB management at SDSC, totaling 18 terabytes of data in more than five million files.
“By organizing images in containers with other images from the same part of the sky, the SRB is helping astronomers assemble seamless mosaics in multiple wavelengths,” said Roy Williams of Caltech’s Center for Advanced Computing Research (figure 3). “Being able to overlay multi-wavelength images for the same area enables new science—for example, you get a good picture of the physics of such things as large galaxies, extended sources, star formation, and objects like the Crab Nebula.” The SRB is proving valuable for the NVO through its flexibility; with a number of different interfaces, containers for efficient access, intelligent pre-filtering, ability to add metadata, and other useful features.

ECOLOGICAL RESEARCH

Environmental scientists are also making increasing use of large data sets, such as airborne and satellite measurements of the Earth. One form is known as hyperspectral data because it contains multiple wavelengths of light that can yield detailed information about the planet’s surface. Hyperspectral data is used in climate models, land use models, and many other applications (figure 4).

The SRB is being used in the NSF’s Long-Term Ecological Research (LTER) project in the Spatial Data Workbench, a collaborative project between the LTER and SDSC that is part of the NPACI Earth System Science thrust area. The LTER network is a collaborative effort involving 24 sites investigating ecological processes over the long term on a nationwide scale. “The SRB allows us to efficiently archive collections of large LTER data sets at SDSC, analyze, and distribute them through the Web,” said John Vande Castle, the LTER’s associate director for technology development. The SRB provides the foundation upon which the components of the Spatial Data Workbench are built. The system contains conventional remote sensing data as well as airborne hyperspectral data consisting of 224 wavelength bands in 10-nanometer increments for LTER research sites.

BIOMEDICAL INFORMATICS

In neuroscience, researchers around the country will soon share high-resolution brain images (figure 5) and other biological data sets of half a terabyte or more over a high-speed communications network as part of the Biomedical Informatics Research Network (BIRN). Funded by the National Center for Research Resources of the National Institutes of Health (NIH), the BIRN project will create a testbed to address biomedical researchers’ need to access and analyze data located at widespread sites. “By integrating data and computational resources with multiple software tools and advanced networking, we expect to make possible a rich new mode of scientific collaboration,” said UC San Diego neurobiologist Mark H. Ellisman, one of the principal investigators for BIRN and leader of the NPACI Neuroscience thrust area. “The ease with which the SRB provides access to large, distributed data collections makes it a key component in enabling exciting advances in the study of brain disorders.”

ANATOMY COLLECTIONS

The NIH-funded Visual Embryo project is providing a resource that enables medical professionals to collaborate with biological scientists in a visual form on questions involving the development of the human...
embryo (Figure 6). The project also gives educators and the general public access to images in the Human Developmental Anatomy Center embryo collections.

“The project is applying leading-edge information technologies in computation, visualization, and networking to create advanced collaborative capabilities,” said principal investigator Mark Pullen, a professor of computer science at George Mason University. “For image management in this demanding collaborative environment, the SRB has proven to be an extremely valuable tool.” The SRB is being used to ingest and manage more than 400 gigabytes of images in some 30,000 files, which are stored in the HPSS archive at SDSC and made available for controlled-access collaborations through the SRB.

E-SCIENCE GRID

As Grid architectures mature, data grids are becoming an essential part of modern scientific infrastructure. In the largest SRB deployment yet, United Kingdom researchers have chosen it as a core component of a data grid architecture that will connect users at more than 10 UK sites. Recently, SDSC researcher George Kremenek demonstrated the use of the SRB to connect Queen’s University of Belfast, the University of Edinburgh, and Rutherford Laboratory in a data grid. “We are exploring the advanced capabilities of the SRB and have hired a senior scientist to help with the support and future developments,” said Kerstin Kleese-van Dam of the CLRC-Daresbury Laboratory, e-Science Centre. In addition to distributing the SRB in a Grid Starter Kit, the UK Grid Support Centre will also provide support for SRB users throughout the UK data grid community. “Previously, all SRB support was provided by SDSC staff,” said Rajasekar. “So this is an important step because it will enable large-scale production use of the SRB.” —PT

—PT

FIGURE 6: VISIBLE EMBRYO
Digitized, very high-resolution embryo images allow medical professionals, biologists, and educators to visualize human embryo development. The SRB data grid helps them manipulate image data collaboratively for improved diagnoses, clinical case management, and medical education. This 280-megabyte embryo montage was produced by aligning multiple tiled image files.

FIGURE 7: A FLEXIBLE TOOL FOR MANAGING DATA
Responding to user requests, SDSC researchers have continued to develop additional ways to access and manage data stored in SRB collections. “Everything in a user’s SRB collections is accessible from any of three interfaces—the SRB client command line, the Web-based mySRB, and the Windows Explorer-like inQ,” said SDSC researcher Charles Cowart. Multiple interfaces enable users to use the SRB more flexibly and to access its many powerful features.
Neuroscience Portal Marshals Rich Network of Technologies

The driver need not manufacture the car, and the pianist need not be a piano tuner, wood turner, or varnisher. For the driver or the pianist, the underlying technology is invisible; it’s something they don’t have to worry about. In the same way, scientists want to use sophisticated computational tools—the computational infrastructure—transparently. “That means access to data, computation, and visualization—together, sequentially or concurrently—from the same simple interface,” said Mark H. Ellisman, leader of the National Partnership for Advanced Computational Infrastructure (NPACI) Neuroscience thrust area and director of the National Center for Microscopy and Imaging Research (NCMIR) at UC San Diego. “That has been the major goal of our NSF-funded Telescience alpha project, to build a transparent interface: the Telescience Portal.”

ACCESSING TOOLS AND APPLICATIONS

The Telescience Portal is an application environment supplying centralized access to all the tools and applications necessary for performing end-to-end electron tomography. It is Web-enabled, so it can be reached with a single login from any Internet-capable location. “One simple Web interface allows the user to accomplish many scientific tasks that invoke many kinds of software, yet the user need not be expert in any of the software,” Ellisman said. “The portal guides one through a session, launching whatever software is needed as it is called for.” The spiny dendrite image (Figure 1) is an example of work done on the portal.

The portal, a rich network of technologies, “For us and our many collaborators, these portals exemplify the advantages of a full computational infrastructure and can serve as models for others to build on,” said Ellisman, a professor of neuroscience at the UCSD School of Medicine. “This mix of tools in service to science shows the way into the future.”

A DECADE OF PREPARATION

NCMIR is a research facility dedicated to advancing technologies for microscopy and three-dimensional imaging of biological materials. It is supported by NIH’s National Center for Research Resources. Ellisman and colleagues have worked with scientists at SDSC for more than a decade to improve the computational technologies associated with advanced microscopies. NCMIR’s central instrument is a state-of-the-art 400-keV intermediate-voltage transmission electron microscope, equipped with charge-coupled device detectors and capable of operating under complete computer control. Because of its higher accelerating voltage, the instrument is able to penetrate thicker specimens than can be imaged with a conventional electron microscope, facilitating 3-D reconstruction.

NCMIR scientists have made their microscopes capable of remote operation over the Internet, and they have joined with colleagues at Osaka University in Japan to do the same for a more powerful 3 MeV ultrahigh-voltage electron microscope, a unique, world-class resource. They have also inaugurated another international telesience collaboration—with the National Center for High-Performance Computing in Taiwan.
nisms for the data flow, and graphical user interfaces have a uniform “look and feel” specifically tailored to the needs of the biological researchers. The interface has also been adapted for wireless devices (Figure 3), realizing the idea of ubiquitous access and control.

“What we’ve pulled together goes beyond what might appear as ‘the portal’ to any individual user,” said Ellisman. “We see it as a partnership that unites visualization, computation, databases and digital libraries, network connectivity, remote instrumentation, and opportunities for collaboration, education, and outreach.”

BIOMEDICAL TESTBED

BIRN, the Biomedical Informatics Research Network, is an ambitious project of the NIH’s National Center for Research Resources aimed at creating a testbed to address biomedical researchers’ need to access and analyze data at many levels located at diverse sites throughout the country. The plan for the testbed is to focus on research involving neuro-imaging, taking advantage of the relative sophistication of this community in the use of information technology. The participants include NIH General Clinical Research Centers and co-located Biomedical Technology Research Resources at UCSD, UCLA, Caltech, Duke University, and two hospitals affiliated with Harvard Medical School. The BIRN Coordinating Center is at the National Biomedical Computation Resource at SDSC and is led by Ellisman and colleagues at NCMIR and SDSC. The BIRN sites are linked in a data grid via Internet 2 installations, and each site has a special complement of visualization, dedicated storage, and computational equipment in addition to its own specialized resources.

“Our portal is modeled on the Telescience Portal,” said BIRN Portal developer David Lee of NCMIR. “But it is specialized for the work of BIRN. Like the Telescience Portal, the BIRN Portal unites application-specific interfaces to complicated technologies, with the same layered structure.” For BIRN, portal developers are adding personal, group, and collaborative workspaces to the tools already available. Many more applications are to be integrated into the portal architecture to enable collaborative work among the BIRN investigators.

“Our portal work is creating infrastructure to enable new science, new technology, and new knowledge of value to disparate groups of users,” said Ellisman. “We’ve brought together activities that engage computer scientists, biomedical scientists, and basic scientists in various fields. The NPACI developments are now being used in BIRN and will be used in a variety of activities within the PACI program and beyond. This is a groundbreaking model, with far-reaching effects and benefits for many disciplines, and we are confident that it will be influential in the design and development of commercial and industrial analogues.” —MM

FIGURE 2: DECEPTIVELY SIMPLE ON THE SURFACE
A structure of modular layers supporting the Telescience Portal leads to transparency for the user of the complex infrastructure.

FIGURE 3: WIRELESS NEUROSCIENCE
The Telescience Portal can be accessed from a handheld computer with a wireless connection; the portal interface is simplified for the smaller screen.
GridPort Provides Simple Entrance to Scientific Computing

Portals allow scientists to spend more time conducting research and less time learning to use high-end computers, according to Thomas, leader of the Grid Computing Group at the Texas Advanced Computing Center (TACC) at the University of Texas, Austin, and leader of the Grid Portals Project in the National Partnership for Advanced Computational Infrastructure (NPACI) Resources thrust. “By developing portals and portal technologies that integrate the Grid and NPACI resources into simple portals, where the users may not even know they are on a supercomputer or on the Grid, we have the potential to change the way science is done,” said Thomas. Her goal is to provide an integrated Grid-computing environment, such as a portal, for users to access NPACI services.

The success of the NPACI HotPage and a number of other Grid-computing portals demonstrates that portals have become known as interfaces that enable scientific users to access information and services in Grid-computing environments. Grid portals hide the complexities of Grid technologies from the user and present simplified, intuitive interfaces to underlying resources. Grid portal toolkits that interface to Grid technologies allow developers to build portals quickly and easily. The NPACI Grid Portal Toolkit (GridPort) has facilitated the development and utilization of Grid technologies such as the Globus Toolkit and the San Diego Supercomputer Center (SDSC) Storage Resource Broker (SRB) from within a unified application programming interface. GridPort supports a set of centralized services that allow multiple application portals to share services and a single-login environment.

“NPACI portal efforts are aimed at making HPC systems and Grid resources accessible to researchers across all fields of science—even researchers with little or no HPC or Grid experience,” said Fran Berman, director of SDSC and NPACI. “We’ve been successful in meeting this long-term goal—allowing individual scientists to build simple portals that take advantage of existing computing infrastructure without investing a significant amount of time or money.”

GRID COMPUTING PORTALS

Computer scientist Pavel Pevzner embodies the type of researcher who Thomas hopes to reach using NPACI portal-creation software called the GridPort Toolkit. Pevzner, the Ronald R. Taylor Professor of Computer Science at the University of California, San Diego (UCSD), and his colleagues at UCSD and the University of Southern California set up a portal, to be available January 2003, that runs a program called Euler (pronounced “OI-ler” and named for 18th century mathematician Leonhard Euler) that uses computers to join up bits of DNA fragments into much larger pieces. While Pevzner is at home with the complex mathematics and programming behind Euler, he has no idea what machine is running the application. According to Thomas, there is no reason he should know. What is important to Euler portal users is that they can submit their carefully prepared sequence data,
set a few variables, and press a button to start the process of unscrambling DNA fragments.

“As scientists develop new applications, they want to share their software with the rest of their community,” said Steve Mock, group leader of the Grid Portals Architecture group at SDSC. “When they set up a portal, it not only serves as a distribution point for the application, but also gives other researchers the means to run it on a set of distributed Grid resources.”

The NPACI portals team developed and refined the GridPort Toolkit and the GridPort Client Toolkit to facilitate the construction of Grid portals. The team’s development of the informational NPACI HotPage, which allows NPACI users to access computing resources through a Web interface, led to the development of GridPort, which directs a user’s input through layers of software known as “middleware.” GridPort is an open architecture designed to be flexible and capable of using other Grid services and technologies as they become available. It is based on advanced Web, security, and metasystems technologies to provide secure, interactive services. GridPort provides portal developers with an interface to an infrastructure that integrates Grid technologies developed by NPACI, the National Computational Science Alliance, and the Grid community. These technologies include the SRB, MyProxy, the Globus Toolkit, and the Grid Security Infrastructure.

The team has also introduced a new approach to portals that allows researchers to tailor a Grid interface to their needs without writing or installing specialized Grid software. Researchers using the GridPort Client Toolkit can create a Web-based portal that communicates with GridPort, which is installed on the computing resource. Key to the design is that the portal’s Web pages can be hosted on any Web server where the user has accounts, or even run as local files from a browser on a laptop. Researchers need only create a Web page to communicate with the NPACI Grid via GridPort to other software. After logging in via the Web, users can securely access their accounts on all NPACI systems, run jobs and applications, check system status, and manage files.

“Now, essentially all it takes to create a customized application portal to GridPort is to incorporate a few extra lines of HTML into a Web page,” said Maytal Dahan, a software developer in the Grid computing group at TACC.

GRID COMPUTING PORTALS

People have different definitions of the term “portal,” a buzzword of the Internet era. Some think of it as the page that appears when they launch a Web browser. For others, a portal is a Web page with links and tools related to a specific topic. Some portals cater to users by providing personalized news, weather, calendar, links, plus a search engine. Software companies use the term as a way to describe products linking business solutions, such as customer management, tracking systems, and billing applications. Prior to the Internet, and its tendency to transform functional metaphors into jargon (surf, spider, shopping cart), a portal was simply an entrance.

A horizontal portal, such as Yahoo!, links to a broad selection of different topics, but does not provide enough information to allow for deep understanding. Yahoo! can also be characterized as an information portal, which provides the user with a broad set of information or links to information. The opposite of a horizontal portal would then be a vertical portal. A vertical portal contains extensive data on a single topic, such as www.fifaworldcup.com, which has many layers of very specific information about the World Cup Soccer tournament. Another form of vertical portal is an application portal, which provides an interface to a specific computer program. “The ideal application portal actually improves the original applications by providing a simplified user interface and extra features such as visualization and a parameter sweep,” said Kurt Mueller, production lead for the NPACI and PACI HotPages.

GridPort has contributed to successes in several fields—from chemistry to molecular biology—and provides an interactive window to NPACI and PACI resources through their respective HotPages. For researchers with codes that require HPC resources, the portals reduce the effort required to develop an interface as well as automatically giving them access to other NPACI resources, particularly the data-handling capabilities of the SDSC SRB.

Computational scientists have different motivations for using computational Grid portals, which is reflected in the variety of portal types. Portals based on GridPort include user portals, such as the NPACI and PACI HotPages, in which the user interacts directly with the HPC resources; community model application portals, such as the LAPK portal, which hide the fact that a Grid or HPC resource is being used; remote instrument application portals, such as the UCSD Telescience portal (see story, page 6), in which users control remote equipment and migrate data across institutional boundaries; and systems of portals such as those being constructed by the NSF-funded National Biomedical Computation Resource (NBCR) program including: the General Atomic Molecular Electronic Structure Systems (GAMESS), Adaptive Poisson-Boltzmann Solver (APBS), Assisted Model Building with Energy Refinement (AMBER), and the Combinatorial Extension (CE) Web Portal. In addition, there are a number of remote sites that have used GridPort and installed their own versions of the HotPage.

“We now have to think not just about the high-performance computers and their architecture, but about Grid-based and data-oriented tools and visualization and analysis tools that constitute the running environment,” said Kim Baldridge, senior principal scientist at SDSC and program manager of NBCR. “It is at the intersection of hardware, interfaces, models, and predictions that the best work will be done in quantum chemistry.” Grid-based tools include applications and visualization servers reachable via GridPort.

Computational chemists at SDSC were among the first to build a research portal based on GridPort, the
GAMESS quantum chemistry package, a favorite tool of the computational chemistry community. This portal, developed by Dahan, provides researchers with seamless access to computational resources across the country, allowing them to investigate ab initio (from first principles) quantum chemistry. “I hear often from researchers who are interested in building a portal but feel overwhelmed by task before them,” said Jerry Greenberg, a computational chemist and staff scientist at SDSC, who has helped to make several applications for the NBCR program available using the GridPort Toolkit. “GridPort is great in that once a portal is set up, it is really easy for users to work with their applications.”

FUTURE DIRECTIONS

GridPort is poised to move from computational chemistry and molecular biology to new fields, such as environmental informatics, astrophysics, and fusion. David Stockwell, an assistant research scientist at SDSC, and his colleagues have used it to create a portal to WhyWhere, a Web-based resource that provides quantitative answers to the question “Where is it found and why?” for any species, anywhere on the globe. WhyWhere gives researchers unprecedented access to environmental data, software for analysis and visualization, and high-end computing power, all integrated in an easy-to-use Web interface. Similarly, Cathie Mills, a developer in the SDSC Grid Portal Architectures Group, has been working with Michael Norman, a physics professor at UCSD. They have been working with the Grid Portals Architecture group to build a portal for Enzo, a cosmological structure application that was used to simulate the condensation of the universe’s first star, a result that was featured in the November 16, 2001, issue of Science. GridPort will soon play a significant role in portal systems for the Department of Energy SciDAC program since it will be used as the framework for the development of a Fusion Grid Portal for the Fusion Grid.

Since 1999, the NPACI Portals team has successfully built production application portals and demonstrated the potential of GridPort. The GridPort project has also been key to the advancement of Grid technologies by being early adopters and users, which, in turn, drove the development, debugging, and testing of those technologies within production environments.

According to Thomas, the advent of the Open Grid Services Architecture (OGSA) and Grid Web Services has created new opportunities for making Grid-portal toolkits that are even simpler for developers, and GridPort will incorporate these new technologies. New portal technologies such as Java Server Pages, customization, and portlets are being integrated into GridPort to make it possible to develop portals that allow users to meet their personal needs more effectively. GridPort will incorporate new portal technologies such as Java Server Pages, customization, and portlets to make it possible to develop portals that allow users to personalize their portals as effectively as possible. Thomas says that as a result, Grid portals will be further developed and utilized.

Collaboration has always been key to the project. Along with sharing portal resources within NPACI, the group has worked with the Alliance portal team and NASA’s Information Power Grid, which has resulted in new Grid software such as MyProxy, and Grid Portals, such as the PACI HotPage. More recently, the collaboration effort has resulted in a new version of a HotPage for the TeraGrid, which will be deployed as the largest, most comprehensive computing infrastructure ever created for scientific research. The TeraGrid will link more than 16 teraflops of computing power, more than one petabyte of storage, and advanced visualization environments, all integrated as a grid system. TeraGrid sites include SDSC, NCSA, Argonne National Laboratory, and the Center for Advanced Computing Research at Caltech. —CF

ENTRYWAYS TO MANY SCIENTIFIC FIELDS

Portals have become established as effective interfaces that let scientific users access information and services in Grid-computing environments, as demonstrated by the success of the NPACI HotPage and numerous other Grid-computing portals. Grid portal toolkits that interface to Grid technologies have proven useful, enabling developers to build portals rapidly. The NPACI Grid Portal Toolkit (GridPort) has led to the creation of portals in scientific fields from chemistry and molecular biology to astrophysics.
Computational Anatomy: An Emerging Discipline

The new discipline of computational anatomy uses mathematical analysis to learn how tissues grow, assume new shapes, and “morph” into mature structures that, at first glance, might look like the action in popular music videos. “Our mathematical maneuvers are distinct from the crude morphing seen in a Michael Jackson video,” said Michael I. Miller, director of the Center for Imaging Science at Johns Hopkins University, an NPACI partner in the Neuroscience thrust area. “The video is done without any coordinate system, by simple photometric transformation. Our mathematical formulation deforms structures in a coordinate space, and thus the original structures are entirely recoverable computationally. That is the difference between morphing and morphometrics.” Miller and his colleagues study, quantitatively, how one brain differs from another, and they’re extending their techniques to predict anatomical anomalies, such as tiny tumors.

“Neuroscientists have long attended to questions like these,” Miller said, “and if not for their sorties into this field, computational anatomy would not have become the rapidly emerging discipline that it is today.”

ESTABLISHING NORMS

For the past decade, Miller’s group has been developing computational methods to analyze gross anatomical structures in the human brain, with the objective of creating tools to help neuroscientists and diagnosticians learn from changes in brain substructures. The underlying mathematics are supplied by metric pattern theory, a formalism developed by Ulf Grenander in the Division of Applied Mathematics at Brown University.

“Computational anatomy could be described as a digital textbook of anatomy, with all its variability in healthy humans, adjusted for things like gender, age, and ethnicity, and also in pathological situations that affect anatomy,” said Grenander. “The main difficulty is that anatomical substructures form highly complex systems, with variation being the rule,” Miller said. As he said in an interview published earlier this year1, “If machines can compute structures that are equivalent to the structures we see in the world, then we can begin to understand them. In computational anatomy, we now have equations that describe how tissues can grow and bend and morph and change. These equations seem to generate very realistic structures.”

Johns Hopkins graduate student Mirza Faisal Beg compared the problem to the much simpler problem of determining the “normal” human heart rate. “When we take the pulse of many thousands of people over many years, we see that the norm is a large cluster around 72 beats per minute, plus or minus three to five beats,” said Beg. “We’re just in the beginning stages of norm establishment with brains, because the notion of normed shape is only now being defined mathematically. But it is at the core of the metric basis for shape in computational anatomy.”

Miller has pioneered an approach to volume metric mapping that considers a starting anatomical description as a deformable template. The template anatomy

---

1 G. Taubes (2002): An interview with Dr. Michael I. Miller in ISI’s In-Cytes: www.incites.com/scientists/DrMichaelIMiller.html
is varied via transformations applied to subvolumes, contours, and surfaces. The equations governing the transformations are generalizations of the Euler equations of fluid mechanics. The rules of growth and transformation, while programmed through genes, obey the physical laws of nature represented in the equations. The group of Grenander, Miller, and French mathematicians Alain Trouvé and Laurent Younes, applies this work to constructing metrics on shape and form, with an emphasis on variation in brains.

BRIEF HISTORY OF A NEW FIELD

One of the first attempts to bring mathematical and physical insight to bear on nature’s limitless variety was made by D’Arcy Wentworth Thompson (1860–1948), a zoologist, mathematician, and classis who worked at the Scottish universities of Dundee and St. Andrew’s. It was Thompson who found it significant that surface-to-volume ratios of species decline as those species get larger. Small creatures are governed more by surface forces, larger creatures by gravitational (volumetric) forces. In 1917, Thompson wrote:

In a very large part of morphology, our essential task lies in the comparison of related forms rather than in the precise definition of each; and the deformation of a complicated figure may be a phenomenon easy of comprehension, though the figure itself have to be left unanalysed and undefined. This process of comparison, of recognising in one form a definite permutation or deformation of another . . . lies within the immediate province of mathematics, and finds its solution in the elementary use of a certain method of the mathematician. This method is the Method of Coordinates, on which is based the Theory of Transformations.

Miller noted that Thompson’s work pioneered the principles of growth and form at a time when there were few, if any, computational methods for actualizing them (Figure 2). The theory was juxtaposed to more genetically based theories of variation, until Grenander appreciated the fundamental basis of the transformational approach as the foundation of a metric pattern theory. Other pioneers include Fred L. Bookstein of the University of Michigan and Ruzena Bajcsy of the University of Pennsylvania. “Building the digital textbook along general lines for humans and others is a major undertaking that will require years of hard work,” Grenander said, adding, “if anyone can do it, it is Michael Miller.”

APPLICATIONS OF COMPUTATIONAL ANATOMY

The group applies their mathematical methods to original anatomies derived from various modes of brain imagery, including positron emission tomography (PET), computerized tomography (CT), and magnetic resonance imagery (MRI). Miller and J. Tilak Ratnanather, an assistant research professor in biomedical engineering at Johns Hopkins, and their collaborators, Arthur Toga and Paul Thompson, of UCLA’s Laboratory for Neuro Imaging, an NPACI Neuroscience partner, are studying growth and development in young, normal brains. An example (Figure 1) shows growth in the corpus callosum area of the juvenile brain. The fastest growth is in the isthmus, which carries fibers to areas of cerebral cortex that support language function. As Toga and Thompson have shown, applying mathematical theo-

FIGURE 2: HORN SHAPES

The tubular horns of sheep, goats, oxen and other herbivores can be circular, triangular, or more complicated polygons, and the variable rates of horn growth cause the angles or edges to trace an ensemble of homologous spirals. In some, such as the male muskox, the horn is not developed in a continuous spiral, changing shape as the individual grows, showing that it does not constitute a logarithmic spiral.

ries of deformation in computational anatomy to a grid overlying the youngest anatomy results in a “prediction” of the later anatomy that accords well with what was seen in the older subjects.

Key here is placing similar structures in the same metric space and then “deforming” one structure to another, called the target structure. The deformations have mathematical properties that persist through subsequent structures; and a statistical calculation of sameness and difference can be made across an entire group of similar structures. “You’d like to have some quantitative measure that actually says that one anatomical shape is moving closer to or farther from another. The distance between one structure and another can also be thought of as the square root of the energy required to transform the first structure onto the metric of the second,” Miller said, “with the assumption that normal transformations follow the least-energy path.”

One of Miller’s favorite examples of the uses of computational anatomy comes from work done in the laboratory of NPACI Neuroscience partner David Van Essen at Washington University in St. Louis, MO (Figure 3). Van Essen, one of the earliest pioneers in computational anatomy, has “built a metric space in which monkey brains and human brains can be usefully compared and contrasted,” Miller said. The work of cortical “flattening” pursued by the Van Essen group is a challenge to the mathematicians, said Miller, to further develop metric pattern theory and realize its promise in neuroanatomy so the geometries of the cortical folds can be precisely quantified in terms of their metric shape.

“The study of brain geometry is now the study of all submanifolds of anatomical significance: landmarks, curves, surfaces, and subvolumes, all taken together to form a complete volume,” Miller said. The more dimensions in which a structure is mapped and the more external physical forces and flows considered, the more degrees of freedom are introduced into any calculation. The complete methodology combines mappings that carry all of the submanifolds of points, curves, surfaces, and subvolumes together. The only constraint, according to Miller, is that the original anatomies must be “topologically identical”—an object in one must occur in the other and must change smoothly from one to the other. “The generalized Euler equations of motion for describing shapes of the same topology are extremely computationally intensive,” Miller said.

The image-matching codes developed by Beg to compute these metric distances run on SDSC’s IBM machine, Blue Horizon, and on Zeus (a similar IBM machine) at the Center for Imaging Science at Johns Hopkins. Beg said, “Image-matching codes are naturally amenable to parallelization via decomposition of the image data into smaller pieces, each given to a different processor.” The Center for Imaging Science was chosen for an SDSC Strategic Applications Collaboration project, and Beg worked with Tim Kaiser from SDSC on optimization and efficient partition algorithms.

![Figure 3: Flattening the Macaque Cerebral Cortex](image)

In color, the cortex as it appears in the normal macaque (top) and in the flattened version used for comparisons (bottom). In black and white, the grid or source map used for the deformation of one comparison brain (A), the deformed source map (B), and the vector field of displacements from one to the other (C). The folds (shading) and target registration lines (black lines) on the reference (atlas) map are shown in (D), while (E) shows the difference between the comparison brain and the atlas map with respect to the same folds.

**NEW DIRECTIONS**

The higher resolution of new brain imagery allows the scientists to consider transformations of a given anatomy not only through mappings of coordinate systems, but also through transformations of the luminance values of the imagery. Since these often correspond to the appearance or creation of new structures or substructures, “our methods actually let us catch a tumor as it grows or a brain region like the hippocampus as it shrinks,” Miller said.

“Work like this gives us new insight into the way that brain structures grow and develop,” Miller said, “and the comparison of this normal growth with abnormal structure can be more accurate as a result. Computations like these in the hands of diagnosticians may ultimately lead to earlier detection and treatment of brain tumors.” —MM

**REFERENCES**


Kids Thrive on Science: SDSC and Houston Girl Scouts Extend Successful In-School Program

Some of the girls come into the program with early signs of trouble: inattention, hostility, and even rebellion. But then the program draws them in. “It’s great to see these kids suddenly glue themselves to computer screens, focused, attentive, and engaged,” said Abby Sibley of the Houston-area San Jacinto Girl Scout Council. Since September 2000, thousands of girls in Houston area schools have been engaged, for the first time in their lives, in the experience of enjoying math and science. The science enrichment program in Houston originated in San Diego, and it is one of the most important contributions of the San Diego Supercomputer Center (SDSC) to the national Education, Outreach, and Training program of the Partnerships for Advanced Computational Infrastructure funded by the National Science Foundation (NSF).

SDSC education director Rozeanne Steckler has teamed with the Houston Girl Scouts to bring science and hope to children considered to be “at risk.” As it enters its third year, the Houston outreach program is gaining momentum.

“In our first two years, the program reached about 3,000 girls at 18 schools in the area,” said Gladys Birdwell, director of community outreach for the San Jacinto Council. “We used modules that Rozeanne Steckler brought to us. Each class had a different topic, and the girls, from kindergarten through sixth grade, just loved it,” Birdwell said. The modules were tested and used in county schools by Steckler and the Girl Scout Council of San Diego and Imperial Counties as part of an ongoing, in-school program called Girls are GREAT (Gifted, Resourceful, Extraordinary, Ambitious, and Talented). The lessons include inexpensive science and technology kits, all loaded into 35-gallon plastic buckets, for hands-on study of everything from fossils to butterflies, and solar clocks to robots.

Over the past year, Houston teachers, Girl Scout staff, and volunteers have carried on the original program, while Steckler and the San Jacinto Council piloted an expanded effort under funding from an NSF grant to Steckler’s program. “We bought 20 laptop computers and organized classes that let some 250 kids at five schools use them over a period of three weeks,” said Sibley, who is in charge of the in-school programs and who also is community outreach membership manager for the San Jacinto Council.

Sibley and the teachers she trained, together with Girl Scout Council staff, managed the complicated logistics of transporting the laptops from school to school, from classrooms to libraries, and into use by groups of 20 to 25 girls. The computers were loaded with games and simulations that Steckler selected for their interactivity. “They had Sammy Science, Sim Tunes, Sim Town, Sim Safari, and other programs, which all require imagination as well as plain responses,” Sibley said.

Fears that the laptops would be intimidating vanished immediately. “The girls just took to them instantly, and they were really sorry when the sessions were over,”
Sibley said. The laptop sessions on one day were followed the next by hands-on work with the science and technology kits. Each school had two hour-long sessions per day, one for kindergarten through second grade, the next for third through fifth grades.

For some girls, the sessions were their first encounters with computers. “We were particularly successful in the Baytown area, where the Girl Scouts have a partnership with the Gang Activities Prevention Program,” Sibley said. “We’ve got a small summer program at San Jacinto, Carver, and Harlem elementary schools in Baytown, involving 46 girls, and we’ve been using the laptops in every school we can since the new year began in August. None of this would have happened without Rozeanne and the experience she brought from her programs with the San Diego Girl Scouts. I don’t think I’ve ever seen a working scientist take such a deep interest in fostering scientific awakening in children.”

Sibley, a bilingual community outreach manager who has worked in the Girl Scout Council for a dozen years, also has a background in psychology and 12 previous years of experience dealing with teenagers in the juvenile justice system. “I came to the Girl Scouts because I saw that intervention would only succeed much earlier, when directed to the very young,” she said. More than 80 percent of the inner-city school population is Hispanic, and the rest are Anglo, African-American, and Asian-American.

The program has won rave reviews from school principals. A letter from Principal Agelia Durand and Assistant Principal Vicki Sipe of Woodview Elementary said, “Woodview did not want to let another day go by without thanking the San Jacinto Girl Scouts for the wonderful opportunity you have provided for our girls, parents, and leaders . . . the past three years have been a huge learning experience. When the program first started . . . we had to really push to enroll 40 girls. This year, we have 103.” The Woodview letter goes on to point out that the science and technology kits have enabled cooperative learning in small groups, empowering “future biologists, future engineers, even future saleswomen,” and it concludes, “We have learned that we have a future.”

“The lessons are relevant and fun,” wrote Naomi Orozco Clements, assistant principal of South Houston Elementary School, in a letter. “It has been exceptionally exciting to have the girls working on the laptop computers. I observed them during one of their meetings. They were all involved in a wonderful learning experience!”

The San Diego County version of the program was evaluated in March 2002 by the LEAD (Learning through Evaluation, Adaptation, and Dissemination) Center at the University of Wisconsin, Madison, which evaluates the impact of NPACI-originated educational outreach programs. The LEAD report noted that the program “addresses key issues in the education of girls in science and mathematics.” The report also states, “By providing opportunities for girls to collaboratively explore through hands-on science activities [the program] helps to engage girls in learning, develop more positive attitudes towards science, and increase their self-confidence. [It] serves a particularly important need in poor ethnic minority communities where resources are scarce and science education may not be emphasized.”

“We have learned that a carefully planned program in enthusiastic hands can grow to touch an entire city,” said Steckler. “We’re eager to expand into a third city next year, if the funding we’ve requested comes through. I’ve been particularly impressed by the way Abby has made a really scarce resource—just 20 laptop computers—work for so many girls in such a short time. She’s going to be able to provide them to 20 to 25 schools in the program during the next year, and I have no doubt that the results will call forth similar efforts in other area schools.” —MM
CLIMBING microwave towers on wind-swept mountain peaks is part of the job for research scientist Hans-Werner Braun from the San Diego Supercomputer Center (SDSC) at the University of California, San Diego (UCSD). Braun and his team trek up rugged ridges, climb radio towers, and engage in creative engineering to link remote Native American learning centers, astronomical observatories, nature reserves, and seismic sensor sites to the super-fast wireless communications network they are building in southern California’s backcountry. “Researchers have spied up Internet access, but they haven’t brought access everywhere,” said Braun, principal investigator of the High Performance Wireless Research and Education Network (HPWREN). “We’re focused on connecting remote science and education sites, and exploring ways in which this technology may be applicable for rapidly deployable emergency communications.”

With the help of a $2.3 million National Science Foundation (NSF) grant, Braun and his team of technical whizzes are bringing the 45 megabits-per-second (Mbps) HPWREN connectivity to some of the remotest corners of San Diego County where such high-speed access would otherwise be unimaginable. The high-speed wireless microwave network runs from SDSC on the Pacific coast to the mountains east of San Diego. Digital signals zip between microwave towers, antennas, and relays situated on ridges, in valleys, and on the peaks of local mountains such as Mount Laguna, Mount Soledad, Red Mountain, and Mount Palomar.

HPWREN’s availability in hard-to-reach areas provides us with the ability to conduct more detailed studies of fault zone structure and seismic wave activity, which in turn will provide seismologists around the world with more accurate data. “Researchers have spied up Internet access, but they haven’t brought access everywhere,” said Braun, principal investigator of the High Performance Wireless Research and Education Network (HPWREN). “We’re focused on connecting remote science and education sites, and exploring ways in which this technology may be applicable for rapidly deployable emergency communications.”

ECOLOGY APPLICATIONS

Spanning more than 4,000 acres, the Santa Margarita Ecological Reserve is nestled between the Santa Ana Mountains in northeast San Diego County. Until last year, remote access to real-time field data at the reserve was only a dream. However, scientists working at the biologically diverse area are now connected to the HPWREN backbone. They also have the capability of linking into the same network from their labs throughout the world to access the data being collected by field instruments in the reserve. Biologists monitor animal calls and other wildlife activity such as bat sonar—without being at the reserve. Water quality and hydrology sensors along the Santa Margarita watershed allow hydrologists to remotely measure parameters such as water acidity and pressure. Cameras at Santa Margarita, linked to the network, give researchers glimpses of mountain lions, golden eagles, hummingbirds, and other wildlife. “Field study is not what it used to be,” said Sedra Shapiro, executive director of the Field Station Programs at San Diego State University (SDSU). “We used to think of field biologists with a backpack and binoculars, but now we have very sophisticated computer technology supporting our field studies.”

“Santa Margarita’s connection to HPWREN is changing the ways in which we conduct science at our four field stations,” said Shapiro. “This high-speed connection also gives us the power to efficiently disseminate information to our students, nearby landowners, government agencies, and policy makers.”

REAL-TIME GEOPHYSICS

Today’s seismologists can view activity from the Earth’s interior as it occurs—enabling scientists from around the world to gain extensive insight into the often-elusive patterns of seismic waves. However, current seismic research techniques don’t provide enough information to determine the exact causes and effects of earthquakes.

HPWREN co-principal investigator Frank Vernon, a researcher at the Scripps Institution of Oceanography’s (SIO) Cecil and Ida Green Institute of Geophysics and Planetary Physics (IGPP), is taking real-time data collection and distribution one step further. “HPWREN enables field scientists like myself and others to send and receive continuous real-time data from remote research stations,” said Vernon. “This lets us collect and distribute data sets that we wouldn’t have access to otherwise.”

“For example, current monitoring systems do not have enough coverage for understanding the detailed three-dimensional fault structure of the San Andreas and San Jacinto faults. HPWREN’s availability in hard-to-reach areas provides us with the ability to conduct more detailed studies of fault zone structure and seismic wave activity, which in turn will provide seismologists around the world with more accurate data.”

HIGH-SPEED ASTRONOMY

More than 50 years have passed since astronomers first started observing the night sky with the Palomar Observatory’s 48-inch Oschin telescope, which was
recently connected to HPWREN’s 45 Mbps backbone. Coupled with the Oschin’s sophisticated digital camera, the high-speed network connection allows researchers to transmit near-real-time images from the observatory to astronomers worldwide. The images are primarily used for near-Earth asteroid research and supernova detection.

“This high-speed connection allows us to easily access data-intensive archival systems, which was really cumbersome—and often impossible—with a 56 kilobits-per-second dial-up modem,” said Steve Pravdo, project leader for the Near-Earth Asteroid Tracking (NEAT) program at Caltech’s Jet Propulsion Laboratory in Pasadena. “These archives let us compare current images with past images in near-real-time, which means that we are more efficient in our efforts to discover and confirm potentially dangerous asteroids.”

Images captured by NEAT cameras are also used by researchers with the international collaborative project called the Nearby Supernova Factory, which includes scientists from the Lawrence Berkeley National Laboratory (LBNL), the Laboratoire de Physique Nucléaire et de Hautes Energies de Paris, the Institut de Physique Nucléaire de Lyon, and the Centre de Recherche Astronomique de Lyon. “Our goal is to discover nearby supernovae and study them in detail so that they can be used more effectively as cosmological distance indicators,” said Greg Aldering, a staff scientist at LBNL and project leader for the Nearby Supernova Factory.

“This real-time connectivity is also crucial to astronomers at Mount Laguna Observatory east of San Diego. Remote observing at the observatory has not only broadened its user base, but has also opened significant research and educational opportunities as global astronomy projects proliferate. “SDSU astronomers are routinely using HPWREN to transfer astronomical images and spectra to campus, to other astronomers in the California State University system, to our partners at the University of Illinois, and to other collaborators worldwide,” said Paul Etzel, director of Mount Laguna Observatory. “Data are therefore received by research astronomers in a timely fashion, unlike the previous 56 kilobits-per-second modem connection, which had been our only data line serving the remote, eastern part of San Diego. In fact, prior to HPWREN, it was faster to write the data to tape and drive with them to San Diego or even Illinois.”

HPWREN is also central to the observatory’s formulation and success of future projects. For instance, SDSU is currently building a high-speed frame-transfer Charge Coupled Device (CCD) camera so that astronomers can better study rapid variations in cataclysmic variable stars. This new CCD camera at Mount Laguna Observatory will take two-megabit images every 0.10 to one second and produce approximately 100 gigabits of data during a typical eight-hour on-target observing sequence. HPWREN’s 45 Mbps transfer rate will allow Mount Laguna Observatory astronomers and their collaborators to access this real-time data throughout the world.

REMOTE EDUCATION

In collaboration with the Southern California Tribal Chairmen’s Association (SCTCA), the HPWREN team has been working to connect the learning centers of the 18 Native American reservations in San Diego County to the high-speed Internet. Originally linking the Pala, Rincon, and La Jolla tribes to HPWREN, the research team is now helping the SCTCA connect the remaining tribes to the Tribal Digital Village Network (TDVNet), which is funded by a grant from Hewlett-Packard Co.

“This technology transfer is being conducted by in-kind replacements of the equipment back to HPWREN, and assuming TDVNet ownership of...
what had been deployed earlier,” said Braun. “This transfer of equipment ownership also includes educating TDVNet staff with the expertise that they need to design, build, configure, and operate a high performance wireless data network.”

“HP is very pleased with the progress made by the TDV team in the last few months,” added Sukumar Srinivas, the original Hewlett-Packard executive responsible for the TDVNet grant to the SCTCA. “The level of competence in dealing with the HPWREN technology and the enthusiasm for creating their own infrastructure is very encouraging from the viewpoint of sustainability of this project beyond the HP grant.”

CRISIS MANAGEMENT

In addition to connecting remote field science and education sites, the HPWREN team has also been exploring ways in which low-power wireless technology can be applied to crisis- and incident-management applications. For instance, the HPWREN team recently participated in a UCSD demonstration of a temporary multimedia installation of seismic and visual instrumentation on the Coronado Bridge, which spans San Diego Bay. The installation provided wireless Internet Protocol (IP) access, using off-the-shelf, commercially available 802.11b technology, from the bridge to the UCSD campus, where participants examined incoming data and discussed the feasibility of a permanent networked monitoring system of this type. The system would specifically be used for UCSD research aimed at the crisis management community. The primary goals of the activity were to assess the feasibility and discuss the added value of such a networked monitoring system for UCSD research and involved government agencies.

Another recent HPWREN demonstration involved a group of local San Diego government agencies and the Space and Naval Warfare Systems Center San Diego (SSC San Diego). Specifically, the exercise examined the feasibility of incorporating real-time images and maps into incident management situations—utilizing the DARPA-funded ENCOMPASS software package.

The exercise took place at a National Guard Armory located in the Kearny Mesa area of San Diego. Creating a simulated scenario involving a chemical spill that required assistance from multiple agencies, SSC San Diego and HPWREN researchers demonstrated the use of high performance wireless communications and Web-based interfaces. The network linked SDSC to the California National Guard Armory via a nearby hospital rooftop relay. “This exercise allowed us to evaluate an Internet-based data sharing scheme, where
multiple agencies could view tailored perspectives of the same incident in real-time,” said Steve Murray of SSC San Diego’s Crisis/Consequence Management Initiative team.

“The drill also gave us a chance to recognize the importance of applications like crisis management within the world of wireless communications,” said Braun.

An additional crisis management demonstration included local California Department of Forestry and Fire Protection firefighters. While HPWREN researchers provided the wireless networking expertise, the firefighters coordinated logistics for the rapid-response exercise. The activity focused on the feasibility of using tripod-mounted antennas and other wireless technology as a means of high-speed Internet connectivity during rural fire incidents.

**DIRT ROADS AND HIGH SEAS**

One of HPWREN’s primary collaborators is another research project funded by the NSF: the Real-time Observatories, Applications, and Data management Network (ROADNet). Building upon existing HPWREN infrastructure, ROADNet allows researchers to collect, post, analyze, and retrieve data from seismic stations, lowland river watersheds, bridges, mountainous watersheds, observatories, ocean buoys, and research vessels.

“This prototype repository will allow us to observe many aspects of the environment in near-real-time and provide the data in a very flexible manner to a variety of scientists and decision-makers,” said John Orcutt, ROADNet principal investigator and acting dean of marine sciences and deputy director of SIO. “Historically, there have been many technological and institutional barriers to sharing environmental data. We are working to remove at least the technological barriers to accessing real-time environmental information.”

The HPWREN and ROADNet teams recently collaborated for a wireless demonstration involving SeaLab II, a United States Navy/SIO undersea habitat that operated August through October in 1965. First, the Navy Deep Submergence Unit (USN-DSU) brought their ship, the Kellie Chouest, and the remotely operated vehicle Scorpio to the waters off Scripps pier. “The purpose of our trip was to reconnoiter the SeaLab II site for a possible live broadcast and Webcast through UCSD,” said Kevin Hardy, director of SIO’s Centennial Program and organizer of the test dive. “In doing so, we brought the shared history of the U.S. Navy and Scripps alive, explored cooperative research opportunities between Scripps, UCSD, and the USN-DSU, and we tested new communications technologies embodied in the HPWREN link between the Kellie Chouest and Scripps pier. We succeeded in all three objectives.”

**NETWORK MEASUREMENT AND ANALYSIS**

With applications from field science and remote education to crisis management, one can easily understand why HPWREN offers the perfect testbed for network measurement and analysis. While field scientists and remote educators use the high-speed wireless network for their research and education applications, the HPWREN team busily conducts network measurement and analysis research on both backbone and access links.

The network researchers not only constantly evaluate performance via quantitative methods, but also work closely with the diverse group of HPWREN users, routinely asking them for qualitative feedback. If anything goes wrong, the HPWREN team is determined to get to the bottom of the problem and learn from it.

When strong winds recently damaged part of an antenna support and disrupted the HPWREN connection to Mount Laguna Observatory, Braun and his team didn’t waste any time going up the mountain, finding the problem, and fixing it—fast.

“We’re trying to build value with the network,” explained Braun. “We want users to see what the network offers them, and we want to find out if the network really is a quantum leap above the technology they used before HPWREN—even if it means climbing a few mountains.” —KMB, LM

**CRISIS-MANAGEMENT DEMONSTRATIONS**

In addition to connecting field science and remote education sites, the High Performance Wireless Research and Education Network (HPWREN) team is examining ways in which the technology may be applied to crisis-management situations in San Diego County. The National Science Foundation-funded HPWREN project utilizes the unlicensed microwave spectrum, operating at 2.4 GHz and 5.8 GHz.
Blue Titan and SDSC Developing Web Software

Blue Titan Software, Inc., the leading provider of origin-to-edge Web services networking solutions, recently announced it has been awarded a State of California Trade, Technology and Commerce Division of Science, Technology and Innovation/CommerceNet NGI grant award. The San Francisco-based company’s customers include Ford Motor Company, PricewaterhouseCoopers, and Lehman Brothers. The grant will be used primarily to fund joint efforts with the San Diego Supercomputer Center (SDSC) in the area of real-time Web services-oriented applications.

The NGI Grant Program is designed to nurture the development of business applications that leverage the power, speed, and reliability of the Next Generation Internet, and have a positive economic impact on the California economy. Using the NGI grant, Blue Titan will introduce its Web services networking fabric into SDSC’s data and knowledge systems initiative, accelerating the collection and dissemination of information supporting real-time mission-critical applications. For example, SDSC’s Information Integration Testbed (I2T) project has demonstrated an advanced Web services approach to providing seamless access to multi-agency government information.

“In a recent demonstration using Blue Titan in the I2T architecture, we have shown the ability to access and integrate information across multiple sources, including an Environmental Protection Agency website containing information on toxic locations,” said Chaitan Baru, co-director of SDSC’s Data and Knowledge Systems program. “As Grid Computing embraces Web services and the Open Grid Services Architecture (OGSA), Blue Titan is well-positioned to provide the infrastructure to support robust environments for executing Web services and providing quality of service and service level assurances.”

“Blue Titan’s Web services networking fabric provides the backbone for real-time, highly distributed mission-critical applications,” said Tony Darugar, chief architect of Blue Titan. “The NGI grant validates our unique architectural approach to Web services, and enables us to work closely with SDSC to deploy Web services networking in the real world, and implement emerging Web services standards around security and quality of service.”

Plant Genomics Website at SDSC Attracts Attention

PlantsT, a database that is part of the National Science Foundation’s Plant Genome Program, has celebrated its first anniversary, providing a variety of gene and protein searches and alignment tools, a lively news section, and even plant research movies.

“Researchers trolling for information on the sequences, functions, and phylogenetic relationships of these vital proteins should visit PlantsT, a year-old database of plant and yeast transporter genes run by the San Diego Supercomputer Center,” said a story about the site in the “NetWatch” section in the June 14 issue of Science.

Michael Gribskov, a senior staff scientist at SDSC, serves as co-PI for PlantsT, a database of the functional genomics of plant transporters. Gribskov also serves as co-PI for the Protein Kinase Resource, which earned a “NetWatch” mention in the January 25, 2002, issue of Science.

NetWatch highlights a science-related website each week. A search of the Science archives turned up 15 sites in the SDSC and NPACI domains, as well as projects led by SDSC researchers: these include sites within sds.edu, npaci.edu, and projects based at SDSC, such as National Laboratory for Applied Network Research, the Cooperative Association for Internet Data Analysis, and the Protein Data Bank. (v 6.14)

SDSC and Entropia Harnessing Desktop PCs

SDSC and Entropia, Inc., the leading provider of PC grid computing solutions, reported success with their partnership to provide leading, distributed-computing capabilities to computational science and engineering projects at SDSC. The center has deployed Entropia’s DCGrid, a powerful and cost-effective PC grid-computing platform that provides high-performance computing capabilities by aggregating the unused processing cycles of existing Windows-based PCs. SDSC plans to use the PC grid to enable new biological and molecular research.

“Grids are about using the right tool for the right job, and desktop PC grids have been very effective for many kinds of problems,” said Fran Berman, director of SDSC and the National Partnership for Advanced Computational Infrastructure (NPACI).

“Our partnership with Entropia allows us to complement our work in TeraGrid and NPACI with a strong peer-to-peer component. Entropia’s commitment to the Open Grid Services Architecture ensures that PC grids can be integrated with other high-performance computing resources to provide our users a comprehensive set of resources for their problems.”

SDSC’s first use of DCGrid applied the widely utilized GAMESS (General Atomic Electronic Structure Systems) software to accurately compute molecular structure and properties with the goal of populating databases for a wide variety of purposes, including analysis of bioterrorism agents such as anthrax and smallpox. No source code changes to GAMESS were needed for deployment on DCGrid.

DCGrid represents a unique type of computing resource for SDSC. With DCGrid, SDSC researchers were able to automate a large quantity of strategic, high-throughput calculations, which allowed them to use the more accurate GAMESS code on a wider range of molecular systems than would have been possible on other platforms. “With DCGrid we are able to exploit the strengths and inherent resources of PC grids to open new horizons while still capitalizing on our existing high-end resources,” said Kim Baldridge, program director of Computational Science at SDSC. (v 6.17)

Lifemapper Produces Species Distribution Maps and Models

The Biodiversity Research Center at the University of Kansas, a former NPACI partner, has released the beta version of Lifemapper. Running as a screensaver, the Lifemapper software computes, maps, and provides knowledge of where Earth’s species of plants and animals live, where they could potentially live, and where and how species could spread across different regions of the world.
An international multi-disciplinary project, Lifemapper harvests the 250-year legacy of geographical information associated with the specimens of life on Earth that is stored in the world's natural history museums. Using this information as input data for predictive modeling of species distributions, Lifemapper can help today's researchers to better understand and conserve the Earth's biological diversity. Beyond research, by delivering species distribution maps to the desktops of PC users worldwide, Lifemapper is highlighting the role of natural history museums in documenting biodiversity.

“By donating their spare cycles, people are directly helping to assemble a powerful, predictive digital library of Earth’s biological diversity,” said Jim Beach, assistant director for Informatics at the Biodiversity Research Center at the University of Kansas. (p 6.14)

NSF Awards $3.28 Million to NLANR

The National Science Foundation (NSF) will provide $3.28 million to the National Laboratory for Applied Network Research (NLANR) Measurement and Network Analysis (MNA) group for work to be conducted over the next three years. Since NLANR's founding in 1997, the program’s measurement and analysis activities have benefited thousands of researchers across the nation by improving the performance of the networks they depend on for high-speed access to supercomputing resources and online data libraries.

NLANR is a collaboration that provides technical, engineering, and traffic-analysis support for NSF's High Performance Connections sites (universities, supercomputer centers, and government research laboratories) and the ultrafast networks that link them (the “next-generation Internets”). NLANR teams help administrators maintain these networks and provide scientists, engineers, and scholars at research institutions in the United States with tools needed to use the high-speed data links efficiently.

The mission of NLANR's Measurement and Network Analysis group is to study the operation of these networks, measuring the flow of traffic and analyzing performance issues, to better understand the theoretical and practical behavior of the systems and to deliver maximum end-to-end performance to users.

“We have been building an extensive network analysis infrastructure,” said Hans-Werner Braun, principal investigator on the effort. “Traffic data collection devices have been installed at approximately 150 locations, most of them in the United States but with a handful in foreign countries as well. The infrastructure also encompasses a growing collection of measurement data and analyses, a body of tools and methods, and avenues for sharing information. The ‘human infrastructure’ is equally important—we are engaging in many collaborations with other researchers, in this country and in others, and within and outside of the network measurement community. (p 6.16)

SDSC's Rozanne Steckler Receives 'Spirirt of Preuss' Award

Rozanne Steckler, director of education at SDSC, was given the annual “Spirit of Preuss” award by Preuss School Principal Doris Alvarez at an Awards Assembly held in July. She was acknowledged for her work involving members of the university community in programs that engage middle school and high school students in learning about science and using computers.

The Preuss School, a public middle and high school located on the UC San Diego campus, is dedicated to providing an intensive college preparatory education for motivated low-income students who will become the first in their families to graduate from college. Using a high degree of personal attention and parental involvement, Preuss prepares students to succeed at the University of California and other colleges.

Steckler assisted the school in setting up a system of remote monitoring of student teachers, supplied teachers from the SDSC Education Department for a multimedia class, and brought the Science and Technology educational modules developed at SDSC into the Preuss curriculum. Previous recipients of the award were UC San Diego Chancellor Robert Dynes, Preuss School benefactors Peter and Peggy Preuss, and Cecil Lytle, provost of Thurgood Marshall College at UCSD. (p 6.15)

Subscribe to Online or ENVISION: www.npaci.edu/Press/subscriptions.html
This model of a fluorescein antigen bound to an immunoglobulin protein illustrates how a tool called QMView models and displays important concepts in biochemistry. QMView (quantum mechanical view) is a visualization and analysis program that sits on top of a computational chemistry code called GAMESS (General Atomic and Molecular Electronic Structure System). QMView runs on UNIX workstations supporting OpenGL, and has been recently ported to PCs and Mac OS X. It includes molecular building, visualization and analysis of results computed from GAMESS, new methods for distributed computing (including the DICE interface), and an improved graphical interface for users. The QMView tool, which was developed by computational chemists Kim Baldridge and Jerry Greenberg at the San Diego Supercomputer Center, provides direct access to Protein Data Bank files and enables display of primary protein structure (blue and red), and secondary structures such as ribbons (purple and yellow). Computed electrostatic contour results for this ligand were derived from quantum chemical computations for the ligand (green) with its associated electrostatic cloud (orange) embedded in the protein.