NCSA is a key partner in the National Science Foundation's TeraGrid project, a $100-million effort to offer researchers remote access to some of the fastest unclassified supercomputers as well as an unparalleled array of visualization tools, application software, sensors and instruments, and mass storage devices. NCSA also leads the effort to develop a secure national cyberinfrastructure through the National Center for Advanced Secure Systems Research, a project funded by the Office of Naval Research.

The center leaves its mark through the development of networking, visualization, storage, data management, data mining, and collaboration software as well. The prime example of this influence is NCSA Mosaic, which was the first graphical Web browser widely available to the general public. NCSA visualizations, meanwhile, have been a part of productions by the likes of PBS's NOVA and the Discovery Channel.

Major support for NCSA is provided by the National Science Foundation. Additional funding comes from the state of Illinois, industrial partners, and other federal agencies.
<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
<th>Title</th>
<th>by</th>
</tr>
</thead>
<tbody>
<tr>
<td>02</td>
<td>The Director's View</td>
<td>Building the Third Pillar of Science</td>
<td>Rob Pennington, NCSA Interim Director</td>
</tr>
<tr>
<td>04</td>
<td>Q&amp;A</td>
<td>Managing Complexity</td>
<td>Kevin Rollins, Dell Inc.</td>
</tr>
<tr>
<td>06</td>
<td>Forecasting the fold</td>
<td>Efforts to understand protein folding and unfolding could lead to treatments for a wide range of misfolding diseases.</td>
<td>Katherine A. Caponi</td>
</tr>
<tr>
<td>10</td>
<td>Blades of gas (turbines)</td>
<td>Fluid dynamics experts take gas turbines for a spin on NCSA’s TeraGrid cluster, concentrating on features that help keep the blades cool.</td>
<td>J. William Bell</td>
</tr>
<tr>
<td>14</td>
<td>The big picture</td>
<td>The Palomar-Quest Survey's nightly snapshots of huge regions of the sky might help answer the question of exactly what's going on up there.</td>
<td>Kathleen Ricker</td>
</tr>
<tr>
<td>18</td>
<td>Mini motors</td>
<td>A team of researchers tries to unlock the mechanism of the proteins that drive muscle movement and other essential biological functions, knowledge that could lead to better treatments for high blood pressure and heart disease.</td>
<td>Trish Barker</td>
</tr>
<tr>
<td>22</td>
<td>Giving sensors an edge</td>
<td>Coupling tiny sensors with software-defined radio technology will provide scientists with the power and flexibility to gather more data.</td>
<td>Trish Barker</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kevin Rollins, president and chief operating officer, Dell Inc. p.04</td>
</tr>
</tbody>
</table>
Building the Third Pillar of Science
The recent report by the federal High-End Computing Revitalization Task Force put it as clearly as anyone could: “Computer modeling and simulation of physical phenomena and engineered systems have become widely recognized as the third pillar of science and technology—sharing equal billing with theory and experiment.”

As the minds behind the cyberinfrastructure that supports this modeling and simulation, we know that’s true and has been for a while now. It seems every engineer using a wind tunnel collaborates with a computational fluid dynamics researcher. Chemists recognize that some processes take place on too small or rapid a scale to be captured at the bench. And the data flowing from telescopes and other powerful instruments simply cannot be handled without high-performance computing support.

We also know that we have to make this pillar strong. We must remember that what it upholds—the progress of human knowledge—is more important than what it is. The fastest computer imaginable doesn't, in itself, improve the world, but what it enables is beyond value. An elegantly designed computing system pales when compared to a cured disease. A better connected society trumps the network and software it’s built on. And that is as it should be. Our own achievements are secondary to the achievements we make possible.

Recognizing this, NCSA recently instituted a plan to encourage a more service-oriented cyberinfrastructure. Our users have always been a crucial part of our development and deployment activities. But it’s time to redouble our efforts in order to ensure long-term success for the center and for what we build.

The first step toward a service-oriented cyberinfrastructure is an integrated and accessible system of hardware and software components that span collaborating sites. A new Cyberinfrastructure Partnership between NCSA and the San Diego Supercomputer Center will meet this task. It will also create a common support mechanism for systems at both centers, relying on our combined strengths in middleware infrastructure, user engagement, training and outreach, visualization, resource evaluation, and security. Data management, analysis, mining, and visualization are rapidly becoming as crucial to research as computationally intensive simulation. Thus, we will continue to pay special attention to data-driven science. The National Laboratory for Advanced Data Research, a joint NCSA/SDSC project, will create a cross-center common catalog of data resources and services.

Integration isn’t enough. We also need an in-depth community engagement process. This means working with researchers and groups of researchers—whether they’re from the traditional sciences or from the humanities or social sciences—to help us blueprint systems and design prototypes. That way, we address specific research objectives within the disciplines.

Instead of asking users to fit a square peg into a round hole, we want them to help us design the peg in the first place. Accomplishing this goal in an environment that is defined by its changing nature presents an uncommon challenge. A team approach is vital since no single person, or even institution, has the depth and breadth to undertake such a complex enterprise alone. This approach will allow us to better understand the background and status of our partnering communities, formally define their expectations, and educate them on the state and capabilities of the cyberinfrastructure we build together.

Truly useful cyberinfrastructure requires substantially more than fast computers, leading-edge software, and the best networks. We are obliged to maintain a robust, on-going relationship with the people who are going to use it.

A pillar’s strength is typically in its stubbornness, its unchanging nature in the face of time and its environment. But our users can’t afford inflexibility. It’s incumbent upon us to provide something more. It’s incumbent upon us to build a cyberinfrastructure that reflects our users’ needs, removes burdensome technical issues from their path, and liberates their most provocative and promising ideas.

Rob Pennington
Interim Director
National Center for Supercomputing Applications
Earlier this year, Kevin Rollins, president and chief operating officer of Dell Inc., presented NCSA with the company's Centers for Research Excellence Award. The award recognized NCSA's successful deployment of Tungsten, a 15-teraflop high-performance computing cluster built on 1,474 Dell PowerEdge servers.

Given at the NCSA Private Sector Program's annual meeting, it was preceded by an informal question and answer session with Rollins. NCSA's Interim Director Rob Pennington kicked off the questioning, then turned it over to the audience.

Q: What is Dell's attitude toward collaborative R&D? Is there a sense of "not invented here," or is Dell open to a shared development model?

A: We're somewhat unique in the industry. We spend about a half a billion dollars a year in internal R&D and have over 3,500 engineers around the globe working on our products at any point in time. But we believe that the industry is really too large for one company to try to control all of the R&D. Our model is to use our R&D dollars to integrate the best technologies we can find on behalf of customers so that they have an easier time using those technologies. We focus on promoting open standards so that we can find those best technologies, make them available to the mass computing market, and then bring the very best technologies to bear. We believe in that collaborative R&D model and have numerous partners that allow us to bring the best technology to customers, rather than proprietary-only technology.

Q: How do you maintain the balance between the internal technology development and the external development?

A: The internal growth really comes from having exciting projects to work on. As engineers will know, if there's something very exciting on the forefront, they want to work on that. And so we make sure that our very best engineers get the first crack at the most exciting, new, cutting-edge technologies, whether they be ours or whether they be technology from outside that we're now pulling in.
Q: Dell's stance on the market is very wide. You go all the way from some of the biggest machines on the top 500 list of supercomputers down to laptops and consumer-level products. Do you see Dell maintaining that very wide stance or is that going to shift?

A: Only about 15 percent of our worldwide volume is in the consumer product arena. Most of the R&D and the access that we have to technology flows into data centers, large corporate research, government, and education environments, and we'll continue to do that. Those technologies over time then flow into consumer products quite often, and there are certainly some technologies like wireless which have application not so much in the data center but with the client at the edge of the IT environment. We pursue those, too. But where we are today is really at the core of our strategic thrust and where we spend over 50, 60 percent of our R&D dollars.

Q: We all know there's a major problem in Internet security. Viruses and worms, intruders of various kinds, hackers. Is Dell concerned that to prevent intruders and to prevent viruses that the hardware now being sold will have to be redesigned?

A: One of the benefits of IT and of software and of the Internet is it's been an open architecture so it's something we've all had access to and been able to use. But obviously, when you operate in an open environment like that, you allow for the small minority of the population who wants to foul things up to participate too.

We have not spent as much time on trying to get a hardware solution because most of the viruses are software related and so the solution is generally software related as well. What we can do is find our best preventive measures. I'm afraid that the best solution is going to be to continue to outrun it and continue to upgrade, continue to check and to be as vigilant as we can be, because I think in an open society we're going to have this with us for a while.

Q: What do you consider the most critical unmet needs in the corporate marketplace?

A: With technology evolving so rapidly, you're constantly on the cutting edge if you want the best and the highest performance. And that cutting edge is often a little ragged. And so the notion of service and support, of stability, is a constant challenge for us. Security, manageability is always a challenge.

Right now what we're seeing and what we're looking at and working on with our business partners is how do we provide for more manageability of heterogeneous environments in the data center, more manageability across multiple platforms for storage, for networking, for server farms. We'd like to be able to get to the point of having one user interface for all of those platforms so you could manage those either remotely or on site, rather than having to have a management system for every one of those various technologies within an institution or corporation. That's one of the next great milestones within technology, the notion of how you manage the complexity.

Q: Lately the outsourcing issue has been in the news. From your view, is it possible to make this outsourcing issue be not such an either/or situation but to be a win/win?

A: The fastest growth markets for us in the world now are outside the U.S. We're still growing in the U.S., but a third of the world's population lives in China, and they don't have computers in China to a great extent. From a growth opportunity over the next five or six years, the China marketplace is going to be very, very large. What we've concluded is it would be unlikely for U.S. technology to be able to be sold into China and India if we had no employees there. As we bring jobs into those countries, what happens is the standard of living improves and the populace there has money to spend. What do they spend their money on? They spend it on U.S. technology.

We believe that kind of technology growth, particularly with the improvement of jobs around the world, is going to do nothing more than improve the overall capability and the ability to sell. I personally believe it will result in more jobs of a higher quality in the United States. It certainly has in our case.

Q: Would you address Dell's efforts to give back to the community in terms of education, especially the K-12 range?

A: What we're doing in a direct way is starting early through a program called TechKnow. As a program, it's as simple as providing computers to grade school kids and teaching them how to take them apart and how to use them. At the end of the training, and their parents have to be involved, they get to keep the computer.

This is where we can get our youngest minds more involved with technology sooner. And if it starts there, then by the time they get to college we'll have great engineers because they'll be so comfortable with this technology that they'll do great things for all of us.
Forecasting the fold
Efforts to understand protein folding and unfolding could lead to treatments for a wide range of misfolding diseases.

Proteins are the molecules of life. They form the framework of our muscular systems and other tissues, serve as antibodies, function as hormones, and perform many other vital tasks in the human body. Proteins are also the enzymes that carry out most of the chemical reactions in the human body.

Each protein comes from a different gene sequence. The sequence determines the structure of each individual protein. There are an astronomical number of three-dimensional configurations that proteins can form—scientists estimate that there may be more possible structures than the number of grains of sand on a beach. Because proteins derive their functionality from their structure, which is based on their gene sequence, each gene sequence produces a different kind of protein destined to complete a unique task.

However, surprisingly little is known about the physical structures of proteins and how they are created from those gene sequences. Scientists know that proteins undergo processes called folding and unfolding through which they are built and torn down, but they don’t know how or why naturally occurring proteins consistently display a particular shape, or native state, after the folding process.

Finding out why proteins fold into a consistent native state is important because occasionally the natural folding process breaks down and the proteins form the wrong structures. When that happens, the proteins’ functions suffer—DNA may not replicate properly or drugs sporting proteins as main components may not work. Protein misfolding diseases, such as bovine spongiform encephalopathy (mad cow disease), Parkinson’s, or Alzheimer’s, can occur.

To prevent proteins from misfolding and causing serious problems, scientists need to know more about why they fold and unfold the way they do. If researchers could discover what stimuli cause certain structures to appear, new solutions to protein misfolding diseases, such as structure-based drugs and a better understanding of gene mutations, could be possible.

Carlos Simmerling, a professor of biochemistry at Stony Brook University in New York, and a team of students and colleagues are working to understand why proteins fold and unfold and what happens during the process. Using NCSA’s Tungsten and Platinum supercomputing clusters, they are simulating the folding and unfolding processes of small proteins and comparing their findings to experimental results from other research groups.
Simulating the folding process

How proteins fold and unfold has long been an issue of debate. Simmerling says, "Nobody knows what the unfolded protein looks like." This is due mostly to the speed with which a protein transitions between folded states and the fact that there aren't many unfolded proteins to observe in normal physiological conditions. In fact, the idea that many naturally occurring proteins fold quickly and reliably to their native state despite an enormous number of possible structural configurations is a paradox that has confounded researchers for decades. If each protein worked its way through all the possible variations in its structure before returning to its native state, folding could take billions of years. Yet many proteins can unfold and fold back into their original structure in seconds or less.

In their quest to understand the folding and unfolding process and predict protein structure from amino acid sequences, Simmerling and his research group began by looking at the smallest protein known with normal folding properties. The 20 amino acid sequence Tryptophan cage, or Trp cage, was isolated and developed by Niels Anderson, professor of chemistry and biomolecular structure and design, and his colleagues at the University of Washington at Seattle. Trp cage is a simple version of a protein derived from Gila monster saliva that could have potential for treating misfolding diseases. Other researchers in the field have used simpler models of Trp cage to predict the final structure of the protein, but none of them have previously simulated the full folding process.

The researchers carried out a series of molecular dynamics simulations, with an allocation of 960,000 hours on NCSA's Tungsten and Platinum clusters to look at Trp cage and other small proteins. They predicted the structure of Trp cage from its amino acid sequence and compared their computer simulations to Anderson's results from physical experiments in which Trp cage was folded in a lab and then measured to plot where individual atoms are naturally positioned. The level of accuracy in the groups' predictions created a stir in the field of biochemistry when they published their findings in the Journal of the American Chemical Society in 2002 because no researchers had ever before simulated protein folding in such a fine level of detail. The final structure of the simulated Trp cage was virtually indistinguishable from the experimental data.

Using the supercomputers, the group can now run multiple simulations at the same time. Simmerling says that the advantage of running multiple simulations in parallel is that the simulations can "talk" to each other, communicating information that can be shared and cutting down on computational time. This benefit enables the researchers to study slightly larger molecules on which there are existing physical experiments, rather than the smaller mini-proteins in very simplified water environments they were previously limited to in simulation.

The new simulations, demonstrating how Trp cage folds when submerged in thousands of individual water molecules, allow the researchers learn what structure emerges from the protein in a particular environment and how the protein forms that structure in specific conditions. The team found that because the protein has a hydrophobic core it folds rapidly in water to protect its inner atoms from the surrounding water molecules.

Simulating what causes protein folding may enable researchers to put that knowledge to use. For example, Trp cage's transformation to a particularly stable structure is consistently sparked by submerging it in water at a specific temperature threshold.
information could be important to researchers who design drugs, helping them produce structure-based drugs that are more stable in water-based solutions at specific temperatures, such as normal human body temperature.

Cracking the code
Simmerling says the work of his group and others also will help make resources such as genome databases more readily useable. A genome represents all the DNA sequences for an organism—genetic codes responsible for the sequence of amino acids that constitute proteins. While genome databases currently provide information about the sequences of amino acids, they don’t provide information about the structures of the proteins. Since the functionality of the proteins is determined by the configurations they form upon folding, the lack of structural information makes genomic sequencing difficult to apply to real-world medicine.

A sequenced genome is “like having a book written in another language,” Simmerling says. “You can see patterns, but may have no idea of their meaning.” The simulations produced by Simmerling and his colleagues on NCSA supercomputers may someday allow researchers to translate all those books of genomes and begin to understand why certain amino acid patterns produce diseases. This could lead to the design of structure-based drugs to treat these diseases.

This research is supported by the National Institutes of Health, the Research Corporation, and the National Science Foundation.

Access Online: http://access.ncsa.uiuc.edu/CoverStories/proteins/
For further information: http://morita.chem.sunysb.edu/~carlos/

Team Members
Stony Brook University
Viktor Horak
Melinda Layten
Carlos Simmerling
Bentley Strockbine

University of Florida
Adrian Rollberg

University of Washington
Niels Andersen
Blades of gas (turbines)

Snapshot of the heat transferred from the wall (colors) and turbulence (white isosurfaces) in a rotating bend.
Fluid dynamics experts take gas turbines for a spin on NCSA's TeraGrid cluster, concentrating on features that help keep the blades cool.

Only someone like Danesh Tafti, who's been taking advantage of NCSA's computers for almost 15 years, would consider his work of three years ago starting out small. But consider these details about his models of the airflow around tiny ribs inside gas turbine blades.

Back then, he was using 32 processors per run. Now, he's averaging about 150. Then, he was looking at three ribs. Now, he's up to 10. Then, the models were covered in a grid with two million individual zones. Now, we're talking about 10 or 20 million. Then, the calculations were taking a month on a Pentium-based cluster. Now, he blasts through much larger calculations in a week on NCSA's Itanium-based TeraGrid cluster.

"We've made a really big jump," says Tafti, who is an associate professor of mechanical engineering at Virginia Tech. They aren't big runs for the sake of big runs, though. "We've taken prediction technology in gas turbines a step higher—maybe more than a step. This is a big jump in resolution and quality." He and graduate students Evan Sewall, Aroon Viswanathan, and Samer Abdel-Wahab work with engineers at General Electric and the South Carolina Institute for Energy Studies at Clemson University, which is supported by the Department of Energy. Together, they make sure these improved simulations translate into better turbines in the real world.

"While we do use massively parallel computations of external flows and in other areas of the engine, very little simulation is currently done with these advanced techniques when it comes to internal turbine passages," according to Andy Smith, a mechanical engineer for GE Global Research's fluid mechanics lab in Niskayuna, NY. "But turbine seizure and failure happen at the local level," so the degree of resolution and quality delivered by the team is highly valuable.
Birth of the cool

The temperature of the gas driving a turbine engine often exceeds 1,300 degrees Celsius. That figure is headed nowhere but northward because the hotter the gas the more efficient and powerful the turbine can be. Higher temperatures, however, cause additional wear and tear. They can even destroy the turbine entirely. For example, if the temperature of the blade is increased by less than 10 percent, according to many estimates, the engine will need to be overhauled twice as often.

Systems for cooling the blades are a crucial part of the turbine design process. Typically, cool air bleeds from a compressor into channels that snake back and forth through a blade’s hollow interior. Through convection, heat from the metal blade is transferred to the air. Placid airflow produces a relatively low heat transfer rate: a turbulent flow boosts the rate. Thus, designers typically stud the channel walls with turbulence inducers such as pock marks and tiny ribs.

Tafti likens the process to stirring a cup of tea. “With your spoon, you generate random motion in the fluid and mix in the milk [more rapidly and thoroughly]. Ribs increase the mixing and the ability of the coolant to conduct heat away from the blade surface.”

Turbulence isn’t all milk and honey, though. The churning air can accelerate or unexpectedly slow down—or sometimes break into multiple flows. Meanwhile, “dead regions” can develop. These areas have the same already-heated air passing them again and again, reducing the heat transfer from the blade.

Friction is also a concern. Created as the air moves past the turbulence inducers, it causes the air pressure to drop. Pressure in the blade has to remain higher than the pressure outside. Otherwise, the 1,300 degree gas shooting through the turbine pushes into the blade instead of coolant inside the blade benignly streaming out.

You spin me right round

As with the pressure difference between the inside and outside of the blade, there’s always a balance. Whether it’s between performance and wear or between the benefit of turbulence induced and the downside of friction created, the balance is critical. The team explores the finest points of these features. Using GenIDLEST, a fluid dynamics code that Tafti has been developing for more than a decade, they model a series of ribs in a channel and the air that flows past them. Currently, they focus on nine or 10 ribs, watching the air movement both as it develops and once it has settled into a stable, though still turbulent, flow.

Altering the number, orientation, and design of the ribs shows them differences that crop up in pressure at various points in the channel, average pressure drops over time, and heat transfer rates. It also reveals split flows and dead regions. With this information in hand, designers can refine plans and look for ways to improve efficiency without causing new problems.

Every change influences every feature. A rib that has rounded upper corners, for example, eliminates an unwanted recirculation area behind the rib. Unfortunately, it also reduces the heat transfer rate.

“You spin me right round” Tafti says. Some of the team’s most recent models have further complicated matters by adding Coriolis forces and centrifugal buoyancy. These forces account for the rotating of the blade and the impact that action has on the airflow within the blade. “It’s extremely difficult to determine these things in the lab,” Tafti says. A large rig is spun at high velocity, and very exacting measurements are taken. “Few can really do this,” he says. And even those who can are only able to gauge the aggregate effect, not each force’s separate impact.

The Coriolis force is of particular interest to designers because it tends to repress turbulence and thus heat transfer, especially at the bends in the serpentine channels. “One of the purposes of the bends is simply to channel the flow throughout the blades so it can act as as efficient a heat exchanger as possible. But the act of turning the flow, in and of itself, promotes turbulence, which is a good thing,” GE’s Smith says.
Initial results of these simulations show that the Coriolis force can increase heat transfer by 50 to 60 percent on one side of the channel but can decrease heat transfer by as much as 50 percent on the other side—when compared to simulations in which the blade's rotation is not accounted for. Centrifugal force, meanwhile, can complement or oppose the influence of the Coriolis force, depending on whether the flow is moving toward or away from the turbine's center.

These results, along with others that discuss issues like rib orientation, were presented in June at the American Society of Mechanical Engineers' Turbo Expo 2004. They can be found in the conference's recently published proceedings.

This research is supported by the Office of Fossil Energy at the U.S. Department of Energy's National Energy Technology Laboratory and the South Carolina Institute for Energy Studies.

Access Online: http://access.ncsa.uiuc.edu/CoverStories/blades/
For further information: http://www.hpcfd.me.vt.edu/
http://www.clemson.edu/scies/
http://www.crd.ge.com/

Team members
Samer Abdel-Wahab  Danesh Tafti
Evan Sewall  Aroon Viswanathan
Andy Smith

Wall temperature in a rotating duct. Flow streamlines show recirculation behind the ribs.

Evan Sewall, Danesh Tafti, and Aroon Viswanathan, Virginia Tech.
The big picture

The Samuel Oschin telescope, a 48-inch aperture wide-field Schmidt telescope to which the Palomar-Quest survey camera is currently attached.

Photo credit: Palomo/Caltech.
The Palomar-Quest Survey's nightly snapshots of huge regions of the sky might help answer the question of exactly what's going on up there.

NCASA and quasars—those distant, tantalizing, extremely bright objects in the night sky, thought to be powered by super-massive black holes—go way, way back.

"Many of the simulations which have given credence to the standard model of quasars and active galactic nuclei were done on NCASA systems," says Robert Brunner, assistant professor in the department of astronomy at the University of Illinois at Urbana-Champaign and a research scientist at NCASA.

By looking at what kind of light quasars emit, Brunner and other researchers interested in these objects can determine not only how far away they are but also the physics that governs their extreme luminosity. Furthermore, because they are so bright, quasars can be seen to great distances, allowing astronomers to use them to probe both the physics of the early universe and the formation and evolution of galaxies.

Now Brunner, who leads NCASA's Laboratory for Cosmological Data Mining, is overseeing the processing and storing of data from the Palomar-Quest Survey (PQS), a sky survey that, by imaging the same large region of sky night after night, might help identify gradual changes in the fabric of the universe.

*Quasar 3C273, the nearest and brightest of all observed quasars, as perceived by the 4-meter Mayall telescope on Kitt Peak in Arizona. Spectroscopy studies suggest that the quasar appears to be receding from us at a speed of 48,000 km/hr. By imaging the sky with another such wide-aperture telescope, the Palomar-Quest survey is expected to be able to identify more quasars and help astronomers understand how the structure of the universe may be changing. Photo credit: NOAO/AURA/NSF.*
A wider view
The Palomar-Quest Survey is a collaboration involving the California Institute of Technology, Yale University, Indiana University, and UIUC/NCSA. Charles Baltay, professor of physics at Yale, co-leads the PQS project and is particularly interested in finding and tracking objects that might change dramatically over a period of weeks or years—a characteristic he identifies as “time variability.” Of particular interest, says Baltay, are very high redshift quasars, “which tell us about the early history of the universe,” and supernovae, which “tell you whether the universe is expanding or contracting, and if it’s expanding, whether the expansion is getting slower or speeding up.”

The original Quest survey was a collaboration of several North and South American universities in which a camera mounted on the 1.0-meter Schmidt telescope at the Llano del Hato National Astronomical Observatory in Venezuela was used to take snapshots of the equatorial sky. This camera consisted of only 16 CCDs, or charge-coupled devices, the solid-state silicon chips inside all digital cameras that convert light into electrical impulses that can be recorded to disk as data.

For the PQS, a larger camera, designed and constructed by Yale and Indiana University, has been mounted on the 48-inch Samuel Oschin Schmidt telescope at the Palomar Observatory, located near San Diego and run by the California Institute of Technology. The new camera at Palomar contains 112 CCDs, sufficient to cover a much larger region of the sky. The actual observations are controlled from Yale. However, the drift-scan survey data are processed and stored at NCSA, chiefly by Adam Rengstorf, a postdoctoral researcher in Brunner’s group at UIUC. Says Rengstorf, “Every two minutes and 20 seconds, we get 112 images off the camera—and that continues for up to nine hours a night.” The resulting data flow, says Rengstorf, is just under 7.4 gigabytes an hour.

The next day, the raw data are bundled and transferred to NCSA’s UniTree mass storage system, after which the information is processed in a complex series of steps. Earlier this year, this was done using 32 processors on NCSA’s Platinum cluster. This processing pipeline is being converted to operate on generic grid infrastructure, eventually allowing the processing to utilize the TeraGrid. “You have to unpack all the data, do the bookkeeping and organization, and remove the instrumental signatures from the data,” explains Rengstorf. This includes “flattening” the image to compensate for the possibility that a given CCD does not respond to light uniformly across its surface. “Then you have to detect objects, and once you detect them you have to figure out where they are on the different chips.”

After the positions of these objects are calculated, they are mapped according to the U.S. Naval Observatory astrometric catalogue, a very accurate map of the known sky. The brightness of each star is measured. All of this information is written to an output catalog at NCSA and returned to mass storage for use by members of the collaboration.

Timing is everything
The Palomar-Quest Survey’s emphasis on time variability is an important advance for observational astronomy. It requires imaging a large region of the sky repeatedly over the course of months or years while scanning for changes. This in turn requires using a Schmidt telescope, which has a much wider field of view than most telescopes in use today—including Hubble. The 48-inch Samuel Oschin telescope at Palomar is the second-largest Schmidt telescope in the world.

“People have studied quasar variability before,” says Brunner. “But it’s usually to monitor a small number of sources for 10-20 years, or to match new data against much older data. Actually going out and surveying significant fractions of the entire sky—a tenth, a fifth, a quarter—and doing it over and over again for a number of years—that’s really a new thing.”

Brunner says that the process has worked well from the
very first data transmission in April of this year. “We’re an invisible part of the infrastructure now,” he says. And for PQS, invisibility—running smoothly and uneventfully without a hitch—is crucial. “If you want to catch things which are changing in time,” says Baltay, “you don’t have the leisure of waiting months and months before you process your data, because by that time the thing is gone.” He hopes that in the future data transmission can be speeded up even more dramatically through automation, so that, for example, raw data received at 8:10 p.m. can be sent, processed, and archived 20 minutes later.

Using data from the new camera that was processed at Caltech led, for example, to the discovery in November 2003 of the planetoid Sedna, the most distant solar system object ever observed and the largest Kuiper Belt object after Pluto. More recently, Palomar-Quest Survey data processed at NCSA yielded a potentially valuable cosmological discovery: a relatively rare, redshift 4.07 quasar that was discovered in follow-up, time-critical observations by a team led by Caltech professor George Djorgovski. While Brunner points out that more distant quasars have been identified, he believes that the real value of the find is that it demonstrates what PQS can do: “It’s proof that we’re going to be able to systematically find high redshift quasars using the Palomar-Quest Survey.”

Undiscovered countries
Brunner and Rengstorf are excited about the scientific potential of studying the time variability of quasars and other astrophysical objects. Brunner, whose interest is in cosmological data mining, expects soon to be able to tackle the problem of characterizing quasars according to their variability using data from PQS in conjunction with spectral analysis from other sources. “Optically, quasars and stars might look the same, just little points of light,” says Rengstorf, “but when you take a full spectra, quasars and stars are very distinct. The spectra show you the underlying physics of what’s going on and lets you quantify more precisely how far away the quasar is, what its redshift is.”

PQS operations at NCSA have been funded in part by the National Science Foundation and NASA.

Access Online: http://access.ncsa.uiuc.edu/CoverStories/quest2/
For further information: http://astronomy.ncsa.uiuc.edu/quest2/index.php

Team members
NCSA/UIUC:
Robert Brunner  Michael Remijian
Andrey Gusev  Adam Rengstorf
Britt Lundgren  Nicholas Waggoner
Adam Myers  Brian Wilhite
Nikita Prokopev

Palomar QUEST Survey Principal Investigators:
Charles Baltay, Yale University
S. George Djorgovski, Caltech
Mini motors
A team of researchers tries to unlock the mechanism of the proteins that drive muscle movement and other essential biological functions, knowledge that could lead to better treatments for high blood pressure and heart disease.

Listen to your heart beat. Take a breath. Flex a muscle. All of these actions are the responsibility of myosin, the essential motor protein that drives muscle contraction.

"As our heart beats, as we breathe, as we walk around, we're using myosin," explains Washington State University mathematics professor Edward Pate.

Myosin controls the skeletal muscles responsible for every move you make, the cardiac muscle that pumps your blood, and the smooth muscles that control blood pressure, push nutrients and waste through the intestines, and drive uterine contractions. Myosin's close relative, kinesin, carries out vital functions in non-muscle cells, shepherding chromosomes to their proper places during cell division, for example, and ferrying proteins, lipids, and other molecules from neural cell bodies to distant axons.

Scientists understand the basic outlines of how myosin and kinesin do their jobs, but certain crucial details have yet to be added to the picture. Pate and his collaborators, Roger Cooke at the University of California, San Francisco, and Todd Minehardt, a former University of Colorado, Denver, professor now in the private sector, are using a combination of experimentation and simulation on NCSA's computational systems to zero in on those details. Their work focuses primarily on myosin, but because the two proteins are so closely related, insights about the muscle cell protein can be generalized to kinesin as well.

**Converting chemical energy to motion**

Myosin is a large, asymmetrical molecule; it has a long tail and two globular heads.

Myosin works in tandem with actin, a filamentous protein. At the start of the kinetic cycle, the two are bound tightly together. This tightly bound state is referred to as the rigor complex; in fact, after death, when muscles are locked in rigor mortis, this rigidity is due to the tight locking of actin and myosin.

The tightly bound proteins are awaiting the binding of the energy-rich nucleotide ATP, which is composed of a ribose sugar and three phosphate groups. When ATP comes on the scene, myosin grabs onto the energy source and releases its grip on actin.

---

*Illustration of a myosin molecule, showing its long tail and two globular heads.*

*Actin and myosin in their bound state. These two proteins are tightly bound at the start of the kinetic cycle.*
When this attached ATP molecule is hydrolyzed, the myosin bends into its “pre-power stroke” state, poised to move into action. The hydrolysis of ATP releases a burst of chemical energy, myosin rebinds to actin, and that burst of chemical energy is translated into motion as the myosin filament and the actin filament slide past one another. The products of ATP hydrolysis, ADP and phosphate, are released during this process.

“Basically, actin is like the railroad track, and myosin travels along actin,” says Minehardt, who has been studying the phenomenon since 1999, when he was a UCSF post-doc.

While the main strokes of the process are understood, the exact mechanism that converts the chemical energy released by ATP into motion is not known.

“How do motor proteins take chemical energy and turn it into mechanical force?” Minehardt asks.

Cooke, Minehardt, and Pate are tackling the question of how the proteins produce force and motion using both laboratory experiments and computational simulations, with each technique filling in part of the picture.

Probing the nucleotide site

Obtaining an X-ray crystal structure that shows myosin bound to actin and generating force is virtually impossible because the reaction occurs too quickly to be captured. So Cooke and Pate use electron paramagnetic resonance (EPR) probes to infer what is occurring at the myosin nucleotide binding site.

A variety of ATP analogs, some of which can bind to myosin without being consumed as fuel, are modified with a moiety that includes an unpaired electron. In a strong magnetic field, this unpaired electron will give rise to a distinct spectrum that can be used to monitor the position and motion of the EPR probe.

In this way, the spectra, the mobility of the probe, and, by inference, the behavior of myosin can be compared under different conditions. For example, the spectrum seen when myosin and actin are bound is much different from that seen when they are separate. When myosin binds to actin, the probe is more mobile, implying that the nucleotide site between the protein’s two “heads”, where ATP can be snared, is opening. “This is the first direct observation of that,” Pate says.

Homing in on detail with MD

As tantalizing as the experimental observations are, the inference of myosin’s position and activity lacks the dynamic detail that can be obtained through molecular dynamics (MD) simulations. With an allocation of 200,000 hours on NCSA’s Tungsten cluster, the team attempts to answer those questions.

“We get insight on how these molecules work in a dynamic sense,” Minehardt explains. “We are interested in observing biological systems at physiologically relevant conditions, and you can’t do that in the laboratory. These reactions take place too quickly.”

Using AMBER, a code developed by UCSF’s Peter Kollman and David A. Case of the Scripps Research Institute, the team runs MD simulations that mimic their laboratory experiments.
The spectra obtained from a nucleotide analog EPR probe. From the spectra, the researchers can infer how the probe is moving and therefore how myosin is behaving under different conditions. The decrease in splitting between the low-field peaks (A) and high-field dips (B) in the presence of actin implies increased probe mobility and an opening of the nucleotide pocket.

"This postulates where every atom is in the protein and the bound probe," Pate explains. By solving a very large system of differential equations, AMBER determines the evolution of the system, how the probe and myosin are interacting, and what the spectrum would look like for this postulated movement. The results of the computation are then compared to the experimental results, with the hope that the two match.

In an article published in Science in May 2003, the team detailed one such match. Using both EPR spectroscopy and MD simulation, they demonstrated that the nucleotide-binding site of kinesin undergoes a change from open to closed when the motor protein binds to its microtubule. They were able to show that there is a movement in a particular region of the nucleotide site, called switch 1; this movement primes the motor for the energy-rich nucleotide hydrolysis, which is essential for the production of force and motion.

"The computer is really, really, helping us here. Otherwise, we would have no real way to interpret this data," Pate says.

Improving drug development
By understanding at an atomic level how myosin and kinesin convert chemical energy into motion, the researchers are laying the groundwork for possible future therapies. For example, a drug that could disrupt kinesin's ability to aid cell division could attack cancer. Drugs targeting the smooth muscles in blood vessels could treat hypertension, while other pharmaceuticals could tackle heart disease.

Pate points out that there are also potential industrial uses for these "ultimate nano-motors."

"One of the things we would want to understand is, what is the minimum motor?" he says.

When the secrets of myosin and kinesin have been unraveled, the beat of your heart, the flexing of your muscles, the flow of your blood, and your overall health could be dramatically improved.

This research is supported by the National Institutes of Health.

Access Online: http://access.ncsa.uiuc.edu/CoverStories/myosin/
For further information: http://www.ucsf.edu/cooke/

Team members
Roger Cooke
Todd Minehardt
Ed Pate
Giving Sensors an Edge
by Trish Barker

Coupling tiny sensors with software-defined radio technology will provide scientists with the power and flexibility to gather more data.

NCSA is working to develop a single integrated circuit that combines data-gathering sensors with software-defined radio (SDR). This extensible sensor platform (ESP) would be capable of operating as a multi-band, multi-mode radio, enabling flexible, robust data transmission.

Because of this flexibility, sensors will no longer be constrained by distance or by cumbersome hardware, allowing researchers to gather extensive data they need.

This research is undertaken through the National Center for Advanced Secure Systems Research (NCASSR) and is supported by the Office of Naval Research (ONR).

Team Members

National Center for Supercomputing Applications
Alex Betts
Donna Cox
Matt Hall
Volodymyr Kindratenko
Meenal Pant
David Pointer
Von Welch
Paul Zawada

University of Illinois
Department of Electrical and Computer Engineering
Jennifer Bernhard
Thomas Huang
Ming Liu
Zhouhui Zhang

Enabling Discovery

1 Gathering Data with ESP

Data-gathering sensors are the foot soldiers of science; they are fundamental to innumerable inquiries across a vast range of disciplines. The extensible sensor platform being developed at NCSA could be released to monitor atmospheric conditions as they ride the wind, or attached to migrating birds in order to learn more about their environment, health, and habits. Or they could gather data for climate research or to monitor pollution.

The Impact of Sensors

Data drives discovery, and sensors are the critical tools that gather the data needed to further inquiry in innumerable scientific fields. Putting more and better sensors in the field will mean more and better data for scientists to interpret. The extensible sensor platform being developed at NCSA will enable researchers to gather data from far-flung locations with greater ease.
Transmitting Data with SDR

The ESP is given power and flexibility through the use of software-defined radio. Far from home base, each sensor could configure its SDR as a global positioning satellite receiver in order to pinpoint its location. Then the sensor could check on the available radio receivers in its vicinity, configuring its SDR to function as a cellular telephone, an amateur packet radio, or other available radio receivers/transmitters in order to send its data back to the lab.

More Data Yields Fresh Insights

The lab will receive and compile the data transmitted by the SDR. Using NCSA computational resources, scientists can then analyze, model, and visualize the data gathered by the ESP, generating new scientific insights and innovations.

The Flexibility of SDR

In standard radio devices, signal processing is performed by hardware. Existing technologies for voice, video, and data use different data types and processing techniques. And because these differences are locked in hardware, they are difficult to bridge. Software-defined radio transfers signal-processing functions from rigid hardware to flexible software. With SDR, a single device can allow many modes of communication and can switch between them just by adjusting its software.

The Security of Software

Another advantage of transferring signal processing to software is that it allows the incorporation of IT-based security concepts, such as public key infrastructure, into communications. SDR transmissions could be encrypted to prevent unauthorized people from accessing the data.
SGI Altix system added to NCSA resources

NCSA has purchased a Silicon Graphics shared-memory, symmetric multi-processor (SMP) computing system and storage technology. The SGI Altix system, called Cobalt, consists of 1,024 Intel Itanium 2 processors running Linux, three terabytes of globally accessible memory, and 370 terabytes of SGI InfiniteStorage that will serve as the center's shared file system, accessible by other high-performance computing resources at NCSA. With a peak performance of more than six teraflops, Cobalt will bring the total computing power at NCSA to over 35 teraflops and the disk storage to three-quarters of a petabyte.

Cobalt is scheduled to be fully installed by the end of this year. The target date for the system to be fully available to scientific users is March 1, 2005. NCSA will also make access to the SGI SMP system available through the TeraGrid cyberinfrastructure.

NCSA joins Globus Alliance

NCSA has joined the Globus Alliance, a consortium dedicated to the collaborative design, development, testing, and support of the open source Globus Toolkit. The Globus Toolkit is the de facto standard grid software. It provides key enabling software and services that let people share computing power, databases, and other tools securely online across corporate, institutional, and geographic boundaries without sacrificing local autonomy. It has been deployed broadly worldwide for both science and industry.

NCSA brings to the Globus Alliance both its experience in engaging with scientific communities—the ultimate beneficiaries of grid technology—and its expertise in grid deployment.

In addition to NCSA, other members of the Globus Alliance governing board are Argonne National Laboratory, the University of Southern California's Information Sciences Institute, the University of Chicago, the University of Edinburgh, and the Swedish Center for Parallel Computers.

Exploring black holes

The National Science Foundation has awarded $1.1 million to NCSA and other partners to produce a television program, a planetarium show, and educational materials exploring the scientific understanding of black holes.

The visualization team at NCSA, led by Donna Cox, will collaborate with Thomas Lucas Productions, Inc., the Denver Museum of Nature & Science, and astronomy experts from across the country to explore and visualize recent discoveries about black holes. The project was initiated with $250,000 in seed money from NASA's high-energy GLAST telescope project.

NCSA will produce high-resolution animated visualizations of cosmic phenomena, working with data generated by a computer simulation. Planetarium visitors will witness the birth of stars, the collision of giant galaxies, and a simulated flight through the event horizon of a supermassive black hole.

The large-dome digital planetarium show will debut at the Charles C. Gates Planetarium at DMNS; the planetarium's dome has 11 "footprints," each of which is 1,280 X 1,024 pixels. That adds up to more than 14.4 million pixels that will transport viewers into the heart of our galaxy.

"The television program and the planetarium show are a wonderful way to reach millions of people across the country and to introduce them to the grandeur and mystery of our universe," Cox says. "These projects are entertaining and exciting, but they're also a form of informal science education and a way to make state-of-the-art technology and cutting-edge science accessible to everyone."

The planetarium show is slated to debut early in 2006.
A national cybersecurity research center led by NCSA has received a second year of funding, more than $7 million, from the Office of Naval Research. The National Center for Advanced Secure Systems Research (NCASSR) was launched in 2003 to conduct research leading to the development of next-generation information security technologies.

NCASSR is led by NCSA, with partners at the University of Illinois at Urbana-Champaign, Battelle Pacific Northwest National Laboratory, InfoAssure Inc., the University of Tennessee, and the Naval Postgraduate School. More information is available at www.ncassr.org.

NCASSR research projects include:
- Cyberinformatics: Exploring methods for the discovery of suspicious patterns and events from a wide variety of sensors and data sources.
- Security tools: Developing software to provide security across many layers of a computer network—including validating hardware integrity, secure operating systems, secure group communication key management, and secure multicast and wireless networks.
- Sensors and software-defined radio (SDR): Combining sensors with software-defined radio and designing a low-cost SDR platform in which signal processing is handled by flexible software rather than rigid hardware.
- SCADA protocol authentication: Developing secure mechanisms for authenticating valid control signals and data acquisition and for blocking unauthorized intrusions to protect the Supervisory Control And Data Acquisition (SCADA) systems that are used in key energy infrastructures, such as electric power grids and water pipelines.

More than 50 girls in grades 6-8 attended the free Girls Engaged in Math and Science workshop June 29-30 at NCSA. The girls learned about programming and modelling tools, participated in hands-on activities, and used the Access Grid to talk to scientists at remote locations about their work. Read more about the GEMS project at http://gems.ncsa.uiuc.edu/.
Teachers complete REVITALISE training

The first cadre of teachers to participate in the REVITALISE program, which helps rural educators gain skills and experience in scientific visualization techniques that they can incorporate in their curricula, completed the two-year program at NCSA June 14-25. For more information on the NSF-funded program, go to http://www.ncsa.uiuc.edu/Projects/AllProjects/Projects59.html.

Students gain HPC skills

Four Washington D.C.-area high school students gained high-performance computing skills during a summer program sponsored by NCSA and the Joint Educational Facilities (JEF). With guidance from instructors and mentors from Bethune-Cookman College, Bowie State University, JEF, NCSA, Florida International University, and the University of Illinois at Urbana-Champaign, the students pursued individual research projects in cluster security, video editing, animation, and networking.

The students had access to a 26-processor high-performance computing cluster at TRECC (the Technology Research, Education and Commercialization Center in DuPage County, IL). TRECC is a program of the University of Illinois that is administered by NCSA and is funded by the Office of Naval Research.

VMI 2.0 released

In May, NCSA released version 2.0 of its Virtual Machine Interface (VMI) software. VMI is a middleware communication layer that addresses the issues of availability, usability, and management in the context of large-scale SANs interconnected over wide-area computational grids. With VMI, users are able to run applications on distributed clusters that use different types of interconnects to communicate among processors. The release is available for download from the VMI project website, http://vmi.ncsa.uiuc.edu/. Access to documentation, a bug reporting tool, and support mail lists are also available at the site.
Following are the five scientists whose research teams used the most hours on NCSA high-performance computing systems between August 1, 2003 and July 31, 2004. Included are the names of their research projects and the names of the NCSA systems used.

1. Robert L. Sugar, University of California, Santa Barbara, the MIMD Lattice Calculation (MILC) Collaboration, 1,449,049 hours (Platinum IA32 Cluster, Mercury IA64 Cluster, Tungsten Xeon Cluster and SGI Origin2000)

2. Gregory A. Voth, University of Utah, Fundamental Dynamical Processes in Condensed Matter, 511,881 hours (IBM P690, Platinum IA32 Cluster, Mercury IA64 Cluster)

3. Renyue Cen, Princeton University, Simulating the Clearing of the Cosmological Fog and the Evolution of the Intergalactic Medium, 432,470 hours (IBM P690, Platinum IA32 Cluster, Tungsten Xeon Cluster)


‘MILCing’ NCSA’s Computing Environment

The MIMD Lattice Calculation (MILC) collaboration is a research team with members at seven institutions: the American Physical Society, the University of Arizona, the University of California, Santa Barbara, Indiana University, University of the Pacific, University of Utah, and Washington University. The group studies quantum chromodynamics, or QCD. QCD describes the strong interactions that bind protons and neutrons together to form the nuclei of atoms. The elementary entitites of QCD are quarks and gluons. Quarks are the primary constituents of protons and neutrons and of a host of short-lived particles produced in high energy accelerator experiments and in cosmic rays. Gluons—the strongest superglue imaginable—bind quarks together to form these particles.

Specific areas of inquiry for the MILC collaboration have included studies of the mass spectrum of these strongly interacting particles, the weak decays of these particles, and the behavior of strongly interacting matter at high temperatures. Since quarks and gluons are among the most basic known particles in the universe and the building blocks for most of the matter we observe, understanding their behavior is critical to answering some of the key questions about the fundamental laws of physics and the origins of the universe.

QCD simulations use a four-dimensional grid, called a lattice, to represent time and the three dimensions of space—variables that are continuous in the physical world. By working on ever finer grids, the researchers can perform more detailed simulations and obtain a more realistic representation of the interactions among strongly interacting particles. However, working on fine grids requires massive amounts of computer time.

The researchers are constantly on the lookout for computers powerful enough to handle their calculations. Fortunately, the MILC code is quite portable, so the group is able to work on a variety of platforms.
The recently completed IMAX short film "Hubble: Galaxies Across Space and Time," directed by Frank Summers of the Space Telescope Science Institute in Baltimore, has been booked to play in more than 30 cities in the United States and Canada over the next few years. This scientifically accurate visualization explores a Hubble dataset of some 30,000 galaxies, showing both their 3D distribution in space and their evolving characteristics over time.

Summers built a visualization display wall—a four-by-four array of flat panel monitors driven by a 16-node Linux cluster—to pursue high-resolution astronomy applications with Hubble. The custom setup was crucial to creating the film on a minimal budget, and relied on software and techniques pioneered by NCSA. NCSA’s Pixel Blaster software, written by Paul Rajlich, allowed the production team to pre-visualize the entire movie digitally at 90 percent of IMAX resolution before the expensive step of film recording.

"It proved to be perfect for IMAX in that it provided both the processing power to compute the film, as well as a 21 million-pixel display to view the high-resolution frames. The combination of a 625 million-pixel Hubble image, display wall technology, and cool scientific visualization has brought the universe directly to the public on an IMAX scale," says Summers.