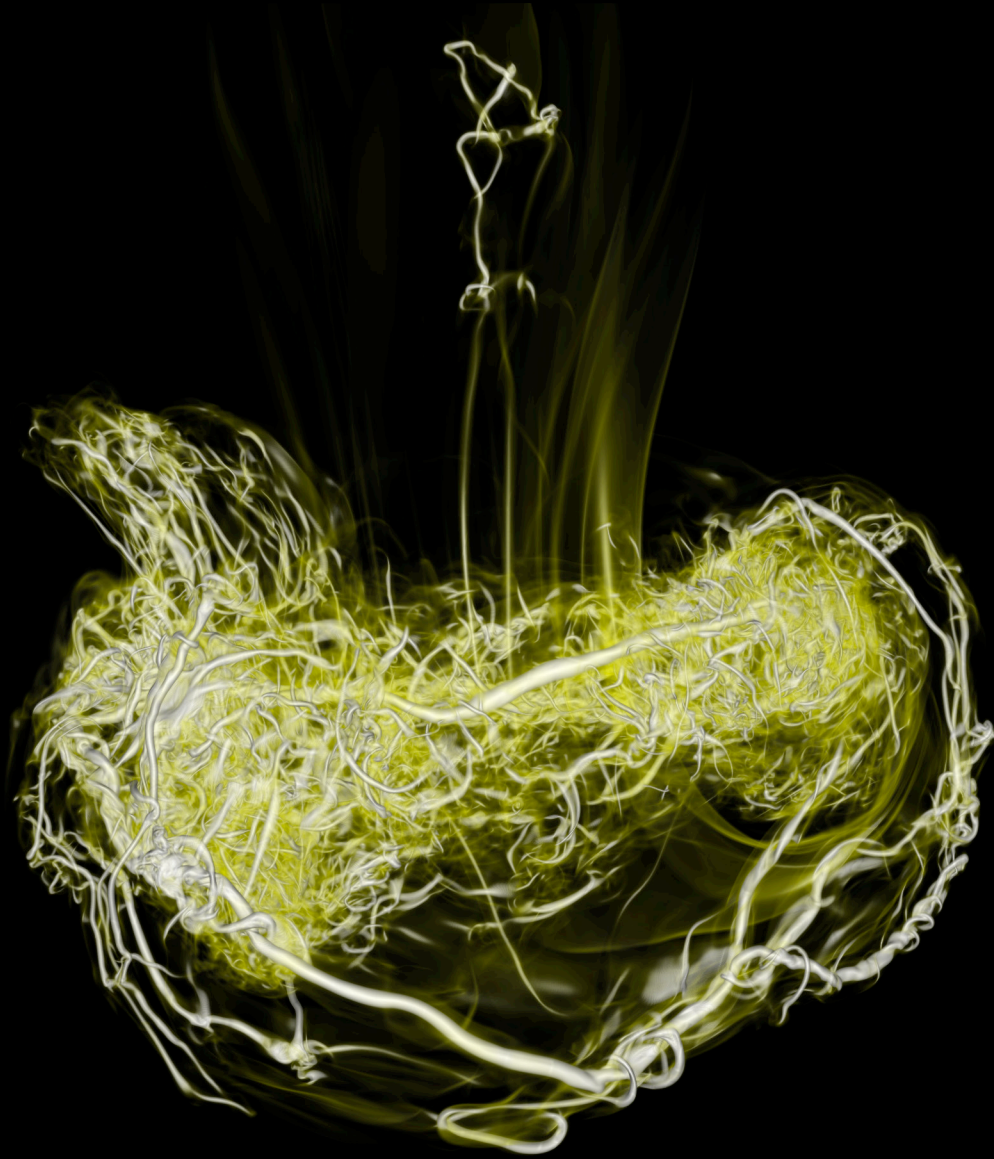


access

NCSA | University of Illinois | Vol 25 No 3 | Fall 2012



SPECIAL AUGMENTED REALITY ISSUE



access

Vol 25 No 3 Fall 2012

Managing Editor

Barbara Jewett

barbaraj@ncsa.illinois.edu

Art Director

Paula Popowski

popowski@ncsa.illinois.edu

Copy Editors

Trish Barker

Nicole Schiffer

Access Editorial Board

Randy Butler

Donna Cox

Thom Dunning

Bill Kramer

John Melchi

Danny Powell

John Towns

Bob Wilhelmson



On the Web

News from NCSA is always available on the Web. Sign up for our RSS feed or email service for the latest developments from NCSA and its partners.

www.ncsa.illinois.edu

National Center for Supercomputing Applications University of Illinois at Urbana-Champaign

1205 West Clark Street

Urbana, IL 61801

217-244-0072

Who we are

The University of Illinois at Urbana-Champaign's National Center for Supercomputing Applications (NCSA), one of the five original centers in the National Science Foundation's Supercomputer Centers Program, opened its doors in January 1986. Over the years NCSA has contributed significantly to the birth and growth of the worldwide cyberinfrastructure for science and engineering, operating some of the world's most powerful supercomputers and developing the software infrastructure needed to efficiently use them.

That tradition continues as the center, Illinois, Cray, and their partners in the Great Lakes Consortium for Petascale Computation develop what is expected to be the first computer dedicated to open scientific research capable of sustaining more than one petaflop, or one quadrillion calculations per second. Called Blue Waters, the system is dedicated to massive simulations and data analysis projects that will improve our society, health, environment, and economic competitiveness. NCSA and the consortium also work with research communities to create the new software technologies, scientific applications, and educational programs needed to take full advantage of this new system.

NCSA also leads the Extreme Science and Engineering Discovery Environment (XSEDE), the most advanced, powerful, and robust collection of integrated advanced digital resources and services in the world. It is a single virtual system that scientists can use to interactively share computing resources, data, and expertise. Scientists and engineers around the world use these resources and services—supercomputers, collections of data, and new tools, for example—to make discoveries that benefit society. The five-year, \$121 million project is supported by the National Science Foundation. XSEDE replaces and

expands on the NSF TeraGrid project. More than 10,000 scientists used the TeraGrid to complete thousands of research projects, at no cost to the scientists.

The center also leaves its mark through the development of networking, visualization, storage, data management, data mining, and collaboration software. 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. Through its Private Sector Program, top researchers explore the newest hardware and software, virtual prototyping, visualization, networking, and data mining to help U.S. industries maintain a competitive edge in the global economy.

Support for NCSA is provided by the National Science Foundation, the state of Illinois, industrial partners, and other federal agencies. For more information, see www.ncsa.illinois.edu.

Cover

Data from simulations conducted on the Blue Waters Early Science System by the Petascale Computing Resource Allocations (PRAC) team led by Stan Woosley of the University of California Observatories. Team members include Chris Malone of the University of California, Santa Cruz, and Andy Nonaka of the Lawrence Berkeley National Laboratory.

The Woosley team is exploring Type Ia supernovae, more specifically the plot of kinetic energy in a supernova. They are focusing on the "bubble" environment inside a star. The environment in which the bubble resides is turbulent, there are a lot of swirls and eddies moving the fluid around when the bubble is ignited. One of the things the team hopes to learn through scientific computing on Blue Waters is if this background turbulent field has any effect on the evolution of the bubble.

The image shown on the front cover, created by Blue Waters visualization experts Rob Sisneros and Dave Semeraro, focuses on the highest values of the magnitude of vorticity to observe the formation of "vortex tubes." This is roughly 0.37 seconds into the evolution of the hot ignition point as it buoyantly rises towards the stellar surface. In this time, the bubble has risen 250 km and has expanded from its initial size of about 2 km to over 120 km across.

THE AR IS

access contents

02 An Expert Opinion

Top problems with the TOP500
William Kramer



04 Q & A

Changing the world
Peter Ungaro

06 T. rex on Your Desk!

You are probably familiar with virtual reality, interactive apps, and handheld computers, but have you heard of their progeny, augmented reality?
by Nicole Schiffer



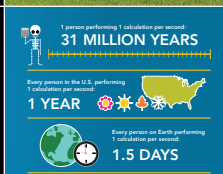
10 Harmonic Convergence

Many institutes dedicated to applying advanced computing to a range of fields and issues share a single building on the University of Illinois campus, and the close quarters aid their collaboration and cross-pollination.
by Trish Barker



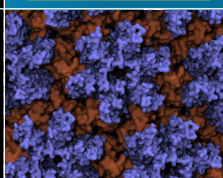
12 Infographic

What is a Petaflop?



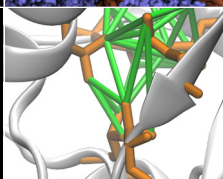
16 Making History

Early results indicate Blue Waters will live up to expectations.
by Barbara Jewett



18 Doing with Less

Researchers used NCSA's Forge to improve computational protein structure modeling.
by Barbara Jewett



21 News & Notes

26 Parting Shot

NCSA's tradition of leading through innovation continues as we showcase augmented reality in this issue of Access. Through a collaboration with daqri, nearly every item listed on this Table of Contents has at least one augmented reality component to demonstrate the uses of 4D. With 4D, you experience three-dimensional content, brought into your world.

To provide a truly seamless experience, we consciously chose not to place a symbol indicating exactly where to point your camera for augmented content. We believe these seamless experiences are the future of publishing, and as leaders in innovation wanted your first seamless AR experience to be with NCSA. We will tell you, however, that you will find these AR experiences on pages 3, 4, 6, 8-9, 10-11, 12, 15, 18, and 25 as well as the front cover. As you move your device's camera over these pages in this issue, the app will recognize markers and trigger the corresponding experience.

daqri's mission is to empower the world to visually connect places and objects with the right information at the right time in a way that is both valuable and beautiful. NCSA is pleased to partner with daqri in this Access experiment. To find out more about daqri and the world of 4D, visit www.daqri.com.

daqri 4D applications are unified by a lightweight, dynamic interface that allows for 360° spatial and kinesthetic interactions.

1 Visit the app store for your Apple or Android device and download the free daqri app. Shortcuts to the app may be found at:
http://go.illinois.edu/daqri_app_apple
http://go.illinois.edu/daqri_app_android

2 Start the application.

3 Point the device's camera toward the page the demonstration is on. You may need to move your camera slightly to more tightly focus on the image if the AR experience does not begin.

4 Experience science in a whole new way! For example, by lifting, lowering, and tilting your device you can change how a molecule looks.

SUE!



TOP PROBLEMS WITH T

THE TOP500 LIST WAS INTRODUCED OVER 20 YEARS AGO to assess supercomputer performance at a time when a major technical challenge was solving dense linear algebra problems on high-end computers. Now, most consider the list a marketing tool or an easy, if simplistic, way to impress politicians, funding agencies, and the naïve rather than having much technical value.

The TOP500 list is based on the floating point computational performance assessed by a single benchmark, Linpack, which performs a dense matrix calculation. Unfortunately, the TOP500 does not provide comprehensive insight for the achievable sustained performance of real applications on any system—most compute-intensive applications today stress many features of the computing system and data-intensive applications require an entirely different set of metrics. Yet many sites feel compelled to submit results to the TOP500 list “because everyone has to.”

It is time to ask, “Does everyone have to?” and more specifically, “Why does the HPC community let itself be led by a misleading metric?” Some computer centers, driven by the relentless pressure for a high list ranking, skew the configurations of the computers they acquire to maximize Linpack, to the detriment of the real work to be performed on the system.

The TOP500 list and its associated Linpack benchmark have multiple, serious problems. This column is too short to deal with all of them in detail. But a few of the issues and possible solutions are briefly listed below.

The TOP500 list disenfranchises many important application areas.

All science disciplines use multiple methods to pursue science goals. Linpack only deals with dense linear systems and gives no insight into how well a system works for most of the algorithmic methods (and hence applications) in use today. The lack of relevance to many current methods will get worse as we move from petascale to exascale computers since the limiting factor in performance in these systems will be bandwidth of memory and interconnects.

Possible improvement: Create a new, meaningful suite of benchmarks that are more capable of representing achievable application performance. Several Sample Estimation of Relative Performance of Programs (SERPOP) metrics are in use today, such as the NERSC SSP test series, the DOD Technology Insertion benchmark series, and the NSF/Blue Waters Sustained Petascale Performance (SPP) test that use

a composite of a diverse set of full applications to provide a more accurate estimate of sustained performance for general workloads. These composite measures indicate critical architectural balances a system must have.

There is no relationship between the TOP500 ranking and system usability.

In a number of cases, systems have been listed while being assembled at factories or long before they are ready for full service, leaving a gap of months between when a system is listed and when it is actually usable by scientists and engineers. This perturbs the list’s claim of historical value and gives misleading reports.

Possible improvement: List only systems that are fully accepted and fully performing their mission in their final configurations.

The TOP500 encourages organizations to make poor choices.

There have been notable examples of systems being poorly configured in order to increase list ranking, leaving organizations with systems that are imbalanced and less efficient. Storage capacity, bandwidth, and memory capacity were sacrificed in order to increase the number of peak (and therefore Linpack) flops in a system, often limiting the types of applications that it can run and making systems harder for science teams to use. For example, for the same cost, Blue Waters could have been configured to have 3 to 4 times the peak petaflops by using all GPU nodes and having very little memory and extremely small storage. This would have made Blue Waters very hard to program for many science teams and severely limited what applications could use Blue Waters, but almost certainly would have guaranteed being at the top of the TOP500 list for quite a while.

Possible improvement: Require sites to fully specify their system capacities and feeds. For example, the amount and speed of memory and the amount and speeds of the I/O subsystems should be reported. This would allow observers to assess how well a system is balanced and would also document how different types of components influence the performance results.

The TOP500 measures the amount of funding for a system—it gives no indication of system value.

The dominant factor for list performance is how much funding a site received for a computer. Who spends the most on a system influences

William Kramer
Deputy Project Director Blue Waters
NCSA

THE TOP500

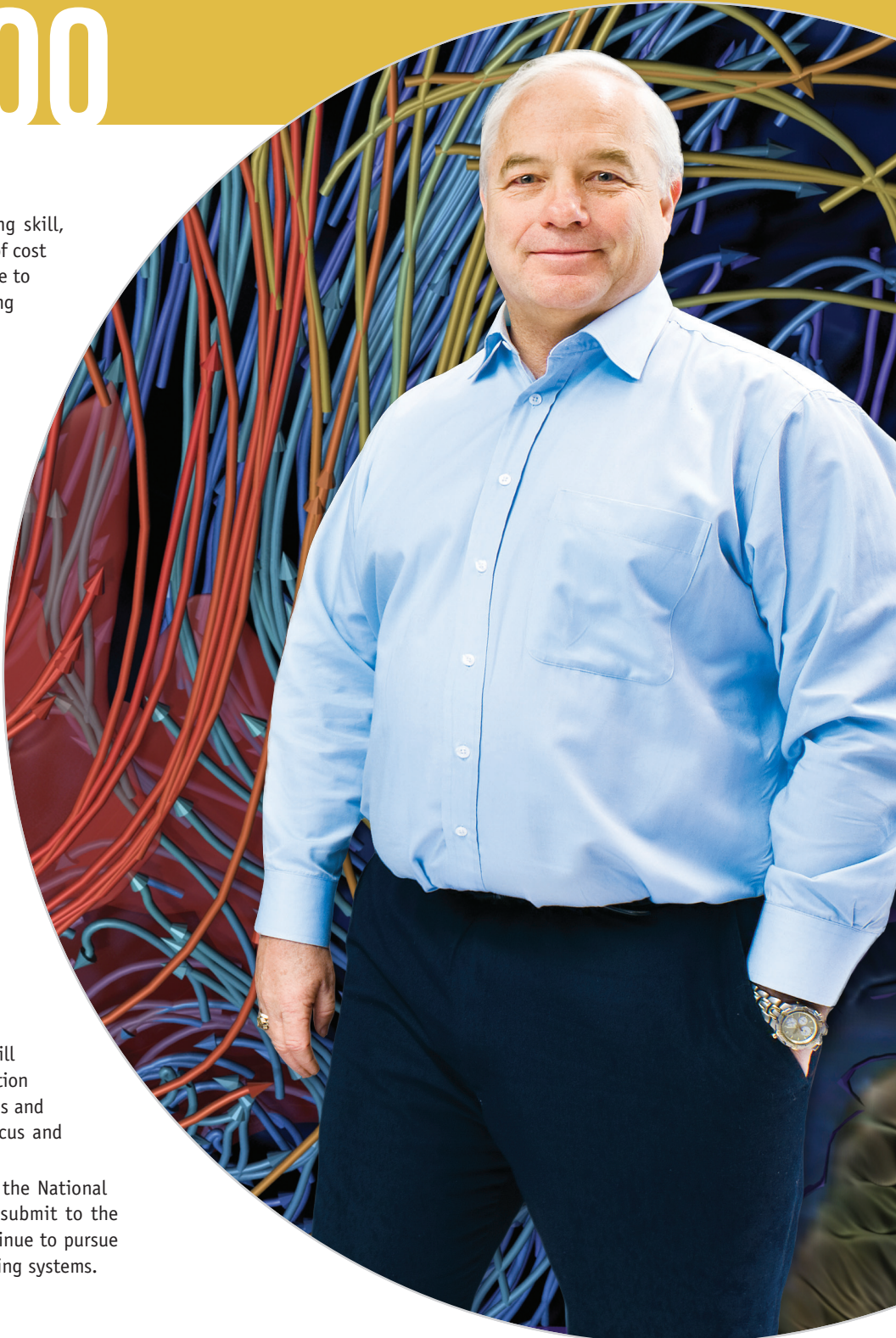
list position as much as (or more than) programming skill, system design, or Moore's Law. Without an expression of cost listed alongside the performance metric it is impossible to understand the relative value of the system, inhibiting meaningful comparisons of systems.

Possible improvement: Require all list submissions to provide a system cost. The cost estimate could be the actual cost paid, a cost estimated from pricing tables, or, at the worst, a component-wise estimate. The cost of a system contract is often publically announced by sites, or IDC (or others) can help calculate a typical "street" selling price for most systems. A cost estimate along with a ranking would provide much more insight and value. Remember, in the past every system listing in the NAS Parallel Benchmark reports was required to have a cost estimate.

Adopting these and other improvements would be steps in the right direction if the list continues. However, it is time the community comes to agreement to entirely replace the TOP500 with new metrics, or multiple lists, that are much more realistically aligned with real application performance. Or the HPC community could just say "No more" and not participate in the list. Many government and industry sites already do this, we just never hear about them (which further limits the use of the list for historical information).

As our HPC community strives for more and more powerful systems, and as we cope with having to implement more exotic architectural features that will make future systems harder to use for sustained application performance, it is critical we have measures to guide us and inform our decision making rather than divert our focus and adversely influence our decisions.

Because of the issues discussed here, and with the National Science Foundation's blessing, Blue Waters will not submit to the TOP500 list this fall or any other time. NCSA will continue to pursue new ways to assess sustained performance for computing systems.





CHANGING THE WORLD

Cray President and CEO Peter J. Ungaro recently spoke with Access' Barbara Jewett about why the company Seymour Cray founded continues to be inimitable.

Q. Let's start by talking about Blue Waters. Did you ever doubt you were doing the right thing in taking on the project?

A. Not for a single second. We're a company that is focused on taking supercomputing technology and applying it at huge scale, so I can't think of a better project that our company is built around than taking on something like Blue Waters and building one of the fastest and most capable systems in the world. When we announced it, a lot of people came up to me and said "You're crazy. This is a crazy project. You are going to do this in such a fast timeline." But this is what Cray is all about. This kind of challenge is what energizes us and why I think we exist in the market today. The other cool thing was, as we started working with the team from NCSA, we quickly discovered we had a great partner. The NCSA thinks very similarly to how Cray thinks as a company. They are focused on their users, and the way the organization views real performance and real usage of the machine is exactly the same way we see things. That made our partnership even stronger and will be a key factor in our combined success.

Q. What do you think about a replacement for the Linpack benchmark as the determiner of placement on the TOP500 list?

A. I'm a huge fan of the TOP500. The list is a great way, maybe even the best way, to keep track of what's going on with the supercomputer market, what's going on at various centers around the world, and how the market and landscape is changing and evolving over time. We owe a lot to the four guys who build the list twice a year. That said, I don't think that using the Linpack benchmark is a good way to tell who has the best or fastest supercomputer because it only measures one statistic. Linpack isn't bad but it is only one metric. It's like trying to determine who's the best pitcher in baseball by seeing who can throw the ball the fastest. But are they accurate? Do they have other pitches they can throw? Do they have mental toughness? No intelligent baseball team would draft a pitcher just because he threw the ball fast—just as I don't think a supercomputer center buys a supercomputer and only looks at one metric. Would Cray disagree with changing the TOP500 list and looking at a number of different metrics? No, we would welcome that. But I know there's a lot of debate, because running the one Linpack benchmark takes a lot of time. These systems are very expensive, and if something takes a lot of time it becomes difficult for people to invest that amount of time. So that's the other side of it. The more metrics you have, the harder it is from a time investment standpoint. You have to take machines out of production just to see how they rank on the TOP500 list.

Q. When you spoke at the NCSA earlier this year, you mentioned that the business side of the technology now has to be considered. What are some of these business aspects?

A. Energy costs are going to be a huge driver. The size of the market that wants to buy these kinds of machines is also going to be a significant driver in the research and development (R&D)

Questions & Answers

investment. The other big thing is what I would call productivity costs, which is driving our R&D agenda today. How do we let people take advantage and make good use of these systems? How do you take a system like Blue Waters, which has both CPUs and GPUs in it, and let people take advantage of both of those technologies, let people leverage them and use them, without becoming experts in each of those areas, without having to do things from scratch and rewrite complete applications? We want scientists to be scientists and use supercomputers to make huge breakthroughs, and not have to become expert parallel computer programmers and spend all their time trying to learn how to use each generation of these machines. This is having a huge impact on the R&D agendas of companies like Cray.

Q. Is Cray going to be more involved with software going forward?

A. Cray is a systems company. We've always been a systems company. And a big part of being a systems company is doing some software, doing some hardware, and then putting all that together. But where we focus, and what I think makes us really unique in the market, is that we spend most of our time in R&D looking at how hardware and software technologies come together, and how we can integrate them in a very tight way so we can give our customers a real supercomputer, not just a bunch of building blocks for them to assemble and figure out for themselves. That tight integration between hardware and software is a big part of what gives us our scalability and performance advantages. As technology changes, we figure out what's the right balance of hardware and software so we can continue to do this very tight integration. Over the past few years, we've put more energy and incremental investments into our software R&D technologies than our hardware, but what we're focused on is how all that integrates together because we are a systems company.

Q. Big data is a hot topic today, whether it's gathered by mega retailers, or it's the data sets, visualizations, and simulations created by researchers. What's Cray's take on big data?

A. In many ways high-performance computing (HPC) and big data are actually all one thing, although there are different components to it: the simulation of the data, the storage and moving around of that data and the management of it, and then the analytics we do on top of that data. As a systems company, our three business units address these components. One is our HPC systems unit, which is where all of our supercomputers are developed. Blue Waters was done in that group. Our second group is the storage and data management unit. The Sonexion™ storage system going into Blue Waters was developed by that team. And we have a third group, YarcData, which is focused on the data analytics side of things, particularly as it relates to graph analytics. As we look out to the future, with simulations being run on systems like Blue Waters, we have to be able to run those simulations, which generate huge data sets that need to be stored and moved around. We then have to be able to do analytics on that data and make sure we are really pulling the knowledge out of that data. It is not just

finding a needle in a haystack but understanding how all this data works together and finding unique and hidden information in the relationships between the data that really gives us that moment where we pull that all together and make a huge step forward.

Q. When you gaze into your HPC crystal ball, what do you see for the future of Cray and the future of supercomputing? And do you think we'll get to exascale any time soon?

A. The biggest thing I see coming is that today we like to focus on what I call simulation engines, which are our traditional supercomputers like Blue Waters. But we also have storage, like what we're doing with our Sonexion™ product, and then we have the data analytics side that we're doing with a product called uRiKA, and I see these three environments coming together on a single platform. A long time ago, Cray announced a technical vision that we call Adaptive Supercomputing, where we build one very scalable infrastructure, then put in different processing technologies and build an adaptive software environment to allow the user to take advantage of how these systems are going to evolve. But how I see this vision evolving is that these three distinct, different areas will all come together under a single infrastructure, with one integrated system. Some people like to call that a data-intensive architecture, which is kind of an integration of HPC and big data. But whatever you want to call it, I see that point out in the not-so-distant future.

Will we have exascale by 2018 or 2019 or 2020? It's really a matter of whether somebody is going to focus their funding on building one of those machines because they believe that doing so will have a profound impact on their mission. Do we have the capability of building one of those machines? I absolutely believe we do. What will its performance, power usage, and capabilities be? That one is still a research problem that we're working on. I am concerned, though, that we're losing time because a lot of the debates that are happening right now are about the importance of exascale and not how to do it, how to structure it, and such. I don't know if exascale is going to be done first in the U.S. It may be done first in another country or two before it gets to the U.S., but I do think that by the end of the decade exascale machines will be in use somewhere ... and hopefully it will be a Cray supercomputer!

Q. Anything else you'd like our readers to know?

A. I want to tell everybody that we are very excited about being a part of what we're doing with the NCSA and Blue Waters, and what we believe that system will mean for the National Science Foundation's scientific community. It is going to be such a huge step forward for the community overall. I can't wait to get the machine completed, into production, and into the hands of users because I can't wait to see what is going to come out of it. Blue Waters has a chance to change the world, literally, in many different scientific areas. We get so excited at Cray just thinking about seeing our supercomputers being used as tools for scientists to change the world. That's a big deal for us and is a big part of why we exist as a company, all the way back from the early days of Seymour.



by Nicole Schiffer



T. REX ON YOUR DESK!



UGMENTED REALITY, OR AR, DIGITALLY superimposes sights and sounds on the images captured by the camera on your smartphone, iPad, or other electronic device, giving the sensation they are really there. Projectors can bring AR to a larger audience, and other devices

can appeal to your senses of smell, touch, and taste in addition to sight and hearing. Possibilities are myriad.

Reality of AR

NCSA's Alan Craig works at the forefront of this new technology. When he asked me if I had seen AR, I figured I had seen something similar enough that the novelty would quickly pass—a common reaction, says Craig.

He laid a picture of a T. rex on regular printer paper on the desk in front of me, handed me an iPad, and told me to point the camera at the picture. In less than a second, a little three-dimensional T. rex popped up. "Tilt the iPad up," he instructed.

The dinosaur stood on the desk like the little Hungarian Horntail dragon Harry Potter drew from a bag!

I reached behind the iPad and poked at the miniature. It didn't react and my finger disappeared behind the 3D image, but this is only a prototype.

"Things seem fake until we interact with them the way we do in the real world," says Craig.

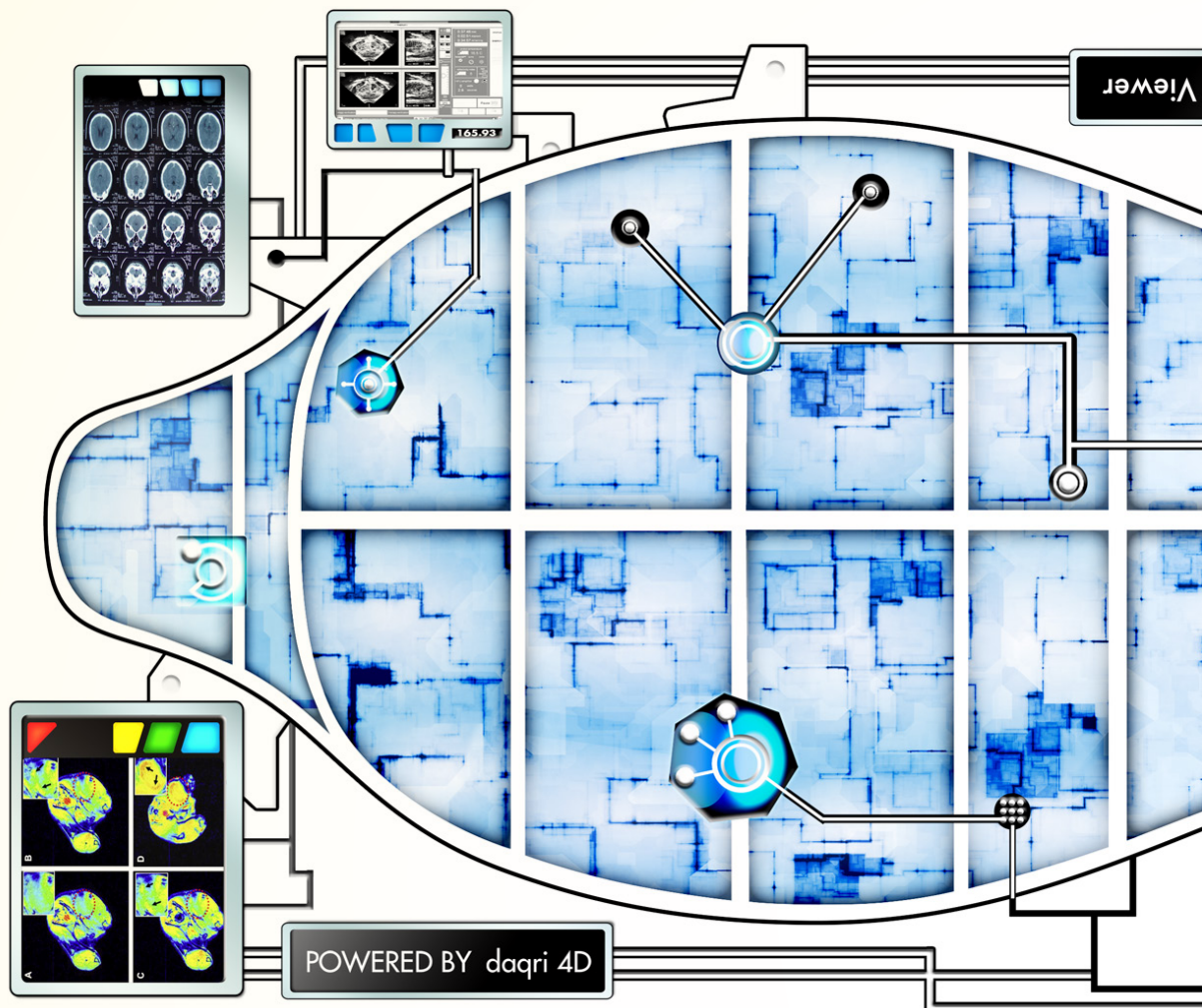
How does AR work?

It starts with an idea, something to communicate, says Craig. He and his team plan how to display the content and then find an existing application to implement their plan or write their own app.

Craig started experimenting several years ago with BlackMagic, an application based on the freely available, open-source ARToolKit. BlackMagic creates what is called a fiducial marker, which is some symbol that a device's camera will recognize, and associates it with a 3D image.

One drawback of ARToolKit and ARToolworks is that they require a customized version for each operating system. Craig said that three different Android phones would require three different implementations because of small variations in the open-source Android operating system.

"This is not a shortcoming of that specific product," says Bob McGrath, a retired NCSA software developer who is working with Craig on the project. "AR is extremely closely tied to the video input and output of the device, and these are quite different on each platform, at least up until now."



AR in use

Where might AR be used? Games, advertising, education, news—the list goes on. Even in its infancy, AR is making inroads in many areas.

For instance, Starbucks has run AR holiday ad campaigns using their cups as the trigger image for the AR app. Marvel used AR to activate videos with sound over the comic book pages of Avengers vs. X-Men, including animating the evolution of some pages from pencil to full-color panels. Even more staid publications like newspapers are getting into the AR game. The Arkansas Democrat-Gazette printed its first enhanced edition on July 1, peppering it with icons that indicate interactive material. Several game systems have taken a shot at incorporating AR. Nintendo 3DS, for example, comes with the AR Games suite.

AR in Education

Craig's main interest is in education. He uses AR to make molecules pop out of a chemistry textbook, simulate an archeological dig for students in a classroom, and create an interactive 3D display of anatomical systems in the human body.

An early project for Craig's team was working with Maurice Godfrey at the University of Nebraska Medical Center to create DNA models for "DNA Days" at Omaha's Henry Doorly Zoo and Aquarium. More than

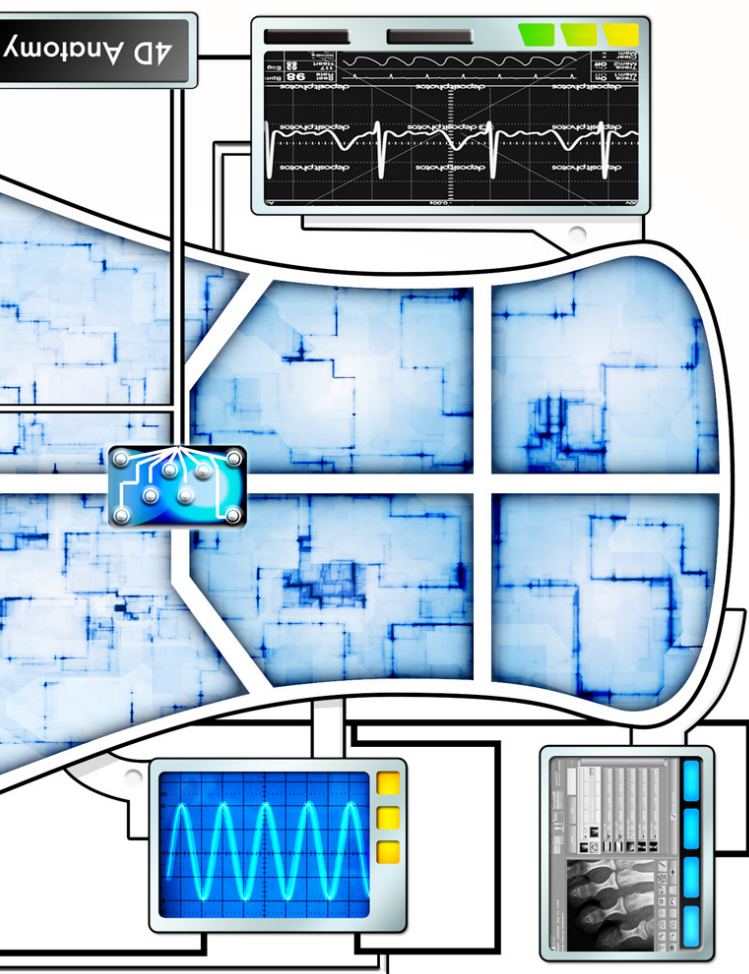
15,000 people used markers in their programs to see 3D images of DNA at a station set up at the zoo. Godfrey says "there was a lot of wow factor there."

After their first booming success, the collaboration continued for a new project, adding color 3D depictions of wildflowers to an ethnobotany workbook. The workbook is designed to educate Native American high school students in Nebraska and South Dakota about native flora and originally included text and black-and-white sketches of the plants.

The first step for the NCSA portion of the team was to talk to botanists about how to accurately portray the plants. Student employee Ryan Rocha created the images using a 3D image and animation software package called Maya, which is also used in scientific visualization at NCSA and in some Hollywood movies.

Close collaboration between the developers and artists was essential throughout the creation of graphics, says McGrath. Better graphics take more memory, sometimes more than smaller devices have. The team aimed for images that were as good as they could be without sacrificing usability.

"Of course, 'good enough' depends on the goals of the user...We will have to adjust graphics to satisfy user communities," McGrath says.



Then they imported the models into the application and used the application to create fiducial markers for each plant. Students used iPads with the custom app installed to access the AR features.

Preliminary feedback is overwhelmingly positive. Classrooms on several reservations, like the Yankton and Rosebud reservations in South Dakota and Winnebago and Omaha reservations in Nebraska, began using the AR version of the workbook this fall, says Godfrey. And he says teachers in Arizona have expressed interest in adapting the workbook to their region. In addition, he exhibited the AR workbook at the National Indian Education Association's annual convention in October.

Still developing

Both Craig and McGrath say one of their primary concerns in AR development is latency—how long does it take for the digital augmentation to appear once the camera sees the marker and how well does the digital image keep up when the user moves the viewing device?

"This is a significant issue for AR because you are trying to do visual tracking and 3D graphics in 'real time,' updating at least 25 frames per second," says McGrath.

Even the amount of memory makes a huge difference. McGrath and Craig found that the application worked much better on McGrath's

MacBook than Craig's, which is identical except for the amount of memory. Some mobile devices also have trouble running the AR app, one more technical detail to work around.

For his current projects, Craig is collaborating with a new company called daqri. daqri provides a platform that is compatible with multiple devices like smartphones, iPads, and iPhones, which will allow Craig to reach a wider audience more easily. McGrath says that the development of graphics is the bottleneck, and daqri aims to make that easier too. daqri promises the potential to create and distribute AR applications very quickly across a wide variety of platforms. This issue of Access was a collaboration with daqri and shows many different (though not all) of the capabilities of augmented reality.

Where is AR going?

AR is growing. Craig expects that people will first encounter AR through education or gaming. He says he thinks it will spread quickly through mobile devices and eventually become embedded in glasses and even contact lenses.

"Getting power to contacts is the hard part," says Craig, "but they have working one-pixel displays now. I think AR will become just a part of our world and sort of expected that it's there. AR is a new medium, and as such, it remains to be seen all the ways it will be adopted and realized."

Craig is in the final stages of authoring a book titled "Understanding Augmented Reality" that will be published by Morgan Kaufmann Publishers in early 2013. □

PROJECT AT A GLANCE

TEAM MEMBERS

Alan B. Craig
Robert E. McGrath
Ryan Rocha
Michael Vila
Gaia Dempsey (Daqri)
Brian Mullins (Daqri)
Jonathan Ross (Daqri)

FUNDING

National Science Foundation
daqri
I-CHASS

FOR MORE INFORMATION

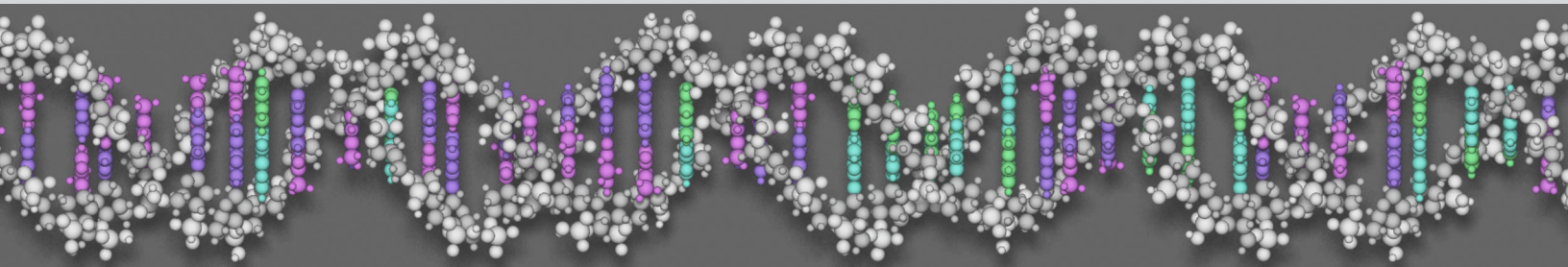
<http://site.daqri.com/>

ACCESS ONLINE

www.ncsa.illinois.edu/News/Stories/AR

HARM CONV

With expert advice from Joanna Shisler, a virologist at the University of Illinois, and with help from Raji Heyrovska from the Academy of Sciences of the Czech Republic, NCSA's Advanced Visualization Laboratory developed Genie, a simple interface to create atomically accurate DNA models based on real gene sequence data. Users can update the rendered appearance of the DNA through variables like bond lengths, atomic elements, and molecular identities. It has been used in an ongoing research project of human papillomavirus (the most common sexually transmitted virus in the United States), and easily handles the virus's nearly 8000 base pair genome. The DNA model shown here is based on some of the base pairs of the HPV-16 strain.



THE NCSA BUILDING IS A HOTBED FOR INNOVATION IN the application of advanced computing technologies. In addition to being the home of NCSA, it's also the headquarters for the Institute for Advanced Computing Applications and Technologies (IACAT), the Illinois Informatics Institute (I3, or I Cubed), eDream (the Emerging Digital Research and Education in Arts Media Institute), and the Institute for Computing in the Humanities, Arts, and Social Science (ICHASS).

These organizations can be thought of as members of an extended family. Having passed its 25th birthday, NCSA is something of a patriarch. ICHASS and eDream both spun off from and have been incubated by the center; IACAT was formed to foster collaborative projects that bring together the center's technology experts and discipline researchers across campus. I3, the only degree-granting family member, married into the clan; they saw what was gathering here and wanted to be a part of it," says NCSA Director Thom Dunning.

While each organization has a specific mission, they share a common interest in the transformative potential of advanced technology.

"They're all focused on one thing, and that's how do we creatively use advanced computing technologies to address a number of different issues. Whether it's issues in the humanities and social sciences, whether it's in media—we're all involved in basically the same type of endeavor," Dunning says.

The organizations connect to and collaborate with one another—and because they share a facility, a project can start with something as simple as bumping into a colleague at the communal coffee pot.

"Just having people together in the same location where they can talk on a regular basis, where they run into one another on a regular basis, that is a very positive aspect of the colocation of NCSA and the other institutes," Dunning says.



KEVIN FRANKLIN

Executive director, Institute for Computing in the Humanities, Arts, and Social Science

"We each are exploring a different flavor of advanced computing. Colocation supports additional creativity and innovation in the space. It's a beautiful environment to be in,

between the technologists that are in the building and the kind of creative talent that's represented in I3 and eDream, I can't think of a better place to be working and for ICHASS to be located."

"ICHASS is the convergence of humanities and social science with advanced and high-performance computing. Our niche is very unique. There aren't any other digital humanities centers that are embedded in a supercomputing center. The synergies with NCSA are extremely important. We can establish ad hoc groups that can be pulled into our interdisciplinary teams. That really enhances the work that we're doing."



BILL GROPP

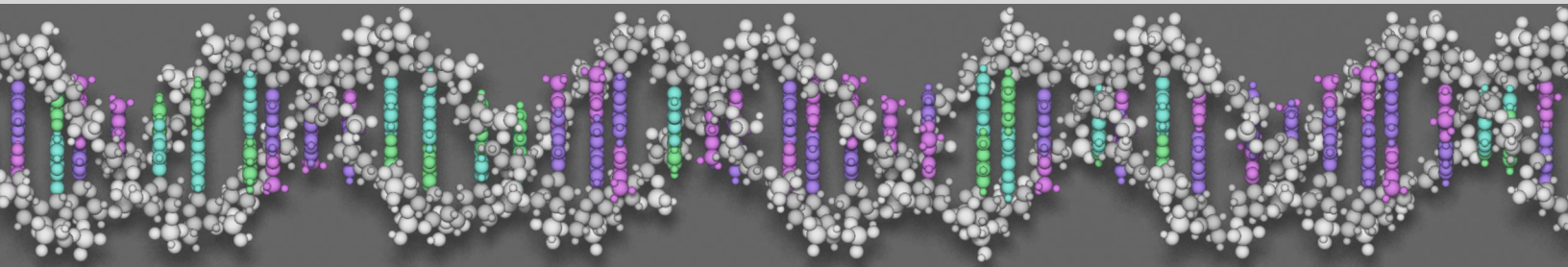
Deputy director for research, Institute for Advanced Computing Applications and Technologies

"Through programs like the IACAT Fellows we can engage faculty and students across the whole Illinois campus. We'd like to do

that in partnership with these other, more specific institutes that are already here," for example by matching faculty needs with the unique capabilities of the other units or bringing I3 students into collaborative projects.

ONIC ERGENCE

by Trish Barker



DONNA COX

*Director, eDream and NCSA's
Advanced Visualization Laboratory*

"We have a common goal to help support and promote creative practice across technology and art. The different institutes are complementary and support different

communities. We don't view our institutes as competitors—rather we strengthen each other's missions and collaborate across these strengths to write proposals and foster interdisciplinary research."

"I have been an artist collaborating with scientists here at NCSA since it opened its doors in 1985. eDream was established to help extend this type of creative practice and support unique creative productions. eDream addresses traditional artists' venues, incorporating new technologies and supporting faculty and visiting artists. ICHASS is complementary, in that it is more focused on the humanities than the arts practice. eDream collaborated with I3 to develop the informatics PhD and support the art and culture concentration. IACAT and eDream converged on a common Creative Computing theme to support faculty and students. These are examples of how the separate institutes collaborate to reinforce their broad interdisciplinary goals."



GUY GARNETT

Director, Illinois Informatics Institute

"We each keep in mind our own separate visions and try to work with the other groups to enhance the overall vision. I3 and eDream are doing all sorts of things together, and it's

sometimes difficult to see what the boundaries are, and that's a good thing, because it means it's truly collaborative. We bring together all of the resources and look at how we can support projects" with collaborating faculty and students, "and being colocated makes that much easier."

"One of the things we're going to start working out is, can we get NCSA staff teaching not only students but also faculty, bringing in high-performance computing as a tool for various classes really early on. I'm always a believer that if you make it available to them, they'll figure out how to use it."



THOM DUNNING

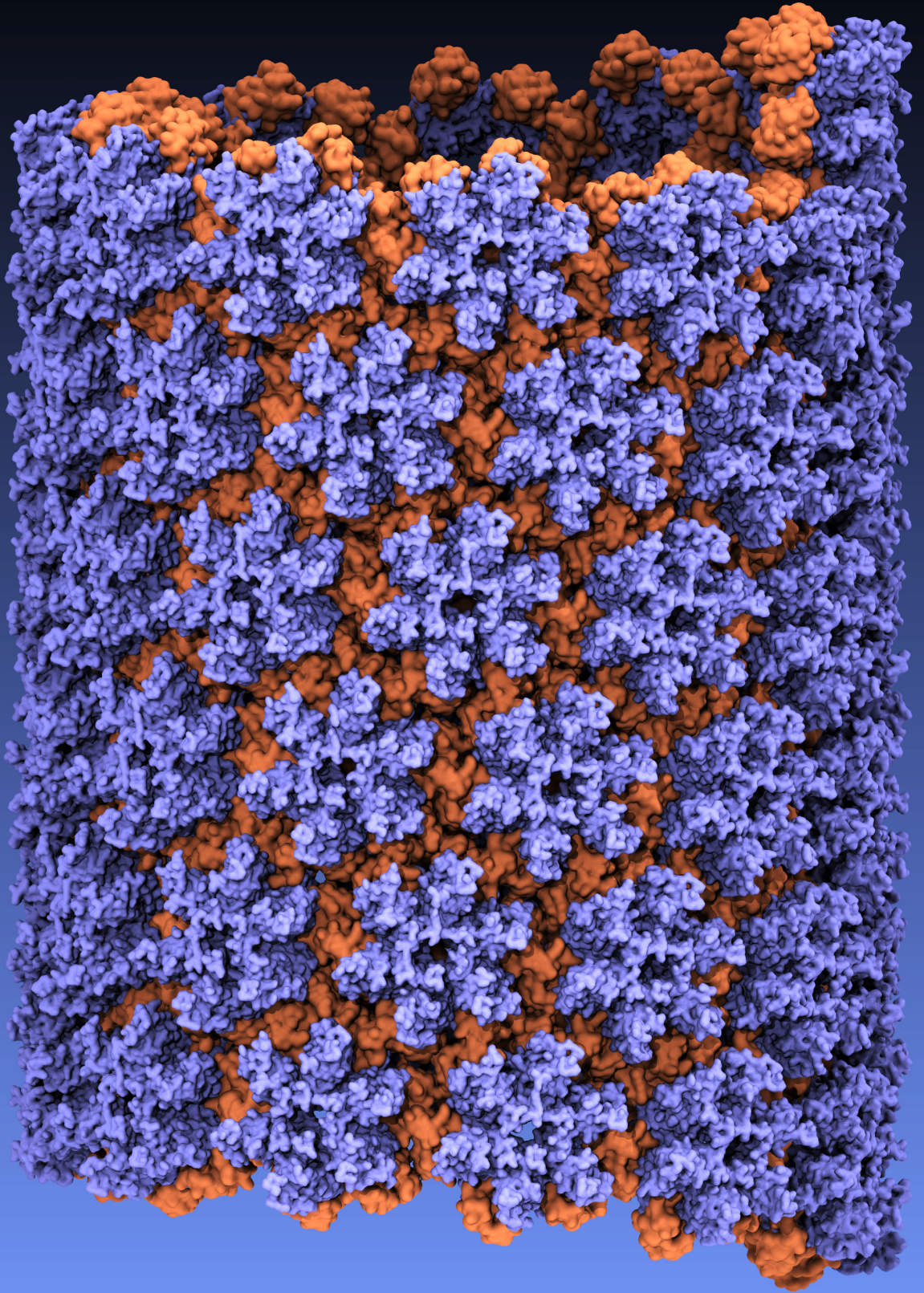
*Director, National Center for Supercomputing
Applications and Institute for Advanced
Computing Applications and Technologies*

"What this collection of institutes is doing provides a signature capability for the University of Illinois. There really isn't any other university

that is doing anything at all like this. Some of them have bits and pieces. But nobody is putting it together the way that we're putting it together here. And that's where I think the really big impact will be."

"[I3, eDream, and ICHASS are] in a building which has a rich history of computing technology. So if an issue or idea comes up, there's someone they can go to for advice and expertise—here's this idea, is it technologically feasible? And if it's not how can we make it feasible? Having the other institutes in this rich environment of computer technologists is valuable. And it exposes NCSA staff to problems that we may not currently be dealing with—our knowledge base is always enriched by looking at different types of problems." □

All-atom model of the hexameric form of the CA protein as found in cylindrical assemblies of HIV capsids in vitro. This model has been determined using the Blue Waters Early Science System, and experimental (x-ray and cryo-EM) measurements.



MAKING HISTORY

by Barbara Jewett

Early results indicate Blue Waters will live up to expectations.

A

BRIEF BUT TANTALIZING TEST RUN OF THE BLUE WATERS sustained petascale supercomputer earlier this year has left researchers eagerly awaiting full deployment. The Blue Waters Early Science System (BW-ESS) was only about 15 percent of the full machine as it consisted of 48 cabinets with 4,512 XE6 compute nodes and 96 service nodes and a Sonexion Lustre Storage Appliance provided two petabytes of disk storage. But that was enough to allow researchers to test the technology used in Blue Waters and identify and fix software bugs and other issues that prevented some codes from successfully scaling.

The results for many of the teams were astounding. Klaus Schulten, a biophysicist at the University of Illinois at Urbana-Champaign, perhaps sums it up best in his report: "Not in our wildest dreams could we have imagined the greatness of the new NSF center machine. We are sure Blue Waters will make science and engineering history."

Schulten's team pursued three projects simultaneously. Their runs studying the HIV capsid yielded some unanticipated results. Enabled by the 2.9 release of NAMD, which includes a high-performance interface to the Cray Gemini network of Blue Waters, the Schulten team was able to explore aspects of the HIV capsid's structural properties in a 100-nanosecond simulation. This simulation revealed that the capsid is stabilized in a manner other than has sometimes been proposed. A manuscript is being prepared for publication detailing their findings.

The team led by Robert Sugar of the University of California, Santa Barbara, used their BW-ESS allocation to cocomplete their calculation of the spectroscopy of charmonium, the positronium-like states of a charm quark and an anticharm quark. They did this using lattice gauge theory, an *ab initio* version of quantum chromodynamics (QCD) that can be simulated numerically. They

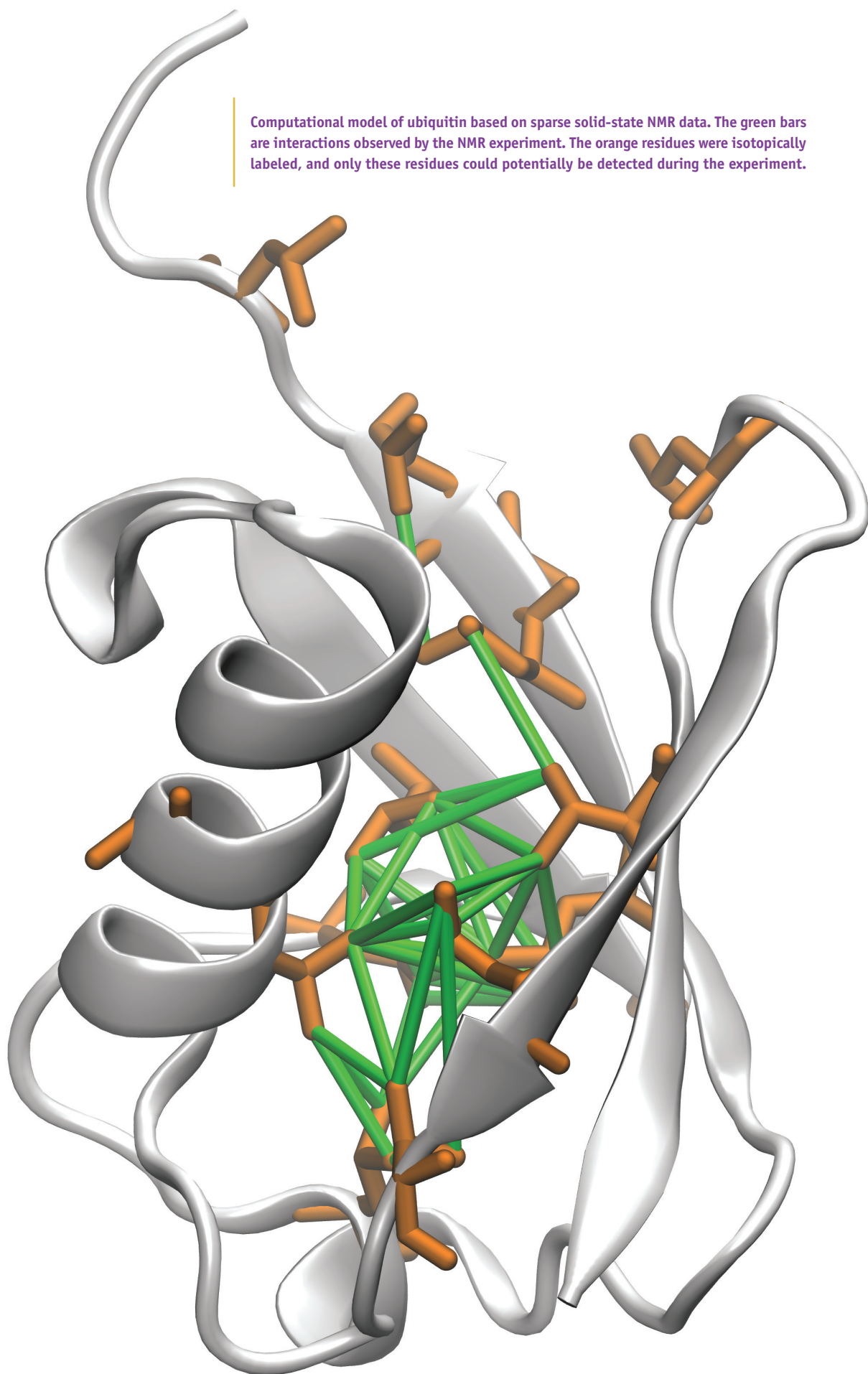
focused on the experimentally well-measured low-lying states, but say in their report that the "ultimate goal is to use the same study to help characterize some of the mysterious excited states that could be molecular states of loosely bound charmed mesons or exotic states with a valence gluon in addition to the standard valence charm quark and anti-quark."

Their simulation produced a wealth of data with, they think, some very promising results, particularly the mass splittings of the $1S$ and $1P$ states. And the team is very pleased that with Blue Waters, statistical errors have been significantly reduced from previous studies. They presented their findings at the Lattice 2012 conference in Cairns, Australia, in June.

As reported in the summer issue of Access, the team led by Stan Woosley of the University of California Observatories assisted the Blue Waters visualization staff in refining user needs in addition to making some Supernova Simulation runs. While it is generally accepted that the exploding star is a white dwarf driven to thermonuclear runaway as the result of the accretion of mass from a binary companion, says Woosley, the nature of the progenitor star and how it ignites and burns are debated.

They accomplished their main goal of simulating on the BW-ESS a large fraction of the ~ 1 second it takes the bubble to reach the surface, using higher resolution than has been previously possible. In addition, they were also interested in how the background turbulence and choice of ignition location affect the bubble's evolution. So they performed five additional simulations. A somewhat surprising result of these additional simulations was that the background turbulence only seems to affect the flame morphology for ignition close to the center. The six simulations produced over 45 TB of data and the team is certain that further analysis will yield additional interesting scientific results. \square

DOING WITH LESS



by Barbara Jewett

Researchers used NCSA's Forge to improve computational protein structure modeling.



PROTEINS, THOSE BIOLOGICAL WORKHORSE molecules that catalyze reactions, hold cells together, move muscles, and duplicate DNA, are nothing without their shape. In fact, that three-dimensional shape, called the protein structure, is the a major determinant of a protein's function.

That's why there's an important need to know what the native structures of proteins are so scientists can better understand how they work, explains Justin MacCallum, a junior fellow at the Laufer Center for Physical and Quantitative Biology at Stony Brook University. His team of researchers—including post-doc Alberto Perez and led by Laufer Center director Ken Dill—used NCSA's Forge and its GPU/CPU architecture in their quest to develop computational techniques for determining protein structure. Understanding the mechanisms of proteins helps advance fundamental biology knowledge and can also lead to practical applications, such as improved drug designs and designs of novel enzymes.

With the retirement of Forge, the work will transfer to Keeneland at the Georgia Institute of Technology. Additionally, the Laufer protein folding team received an award for two million hours on the new Titan GPU/CPU supercomputer at Oak Ridge National Laboratory. Both are XSEDE-allocated resources.

Proteins are made up of chains of amino acids held together by peptide bonds. The chains then fold, turning into their functional

shape, or structure. While the chain sequences are fairly well understood, scientists still have much to learn about how they fold and their resultant structure.

"A lot of people put a tremendous amount of time and effort into exploring protein structures," MacCallum says. "The main technique they use is X-ray crystallography. But it has one really big limitation. And that is, it's very, very difficult to get many proteins to crystallize. So even though there are extremely interesting proteins that we would like to know more about, such as membrane proteins or proteins containing intrinsically disordered regions, it is often difficult to obtain crystals that can then be examined with X-ray crystallography."

MacCallum and his team hope to make it easier to determine structures in cases where crystals cannot be obtained. They are developing theoretical techniques that can be combined with experimental techniques other than crystallography to get protein structures. He uses a hybrid approach that combines detailed molecular dynamics simulations with distant restraints derived from bioinformatics, evolution, and experiments. The restraints serve to restrict the size of the conformational space to be searched and make the computation traceable on current computer hardware. They're hoping, he says, to find that sweet spot where researchers can use experiments to tell them something about the structure, and then use computational modeling to fill in the gaps.

Challenging folds

Currently, it is extremely difficult to get the structure of an unknown protein working only from its sequence. The problem, says MacCallum, is that even with GPUs, supercomputers still aren't fast enough, as most proteins take longer to fold than can be simulated. Computer simulations can be done in the hundreds of nanoseconds to microseconds timescale, but most proteins take milliseconds or longer to fold.

MacCallum's application runs about 100 times faster on GPUs than CPUs, so Forge "really lets us tackle problems we couldn't address before" he notes, adding that what would take years to run on a CPU cluster is reduced to just weeks on Forge.

"If we have a good amount of experimental data that's all consistent, we don't need to have much computer power to get a structure, and we don't need to have very accurate modeling tools to get reasonable structures. On the other extreme, if we have no experimental data, then even with the GPU computing power of Forge we don't have enough compute power to get a structure. But there's some sort of inbetween area where, with enough computing power, you can get by with a lot less data than you used to need in order to get a structure. And that's the area we're exploring. We're trying to see how far we can push, and how little data we can put into these calculations, and still get something reasonable out of it. And the faster the computers are, the less data we need," says MacCallum.

"The cost of the experimental work is much, much higher than the cost of the computer time," he continues. "So we can save on computer costs by making this calculation faster, but we can also save an awful lot on bench time if we can make this work. And that's what we're here for, to help the experimentalist."

By giving the computer even just a little bit of data about the protein, though, the tool MacCallum's team is developing can zero in on the structures that are the most likely outcomes. In some cases, says MacCallum, they've been able to get reasonable structures with surprisingly little data. For example, one protein with which they were working had some sparse data available from solid-state nuclear magnetic resonance (NMR) experiments. That protein folds on a millisecond timescale. By inputting the NMR data into the tool, the protein folded on a 50-nanosecond timescale.

The CASP challenge

Trying to get something from almost nothing is what makes the CASP (Critical Assessment of Protein Structure Prediction) competition so interesting. CASP competitors are given the sequence of a protein and have 21 days to computationally determine a structure. The structures of the proteins have been identified through experimental means, but the structures have not yet been entered into the Protein Data Bank (PDB), so there's no way to see what the correct structure is until the competition ends. Thus the biannual CASP blind prediction competition serves as sort of a benchmarking exercise, says MacCallum, as well as providing information on the current state of the art of protein structure prediction.

For the current CASP10 competition, the sequences of structures recently solved by crystallography or NMR were made available to predictors from April through July 2012. Throughout

the fall, independent assessors in each of the prediction categories are processing and evaluating the tens of thousands of models submitted by approximately 200 prediction groups worldwide, including MacCallum's group. The results will be made public and discussed at the CASP10 meeting to be held in December in Gaeta, Italy. About this same time, the structures of the 114 competition proteins will be entered into the PDB.

The PDB archive is the worldwide repository of information about the 3D structures of large biological molecules, including proteins and nucleic acids. It was established in 1971 at Brookhaven National Laboratory and originally contained seven structures. In 1998, the Research Collaboratory for Structural Bioinformatics at Rutgers became responsible for the management of the PDB, which now contains more than 77,000 structures.

One change to this year's CASP competition was especially useful to the Stony Brook group. For the first time they had some target sequences that came with data. Not real experimental data, says MacCallum, but information that one might be able to get from an experiment. CASP organizers wanted to see if extra information improved the outcomes. That was exciting for MacCallum's team because they've developed their methods around the idea of having extra data to form their predictions. MacCallum says that was a good way to test their tool and they were pleased with the one result they know to date. So they believe they are on the right track.

"After two years of development, we have the tools working pretty well right now," he says. "Obviously, there's still going to be tweaks and improvements to come. We've talked with some experimentalists and hope they'll collaborate with us so we can try our tools on real problems. We've taken a protein where the structure is known and we've proven that we can get the same answer with our technique. Now we want to start applying this to proteins where nobody knows the structure. We hope some really interesting biology and biochemistry will come out of that." □

PROJECT AT A GLANCE

TEAM MEMBERS

Justin MacCallum
Alberto Perez
Ken Dill

FUNDING

National Institutes of Health
Laufer Center for Physical and Quantitative Biology

FOR MORE INFORMATION

<http://laufercenter.org>

ALAN CRAIG JOINS XSEDE AS DIGITAL HUMANITIES SPECIALIST



Alan Craig stands on the beach of the Arctic Ocean at Barrow, Alaska, where he assisted with a camp that helped students learn about climate change during July 2012. The camp was a part of an ongoing collaboration that includes I-CHASS, Ilisagvik College, the American Indian Higher Education Consortium, and the National Museum of the American Indian. (Photo courtesy of Kate Cooper.)

Alan Craig, research scientist at the National Center for Supercomputing Applications (NCSA), recently was named Digital Humanities Specialist for the Extreme Science and Engineering Discovery Environment (XSEDE) project. In this new role, he will bridge the gap between traditional supercomputing users and those in the humanities, arts, and social sciences (HASS). Craig brings 25 years of experience connecting users to high-performance and high-throughput computing (HPC and HTC) resources, and his appointment will allow him to promote and facilitate the use of XSEDE resources among researchers who traditionally have not been users of HPC and HTC.

"I think the fact that XSEDE opened this position attests to the fact that XSEDE takes the needs of the HASS community seriously and that it can benefit both the XSEDE community and the HASS community to work together," he says, noting that his goal in his new appointment is to help users communicate the "hidden story" in their data using visualizations and make their data widely accessible through online publication.

Novel applications of technology, especially in education, permeate Craig's work. He helped pioneer the use of sound to depict

scientific data. Now he is blurring the boundary between the real and digital worlds using augmented reality, which superimposes artificial enhancements on reality in real time using any of the five senses. For more information on his work, see the story on page 6 in this issue of Access.

Craig also serves as the senior associate director for human-computer interaction for the Institute for Computing in Humanities, Arts, and Social Science (I-CHASS) at the University of Illinois and as principal investigator for the National Science Foundation-funded Early Concept Grant for Exploratory Research (EAGER), awarded through I-CHASS to support development of an augmented reality application for mobile devices like smartphones and iPads. For a new project, Craig will create a multimedia database using XSEDE's Gordon, a data-intensive supercomputer at the San Diego Supercomputer Center that went into production at the beginning of 2012. The database will facilitate academic analysis of digital media.

NCSA, KISTI FORM JOINT LAB



KISTI President Dr. Young-Seo Park (left) and NCSA Executive Director Danny Powell

NCSA and the Korea Institute of Science and Technology Information (KISTI) will undertake collaborative research and development projects through a newly formed joint laboratory.

An agreement forming the KISTI-NCSA Joint Lab was signed in Korea in May. The joint venture will focus on developing technology related to supercomputing, potentially including application codes, research and education tools and interfaces, and programming environments.

"This is an expansion of the fruitful relationship that NCSA and KISTI have had going back many years," explained NCSA Executive

director Danny Powell. “The Joint Lab will take our collaboration to a new level and will spur significant cooperative efforts.”

The collaborators hope to have various KISTI staff members stationed at NCSA. Staff from the two centers also will gather for annual meetings, alternating between Korea and the United States.

SAAHPC 2012 BEST PAPER WINS PROCESSOR

The award for Best GPU Paper at the 2012 Symposium on Application Accelerators in High-Performance Computing went to “Power Aware Computing on GPUs” by the University of Tennessee team of Kiran Kasichayanula, Dan Terpstra, Piotr Luszczek, Stan Tomov, Shirley Moore and Gregory Peterson. The winners received an NVIDIA Tesla C2075 GPU computing processor, courtesy of NVIDIA.

ADDITIONAL ILLINOIS FELLOWS SELECTED

The Institute for Advanced Computing Applications and Technologies (IACAT) has selected six additional researchers from the University of Illinois at Urbana-Champaign to receive one-year fellowships to pursue collaborative projects with the NCSA.

Seven fellows were previously selected to work on projects in animal science, chemistry, astronomy, and other areas. The additional fellowships were made possible by funding from the offices of the University’s Provost and the Vice Chancellor for Research.

The six new fellows and their projects are:

- **Weng Cho Chew** (Electrical and Computer Engineering), working with **Narayan Aluru** (Mechanical Science and Engineering) and NCSA’s **John Towns**, aims to improve two algorithms used to model Casimir forces, which arise between objects at very small length scales and can affect performance of micro- and nanoelectromechanical systems. The team’s goal is to increase by orders of magnitude the size of the problems that can be solved.
- **Robert B. Haber** (Mechanical Science and Engineering) will work with NCSA’s **Volodymyr Kindratenko** to explore the use of graphics-processing units (GPUs) to accelerate the performance of a solver used to model hyperbolic systems.
- **Xiuling Li** (Electrical and Computer Engineering), in collaboration with **K. Jimmy Hsia** (Mechanical Science and Engineering) and **Seid Koric** (NCSA), will model the rolling process of strain-induced rolled-up micro- and nanostructures. Understanding of this process could provide guidance for the design and assembly of nanostructures into systems that can adapt to environmental changes in heat, optical radiation, and mechanical actuation.
- **Robert Markley** (English) will team with **Kenton McHenry**, leader of NCSA’s Image, Spatial, Data Analysis Group, and **Michael**

Simeone of the Institute for Computing in the Humanities, Arts, and Social Science (I-CHASS) to build on a previous computational analysis of 17th and 18th-century maps of the Great Lakes region, adding approximately 400 maps to the data set to test a hypothesis that the maps encode short-term meteorological and long-term climatological data in the form of variations in cartographic depictions of coastlines, islands, and water passages among the lakes.

- **Surangi W. Punyasena** (Plant Biology) will work with **David Tcheng** (Illinois Informatics Institute) and NCSA’s **Michael Welge**. The team plans to use advanced machine learning algorithms to develop a high-throughput automated system to count pollen and spore samples from fossils. This will make it feasible to create large data sets, which is impractical when counts are taken by individual highly trained experts using light microscopes.
- **Ted Underwood** (English) will work with **Loretta Auvil** and **Boris Capitanu** from the Illinois Informatics Institute and NCSA’s **Mark Straka** to scale up an algorithm for topic modeling—a technique for mapping the topics that could have produced a given collection of documents. The research team previously established that this method can reveal significant patterns in literary history.

IFORGE CLUSTER UPGRADED



iForge, the high-performance computing system designed for NCSA’s industry partners, was recently upgraded with new processors and increased capacity.

128 dual-socket Intel blades (Dell PowerEdge M620) replace the previous 116 dual-socket Intel blades (Dell PowerEdge M610), three quad-socket Intel blades (Dell PowerEdge M910), and four quad-socket AMD servers (Dell PowerEdge C6145). Two quad-socket AMD nodes (Dell PowerEdge C6145) from the previous configuration are being retained. The new blades also feature more memory than their predecessors (128 GB/node), more memory bandwidth, and higher memory clock speed.

This swap increases the capacity of the dual-socket blade infrastructure from 1,392 cores to 2,048—approximately a 50 percent capacity upgrade. The new servers feature Intel's Sandy Bridge CPU architecture, which offers a 33 percent to 80 percent speedup compared to the previous "Westmere" CPUs.

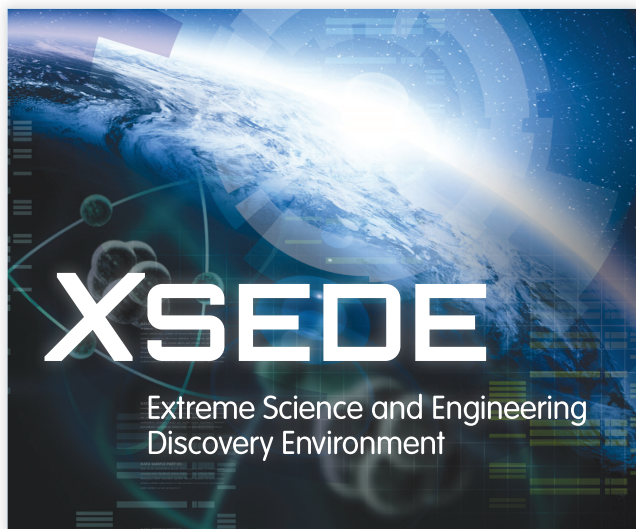
All servers in iForge will retain the 40 gigabit/sec QDR Infiniband fabric provided by Mellanox; the two quad-socket AMD nodes have been upgraded to feature dual QDR Infiniband connections.

iForge servers are also now "stateless," meaning they run without local hard disks. Instead, users store files and launch applications from a 700 terabyte GPFS parallel filesystem.

iForge also has been updated with Red Hat Enterprise Linux 6.2, and various software development tools (compilers, math libraries, etc.) have also been upgraded.

For more information about iForge and the NCSA Private Sector Program, see industry.ncsa.illinois.edu or contact Merle Giles, mjgiles@ncsa.illinois.edu.

CAMPUS CHAMPIONS FELLOWS NAMED



The National Science Foundation's Extreme Science and Engineering Discovery Environment (XSEDE) program has named the first Fellows for its Campus Champions (CC) program, pairing a Champion with a member of XSEDE's Extended Collaborative Support Services (ECSS) staff to work on real-world science and engineering projects for about one year.

Campus Champions are volunteers who advise researchers on the use of high-end cyberinfrastructure (including XSEDE resources) at their respective campuses. The goal of the CC Fellows program is to increase expertise on campuses by including CCs as partners in XSEDE's ECSS projects.

Four Fellows have been selected to participate in the first round of the program:

- **Dirk Colbry**, a research specialist at the Institute for Cyber-Enabled Research at Michigan State University, paired with **Ritu Arora**, a research associate in high-performance computing at the Texas Advanced Computing Center (TACC). Colbry brings many years of high-end computing expertise to the program and hopes to learn more about large-scale data analysis. Colbry and Arora are supporting Virginia Kuhn's interactive large-scale video analytics project. Kuhn is associate director of the Institute for Multimedia Literacy and assistant professor in the School of Cinematic Arts at the University of Southern California. The project focuses on indexing, tagging, and searching vast media archives in real time, applying a hybrid process of machine analytics and crowd-sourcing tagging. The San Diego Supercomputer Center's STET data-intensive Gordon supercomputer will be used as a resource for this project.
- **Nasseer Idrisi**, an assistant professor at the University of the Virgin Islands' Center for Marine and Environmental Studies, paired with **Kwai Wong**, a computational scientist with the National Institute for Computational Sciences (NICS). They are supporting principal investigators John Bryant Drake and Joshua Fu at the University of Tennessee Knoxville, both of whom are doing climate simulation. Working with Wong, Idrisi will expand his knowledge in parallel programming while contributing a domain viewpoint to the projects.
- **Liwen Shih**, professor and computer engineering chair at the University of Houston-Clear Lake, paired with **Yifeng Cui**, a research scientist at SDSC. Their project focuses on Cui's work on physics-based seismic research in collaboration with SDSC and the Southern California Earthquake Center (SCEC). Thomas Jordan at USC is the SCEC principal investigator. Phil Maechling is SCEC's information technology architect and also a member of the XSEDE Advisory Board.
- **Jack Smith**, a research staff member with Marshall University and cyberinfrastructure coordinator at the West Virginia Higher Education Policy Commission, paired with **Yaakoub El Khamra**, a research engineering/scientist associate with TACC. Smith has extensive experience in many programming areas, as well as chemistry and material science, and would like to use knowledge gained in the Fellows program to grow research programs in new areas. They are working on the projects of Ronald Levy (Rutgers) and Thomas Bishop (Louisiana Tech). Both involve the use of SAGA software to manage large numbers of loosely coupled bioinformatics calculations.

Accepted Fellows, with the support of their home institution, make a 400-hour time commitment and are paid a stipend to allow them to focus time and attention on these collaborations. The program also includes funding for two visits, each ranging from one to two weeks, to an ECSS site to enhance the collaboration. Funding is provided to attend and present at a Fellows symposium at the annual XSEDE conference.

Further information on the Campus Champions Fellows program is online at www.xsede.org/ccfellows.

COOPERATION AND IDEAS



The first CHANGES workshop was a high-level platform to discuss the latest trends in supercomputing, sophisticated information techniques, interdisciplinary applications, e-Science and its applications in China, Germany, and the United States. CHANGES stands for CHinese-AmericaN-German E-Science and cyberinfrastructure. Partners are CNIC-CAS from China, NCSA-Illinois from the United States, and Jülich Supercomputing Centre at Forschungszentrum Jülich from Germany. Besides the presentations, the workshop provided a forum for trilateral cooperation on student exchange and mutual research projects.

NEW PROJECT FOR SCIENCE CYBERSECURITY

NCSA's cybersecurity group is collaborating with Indiana University to form the Center for Trustworthy Scientific Cyberinfrastructure (CTSC) using a three-year, nearly \$4.3 million grant from the National Science Foundation.

CTSC will help NSF cyberinfrastructure projects improve security by offering assistance in identifying problems and designing and implementing solutions. Six projects are already identified for the first year, and future projects will be evaluated through an open call.

The center will also expand education and training opportunities in cybersecurity. Training materials that include best practices guides, case studies, and lessons learned throughout the project will target project managers, developers, and operations staff. In collaboration with the University of Illinois, the center will support students with new cybersecurity course materials, especially for scientific computing, that will be available to other institutions.

Von Welch, deputy director of Indiana University's Center for Applied Cybersecurity Research, is CTSC principal investigator.

Welch was formerly at NCSA. Randy Butler, director of NCSA's Cybersecurity Directorate and CTSC deputy director, will lead training and education activities. Team members include NCSA's Jim Basney, who specializes in identity management, Barton Miller of the University of Wisconsin-Madison, a software security expert, and Jim Marsteller from the Pittsburgh Supercomputing Center who works in operational security. Scott Koranda from the University of Wisconsin-Milwaukee supplies an inside "scientific needs" view from previous work with the Laser Interferometer Gravitational-Wave Observatory (LIGO), one of the initial six projects.

ILLINOIS PROFESSORS RECEIVE FERNBACH AWARD



Laxmikant "Sanjay" Kale



Klaus Schulten

Two University of Illinois professors—Laxmikant "Sanjay" Kale and Klaus Schulten—have been named the recipients of the 2012 IEEE Computer Society Sidney Fernbach Award, "for outstanding contributions to the development of widely used parallel software for large biomolecular systems simulation."

Established in 1992 in memory of high-performance computing pioneer Sidney Fernbach, the Fernbach Award recognizes outstanding contributions in the application of high-performance computers using innovative approaches. Kale and Schulten accepted the award at SC12 in Salt Lake City.

Kale is a professor of computer science, director of the Parallel Programming Laboratory, and a senior investigator for the Blue Waters project. His parallel computing work focuses on enhancing performance and productivity via adaptive runtime systems, with research on programming abstractions, dynamic load balancing, fault tolerance, and power management. These research results are embodied in Charm++, a widely distributed parallel programming system.

He has collaboratively developed applications for biomolecular modeling (NAMD), computational cosmology, quantum chemistry,

rocket simulation, and unstructured meshes. He is a co-winner of the 2002 Gordon Bell award. Kale and his team won the HPC Challenge Best Performance award at Supercomputing 2011 for their entry based on Charm++.

Schulten is a Swanlund Professor of Physics, directs the Center for Biomolecular Modeling at the Beckman Institute, and co-directs the Center for the Physics of Living Cells in the Department of Physics. His research, focused on molecular assembly and cooperation in biological cells, requires large-scale computing. He was the first to demonstrate that parallel computers can be practically employed to solve the classical many-body problem in biomolecular modeling. Thousands of researchers worldwide use his group's software in molecular graphics (VMD) and modeling (NAMD) on personal computers as well as at the world's leading supercomputing centers.

Presently his group is developing a new computational method that assists biologists in solving the structures of the very large macromolecular complexes forming the machinery of living cells.

SPIN AT NCSA



Over 100 University of Illinois undergraduate students turned out to learn more about a new student internship program being launched at NCSA. Students Pushing Innovation, known as SPIN, will provide paid internships for undergraduates to work with NCSA staff on a variety of projects for the Spring 2013 semester. Students had the opportunity to discuss potential projects with the NCSA staff who will serve as SPIN mentors.

PETASCALE DAY

Petascale refers to computing and data in the quadrillions, like the more than 11 quadrillion calculations per second Blue Waters can perform and the more than 380 quadrillion bytes available in NCSA's new tape archive.



In scientific notation, 1 quadrillion is 10 to the 15th. So on 10.15 (October 15) NCSA celebrated PETASCALE DAY! A variety of informative and fun activities for NCSA staff and the local community were held, ranging from a lightning talk on "What is a Quadrillion?" to science presentations, 3D scientific visualization demonstrations, tours of Blue Waters, and free movie screenings—of the 2012 documentary "Dynamic Earth" and the 1983 classic "War Games."

DUNNING TO RETIRE IN 2013

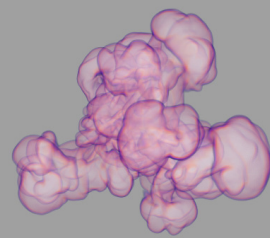
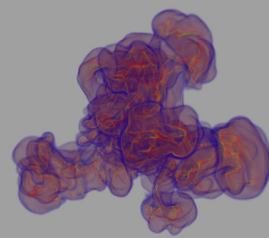
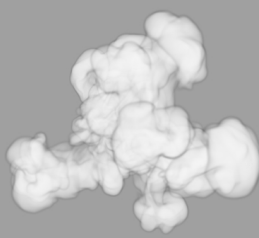
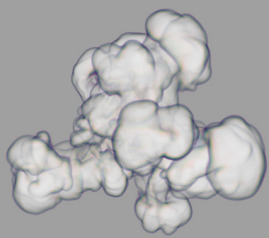
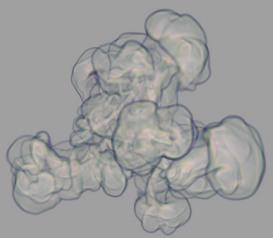


After nearly eight years at the helm of NCSA, Thom H. Dunning Jr. has announced that he will retire as the center's director in 2013. He will remain as NCSA's leader until the University of Illinois completes a national search for his successor, and then plans to continue as a professor in the University's Department of Chemistry.

In the past eight years, NCSA has attracted more than \$394 million in federal grant funding, and is currently leading the two largest computing projects ever funded by the National Science Foundation: the Blue Waters project and the Extreme Science & Engineering Discovery Environment (XSEDE) project.

Dunning is NCSA's fourth director. Larry Smarr led the center from its inception in 1985 to 2000, Dan Reed from 2000-2003, and current NCSA chief technology officer Rob Pennington served as interim director in 2004.

DUO IMAGE

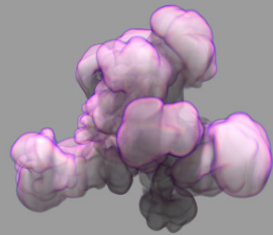
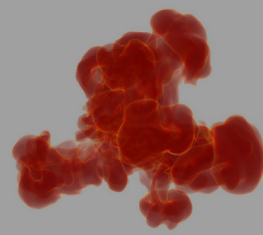
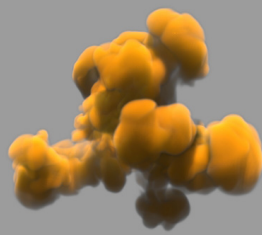
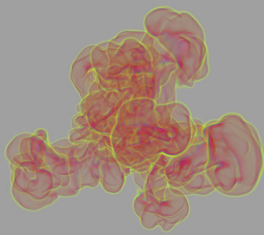
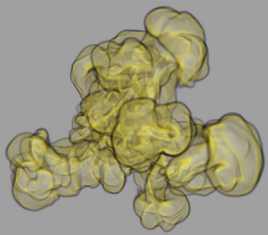
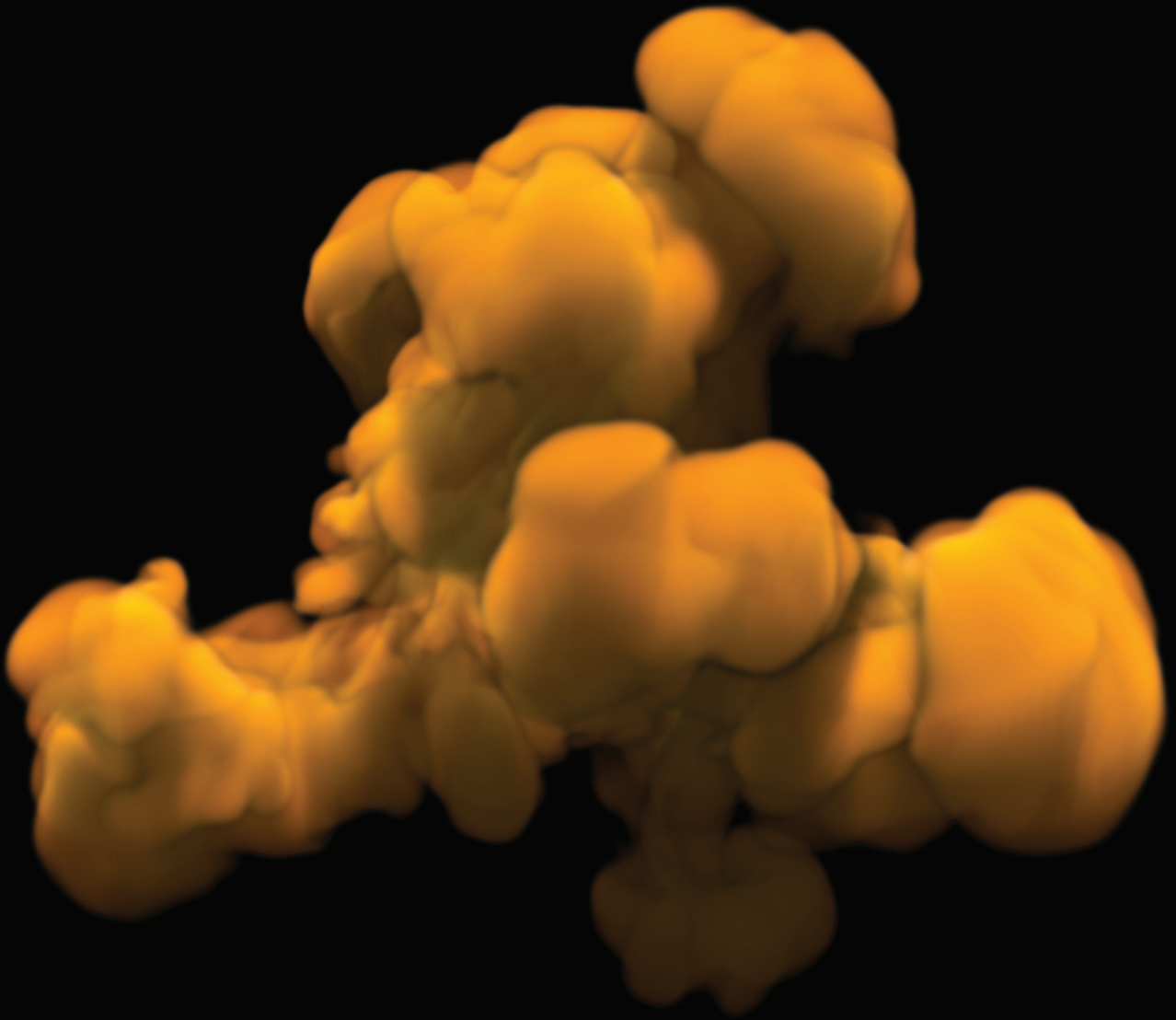
**F**

FROM THE TIME NCSA OPENED ITS DOORS, SCIENTIFIC visualization has played an important role. The visualization team of David Bock and Mark Van Moer provides technical expertise and assistance to computational scientists and researchers desiring to represent and analyze data generated by computational models. This facilitates the understanding of the simulated phenomena.

Their expertise has long been available to those using NCSA computational resources as they, along with Dave Semeraro and Rob

Sisneros, comprise NCSA's Advanced Digital Services Visualization Group. Since July 2011, Bock and Van Moer have been part of the XSEDE project's Extended Collaborative Support Service, expanding the base of those who can take advantage of their special skills.

Bock and Van Moer work with users to identify the best tools for their needs and assist in effective operation of visualization software. In addition, they also develop visualization solutions using custom visualization software to generate publication-quality imagery and animations for posters, papers, and journal articles.



For example, Bock has been working with astrophysicist Dean Townsley of the University of Alabama to visualize his simulations of a supernova. Using a custom volume rendering system he developed, Bock employs a variety of techniques to provide better understanding of the data for the researchers. Shown on this page are images from a data set where Bock investigated volume representations by experimenting first with various color and alpha transfer functions without taking into account physically-based lighting using default absorption, scattering, and emission values, then added lighting and self-shadowing to the

volume rendering scene. The large image above uses the same scene above but with the Rayleigh phase function applied to approximate the light scattering behavior, and changing the color map to orange. □



National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign

1205 West Clark Street
Urbana, IL 61801

www.ncsa.illinois.edu

