XSEDE

ACCELERATING **SCIENTIFIC** DISCOVERY

XSEDE

Extreme Science and Engineering
Discovery Environment

ON THE COVER Large-scale view of the simulation volume, centered on the most massive galaxy cluster in the simulation at the present cosmic time. Dark matter density is shown in blue and purple, and the velocity of normal matter (gas) is shown in red and orange. Image credit: Illustris Collaboration. *Full story on page 2.*

XSEDE Partner Institutions

Cornell University Center for Advanced Computing

Indiana University

Jülich Supercomputing Centre

Michigan State University

National Center for Atmospheric Research

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

National Institute for Computational Sciences - University of Tennessee Knoxville/Oak Ridge National Laboratory

Ohio Supercomputer Center - The Ohio State University

Pittsburgh Supercomputing Center - Carnegie Mellon University/ University of Pittsburgh

Purdue University

Rice University

San Diego Supercomputer Center - University of California, San Diego

Shodor Education Foundation

Southeastern Universities Research Association

Texas Advanced Computing Center - The University of Texas at Austin

University of California, Berkeley

University of Chicago

University of Southern California

University of Virginia

XSEDE provided advanced digital services—encompassing supercomputers, a tape system, fast network connection, and efficient transfer of central-processing-unit time from machine to machine—that accelerated completion of a galaxy formation simulation by approximately a year.

THE ILLUSTRIS PROJECT

An international team of researchers used resources from XSEDE to develop components that would serve as the basis for "Illustris," the most ambitious simulation of galaxy formation ever done. The project is described in a paper published on May 8, 2014, in the journal *Nature*.

This work has received worldwide news media coverage, including The New York Times, Los Angeles Times, The Washington Post, CNN, BBC News, Der Spiegel (in Germany), Le Figaro (in France), The Times of India, and many others.

Mark Vogelsberger of the Massachusetts Institute of Technology (MIT) led the project, having initiated it while he was still a postdoctoral researcher at the Harvard-Smithsonian Center for Astrophysics (CfA). "Illustris represents the most detailed model of the universe," he says. "It took our international team, including major contributions from Volker Springel at the Heidelberg Institute for Theoretical Studies (HITS) in Germany, about five years to prepare and execute the simulation."

An excerpt from a news release issued by CfA conveys the commitment and computational power necessary for an undertaking of this magnitude: "The actual calculations took three months of 'run time,' using a total of 8,000 CPUs running in parallel. If they had used an average desktop computer, the calculations would have taken more than 2,000 years to complete."

Like a Time Machine

The CfA news release also expresses the unique perspective that Illustris offers: "Since light travels at a fixed speed, the farther away astronomers look, the farther back in time they can see. A galaxy one billion light-years away is seen as it was a billion years ago. Telescopes like Hubble can give us views of the early universe by looking to greater distances. However, astronomers can't use Hubble to follow the evolution of a single galaxy over time."

Illustris allows one to journey back and see in high detail our universe 12 million years after the Big Bang and then watch the cosmos evolve over a period of 13.8 billion years. The simulated volume captures tens of thousands of galaxies with 12 billion resolution elements (pixels) in a cube running 350 million light-years (106.5 megaparsecs) across.

"We can pause the simulation and zoom into a single galaxy or galaxy cluster to see what's really going on," says Shy Genel of CfA.

This simulation is distinguished not only by how comprehensive it is but also by the fact that it creates a mixed population of elliptical and spiral galaxies and shows small-scale evolution of gas and stars.

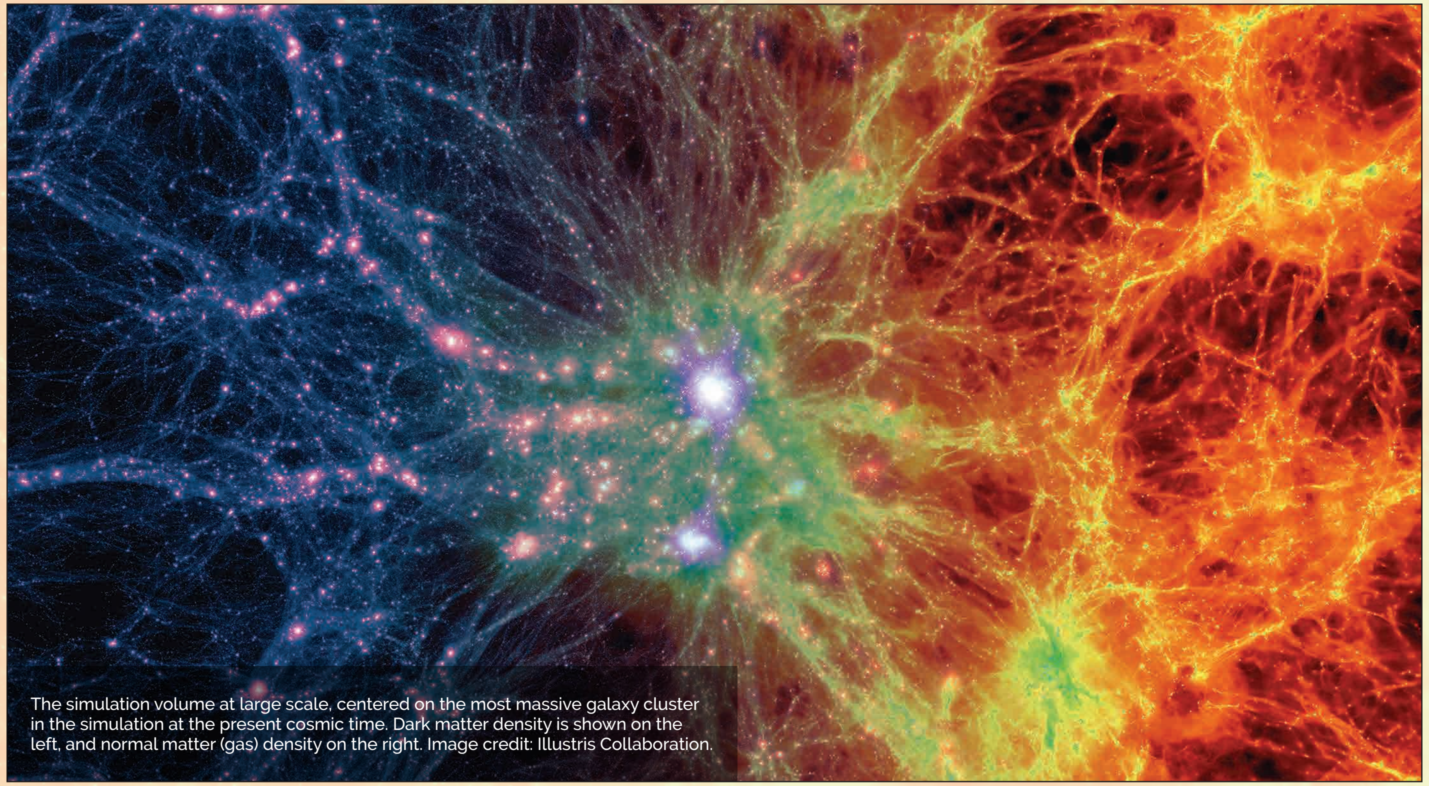
"It is crucial to have these self-consistent models which include the dark components and baryonic physics in one complete simulation," Vogelsberger says.

Important Early-phase Simulations

The XSEDE advanced computing resources Kraken at the National Institute for Computational Sciences (NICS), and Ranger and Stampede at the Texas Advanced Computing Center (TACC) supported three sets of simulations that were essential to the project.

The first contribution of XSEDE came in the form of a series of small-volume simulations that were run during the development phase of the physical model; these were critical to understanding different aspects of how the researchers' formulations of astrophysical processes affect the resulting simulated galaxy populations. In the end, the fiducial physical model—that is, the fixed basis of reference or comparison—that would be used for the production simulations was developed based on these tests.

The second aspect of XSEDE's assistance was a series of simulations identical to the "flagship" one, which was accomplished with resources from the Partnership for Advanced Computing in Europe (PRACE) except employing a smaller number of pixels and thus resulting in coarser



The simulation volume at large scale, centered on the most massive galaxy cluster in the simulation at the present cosmic time. Dark matter density is shown on the left, and normal matter (gas) density on the right. Image credit: Illustris Collaboration.

details. These simulations, however, were completed about six months before the flagship and produced very similar results, Genel says. The similarity, he adds, allowed the researchers to develop and test post-processing analysis tools, and by applying them on these lower-resolution simulations, to already have a good idea of what the results of the flagship simulation would look like. This in turn enabled them to move the project in what they considered to be the most exciting directions.

The third area of support consisted of simulations that follow the same large cosmological volume as the other production runs, but with physics that were less complete, namely excluding the formation of stars and black holes, or normal matter altogether—keeping only dark matter, the mysterious and invisible substance that comprises an estimated 23 percent of the cosmos. Through a direct comparison of this set of simulations with the others, the researchers were able to gain insight into the effects of the components.

"Each of these sets of early-phase simulations was very important to our project as a whole," Genel says. "Without having that first and second set of simulations, our project would have advanced significantly slower—probably on the order of a year or so."

He explains the project's use of MPI and OpenMP (an application program interface): "The memory consumption of our code that runs the simulation changes during the run as cosmic structures become more and more complex. When our simulation approaches the current age of the universe, we sometimes need to increase the total memory available to it by increasing the number of CPUs. Using a combination of MPI and OpenMP, we are able to also efficiently use the compute power that comes with those additional CPUs, not only their internal memory."

The integrated nature of XSEDE's advanced digital services was advantageous to the Illustris project. "On top of using the machines for the calculations themselves, we also used the tape system, as well as the fast connection of the XSEDE network, which was very handy for us," Genel says. "Also, the flexibility of transferring allocated CPU time from one machine to the other is a feature we used several times, which has contributed to the global efficiency of using the CPU time we were granted."

XSEDE Campus Champions help to not only improve a scientific research team's work, but help grow the base of computational science users through outreach and networking across the country.

ENGINE OF COLLABORATION

Curtis Marean had a problem.

The Arizona State University archeologist was leading a massive international collaboration studying early human settlements in the Cape Floral Region of South Africa. Evidence suggested that extreme aridity about 50,000 to 150,000 years ago reduced *Homo sapiens* to 15,000 or fewer survivors in that tiny region. Surviving that near-extinction changed us somehow, made us capable of unprecedented levels of cooperation and advanced planning.

But how did the emergence of "modern behavior" happen?

"Our project began as a straight archeological dig," Marean says. "Then I realized that we needed much better climate and environmental contextual data to understand the archeological record we were excavating."

XSEDE Campus Champion Eric Shook was instrumental in helping Marean's team create those data. An assistant professor of geography at Kent State University, Shook helped the collaborators adapt their models for supercomputers. But just as importantly, he helped a group of eminent field researchers become enthusiastic and sophisticated consumers of high-performance computing (HPC).

Preaching the Gospel of HPC

Marean's story isn't unusual, according to Kay Hunt of Purdue University, project coordinator for the Campus Champions Program since its inception in 2008. Sometimes, help from a peer can make all the difference in becoming an HPC pro.

In Campus Champions, she says, "We identify a local person at a campus or institution who's going to let colleagues know that XSEDE offers resources at no cost and how to get access, serve as the point person to get them started and give feedback to XSEDE to make the whole experience better."

The Campus Champions Program has also expanded to include Student Champions, who help reach students; Regional Champions, who coordinate Campus Champions across seven national regions; and Domain Champions, who concentrate on specific scientific disciplines. But the Campus Champions remain the anchor, preaching the HPC gospel.

Champion	Mentor	Institution
Nitin Sukhija	Trey Breckenridge	Mississippi State University
Shawn Coleman	Jeff Pummill	University of Arkansas
Genaro Hernandez	Paul Schou	University of Maryland-Baltimore County
Tarun Kumar Sharma	Liwen Shih	University of Houston-Clear Lake

Champion	Region #	Institution
Ben Nickell	1	Idaho National Labs
Ruth Marinshaw	2	Stanford University
Aaron Bergstrom	3	University of North Dakota
Dana Brunson	4	Oklahoma State University
Mark Reed	5	University of North Carolina
Scott Hampton	6	University of Notre Dame
James Cuff	7	Harvard University

Champion	Domain	Institution
Rob Kooper	Data Analysis	University of Illinois
Mao Ye	Finance	University of Illinois
Tom Cheatham	Molecular Dynamics	University of Utah
Brian Couger	Genomics	Oklahoma State University
Virginia Kuhn	Digital Humanities	University of Southern California
Michael Simeone	Digital Humanities	Arizona State University
Sudhakar Pamidighantam	Chemistry and Material Science	Indiana University

Campus Champion and Regional Champion Dana Brunson of Oklahoma State University calls herself "the cyberevangelist for the campus." Her job, running the university's HPC center, stresses user support.

At her first HPC conference—SC08—Jeff Pummill at the University of Arkansas, a Campus Champion and coordinator of the Regional Champion Program, introduced her to Hunt, who immediately recruited her.

"The program helps me offer local people any kind of resource they need," Brunson says. "I need never say, 'Oh, I'm sorry; you can't do that research because we don't have those resources.'"

Added Organization, Enhanced Opportunity

Pummill started using HPC in 2005, when the University of Arkansas acquired its first Top500 cluster. Impressed by the Campus Champions Program's promise for breaking down academic silos, he became an "early adopter"—leading to his role as a part-time XSEDE staffer heading the Regional Champions initiative.

The Campus Champions Program had grown so well and so quickly that the need for an additional level of organization had become clear, Pummill says. Regional Champions serve as the program's eyes and ears, ensure Campus Champions get what they need, disseminate information about new resources and encourage inter-campus collaborations.

The first region to become active—Region 7, or New England—recently had its inaugural meeting. Twenty-two champions attended, representing 14 of the 17 campuses in the region.

"We were amazed by the turnout," Pummill says. Next he hopes to leverage social media to help the new Regional Champions foster communication and collaboration between their Campus Champions.

A Focus on Scientific Disciplines

The Pittsburgh Supercomputing Center's (PSC's) Sergiu Sanielevici, manager of XSEDE's Novel and Innovative Projects Program, got the idea of "domain champions" at a European Grid Initiative meeting session about their own XSEDE champions-inspired program.

"I discovered that their program was organized not around individual universities, but fields of science," he says. "They were looking for people who could talk with people in their field, championing HPC use among colleagues. It was an interesting idea."

The Domain Champions Program was set up, initially as an experiment, complementing the Campus Champions. In its first year, it has made good progress bringing XSEDE and advanced computing to the attention of prospective user communities, particularly in bioinformatics and the digital humanities

"We're looking for people who have influence in their research community, and have made productive use of XSEDE resources," Sanielevici explains. "And who are willing to spend the extra effort to reach out and make the availability and advantages of XSEDE known to people in that community."

Unprecedented Collaboration

Shook realized that his archeological collaborators, while rich in many scientific disciplines, needed particular help transitioning to being HPC users. Marean agreed. As Shook says, "I became a little bit more of a communicator in terms of how XSEDE resources could be used and how all of these pieces could be put together on a computational platform."

Shook had joined Marean's collaboration thanks to an XSEDE program that offered him additional tools: the Extended Collaborative Support Service (ECSS) Campus Champions Fellows. As a participant, he could call on his ECSS mentor, David O'Neal from PSC, for assistance.

The group successfully moved their climate model to PSC's very-large-memory Blacklight system. Shook also suggested adding the San Diego Supercomputer Center's Trestles system to amplify the scope of their agent-based model of humans, climate and food sources interacting. Marean's team thus secured a full XSEDE research allocation—with Shook as a co-principal investigator.

"In my opinion there are three evolved behavioral traits that separate modern humans from other animals: advanced cognition, a dependence on social learning, and everyday and prolonged cooperation between unrelated individuals," Marean says. "Solving the great questions of science, such as how and why hyper-prosociality evolved, rests on that self-same evolved proclivity, which we see in XSEDE and Campus Champions themselves."

XSEDE ECSS supported University of Texas Center for Transportation Research scientists to visualize and simulate solutions for traffic issues in Austin, one of the fastest growing cities in the U.S.

AS AUSTIN GROWS, SO DO ITS TRAFFIC WOES

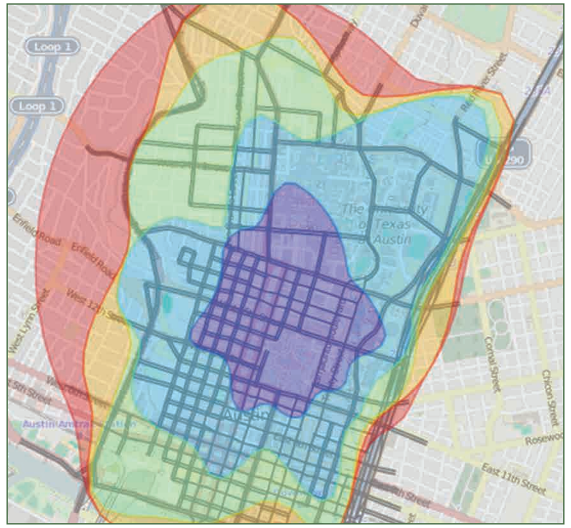
XSEDE ECSS supports UT's Center for Transportation Research to visualize, simulate solutions for traffic issues.

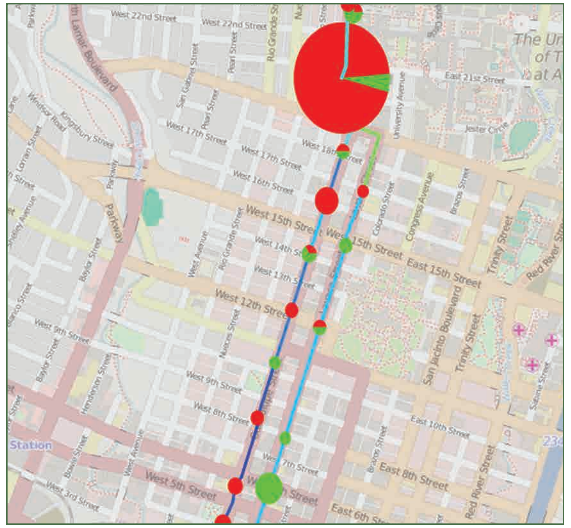
Ask any Austinite what they enjoy least about the city, and many will mention the escalating traffic issues. According to Forbes, Austin is one of the fastest growing cities in the U.S., but without a transportation infrastructure equipped to handle the explosion of new transplants, it also ranks as one of the cities with the worst traffic congestion.

The Network Modeling Center (NMC), a group of researchers within the Center for Transportation Research (CTR) at The University of Texas at Austin, is using advanced transportation models to help transportation agencies understand, compare, and evaluate alternative solutions and development strategies. However, the complex impact of alternative solutions on everyday traffic is not easy to predict.

Dynamic traffic assignment (DTA) models allow the NMC to realistically simulate the minutiae of traffic in Austin at specific locations and during specific time intervals. Although similar traffic models have been used for the past 50 years, DTA models are novel in that they account for how traffic changes at distinct time periods, for instance how congestion builds up during the morning rush hour and then slowly dissipates.

Running these DTA models left the NMC with "pages and pages of text," so they reached out to Greg Abram, a research scientist at the Texas Advanced Computing Center (TACC), through XSEDE's Extended Collaborative Support Service (ECSS) program. The program pairs expert staff members with researchers for an extended period of weeks up to a year. ECSS staff such as Abram provide expertise in many areas of advanced cyberinfrastructure and work with a research team to advance their work through that knowledge. In this case, Abram helped the researchers develop an interactive, web-based visualization tool. The visualization tool offers many unique ways to view the Austin transportation network and allows researchers to gain additional clarity into the inner-workings of Austin traffic.





The MNC's web-based visualization tool displays simulations from DTA models. The left image shows the model's estimation of the time it takes to reach different sections of downtown during the morning peak. The right image shows the model's estimation of bus passengers boarding and disembarking at each stop on a specific route. Credit: The Network Modeling Center, The University of Texas at Austin

Through an XSEDE allocation, they also began working with research scientists Weijia Xu and Amit Gupta at TACC to improve the performance of their DTA models. Xu and Gupta used TACC's Stampede supercomputer to test and develop algorithms to decrease the model simulation time.

"Sometimes our models take days to run, and if you're a decision maker that's an eternity," Natalia Ruiz-Juri, assistant director of the NMC says. "By working with the TACC team, we have seen some of our computations run five to 10 times faster."

But even with the initial improved performance, the team realized they could benefit from additional support from TACC through XSEDE's ECSS.

"The main focus of our ECSS support is a continuation of what we have already accomplished," Xu says. "We want to create new algorithms that expand the current workflow and make the models even more efficient."

Reducing model run times will enhance the group's ability to model large-scale and future traffic scenarios and congested networks. To improve the visualization tool, Abram will continue working to fine tune and add new features, like representing real-time traffic data using TACC's visualization resource, Maverick. Through ECSS, the NMC also gained access to TACC systems Ranch and Corral for increased storage capacity.

"The ECSS program has been a game changer for our research. The way I'm planning my future research steps is shaped by what I now know we can accomplish through this collaboration," Ruiz-Juri says. "It's been helpful to work with researchers that are familiar enough that they understand what we want to accomplish, but also have a fresh outlook."

One key element of XSEDE's success has been its ability to train highly-specialized scientists in the basics of using digital research resources, like supercomputers, which are now essential in many disciplines..

IT ISN'T ROCKET SCIENCE

We'll start out with a slight exaggeration: brain surgeons and rocket scientists think programming a supercomputer is difficult.

Which is a shame, says John Urbanic, parallel computing specialist at the Pittsburgh Supercomputing Center (PSC) and a lead instructor in XSEDE's programming workshop series.

"Some of the technology—in particular MPI—has a reputation for being very intimidating and difficult to use," Urbanic says. "That can turn off researchers who just want to 'do their projects.' But at every workshop, we take 150 people and convince them otherwise."

In many fields, using supercomputers is no longer a "nice to have": it's required to remain at the cutting edge of research. XSEDE's workshops provide a lifeline for gaining those skills.

"Many people in HPC might equate XSEDE with 'supercomputers' and nothing more," says Vince Betro, former computational scientist at JICS/NICS and a training manager, campus champions leadership team member, and user engagement feedback coordinator for XSEDE. "We'd like to change that."

"What good are supercomputers without the expertise to properly use them?" Betro adds. "And what good is training if you aren't there to receive it?" XSEDE's workshops provide definitive solutions to both issues, offering opportunities to attend in person or, through sponsoring campuses, remotely yet with on-site assistance. Remote, asynchronous learning is also available.

Spanning a range of needs

The XSEDE workshops span the major languages and software tools necessary for HPC, including MPI, OpenMP, OpenACC and SPARQL. The idea, says Urbanic, is to offer practical workshops to get researchers up to speed in a realistic computational setting. XSEDE enables the classes to leverage some of the premier supercomputers in the nation.

"We use whichever machine is the best platform to teach students," Urbanic says. The workshops give OpenACC students access to Blue Waters at the National Center for Supercomputing Applications (NCSA). MPI classes employ Stampede at the Texas Advanced Computing Center (TACC). "And we teach OpenMP on the world's best shared memory machine, which historically was Blacklight here at PSC, and a role that will be filled by the upcoming Bridges system."

XSEDE's training offerings enable learners to gain skills needed to be effective HPC researchers and practitioners. To recognize learners who master these skills, XSEDE has begun a program based on the Mozilla Open Badges infrastructure to issue digital badges. The first badges were issued to participants in PSC's OpenACC workshop in February 2015 and to XSEDE staff members who participated in a series of staff training events. More badges will be offered in the future to address the competencies identified in XSEDE's training roadmaps.

Avenues of access

A popular way to access the workshops is through XSEDE's asynchronous online learning systems, Cyberinfrastructure Tutor (CI-Tutor) and the Cornell Virtual Workshop (CVW).

"These systems offer self-paced courses that are free to anyone and have been accessed by learners worldwide," says Sandie Kappes, training project coordinator at NCSA and XSEDE instructional designer. "They've been used by graduate students, corporate and academic researchers, and even high school students. In addition, university professors have supplemented their course materials by referring their students to them."

Most courses include quizzes to let learners self-assess their understanding of the materials and exercises to practice what they have learned. In addition, CVW provides access to lecture videos, quizzes, and homework used in two semester-long courses, 'Applications of Parallel Computers' and 'Engineering Parallel Software,' taught at the University of California, Berkeley. These materials were made available to students within CVW during the synchronous delivery of these courses, and were then archived for asynchronous access. The courses are supported by course infrastructure, resources and homework provided by XSEDE and UC Berkeley. Partner institutions provide a local instructor who is responsible for advising and grading.

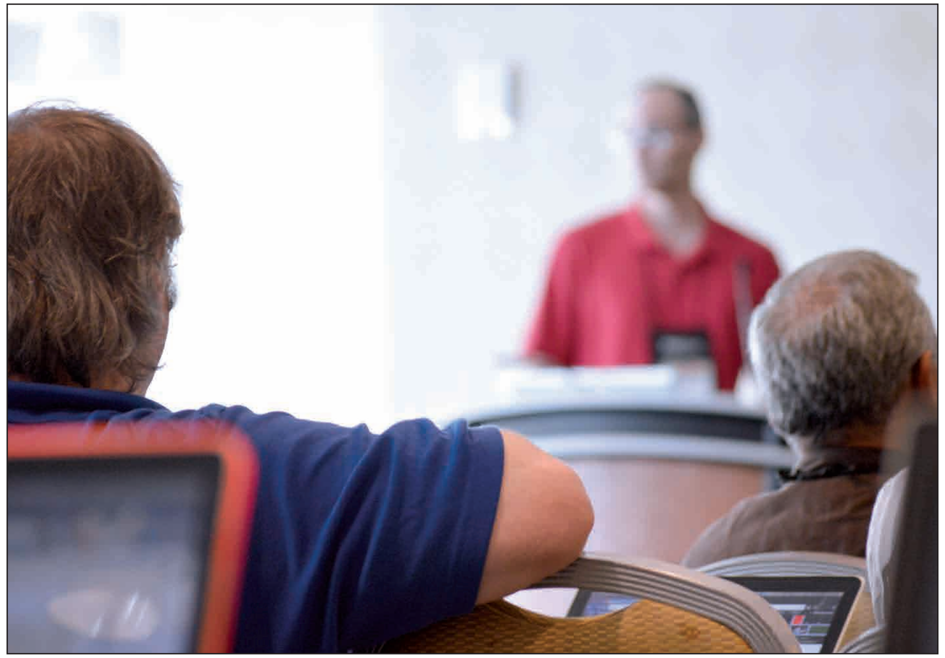
"Cornell's training and consulting staff will continue to develop and deploy Cornell Virtual Workshops for the XSEDE and Stampede programs," says Susan Mehringer, assistant director, consulting at Cornell and head of the CVW. "We'll also develop the cloud computing training for the recently announced Jetstream program."

Rave Reviews

Perhaps the best gauge of the workshops' success is the feedback XSEDE gets from attendees.

"These workshops are well known for being of the best quality," agrees Galen Collier, a computational scientist at Clemson University in South Carolina. "Everybody looks forward to these classes."

While Collier has his own research projects, his



chief role is to help other researchers on the Clemson campus get up to speed in using supercomputing resources—whether Clemson's own Palmetto cluster or XSEDE resources such as Blacklight. He says that PSC's workshops help him use his own time more effectively.

"The course was really good because it had a lot of practical applications," says Mary Krauland, a research programmer at the University of Pittsburgh's Public Health Dynamics Laboratory. "In the practical component we actually got to do the programming ourselves on several problems. It was a really nice intro." Krauland models disease transmission as part of the Models of Infectious Disease Agent Study—MIDAS—at Pitt, PSC and other institutions.

"The workshop made MPI programming a lot less intimidating by focusing on the most essential functionalities," adds Berhane Temelso, a postdoctoral associate at Bucknell University. He models how clusters of molecules in the atmosphere grow into aerosols, with applications ranging from cloud formation to pollution effects to climate change. "It was … a great workshop with the goal of demystifying MPI and encouraging users to write parallel code."

"We understand that XSEDE is one important part of a larger ecosystem of HPC and computational science cyberinfrastructure," says Betro. "It's important to remember that this is true not only in terms of supercomputers but also the people who make science happen: both researchers and enablers!"

Links to XSEDE's online courses can be found on the XSEDE User Portal at https://portal.xsede.org/web/xup/online-training or http://www.psc.edu/index.php/xsede-hpc-series-all-workshops. For those with an XSEDE User Portal account, the courses can be accessed directly from these links without requiring an additional login. Access is also available directly from Ci-Tutor at www.citutor.org and CVW at https://www.cac.cornell.edu/VW/.

Doug Jennewein has helped establish better University of South Dakota computing infrastructure through a variety of XSEDE training sessions, evangelism, and Campus Champion work, allowing local scientists the opportunity to use more and better digital research services.

JENNEWEIN'S JOURNEY

Computer scientist Doug Jennewein's working environment is, in more ways than one, a great frontier. Situated at the University of South Dakota (USD) on America's wind-swept plain, he is in a role that is allowing him to pursue his passion and be a pioneer in growing the use of computationally-assisted research.

"I've always been interested in science and technology, usually as separate fields, but working in research computing has provided the opportunity to scratch both itches, so to speak, which is very exciting," says Jennewein, USD's research computing manager.

When he was an undergraduate at USD, a friend's talk of building a "Beowulf cluster" intrigued him. Later as a research analyst in the computer science department in 2004, Jennewein deployed a proof-of-concept cluster, which in turn led to the department securing funding for a small cluster that was brought online in 2006 and expanded in 2008, 2009, and 2012 with grant and institutional support.

Doug Jennewein, research computing manager at the University of South Dakota. Courtesy: USD Marketing & University Relations

Just beginning his job as research computing manager in 2012, he was looking for lots of advanced computing training materials, especially what might be available online. High-performance computing (HPC) center management and growth were of interest to him, but even more, he wanted to learn about science engagement, parallel programming, and file system I/O, as well as specific programming languages such as R and Fortran, which are common in the research setting but not so much in mainstream computing.

He recalled having heard in 2011 through a newsletter of the Great Plains Network (a consortium of universities in the Midwest) about the transition of TeraGrid to XSEDE and a discussion of the new XSEDE portal. That portal turned out to be the entryway to exactly the type of training he was after.

"There's just a vast variety of topics that has been covered by XSEDE, and that's been very valuable to me as a relatively early career professional in a remote rural state," Jennewein says.

He approached XSEDE's online webinars to learn about unfamiliar topics, such as certain programming languages, but also for enlightenment concerning best practices for conducting training and outreach. He wanted to observe how XSEDE conducts a class and how it presents complicated subjects to a new audience. "I did this for my own benefit and to improve my own training and outreach here on campus," he says.

In addition to what to he's gleaned from the webinars, he's gained insight from the cluster user guides and application user guides. "Both told me about the resources at XSEDE but also showed me kind of a large-scale reference implementation of how to organize and present user guides and documentation," he explains.

Historic Old Main, the oldest building on the University of South Dakota campus, Vermillion, SD. Courtesy: USD Marketing & University Relations

Fostering the Growth of Advanced Computing

Jennewein's current objectives are to engage USD researchers and determine ways he can support them in leveraging HPC in their projects, while in the process furthering his knowledge of domain sciences.

Although USD's computing profile has largely been small task-parallel jobs, with limited use of MPI and distributed memory, Jennewein aspires to leverage his position to foster the growth of advanced computing at the university.

One of his strategies is what he refers to as "research computing house calls," which entails visiting with researchers and their teams in person to determine their needs and to demonstrate interest in and support of their respective scientific domains. The "research computing house calls" can be initiated through his response to a help desk request or by attending a departmental seminar or workshop.

Currently, the largest community involved in computationally-assisted research at USD is high-energy physics, but user groups from the materials chemistry and bioinformatics disciplines also need advanced computing support. The bioinformatics community constitutes a prime set of current and prospective users because USD is home to the Sanford School of Medicine.

"As new faculty come to campus, they expect computational resources, especially bioinformatics folks involved in high-throughput sequences or next-generation sequencing," Jennewein says. "Particularly for those use cases, we are looking at XSEDE as a scale-out option."

He adds that XSEDE makes itself an attractive choice to the USD bioinformatics researchers and others by affording them an on-campus liaison (Jennewein) for assistance in writing a compute allocation proposal to gain access to national-level advanced digital resources.

Recently, Jennewein discovered a way to further solidify his relationship with XSEDE by becoming a Campus Champion.

Rolling out the Champions program at USD is one of Jennewein's main goals for 2015. "I definitely look forward to that," he says. "Also, further interactions with the other Champions." And he will be equipped with more knowledge to share during "research computing house calls" on the Great Plains.

Researchers were assisted by XSEDE ECSS, which provided high quality support, code optimization and visualizations of the flow structures inside their detailed wind-farm simulations.

UNLOCKING THE WIND'S POWER

Researchers at Johns Hopkins University (JHU) have developed high-resolution computer simulations that take into account how the air flows within and above large wind farms in unprecedented detail—and challenge current industry thinking at the same time.

The study, published recently in the *Journal of Renewable and Sustainable Energy*, noted that these simulations counter beliefs that the highest power output comes when wind-farm turbines are arranged in a checkerboard pattern. Now, after using the Trestles supercomputer at the San Diego Supercomputer Center to conduct detailed simulations, researchers have found that the highest power results when the lateral offset of turbines is such that they are just outside each other's wakes.

Moreover, these detailed computer simulations using XSEDE resources take into account the total interaction between the wind-farm and the atmosphere. JHU researchers were also assisted by XSEDE's Extended Collaborative Support Services (ECSS), which resulted in highly detailed visualizations of the flow structures within the wind-farm simulations. In addition, XSEDE provided support for the researchers to further optimize their codes with the aim of further increasing the accuracy of future simulations.

"The effect of spacing and relative positioning of the turbines on the wind farm is crucial for good wind-farm design," says Richard Stevens, who conducted the research with Charles Meneveau and Dennice Gayme at JHU.

"Wind-farm designers typically rely on simpler computer models that predict the wake effects caused by the turbines. While such models work well for smaller wind farms, they become less accurate for larger wind farms, where the wakes interact with one another as well as with the atmospheric wind."

After comparing the results from very detailed wind-farm simulations with existing industry models, the research team developed a new model, called the "coupled wake boundary layer model," which can more effectively predict wind-farm performance. These new results are described in a recent paper also published in the *Journal of Renewable and Sustainable Energy*.

"The benefit of high-fidelity computer simulations such as those performed using Trestles is that the flow in the wind farm can be studied in detail to get insight about the main physical mechanisms that are important for the performance of very large wind farms," says JHU's Meneveau.

As wind energy becomes more important around the globe as a source for clean, renewable power, researchers are turning to advanced supercomputers to create detailed simulations to study the interactions between wind flow and other atmospheric conditions.

HARNESSING THE CYBER WIND

XSEDE helped bring the Cyber Wind Facility to fruition by giving researchers "one stop shopping" for two supercomputers: PSC's Blacklight for large-memory analysis, and TACC's Stampede for massively parallel, high-speed computation.

Wind energy commands great promise for power generation. But to turn a profit, wind farms need to reach reliability levels better than 90 percent over a 20-year operational life. When you're talking about getting moving parts that approach and exceed the size of a 747 to withstand the buffeting of turbulent winds that can vary by 40 or 50 percent from the average wind speed, that's a tall order.

"Two issues we focus on are power capture and reliability," says James Brasseur of Pennsylvania State University. "Since it is very expensive to replace components sitting 80 to 100 meters above the ground, a failure rate of just a few percent can eliminate profitability."

Brasseur, graduate student Ganesh Vijayakumar and their colleagues at Penn State are building a system to analyze this problem: The Cyber Wind Facility, a very-high-resolution computer model of a wind turbine in turbulent atmospheric winds.

Multiple needs, multiple resources

XSEDE played a major role in moving the project forward by providing three supercomputers with different strengths for different phases of the project.

To put together a detailed computer model of all the forces that drive a wind turbine, the researchers first had to pull them apart. They used the Pittsburgh Supercomputing Center's Blacklight and its large shared memory for the critical task of dividing 160 million computational cells over the thousands of processors required to run the massive simulation. Blacklight's large memory also allowed them to analyze the huge datasets produced in their larger-scale simulations.

In an early phase of the research, the researchers also used Kraken at the National Institute for Computational Sciences at the University of Tennessee. In the next phase, they moved their simulations to the Texas Advanced Computing Center's Stampede, whose half a million processing cores produce a blistering peak performance of nearly 10 petaflops (nearly 10 quadrillion FLOPS). Stampede's massive computational speed allowed the group to carry out the final simulations quickly and in great detail.

"Such a high-fidelity simulation involves both computing and massive data-intensive elements," Vijayakumar says. "Blacklight was invaluable for the data-intensive parts of the simulation and the analysis of the output, Stampede for the massive computation needed for the final simulation." The one-two punch of data-intensive and computationally intensive systems provided a unique ecosystem of resources that XSEDE's unified allocation and support systems made as easy to access as a single center, the researchers note.

Detailed simulations, which require the use of advanced supercomputers, allow researchers to study the flow in wind-farms in great detail. This image shows a visualization of the flow in a very large wind-farm obtained from a high-resolution simulation. The blue regions indicate the low velocity wind-speed regions (wakes) formed behind the turbines. Visualization by David Bock (National Center for Supercomputing Applications) and XSEDE's Extended Collaborative Support Services.

This international research project involved the use of many XSEDE resources and expertise to re-establish lines of genomic relationships between over 10,000 species of birds by running and optimizing a new code called ExaML.

RESEARCHERS RETHINK HOW OUR FEATHERED FRIENDS EVOLVED

A widely published global genome study using XSEDE resources and expertise on how avian lineages diverged after the extinction of dinosaurs may have ruffled a few feathers—who knew that flamingoes are closely related to pigeons? Or that while falcons may seem similar to eagles, they're actually more closely related to parrots and songbirds?

A four-year project, called the Avian Phylogenomics Consortium, has published eight papers in the journal *Science* plus an additional 20 papers in other journals. The project meant redrawing the family "tree" for nearly all of the 10,000 species of birds alive today. Researchers by compared the entire DNA codes, or genomes, of 48 species as varied as parrot, penguin, downy woodpecker, and Anna's hummingbird.

On a broader scale, understanding evolution is fundamental to understanding biology, and reconstructing the evolutionary history of birds presented an enormous challenge that was only doable through the use of advanced computational resources. Moreover, resolving the timing and phylogenetic relationships of birds is important not only for comparative genomics but also to inform us about human traits and diseases. For example, the study included vocal-learning species, such as songbirds, parrots and hummingbirds, which serve as models for spoken language in humans and have been proving useful for insights into speech disorders.

"Characterization of genomic biodiversity through comprehensive species sampling has the potential to change our understanding of evolution," wrote Guojie Zhang of the Beijing Genomics Institute and the University of Copenhagen; Erich D. Jarvis, associate professor of neurobiology at the Howard Hughes Medical Institute and Duke University; and Tom Gilbert also of the University of Copenhagen, and one of the study's coordinators, in an introduction to a special issue of the journal *Science*.

Multiple XSEDE Resources

A massive undertaking in and of itself, the high-performance computing aspect of the project employed resources throughout XSEDE to analyze the complete genomes. XSEDE HPC systems used included Ranger, Lonestar, and Stampede at the Texas Advanced Computing Center (TACC); Nautilus at the National Institute of Computational Sciences (NICS); and the data-intensive Gordon system at the San Diego Supercomputer Center (SDSC).

In all, the project involved more than 200 researchers at 80 institutions in 20 countries, with related studies involving scientists at more than 140 institutions worldwide.

The genome-scale phylogenetic analysis of the 48 bird species presented computational challenges not previously encountered by researchers in smaller-scale phylogenomic studies that analyzed only a few dozen genes. The inclusion of hundreds of times more genetic data per species allowed the researchers to realize the existence of new inter-avian relationships.

With approximately 14,000 genomic regions per species, the size of the datasets and the complexity of analyzing them required not only advanced resources but new computing methods, one of which was led by computer scientists Tandy Warnow, an adjunct professor at The University of Texas at Austin and professor at the University of Illinois at Urbana-Champaign; and Siavash Mirarab, a graduate student at The University of Texas at Austin.

"Part of the challenge is the fact that many species of birds seem to have developed in a relatively short period of time after the events that lead to the extinction of dinosaurs," says Mirarab, co-lead author on one major Science-published paper. "In this research, we used whole genome sequencing and a wide range of computational methods, including some new ones developed for this study to reconstruct the evolutionary history of birds from an unprecedented dataset. Analyzing a dataset of this size required tremendous computational powers. We needed to develop new methods, and testing these new methods required lots of computation."

By testing one new technique, called statistical binning, on simulated data sets, the TACC team demonstrated that this approach was more accurate than previous techniques. The focus of a second paper by Mirarab and colleagues, validating and using this technique, consumed more than 1 million hours of compute time on TACC resources as well as on the Condor cluster at The University of Texas at Austin. "Running all these analyses was doable only because of the exceptional computational resources we had access to," says Mirarab.

Many of the computations using SDSC's Gordon cluster were done by Andre J. Aberer, a bioinformatics researcher at the Heidelberg Institute for Theoretical Studies (HITS), with the assistance of SDSC Distinguished Scientist Wayne Pfeiffer. They ran a new code called ExaML (Exascale Maximum Likelihood) to infer phylogenetic trees.

Developed by Alexandros (Alexis) Stamatakis, head of the Scientific Computing Group at HITS, ExaML couples the popular RAxML search algorithm for inference of phylogenetic trees using maximum likelihood with an innovative MPI parallelization approach. This yielded improved parallel efficiency, especially on partitioned multi-gene or whole-genome datasets.

"I had previously collaborated with Alexis on improving the performance of RAxML," says SDSC's Pfeiffer. "He described the goals of the Avian Genome Consortium, and we agreed that Gordon, with its just-released fast processors, could provide much of the computer time needed for this ambitious project. In the end, more than 400,000 core hours of computer time were consumed on Gordon."

"After doing initial analyses on our institutional cluster, we rapidly realized that comprehensive analysis of the more challenging data sets being considered would require supercomputer resources," says Aberer. "Access to Gordon was thus invaluable for achieving results in a timely manner."

The Nautilus supercomputer at NICS was used to generate phylogenetic trees at the chromosome level for all the bird genomes in the analyzed dataset. "We used Nautilus's large shared memory to run ExaML and RAxML tools for phylogenetic inference for all the bird genomes," says Bhanu Rekepelli, a researcher at NICS at the time these analyses were done. "These runs required 600 gigabytes of memory and 2.5 hours of computation for one iteration for one chromosome—overall, it took a month of simulations to complete the analysis for all the chromosomes."

The XSEDE ECSS program provided a University of Tennessee-Knoxville researcher with expertise and computing resources for DNA sequence data analyses in an effort to understand ecological changes from an oil spill in the Gulf of Mexico.

BOOSTING ENVIRONMENTAL RESEARCH

On April 20, 2010, the Deepwater Horizon oil rig in the Gulf of Mexico exploded and sank, claiming the lives of 11 people and massively assaulting the ocean and connected ecosystem. After the blast, oil gushed from the bottom of the sea at the rig site for 87 days until it was finally capped.

Even before the oil came ashore, researcher Annette Engel and her students from the University of Tennessee, Knoxville, were in the Gulf Coast marshes taking samples, and during all the years since, have been returning to the same areas for that purpose. Their efforts have yielded the most comprehensive dataset in existence for microbial community change in the marshes, both oiled and not.

Annette Engel (center), associate professor of earth and planetary sciences at the University of Tennessee, Knoxville, samples Louisiana marsh sediment with other researchers from the Coastal Waters Consortium. Image credit: Audrey T. Paterson, UT Knoxville.

What's happening to the microbes is important because they are baseline organisms that serve as an interface between the physical and chemical conditions of a habitat and animals and other life that compose an ecosystem. Microbial DNA sequences performed through Engel's project have revealed fundamental changes in the types of bacterial communities, as well as the overall ecosystem of the Gulf Coast area.

"By determining the microbial response and investigating how long it takes for communities to return to pre-spill membership, if they ever will, we may be able to predict how the overall marsh food—composed of plants, insects, birds, fish, and other animals that require Gulf Coast marsh habitats—may have responded and will continue to respond in the years to come," Engel says.

Overcoming the Limitations

Although advances in DNA sequencing technology make it possible to obtain hundreds of thousands to millions of sequence reads from a single sample, Engel found that analyses of the large datasets to determine fluctuations within microbial taxonomic groups across time and space were limited by the capabilities of her lab computers. She discovered, however, that a solution was available to overcome the impasse.

"I was really happy that through a series of circumstances we were able to connect with NICS [the National Institute for Computational Sciences] and Bhanu Rekepalli and learn about all the different opportunities that they offered to non-traditional high-performance computing [HPC] users," she says.

Rekepalli introduced Engel to XSEDE's Extended Collaborative Support Service (ECSS) program. ECSS enables researchers to be paired with expert staff members for an extended period while using XSEDE's computers. ECSS staff provide expertise in many areas of advanced cyberinfrastructure and can work with a research team to accelerate and enhance their efforts through that knowledge.

As an XSEDE service provider, NICS was able to make HPC consultation services from its computational scientists Rekepalli and Junqi Yin and an allocation on the NICS-managed Nautilus—a robust, multicore shared-memory system—available to Engel's project.

Rekepalli met with Engel and her team to develop a work plan and provide her students with basic HPC training, which included remote login, module operation, and job submission. They also discussed the sequence data pipeline analysis that would be used in the project.

Engel's Ph.D. student, Chanda Drennen, has handled most of the HPC tasks on the project, working directly with Yin, who has helped her use the advanced computing in analyzing the DNA sequence data.

Drennen's role in the project corresponds with the mission of the Coastal Waters Consortium (CWC), which through the Gulf of Mexico Research Initiative is providing the funding for Engel and her team. CWC's mission is to have students directly involved, or driving, the research. For Drennen, the ability to perform the data analyses using HPC is essential to the completion of her dissertation.

NICS instrumented and optimized several commands in the open-source bioinformatics application called mothur for an operational taxonomic unit analysis pipeline with OpenMP on shared memory architecture, and the result was an acceleration of processing time by an order of magnitude as compared with what was possible on the desktop.

Thanks to XSEDE, Engel and her students are now able to perform operational taxonomical analyses on hundreds of thousands of sequencing data, with clearer and more comprehensive results. And going forward, the collaboration will continue to accelerate and enhance the quality of discoveries they can add to the body of knowledge about the effects of the oil-spill disaster on coastal environments.

The Coastal Waters Consortium has been studying marsh ecosystem processes and health since the Deepwater Horizon oil spill in 2010. Image credit: Annette Engel, the University of Tennessee, Knoxville.

Campuses across the country can use XSEDE resources through the XSEDE Campus Bridging initiative and the XSEDE Compatible Basic Cluster, which helps expand the capability and purpose of each campus to assist its local researchers.

CAMPUS BRIDGING GETS THE RESEARCH DONE AT MONTANA STATE UNIVERSITY

For Pol Llovet, it all comes down to good strategy.

As associate director of the Montana State University Research Computing Group in Bozeman, Llovet manages MSU's cyberinfrastructure. Whether it's talking with scientists to figure out how technology can best fit their needs or working deep in the guts of supercomputers to tweak code, Llovet is dedicated to making sure the research gets done in the fastest, most efficient way possible. XSEDE and its Campus Bridging initiative are vital to his mission.

The Campus Bridging effort aims to facilitate use of local and national cyberinfrastructure in ways that are easy and useful to researchers and educators—a bridge from the campus to the regional, national and international levels.

"Our strategy requires XSEDE," says Llovet. "We know we're never going to be a Top 500 supercomputing site, but we still aim to provide everything that our researchers need. We start locally, we do a local proof-of-concept to see what resources they need. If the researcher's work can be done on our cluster, we just do it. If they need more, because we've made the cluster XSEDE compatible, we can take the work we've done and move it to our XSEDE allocation, and move very quickly. As you can see, Campus Bridging is critical—otherwise this strategy doesn't work."

In July 2014, XSEDE partners Indiana University and Cornell University announced the XSEDE Compatible Basic Cluster (XCBC) software suite, designed to help researchers ranging from big data scientists to people running small campus clusters—people just like Llovet.

The XCBC software suite lets a campus create a cluster with open-source software tools that match the software of the most commonly used systems within XSEDE. That way, a command that a researcher might use on an XSEDE cluster works the same way on a local cluster. XCBC also includes software to aid data movement and integration with XSEDE, particularly software for use on Globus Online, currently XSEDE's most widely used campus-bridging data movement tool.

Although Montana State University is a mid-size university with around 15,000 students and 550 full-time faculty members, its research output is impressive. In fact, MSU is among the nation's top tier of research universities, recognized by the Carnegie Foundation for the Advancement of Teaching as one of 108 research universities with "very high research activity."

Ben Poulter is proud to be among that group. As a professor in the MSU Department of Ecology, Poulter investigates the role of climate and humans on terrestrial ecosystems using vegetation models, remote sensing, forest inventory and physiological measurements. Terrestrial ecosystems include forests, boreal forests, tropical forests, temperate forests, grasslands and semi-arid systems. More specifically, his lab studies the terrestrial carbon cycle, which plays a key role in absorbing carbon dioxide and thereby mitigating fossil fuel emission.

"A lot of the work we do is motivated to improve our contribution to an annual assessment of the global carbon budget published by the Global Carbon Project each year," says Poulter. "That's quite relevant for climate change science where we're trying to understand what is a safe level of CO_2 that could be emitted. The challenge in trying to mitigate greenhouse gases is that we're not sure how the terrestrial biosphere will respond to climate change. As the planet gets warmer, as ecosystems get more stressed due to high temps and drought, they could stop acting as carbon sinks and start releasing CO_2 back into the atmosphere."

Poulter's lab uses simulations called dynamic global vegetation models to understand this carbon cycle. He must use high-performance computing to analyze his data—and that's where Llovet and the Research Computing Group come in. Llovet introduced Poulter to XSEDE and its resources.

"We would not be able to do much of our work without access to XSEDE," says Poulter. "We run simulations, for example, where we divide the world into 70,000 grid cells, and for each grid cell we're simulating physiological processes and demographic processes over 1,000 to 2,000 years at a half hour or daily time stamp. By using XSEDE tools and machines, we can compress what might be several days or weeks of simulation time on a single computer into a job that only takes a few hours."

Poulter is so pleased with XSEDE and its ease of use that he's passing the word on to the Big Data scientists of tomorrow—his students. During one recent final project for a graduate course in ecological modeling, Poulter encouraged his students to use XSEDE to run the analyses. As a result, his students now see HPC as something they can do themselves, without extensive IT training. "Now that they know the fundamentals of high-performance computing, it opens up a lot more opportunities for them to be able to conduct their research. Hopefully, I've set them up for successful thesis research and beyond," he says.

Students, faculty, staff, and everyone else at Montana State University is reaping the benefits of Llovet's solid XSEDE strategy.

"The cool thing about XSEDE and its Campus Bridging Initiative is that they allow me to help researchers, right now," says Llovet. "I can discuss their technology needs, help them get going, improve their analyses, broaden the scope, and do it today with top-class national resources."

Montana State University campus.
Credit: Montana State University

XSEDE provided supercomputer access to a Georgia Tech research team that allowed greater understanding of the fluid mechanics physics involved in researching data coordinates of millions of molecules of in-air pollutants.

TRACKING POLLUTANTS BACKWARDS

A pernicious mix of agents that harms human health and the ecosystem, air pollution is chock-full of substances that interact in complex ways. Understanding the complexity so that public strategies and policies to combat the problem can be devised requires some seriously scientific detective work, and XSEDE is equipped to help.

XSEDE provides advanced computing resources that a research group has used to track the flow of molecules backwards in time through the atmosphere to discover the effect that turbulence, or chaotic flow, has in assembling air pollution's noxious brew. After all, to fight air pollution, we need to understand how it forms.

"This research has shown that dispersion happens faster going backwards than forwards, which means that the disorderly motions in turbulent flow are actually more efficient in bringing various entities together, from far-away spots, than spreading them apart," says Georgia Tech Professor P. K. Yeung. "This observation is important in explaining how chance accumulations of pollutant material may occur locally with detrimental consequences."

Yeung and his students, along with his collaborator, Brian Sawford of Monash University in Australia, have been studying the motion of infinitesimal fluid parcels as well as diffusing molecules, thanks to compute allocations on XSEDE's Stampede at the Texas Advanced Computing Center (TACC) and Kraken at the National Institute for Computational Sciences (NICS).

With help from TACC's Lars Koesterke through XSEDE's Extended Collaborative Support Service (ECSS) program, Yeung and his students are learning how to use Intel® Many Integrated Core co-processors effectively in the analyses of an indeterminately large number of pairs of fluid elements and molecules. XSEDE's support has allowed the group to produce results under the condition of a high Reynolds number better than other groups in the field. The Reynolds number is essentially a means of quantifying the degree of chaos in the turbulence.

The group's latest results are described in a paper titled "Characteristics of backward and forward two-particle relative dispersion in turbulence at different Reynolds numbers," newly submitted for journal publication. Other elements of their work also were presented at the 19th Australasian Fluid Mechanics Conference (December 2014) in Melbourne, Australia, and at the XSEDE'14 Conference (July 2014) in Atlanta.

As Yeung explains, the wind that carries pollutants is more than likely to be turbulent, which means that creating accurate simulations requires a complete record of the position of coordinates of many millions of molecules spanning a long time period. Such simulations and the tremendous details they contribute to the effort to combat air pollution are possible only with the support of XSEDE-caliber advanced computing resources.

Georgia Tech researchers used XSEDE supercomputers Kraken and Stampede for code development and production, speeding up their normal processes, in order to understand and predict signs of black holes swallowing stars.

COSMIC SLURP

Somewhere out in the cosmos, a flash of light explodes from a galaxy's center. A star orbiting too close to the galaxy's central supermassive black hole has been torn apart by the force of gravity, heating up its gas and sending out a beacon to the far reaches of the Universe.

What would such a beacon look like? And what could it tell us about black holes—some of the most powerful and mysterious forces in the Universe?

"Black holes by themselves do not emit light," says Tamara Bogdanovic, assistant professor of physics at the Georgia Institute of Technology. "Our best chance to discover them in distant galaxies is if they interact with the stars and gas that are around them."

In recent decades, with improved telescopes and observational techniques designed to repeatedly survey the sky, scientists noticed that some galaxies that previously looked inactive would suddenly light up at their very center.

"This flare of light was found to have a characteristic behavior as a function of time," Bogdanovic says. "It starts very bright and its luminosity then decreases in time in a particular way.

Astronomers have identified those as galaxies where a central black hole just disrupted and "ate" a star.

"It's like a black hole putting up a sign that says 'Here I am,'" she says.

Bogdanovic tries to predict the dynamics of major cosmic events, such as a black hole devouring a star—also known as a "tidal disruption"—through a mix of theoretical and computer-based approaches.

Using the Stampede supercomputer at the Texas Advanced Computing Center and Kraken at the National Institute for Computational Sciences (both part of XSEDE), she and her collaborators modeled the dynamics of these powerful forces and charted their behavior using numerical simulations.

Tidal disruptions are relatively rare cosmic occurrences, occurring only once every 10,000 years in a Milky Way-like galaxy. The flare of light they produce, on the other hand, can fade away in only a few years. Because it is such a challenge to pinpoint tidal disruptions in the sky, astronomical surveys that monitor vast numbers of galaxies simultaneously are crucial.

So far, only a few dozen of these characteristic flare signatures have been observed and deemed "candidates" for tidal disruptions, but with data from PanSTARRS, Galex, the Palomar Transient Factory and other upcoming astronomical surveys becoming available to scientists, Bogdanovic believes this situation will change dramatically.

"As opposed to a few dozen that have been found over the past 10 years, now imagine hundreds per year—that's a huge difference!" she says. "It means that we will be able to build a varied sample of stars of different types being disrupted by supermassive black holes."

In a paper published in the Astrophysical Journal in 2014, Bogdanovic, working with Roseanne Cheng (Center for Relativistic Astrophysics at Georgia Tech) and Pau Amaro-Seoane (Albert Einstein Institute in Potsdam, Germany), simulated the tidal disruption of a red giant star by a supermassive black hole.

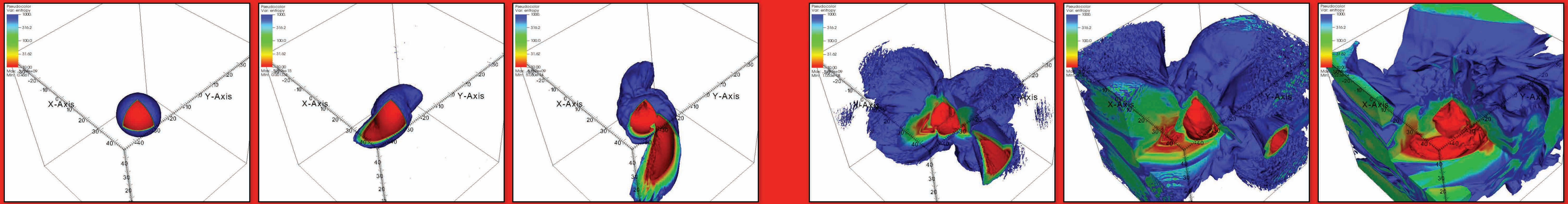
The paper came on the heels of the discovery of an observed tidal disruption event in which a black hole disrupted a helium-rich stellar core, thought to be a remnant of a red giant star, 2.7 billion light years from Earth.

The team ran simulations on Georgia Tech's Keeneland supercomputer, as well as Kraken and Stampede, to reconstruct the chain of events by which a stellar core might evolve under the gravitational tides of a massive black hole.

"Calculating the messy interplay between hydrodynamics and gravity is feasible on a human timescale only with a supercomputer," Roseanne Cheng says.

The research shows how simulations complement and constrain theory and observation and help decode the signatures of observed tidal disruption events.

"The most recent data on tidal disruption events is already outpacing theoretical understanding and calling for the development of a new generation of models," she says. "The new, better quality data indicates that there is a great diversity among the tidal disruption candidates. We are yet to understand what causes these differences in observational appearance, and computer simulations are guaranteed to be an important part of this journey."



This series of images shows the evolution of a white dwarf star as it is disrupted by a massive black hole. The star is flexed by the tidal field of the black hole and develops strong shocks on its surface (blue and green) but its remnant core (red) survives disruption. Image credit: Tamara Bogdanovic, Georgia Institute of Technology.

A Caltech research team performed extremely detailed, fully three-dimensional simulations of supernova explosions thanks to ECSS support, which helped optimize code on XSEDE-allocated Stampede and NCSA's Blue Waters.

SPACE LAB

If you were to go back far enough into the Earth's cosmic ancestry, you might be surprised to find it all started with a supernova explosion. These explosive cosmic events are like laboratories in space, generating elements that enable the creation of life later on; in fact, most of what makes up the Earth, including us humans, evolved from these fundamental elements. This is why simulating the process of a star going supernova is so important—it could potentially be the key to unlocking some of the bigger mysteries of how we came to be in the universe.

Philipp Mösta, postdoctoral scholar at Caltech, Christian D. Ott, professor of astrophysics at Caltech, and fellow researchers working with Peter Diener, research professor at the Center for Computation and Technology of Louisiana State University, are studying extreme core-collapse supernovae. These events make up only one percent of all supernovae that are observed but are the most extreme in terms of the energy emitted into the universe.

In the past, simulation work within their field was mostly done in two dimensions (2D), and codes did not always include all the physics, neglecting general relativistic effects, for example. For the first time ever, Mösta and his fellow researchers are running fully general relativistic three-dimensional (3D) simulations, and new details are emerging. Their simulations could change the landscape of computational research in their field moving forward.

"It will probably indicate to other groups who, so far, have focused on performing simulations with symmetries imposed, that they will have to move to full 3D simulations as well, which will ultimately strengthen our community," says Mösta.

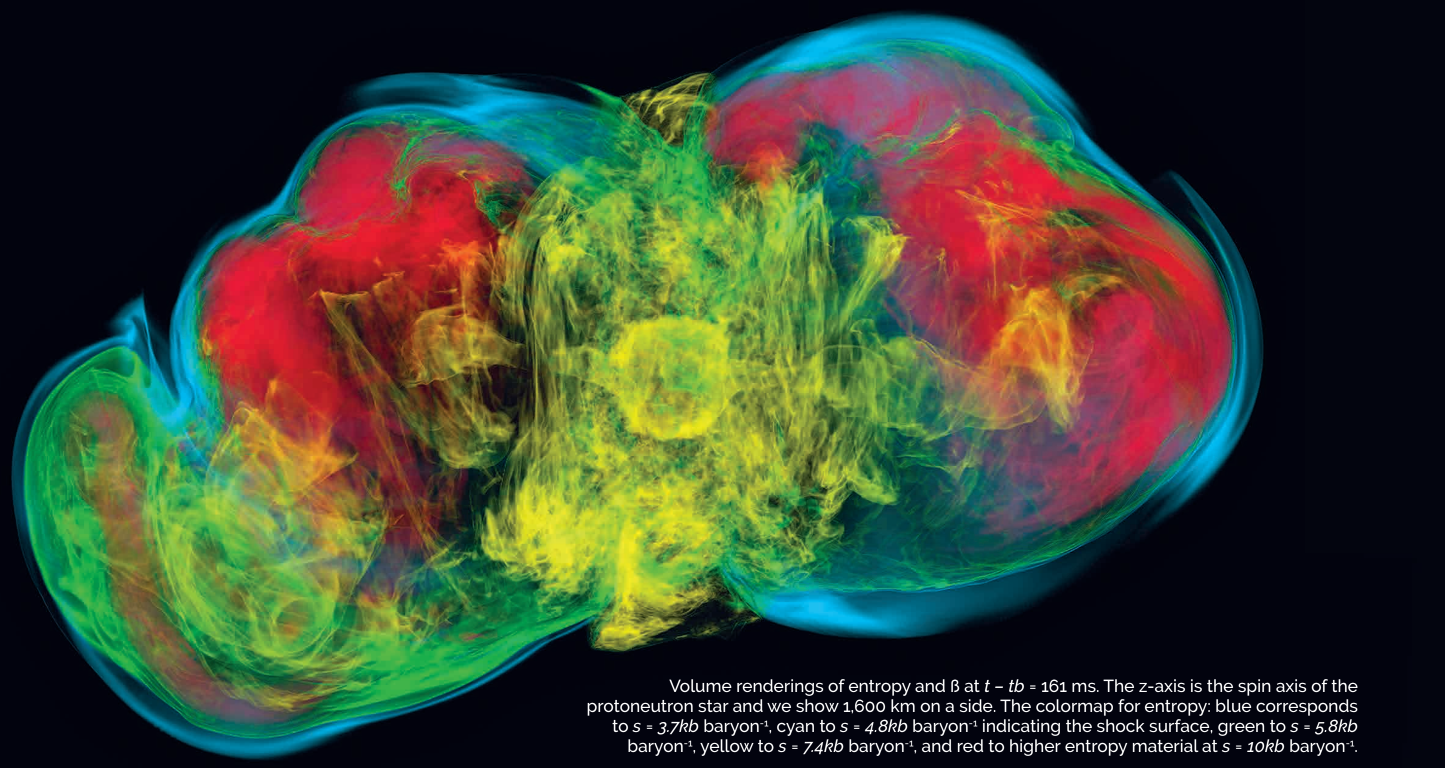
This journey through dimensions began with an XSEDE allocation on the Stampede supercomputer at the Texas Advanced Computing Center (TACC) at The University of Texas at Austin back in 2013.

Working with Lars Koesterke, research associate in the High Performance Computing Group at TACC, along with XSEDE's Extended Collaborative Support Services (ECSS), the team got assistance in optimizing the performance of their code. "We had to pair fast individual compute core performance with excellent communication throughput through a fast network," explains Ott. "This required a thorough optimization of our code toward the strength of modern architectures such as Stampede and Blue Waters." Thanks to this support, the team was able to create a fully functional version with offloading onto the Xeon PHI co-processors.

For the team, this meant larger simulations were possible without fear of burning through their allocation too quickly. "We were able to perform the first fully general relativistic 3D simulations without any symmetries and the difference in comparison to 2D was drastic," Ott continues. "We now know if we want to predict what the signature of these extreme supernova explosions might look like, we need to do it in full 3D."

At first, the team had to cut out a part of the inner star. "At the end of the life of one of these massive stars that we're simulating there's an iron core in the center, about 1.5 times as massive as the sun, just consisting of iron," explains Ott. "That's what is undergoing the collapse to a neutron star, and it's this collapse that provides the energy that then ultimately powers the explosion."

For the other 99 percent of observed supernova events, the main proposed mechanism driving the explosions is energy deposition by neutrinos behind the shockwave. However, this couldn't possibly deliver the amount of energy that is observed for these hyper-energetic supernovae. The mechanism studied by Mösta and Ott relies on a combination of magnetic fields and strong rotation. Extra



Volume renderings of entropy and ß at $t - t_b$ = 161 ms. The z-axis is the spin axis of the protoneutron star and we show 1,600 km on a side. The colormap for entropy: blue corresponds to $s = 3.7 k_b$ baryon^{-1}, cyan to $s = 4.8 k_b$ baryon^{-1} indicating the shock surface, green to $s = 5.8 k_b$ baryon^{-1}, yellow to $s = 7.4 k_b$ baryon^{-1}, and red to higher entropy material at $s = 10 k_b$ baryon^{-1}.

magnetic field behind the shockwave is built up by extracting rotation from the collapsed core of the star. This magnetic field exerts extra pressure, which leads to a bipolar explosion in the form of two jets that are pinching out along the rotation axis of the star—or so it was anticipated on the basis of 2D simulations.

These initial general-relativistic magneto-hydrodynamics (GRMHD) simulations were performed on Stampede. However, when the shockwave launched in the explosion propagates farther out into the core, the computational demand of the simulations increases dramatically, requiring the team to scale up. The decision was then to move from Stampede to NCSA's Blue Waters supercomputer, allowing them to expand to 30,000 cores.

While running their simulations on Blue Waters, the team worked with NCSA visualization programmer Robert Sisneros to improve the performance of their analysis scripts to read and manage their data. This included integrating the launch of VisIt—an open-source, interactive, and scalable visualization, animation, and analysis tool. "While the analysis itself is no different when deploying a parallel job to the batch system on Blue Waters, there is less flexibility in the ways components can be launched," explains Sisneros. "Their codes needed a few updates to correctly manage launching parallel components, as well as a few other changes due to subtle differences among various versions of VisIt."

Mösta and Ott credit their success to having computational time on both of these powerful petaflop supercomputers, along with access to knowledgeable support staff. "Both Stampede and Blue Waters provided the computational power we needed to prove we could successfully perform 3D simulations, and more importantly, to prove the importance and impact of the simulations," says Mösta.

"What we've shown is that the jets that appear stable in 2D are actually unstable in 3D. They twist, rotate, and become unstable due to a phenomenon that is called the magneto-hydrodynamic kink instability. This instability of the magnetic field itself is the same that is also seen in fusion reactors that are using magnetic fields to confine the plasma," says Mösta.

Mösta and Ott are using their new simulations to predict how these extreme explosions will work in full 3D. "This will allow us to take the next step, connecting our results to observations by predicting which elements are produced and how are they distributed when we look at the remnant of the supernova," says Mösta.

The team believes these results will generate needed momentum to move forward in developing the next generation of codes dedicated to performing fully 3D simulations. Soon including all of the relevant physics while running on a large number of compute cores will become the norm for simulation work in their field.

XSEDE

Extreme Science and Engineering
Discovery Environment

xsede.org | info@xsede.org

XSEDE is supported by the
National Science Foundation.