Expanse is Coming!
Computing Without Boundaries

SDSC was awarded a five-year National Science Foundation (NSF) grant valued at $10 million to deploy Expanse, a new supercomputer designed to advance research that is increasingly dependent upon heterogeneous and distributed resources. “Having access to supercomputers such as Expanse and Comet has become a necessity for researchers across an ever-growing number of science domains at institutions throughout the United States,” said UC San Diego Chancellor Pradeep K. Khosla. “SDSC is nationally recognized for building high-performance computational systems specifically designed for scientific research. The NSF sought a national leader to help the organization deliver on its goal to make such resources available to a broader user base for the purpose of advancing scientific discovery.

Continued on page 3
Our next General Staff Meeting won’t happen until October, but I’m delighted to report that we have been landing grant after grant during the last couple of months, thanks to the hard work of many SDSCers.

In June the National Science Foundation awarded SDSC a $10 million+ grant to deploy an all-new supercomputer we’re calling Expanse – a fitting name for this new research ecosystem. Expanse will eventually succeed Comet, which will continue through March 2021.

When it enters production during the latter half of 2020, Expanse will open up new modes of research computing made possible by its powerful capabilities and ease of connecting to outside software, such as direct scheduler-integration with major cloud providers, and using high-speed networks to facilitate data movement to and from the cloud. As PI for this project, I’m confident that Expanse will help researchers push the boundaries of computational research.

SDSC was also successful in procuring several other NSF awards, including three cloud-related grants as we implement our hybrid cloud strategy. A few highlights:

- A five-year, $5 million grant to SDSC and UC San Diego’s IT Services Division, the University of Washington’s eScience Institute, and UC Berkeley’s Division of Data Science to develop CloudBank, a set of managed services designed to simplify access to public clouds among the computer science community.
- A two-year grant worth about $400,000 to deploy a new system as part of our Triton Shared Computing Cluster (TSCC), which primarily serves UC San Diego researchers and our industry partners. Triton Stratus will provide researchers with improved facilities for accessing emerging computing tools and scaling them to commercial cloud computing resources.
- An Internet 2 grant funded by the NSF that allows users of our popular CIPRES gateway access to powerful processors from commercial cloud provider AWS to accelerate discovery. Access is free during the one-year period of the award.

I’m also pleased to report that as of July 1, SDSC reports to the Executive Vice Chancellor’s Office, with Senior Associate Vice Chancellor Robert Continetti providing direct oversight. SAVC Continetti also provides oversight to other interdisciplinary units on campus including the Halicioglu Data Science Institute, Qualcomm Institute, and the UC San Diego Design Lab. At the same time I’d like to thank Vice Chancellor for Research Sandra Brown for her guidance throughout the time SDSC reported to her office.

The months and years ahead will be very busy but rewarding ones for all of us. And as we all know, that’s a good thing!

Michael L. Norman
SDSC Director
UC San Diego, UC Berkeley, U Washington Announce ‘CloudBank’ Award

UC San Diego, UC Berkeley, and the University of Washington have been awarded a five-year, $5 million NSF grant to develop CloudBank, a suite of managed services to simplify public cloud access for computer science research and education. CloudBank will serve both novice and advanced users, providing a comprehensive set of user-facing and business operations functions and services to the broad computer science research and education community. Project participants include SDSC and UC San Diego’s Information Technology Services Division, the University of Washington’s eScience Institute, and UC Berkeley’s Division of Data Science and Information. Initially, CloudBank will provide access to Amazon AWS, Google GCP, and Microsoft Azure. Others may become available over time. An FAQ on CloudBank is available at cloudbank.org.

“Public cloud has become an essential resource for computer science research and education, but with the rapid growth in the diversity of resource offerings, users increasingly encounter pain points to adoption that limit the potential of these resources in their work,” said SDSC Director Michael Norman, Principal Investigator for the project. “CloudBank will address these pain points by providing ‘on-ramp’ support that helps researchers overcome challenges such as managing cost, translating and upgrading research computing environments to an appropriate cloud platform, and learning cloud-based technologies that accelerate and expand research.”

Read more at https://qrgo.page.link/2DuzQ

NSF Awards $10 Million to SDSC to Deploy ‘Expanse’ Supercomputer

“The name of our new system says it all,” said SDSC Director Michael Norman, Principal Investigator (PI) for Expanse, and a computational astrophysicist. “As a standalone system, Expanse represents a substantial increase in the performance and throughput compared to our highly successful, NSF-funded Comet supercomputer. But with innovations in cloud integration and composable systems, as well as continued support for science gateways and distributed computing via the Open Science Grid, Expanse will allow researchers to push the boundaries of computing and answer questions previously not possible.”

Read more at https://qrgo.page.link/1734v
Alyssa joined SDSC in September 2018 as a technical project manager for the Center’s Research Data Services (RDS) division, after spending three years as a consultant with Booz Allen Hamilton. She has expertise in various areas of project management resources and workflows, including Agile, a development methodology based on an iterative and incremental approach. Alyssa is also familiar with the technical certification process, including networking, and security and cybersecurity requirements. Alyssa earned her bachelor’s degree from UC Irvine in Managerial/Business Economics. Innovators caught up with Alyssa as she’s about to celebrate her one-year anniversary with SDSC.

Q: I understand that you have a DOD clearance for classified research. Can you tell us about that?
Yes, it was for my previous job as a government consultant. My co-worker encouraged me to do so just to get my foot in the door when I was still new to the DOD requirements. It was a prerequisite to get started on my contract work.

Q: You are also well versed in the Agile methodology. Can you describe what that is for the uninitiated?
Agile is a framework that allows greater flexibility in one’s workday, because it accounts for when things don’t go according to plan. Agile is unlike Scrum, where one must adhere to a strict rule set. In my previous role, I worked within the Waterfall framework where phases took place in a firm linear order. While at my previous position I saw a newsletter with a line that read “Are you tired of Gantt charts?” And I said absolutely! So I took an Agile ICP (ICAgile Certified Professional) course because I was tired of hitting the same speed bumps each time. I didn’t want to do keeping doing that – I wanted to go around them!

Editor’s note: A Gantt chart, named after Henry Gantt, illustrates a project schedule. Developed about 100 years ago, it is still widely used today, listing the tasks to be performed on the vertical axis and time intervals on the horizontal axis.
Q: You also participated in this year's Campus LISA (Large Installation System Administration) conference as a panel speaker to present how Agile methodologies work across a variety of different teams.

I was invited by my colleague Bill Homan, who's a member of the Campus LISA committee, to participate because he knew that I had expertise working with Agile. The panel talked about how project management works in each of our areas, and how Agile can work in one's organization while remaining flexible to achieve one's ends or goals. Life happens, after all – people go on vacation or have babies, so the Agile methodology takes those things into account when people are working in a team environment.

Q: Can you describe some projects during your time with RDS where your expertise comes into play?

I supported the PETTT (Productivity Enhancement, Technology Transfer, and Training) project, which looks into compute options for the Air Force Research Lab. My entire toolkit from the DOD translated over to this project. I assisted by chronicling and reporting deliverables by a certain date to help alleviate the workload for other RDS members.

Another example is our Universal Scale Storage, or large-scale storage, option. For anything above 25 terabytes of data, that's where Ryan Nakashima and I step in to introduce a larger-scale option that's more price-efficient. Once you're at 25TB, at that point it's cheaper to buy 200TB for the same monthly recharge. I mostly handle the billing and provisioning, and also contract and associated paper work.

Q: What attracted you to SDSC or UC San Diego?

I applied 'cold' to the university, so I had no real connections here. When I came across the job posting for a technical project manager, to me it was appealing because I was still a private employee, and this looked like a perfect opportunity to join as a state-funded employee. SDSC was doing teraflops of data processing – that made me hungry! My last job did not always challenge me. Here I could really sink my teeth into some of the computational and cloud work. Plus, it's pretty awesome to have 'Supercomputer Center' on your business card!

Q: You did go to UCI so maybe that was in part a connection?

I did work in the Experimental Social Science Laboratory (ESSL) while at UCI under Dr. Michael McBride, who was a mentor for me for many years. I had a very positive experience with the culture of researchers, and how everybody was very open to exchanging ideas and receiving them. That was something I could flourish in because before, my work environment was mostly very rigid and a 'this is how we do things' approach. There was no room for flexibility.

Q: What attracted you to the area of technical project management to begin with?

Seeing how blatantly disorganized some of my clients were! I take satisfaction in having an established workflow and knowing exactly where something is going to after a certain amount of steps. I saw the level of dysfunction and a lot of that work was so unnecessary. I won't call it an epiphany, but I had a moment when I thought that I really had to try something different. Applying to SDSC was my 'reach' position.

Q: I must ask: Are you just as organized at home?

I try not to bring my work processes home, but I do think in a very compartmentalized way. Like my closet – it's organized in specific sections, so I can probably draw out each piece of clothing in my head!

Q: One final but always important question: What do you enjoy doing when not managing projects here at SDSC?

I enjoy board games, arcade 80s-style games, and other video games. I have a pretty nice gaming console and an area that's a nerd cave, as my friends affectionately call it. My boyfriend and I like to play online, and currently we play Final Fantasy XIV, which is a massively multi-player online role-playing game. It has a global server so one is playing with hundreds or thousands of people across the world. I also like going to Coin-Op or Coin Haus!

Find out more about SDSC’s Research Data Services Division at https://qrgo.page.link/RHRVK
Researchers at the New Jersey Institute of Technology (NJIT) used SDSC’s Comet supercomputer to create detailed simulations of graphene-water interactions to determine if graphene is a good candidate for delivering medicine to a specific part of the body. A single layer of graphene is 97% transparent, stronger than steel, almost weightless, and can conduct electricity – making it an ideal candidate for use in many other applications. “One of the critical issues is how the graphene flakes behave when they are placed inside a water droplet, said NJIT Mechanical and Industrial Engineering Professor Dibakar Datta. “Doing experiments to understand the graphene-water interaction is expensive and requires a great deal of labor, so to meet this challenge, we performed computer simulations to gain fundamental insight.”

Read more at https://qrgo.page.link/QJPkT

Comet Supports UC Riverside Study of San Andreas Fault System

SDSC’s Comet supercomputer was used in a study of the San Andreas Fault led by Christodoulos Kyriakopoulos, a research geophysicist at UC Riverside. The study provides seismologists and geologists with a new understanding of a complex set of faults in the region that has the potential to impact the lives of millions of people in the U.S. and Mexico. “The main findings of our work concern the dynamic interactions of a postulated network of faults in the Brawley seismic zone in Southern California,” said Kyriakopoulos, lead author of a study published in the *Journal of Geophysical Research: Solid Earth.*

While the study provides seismologists and geologists with a new understanding of a complex set of faults in the region that has the potential to impact the lives of millions of people in the U.S. and Mexico, Kyriakopoulos says some of the findings point to the possibility of a multi-fault earthquake in Southern California, which could have dire consequences.

Read more at https://qrgo.page.link/bvM7e
Using Machine Learning Models to Better Predict Bladder Cancer Stages

The invasive and expensive diagnosis process of bladder cancer, which is one of the most common and aggressive cancers in the United States, may be soon helped by a novel non-invasive diagnostic method thanks to advances in machine learning research at SDSC, Moores Cancer Center, and CureMatch Incorporated. Research scientists Igor Tsigelny and Valentina Kouznetsova have been working on the development of a machine-learning (ML) model that looks at a patient’s metabolites and their chemical descriptors. The model accurately classifies the stages of bladder cancer in a patient, according to the researchers. Tsigelny is the lead author on a recently published study about the research in the Metabolomics journal. More than 81,000 Americans were diagnosed with bladder cancer in 2018 and of those, more than 17,000 died from the condition, according to statistics from the American Cancer Society.

Read more at https://qrgo.page.link/CuS48

SDSC’s Sherlock Division Debuts Innovation Accelerator Platforms

SDSC’s Sherlock Division has announced the launch of Innovation Accelerator Platforms within its Sherlock Cloud infrastructure and its newest offering, Vyloc Cloud. These turn-key solutions provide quick access to on-demand, elastic platforms for secure processing, visualization, and storage of a wide range of data. Formerly known as the Health Cyberinfrastructure Division, Sherlock recently unveiled Vyloc Cloud, a managed cloud capability for open (non-protected) data. Using its public cloud platforms expertise, Sherlock built Vyloc Cloud on top of these platforms (e.g., AWS and Azure) to facilitate needed micro-services and applications that provide an infrastructure to support big data, analytics, and data science use cases. The addition of Vyloc Cloud to its portfolio lets Sherlock offer these platforms to customers that conduct research with both protected data and open data.

Read more at https://qrgo.page.link/dXkBa

Check out SDSC’s Research Videos, Social Media Feeds!

Did you know that SDSC has a YouTube channel? Just click on the YouTube icon at the bottom of the SDSC home page under the ‘Get Connected’ section or go to…

www.youtube.com/SanDiegoSupercomputerCenter

There you’ll find a wide variety of topical videos, with playlists for Events, Training, Grand Challenges, and more. Except for the conference presentations, most videos are less than four minutes long and communicate complex research challenges in a clear and compelling way, thanks to our resident videographer, Jon Chi Lou.
Industry’s Window into HPC Innovation

Multiple grant awards from the National Science Foundation this summer have continued SDSC’s thirty-plus year legacy of innovation in High Performance Computing. One of those awards is for CC* Compute: Triton Stratus, a two-year grant worth almost $400,000 to deploy enhancements to the campus’s existing Triton Shared Computing Cluster (TSCC) HPC platform.

Triton Stratus will provide researchers with improved facilities for utilizing emerging computing paradigms and tools, namely interactive and portal-based computing, and scaling them to commercial cloud computing resources. Researchers, especially data scientists, are increasingly using tools such as Jupyter notebooks and RStudio® to implement computational and data analysis functions and workflows.

Triton Stratus will establish a pool of new compute nodes dedicated to Jupyter and RStudio, providing researchers with higher performance technology and better system response times. Moreover, the project will investigate and deploy techniques for packaging tested notebooks and sending or “bursting” them to commercial cloud services for greater scale or throughput.

These tools are part of a general trend in research computing toward web-based and graphical interfaces for accessing HPC systems, especially for attracting newer generations of researchers and data scientists. Triton Stratus will allow exploration of the emerging hybrid model of on-premise cluster computing resources, coupled with commercial cloud computing services.

Yours truly is the Principal Investigator (PI) for the project, along with SDSC’s Robert Sinkovits, Subhashini Sivagnanam, and Mary Thomas as co-PIs.

Triton Stratus leverages the strong commercial momentum in cloud computing and underlying technologies. The worldwide public cloud services market is projected to grow 17.5 percent in 2019 to total $214.3 billion, up from $182.4 billion in 2018, according to research and advisory company Gartner, Inc. The fastest-growing market segment will be cloud system infrastructure services, or infrastructure as a service (IaaS), which is forecast to grow 27.5 percent in 2019 to $38.9 billion, from $30.5 billion in 2018. The second-highest growth rate of 21.8 percent will be achieved by cloud application infrastructure services, or platform as a service (PaaS).

The project will permit SDSC to investigate and understand a number of dimensions of hybrid on-premise and cloud computing, including the right balance of on-premise and cloud resources, the best modes for scaling or bursting to cloud platforms, and new models of interactive research computing. We expect these avenues of investigation, and others, to be of great interest to industry as well. To learn how you can get involved, please contact us directly or visit our website.

Ron Hawkins
SDSC Director of Industry Relations