Expanse Enters Production!

SDSC’s new Expanse supercomputer formally entered service for researchers early last month, following a review by the National Science Foundation (NSF), which awarded SDSC a grant in mid-2019 to build the innovative system. At over twice the performance of Comet, SDSC’s current petascale supercomputer, Expanse supports SDSC’s theme of ‘Computing without Boundaries’ with powerful CPUs, GPUs, and a data-centric architecture that supports a wide range of scientific workloads including experimental facilities, edge computing, and public clouds.

[Continued on page 5]
With 2020 now behind us I want to once again thank the entire SDSC staff for your continued high levels of dedication and resilience during the last 10 months. For me, the year really put in perspective the fact that we are in a sense an extended family that does exactly what all families do: help each other when needed and pause to reflect and be thankful for all the positives in our lives.

The final quarter of the year was, as in previous years, very busy, yielding many positives for SDSC. We had a very successful virtual debut at the Supercomputing 20 conference in November, celebrating the Center’s 35th anniversary with the launch of our new Expanse supercomputer and generating a lot of interest in our new CloudBank portal.

The preparation by the External Relations group was exceptional and we had many more folks watching our presentations than in previous years! Maybe a hybrid model of live/virtual is the way forward for these kinds of events.

SDSC not only picked up two HPCwire ‘Best’ awards during SC20, but also played a collaborative role in research that earned Rommie Amaro and her team the prestigious Gordon Bell Special Prize for COVID-19 research. Kudos to our own Mahidhar Tatineni, who leads our User Services group for coordinating access to Comet to optimize codes before the team scaled up to larger systems. More about that on Page 3, and elsewhere in this issue you will find a staff profile on Mahidhar, who has done outstanding work getting Expanse up and running. More awards news can be found on Page 4 of this issue so congratulations to every one of you!

Also on Page 4 you’ll read about SDSC’s first team to participate in the Student Cluster Competition held during SC20. The six-student team, called the SDSC/UCSD Superscalars, ranked fourth among 19 teams during the multi-day virtual computing challenge. Kudos to the entire team and to Mary Thomas in particular, who leads our HPC training program. I also commend Mahidhar Tatineni, Martin Kandes, Nicole Wolter, Jeff Sale, and Robert Sinkovits for serving as mentors to the student team.

Wishing all of you a bright 2021, and please remember that our family is here for you as well!

Michael L. Norman
SDSC Director
SDSC Receives Two ‘Best’ HPCwire Awards, Cited in Gordon Bell Special Prize for 2020

SDSC received two HPCwire awards during the SC20 conference, including ‘Best Use of HPC (High-Performance Computing) in the Cloud’, and ‘Best Use of HPC in Energy.’

Editors’ Choice Award: Best Use of HPC in the Cloud (Use Case): Researchers from the IceCube Neutrino Observatory and SDSC leveraged 51,000 cloud GPUs simultaneously available via commercial providers AWS, Azure, and Google to perform the largest GPU ‘cloud burst’ in history to process data from neutrino sensors buried in the ice of the South Pole.

Research partners in the first experiment for which the award was given included:
• Frank Würthwein and Igor Sfiligoi (SDSC)
• Benedikt Riedel and David Schultz (IceCube Neutrino Observatory, University of Wisconsin – Madison)
• Open Science Grid, HTCondor, Internet2, and Pacific Research Platform
• The National Science Foundation

Editors’ Choice Award: Best Use of HPC in Energy:
Researchers from the Georgia Institute of Technology and the Hanoi University of Science and Technology in Vietnam used the Comet supercomputer to identify four lead-free candidates for a more efficient, less expensive alternative to the silicon that is typically used in solar panels. Collaborators included Huan Tran (Georgia Institute of Technology) and Ngoc Tuoc (Hanoi University of Science and Technology).

The Gordon Bell Special Prize
SDSC also played a supporting role in findings that earned this year’s prestigious Gordon Bell Special Prize for supporting COVID-19 research, also announced during SC20. Mahidhar Tatineni, SDSC’s lead for the User Services group and a co-author of the research paper (see page 6), coordinated the use of Comet by a team led by UC San Diego Computational Chemist and Biophysicist Rommie Amaro. The team was recognized for their study exploring the movement of SARS-CoV-2’s spike protein to understand how it gains access to the human cell.

Comet and other resources were used to optimize codes before the team moved to larger HPC systems at Argonne National Laboratory and Oak Ridge National Laboratory to conduct full-scale simulations. The study, called ‘AI-Driven Multiscale Simulations Illuminate Mechanisms of SARS-CoV-2 Spike Dynamics’, was in Proceedings of SC20, Virtual Event, November 16-19, 2020.

Read more at https://qrgo.page.link/ktE75

Read more at https://qrgo.page.link/LXm29
SDSC Student Team Shines at Supercomputing 2020 Conference

A team fielded for the first time by SDSC competed in the Student Cluster Competition in November 2020 at the annual International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC20) ranked fourth place overall among 19 teams during the multi-day challenge, held in a virtual format.

The six-student team, called the SDSC/UCSD Superscalars, ran all applications during the competition, which included three benchmarking apps (HPL, HPCG, and IO-500), three science apps (CESM, GROMACS, and the MEMXCT reproducibility application), and a “mystery application.” In addition to placing in the top quartile overall, the Superscalars achieved the highest score on the mystery application, ranked third in HPCG with 4.056 teraflops (a processor’s capability to calculate one trillion floating-point operations per second), fifth for the reproducibility application, and sixth for the IO-500 with 16 gigabytes per second (GB/s). They also submitted full or partial results for every one of the applications in the competition.

The team members included Co-Captains Jacob (Xiaochen) Li (Computer Science and Engineering) and Max Apodaca (Electrical and Computer Engineering Department). Both departments are in UC San Diego’s Jacobs School of Engineering. Other team members were Zihao Kong (Electrical and Computer Engineering Department), Arunav Gupta (Data Science, Mathematics), Hongyu Zhou (Mathematics), and Hongyi Pan (Computer Science & Engineering).

“Our team did extremely well, considering that they were the first team sponsored by SDSC,” said Mary Thomas, a computational data scientist with SDSC and lead for the Center’s HPC training program. “Not ones to rest on their laurels, they’re already organizing to submit to the SCC21 competition!”

In addition to Thomas, SDSC team mentors included Mahidhar Tatineni, Martin Kandes, Nicole Wolter, Robert Sinkovits, and Jeff Sale, as well as mentors from SDSC industry partners including Lewis Carroll (AMD), Paul Yu (Microsoft/Azure), and Abe Stern (NVIDIA).

Read more at https://qrgo.page.link/L15YX
SDSC’s SCC20 team page: https://hpc-students.sdsc.edu/scc20

More Awards!

Several Value Awards were given to SDSC staffers during the Winter General Staff e-Meeting on December 9th. In order of their announcement, ‘Innovation Awards’ went to Cindy Wong in External Relations for her outstanding work as the technical coordinator for several large virtual conferences and events, including the SC20 conference in November; and to Mahidhar Tatineni, Lead for User Services in the Data-Enabled Scientific Computing division, for providing his HPC and machine learning expertise as a participant and co-author of a study led by UC San Diego Computational Chemist and Biophysicist Rommie Amaro that won the prestigious Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research in November during the SC20 virtual conference.

The ‘Leadership Award’ went to Julie Gallardo in the Business Office for leading the SDSC Fiscal Team through the campus Enterprise System Renewal (ESR) process and developing tools and systems for both SDSC and campus to support that transition.

The ‘Level Up Award’ went to Ryan Nakashima, research services manager with SDSC’s Research Data Services Division, and five of his direct reports – Alex Andrieu, Jenny Nguyen, Victoria Nguyen, Victor Chan, and Henry Chan. Ryan created an IT research internship program for students to build their workplace skills while creating a path for SDSC to have access to such experienced employees.

The ‘Team Award’ went to SDSC’s Recharge Billing Team led by Csilla Csori, which includes Alyssa Arce, Mike Gorney, Chris Battistuz, Steven Yeu, Ryan Nakashima, Jasjot Sumal, Mike Dwyer, and Julius Eshabarr. That cross-departmental team did an outstanding job of managing a comprehensive restructuring of SDSC’s billing process to comply with campus use of the Oracle Financial Cloud.

Congratulations to all recipients!
Expanse Enters Production!
[Continued from page 1]

“The name of our new system says it all,” said SDSC Director Michael Norman, principal investigator for Expanse and a computational astrophysicist. “With innovations in cloud integration and other features such as composable systems, as well as continued support for science gateways and distributed computing via the Open Science Grid (OSG), Expanse will allow researchers to push the boundaries of computing and substantially reduce their times to discovery.”

A key innovation of Expanse is its ability to support composable systems, which can be described as the integration of computing elements such as a combination of CPU, GPU, and other resources into scientific workflows that may include data acquisition and processing, machine learning, and traditional simulation. Expanse also supports integration with the public cloud providers, leveraging high-speed networks to ease data movement to/from the cloud, and a familiar scheduler-based approach.

Read more at https://qrgo.page.link/JuY1G

National Science Foundation Forms Throughput Computing Partnership

The National Science Foundation (NSF) awarded $22.5 million to a partnership between the Center for High Throughput Computing (CHTC) at University of Wisconsin – Madison and the Open Science Grid (OSG) to advance open science via the practice of distributed High Throughput Computing (dHTC). The project seeks to harness the computing capacity of thousands of computers assembled in a network of campus clusters to cut time to science result that might take years, to be done in days, especially for applications that are parallel by design.

The Partnership to Advance Throughput Computing (PATH) award is for five years and will fund more than 40 individuals across participating institutions. SDSC is a partner via Frank Würthwein, the Center’s lead for High-Throughput Computing and a physics professor at UC San Diego. Würthwein also is executive director of OSG, a national cyberinfrastructure funded by the NSF to advance the sharing of resources, software, and knowledge.

Read more at https://qrgo.page.link/kN8Jo

Comet Calculations Boost Understanding of Our Immune System

Two SDSC researchers – Madhusudan Gujral, a senior bioinformatician, and Robert Sinkovits, SDSC’s director of scientific computing applications – recently contributed to a study led by Vanderbilt Vaccine Center of Vanderbilt University Medical Center on T cell receptors, which play a vital role in alerting the adaptive immune system to mount an attack on invading foreign pathogens including the Coronavirus SARS-CoV-2.

SDSC’s Comet supercomputer was used to perform complex calculations on the receptor sequence data from sorted human T cells to allow scientists to better understand the size and diversity receptor repertoire in healthy individuals. The team’s findings were published last month in Cell Reports as a follow-up study to earlier findings about B cells published in the journal Nature.

Read more at https://qrgo.page.link/VUfK9

Comet Aids in Predicting Oil Dispersal During Spills

According to the National Oceanic and Atmospheric Administration (NOAA), thousands of oil spills occur each year in the U.S. While most incidents involve less than one barrel, the spills have wreaked economic and environmental devastation for decades. To better understand the fate of oil droplets for effective countermeasures, researchers recently created simulations using supercomputers, including SDSC’s Comet system, enabling Cosan Daskiran, a postdoctoral researcher and senior engineer at the New Jersey Institute of Technology, to model studies on how oil dilutes under specific conditions.

“We used supercomputers to create high-fidelity, large eddy simulations of underwater oil blowout in water crossflow conditions,” said Daskiran. “The main goal was to understand the fluid dynamics and estimate the trajectory of different-sized oil droplets, which is important for the countermeasures following oil spill incidents.”

Read more at https://qrgo.page.link/o2KQB
Meet Mahidhar Tatineni
Director of SDSC’s User Services Group

By Jan Zverina

As the Lead for SDSC’s User Services group, Mahidhar (Mahi) Tatineni has several key roles. As with SDSC’s ‘Comet’ supercomputer, Mahi contributed significantly to the early benchmarking process for ‘Expanse’, which went into operation in late 2020 following early-user testing to validate its innovative architecture. Mahi joined SDSC in 2005 after working at UCLA as a staff research associate and lecturer. He received his M.S. and Ph.D. in Aerospace Engineering from UCLA and has completed numerous optimization and parallelization projects on several SDSC supercomputers.

Q: Running early performance and benchmarking tests on new systems such as Expanse is just one part of your role as head of the User Services Group. What’s the depth and breadth of the group?

High-performance computing and data analytics underpin research in a broad range of academic domains. The breadth of the group reflects this diversity of applications on SDSC’s supercomputers. The User Services staffers have advanced degrees with HPC experience in chemistry, mechanical and aerospace engineering, and computer science. The group provides a broad range of services including user support for diverse workloads on SDSC’s HPC systems; build, performance testing, and deployment of a large HPC and data analytics applications stack; in-depth training/workshops on HPC resources; and contributing to several successful NSF and NIH research proposals.

Q: You’re a co-author on a study led by UCSD Computational Chemist and Biophysicist Rommie Amaro that won the coveted Gordon Bell Special Prize for COVID-19 research. Can you describe your role in that research and your response when the prize was announced?

The COVID-19 study led by Rommie Amaro is a shining example of multi-institutional collaborative research using HPC resources at large scales. The components of the research...
leveraged different resources at several sites (SDSC, TACC, ANL, and ORNL), depending on specific application requirements. My role in the research was to help optimally set up and manage these large-scale runs on Comet and provide operational support. The simulations completed by Rommie’s team would normally be beyond the computational capacities of even the largest computers, but with a novel artificial intelligence- (AI) based approach they were able to compute them. It was a delight to see the problem being solved in COVID-19 research and being recognized as such with the Gordon Bell Special Prize.

Q: Can you tell us about your career path, from being interested in aerospace engineering at UCLA to joining a team that designs and builds national multi-million-dollar HPC systems such as Comet and Expanse?

I’ve been interested in using HPC resources for research ever since my undergrad years back in India. At UCLA I worked on development of high order numerical methods with applications in hypersonic flow simulation, and while there our research group built several clusters using commodity hardware, which started my involvement with the design and testing of HPC systems. It was a natural transition from there to helping with the HPC systems at SDSC and I liked the Center’s national reach. When I joined SDSC the frontline system, DataStar, was based on IBM’s Power 4 processors, which was state-of-the-art at that time. Fast forward more than a decade and it’s incredible to see the advances in HPC with Comet and now Expanse!

Q: For you, and from a user’s perspective, what’s the most impressive feature about Expanse that potential users should be aware of?

Expanse has many innovative features that make it a great resource for HPC-enabled research. If I had to pick one aspect, the AMD EPYC 7742 Processor would be its most important feature. Its innovative design and per core performance make Expanse between 1-1.8x faster than Comet, meaning overall throughput will easily be doubled.

Q: How would that benefit researchers?

The increased per core performance and nearly 2X the core count means that users can greatly reduce their time to completion of their research. Additionally, each rack on Expanse has more than 4X the cores when compared to Comet, allowing users to scale up to much larger core counts within a rack.

Q: One final question: What do you like to do outside your (home) office when it’s time to remotely unplug from SDSC?

When I’m not working on supercomputers, my favorite activity is hiking with family and friends in the California hills and mountains. My hikes have ranged from the coastal trails in San Diego to the nearby mountains and even the glaciers and peaks in the Sierras. When I’m not hiking, I’m either watching sports or messing around with my electric guitar and a new Spark amplifier, which interestingly has AI-enabled features!
Beginning the Voyage to HPC and AI Integration

The previous installment of this column discussed SDSC's latest innovative High-Performance Computing (HPC) resource, Expanse, which entered production at the end of 2020. Expanse will provide a modernized foundation for SDSC's ongoing industrial partnership program, delivering ample and scalable compute power in support of industry-sponsored projects and exploring new concepts in high-performance computing such as composable computing with Kubernetes, hybrid HPC-and-Cloud, and the efficiency of direct-to-chip liquid cooling.

Also, in 2020 SDSC received an award from the National Science Foundation for the Voyager system, targeted at exploring artificial intelligence (AI) processors in science and engineering. Voyager will address the increasing employment of AI methods in scientific computing and the joining of conventional modeling and simulation with AI techniques in a complementary fashion. The Voyager experiment targets scientific use cases in astronomy, chemistry, materials science, genetics, microbiome studies, satellite image analysis, computer science, and more.

SDSC will work with its integration partner, Supermicro Inc., to assemble and deploy Voyager with unique training and inference processors for initial availability in October 2021. The Voyager program supports strong industrial engagement as an integral component of the project.

Please stay tuned for announcements in the coming months as to how you can participate. With Voyager joining Expanse on the data center floor later in 2021, SDSC will have a strong foundation for supporting and engaging with industrial partners in the years to come. We look forward to being in touch about opportunities related to those systems, and to hosting you at the Center when conditions permit. In the meantime, to contact us with your questions or to learn more, please visit SDSC's Industry page, email us, or give us a call.

Ron Hawkins
SDSC Director of Industry Relations

Note: The Expanse and Voyager systems are made possible through support from the National Science Foundation via award numbers 1928224 and 2005369, respectively.