SDSC Shares Spotlight in Top Supercomputing Achievement Award – Twice!

SDSC shared top honors at last month’s International Conference for High Performance Computing, Networking, Storage and Analysis (SC16) in Salt Lake City, Utah, with two HPCwire ‘Top Supercomputing Achievement’ awards that recognized the use of the Center’s Comet supercomputer to help verify Einstein’s theory of gravitational waves. The awards, won in both the online publication’s annual Readers’ Choice and Editors’ Choice categories, also recognized the Open Science Grid (OSG), the National Science Foundation’s Extreme Science and Engineering Discovery Environment (XSEDE), the Holland Computing Center at University of Nebraska-Lincoln (UNL), the National Center for Supercomputing Applications (NCSA), and the Texas Advanced Computing Center (TACC). SDSC also won an HPCwire Editors’ Choice Award for ‘Best Use of High-Performance Computing in the Cloud.’ That award was for the collaboration between SDSC and Mellanox to create a virtual off-loading computing processor used in Comet to optimize the processing of partner workloads.

Read more at https://goo.gl/oFZalp
See page 6 for more SC16 Highlights!
I’m proud to say that 2016 was a landmark year for SDSC. We generated a combined $22.5 million in research grants during the fiscal year ended June 30, and that’s not counting some significant awards made during the second half of 2016. Moreover, SDSC’s overall success rate on federal proposals was 45 percent, more than twice the national average for computer science and engineering proposals at the National Science Foundation. We spent much of 2016 developing a comprehensive strategic plan to guide SDSC during the next several years. The plan focuses on three strategic priorities:

**Develop and deliver versatile, advanced computing platforms and technologies:** This priority builds on SDSC’s core strengths in advanced computing, which has become essential to nearly every field of research. We plan to remain at the forefront of how more researchers can access these systems, and how future systems will be designed to better serve the research community.

**Deliver data platforms to ingest, process, and analyze Big Data:** Large scientific experiments, gene sequencers, machine and human sensors, and the Internet of Things (IoT) all have one thing in common: they are extremely data-intensive and part of what’s called ‘big data.’ They also pose major technological, applications, and workforce challenges in developing end-to-end solutions to answer today’s most pressing scientific questions. This application-driven priority aligns with ones at the national level, such as ‘smart’ connected communities; developing new research tools; wearable biomedical devices; and research innovations at the nexus of food, energy, and water systems.

**Partner with the life sciences research community in the development and delivery of advanced computing and data platforms, information discovery, and data management services:** SDSC has established itself as a valuable partner with those conducting research in biomedical sciences, including genomics, brain research, and personal health monitoring systems. Our biomedical computing expertise and resources already include computational platforms, data management solutions, application performance optimization, scientific workflow software, and even compliance environments for computing on protected health data. We view life sciences as a major area of opportunity, such as new discovery platforms that integrate large but disparate biomedical data to create a comprehensive knowledge base for biomedical scientists. Such initiatives will in turn increase our value to campus, the UC system, the state of California, and of course the national research community.

In early 2017 we will outline specific goals and the projects that support our strategic plan. SDSC’s future remains bright and this ‘roadmap’ will ensure that we take full advantage of the opportunities that lie ahead.

Happy holidays and a peaceful and prosperous New Year to all! I’m sure that many exciting projects will make SDSC an invigorating place to be in 2017!

Michael L. Norman
SDSC Director
Homeland Security Awards CAIDA $1.4M to Help Prevent Internet Cyberattacks

The Center for Applied Internet Data Analysis (CAIDA) at SDSC was awarded a $1.4 million grant from the U.S. Department of Homeland Security to demonstrate and illuminate structural and dynamic aspects of the internet infrastructure relevant to cybersecurity vulnerabilities. The grant will provide funds to CAIDA to expand the scale and capabilities of its secure measurement platform Archipelago (Ark), which supports large-scale active measurement studies of the global internet.

"(This) project will deliver groundbreaking capabilities in internet measurement and identification of infrastructure vulnerabilities," said Ann Cox, program manager for Internet Measurement and Attack Modeling, a DHS program that supports researchers in academia and the cybersecurity community developing solutions in areas of resilient systems, modeling of internet attacks and network mapping and measurement.

Read more at https://goo.gl/qI3lPN

SDSC, Dell EMC, and Intel Announce Life Science Applications Initiative

Spurred by the increasing reliance of life sciences researchers in the academic and private sectors on computational methods and data-enabled science, SDSC has inaugurated a new life sciences computing initiative focused on improving the performance of bioinformatics applications and related analysis pipelines on current and future computing systems. The initial work, supported by Dell EMC and Intel, focuses on benchmarking and profiling selected genomic and Cryo-EM analysis pipelines and developing targeted recommendations for technical architectures to service those pipelines.

"UC San Diego is a pillar in one of the most vibrant life science research ecosystems in the world, and life science computing has always been a major focus of SDSC," said SDSC Director Michael Norman. "We view this partnership as a unique opportunity to bring the combined experience of SDSC, Dell EMC, and Intel to bear on one of the most important areas of data-intensive research today."

Read more at https://goo.gl/0YIY4j
Amit Majumdar is Director of SDSC’s Data Enabled Scientific Computing division and an Associate Professor in the Department of Radiation Medicine and Applied Sciences at UC San Diego, where he conducts collaborative research involving high-performance computing (HPC) applications for adaptive radiation therapy. Majumdar received his B.S. from Jadavpur University in Calcutta, India, in 1985; his M.S. from Idaho State University in 1988; and his Ph.D. from the University of Michigan in 1996. He joined SDSC in 1997.

Q: How is DESC structured and what capabilities does it offer users?
A: DESC is responsible for the High Performance Computing Systems group, User Services group, Scientific Computing Applications group, and Scientific Visualization group. SDSC’s XSEDE program is coordinated from the DESC division and the recently formed Advanced Technology Lab is also part of DESC. DESC staff members have degrees in domain sciences in areas such as chemistry, computer science, physics, applied mechanics, astrophysics, bioinformatics, and various branches of engineering. All staff members have expertise in high-performance computing. They design and maintain HPC systems and provide user support for our supercomputers, and also work on funded research projects involving various scientific applications.
Q: What kind of synergy comes from this type of collaboration?
A: Our group is nationally recognized for its contributions to user support and the XSEDE project. We collaborate with researchers from various U.S. universities and other HPC centers. Many DESC staff are PIs and Co-PIs of their own research projects funded by the National Science Foundation (NSF) and other government agencies and industries.

Q: What is an example of outcomes from this independent research?
A: One would be related to what are known as science gateways, or a set of tools, applications, and data that are integrated via a web-based portal or a suite of applications, customized to meet the needs of a specific community. I'm personally involved as PI of the Neuroscience Gateway (NSG), which offers neuroscientists access to supercomputers as well as providing a community forum for collaboration.

Q: Can you tell us more about the Neuroscience Gateway and how it is being used?
A: In 2015, the NSF and the United Kingdom's Biotechnology and Biological Sciences Research Council awarded funding for the second time for the Neuroscience Gateways project led by SDSC. This is a collaborative project between SDSC, Yale University, and University College London, and is contributing to the national BRAIN initiative announced by the Obama administration in 2013 to advance researchers' understanding of the human brain. The gateway allows on-demand, automated communication between HPC resources and neuroscience researchers, and projects such as Open Source Brain (OSB), ModelDB, and the Neuroscience Information Framework (NIF). We already have a user community of over 400, and their yearly usage of supercomputing time is reaching 10 million core hours across HPC systems at SDSC, as well as the Texas Advanced Computing Center, the Pittsburgh Supercomputing Center, and Indiana University.

Q: Regarding your appointment in UC San Diego's Department of Radiation Medicine, how can HPC play a role in radiation therapy?
A: To optimize patient treatment, radiation oncologists need to solve three computationally intensive tasks: image analysis of tumor location, dose calculation, and optimal delivery so that radiation reaches only the tumor and not nearby critical organs and normal tissue. Currently, there are efficient computer algorithms and associated codes used to solve these in a clinical environment. Our idea is to use the power of HPC to speed up some of these calculations. We're exploring the use of small graphics processing units (GPUs) as the parallel computing framework for real-time treatment planning in a clinical environment. Another area is application of data analytics to look at patients' radiation therapy data and associated recovery results.

Q: How has high-performance computing changed since you joined SDSC?
A: It has changed by orders of magnitude. Just consider the number of cores we now run on. When I started 19 years ago, 32 cores were considered supercomputing. Now we run on 300,000 to 500,000 cores. That's four orders of magnitude in less than two decades in scientific codes scaling at that level! Also, science gateways have made it easy for many more users from more disciplines to use HPC resources, including disciplines which traditionally didn't use HPC. Data Science applications and HPC are working hand in hand now, and that is a very exciting area of research and development.

“Our group is nationally recognized for its contributions to user support and the XSEDE project. We collaborate with researchers from various U.S. universities and other HPC centers.”

Visualization of 3-D Cerebellar Cortex model generated by researchers Angus Silver and Padraig Gleson from University College London. The Neuroscience Gateway was used for this simulation.
The ‘Book of Norman’ and Other Highlights

SC16, the 28th annual conference of high-performance computing, attracted more than 11,100 registered attendees during its six-day run. There were almost 350 exhibitors from industry, academia, and research organizations from around the world, filling 150,000 square feet of space at the Salt Palace in Salt Lake City, Utah.

In keeping with our tradition of having a little fun during the conference, SDSC Director Mike Norman kicked off the Center’s presentations by introducing ‘The Book of Norman’, which contains his Top Tenets of Supercomputing and Data Science Management. While the book may not be as popular as another one with a somewhat similar title, here are Mike’s tenets:

**Top Tenets of Supercomputing**
- First and foremost, supercomputers must support Productive research
- Today’s systems should be highly Reliable
- They should be Open
- They should be Versatile
- They should be backed by Expertise
- Supercomputers should support what’s Next

**Top Tenets of Data Science Management**
- Platforms should use latest technologies for Large-scale storage and data protection
- Include an advanced, End-to-end cyberinfrastructure
- Pioneer innovative predictive Analytics
- Support basic and advanced courses in Data science
- Attract the nation’s top Experts in all things related to data
- Lead efforts aligned strongly with campus, state, and national Research priorities

You’ll note that the capitalized letters spell P-R-O-V-E-N L-E-A-D-E-R. SDSC backed up those words with a series of presentations, including a panel discussion about the recent $15 million National Science Foundation (NSF) grant for SDSC to lead the formation of a Science Gateways Community Institute to accelerate the development of highly functional, sustainable science gateways that meet the needs of researchers across the full spectrum of NSF directorates. SDSC Associate Director Nancy Wilkins-Diehr is the principal investigator for that project.

SDSC also held a series of short tutorials covering topics such as machine learning for face recognition. The tutorials were led by HPC Applications Specialist Andrea Zonca, and participants had an opportunity to win one of 16 Raspberry Pi 3 mini-clusters. All in all, SDSC’s display area drew a steady stream of visitors, making it one of the best SC conferences ever for the Center.
Comet, Gordon Assist in Finding Promising Drug Leads to Combat Heart Disease

Using a unique computational approach to rapidly sample proteins in their natural state of gyrating, bobbing, and weaving, a research team from UC San Diego and Monash University in Australia has identified promising drug leads that may selectively combat heart disease, from arrhythmias to cardiac failure. Reported in the September 5, 2016 Proceedings of the National Academy of Sciences (PNAS) Early Edition, the researchers used the computing power of SDSC’s Gordon and Comet to perform an unprecedented survey of protein structures using accelerated molecular dynamics or aMD – a method that performs a more complete sampling of the myriad shapes and conformations that a target protein molecule may go through.

Read more at https://goo.gl/DXMZO9

Los Angeles Fire Dept. and UC San Diego WIFIRE Team Join Forces to Fight Wildfires

The Los Angeles Fire Department, challenged by yet another series of late summer wildfires, successfully tested a new web-based tool developed by UC San Diego researchers to perform data-driven predictive modeling and analyses of fires that have a high potential for rapid spread. Called Firemap and developed by the WIFIRE collaboration headed by SDSC Chief Data Science Officer Ilkay Altintas, the new tool enables a ‘what-if’ analysis of fire scenarios ahead of time as well as in real-time fire forecasting. WIFIRE’s Firemap data resource also provides easy access to information on past fires, past and current weather conditions as well as weather forecasts, satellite detections as fast as they are received, and information on vegetation and landscapes from a variety of sources. The goal of WIFIRE, the result of a multi-year National Science Foundation grant, is to make data and predictive models readily available so that the direction and rate of fire spread can be known as early as possible to assist in rescue and containment efforts.

Read more at https://goo.gl/qEQS9R

New Drug Candidate May Reduce Deficits in Parkinson’s Disease

An international team led by UC San Diego researchers has employed a novel computational approach to design and create a new compound that in laboratory studies has reduced deficits and neurodegenerative symptoms that underlie Parkinson’s disease. In a study published in the September 27 Advance Access issue of Brain, the researchers describe how their compound prevents the binding and accumulation of alpha-synuclein or α-syn in neuronal membranes, now considered a hallmark of Parkinson’s disease and a related disorder called dementia with Lewy bodies. “We’ve demonstrated a novel computational approach to design potential therapies for Parkinson’s disease and related disorders,” said the study’s co-first author Igor Tsigelny (left), a research scientist with SDSC, the UC San Diego Moores Cancer Center, and Department of Neurosciences.

Read more at https://goo.gl/iVkykV
During the recent presidential campaign, subjects of much debate and interest were the intertwined topics of U.S. manufacturing and trade policy. Campaign rhetoric aside, manufacturing actually continues to play a key role in the U.S. economy and is closely linked to R&D and technology development.

A study published in 2015 by Deloitte Touche Tohmatsu found that "advanced manufacturing industries" accounted for 70% of the 40 million jobs and $2.7 trillion in economic output generated by in the U.S. The study further found that the Internet of Things (IoT) and Predictive Data Analytics were two of the most promising technologies for driving innovation and competitiveness in U.S. advanced manufacturing activities.

IoT and predictive analytics can work hand-in-hand to optimize advanced manufacturing systems, leading to increased efficiencies, throughput, and quality control. IoT permits the deployment of thousands of low-cost sensors throughout a manufacturing line and even on the manufactured components themselves. Continuous measurements can be streamed and collected in a "data lake," where predictive analytics and machine learning techniques can be applied to reveal hidden insights on how to improve the processes.

Through its membership in the Smart Manufacturing Leadership Coalition (SMLC) based in Los Angeles, SDSC researchers are ready to contribute their expertise to the development of new smart manufacturing technologies and systems. In June of this year, the SMLC received an award from the Department of Energy for the Clean Energy Smart Manufacturing Innovation Institute (CESMII), a $140 million, five-year public-private partnership and the most recent award in a national network of Manufacturing Innovation Institutes. CESMII will focus on accelerating the development and adoption of advanced sensors, data analytics, and controls in manufacturing while reducing the cost of these technologies by half and improving the efficiency of U.S. advanced manufacturing.

As this initiative gathers momentum in 2017, SDSC looks forward to working with the DOE and the many academic organizations and companies that are part of CESMII to deliver new innovations in smart manufacturing technology. If smart manufacturing, IoT, and predictive analytics are important to the future of your business, please feel free to contact us for more information or to explore potential collaborations.

Ron Hawkins
SDSC Director of Industry Relations