

## SDSC to Host “Grand Challenges in Data-Intensive Discovery” Conference

*Oct. 26-28 Event to Explore Opportunities for Flash Memory-based ‘Gordon’ Supercomputer to Debut in 2011*

August 2, 2010

Jan Zverina

The San Diego Supercomputer Center (SDSC) at the University of California, San Diego, will host a special conference in late October as it prepares to deploy a unique data-intensive, high-performance computing (HPC) system called *Gordon* in mid-2011.

Called "Grand Challenges in Data-Intensive Discovery", the multi-disciplinary conference will be held October 26-28, 2010 at SDSC, located on the UC San Diego campus. Attendees can register for the conference beginning August 2 by visiting [www.sdsc.edu/gordongrandchallenge](http://www.sdsc.edu/gordongrandchallenge).

"Science has entered a data-intensive era, driven by a deluge of data being generated by digitally based instruments, sensor networks, and simulation devices," said Michael Norman, interim director of SDSC. "Hence, a growing part of the scientific enterprise is associated with analyzing such data, placing special demands on computer architectures because the associated calculations have frequent I/O accesses, large memory requirements, and often limited parallelism.

"We believe this conference will greatly benefit those doing research in data-intensive fields," added Norman. "*Gordon*, and future systems like it, will open new opportunities across numerous areas of research, and it is important that such resources are utilized to their utmost capabilities." The goal of the GCDID conference is to provide an opportunity for attendees to share their expertise while exchanging ideas about the computational challenges and concerns common to data-intensive problems. Specifically, the conference is structured to facilitate discussion to help:

Articulate and clarify "Grand Challenges" in data-intensive research across a broad range of disciplines, including arts, astronomy, biology, computer science, earth sciences, economics, engineering, humanities, medicine, neuroscience, social sciences, and data-related technologies

Identify applications and disciplines that can benefit from *Gordon's* unique architecture and capabilities, including those that have not been part of the traditional HPC community

Identify common technical needs across disciplines and relevant software solutions

Recognize opportunities for leaders in data-intensive science to take advantage of SDSC's available expertise in this area

Speakers scheduled to address the GCDID conference, according to topic, include:

*Visual Arts* - Lev Manovich, UC San Diego

*Needs and Opportunities in Observational Astronomy* - Alex Szalay, Johns Hopkins University

*Transient Sky Surveys* - Peter Nugent, Lawrence Berkeley National Laboratory

*Large Data-Intensive Graph Problems* - John Gilbert, UC Santa Barbara  
*Algorithms for Massive Data Sets* - Michael Mahoney, Stanford University  
*Needs and Opportunities in Seismic Modeling and Earthquake Preparedness* - Tom Jordan, University of Southern California  
*Economics and Econometrics* - James Hamilton, UC San Diego  
*Needs and Opportunities in Fluid Dynamics Modeling and Flow Field Data Analysis* - Parviz Moin, Stanford University  
*Needs and Emerging Opportunities in Neuroscience* - Mark Ellisman, UC San Diego

*Data-Driven Science in the Globally Networked World* - Larry Smarr, UC San Diego  
*Accelerating Data - Intensive Science with Dash and Gordon* - Allan Snively, San Diego Supercomputer Center  
*Data Challenges in Biomedical Informatics* - Lucila Ohno-Machado, UC San Diego  
*Needs and Opportunities in Computational Biology* - Rick Stevens, Argonne National Laboratory. \*

\* to be confirmed

Contributed talks, panel discussions, sharing of early user experiences, and posters will complement the plenary speakers listed above. The schedule will be posted on the conference website shortly. "With a peak speed of 245 teraflops, very large shared memory nodes, and a quarter-petabyte of flash SSD (solid state drive) memory, *Gordon* will vastly accelerate large database and data mining applications," said Norman. "This system will reduce solution times and yield results for applications that now tax even the most advanced supercomputers, while helping to make sense of the avalanche of data generated by the digital devices of our era."

SDSC announced last November that it won a five-year, \$20 million grant from the National Science Foundation (NSF) to build and operate *Gordon*, the first high-performance supercomputer to employ a vast amount of flash memory to help speed solutions now hamstrung by slower spinning disk technology. *Gordon* will be a peer-reviewed, allocated resource on NSF's TeraGrid, and be available to any U.S. researcher. TeraGrid is the nation's largest open-access scientific discovery infrastructure.

SDSC has now deployed *Dash*, a smaller prototype of *Gordon* that gives prospective users an opportunity to explore *Gordon*'s unique architectural features. *Dash* also gives computer scientists and systems architects a way to address some of the key challenges that must be met as part of *Gordon*'s deployment. Attendees will have an opportunity to hear from early *Dash* users during a panel session focused on early user experience.

**About SDSC** As an Organized Research Unit of UC San Diego, SDSC is a national leader in creating and providing cyberinfrastructure for data-intensive research, and is celebrating its 25th anniversary this year as one of the National Science Foundation's first supercomputer centers. Cyberinfrastructure refers to an accessible and integrated network of computer-based resources and expertise, focused on accelerating scientific inquiry and discovery. SDSC is a founding member of TeraGrid, the nation's largest open-access scientific discovery infrastructure.

Media Contacts: Warren R. Froelich, SDSC Communications, 858-822-3622 Jan Zverina, SDSC Communications, 858-534-5111