CCT Professor Receives NSF Award to Develop Innovative Data Scheduler

Professor Tevfik Kosar received a half-million dollar grant from the National Science Foundation to support his work on Stork Data Scheduler, an innovative computing tool that helps researchers access and transfer large data sets easily and efficiently.

The grant award provides Kosar, a professor with the LSU Department of Computer Science who holds a joint appointment with the CCT, with funding for three years to further develop and enhance the Stork Data Scheduler.

As computational science applications expand and become increasingly complex, researchers using these applications are generating larger and larger amounts of data, sometimes up to hundreds of terabytes and even petabytes. Sharing, disseminating, and analyzing these large data sets is a growing challenge for researchers, who need to collaborate but are unable to move so much information quickly or effectively.

Even though many researchers now have access to regional high-speed, fiber optic networks such as the Louisiana Optical Network Initiative, or LONI, many users cannot obtain even a fraction of the theoretical speeds these networks promise because of data overload, which slows transmission and causes a bottleneck in computational performance and reliability.

Kosar’s project, funded through the National Science Foundation’s Strategic Technologies for Cyberinfrastructure Program, aims to ease data bottlenecks, which will improve high-performance computing systems’ performance. Stork, so named because it delivers data, is a batch scheduler program that makes it easier for researchers to share, store and deliver data across these systems.

Using Stork, researchers can transfer very large data sets with only a single command, making it one of the most powerful data transfer tools available. Stork is compatible with advanced high-performance computing toolkits, and researchers can use the software to access the power of these large systems and use them more effectively.

“The Stork data scheduler makes a distinctive contribution to the computational research community because it focuses on planning, scheduling, monitoring and management of data,” Kosar said. “Unlike existing approaches, Stork treats data resources and their related tasks as primary components of computational resources, not simply as side effects. This will lead to quicker and more effective collaboration among researchers.”
Researchers consider the Stork Data Scheduler a highly transformative project because of its potential to dramatically change how scientists perform their research and to rapidly facilitate sharing of experience, raw data, and results. Future applications could rely on Stork to manage storage and data movement reliably and transparently across many systems, eliminating the unnecessary failure of distributed tasks.

The Stork team made the first version (Stork 1.0) available for download through the Stork project Web page, www.storkproject.org, in December 2008. Stork is open source, and users can download it for free.

Data storage and management is Kosar’s research specialty at the University. In 2006, he received a $1 million grant from NSF to create advanced data archival, processing and visualization capabilities across the state through the PetaShare project (www.petashare.org).

Kosar received a National Science Foundation CAREER Award in January 2009 for his research addressing the problems of distributed data storage and transfer. Through his work on the CAREER project, titled “Data-aware Distributed Computing for Enabling Large-scale Collaborative Science,” Kosar is developing the theory and foundations of new computing systems that manage data more effectively with automated processes. These processes enable scientists to spend more time focusing on their research questions and less time dealing with data. This project is funded for five years at $400,000.

Kosar’s recent STORK grant expands on the models and algorithms created through his CAREER grant work, implementing them in a scheduling software program that will be available for production and distribution.

“Dr. Kosar has consistently demonstrated creative and innovative solutions for large scale data management. We are both proud and excited that he has received funding to continue his work, which will benefit other University researchers and the broader scientific community,” said CCT Interim Director Stephen David Beck.

Pats on the Back:
• Congratulations to the CCT faculty and staff who organized a successful Cluster 2009 conference in New Orleans last week. Daniel S. Katz served as general chair, and Thomas Sterling was program chair for this event.

CCT in the News:
• LSU PROFESSOR RECEIVES NSF AWARD TO DEVELOP STORK DATA SCHEDULER
  Source: HPC Wire
  http://www.hpcwire.com/topic/middleware/LSU-Professor-Receives-NSF-Award-to-Develop-Stork-Data-Scheduler-56964947.html?viewAll=y

Lectures This Week:
• Xiaoliang Wang, an LSU researcher, will speak as part of the Computational Mathematics Series on “Noise-Induced Transition for the Kuramoto-Sivashinsky Equation” today at 3 p.m. in Johnston 338.

Please Note:

• The next HPC training, Accounts Allocation and Management, will take place Thursday, Sept. 10 from 1:30 p.m. to 3:30 p.m. in 338 Johnston and on the Internet. Please visit: http://www.hpc.lsu.edu/training/ for more details and to register.

• Future ALL CCT meetings for the Fall 2009 semester will take place Sept. 23, Oct. 21, Nov. 11 and Dec. 16. All meetings are at 3 p.m. in Johnston 338 unless otherwise announced. Please make every effort to attend these important meetings.

• Registration is now open for the Supercomputing 2009 Education Program at the conference in Portland, which will take place Nov. 14-17. The Education Program helps educators and students learn more about computational science topics and gives educators ideas to bring these topics into their classrooms. The program is open to undergraduate faculty, undergraduate and graduate students, and high school teachers. To register or for more information, please visit http://computationalscience.org/sc09.

• The SC09 Student Contest Program is accepting team registrations. This is a competitive programming event, where teams of no more than five students will be given eight to 12 problems from various scientific problem domain areas. The competition will take place Monday, Nov. 16 at the SC09 conference in Portland, Oregon. Awards will be announced on Tuesday, November 18 at an SC09 Education Program plenary session. Register your team today, http://sc09.sc-education.org/conference/studentcomp_signup.php. Deadline to register is Thursday, October 1, 2009.

• The Baton Rouge Area Convention and Visitors Bureau keeps an updated list of events and social activities for the area, including the Live After Five concert series schedule, on its Web site, http://visitbatonrouge.com. The bureau also produces an electronic newsletter of upcoming events. Please visit the site for ideas on places to take visitors or to find new things to do in the area.

• Please remember to send your news concerning grants, awards, conferences, or other pertinent information that should be communicated to CCT to PR Manager Kristen Sunde at ksunde@cct.lsu.edu.

Upcoming Grant Deadlines:

Note: Please see the CCT deadline Web site, as many NSF deadlines are listed here: http://www.cct.lsu.edu/about/grants/deadlines/events.php

• Social Computational Systems (SoCS)
September 21, 2009 10:00 a.m.
At most $250,000.00 available

- CreativeIT
  October 13 2009 10:00 a.m.
  A portion of $7,000,000.00 available