For more information on the lecture, please contact Dr. Othmane Bouhali.

RSVP

othmane.bouhali@qatar.tamu.edu

tasc
Texas A&M at Qatar
Advanced Scientific Computing

TEXAS A&M UNIVERSITY at QATAR

Server-side Log Data Analytics for Supercomputer I/O Workload Characterization

Xiaosong Ma
Principal Scientist, Qatar Computing Research Institute, HBKU

2 May, 1–2 p.m., LH143

Inter-application I/O contention and performance interference has been recognized as a severe problem. In this work, we demonstrate that high I/O variance co-exists with the fact that individual storage units remain under-utilized for the majority of time. This motivates us to propose IOAI, a system that performs automatic application I/O characterization and I/O-aware job scheduling. IOAI analyzes existing I/O traffic and batch job history logs, without any application-related prior knowledge or user/developer's involvement. It identifies the small set of I/O-intensive candidates among parallel applications running on a supercomputer and subsequently mines their I/O patterns, using more detailed per-I/O-node traffic logs. Based on such auto-extracted information, IOAI provides online I/O-aware scheduling recommendations to steer I/O-intensive applications away from heavy ongoing I/O activities.

We evaluate IOAI on the same supercomputer, using both real applications (with extracted I/O patterns validated by contacting users) and our own pseudo-applications. Our results confirm that IOAI is able to (1) identify I/O-intensive applications, plus their detailed I/O characteristics, and (2) significantly reduce these applications' I/O performance degradation/variance by jointly evaluating outstanding applications' I/O pattern and real-time system I/O load.

Xiaosong Ma

Xiaosong Ma is currently a principal scientist at Qatar Computing Research Institute, HBKU. Previously, she was associate professor in the Department of Computer Science at North Carolina State University, as well as a joint faculty member at Oak Ridge National Laboratory. Her research interests are in the areas of storage systems, parallel I/O, high-performance parallel applications, cloud computing, and self-configurable performance optimization. She received the U.S. DOE Early Career Principal Investigator Award in 2005, the NSF CAREER Award in 2006, an IBM Faculty Award in 2009, and a NetApp Faculty Fellowship, an HPDC 20-Year Best Paper Award, as well as a University of Illinois Department of Computer Science Alumni Distinguished Educator Award in 2012. Xiaosong received her Ph.D. in computer science from the University of Illinois at Urbana-Champaign in 2003 and her B.S. in computer science from Peking University (China) in 1997.