THE CHINESE UNIVERSITY OF HONG KONG
Department of Information Engineering
Seminar

Data Access of Big Data: Solving the system bottleneck
By
Prof. SUN Xian-He
Illinois Institute of Technology, USA

Date : **14th May, 2019 (Tue)**
Time : **11:00am – 12:00nn**
Venue : **Room 833, Ho Sin Hang Engineering Building**
The Chinese University of Hong Kong

Abstract

We have entered the era of big data. However, the performance improvement of disk-based storage systems has been much slower than that of memory and network, creating a significant I/O performance gap. To reduce the performance gap, storage subsystems are under extensive changes, adopting new technologies and adding more layers into the memory/storage hierarchy. With a deeper memory hierarchy, the data movement complexity is increased significantly, making it harder to utilize the potential of the deep memory-storage hierarchy (DMSH) architecture. In this talk, we first show that storage data access is the most concerned performance bottleneck, even in a distributed mobile environment. Then, we present the development of *Hermes*, an intelligent, multi-tiered, dynamic, and distributed I/O caching system that utilizes DMSH to significantly accelerate I/O performance. Hermes is a US NSF supported large software development project. It extends HPC I/O stacks to integrated memory and parallel I/O systems, extends the widely used Data Format (HDF) and HDF5 library to understand users' need and to achieve application-aware optimization in a DMSH environment. We will introduce the Hermes' design and implementation; discuss its uniqueness and challenges; and present some initial design and implementation results.

Biography

Dr. **Xian-He Sun** is a University Distinguished Professor of Computer Science at the Department of Computer Science in the Illinois Institute of Technology (IIT). He is the director of the Scalable Computing Software laboratory at IIT and a guest faculty in the Mathematics and Computer Science Division at the Argonne National Laboratory. Before joining IIT, he worked at DoE Ames National Laboratory, at ICASE, NASA Langley Research Center, at Louisiana State University, Baton Rouge, and was an ASEE fellow at Navy Research Laboratories. Dr. Sun is an IEEE fellow and is known for his memory-bounded speedup model, also called Sun-Ni's Law, for scalable computing. His research interests include data-intensive high-performance computing, memory and I/O systems, software system for big data applications, and performance evaluation and optimization. He has over 250 publications and 6 patents in these areas. He is the Associate Chief Editor of IEEE Transactions on Parallel and Distributed Systems, a Golden Core member of the IEEE CS society, a former vice chair of the IEEE Technical Committee on Scalable Computing, the past chair of the Computer Science Department at IIT, and is serving and served on the editorial board of leading professional journals in the field of parallel processing. Dr. Sun received the Overseas Outstanding Contributions Award from the China Computer Federation (CCF) in 2018. More information about Dr. Sun can be found at his web site www.cs.iit.edu/~sun/.

** ALL ARE WELCOME **

Host: Prof XING Guoliang (Tel: 3943-8474, Email: guoliang@ie.cuhk.edu.hk)
Enquiries: Information Engineering Dept., CUHK (Tel.: 3943-8385)