

ISTeC

The Information Science and Technology Center



Colorado
State
University

ISTeC.ColoState.edu



Colorado State University's
Information Science and Technology Center (ISTeC)
presents two lectures by

David Hung-Chang Du

Qwest Chair Professor of Computer Science and
Engineering at University of Minnesota, Minneapolis

ISTeC Distinguished Lecture

In conjunction with the Computer Science Department,
and Electrical and Computer Engineering Department

"Intelligent Storage in Big Data Era"

Monday, March 23, 2015
Reception with refreshments: 10:30 am
Lecture: 11:00 am – 12:00 noon
Location: Morgan Event Hall



Computer Science Department, and Electrical
Engineering Department Special Seminar Sponsored by ISTE C

"Efficient Usage of Two Emerging Memory/Storage Technologies: NVRAM and Shingled Write Recording Drives"

Monday, March 23, 2015
Lecture: 2:00 pm – 3:00 pm
Location: Engineering E105

ISTeC (Information Science and Technology Center) is a university-wide organization for promoting, facilitating, and enhancing CSU's research, education, and outreach activities pertaining to the design and innovative application of computer, communication, and information systems. For more information please see ISTeC.ColoState.EDU.

Abstracts

Intelligent Storage in Big Data Era

The Internet today has grown to an enormously large scale. Devices large and small are connected globally from anywhere on the earth. Therefore, we can argue that we are in a network centric era. With the rapid advancement of technology, we now also have cheap and small devices with high computing power and large storage capacity. These devices are designed to improve our daily life by monitoring our environment, collecting critical data, and executing special instructions. These devices have gradually become a dominant part of our Internet. Many imaging, audio and video data are converted from analog to digital and digital data are generated at an alarming rate. As a result, unprecedented amount of data are available. How to manage and look for the desired information becomes a great challenge. How to preserve these data becomes a crisis. We can certainly also say that we are in a data centric era. In fact we are currently witnessing the emerging of these two into a new era of big data. In this talk, we will examine the challenges and research trends in storage research for this new big data era. Several other major developments like cloud computing, mobile computing, new memory/storage technologies are certainly creating big impact in this new era. In this talk, we will present a vision of content addressable future Internet. What are the essential changes in data representation, information retrieval, storage systems and networking design will be discussed. We believe an object-oriented intelligent storage is an essential part of the solution to this new computing and communication environment. We will also present a number of research projects that are currently under investigation in our NSF I/UCRC Center on Intelligent Storage. These projects include data deduplication, long-term data preservation, data center power management, and flash memory based solid state drives.

Efficient Usage of Two Emerging Memory/Storage Technologies: NVRAM and Shingled Write Recording Drives

The past decade has witnessed tremendous advances in computing, wired and wireless communication and storage technologies. It is also important that remarkable cost reductions have made large computing and storage capacity available to increasing number of consumers. With the unprecedented connectivity provided by Internet, many new applications have emerged and are being developed. A huge amount digital data has been created and become available for accesses to satisfy the demand of these new applications. It becomes an increasing challenge to even store this huge amount available data for our daily use. To answer the challenges in both computing and storage in this big data era, we more heavily depend on emerging memory/storage technologies. In this talk, we will cover the usage of Non-Volatile memory (NVRAM) and a new storage technology called Shingled Writing Recording.

All computers today have DRAM (Dynamic Random Access Memory) as main memory due to its fast access speed. Main memory in a computer is used for holding executable programs as well as caching data from slower storage devices. Since DRAM is volatile, the data updated has to be quickly flushed to storage devices like hard disks. Otherwise, the computer may run the risk of losing data if the computer or power failed. Because of this, the caching algorithms in the past only consider read cache since writes will be written to disks within a very short duration like less than 30 seconds. With the emerging non-volatile memory (NVRAM) becoming possible main memory, we have to consider caching for both reads and writes. In this talk, we will discuss a new caching algorithm called H-ARC. H-ARC is an extension of ARC caching algorithm and H-ARC will consider both read and write cache together. Our evaluation shows H-ARC can improved both read and write performance at the mean time it can dramatically reduce the traffic between main memory and storage devices like disk drives.

The big data era has brought tremendous potentials and benefits to the society. At the same time, it creates a crisis of “drowning in data”. Individuals cannot identify and locate the right information for their needs and organizations are struggling to manage and preserve the existing data. A complete solution to meet this challenge calls for new storage architectures, different storage systems design, new data models, new information access methods and new ways to deliver information. In this talk, we introduce a promising technology called Shingled Magnetic Recording (SMR) for storing large volume of data in magnetic disk drives (called Shingled Write Disks). In a SMR drive, data to be written sequentially with current track overlapped with some of the previous tracks to increase areal density of the drive. We will discuss the usage and several design issues of SMR drives.

Speaker Biography

Dr. Du is currently the Qwest Chair Professor of Computer Science and Engineering at University of Minnesota, Minneapolis. He has served as a Program Director (IPA) at National Science Foundation (NSF) CISE/CNS Division from March 2006 to September 2008. At NSF, he was responsible for NeTS (networking research cluster) NOSS (Networks of Sensor Systems) Program and worked on Cyber Trust (Internet Security) Program. He is also the Director of a NSF I/UCRC Center on Intelligent Storage (CRIS). CRIS has been sponsored by more than 10 companies including Seagate, NetApp, Symantec, HP, Dell, SGI, LSI, Xyratex, NEC Labs, HGST, Los Alamos National Lab, etc. Dr. Du received a B.S. degree from National Tsing Hua University in 1974, an M.S. and Ph.D. degree from University of Washington (Seattle) in 1980 and 1981 respectively.

He joined University of Minnesota as a faculty since 1981. Dr. Du has a wide range of research expertise including multi-media computing, mass storage systems, high-speed networking, sensor networks, cyber security, high-performance file systems and I/O, database design, and CAD for VLSI circuits. He has authored and co-authored over 260 technical papers including 110 referred journal publications in these research areas. He has graduated 54 Ph.D. and more than 100 M.S. students in the last 30 years. Dr. Du is an IEEE Fellow (since 1998) and a Fellow of the Minnesota Supercomputer Institute. He is currently serving on the Editorial Boards of several international journals. He has also served as Conference Chair and Program Committee Chair for several major conferences in multimedia, networking, database, parallel/distributed computing and security areas. Currently he is the General Chair of the 30th IEEE Symposium on Security and Privacy (2009) and Program Committee Co-Chair for the 37th International Conference on Parallel Processing (2009), General Chair of ICDCS (2011) and General Chair for ICPP (2014). He has had research grants from many federal funding agencies including NSF, DARPA, ONR, and DOE. He has a strong tie with many industrial researchers and has collaborated with a number of companies.

To arrange a meeting with the speaker, please contact Prof. Liuqing Yang, lqyang@mail.colostate.edu.