ECE Seminar

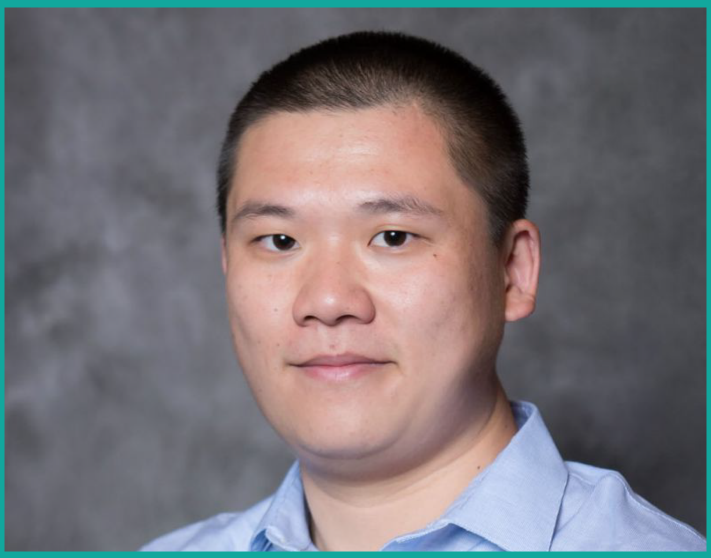
2019 ECE SPEAKER SERIES

Dingwen Tao

Assistant Professor

Department of Computer Science

University of Alabama



Fri. Nov. 1, 2019 @ 1 PM

8 St. Mary's Street, PHO 339

Faculty Host: Martin Herbordt

Light refreshments will be available outside of 339 at 12:45 PM

KEEPING-UP WITH SCIENTIFIC DATA EXPLOSION IN THE ERA OF EXASCALE COMPUTING

Abstract: The next generation of supercomputers will be exascale high-performance computing (HPC) systems, which are capable of at least 1018 floating-point operations per second, or a factor of 10 times faster than the nation's most powerful supercomputers in use today. The systems will help researchers tackle increasingly complex problems through modeling large-scale systems, such as nuclear reactors or global climate, and simulating complex phenomena. In order to achieve success, these systems must be able to reliably store enormous amounts of high-precision data and perform I/O at an extremely high rate. However, there are serious challenges to build a parallel file system with 10 times performance improvement. Thus, to overcome the gap between computation speed and file system's I/O speed/capacity, HPC researchers have to develop more intelligent and effective ways to reduce data size without losing important information. Data compression provides a good solution for reducing the data size. Although loss less compression can retain all the data information, it suffers from very limited compression ratio. To this end, we have developed a novel lossy scientific data compression framework called SZ, as well as a series of optimization techniques on different hardware (such as GPUs and FPGAs) for various scientific applications (e.g., quantum chemistry, cosmology). These techniques can significantly reduce the data size while maintaining a high fidelity for post-analysis through various accurate error control schemes. On the other hand, lossy compressor developers and users are missing a tool to understand the data alteration after compression in a systematic and reliable way. To address this gap, we have designed and developed generic frameworks (named Z-checker and Foresight), which can be used to systematically evaluate, analyze, and visualize the compression impacts on both application execution and postanalysis, taking into consideration domain features.

Bio: Dr. Dingwen Tao received his bachelor's degree in mathematics from University of Science and Technology of China in 2013 and his Ph.D. degree in computer science from University of California, Riverside in 2018. Currently, he is an assistant professor of computer science in The University of Alabama (UA) and leading the High Performance Data Analytics & Computing Lab (HiPDAC) at UA. Prior to joining UA, he worked at Brookhaven National Laboratory, Argonne National Laboratory, and Pacific Northwest National Laboratory. His research interests include high-performance computing, parallel and distributed computing, big data analytics, scientific data reduction and visualization, resilience and fault tolerance, and large-scale machine learning and deep learning. He has published over 25 peer-reviewedhigh-quality papers in prestigious ACM and IEEE conferences and journals, such as ACM ICS, HPDC, PPoPP, SC, IEEE BigData, CLUSTER, IPDPS, MSST, IEEE TPDS, SAGE IJHPCA. Dr. Tao's research has been supported by U.S. DOE and NOAA.

BU Department of Electrical & Computer Engineering