

# Spring 2016 Seminar Series

**DEDUPLICATION-AWARE ECOSYSTEM: A BOTTOM-UP APPROACH**

WEDNESDAY, MARCH 23, 2016

1:00 PM – HEC 356

In the era of big data, the rapid growth in data volume and complexity requires highly efficient schemes to reduce the amounts of data. Deduplication schemes can remove the redundant data, which is helpful to obtain space savings and improve network bandwidth efficiency. In this talk, I will present an overview of data deduplication schemes in a bottom-up manner. First, in the chunk-level deduplication, the locality and similarity are explored and exploited to support fast indexing to identify possibly duplicate data. Moreover, in the application level, similar images often consume large amounts of storage capacity, which can be efficiently addressed by feature-based compression approaches. Therefore, the deduplication schemes construct a bottom-up ecosystem to significantly improve system performance.

**DR. YU HUA**

Huazhong University of Science and Technology



Dr. Yu Hua is an Associate Professor in Huazhong University of Science and Technology, China. He obtained his B.E and Ph.D degrees from Wuhan University, respectively in 2001 and 2005. He was Postdoc Research Fellow in University of Nebraska-Lincoln in 2010-2011. His research interests include massive storage systems, storage class memory and data center networks. He publishes more than 60 papers in major journals and conferences, including IEEE Transactions on Computers (TC), IEEE Transactions on Parallel and Distributed Systems (TPDS), USENIX ATC, USENIX FAST, SC, HPDC, INFOCOM, ICDCS and MSST. He serves for multiple international conferences, such as INFOCOM, RTSS, ICDCS, MSST, ICNP and ICPP. He is a senior member of IEEE.

*Hosted by: Dr. Jun Wang*

