



北京大学海外名家讲学计划

DEDUPLICATION STORAGE SYSTEMS: REDUCING POWER AND SPACE

Kai Li

Paul M. Wythes & Marcia R. Wythes Professor
Princeton University

2015年8月26日 星期三 10:35am

中关村新园1号楼集贤厅



ABSTRACT: Since deduplication storage system was introduced over ten years ago, it has changed the landscape of the storage industry. Disk-based deduplication storage systems can reduce the footprint of backup data by 10-30X. Today's flash-based deduplication storage systems can reduce most of their popular workloads by about 5X. Such reductions also dramatically reduce the energy requirements in data centers. This talk gives an overview of the basic deduplication techniques for building storage systems and a few case studies of commercially available deduplication storage systems.

BIOGRAPHY: Kai Li is a Paul M. Wythes & Marcia R. Wythes Professor at Princeton University, where he worked as a faculty member in the computer science department since 1986. He received his Ph.D. degree from Yale University, M.S. degree from Chinese Academy of Sciences, and B.S. degree from Jilin University. His research expertise is in operating systems, parallel and distributed systems, de-duplication storage systems, large display systems, and data analysis and search of large feature-rich data. He pioneered Shared Virtual Memory or Distributed Shared Memory (DSM) which allows shared-memory programming model on clusters of computers, which won the ACM SIGOPS Hall of Fame Award in 2012. He proposed user-level DMA mechanism for efficient cluster communication has evolved into the RDMA in the Infiniband standard. He co-led the ImageNet project which constructed a large image database and propelled deep learning methods. In 2001, He co-founded Data Domain, Inc. (now an EMC division) and led the innovation of deduplication storage system products to replace the tape libraries for backup and archive at enterprise data centers. Since 2008, Data Domain product line has maintained over 60% of the deduplication backup storage market. He is an ACM fellow, IEEE fellow and a member of National Academy of Engineering.

主办单位：北京大学国际合作部、北京大学高能效计算与应用中心