電腦科學系
Department of
Computer Science

香港城市大學
City University of Hong Kong

SP-Cache: Load-balanced, Redundancy-free Cluster Caching with Selective Partition

SPEAKER **Prof Wei WANG**

Assistant Professor
Department of Computer Science
and Engineering
Hong Kong University of Science and
Technology
Hong Kong

DATE 12 February 2019 (Tuesday)
TIME 10:30 am - 11:30 am
VENUE CS Seminar Room, Y6405
6th Floor, Yellow Zone
Yeung Kin Man Academic Building
City University of Hong Kong
83 Tat Chee Avenue
Kowloon Tong

ABSTRACT

Big data clusters increasingly employ in-memory solutions to improve I/O performance. However, the routinely observed file popularity skew and load imbalance create hot spots in a cluster, which significantly degrade the benefits of in-memory caching. Common approaches to tame load imbalance include copying multiple replicas of hot files and creating parity chunks using storage codes. Yet, these techniques either suffer from high memory overhead due to cache redundancy or incur non-trivial encoding/decoding complexity. In this talk, I will present an effective approach to achieve load balancing without cache redundancy or encoding/decoding overhead. Our solution, which we call SP-Cache, selectively partitions files based on their popularity and evenly caches those partitions across the cluster. We develop an efficient algorithm to determine the optimal number of partitions for a hot file---too few partitions are incapable of mitigating hot spots, while too many are susceptible to stragglers. We have implemented SP-Cache in Alluxio, a popular in-memory distributed storage for data-intensive clusters. Real cloud deployment and trace-driven simulations show that, compared to the state-of-the-art solution, SP-Cache reduces the file access latency by up to 40% in both the mean and the tail, using 40% less memory.

BIOGRAPHY

Prof Wei Wang is currently an Assistant Professor in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology (HKUST). He is also affiliated with HKUST Big Data Institute. Wei received Ph.D. from the University of Toronto in 2015, and M.Eng. and B.Eng from Shanghai Jiao Tong University, all in the Department of Electrical and Computer Engineering. His research interests cover the broad area of distributed systems, with special emphasis on big data and machine learning, cloud computing, and computer networks in general. He is a recipient of the 2015 Chinese Government Award for Outstanding PhD Students Abroad and the Best Paper Finalist Award at the USENIX ICAC 2013. Wei was recently named as the Distinguished TPC member of IEEE INFOCOM 2018.

All are welcome!

In case of questions, please contact Dr Henry Xu at Tel: 3442 4840, E-mail: henry.xu@cityu.edu.hk, or visit the CS Departmental Seminar Web at http://www.cs.cityu.edu.hk/.