

ZHANG Hao (Dr.)

zhanghaohit@gmail.com | (65) 8341-3126

<https://zhanghaohit.github.io/>

Education

National University of Singapore (NUS), Singapore

Jul. 2012 – May 2017

Ph.D., Computer Science, School of Computing

- Research Interests: in-memory database systems, database system architectures, distributed systems, performance issues
- Dean's Graduate Research Excellence Award
- Research Achievement Award
- Topic: In-memory Large Scale Data Management and Processing

Harbin Institute of Technology (HIT), China

Sep. 2008 – Jul. 2012

Bachelor, Computer Science and Technology, School of Computer Science and Technology

- Cumulative Grade Point Average: 93.1/100; Ranking: 3th/174, Top 1.72%
- National scholarships and awards
- Thesis Title: Feature Extraction for Advertising Promotion

I-Shou University, Taiwan

Sep. 2010 – Jan. 2011

Exchange student, Department of Computer Science

Work/Internship Experience

Ph.D./Research Assistant

Mar. 2016 – Jul. 2017

Company: Department of Computer Science, National University of Singapore

Globally Addressable Memory: Towards Shared Memory Model on RDMA Network (Ph.D. Research)

- Proposed an efficient DSM framework to achieve a shared memory model on distributed systems based on generic RDMA networking.
- Designed an efficient cache coherence protocol based on PSO consistency model.
- Built a transaction engine, DHT and graph processing engine on top of it.

UVMM: User-space Virtual Memory Management for In-memory Databases (Ph.D. Research)

- Investigated various “Caching/Anti-Caching” approaches that are designed for use in both user-space and kernel-space, their tradeoffs in terms of performance and usability.
- Conducted an experimental study to analyze the components of “Caching/Anti-Caching” process, in terms of the overhead, eviction accuracy, I/O efficiency, CPU utilization, etc.
- Proposed a general “Anti-Caching” approach – UVMM, trying to combine the advantages of being in user-space and kernel-space.

MemepiC: Towards a Unified In-memory Big Data Management System (Ph.D. Research)

- Designed a unified in-memory data management system, which not only provides low-latency storage service, but also integrates in-memory data analytics functionality.
- Designed a RDMA-based messaging protocol, which combines both one-sided and two-sided RDMA verbs.
- Identified major types of syscalls in a typical in-memory system, and proposed detailed solutions on how to remove them from the critical path.

epiC: an Extensible and Scalable System for Processing Big Data

- Proposed an extensible system to tackle the Big Data's data variety challenge, by introducing a general Actor-like concurrent programming model, independent of the data processing models.
- Like Hadoop, programs written in this way can be automatically parallelized and the runtime system takes care of fault tolerance and inter-machine communications.

Quality Assurance Intern

Oct. 2011 – Apr. 2012

Company: Commerce Search Group, Baidu Inc., Beijing

- Quality assurance of advertisement recommendation systems
- Development of auto-test frameworks
- Recognized by project management for high productivity, excellent self-learning, problem solving, and good communication and coordination capabilities

Professional Skills

- Language Proficiency: English, Mandarin/Chinese
- Programming Skills: C, C++, Java, Python, PHP, JavaScript, etc.
- Software Skills: Linux, MySQL, Hadoop, HBase, Spark, Redis, Memcached, etc.

Professional Activities

- Reviewer of IEEE Transactions on Knowledge and Data Engineering (TKDE).
- Reviewer of Big Data Research Journal.
- Reviewer of World Wide Web Journal.
- In-memory Data Management Systems -- Challenges and Opportunities, talk at Renmin University, China.

Selected Publications ([Google Scholar Citation](#))

- **H. Zhang**, Q. Cai, G. Chen, W. Guo, R. Liu, B. C. Ooi, K.-L. Tan, Y. M. Teo. Globally Addressable Memory: Towards Shared Memory Model on RDMA Network.
- Q. Cai, **H. Zhang**, G. Chen, B. C. Ooi, K.-L. Tan, W.-F. Fai. MemepiC: Towards a Unified In-Memory Big Data Management System. (Under Review: Transactions on Big Data)
- A. Dinh, J. Wang, S. Wang, G. Chen, W.-N. Chin, Q. Lin, B. C. Ooi, P. Ruan, K.-L. Tan, Z. Xie, **H. Zhang**, M. Zhang. UStore: A Distributed Storage With Rich Semantics, arXiv preprint arXiv:1702.02799, 2017.
- **H. Zhang**, G. Chen, B. C. Ooi, K.-L. Tan, M. Zhang. In-Memory Big Data Management and Processing: A Survey. IEEE Transactions on Knowledge and Data Engineering (TKDE), Vol. 27, No. 7, 1920 – 1948, July 2015.
- K.-L. Tan, Q. Cai, B. C. Ooi, W.-F. Wong, C. Yao, **H. Zhang**. In-memory Databases – Challenges and Opportunities -- From Software and Hardware Perspectives. ACM SIGMOD Record, Special Issue on Visionary Ideas in Data Management, Vol. 44, No. 2, 35 – 40, June 2015.
- D. Loghin, B. M. Tudor, **H. Zhang**, B. C. Ooi, Y. M. Teo. A Performance Study of Big Data on Small Nodes. 41st Int'l Conference on Very Large Data Bases (VLDB), 762 – 773, 2015.
- **H. Zhang**, G. Chen, B. C. Ooi, W.-F. Wong, S. Wu, Y. Xia. "Anti-Caching"-based Elastic Memory Management for Big Data. 31st IEEE International Conference on Data Engineering (ICDE), 1268 – 1279, 2015.
- **H. Zhang**, B. M. Tudor, G. Chen, B. C. Ooi. Efficient In-memory Data Management: An Analysis. 40th Int'l Conference on Very Large Data Bases (VLDB), 833 – 836, 2014.