Chen Luo

Department of Computer Science, University of California, Irvine – Irvine – CA
949-372-8206 • cluo8@uci.edu

Research Interest

Database Storage Management, LSM-trees

Education

University of California, Irvine, CA *Sept. 2016–Dec. 2020*
Ph.D. in Computer Science. Supervisor: Michael J. Carey GPA: 4.00/4.00
Thesis: On Optimizing LSM-based Storage for Big Data Management Systems

Tsinghua University, China *Sept. 2013–July 2016*
M.Eng. in Software Engineering GPA: 94.1/100 (2 out of 136)

Tongji University, China *Sept. 2009–July 2013*
B.Eng. in Software Engineering GPA: 4.72/5 (2 out of 169)

Research Experience

Efficient Maintenance and Exploitation of LSM-based auxiliary structures [PVLDB 2019]
- Designed efficient point lookup methods to improve the applicability of LSM-based secondary indexes
- Proposed new maintenance strategies for LSM-based secondary indexes and range filters that substantially improved ingestion performance

Minimizing LSM-tree's Write Stalls via Novel Merge Scheduling [PVLDB 2020]
- Proposed a two-phase evaluation approach to evaluate write stalls of various LSM-tree designs
- Designed a novel greedy merge scheduler to minimize write stalls of LSM-trees
- Key insight is that LSM-trees can achieve both high write throughput and low percentile latencies

Adaptive Memory Management of LSM-trees
- Proposed a new memory management architecture for LSM-trees, including a new memory component structure, novel flush policies, and a memory tuner for the write memory and the buffer cache
- Extensive experiments on YCSB and TPC-C benchmarks demonstrated the effectiveness of the proposed techniques in reducing I/O costs

Internship Experience

Research Intern, Microsoft Research, Redmond *June 2019–Sep. 2019*
- Mentor: David Lomet
- Built a customized SSD controller to support batched writes and variable-size pages
- Collaborated with David Lomet on designing and optimizing a novel cleaning algorithm for log-structured stores that substantially reduced cleaning overheads

Research Intern, IBM Almaden Research Center *June 2017–Sep. 2017*
- Mentor: Pinar Tozun, Yuanyuan Tian
- Implemented a unified multi-zone indexing method for evolving data in HTAP systems **[EDBT 2019]**

Software Developer Intern, eBay China Development Center *July 2012–Mar. 2013*
- Participated in the development of eBay's web application framework
- Redeveloped the internal web traffic analytics system with MapReduce

Publications

[1] **Chen Luo**, Michael J. Carey. LSM-based storage techniques: a survey. *VLDB Journal*, 29 (1), pp. 393–418, 2020

[2] **Chen Luo**, Michael J. Carey. On performance stability in LSM-based storage systems. *PVLDB*, 13(4), pp. 449–462, 2019

[3] **Chen Luo**, Michael J. Carey. Efficient data ingestion and query processing for LSM-based storage systems. *PVLDB*, 12(5), pp. 531–543, 2019

[4] **Chen Luo**, Pinar Tozun, Yuanyuan Tian, Ronald Barber, Vijayshankar Raman, and Richard Sidle. Umzi: Unified Multi-Zone Indexing for Large-scale HTAP. *EDBT*, pp. 1–12, 2019

[5] Taewoo Kim, Alexander Behm, Michael Blow, Vinayak Borkar, Yingyi Bu, Michael J Carey, Murtadha Hubail, Shiva Jahangiri, Jianfeng Jia, Chen Li, **Chen Luo**, Ian Maxon, and Pouria Pirzadeh. Robust and efficient memory management in Apache AsterixDB. *Software: Practice and Experience*, 2019

[6] **Chen Luo**, Fei He, Fei Peng, Dong Yan, Dan Zhang and Xin Zhou. PSpec-SQL: Enabling Fine-Grained Control for Distributed Data Analytics. *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 2019

[7] **Chen Luo**, Fei He, and Carlo Ghezzi. Inferring software behavioral models with MapReduce. *Science of Computer Programming*, 145, pp. 13-36, 2017

[8] **Chen Luo**, Michael J. Carey. Breaking down memory walls: adaptive memory management for LSM-based storage systems. *Submitted to VLDB 2021*. Preprint: `https://arxiv.org/abs/2004.10360`

[9] David B. Lomet, and **Chen Luo**. Efficiently reclaiming space in a log structured store. *Submitted to SoCC 2020*. Preprint: `https://arxiv.org/abs/2005.00044`

[10] Jae Young Do, **Chen Luo**, and David B. Lomet. Programming an SSD controller to support batched writes for variable-size pages. *Under Preparation*