Chen Luo

San Mateo – CA • 949-372-8206 • cluo8@uci.edu • luochen01.github.io

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 2019]

- 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 for LSM-trees [PVLDB 2020]

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

Industrial Experience

Software Engineer, Snowflake Inc. *Jan. 2021–Present*

- Manager: Leonidas Galanis
- Working on extending Snowflake to support OLTP workloads

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 **[ICDE 2021]**
- 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
- Designed and implemented a unified multi-zone indexing method for evolving data in HTAP systems **[EDBT 2019]**

Selected Publications

[1] **Chen Luo**, Michael J. Carey. Breaking down memory walls: adaptive memory management for LSM-based storage systems. *PVLDB*, 14(3), pp. 241–254, 2021

[2] David B. Lomet, **Chen Luo**. Efficiently reclaiming space in a log structured store. *ICDE*, 2021

[3] Jae Young Do, **Chen Luo**, and David B. Lomet. Programming an SSD controller to support batched writes for variable-size pages. *ICDE*, 2021

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

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

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

[7] **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