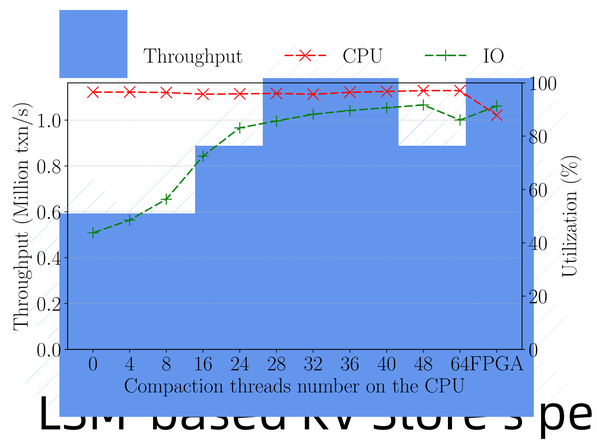
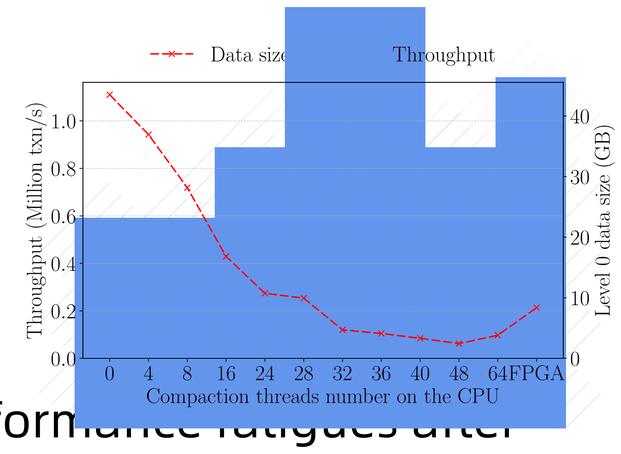
FPGA-Accelerated Compactions for LSM-based Key-Value Store

Teng Zhang, Jianying Wang, Xuntao Cheng, Hao Xu, Nanlong Yu, Gui Huang, Tieying Zhang,
Dengcheng He, Feifei Li, Wei Cao, Zhongdong Huang and Jianling Sun
Alibaba Group
Alibaba-Zhejiang University Joint Institute of Frontier Technologies, Zhejiang University

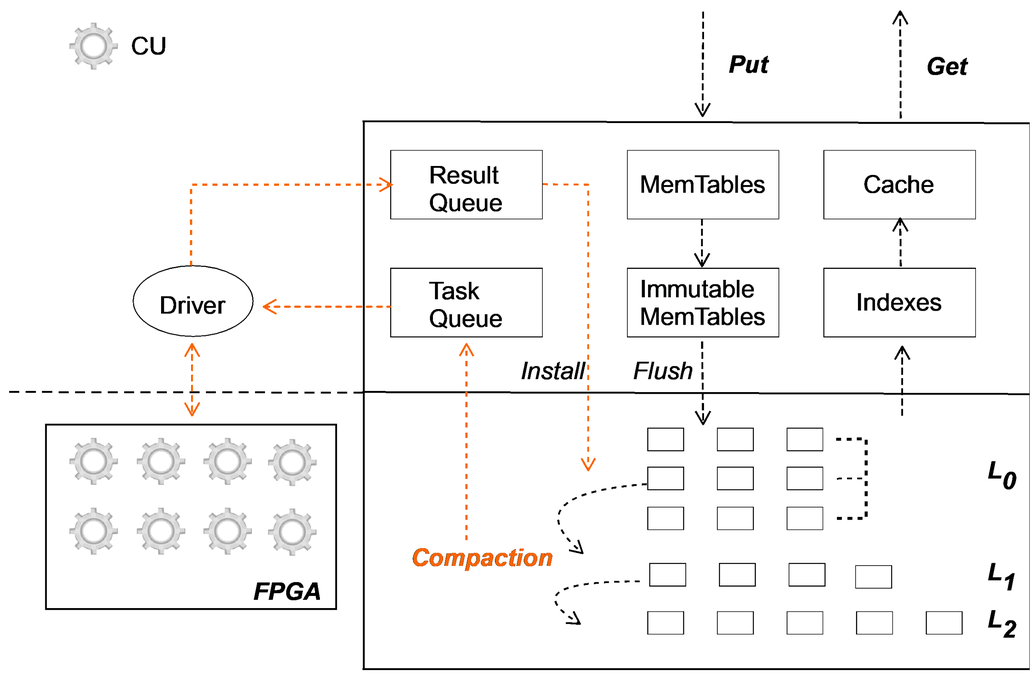
Problems





- LSM-based KV Store's performance fatigues after serving WPI workloads for a long time.
- Shattered L_0: Single lookup may incur multiple I/Os.
- Shifting Bottleneck: CPU resource contention for short KVs.

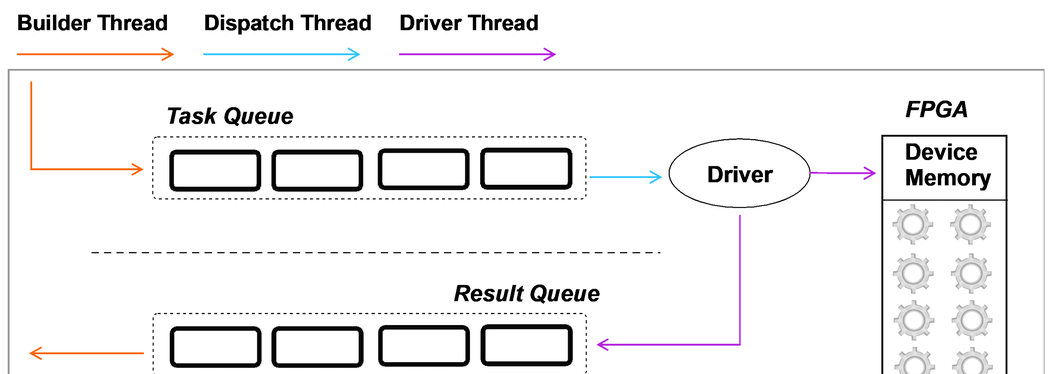
FPGA Offloading



Compaction Unit

- 4-way Decoder
- Controller to coordinate the pace of each module
- Workload-dependent merging selectivity
- Analytical Model: within 13% error

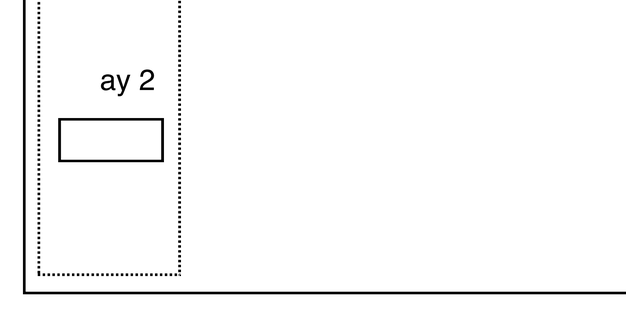
Compaction Tasks Management



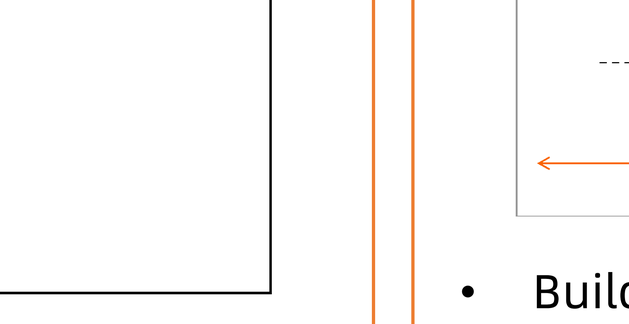
- Builder Thread (construct tasks by partitioning extents into similar size)
- Dispatch Thread (round-robin)
- Driver Thread (redo failed tasks, 0.03% in practice)
- Separate Path for Data and Instructions Transfer

Evaluations

2~5x Speedup comparing to single CPU thread

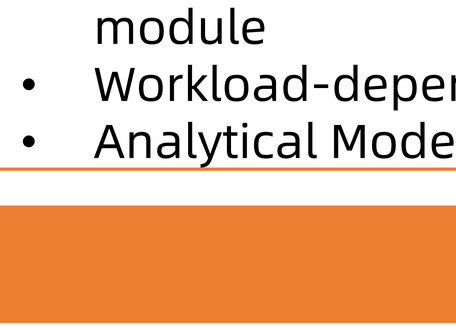


Key = 8B

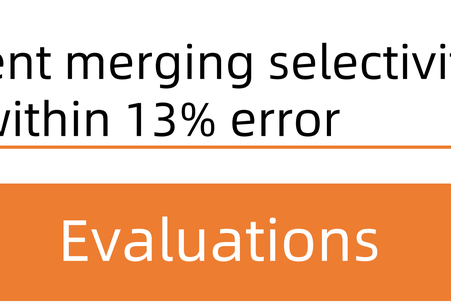


Key = 64B

DBBench workloads, 23.3% throughput improvement, 31.7% efficiency (transactions per watt) improvement



TPS



CPU

IO

Industrial applications

- E-commerce
- Advertisement
- Instant Messaging

Ongoing work

- Individualized Benchmark Service
- Auto Tune for DBMS
- Distributed RDBMS
- ML for hot cold data separation
- Scheduling of compactions

Email

X-Engine/DAMO