

# X-Engine: An Optimized Storage Engine for Large-scale E-commerce Transaction Processing

Gui Huang, Xuntao Cheng, Jianying Wang, Yujie Wang, Dengcheng He, Tieying Zhang, Feifei Li, Sheng Wang, Wei Cao, and Qiang Li  
Alibaba Group



Email

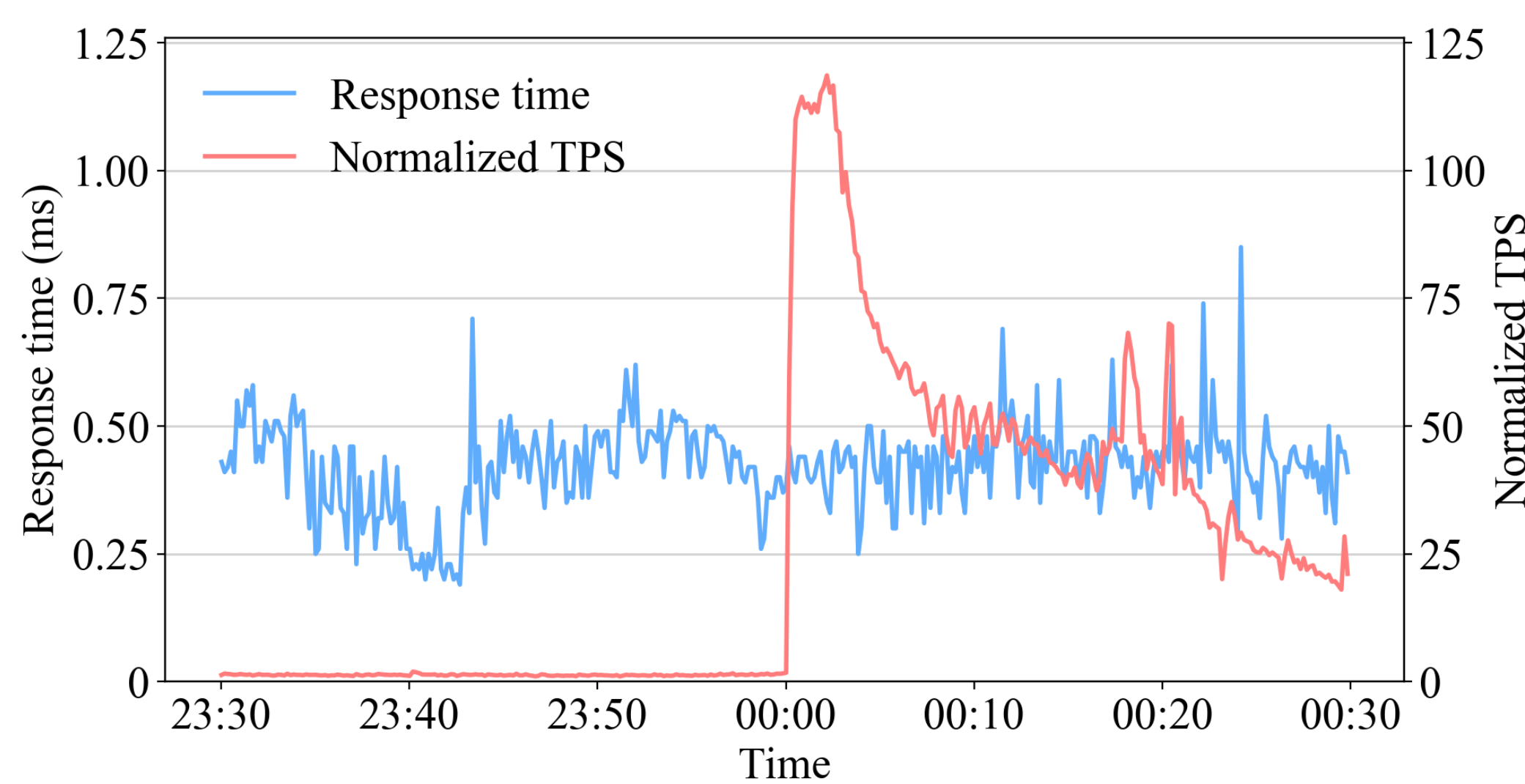


DingTalk



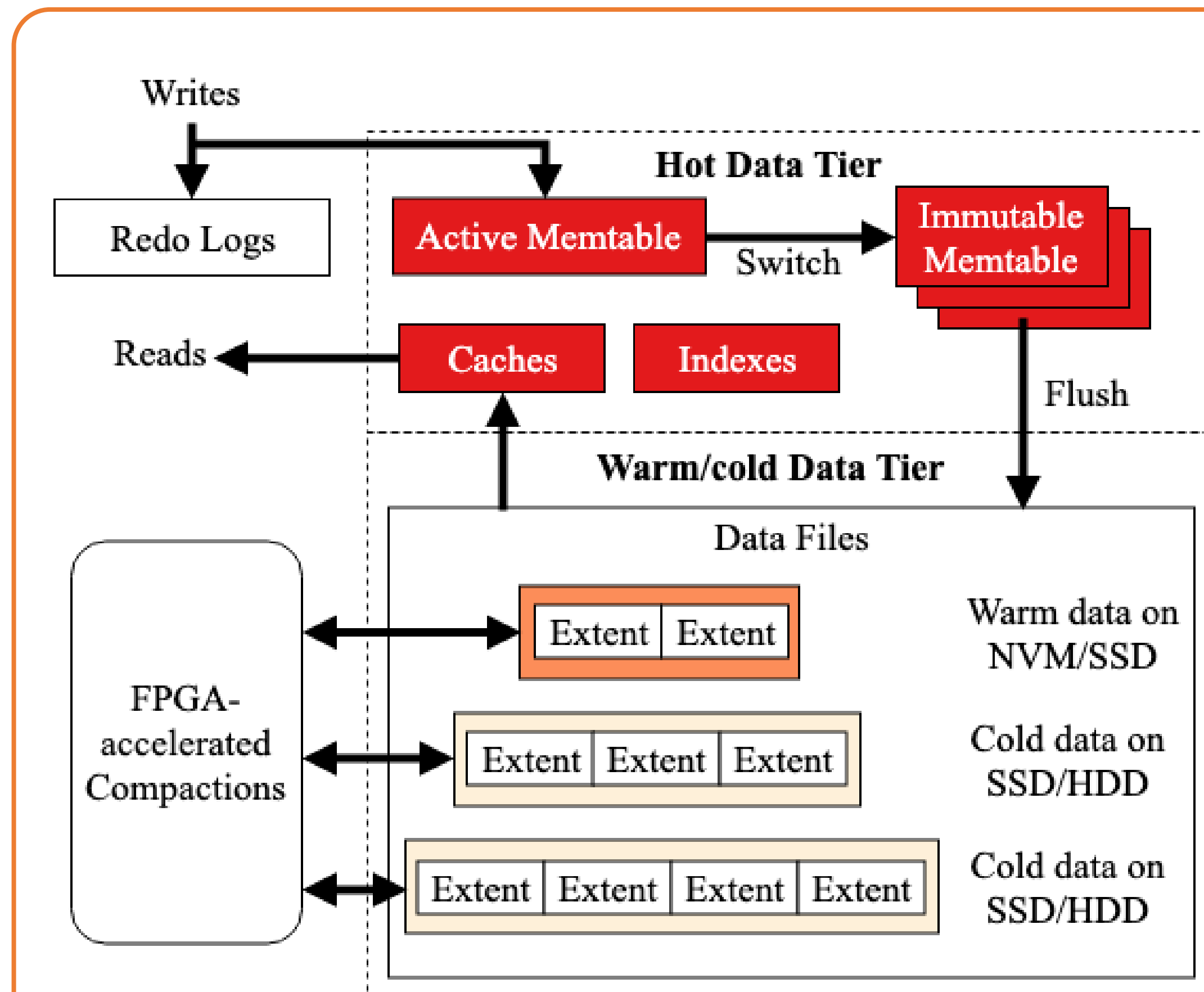
Paper

## Problems



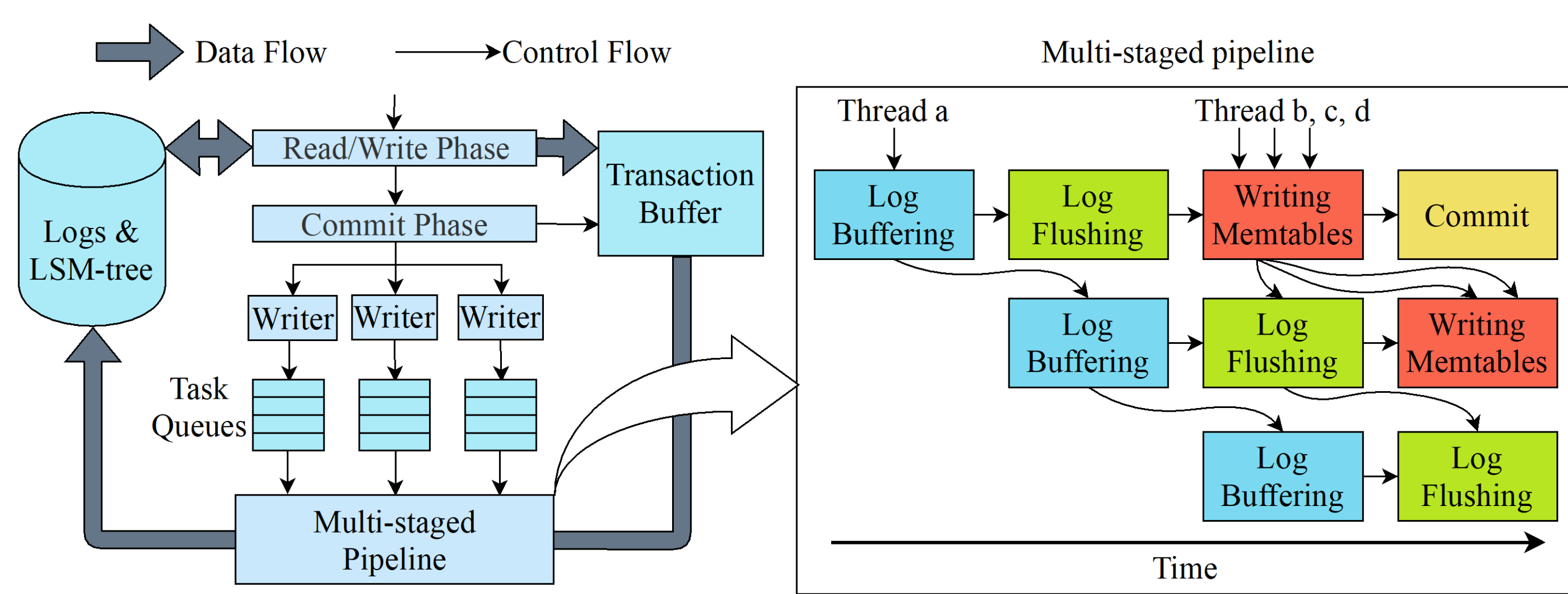
- 122 x TPS spike on the start of the Singles' Day Shopping Festival
- R/W ratio drops to 3:1 from 10:1. More inserts/updates than usual
- Business-critical response time

## Architecture



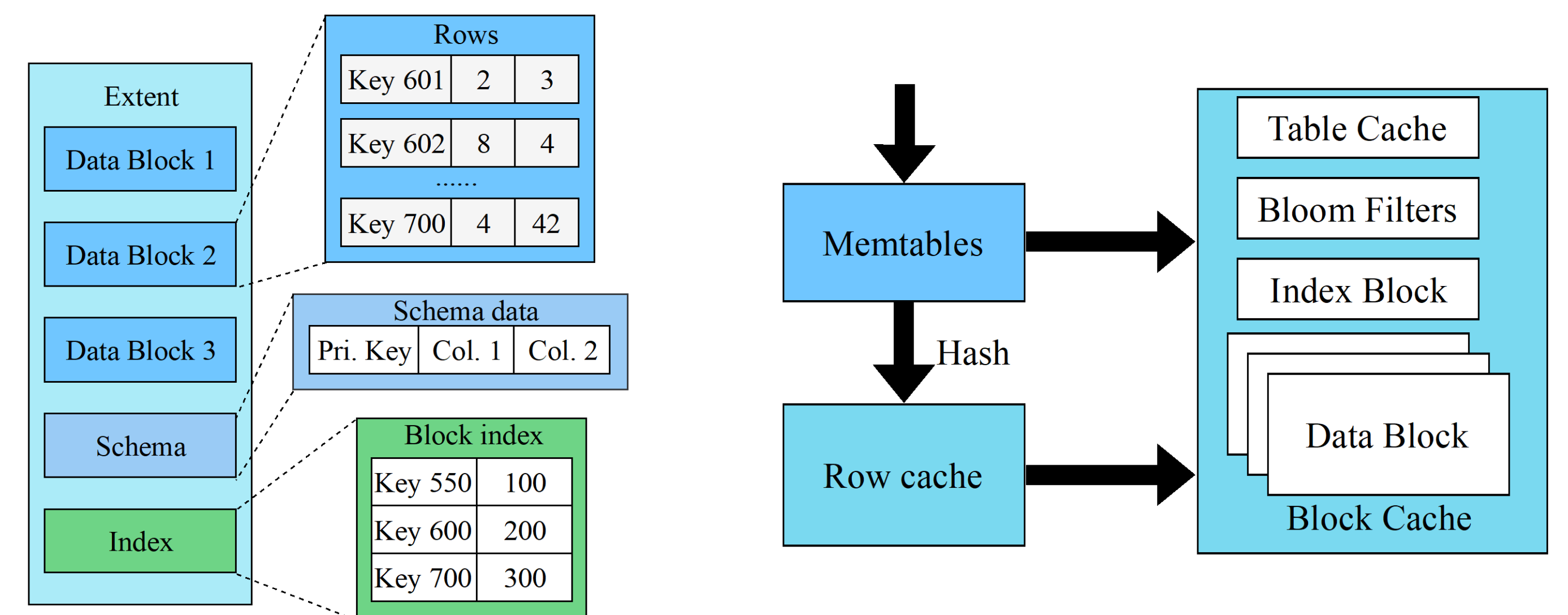
- Tiered storage with the LSM-tree suits the data locality in e-commerce transactions
- Caches compensate the read performance
- Reduced extent sizes increase data reusing during compactions
- Using CPU for transactions, and FPGA for compactions

## Optimizing the write path



- Asynchronously buffering changes
- Tuned number of writer threads
- Separating disk I/Os and memory writes
- Fine tuning of the thread-level parallelism
- Using fewer threads to saturate the I/O

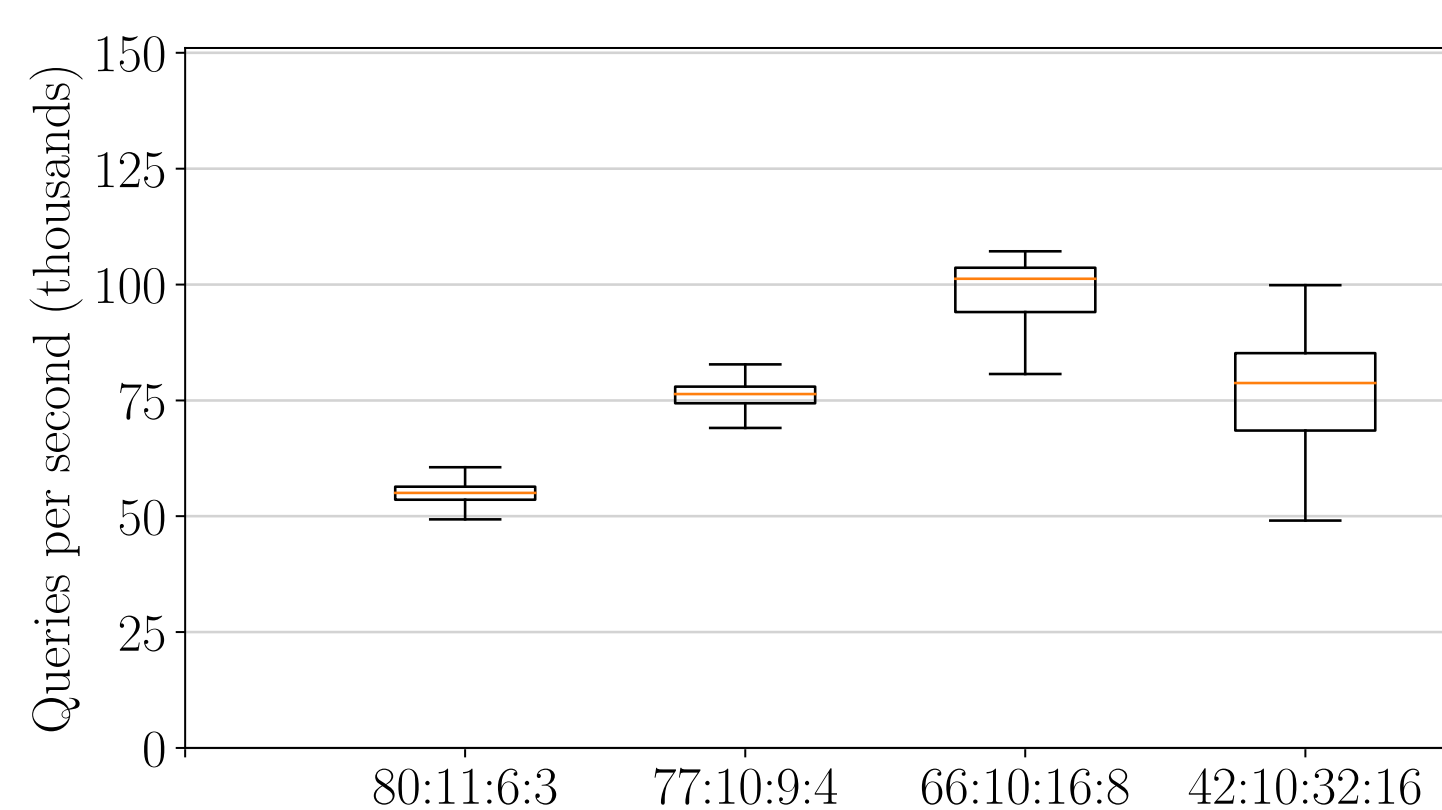
## Optimizing the read path



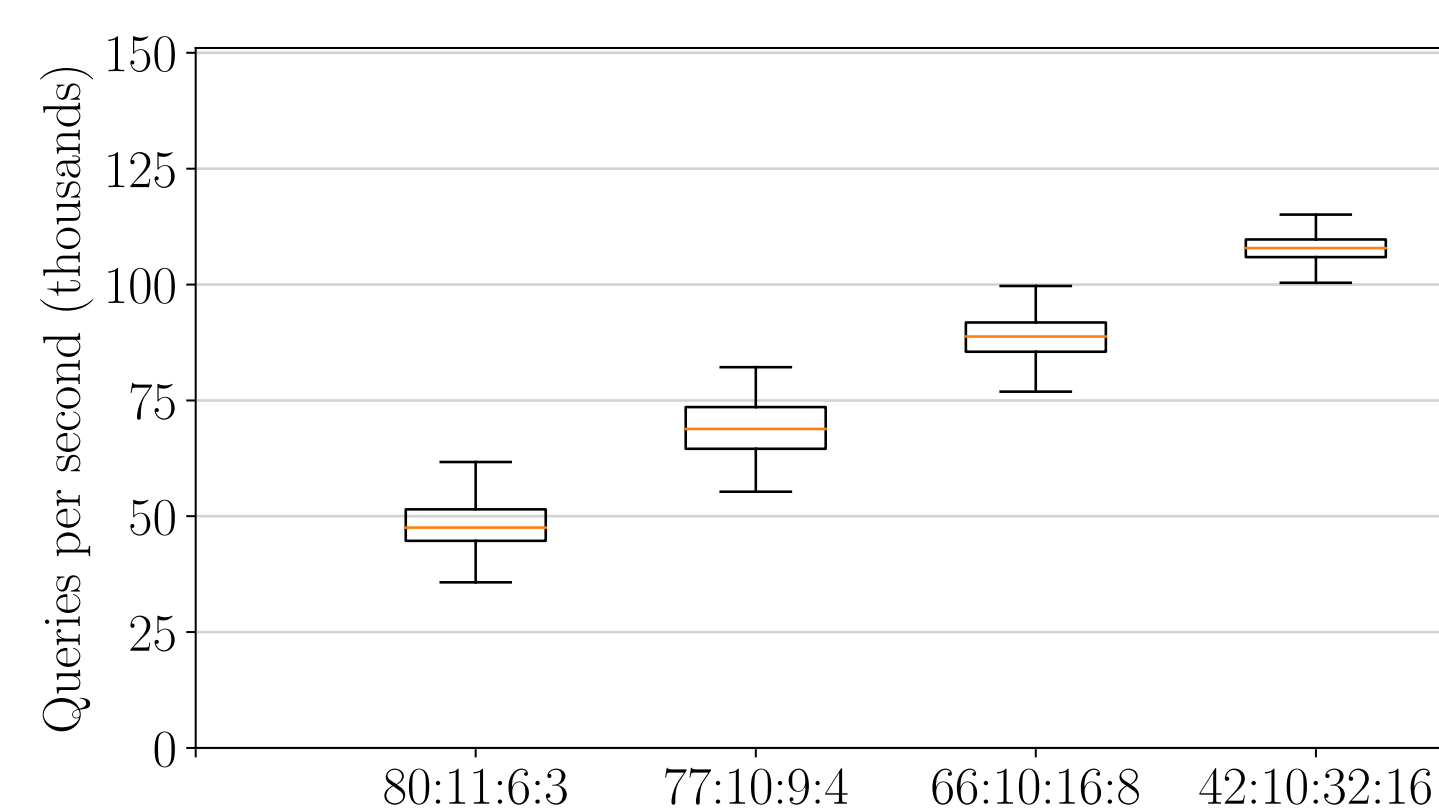
- Small-size extents with schema and indexes
- Multi-version skiplists for the memtable
- Row caches for point queries
- Block caches for range queries
- Copy-on-write multi-version metadata index

## Evaluations

Workload: mixing point/range lookups, updates, inserts

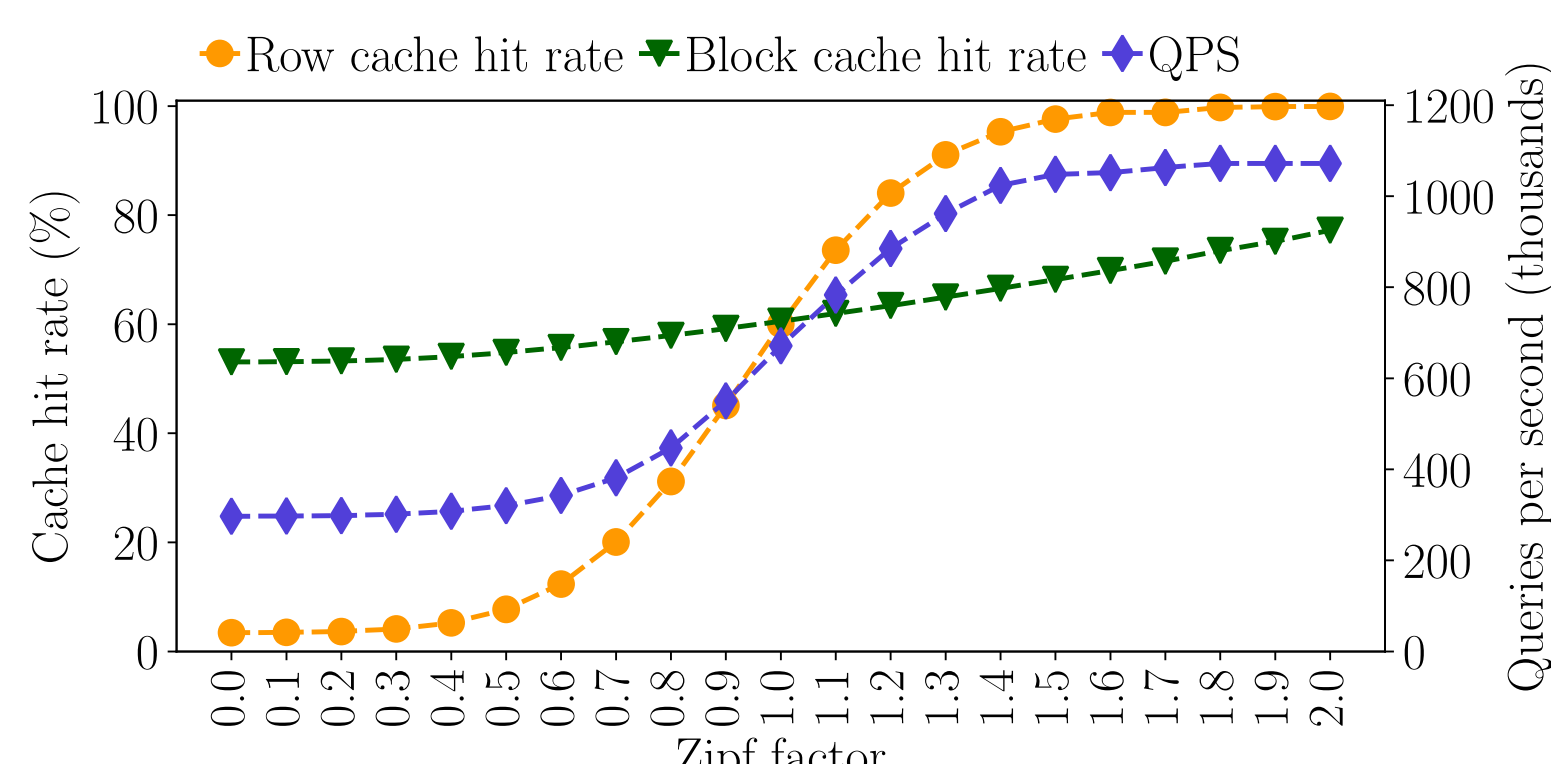


MySQL(InnoDB)

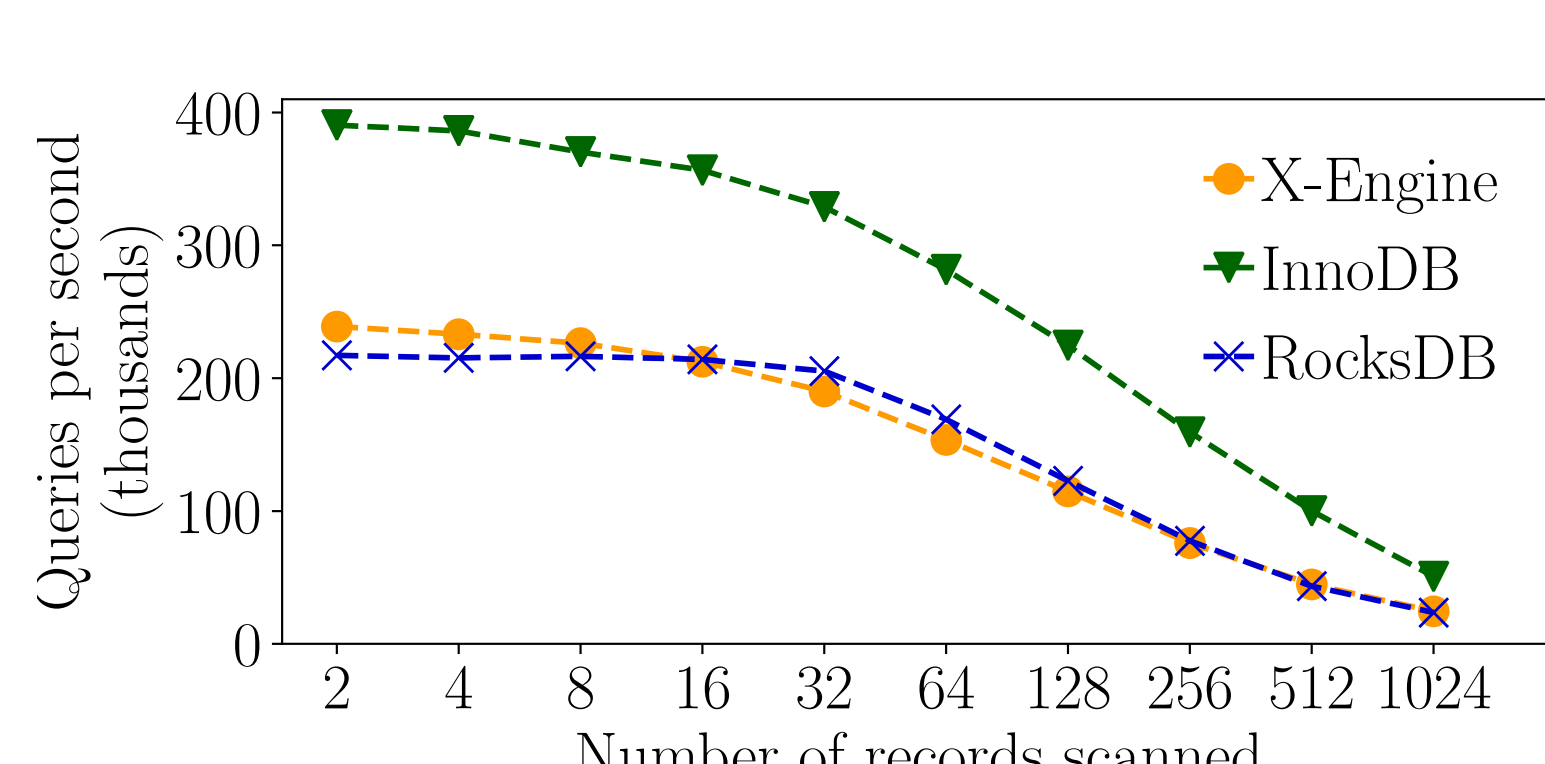


MySQL(X-Engine)

Workload: skewed point queries, range queries



Cache efficiencies



Range query efficiency

## Industrial applications



- Payment
- Logistics
- Trade history archives
- Advertisements
- Messages
- Cloud-native DB administration

## Ongoing work

- Distributed read/write split
- Hybrid transactional/analytical processing
- ML for data temperature detection
- ML for record prefetching & caching
- Scheduling of compactions
- Emerging hardware
- Security