







Leaper: A Learned Prefetcher for Cache Invalidation in LSM-tree based Storage Engines

Lei Yang¹, Hong Wu², Tieying Zhang², Xuntao Cheng², Feifei Li², Lei Zou¹,
Yujie Wang², Rongyao Chen², Jianying Wang², and Gui Huang²

Peking University¹ Alibaba Group²

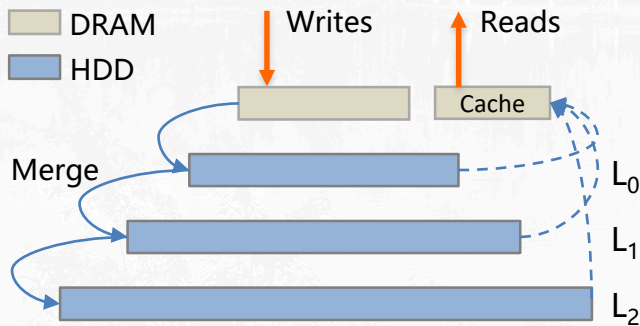


Outline

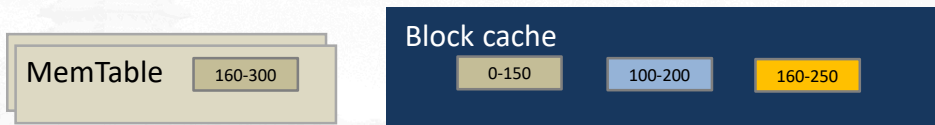
- 01  Background
- 02  Cache Invalidation Problem
- 03  Our Approach
- 04  Experimental Results

Background

- LSM-tree
 - Write
 - Merge operations
 - Flush
 - Compaction
 - Read
 - Cache

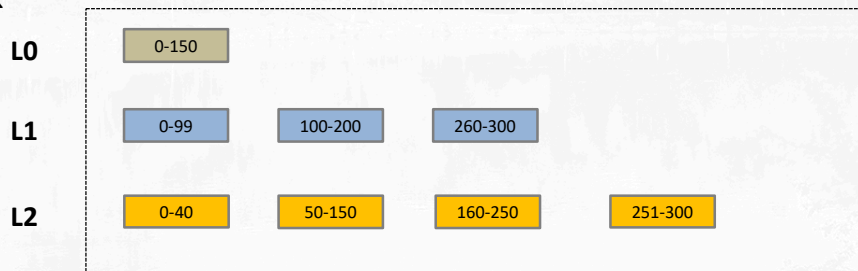


Cache Invalidation Problem

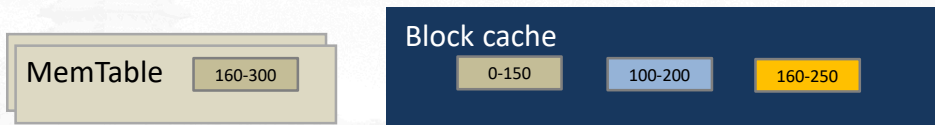


Memory

Disk

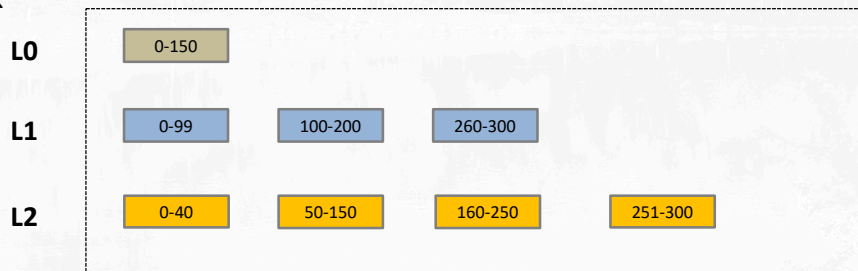


Cache Invalidation Problem



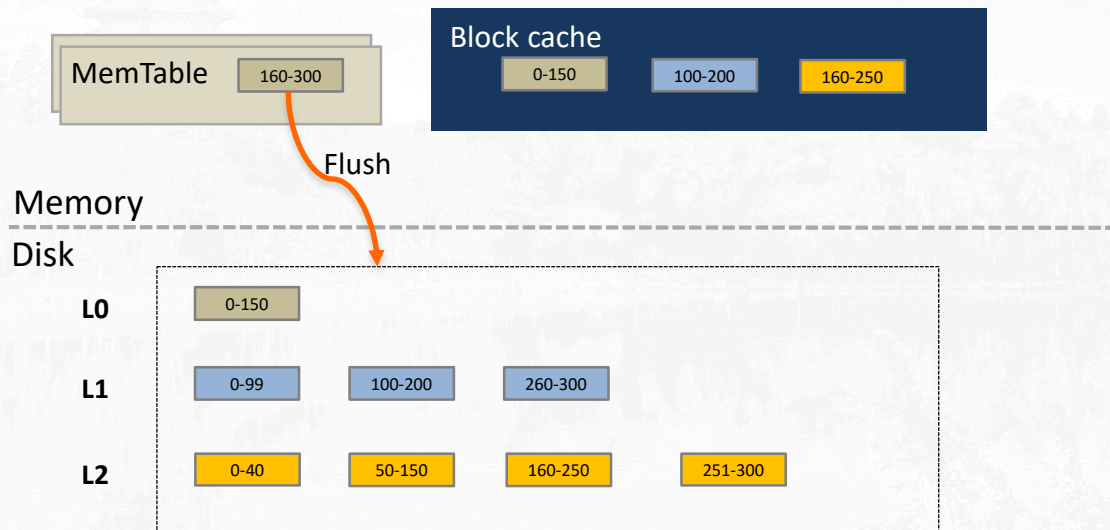
Memory

Disk

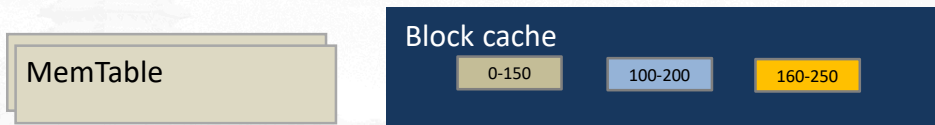


Read Request	Result
Key 285	✓
Key 205	✓
Key 155	✓
Key 10	✓

Cache Invalidation Problem

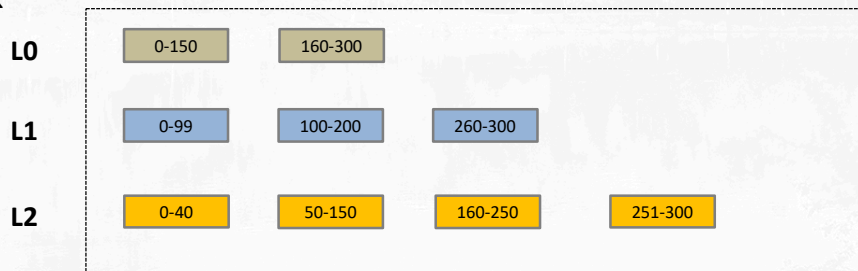


Cache Invalidation Problem

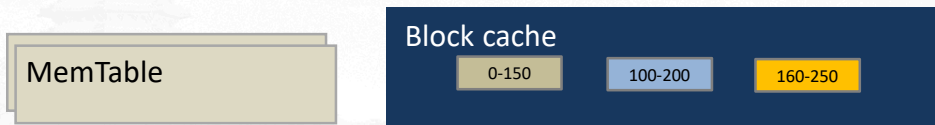


Memory

Disk

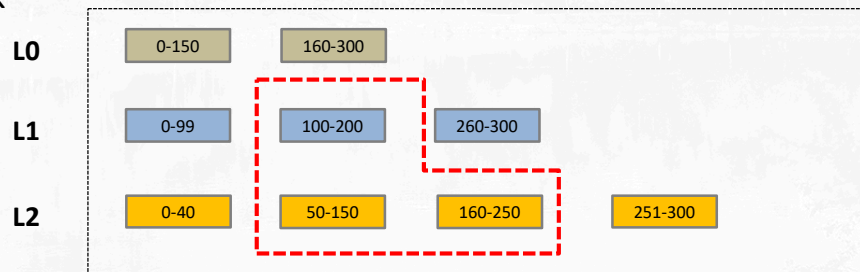


Cache Invalidation Problem



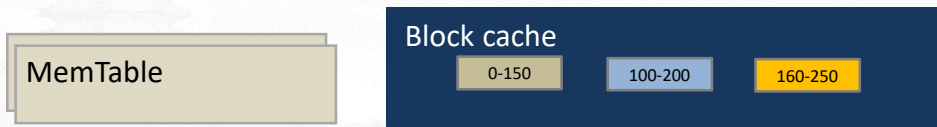
Memory

Disk



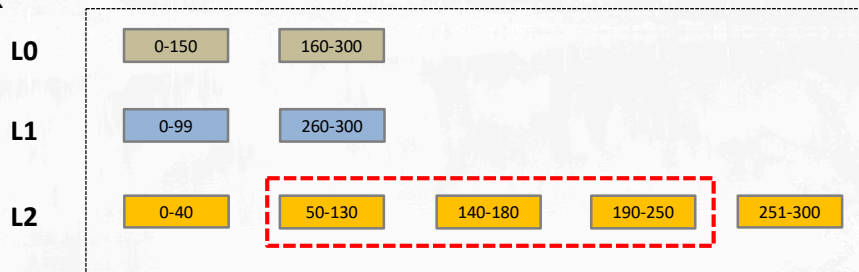
Compaction

Cache Invalidation Problem

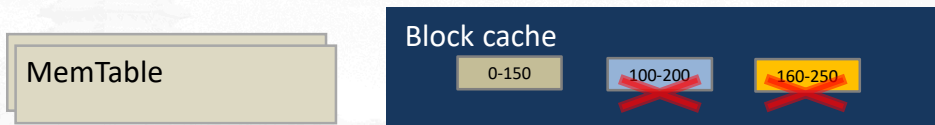


Memory

Disk

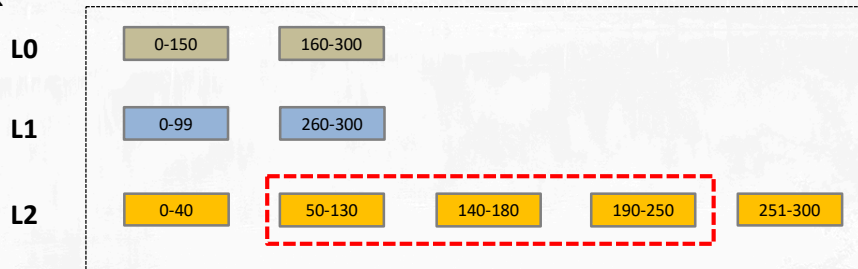


Cache Invalidation Problem

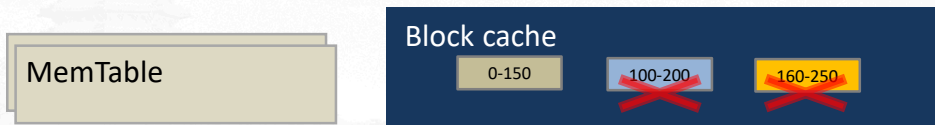


Memory

Disk

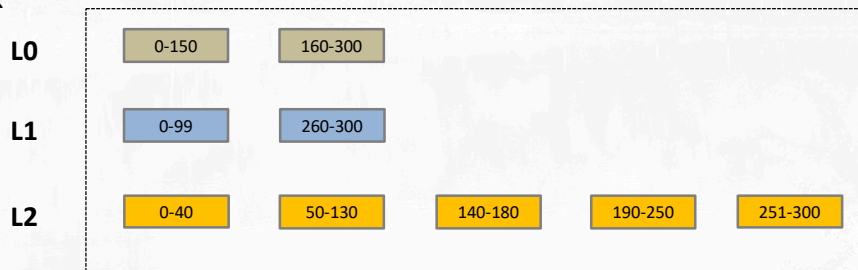


Cache Invalidation Problem



Memory

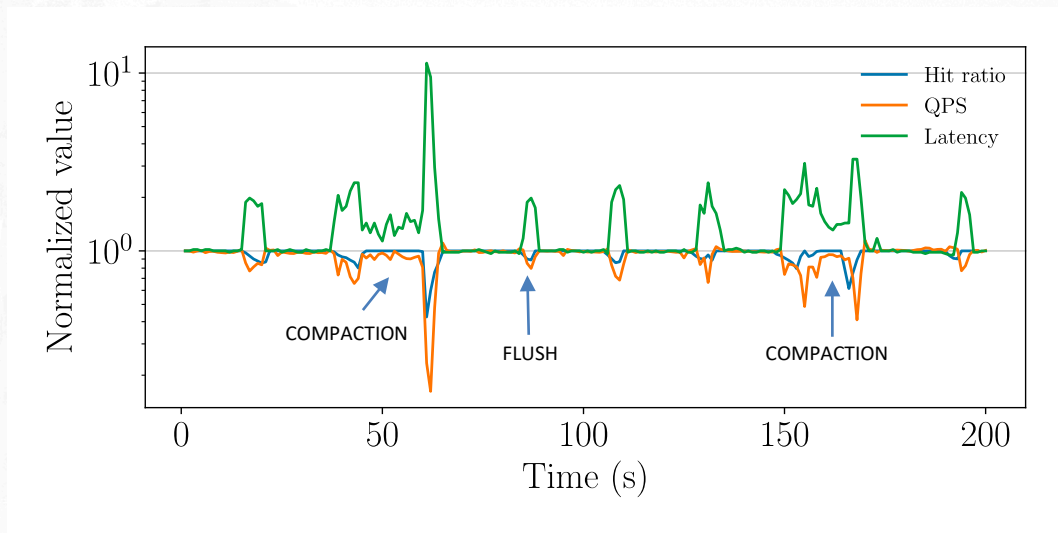
Disk



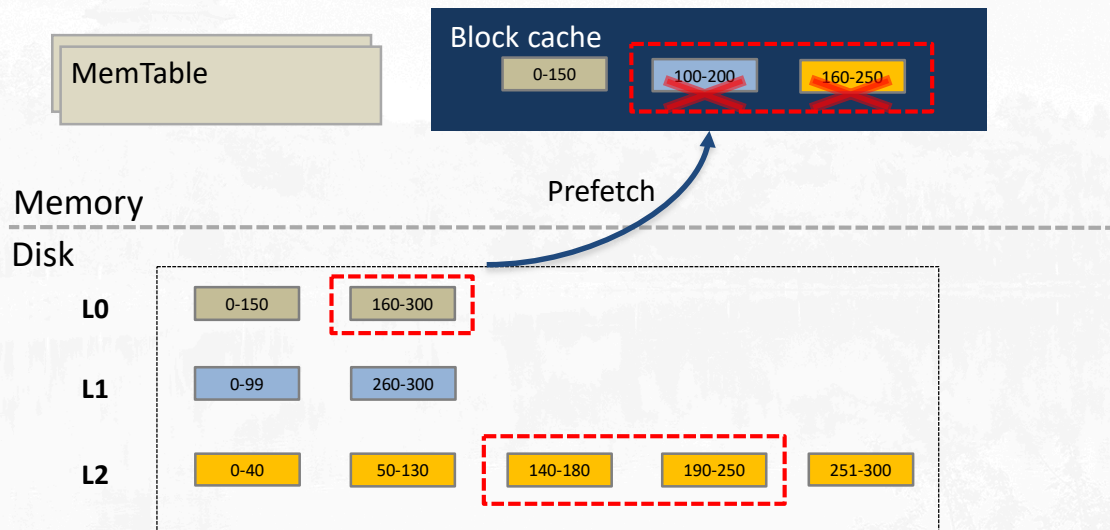
Read Request	Result
Key 285	X
Key 205	X
Key 155	X
Key 10	✓

Cache Invalidation Problem

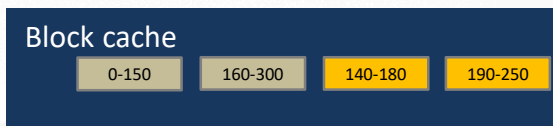
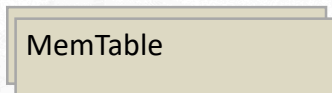
Performance Jitter



Intuition

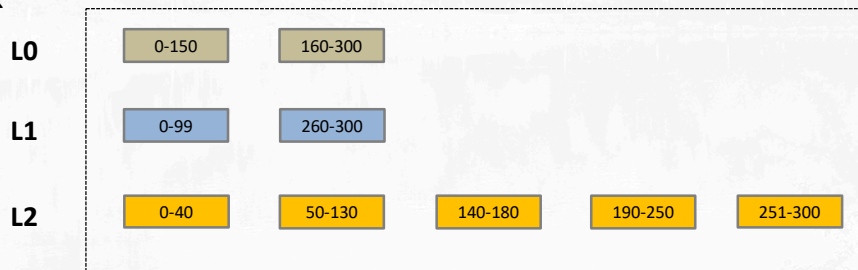


Intuition



Memory

Disk



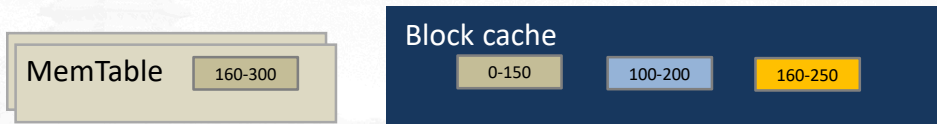
Read Request	Result
Key 285	✓
Key 205	✓
Key 155	✓
Key 10	✓



Learned Prefetcher

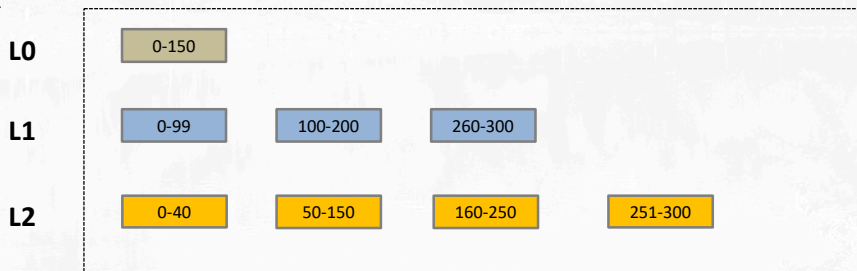
- How to Learn?
- How to Prefetch?

How to Learn



Memory

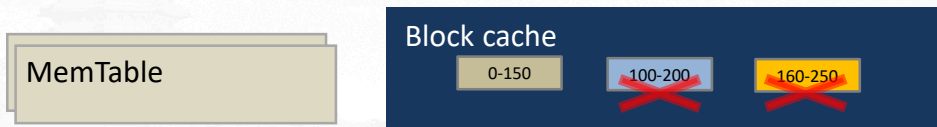
Disk



Read Request	Result
Key 285	✓
Key 205	✓
Key 155	✓
Key 10	✓

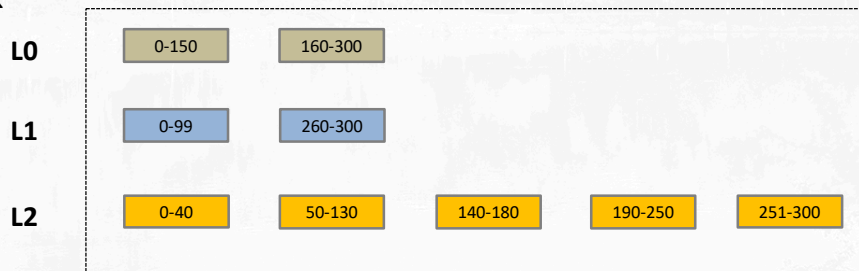
Blocks	Access(es)
160-300	1
160-250	1
100-200	1
0-150	1

How to Learn



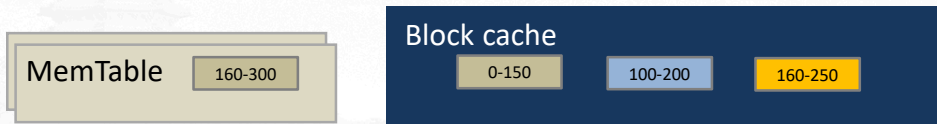
Memory

Disk



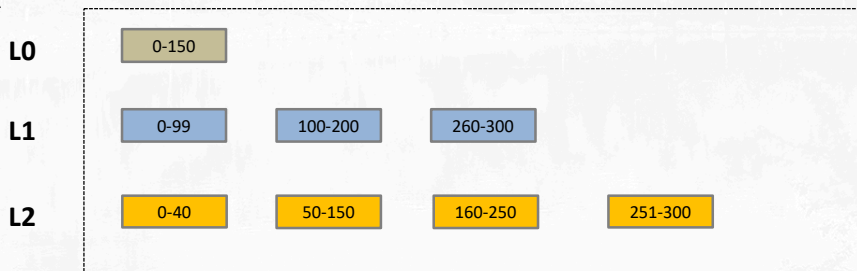
Blocks	Access(es)
160-300	1
160-250	1
100-200	1
0-150	1

How to Learn



Memory

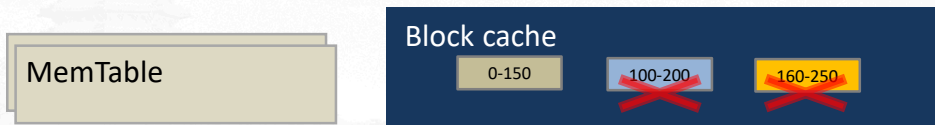
Disk



Read Request	Result
Key 285	✓
Key 205	✓
Key 155	✓
Key 10	✓

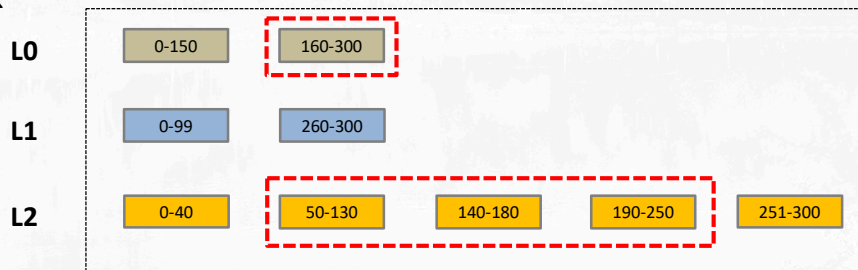
Key Ranges	Access(es)
[0, 50)	1
[150, 200)	1
[200, 250)	1
[250, 300)	1

How to Learn



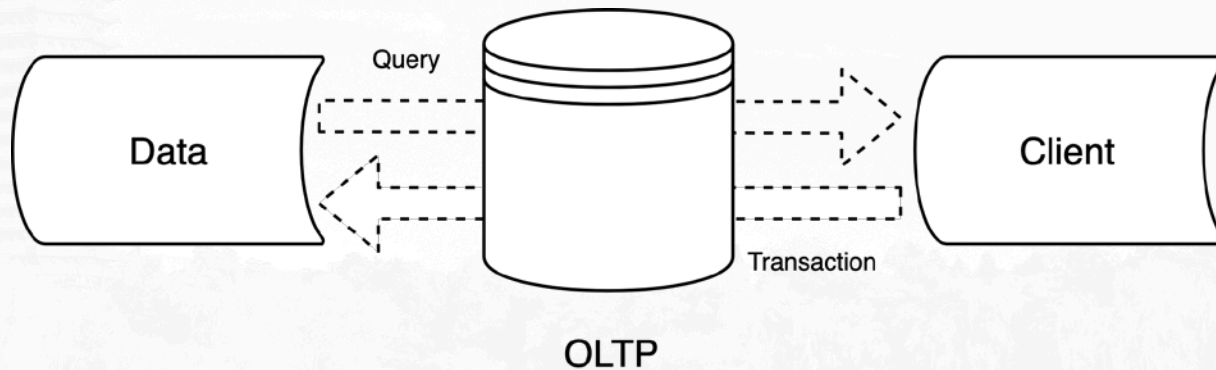
Memory

Disk

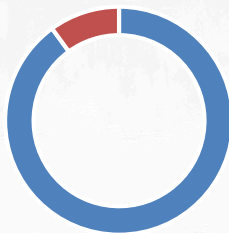


Key Ranges	Access(es)
[0, 50)	1
[150, 200)	1
[200, 250)	1
[250, 300)	1

How to Prefetch

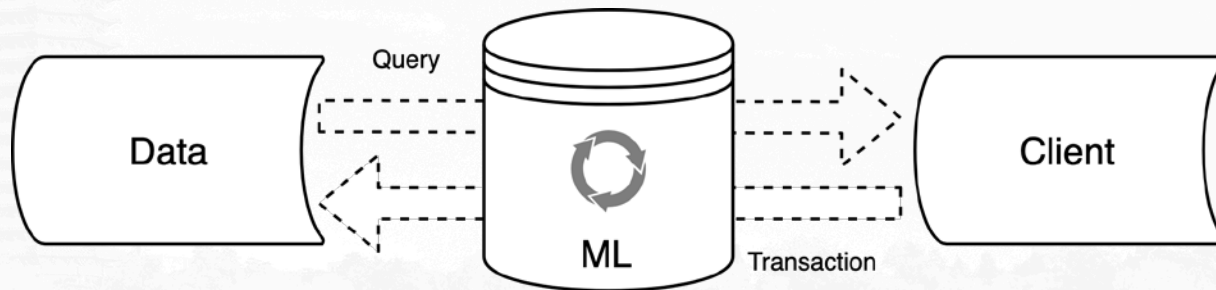


Hardware resource utilization



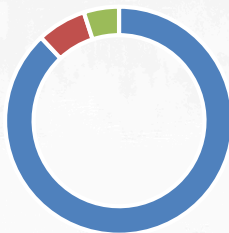
■ Used ■ Not used

How to Prefetch



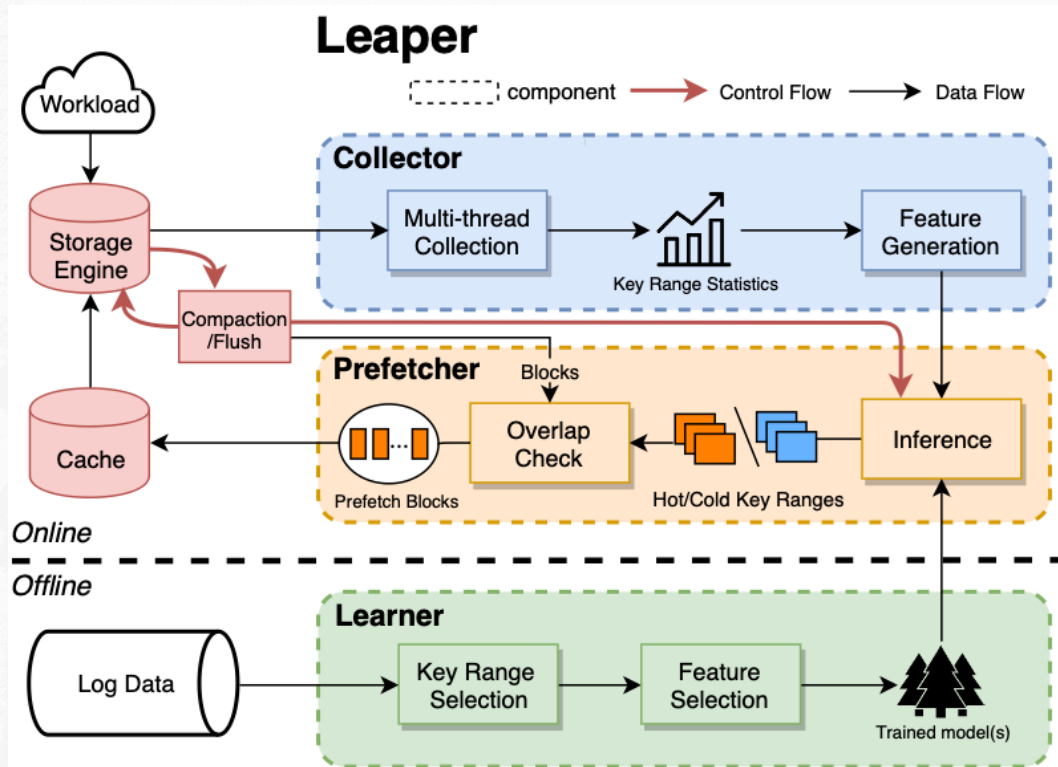
OLTP

Hardware resource utilization



■ DB used ■ Not used ■ ML used

System Overview

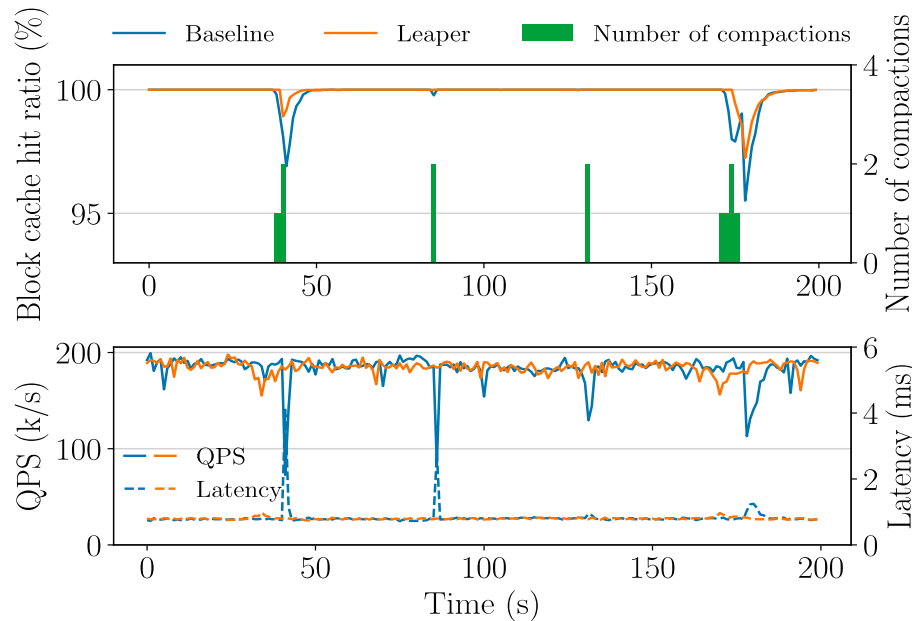


Experimental Setup

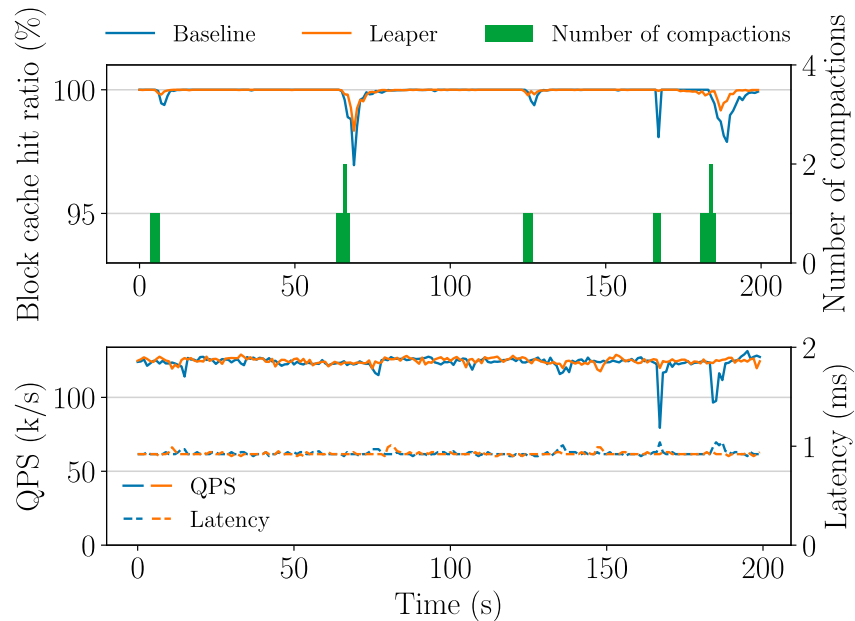
- Testbed
 - 2x 24-core Intel Xeon Platinum 8163 CPUs
 - 512G Samsung DDR4-2666 DRAM
 - 2x Intel SSDs
 - X-Engine with MySQL 5.7
- Workload
 - E-commerce workload from Tmall
 - Instant messaging workload from DingTalk
 - Synthetic workloads
- Baseline
 - Incremental Warmup Algorithm

Type	Lookups	Updates	Inserts	R/W ratio	Table size	Skewness
E-commerce	85%	10%	5%	6:1	10m	0.3
Instant messaging	40%	35%	25%	2:3	8m	0.9
Synthetic	75%	20%	5%	3:1	20m	0.5

Experimental Results



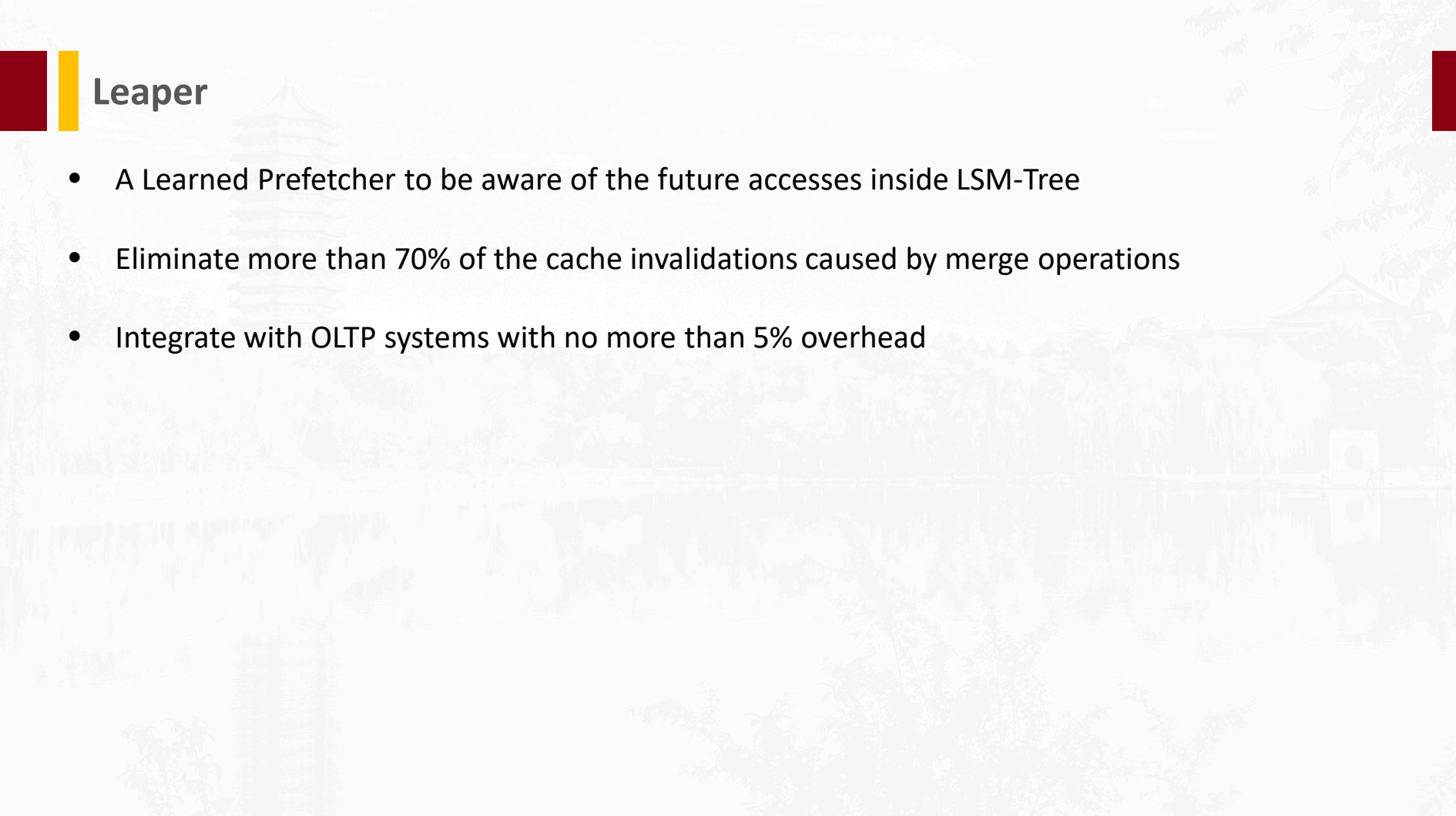
E-commerce



Instant messaging



Leaper

- A Learned Prefetcher to be aware of the future accesses inside LSM-Tree
 - Eliminate more than 70% of the cache invalidations caused by merge operations
 - Integrate with OLTP systems with no more than 5% overhead
- 

Thanks!

Q & A



北京大学
PEKING UNIVERSITY

