# Hailstorm

## Disaggregated Compute and Storage for Distributed LSM-based Databases

**Laurent Bindschaedler**, Ashvin Goel, Willy Zwaenepoel

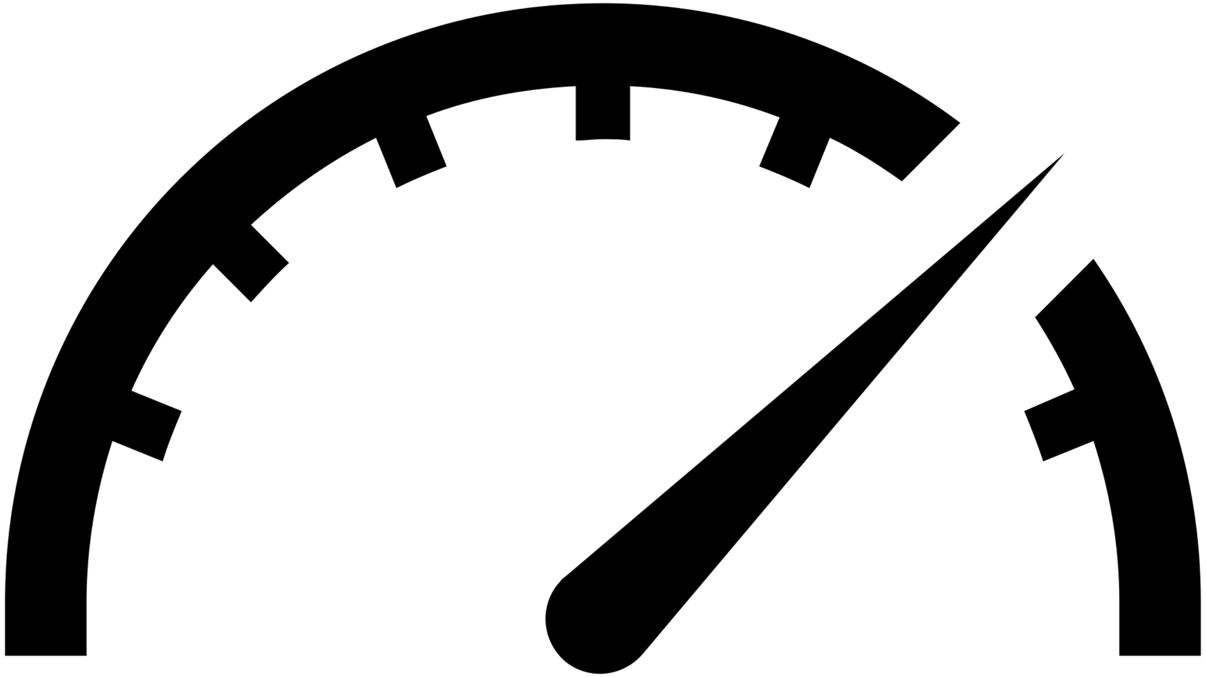
EPFL

UNIVERSITY OF TORONTO

THE UNIVERSITY OF SYDNEY

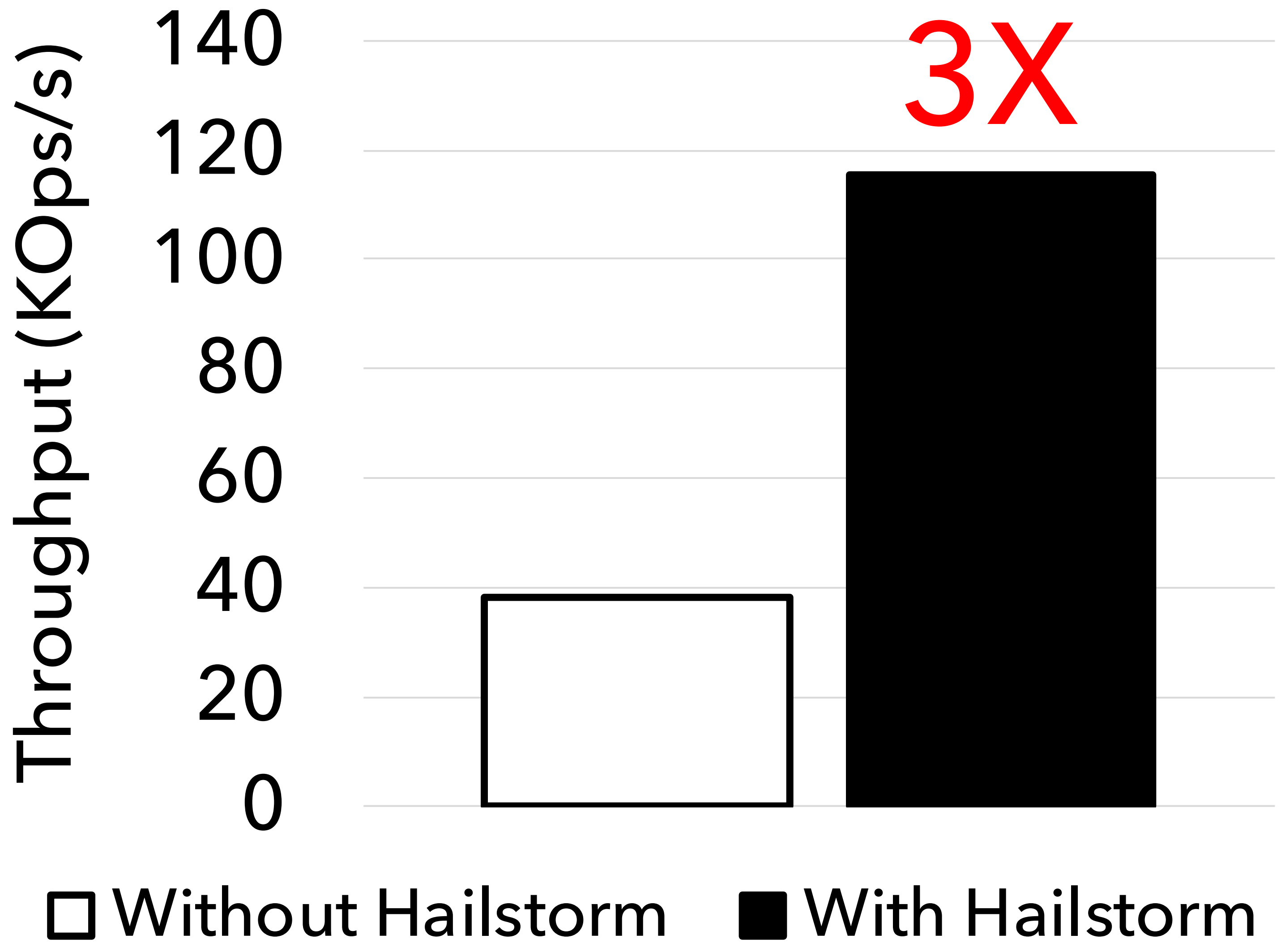
# Hailstorm Improves Compute and Storage Load Balance in Distributed Databases

⇒ higher **throughput** and better **resource utilization**



Video of Presenter

# Preview – Production Trace on MongoDB



Video of
Presenter

# Outline

1. Background
   A. Distributed Databases
   B. Load Imbalance in Distributed Databases
   C. Shard Rebalancing
2. Hailstorm Architecture
3. Evaluation
4. Conclusion

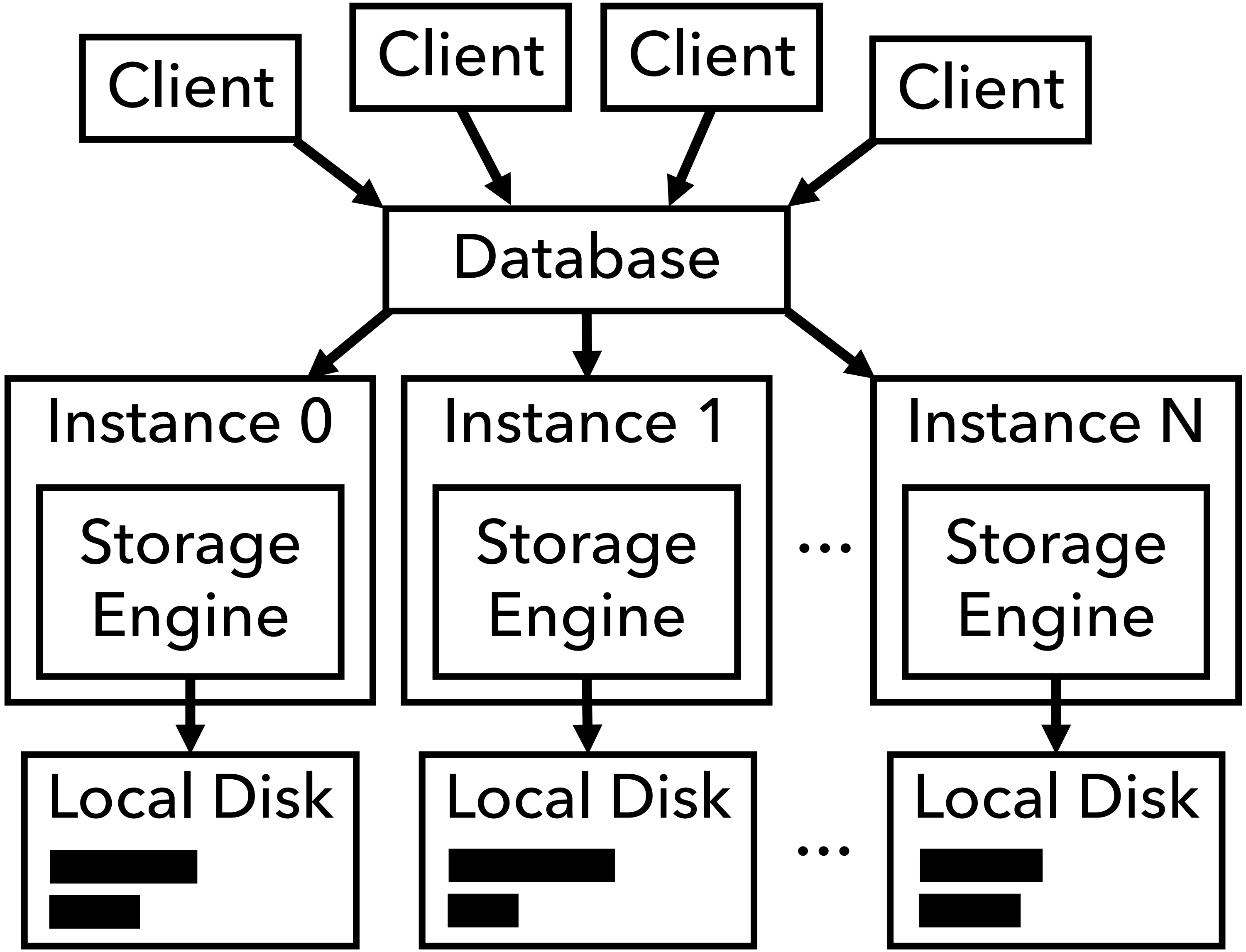
Video of Presenter

# Outline

1. Background
   - A. Distributed Databases
   - B. Load Imbalance in Distributed Databases
   - C. Shard Rebalancing
2. Hailstorm Architecture
3. Evaluation
4. Conclusion

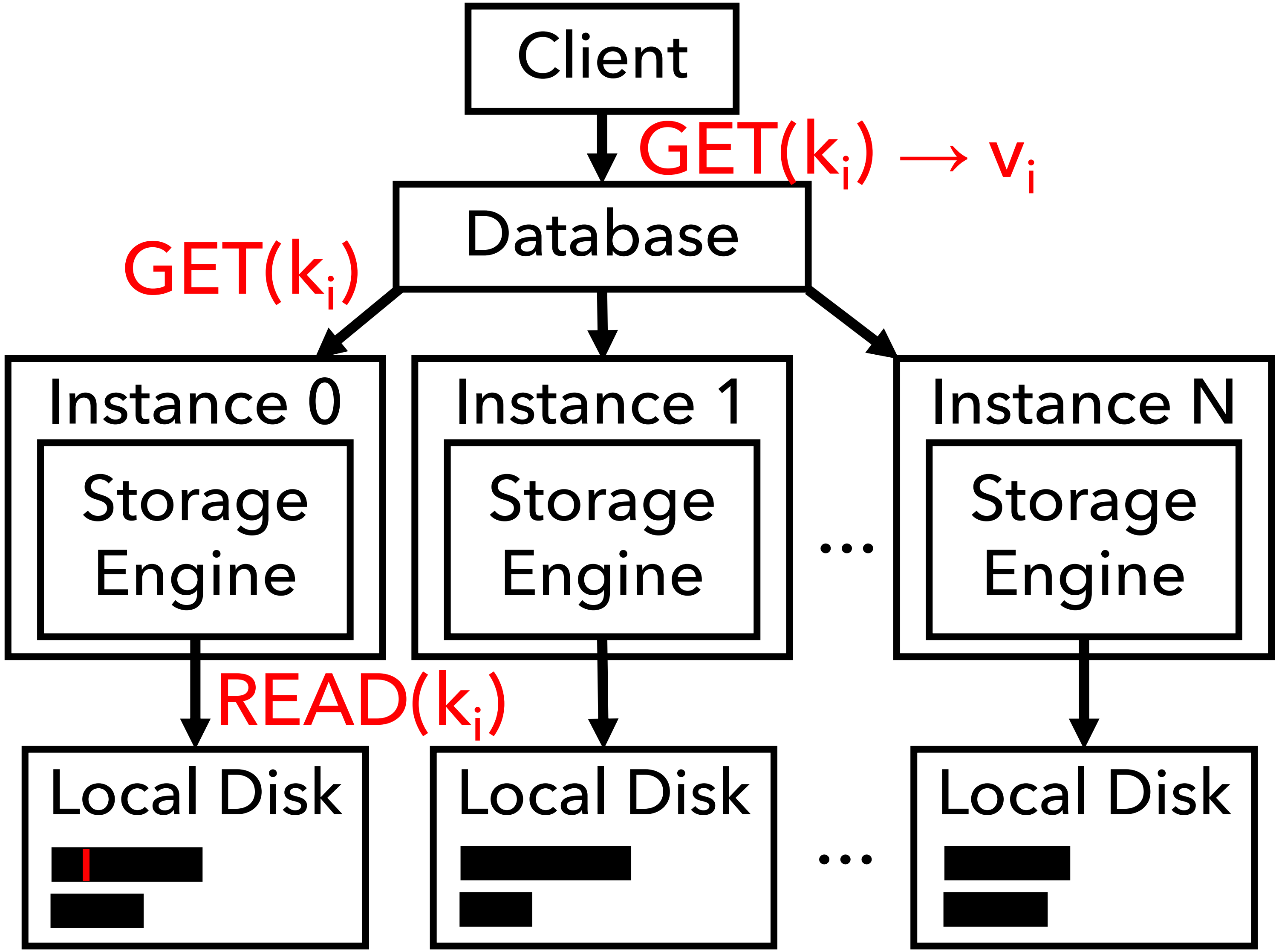
Video of
Presenter

# Background – Distributed Databases



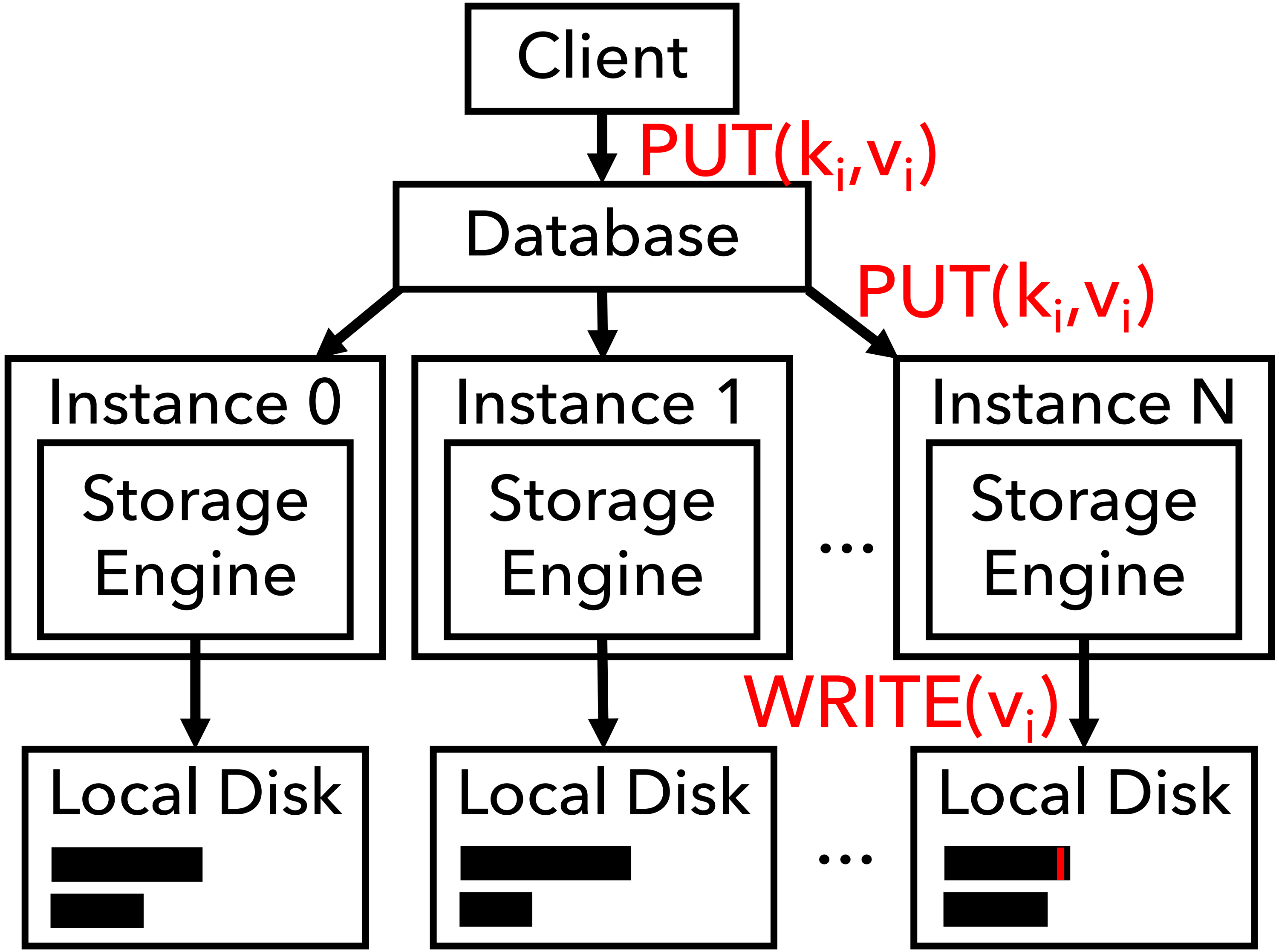
Video of Presenter

# Background – GET Operation



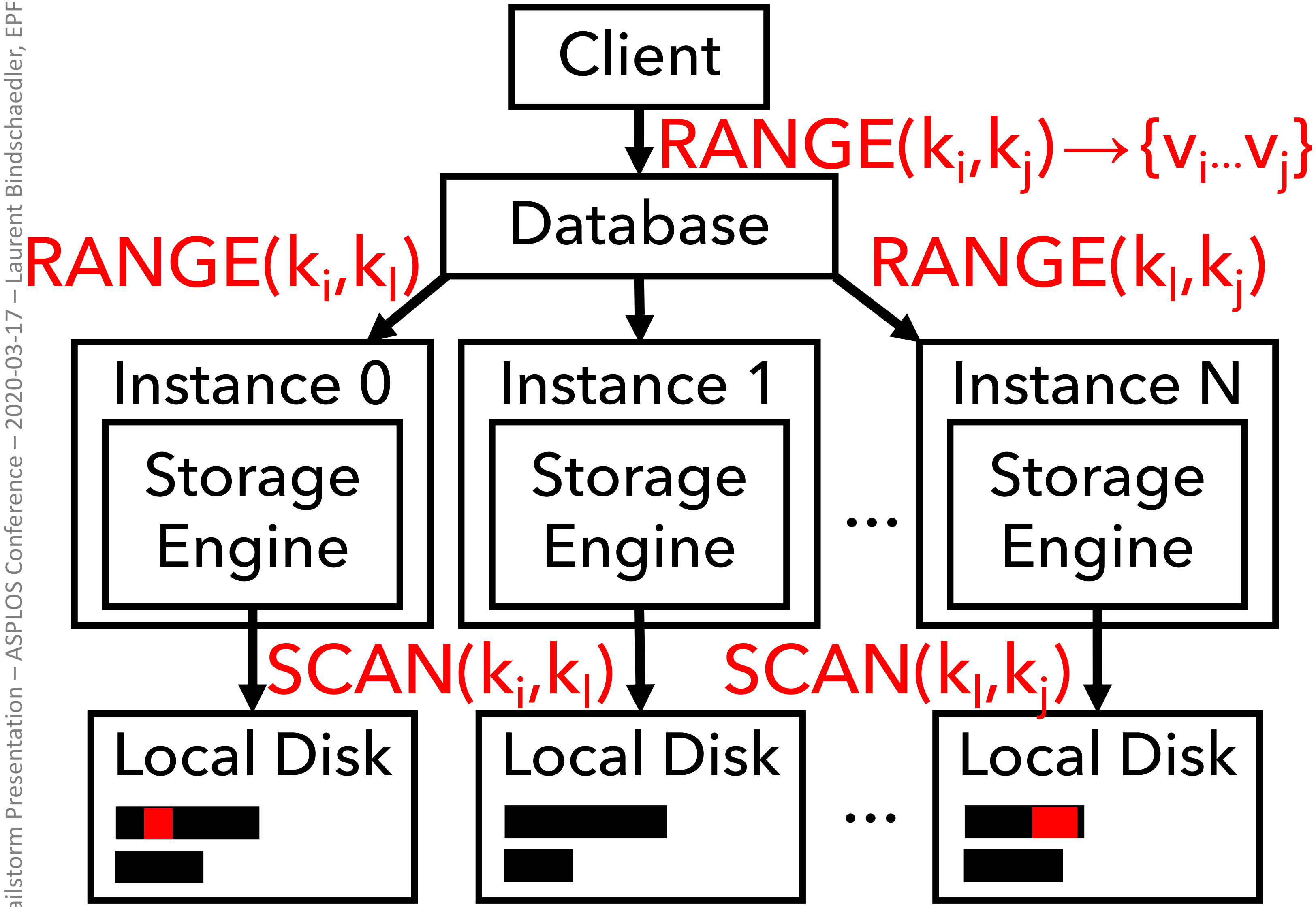
Video of Presenter

# Background – PUT Operation



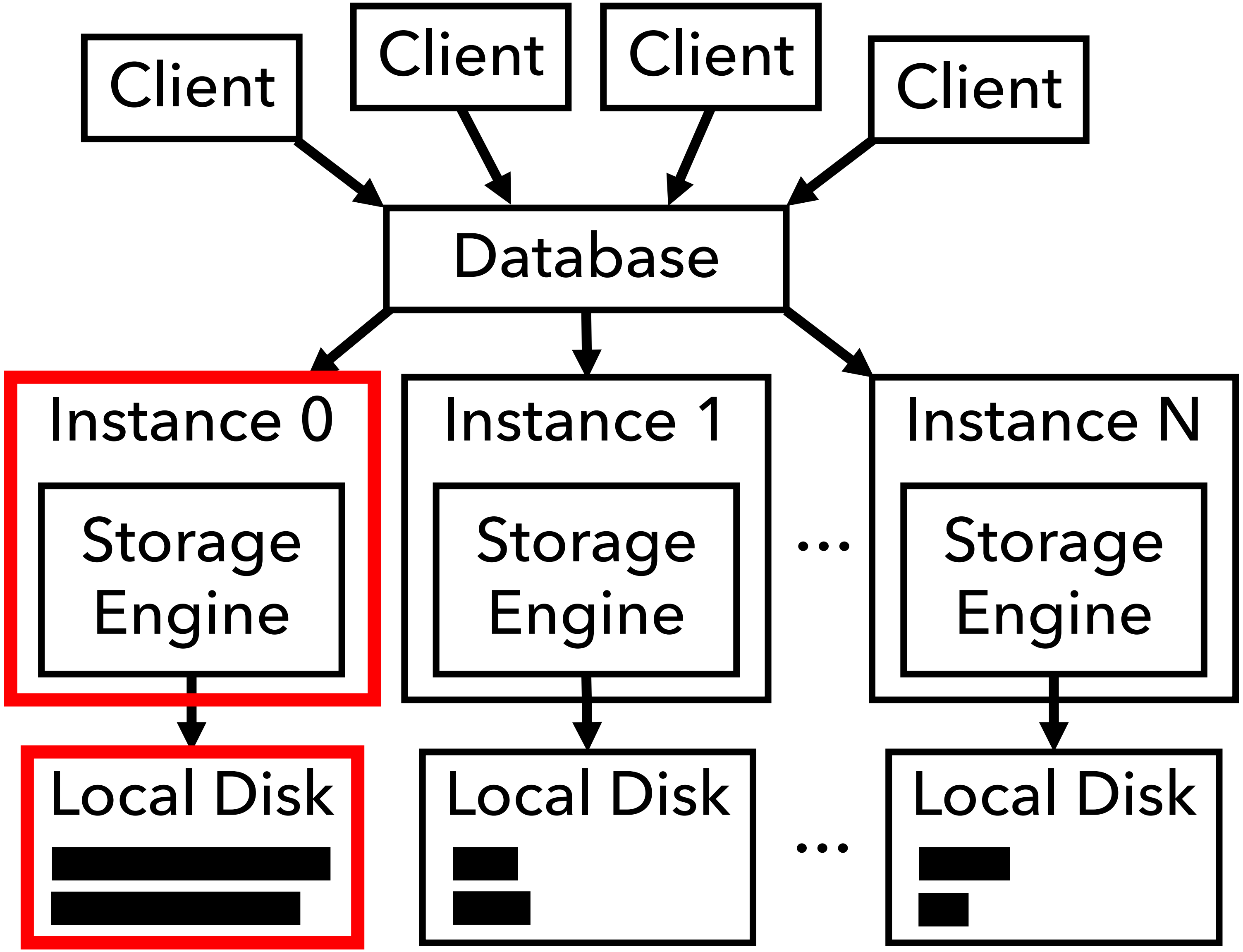
Video of Presenter

# Background – RANGE Operation



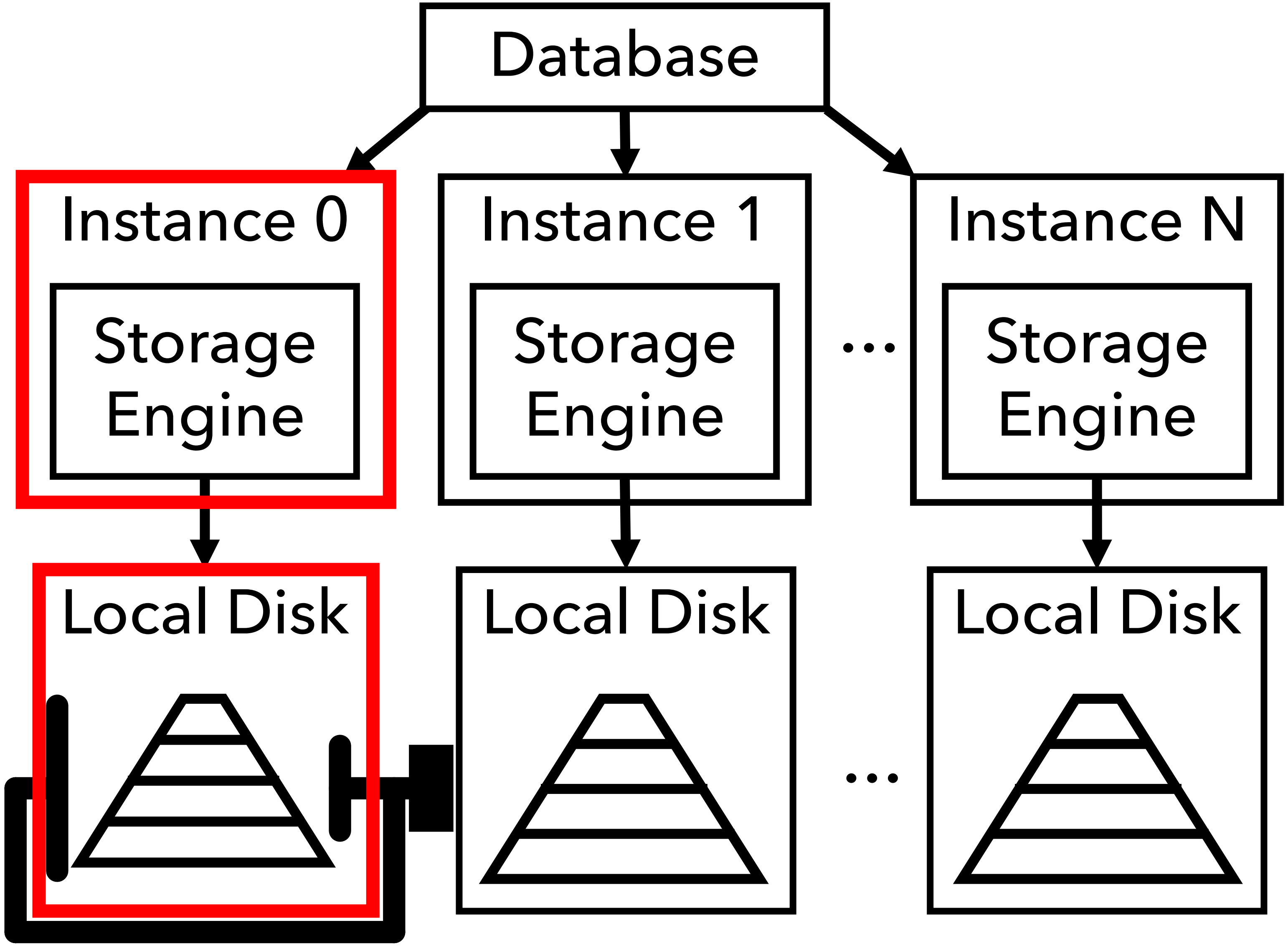
Video of Presenter

# Load Imbalance Cause #1: Skew



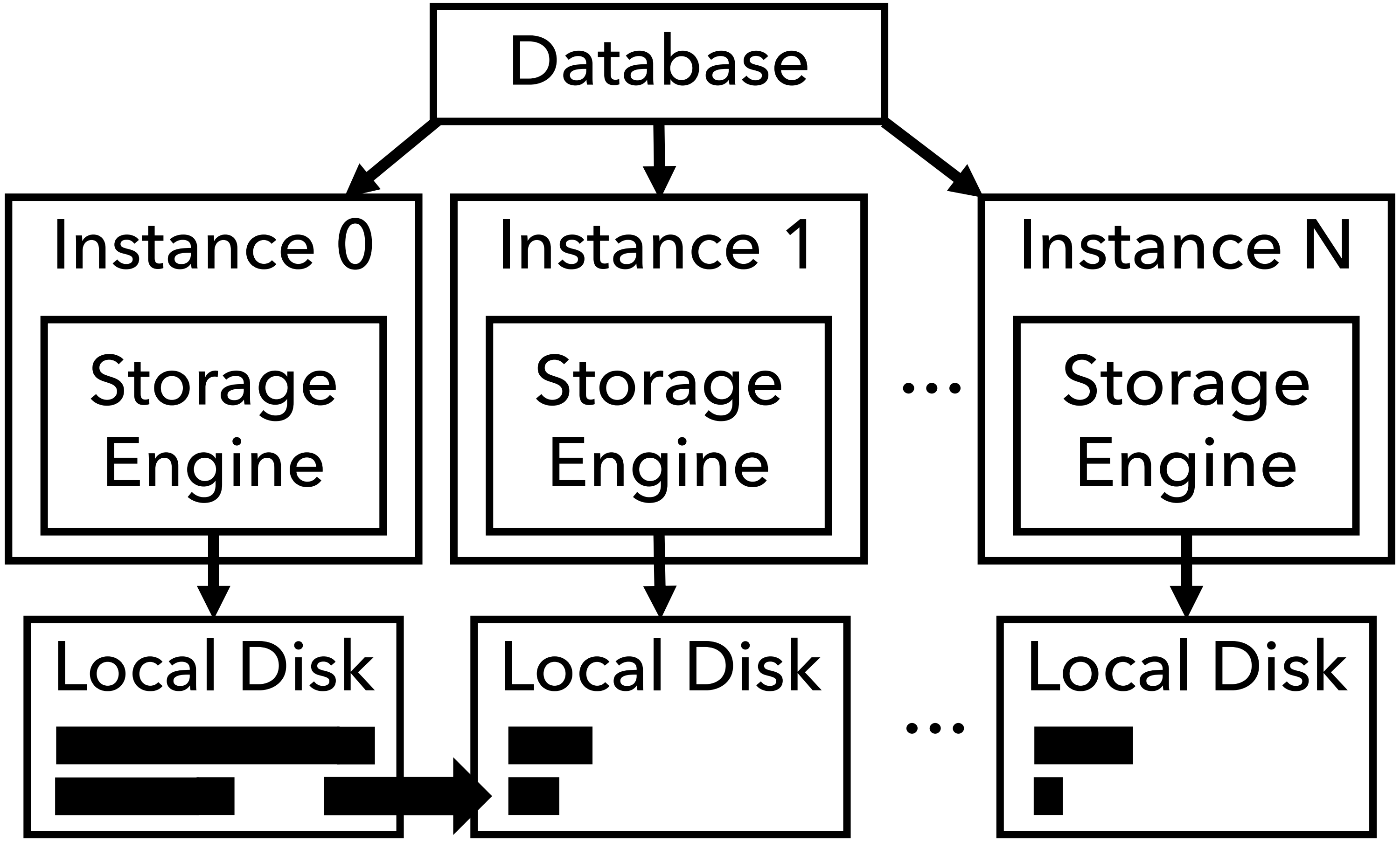
Video of Presenter

# Load Imbalance Cause #2: LSM Compaction



Video of Presenter

# Background – Shard Rebalancing

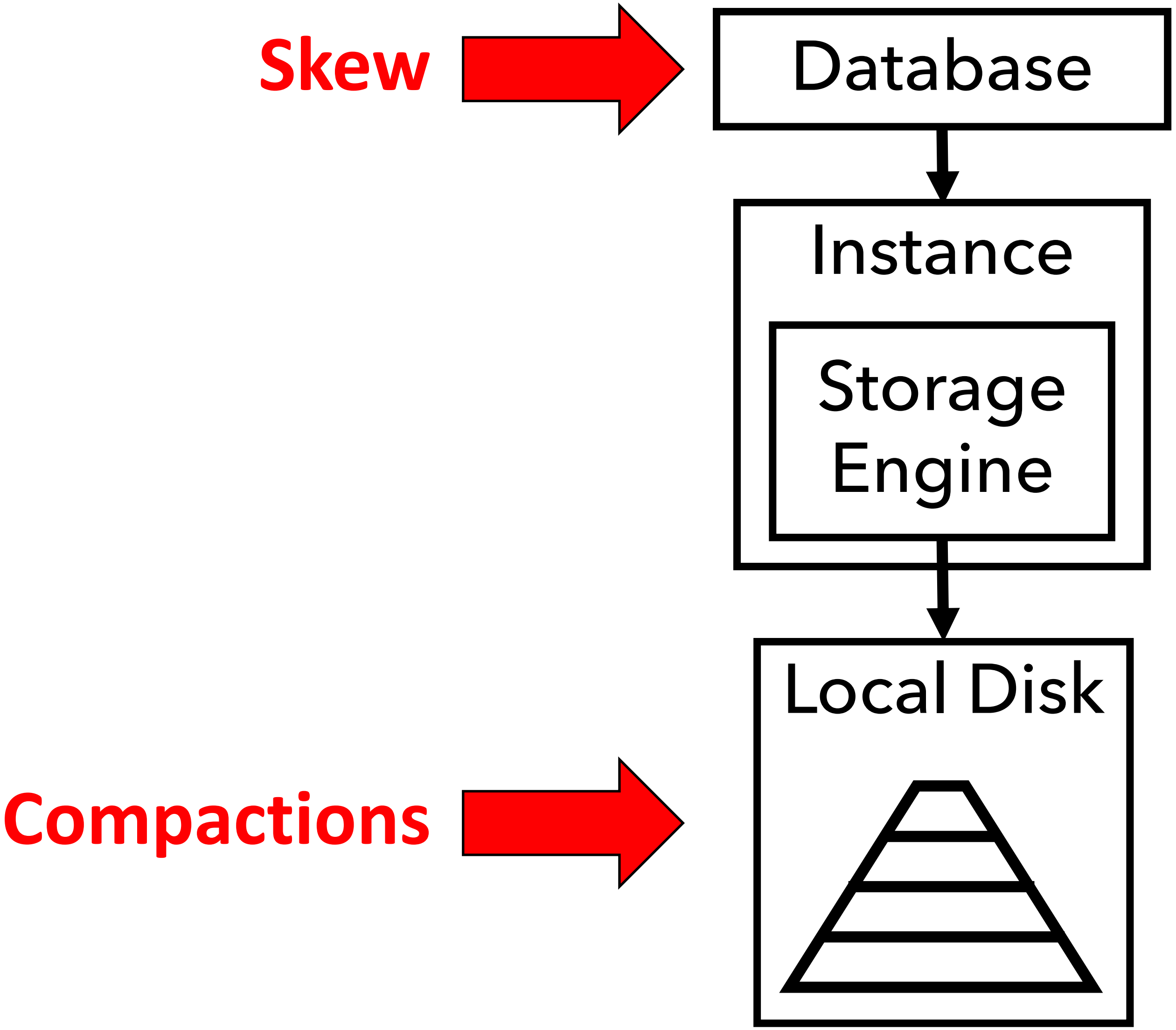


⇒ **Increase load on overloaded instances**

⇒ **Slow & often too late**

Video of Presenter

# CPU and Storage Load Imbalance in LSM Databases



Video of Presenter

# Outline

1. Background
   A. Distributed Databases
   B. Load Imbalance in Distributed Databases
   C. Shard Rebalancing

2. Hailstorm Architecture

3. Evaluation

4. Conclusion

Video of
Presenter

# Hailstorm: Disaggregate Storage & Compute

Scale each resource independently

## Fine-grained storage pooling

- Pool disks within a rack
- Split data in blocks and spread blocks uniformly

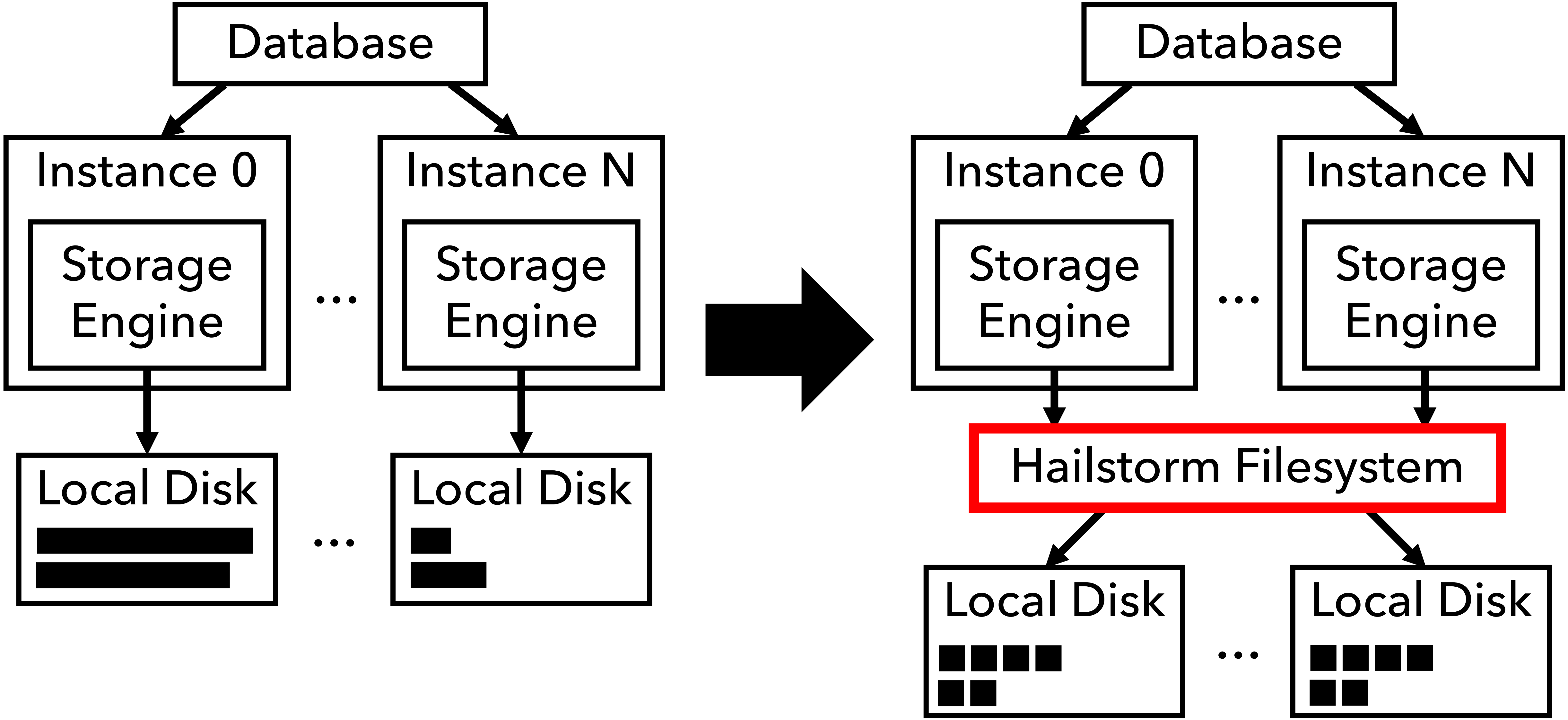
⇒ Achieve storage load balance

## Compaction offloading

⇒ Improve compute load balance

Video of Presenter

# Storage Architecture: Hailstorm Filesystem



# The Hailstorm Distributed Filesystem

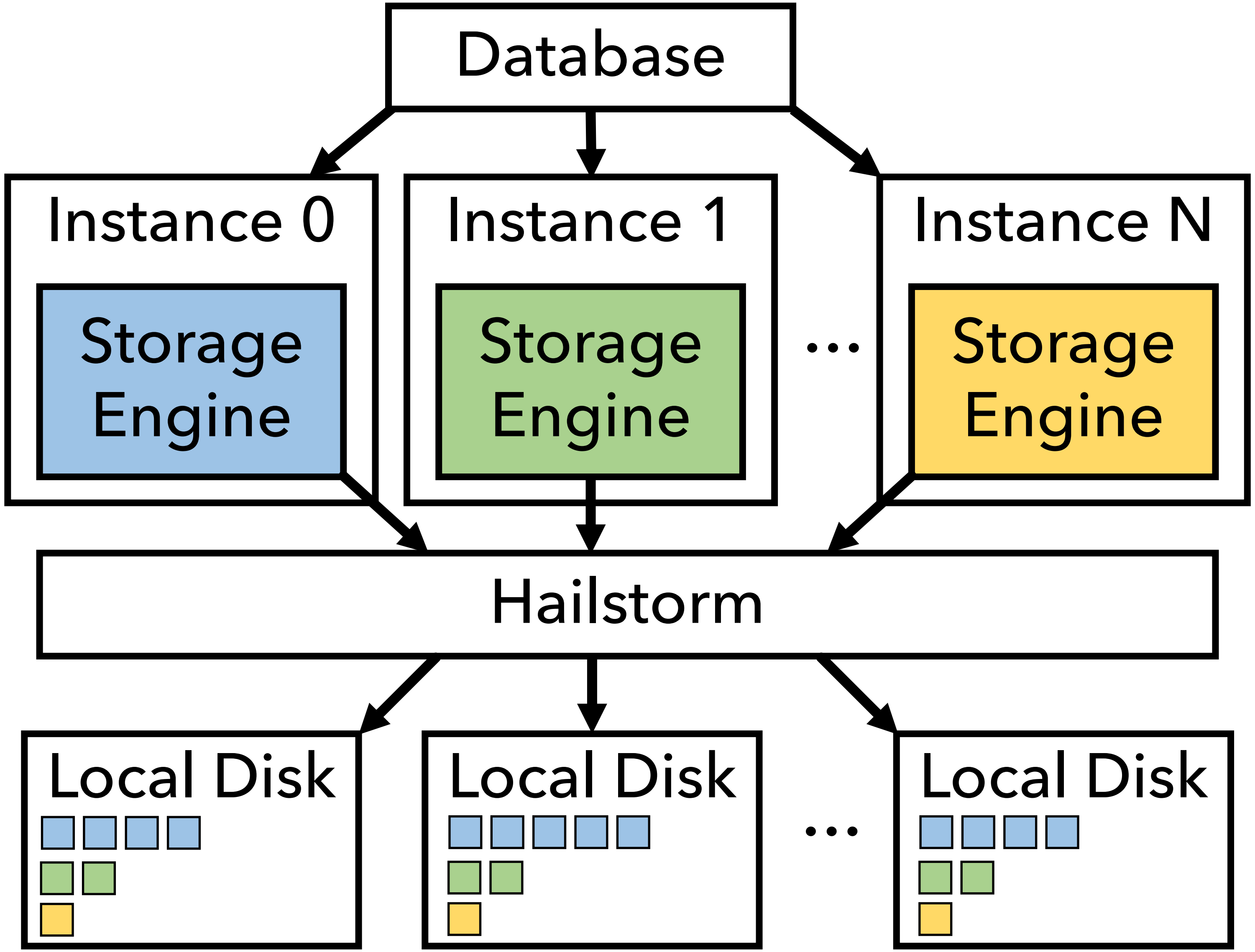
Drop-in replacement for local filesystem
⇒ Supports fine-grained storage pooling & compaction offloading

Blocks are spread in a deterministic order
⇒ Storage engines locate and access data independently

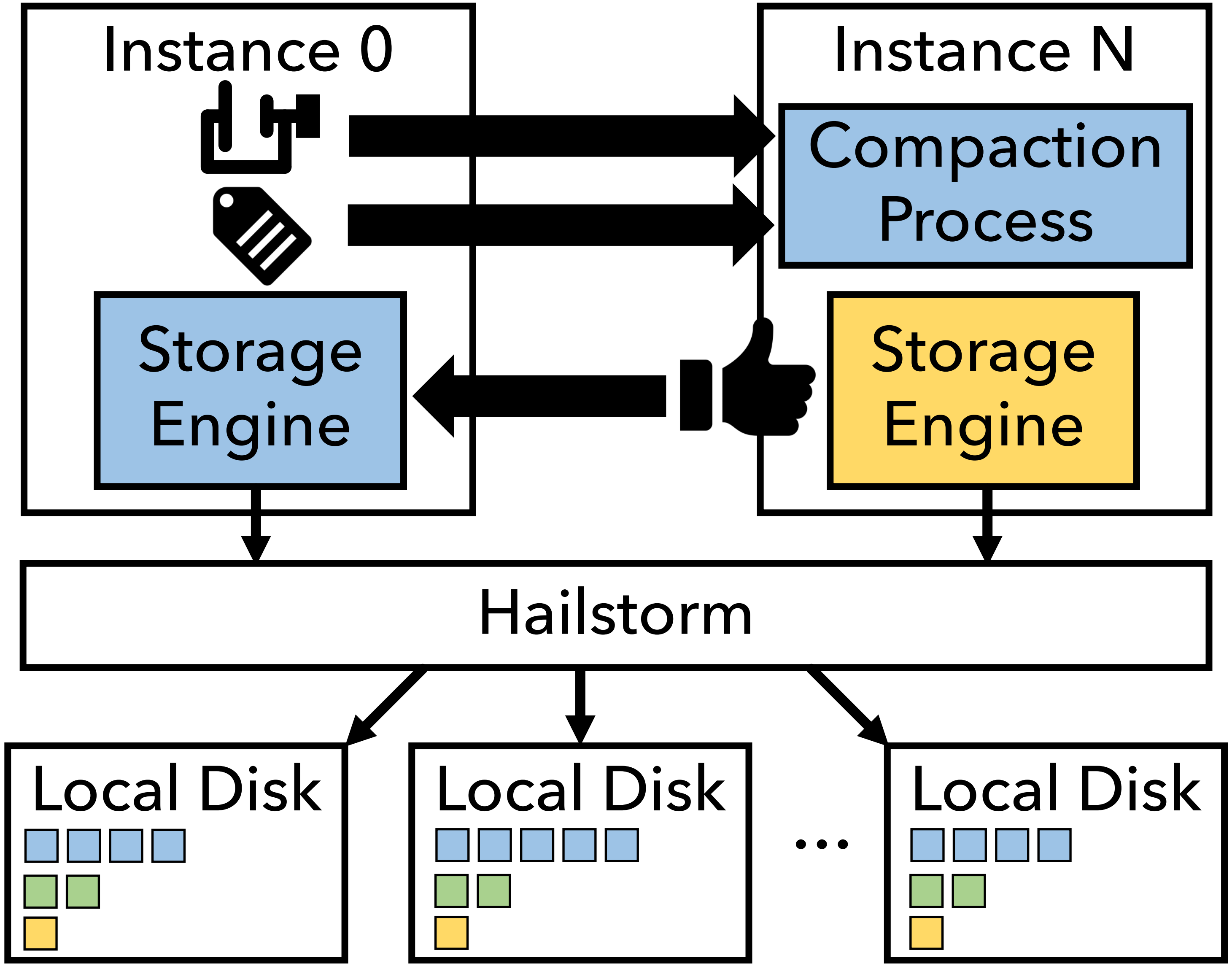
Compaction offloading is efficient
⇒ Blocks are everywhere
⇒ Just need file metadata to locate blocks

Video of Presenter

# Disk Pooling & Data Block Spreading



# Compaction Offloading



# Outline

1. Background
   A. Distributed Databases
   B. Load Imbalance in Distributed Databases
   C. Shard Rebalancing
2. Hailstorm Architecture
3. Evaluation
4. Conclusion

Video of Presenter

# Evaluation

8 16-core machines
32GB of RAM, Intel S3500 SSD, 40GigE switch

MongoDB (key-value)
TiDB (SQL ACID)

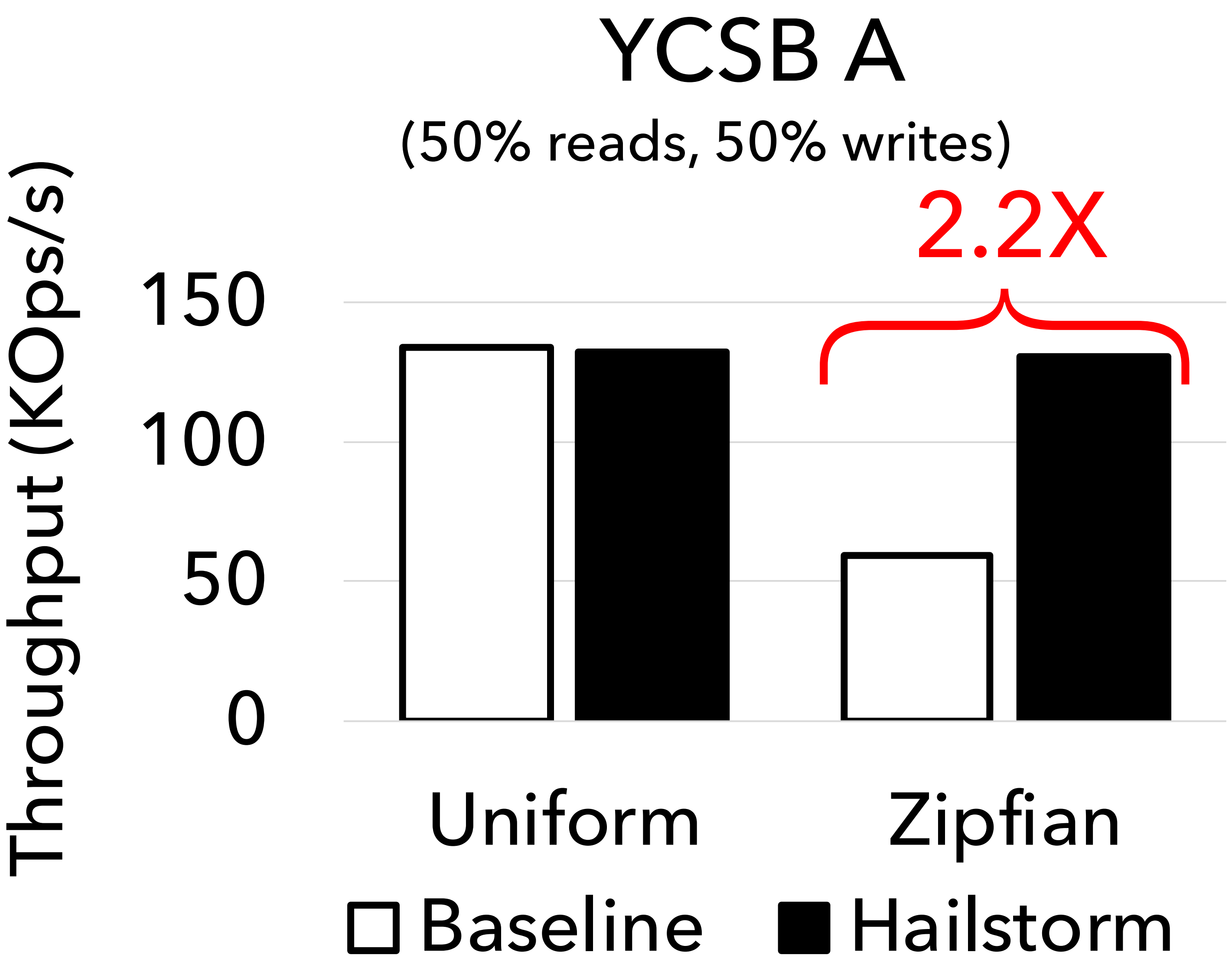
RocksDB storage

Baseline vs Hailstorm

Uniform vs Zipfian (skewed)

Video of Presenter

# MongoDB – Write-Intensive YCSB Throughput

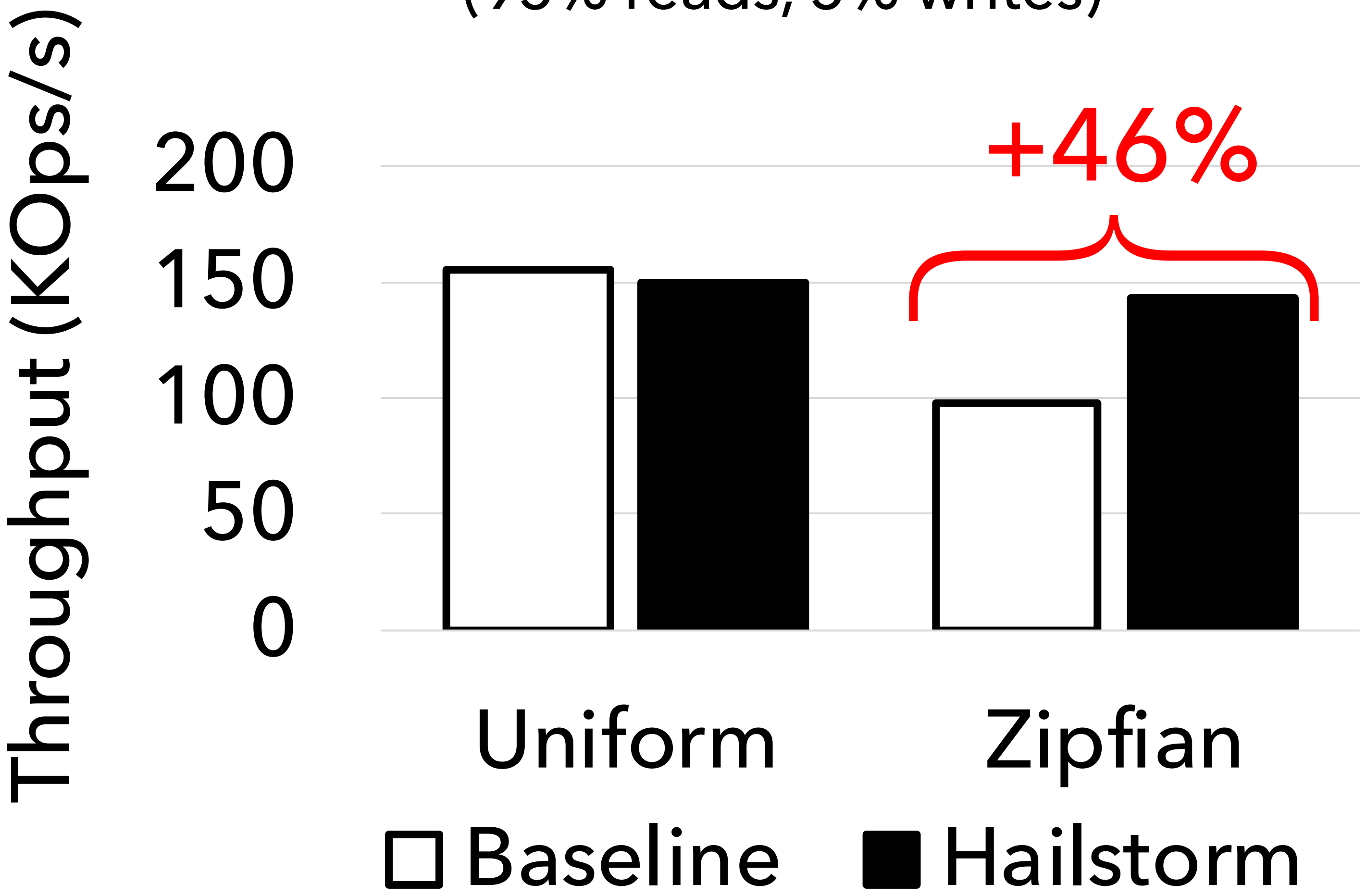


Video of Presenter

# MongoDB – Read-Intensive YCSB Throughput

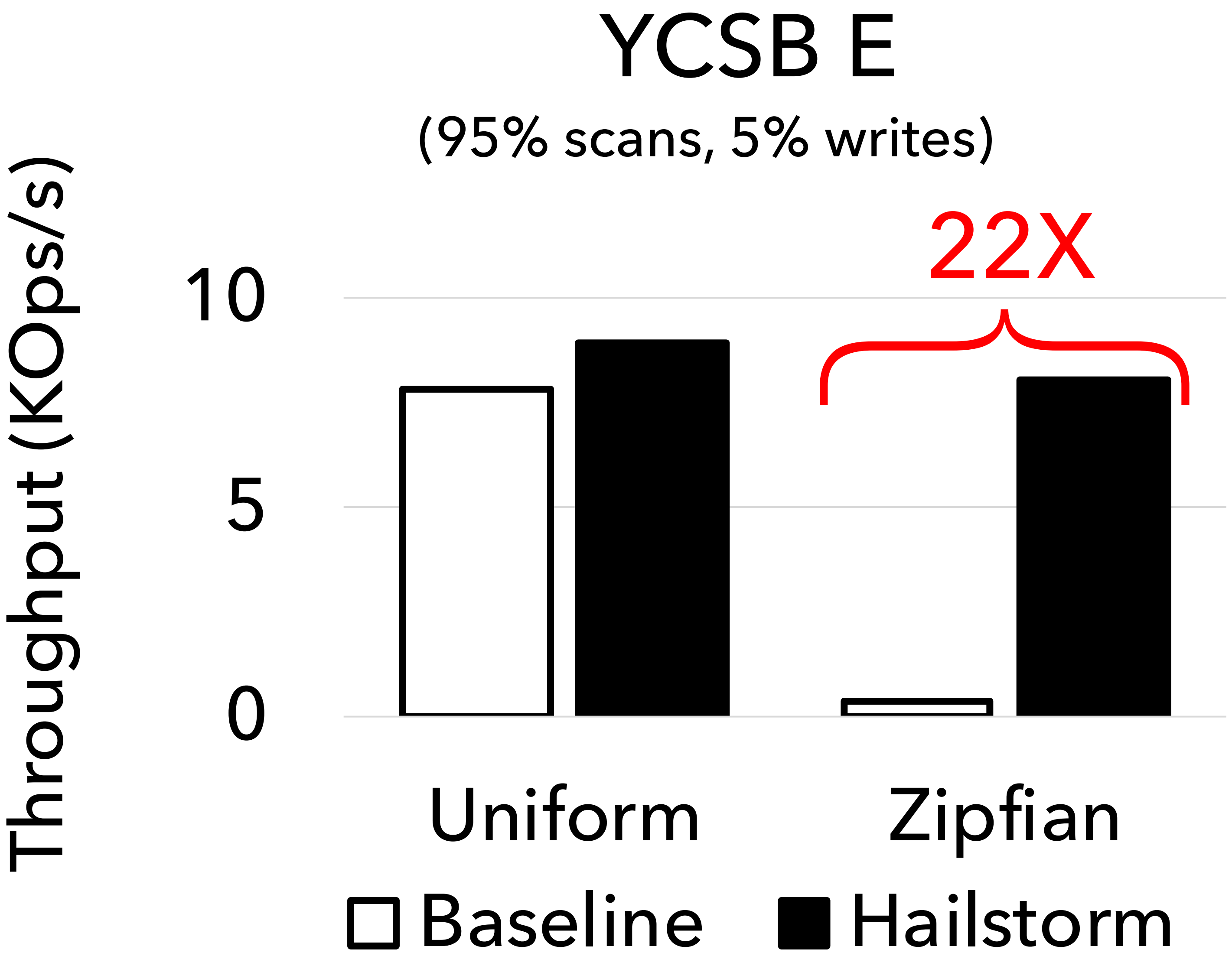
YCSB B

(95% reads, 5% writes)



Video of Presenter

# MongoDB – Scan-Intensive YCSB Throughput



Video of Presenter

# TiDB – TPC-C & TPC-E

| Configuration | | Standalone TiDB | TiDB over Hailstorm |
|---|---|---|---|
| TPC-C | tpmC | 32,184 | 50,178 |
| | $ / tpmC* | 3.10 | 2.00 |
| TPC-E | tpsE | 277.3 | 408.1 |
| | $ / tpcE* | 360.60 | 245.05 |

1.5X improvements

*Estimated total system cost: $100,000.

Video of Presenter

# Additional Results in Paper

- Complete YCSB benchmark
- Throughput over time
- Response latency
- Large datasets
- Shard rebalancing
- Comparison with HDFS
- Performance breakdown
- Hailstorm for B-trees

Video of
Presenter

# Conclusion

Hailstorm improves load balance in LSM-based databases

⇒ Higher throughput & resource utilization

Hailstorm works with existing databases

Key idea: **compute and storage disaggregation**

- Fine-grained storage pooling
- Compaction offloading

Video of Presenter

# Thank you and stay tuned for updates!

# Hailstorm

## Disaggregated Compute and Storage for Distributed LSM-based Databases

Video of
Presenter