

Larger-than-Memory Data Management on Modern Storage Hardware for In-Memory OLTP Database Systems

Lin Ma, Joy Arulraj, Sam Zhao, Andrew Pavlo, Subramanya R. Dulloor, Michael J. Giardino, Jeff Parkhurst, Jason L. Gardner, Kshitij Doshi, Col. Stanley Zdonik

Larger-than-Memory Databases

Crux:

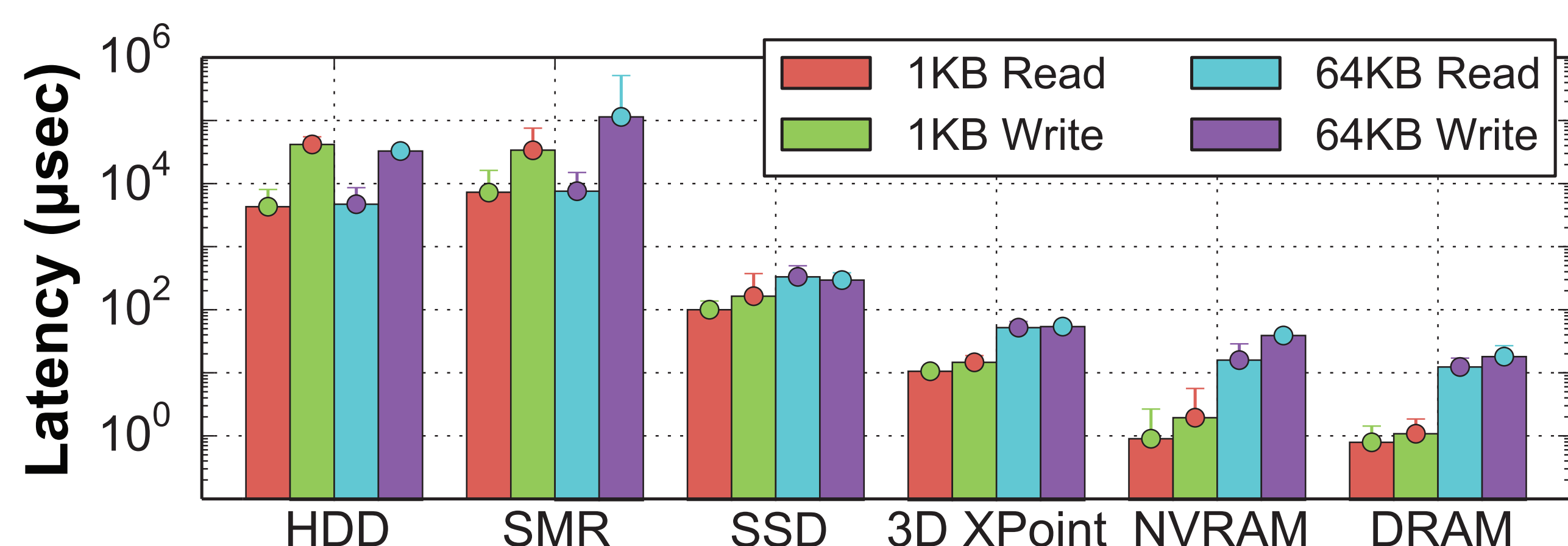
- OLTP workloads exhibit skewed access patterns
- Move cold data to cheaper secondary storage
- Deliver high performance for transactions that operate on hot in-memory tuples

In-Memory DBMSs that support cold data management:

- Anti-caching for H-Store, Microsoft's Project Siberia, EPFL's VoltDB, Apache Geode, MemSQL

Storage Technologies

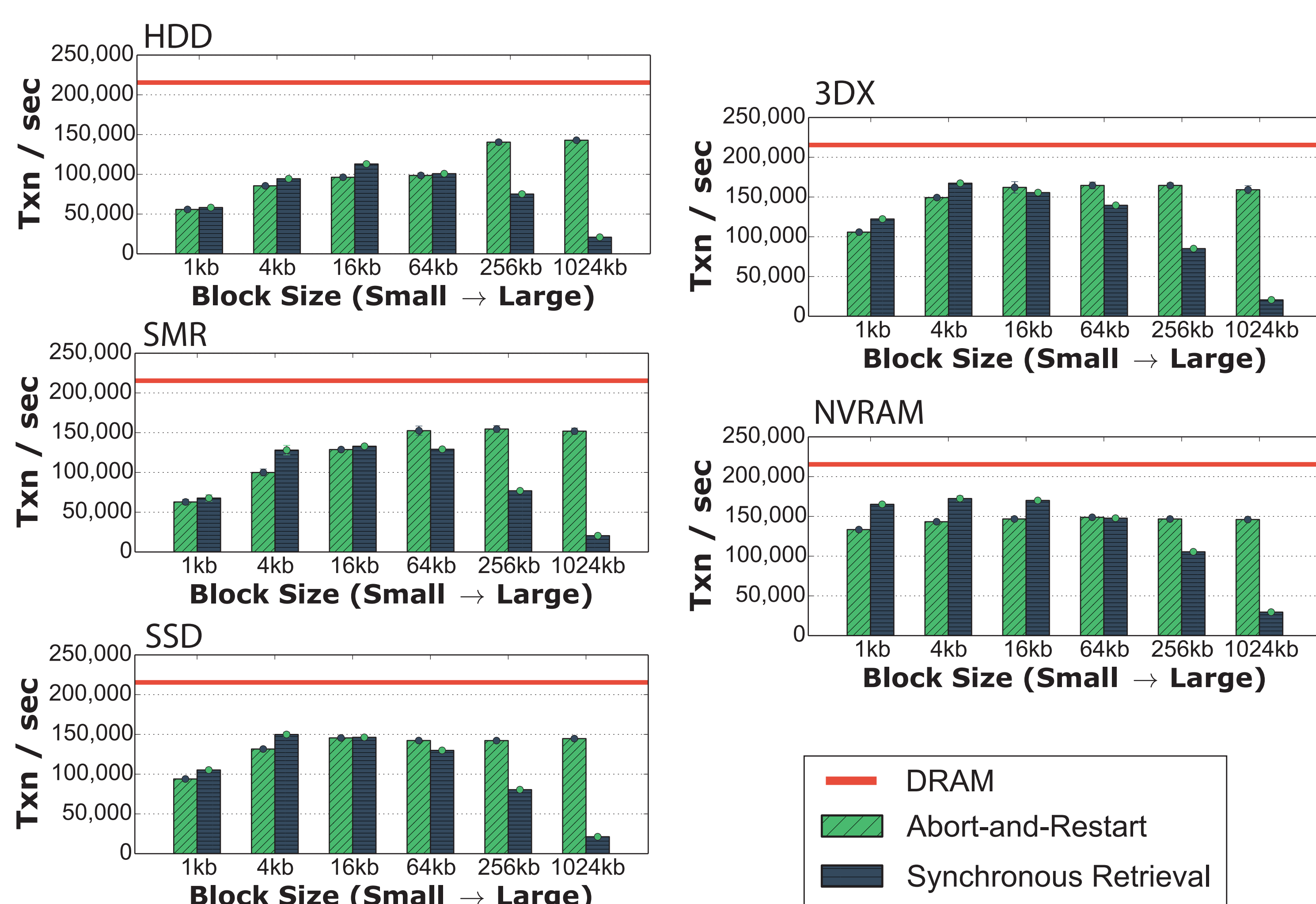
- Read/write latency of different storage devices:



Storage devices evaluated using microbenchmark to simulate reading/writing cold tuple workload in an in-memory DBMS.

Cold Tuple Retrieval

- Throughput for the YCSB workload in H-Store with anti-caching (10GB database with 1.25GB DRAM):



Cold Data Management Policies

Hardware independent policies:

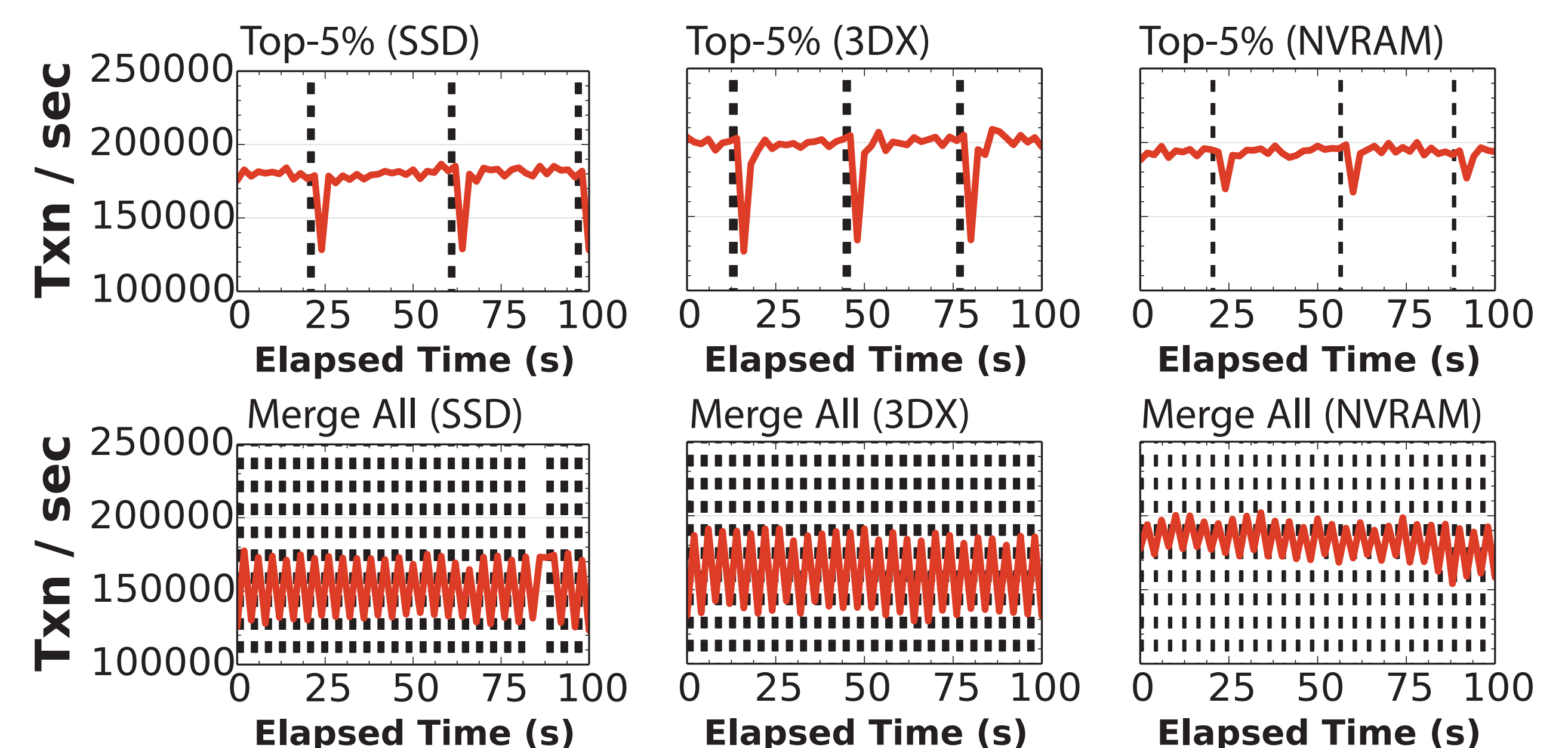
- Cold tuple identification
- Evicted tuple meta-data
- Eviction timing

Hardware dependent policies:

- Cold data retrieval
- Merging threshold
- Access methods

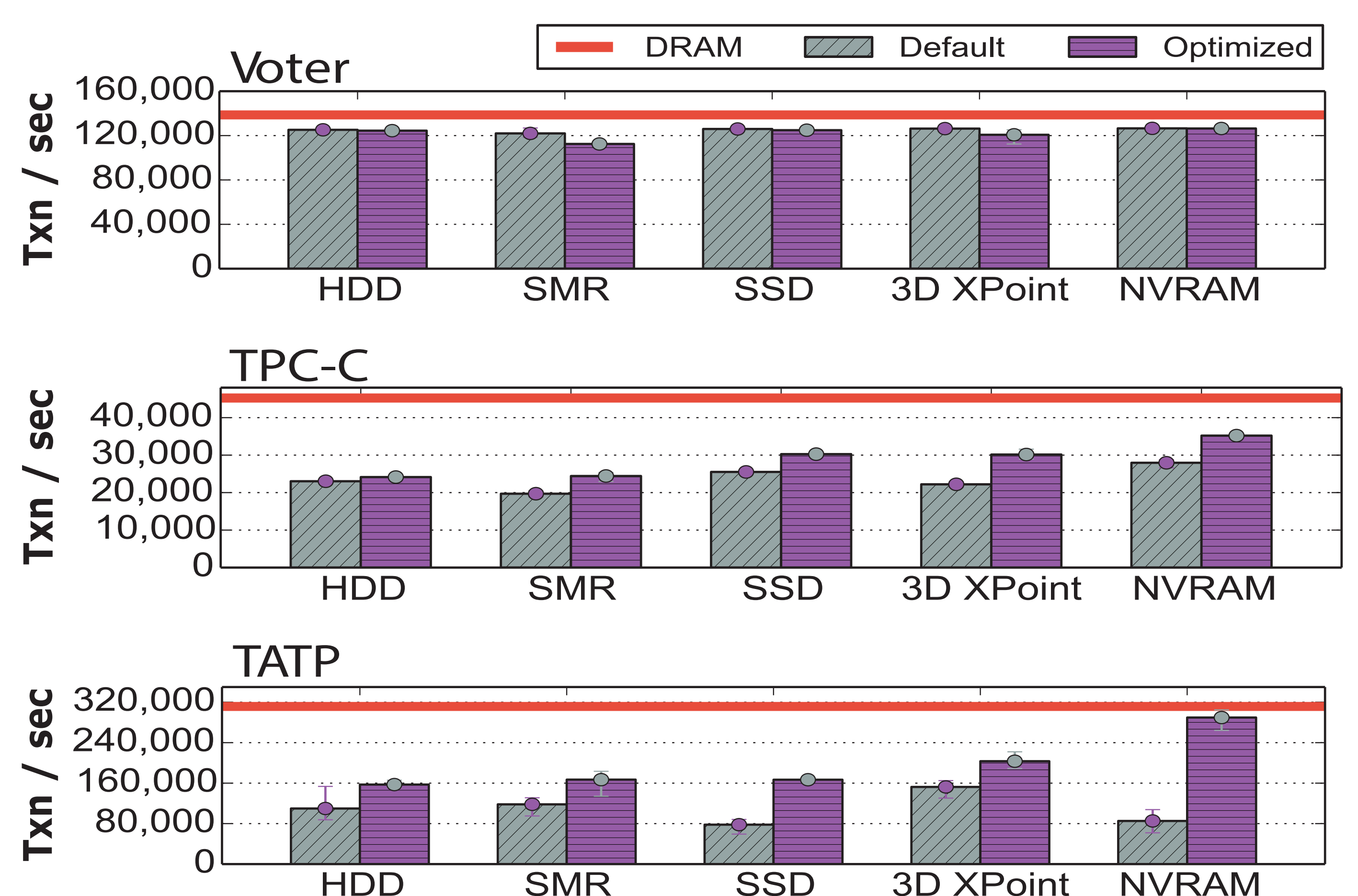
Merging Threshold

- Put an accessed cold tuple into a temporary buffer or merge it back into the table based on access frequency:



Throughput for YCSB in H-Store under different merge threshold policies. Vertical bars show tuple eviction to secondary storage. The y-axis shows real-time throughput of the DBMS every second.

Additional Workloads



Throughput measurements for H-Store with anti-caching when using the optimal hardware-dependent policy configuration for each storage device compared to a default configuration