A Memory-Efficient Persistent Key-Value Store on eNVM SSDs

Arup De & Zvonimir Bandic

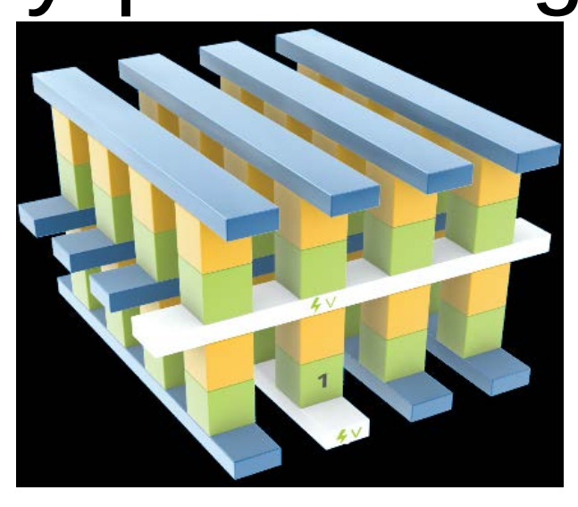
Abstract

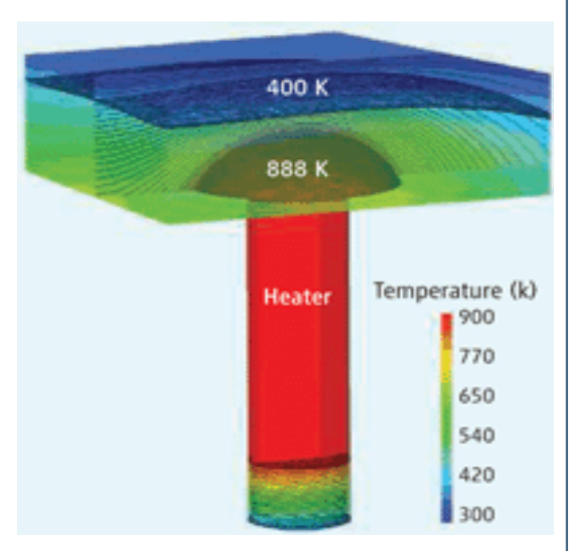
Emerging fast byte-addressable non-volatile memory (eNVM) technologies such as ReRAM and 3D Xpoint are projected to offer two orders of magnitude higher performance than flash. However, the existing solid-state drive (SSD) architecture optimizes for flash characteristics and is not adequate to exploit the full potential of eNVMs due to architectural and I/O interface (e.g., PCIe, SATA) limitations. To improve the storage performance and reduce the host main memory requirement for KVS, we propose a novel SSD architecture that extends the semantic of SSD with the KVS features and implements indexing capability inside SSD. It has in-storage processing engine that implements key-value operations such as get, put and delete to efficiently operate on KV datasets. The proposed system introduces a compute channel interface to offload key-value operations down to the SSD that significantly reduces the operating system, file system and other software overhead. This SSD achieves 4.96 Mops/sec get and 3.44 Mops/sec put operations and shows better scalability with increasing number of key-value pairs as compared to flash-based NVMe (flash-NVMe) and DRAM-based NVMe (DRAM-NVMe) devices. With decreasing DRAM size by 75%, its performance decreases gradually, achieving speedup of 3.23x as compared to DRAM-NVMe. This SSD significantly improves performance and reduces memory by exploiting the fine grain parallelism within a controller and keeping data movement local to effectively utilize eNVM bandwidth and eliminating the superfluous data movement between the host and the SSD.

Introduction

- Key-value store is a fundamental building block for many enterprise applications
  - Social Networks
  - Online shopping
  - Inline storage deduplication

- Key-value store
  - Supports simple operations: Get, Put and Delete
  - Preferred over traditional relational DBs for its superior scalability, performance and simplicity
  - Often implemented through an in-memory index structure which points to key-value pairs in storage
  - Popular management solution for large volume of records
- Emerging NVM technologies are very promising
  - Byte-addressable
  - High density
  - Low standby power
  - DRAM-like performance



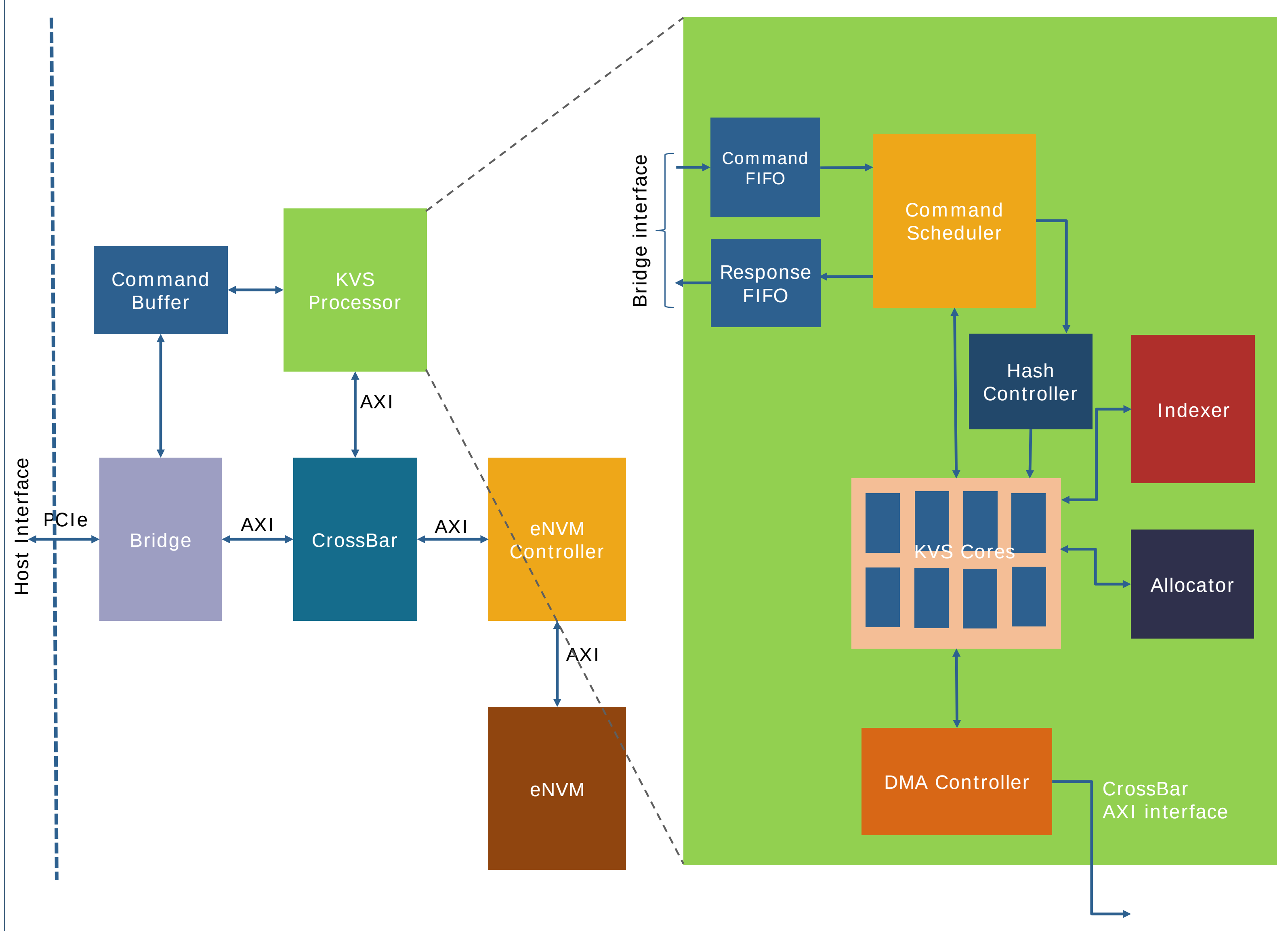


Challenges

- The existing SSD architecture optimizes for flash characteristics and is not adequate to exploit the full potential of emerging NVM technologies due to architectural and I/O interface (e.g., PCIe, SATA) limitations
- The main memory size imposes a challenging problem in scalability and performance of key-value stores due to relatively slow growth of DRAM capacity as compared to rapidly growing key-value datasets
- Key-value store has random accesses to the storage and the existing memory/storage hierarchy is not adequate for this type of application
  - Cache miss, TLB flush
  - Poor host CPU utilizations
  - Large DRAM usage for caching and metadata management

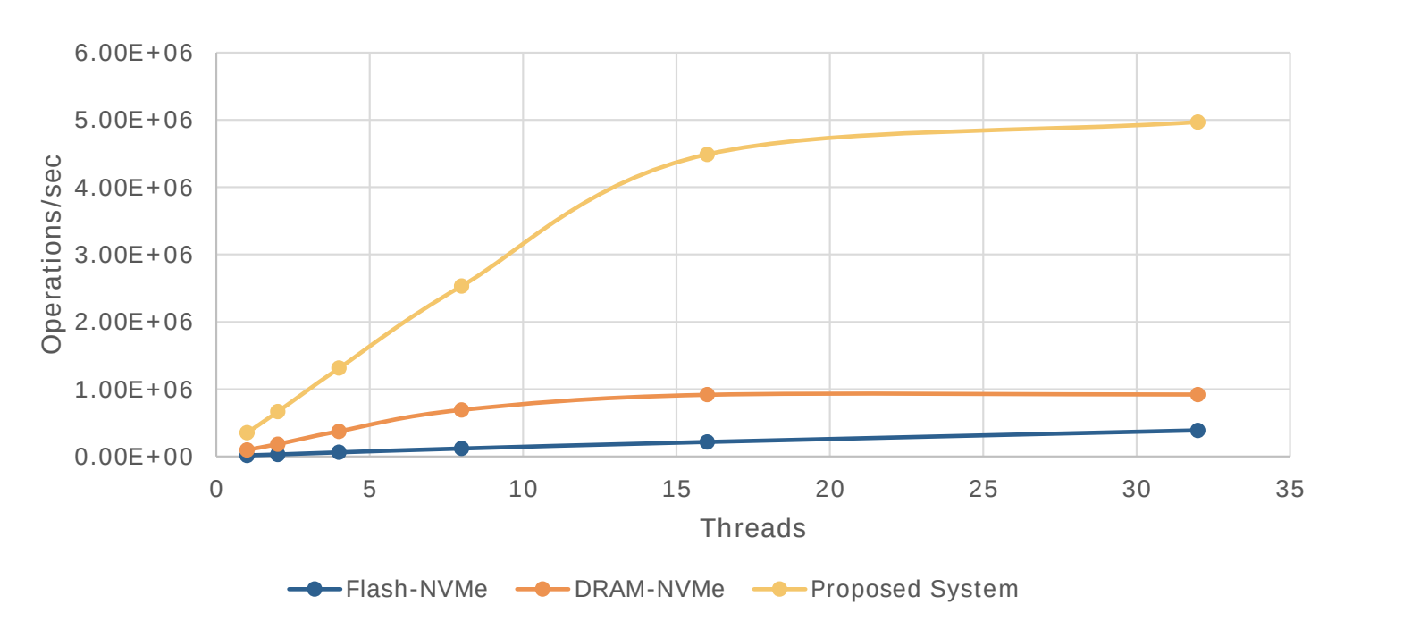
The Proposed System Architecture

- The proposed system extends the conventional storage interface (read and write) with key value store interface (get, put, delete)
- It leverages the high internal bandwidth (8 to 10x more than the external I/O interface) to improve overall system performance and reduce the memory requirement
  - Improve CPU utilizations
  - Reduce host memory usage
- It introduces "Compute Channel" interface to offload key-value operations down to the SSD that significantly reduces the OS and other software overhead

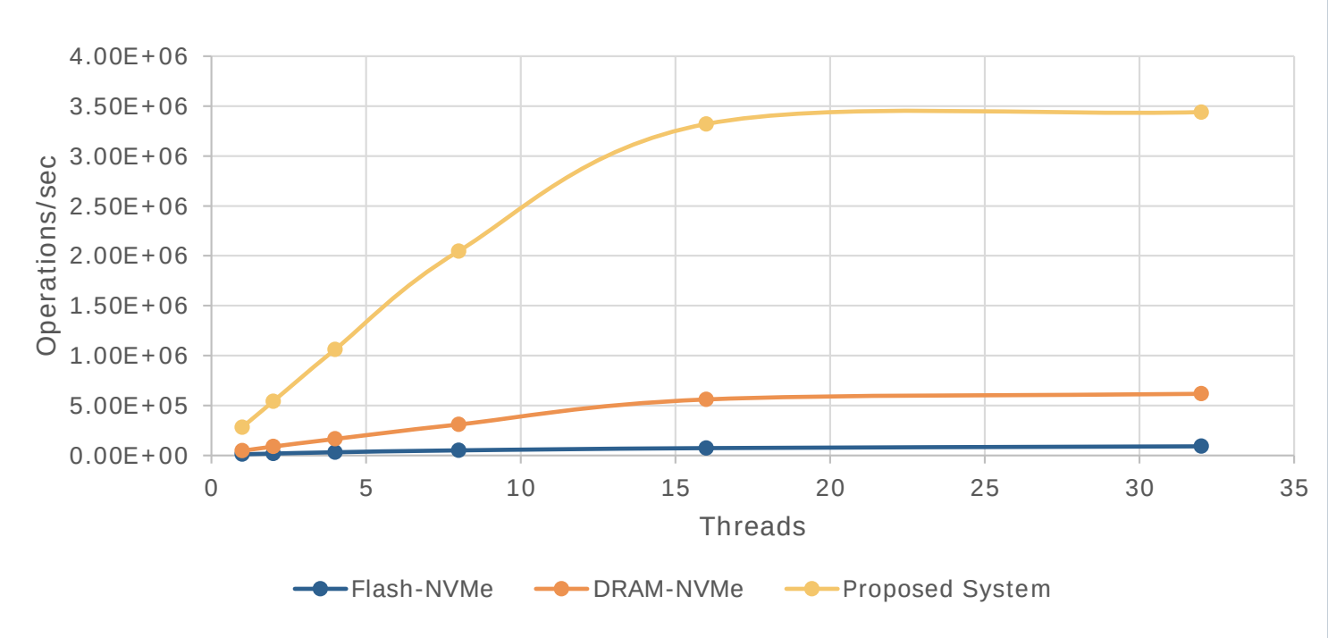


Results

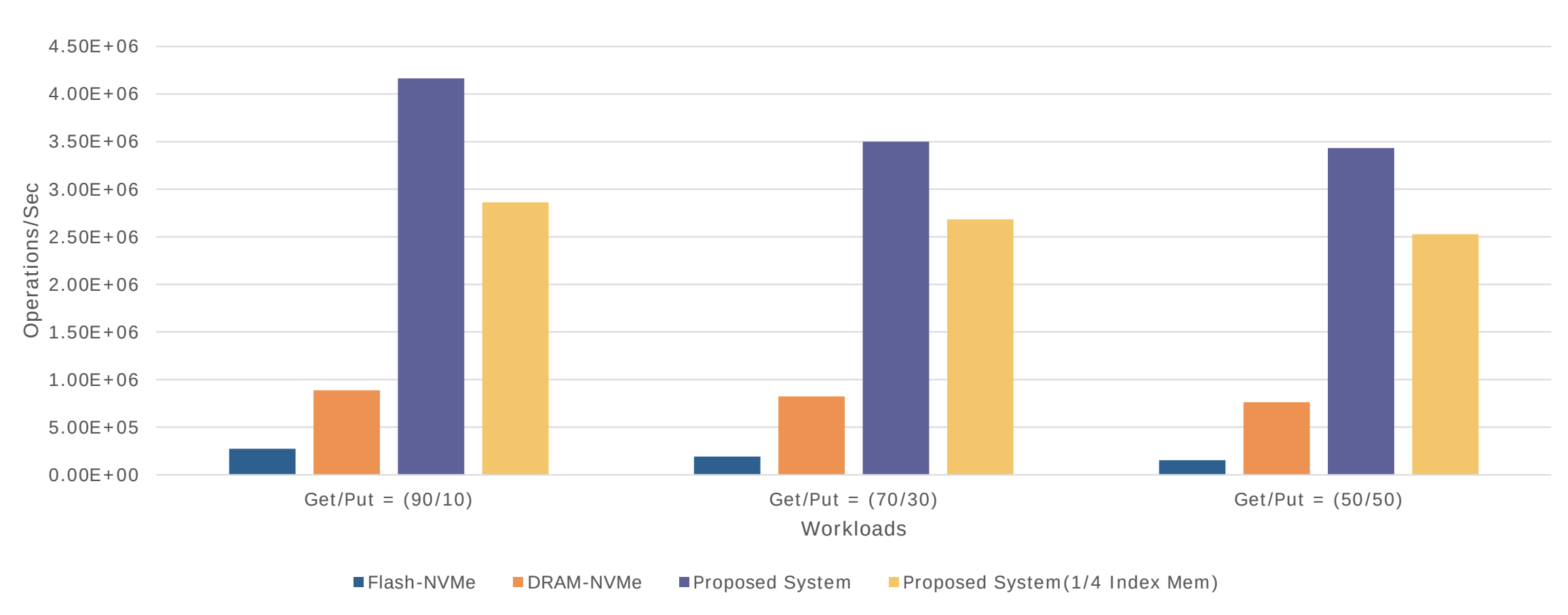
Get Performance (32 B key and 512 B Value)



Put Performance (32 B key and 512 B Value)



Workloads Performance(32 B key and 512 B Value)



Conclusion

- Propose a memory-efficient key-value store for next-generation SSDs
- Extended semantic of SSD with key-value store features
- Significantly reduced the host CPU and DRAM usage for key-value data processing
- Demonstrated in a prototype storage system with adequate software and hardware support