Jin-Soo Kim
(*jinsoo.kim@snu.ac.kr*)

Systems Software &
Architecture Lab.

Seoul National University

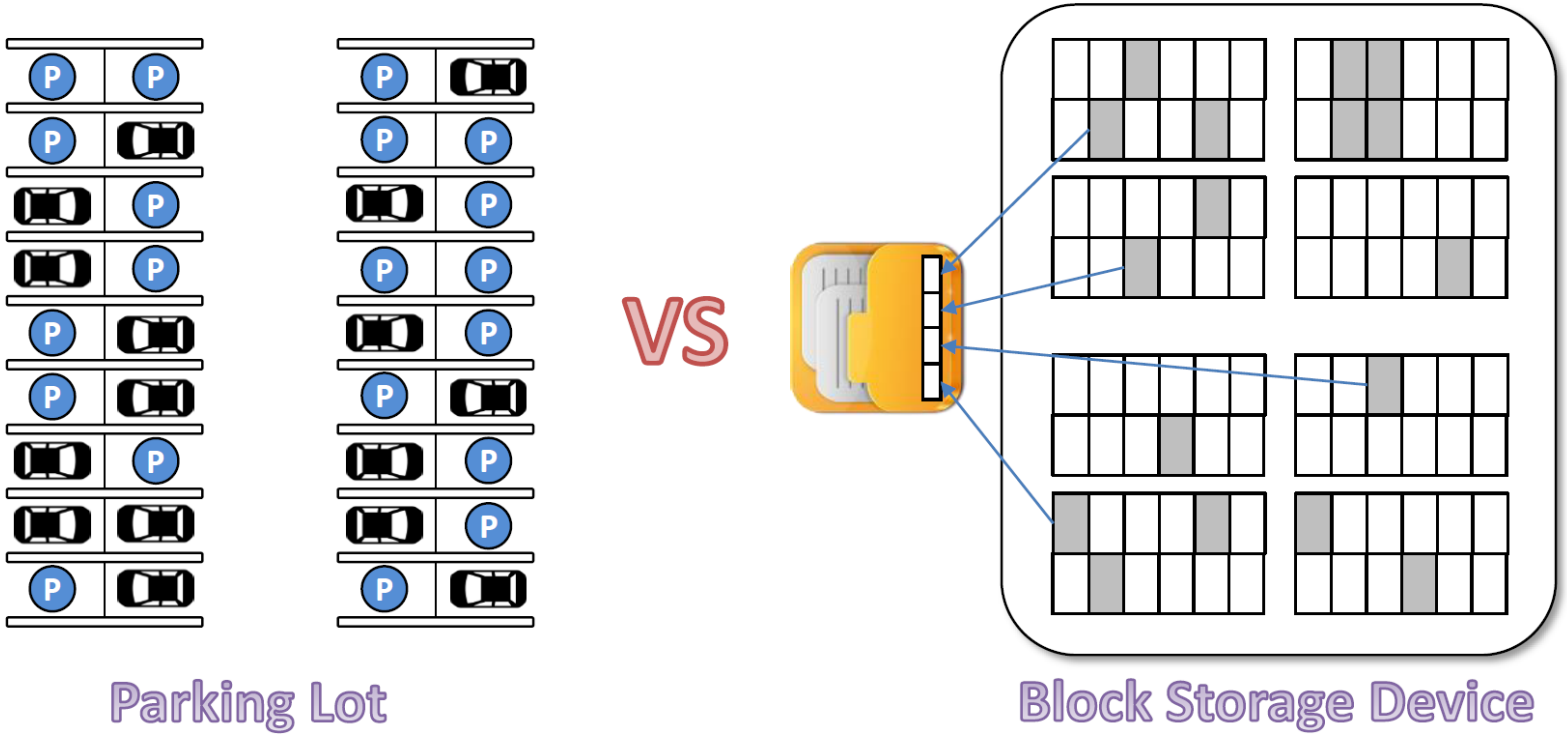
Spring 2026

New SSDs

Key-Value SSD (KVSSD)

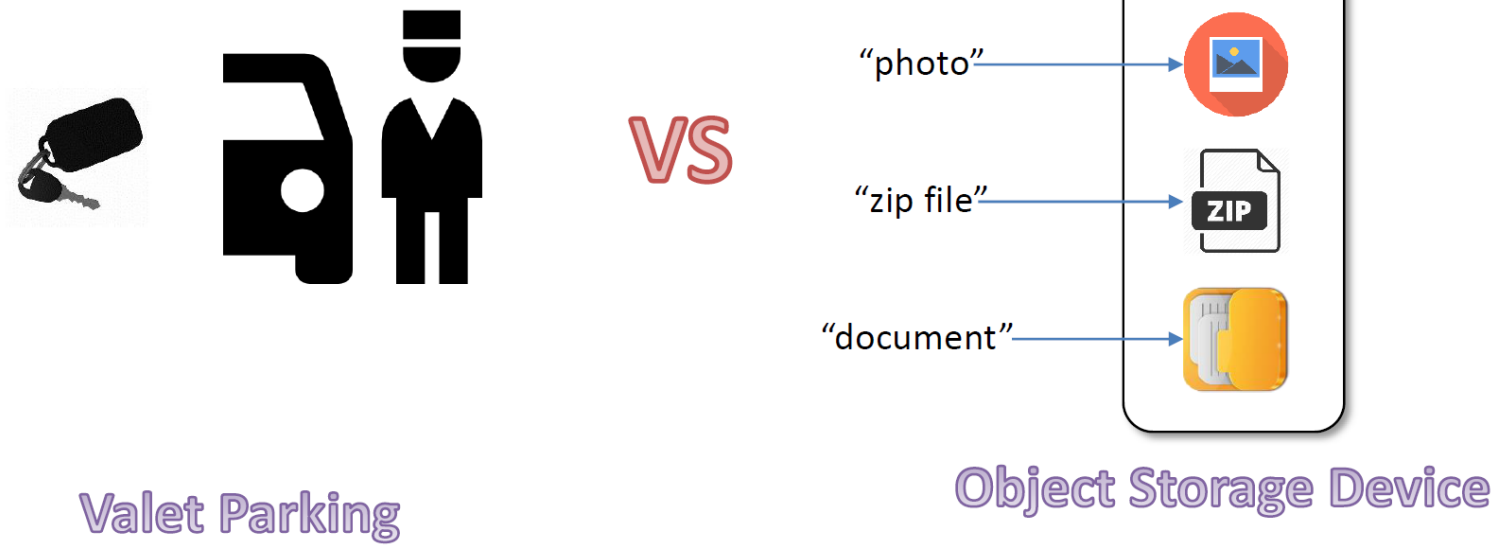
Block: Parking Lot Structure

- A driver (host) is responsible for parking (data management)

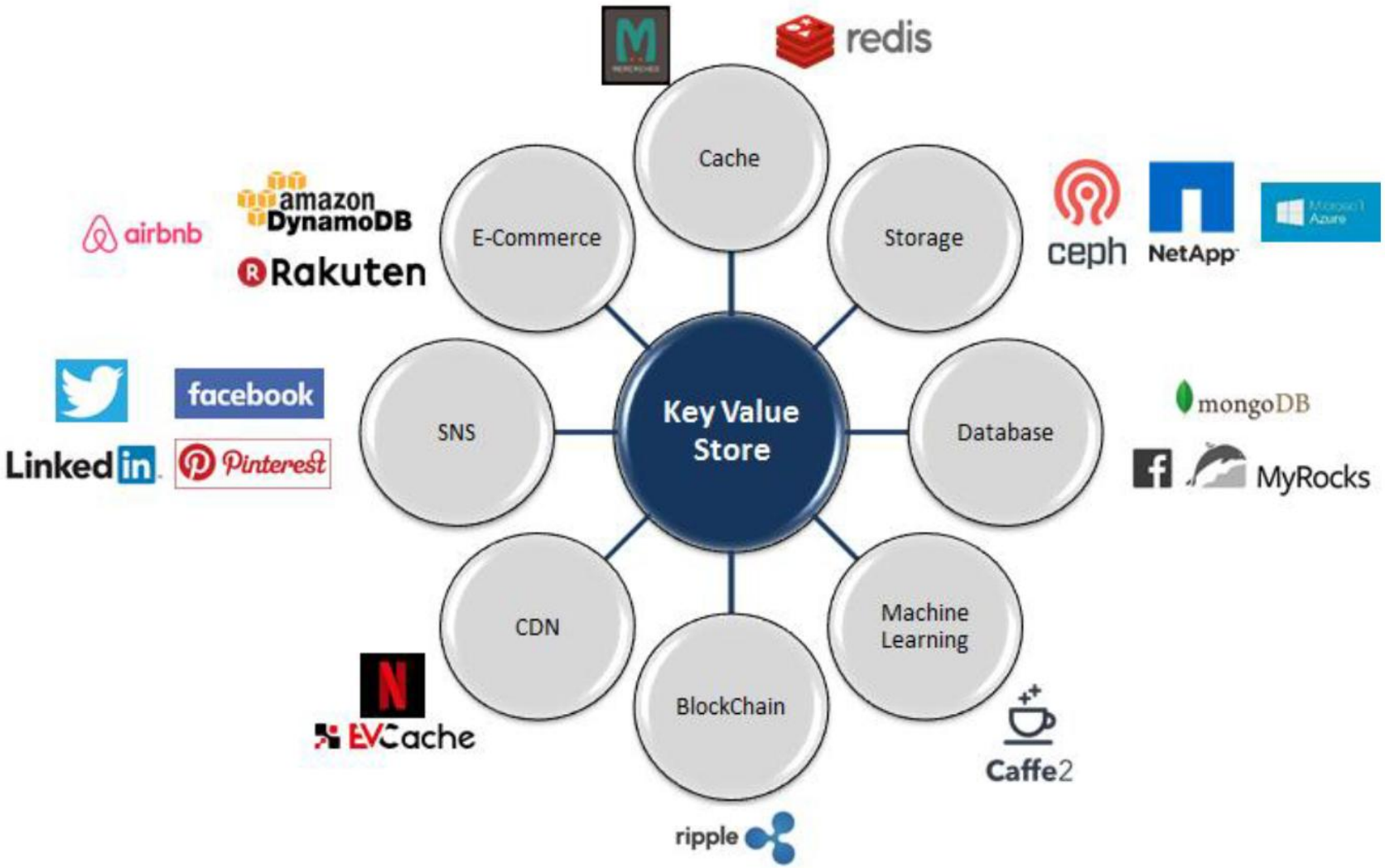


Object: Valet Parking

- A parking facility (storage) is responsible for parking (data management)



KV Stores Common in Systems at Scale



Standardization

- SNIA Key Value Storage API
  - v1.0 (2019) and v1.1 (2020)
  - `kvs_retrieve_kvp()`, `kvs_store_kvp()`, `kvs_delete_kvp()`, `kvs_exist_kv_pairs()`, `kvs_create_iterator()`, `kvs_iterate_next()`, `kvs_delete_iterator()`, …
  - Keys starting with a specific prefix can be retrieved using an iterator

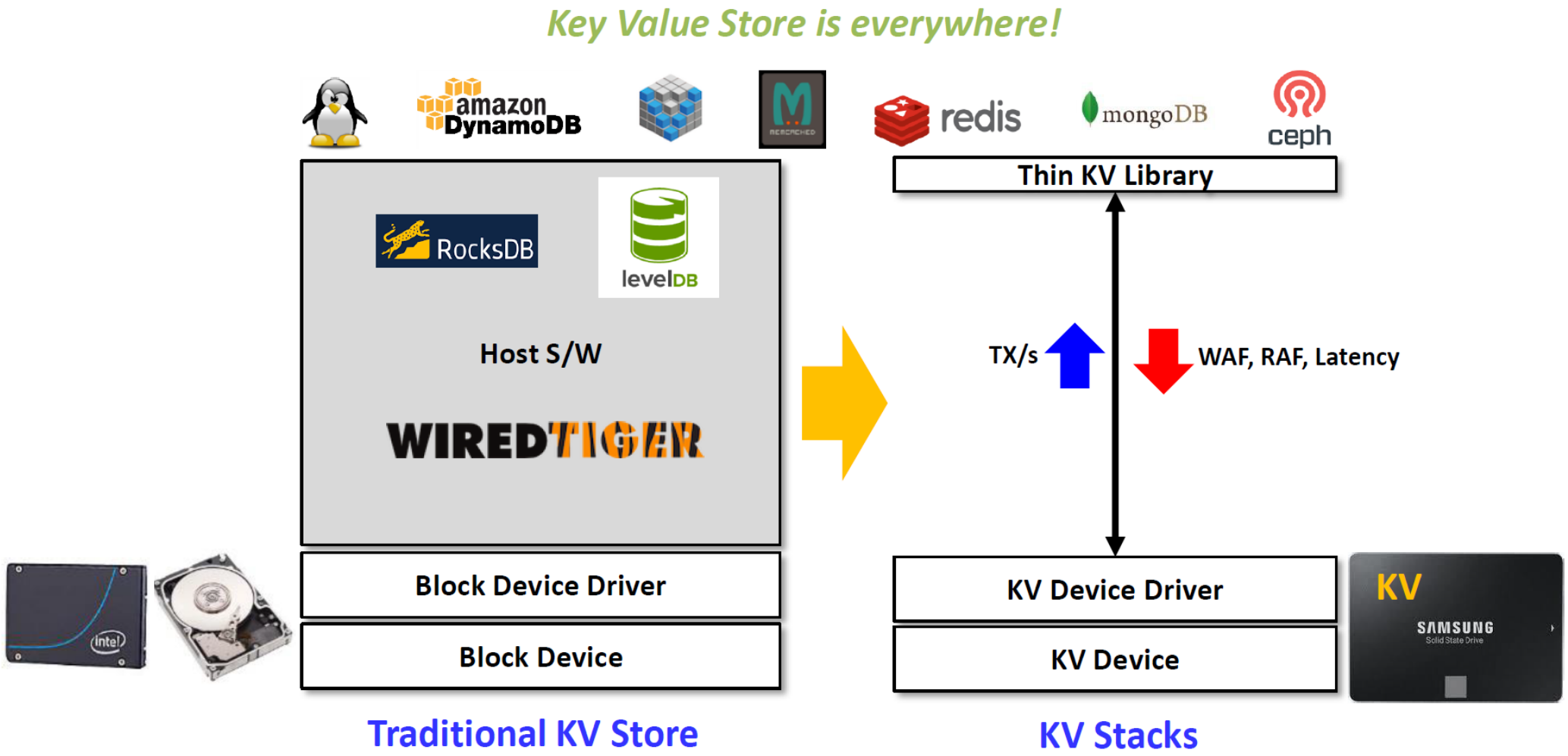
- NVMe Key Value Command Set Specification
  - Revision 1.0c (part of NVMe 2.0 specification, 2022)
  - `Store`, `Retrieve`, `Delete`, `Exist`, `List`
  - Key size: max 16 bytes
  - The List command retrieves a list of KV keys starting at the KV key specified

RocksDB vs. KVSSD

- RocksDB
  - Originated by Facebook and actively used in their infrastructure
  - Most popular embedded NoSQL database
  - Persistent Key-Value store
  - Optimized for fast storage (e.g., SSD)
  - Uses Log-Structured Merge (LSM) Tree architecture

- KV Stacks on KVSSD
  - Benchmark tool directly operates on KVSSD through KV Stacks

Key Idea


Key Value Store is everywhere!
amazon DynamoDB
redis
mongoDB
ceph
RocksDB
levelDB
Host S/W
WIREDTIGER
Block Device Driver
Block Device
Thin KV Library
TX/s
WAF, RAF, Latency
KV Device Driver
KV Device
KV
SAMSUNG
Solid State Drive
Traditional KV Store
KV Stacks

KVSSD Software Stack

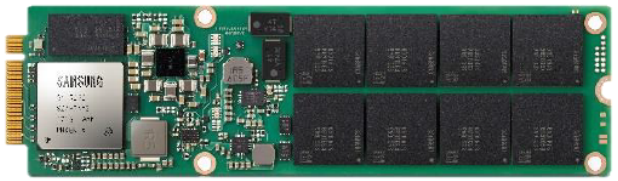
Key Value Library & Tools

Cache | AIO | Multi-Queue | Multi-Device | Memory Manager | Tools

KV Abstract Device Interface (ADI)

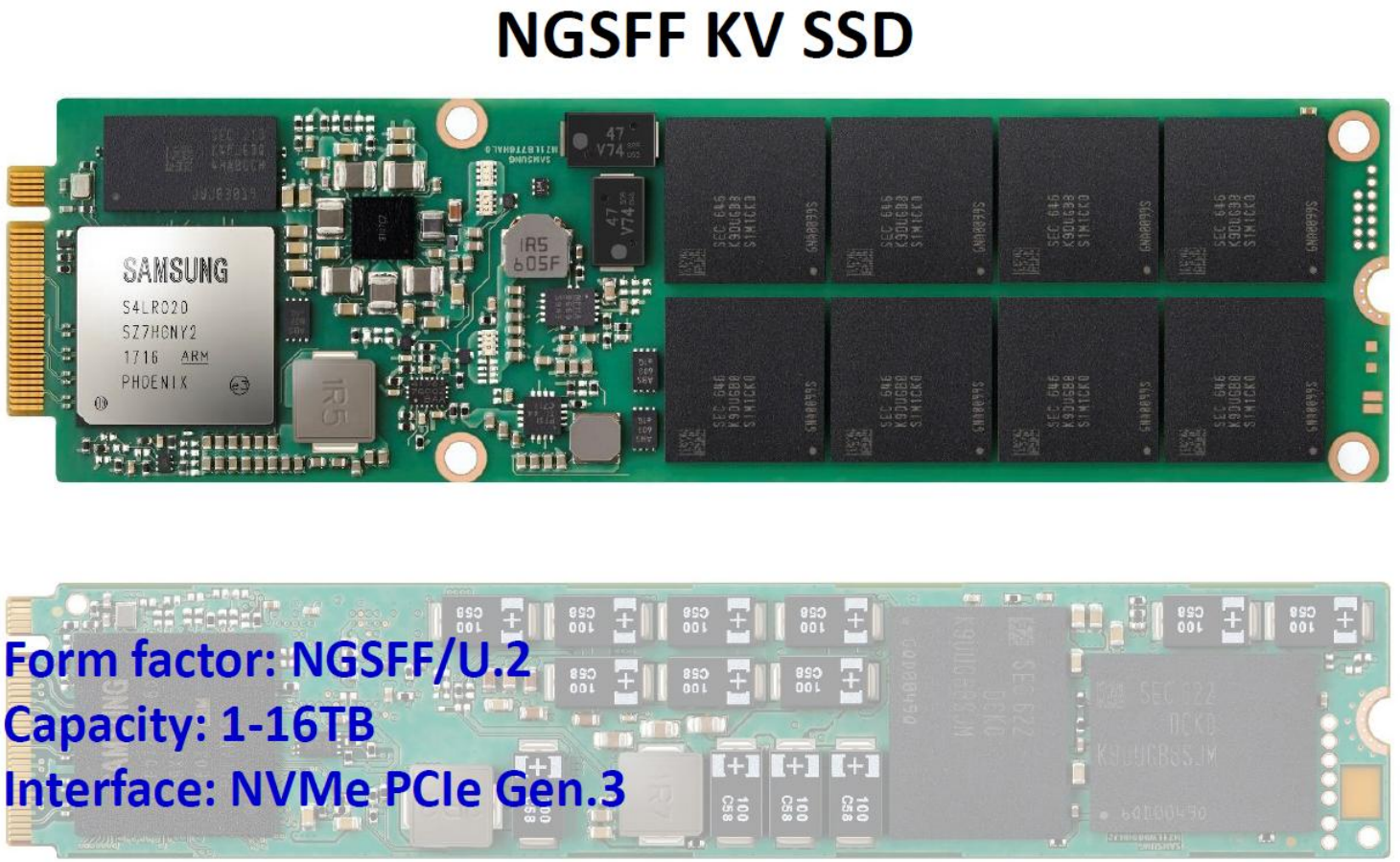
store/retrieve/delete/exist | KV Pair | namespace

Linux Kernel Device Driver | **Linux User-space Device Driver** | **Windows Device Driver**



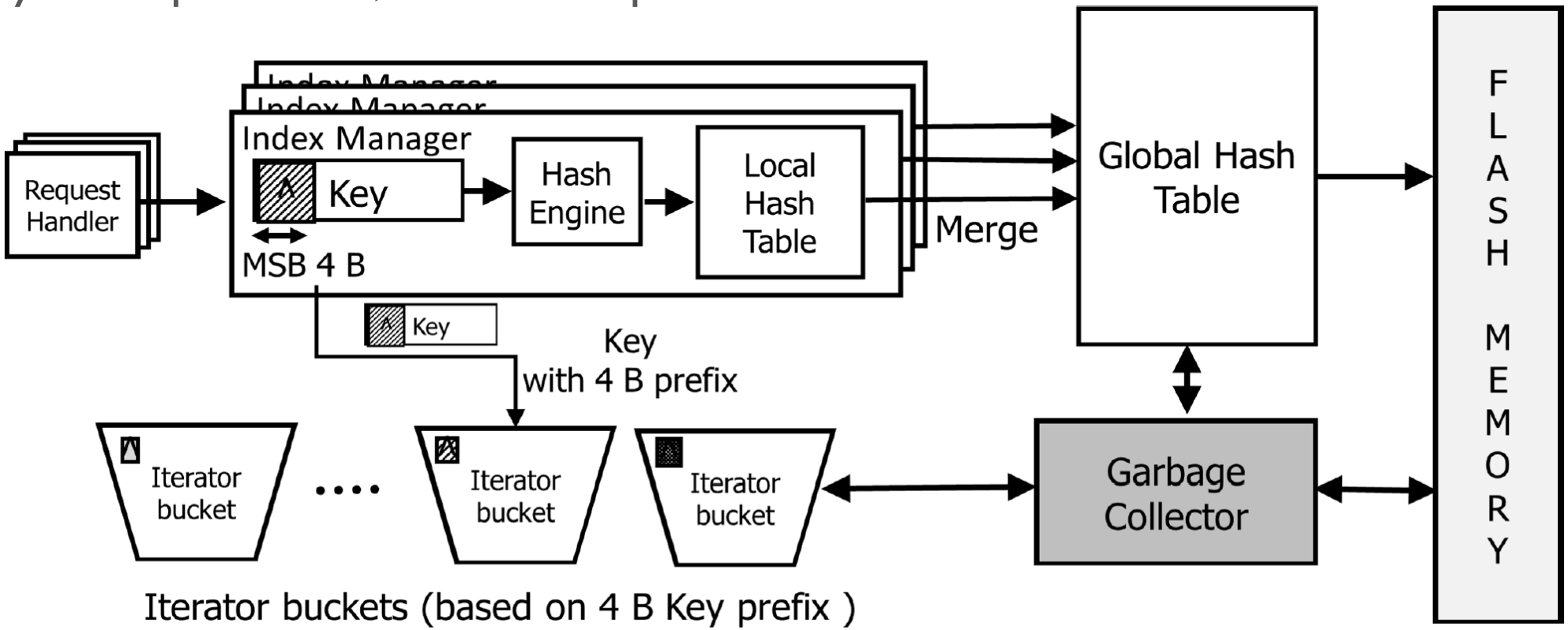
KVSSD Prototype

- Samsung KV-PM983



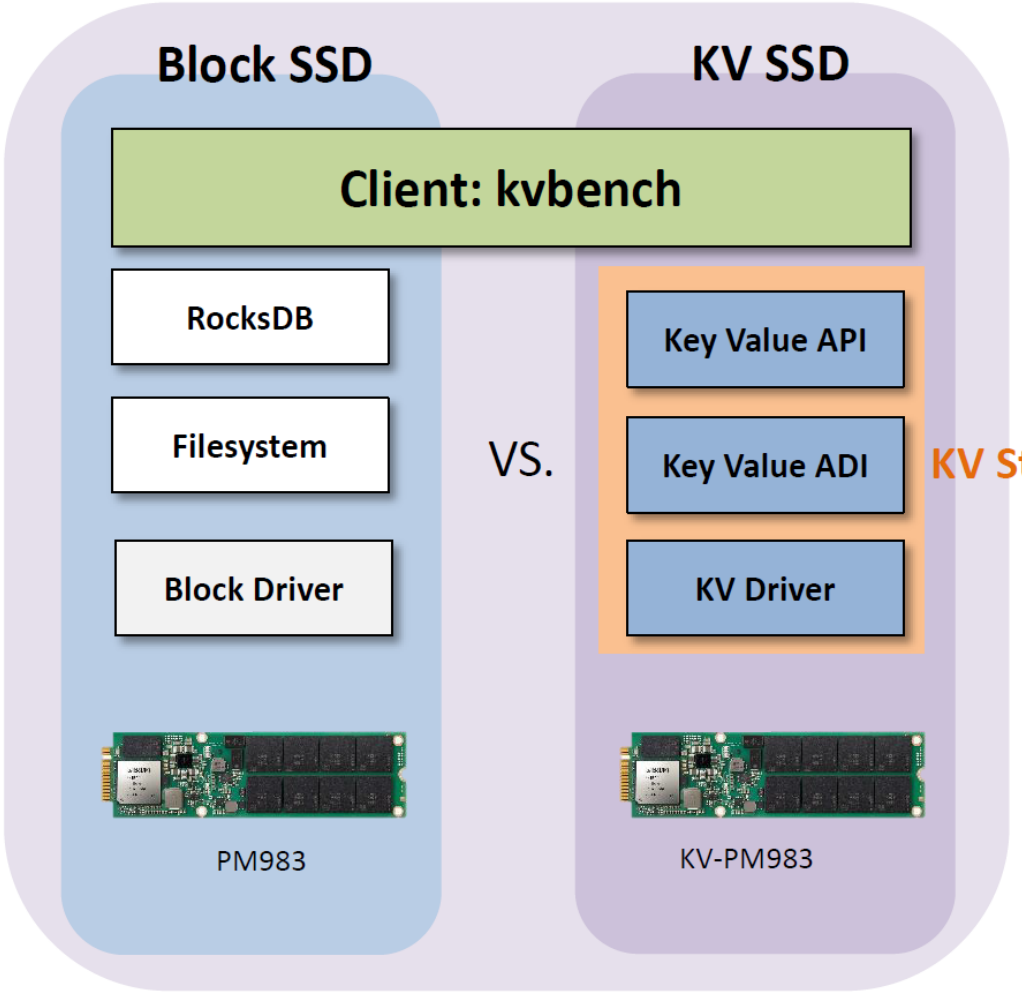
KVSSD Design

- https://github.com/OpenMPDK/KVSSD
  - Key size: up to 255B, Value size: up to 2MB



Source: Y. Kang et al., Towards Building a High-Performance, Scale-In Key-Value Storage System, SYSTOR, 2019.

RocksDB vs. KVSSD: Evaluation

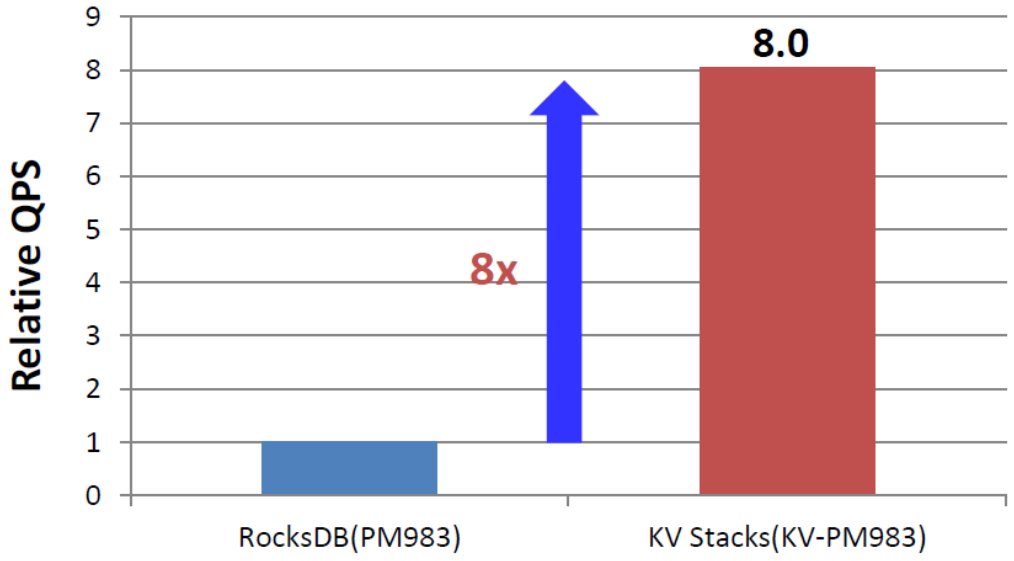


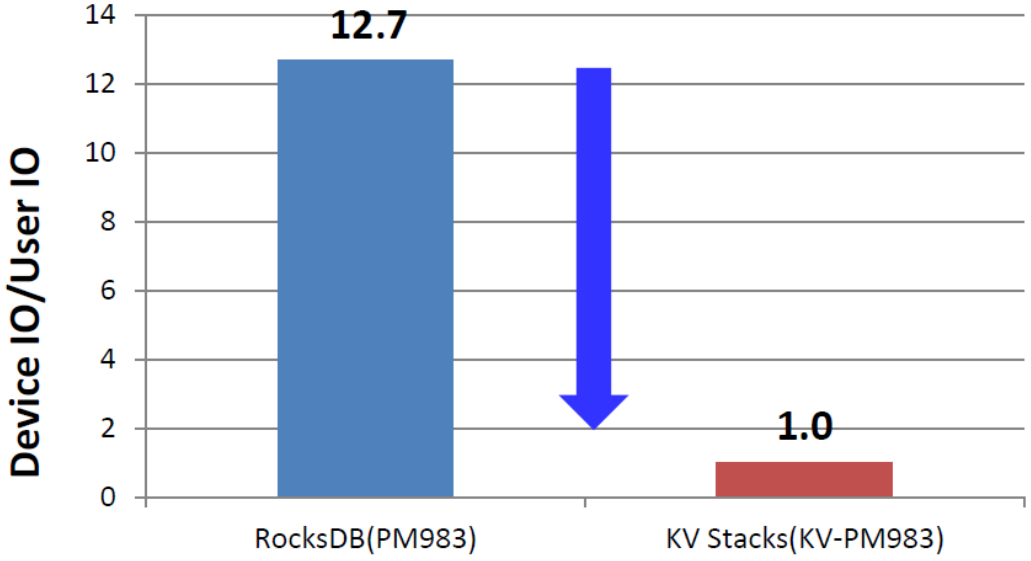
- **Better Performance**
  - Lean software stacks
  - Overhead moved to device
- **IO Efficiency**
  - Reduction of host traffic to devices

Hardware	Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz 96 GB RAM PM983(Block) & KV-PM983 SSD
Software	Ubuntu 16.04 RocksDB v5.0.2 on XFS 50M records, 16B Key, 4KB value

RocksDB vs. KVSSD: Results

- Random PUT performance
  - 8x more QPS (Query Per Second) with KV Stacks than RocksDB on block SSD
  - 90+% less traffic goes from host to device with KV SSD than RocksDB on block device





** Workload: 100% random put, 16-byte keys of random uniform distribution, 4KB-fixed values on single PM983 and KV-PM983 in a clean state*

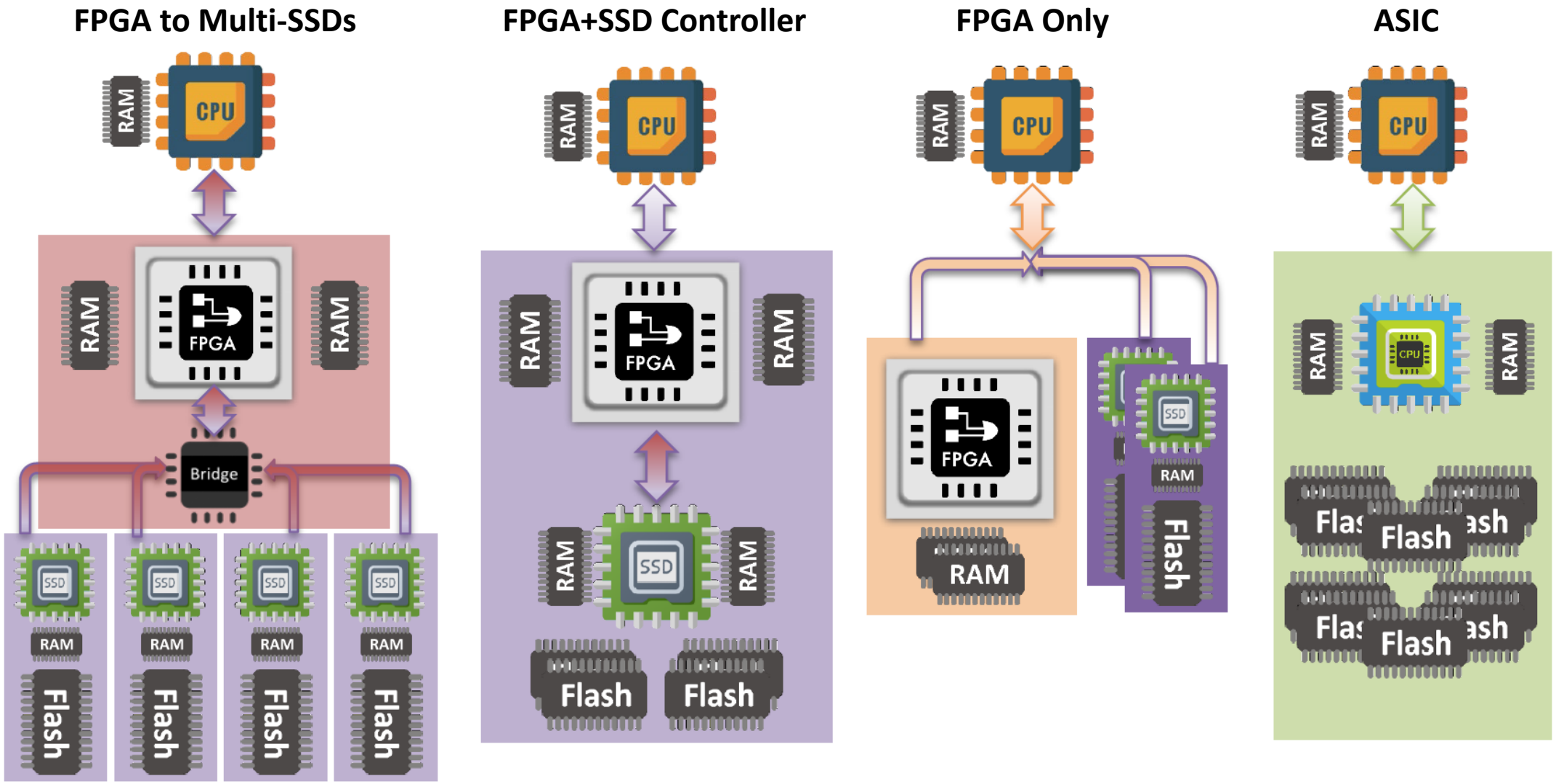
Summary: KV-SSD

- ***What's good?***

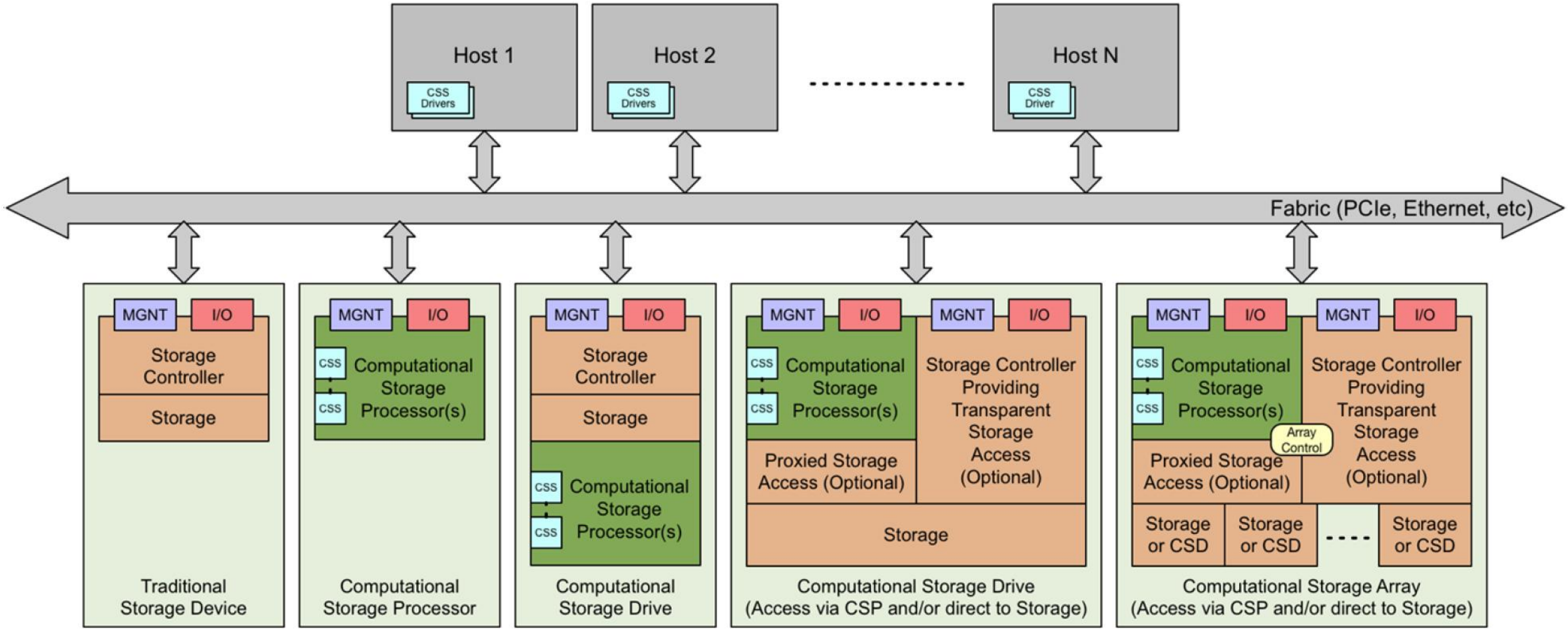
- ***What's bad?***

Computational Storage

SNIA: Computational Storage Instances

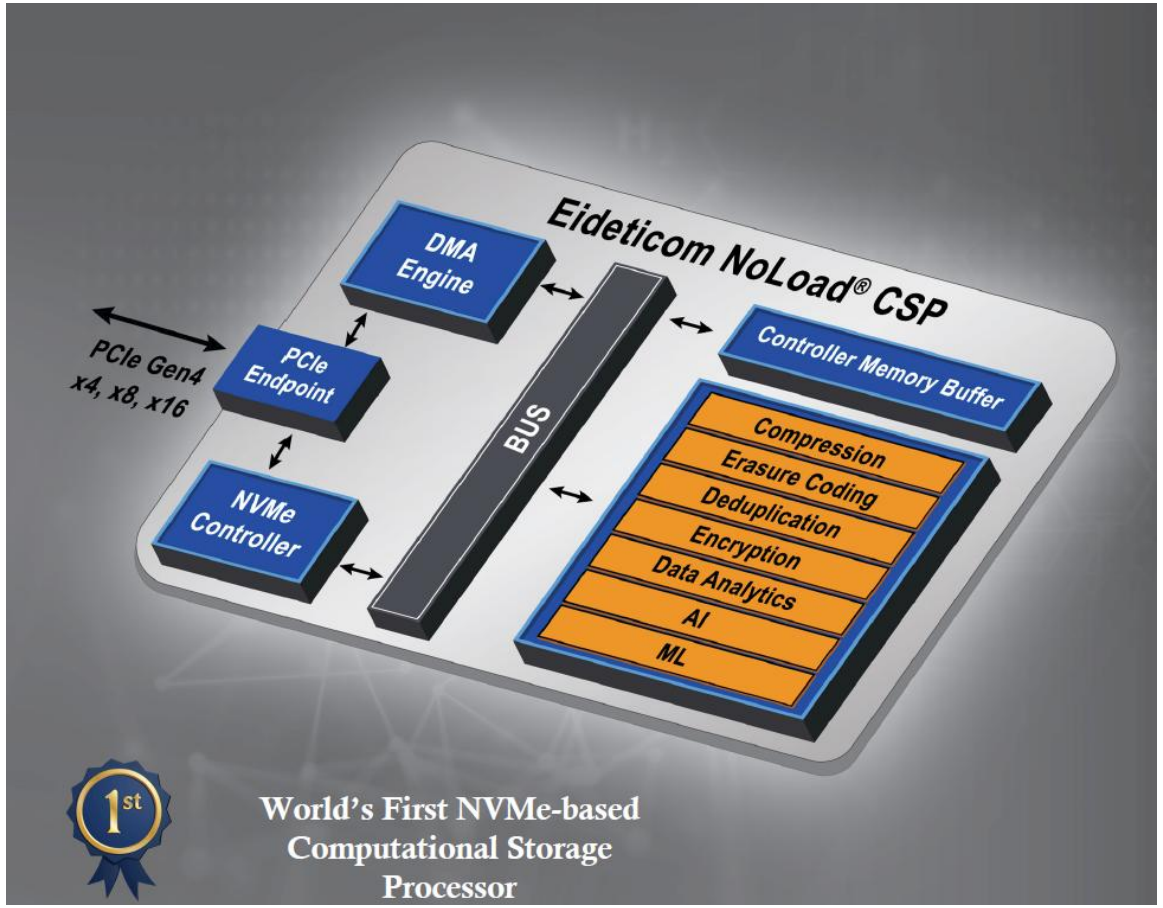


SNIA: Computational Storage Devices (CSx)

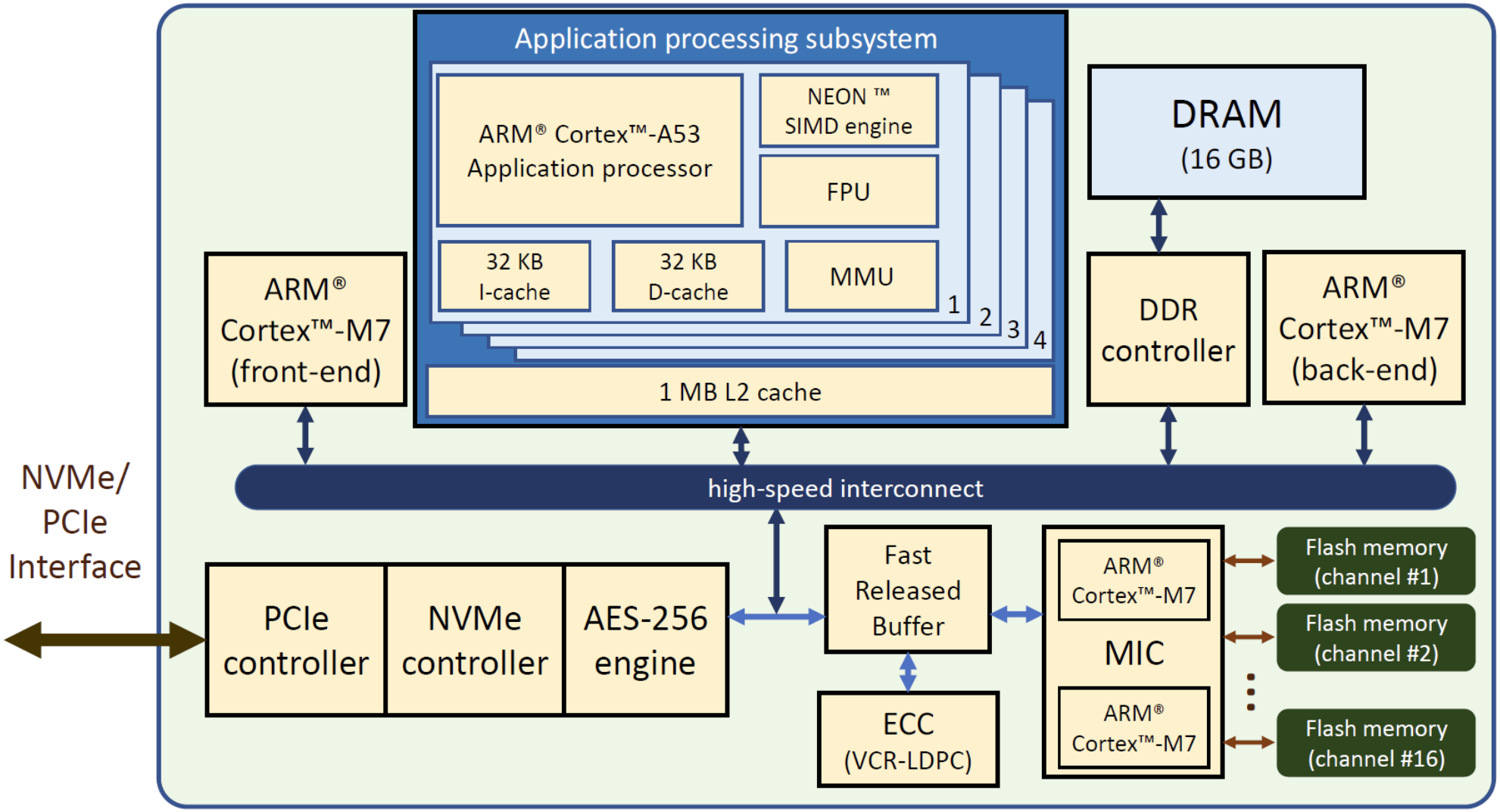


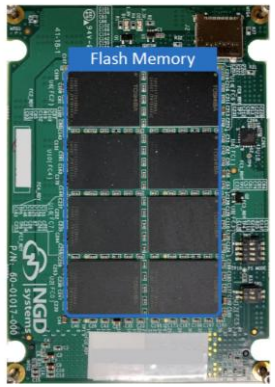
Eideticom NoLoad CSP

- NVMe computational accelerators
  - Compression
  - Encryption
  - Erasure coding
  - Deduplication
  - Data analytics
  - AI and ML



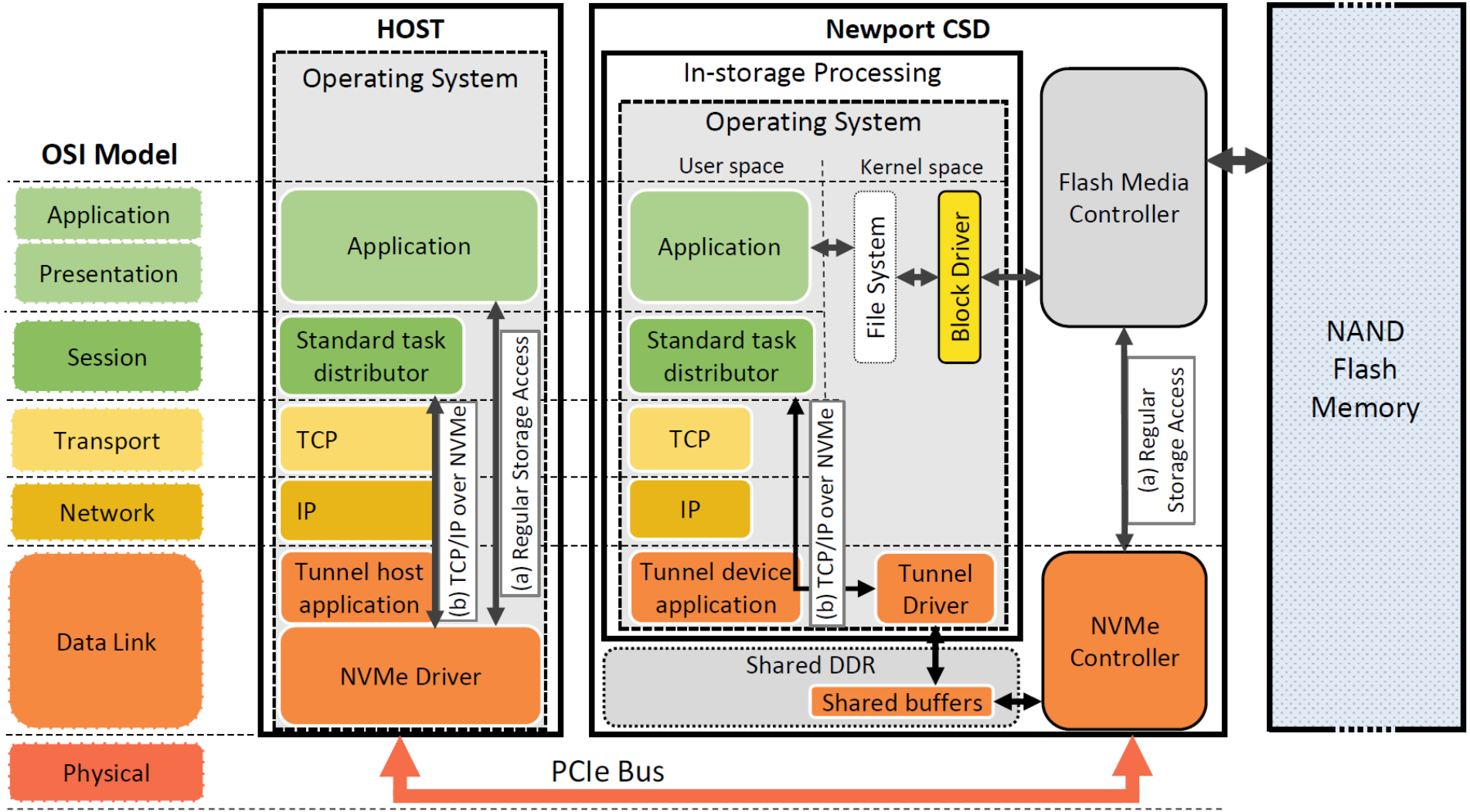
NGD Newport SSD





Source: J. Do et al., "Cost-effective, Energy-efficient, and Scalable Storage Computing for Large-scale AI Applications, ACM ToS, 2020.

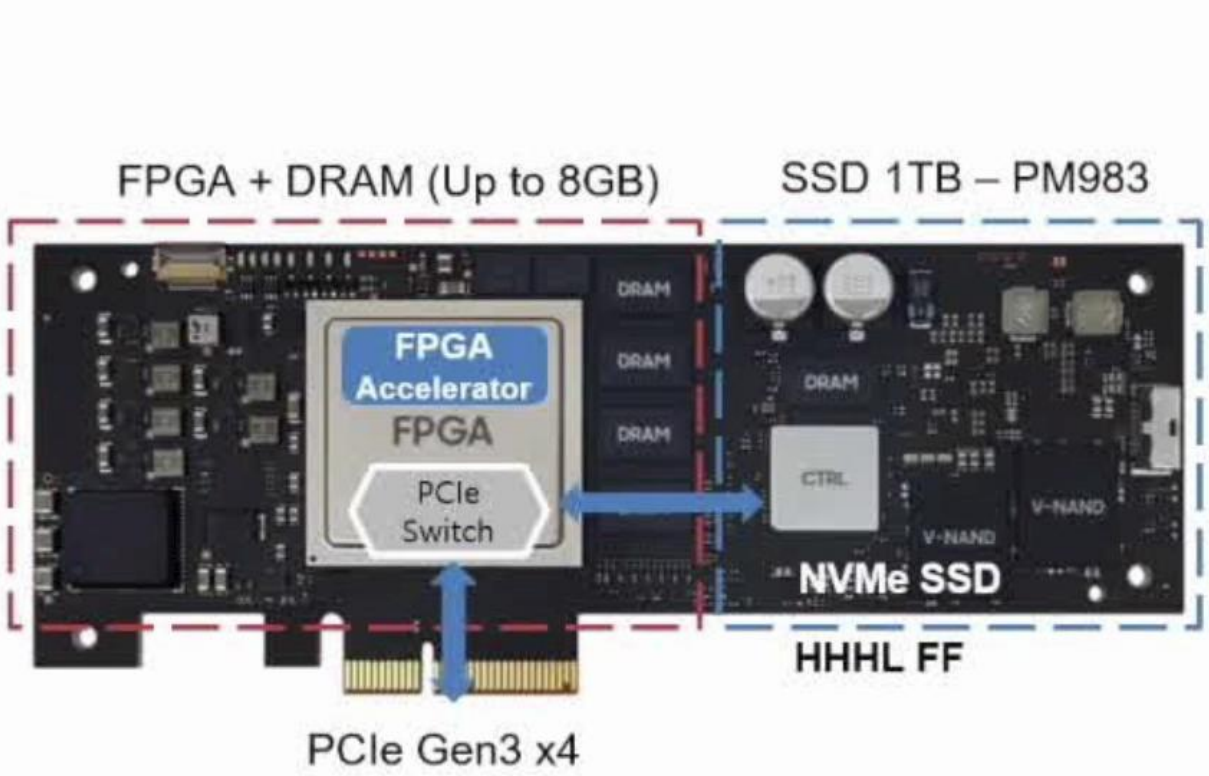
NGD Newport SSD: Software Stack

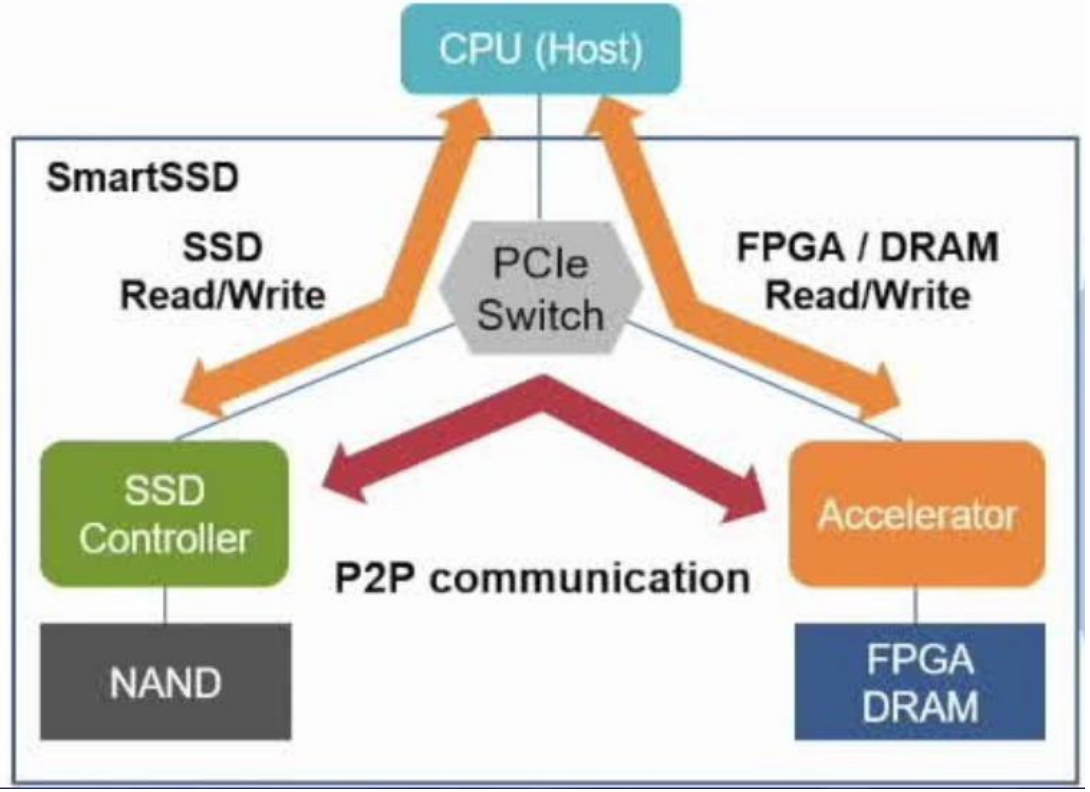


Source: J. Do et al., "Cost-effective, Energy-efficient, and Scalable Storage Computing for Large-scale AI Applications, ACM ToS, 2020.

Samsung SmartSSD 1.0

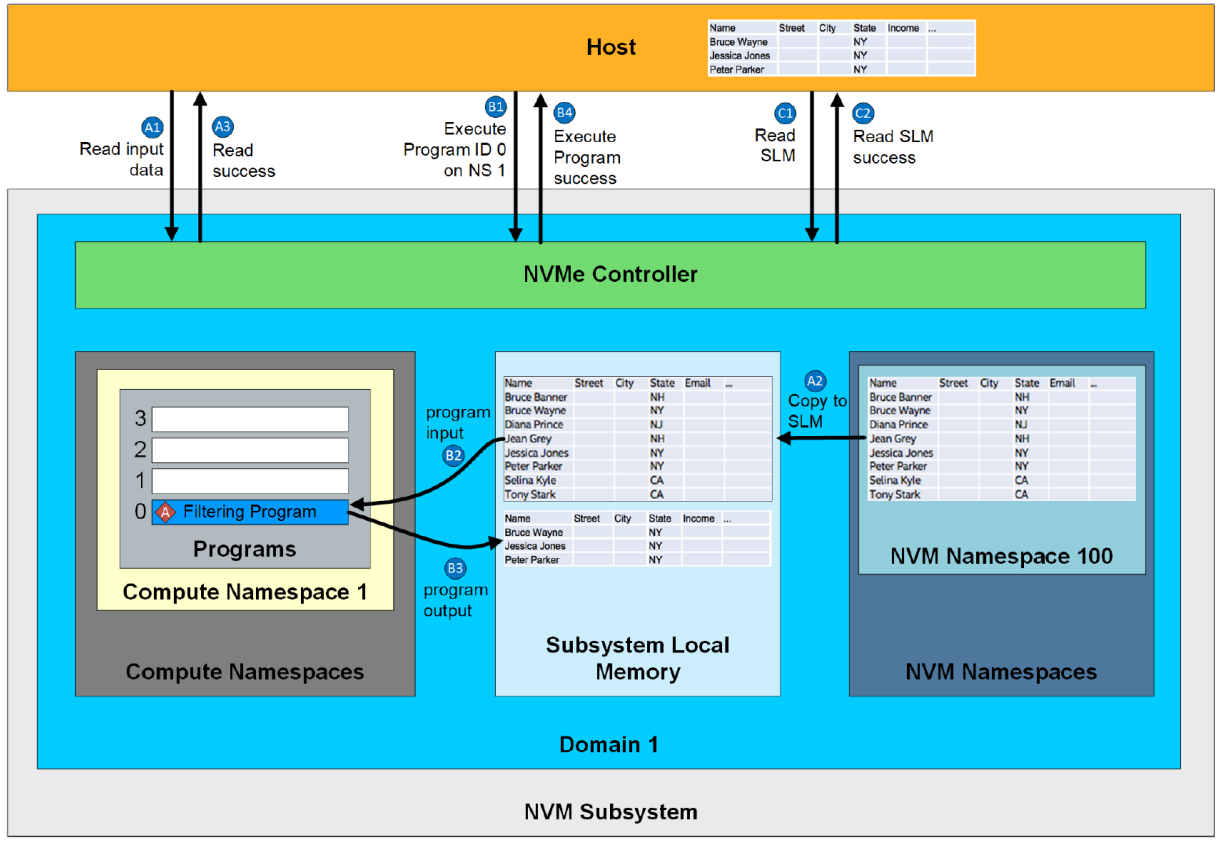
- PM983F
- Xilinx FPGA





NVMe TP4091

- NVMe Computational Storage Architecture
  - Compute Namespaces
  - Subsystem Local Memory (SLM)
  - NVM Namespaces
- Commands
  - Load program
  - Execute program
  - Activate program, ...
- Programs operate on SLM
  - Fixed function programs
  - Downloadable eBPF programs



Towards Storage Intelligence