Offloading xPU Storage Compute Tasks to SSD

Mahinder Saluja
Director of Technology and Storage Pathfinding, SSD BU
KIOXIA America, Inc.

FMS
the Future of Memory and Storage

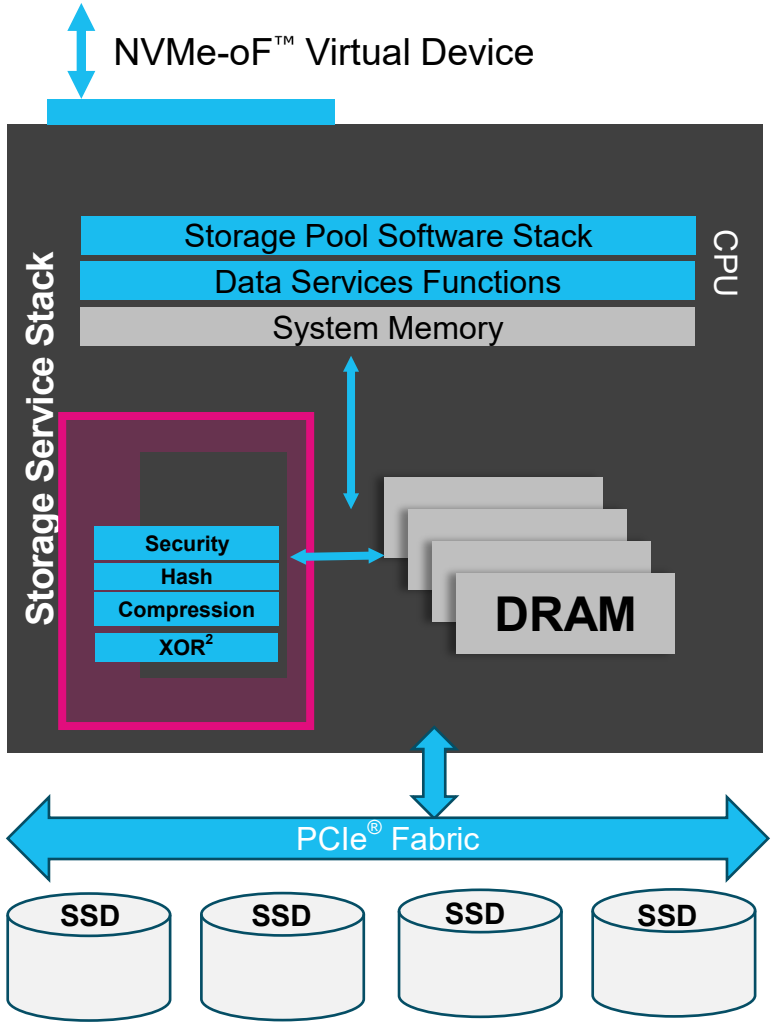
KIOXIA

Agenda

- xPUs in Storage Systems
- Storage System Data Services Overview
- xPUs to Leverage SSD Offloads
- PoC Results
- Summary

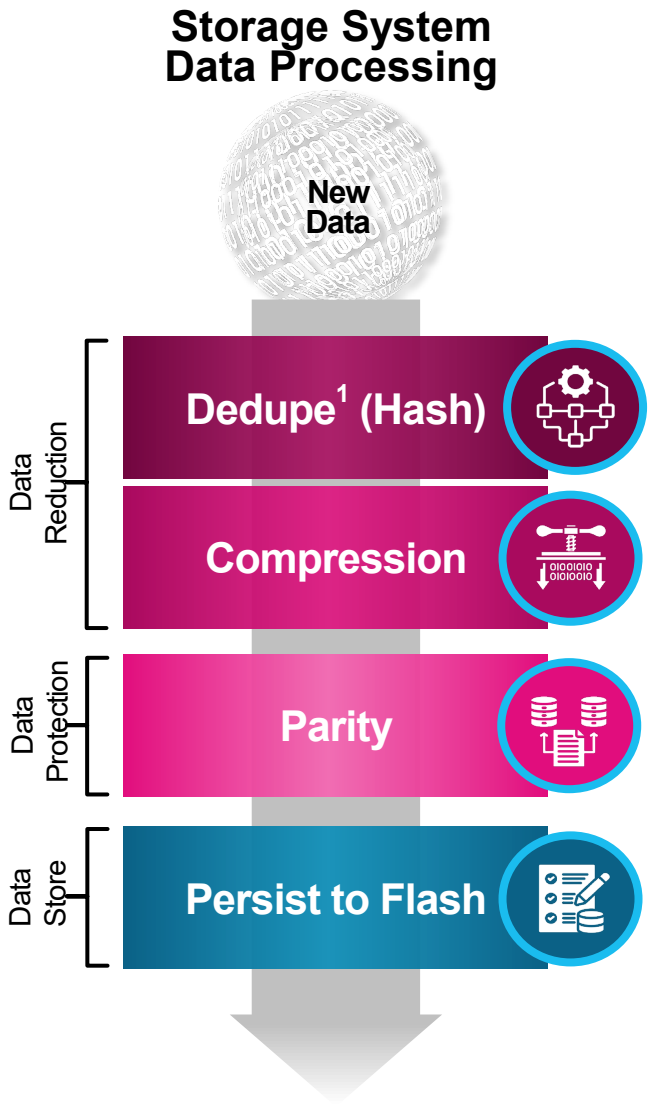
Role of xPUs in Storage Systems

- xPU[1]s are now essential for offloading data services, networking, security and virtualization functions
- Data services like compression, hashing (deduplication) and parity are compute-intensive storage functions
- xPUs risk becoming the next bottleneck with increasing:
  - xPU bandwidth is PCIe® lanes limited
  - Fabric bandwidth: 50 GB/s → 200 GB/s (400 Gbps -> 1.6 Tbps)
  - SSD throughput: ~16 GB/s → ~32 GB/s (with PCIe Gen evolution)
- Proposing an SSD-based, fixed-function compute offload architecture that:
  - Offloads compute intensive data services from xPUs and traditional CPUs
  - Enhances scalability, efficiency, and performance



1. An auxiliary processing unit (xPU) can refer to a Data Processing Unit (DPU) or other processing units that offload and accelerate specialized tasks more efficiently than a general-purpose CPU. 2. Exclusively-OR (XOR). PCIe is a registered trademark of PCI-SIG. NVMe-oF is a trademark of NVM Express, Inc. in the United States and other countries. Images and/or graphics within this slide are the property of KIOXIA America, Inc. (KAI) and are reproduced with the permission of KAI.

Storage System Data Services Overview



- **Data ingestion in a storage system involves a series of compute intensive operations**
- **With increased SSD performance, data management tasks require increased processing power usually solved by:**
  - *General purpose CPU cores*
    - Current compute architecture and memory hierarchy increase total cost of acquisition and total cost of ownership
    - Overprovisioning cores, CPU caches and memory
  - *Accelerators and DPU[2]s*
    - Performance limited by number of PCIe® lanes assigned or system memory
    - Does not scale with added SSDs
    - Consumes additional power

KIOXIA

1. Deduplication (dedupe) hashes are unique identifiers generated from data blocks using hashing algorithms to identify duplicate data within a system. 2. Data Processing Unit (DPU). All images and/or graphics within this slide are the property of KIOXIA America, Inc. (KAI) and are reproduced with the permission of KAI. PCIe is a registered trademark of PCI-SIG

Host Orchestrated Compute Offload Building Blocks

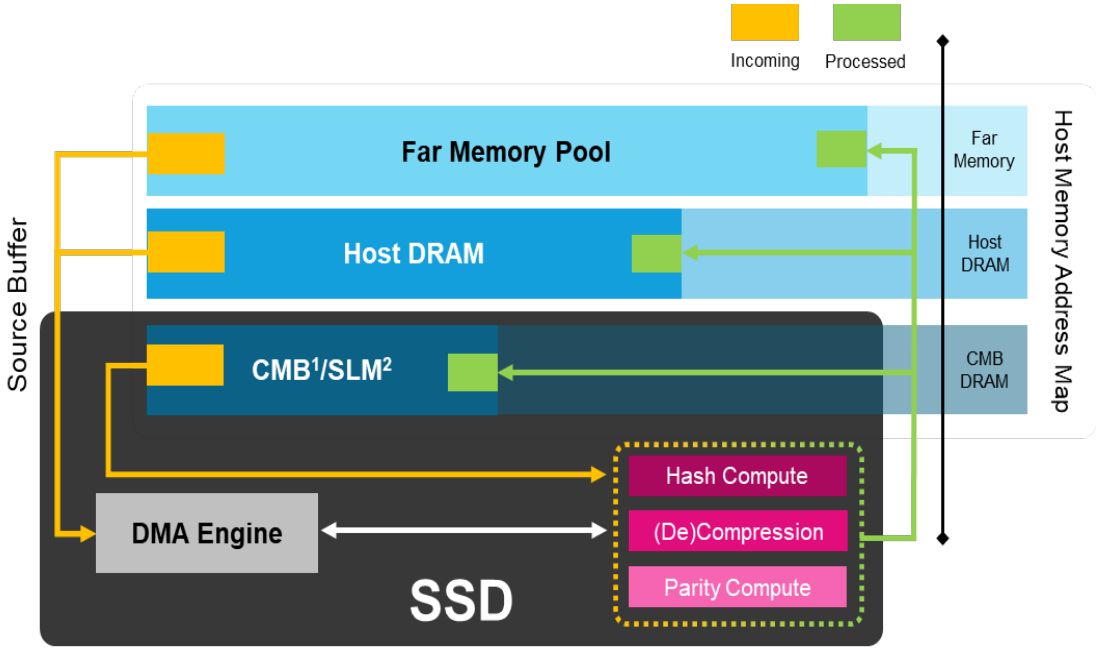
Power-efficient compute engines

DRAM bandwidth saving

Host orchestrated standard based

Applications of Offload Engines

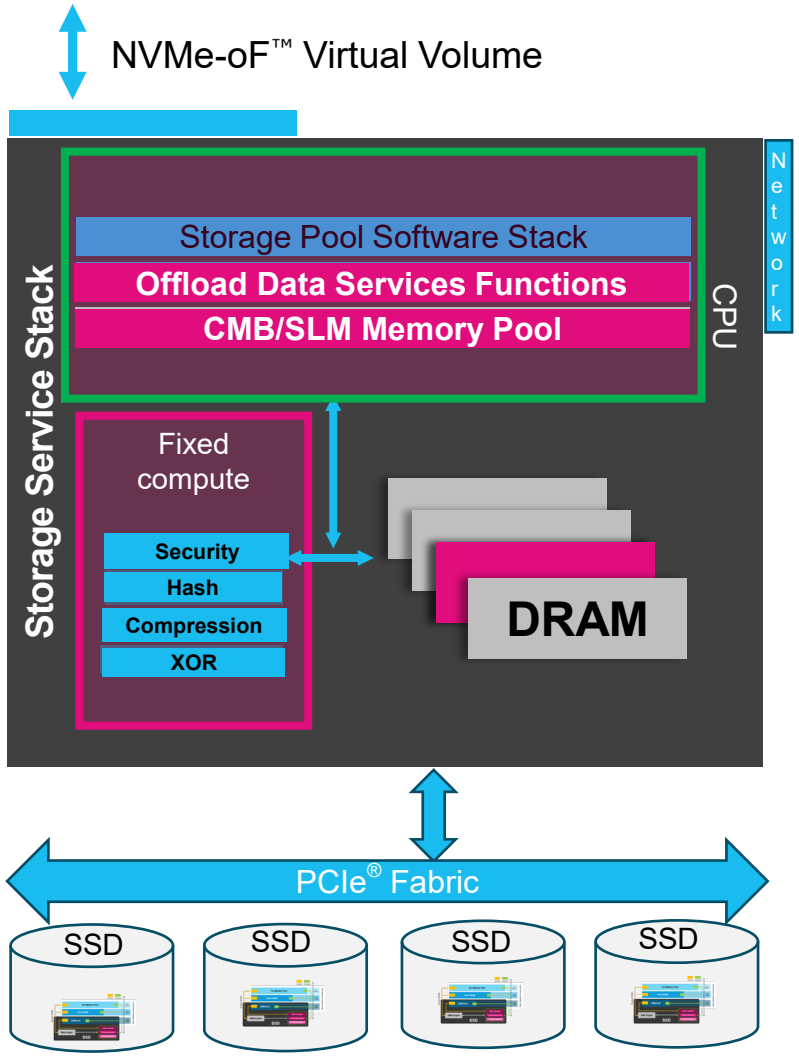
- **Hash/CRC[3]:** Dedupe, Object/File signature/scrubbing, buffer integrity
- **(De)Compression:** Compression with levels, decompress and filter
- **Parity Compute**: Erasure code (EC), compare, Data scrubbing, RAID Rebuild



Host places data in appropriate buffer (host memory/ far memory/ CMB[1]) → Host issues command to the SSD → Compute engine on the SSD executes the command → Processed data is placed into the destination buffer → Command completion returns key attributes

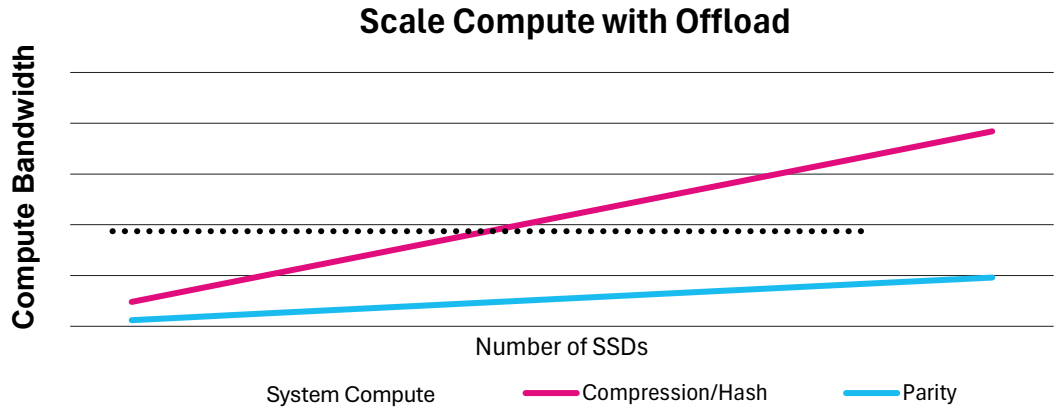
1. Controller memory buffer (CMB). 2. Subsystem local memory (SLM). 3. Cyclic redundancy check (CRC). All images and/or graphics within this slide are the property of KIOXIA America, Inc. (KAI) and are reproduced with the permission of KAI.

xPUs Leverage SSD Offload

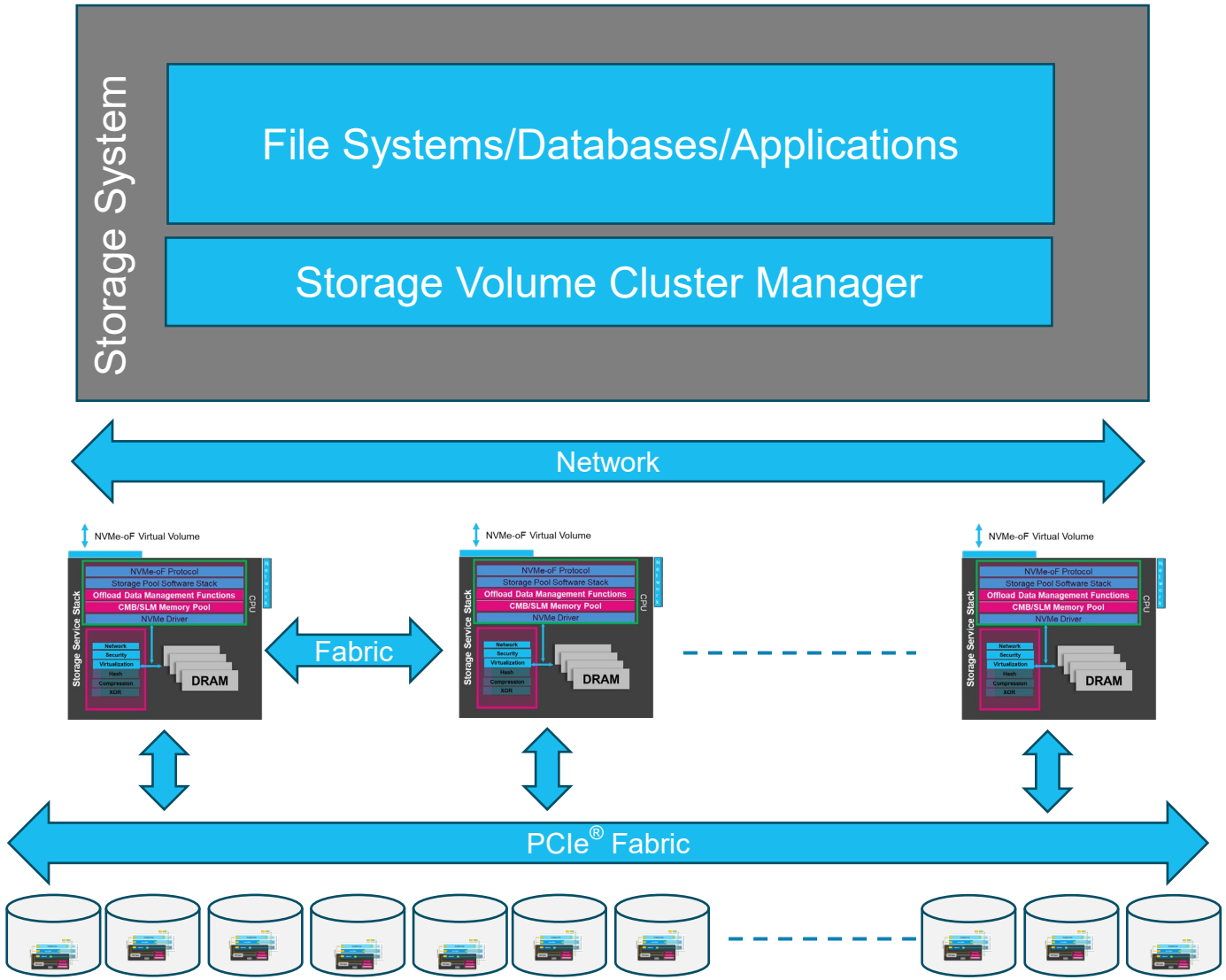


- Offload xPU compute and memory to SSD
- xPU orchestrate offload data management compute functions on SSD
- Storage pool continue to run on xPU
- xPU scale storage functions linearly with added SSDs

- **With Offload Capable 24 SSDs** (for illustration, not official specs.)
  - CMB/SLM memory bandwidth @ 10 GB/s/ SSD :240 GB/s
  - Parity compute bandwidth @ 2 GB/s/ SSD : 48 GB/s
  - (De)compression @ 10 GB/s :240 GB/s



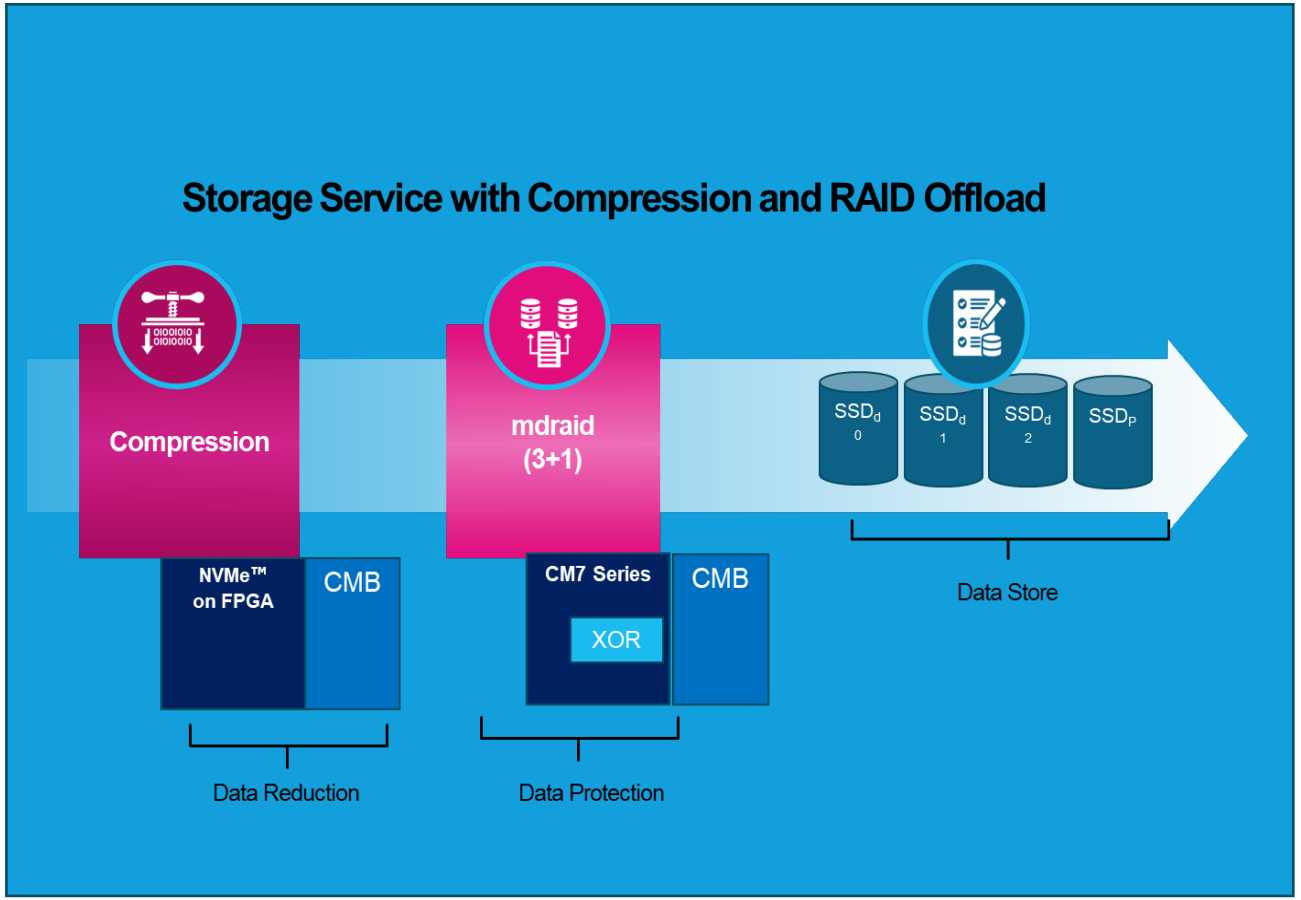
Storage System with DPU Offload to SSD



- More resources for storage systems to value-add services like data vectorization, etc.
- Scale performance and capacity with every SSD without additional compute
- xPU teams up with SSD offload paving the way for headless storage systems[1]

1. A headless storage system is a solution where the user interface (UI) or frontend is separated from the backend, which handles the storage and management of data.

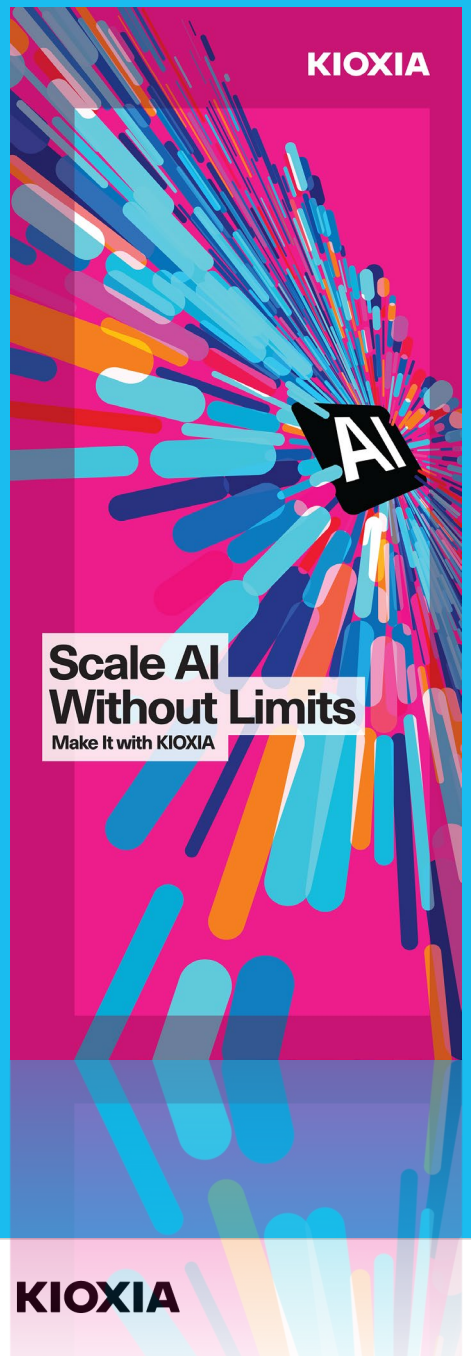
Results from Data Pipeline Offload Proof of Concept (on Host with CPU)



No Offload (CPU) vs Offload

	No Offload	Offload	% Benefit
Write Bandwidth	~140 MB/s	~140 MB/s	-
Compression Ratio	2.5x	2.5x	-
Compression (gzip) CPU Core	200%	~1%	~199%
RAID CPU Resources	4%	4%	-
DRAM Bandwidth	~600 MB/s	~160 MB/s	3.8x

Summary



- xPUs risk becoming the next bottleneck with increased SSD and network performance
- SSD offload scales linearly with every added SSD into the cluster
- xPU teams-up with SSD offload, paving the way for headless storage systems

See the offload proof of concept demonstration at the KIOXIA booth!

KIOXIA booth #307

KIOXIA