Feeding the Beast

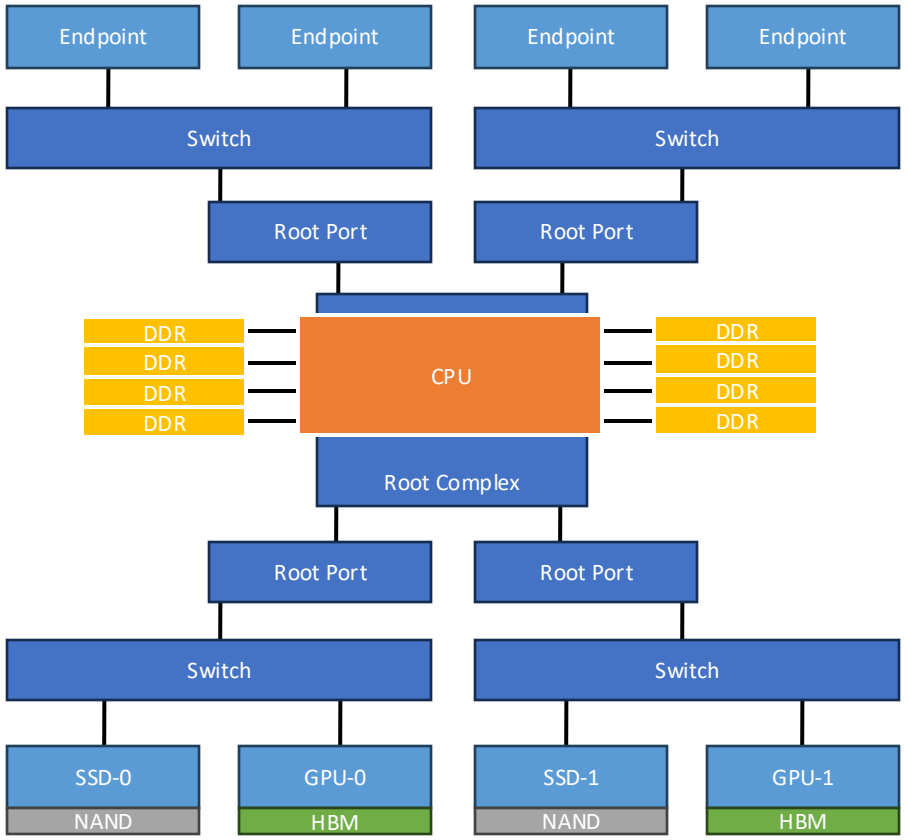
Bridging **NVMe** Storage and **GPUs** while Preserving **File** Semantics

Simon A. F. Lund | Principal Engineer Samsung Semiconductor

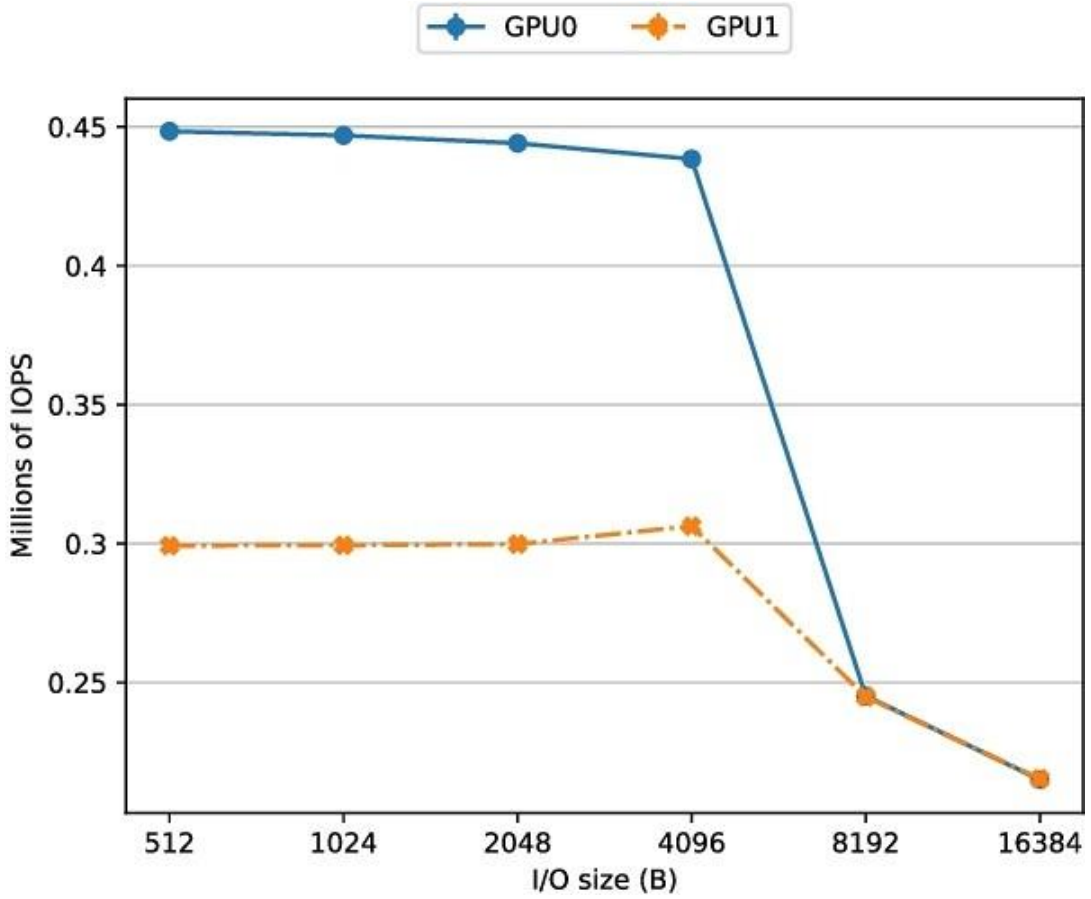
Systems Architecture: context

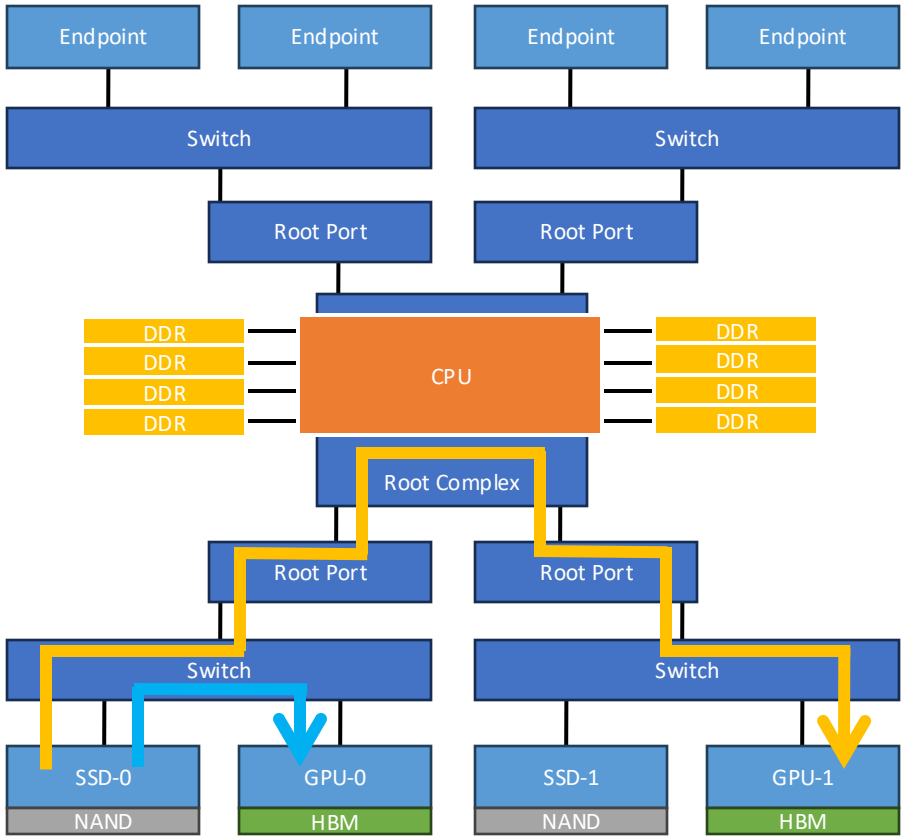
- **NVMe** Storage with **NAND** media
  - **PCIe**, TCP, RDMA/RoCe, Infiniband
  - **128** to **8K** Queues per controller
- **CPU** with **DDR** memory
  - **128** Threads
- **GPU** with **HBM** memory
  - **100K+** Threads

BW: NAND < DDR < HBM



Challenges: data access latency **HW**



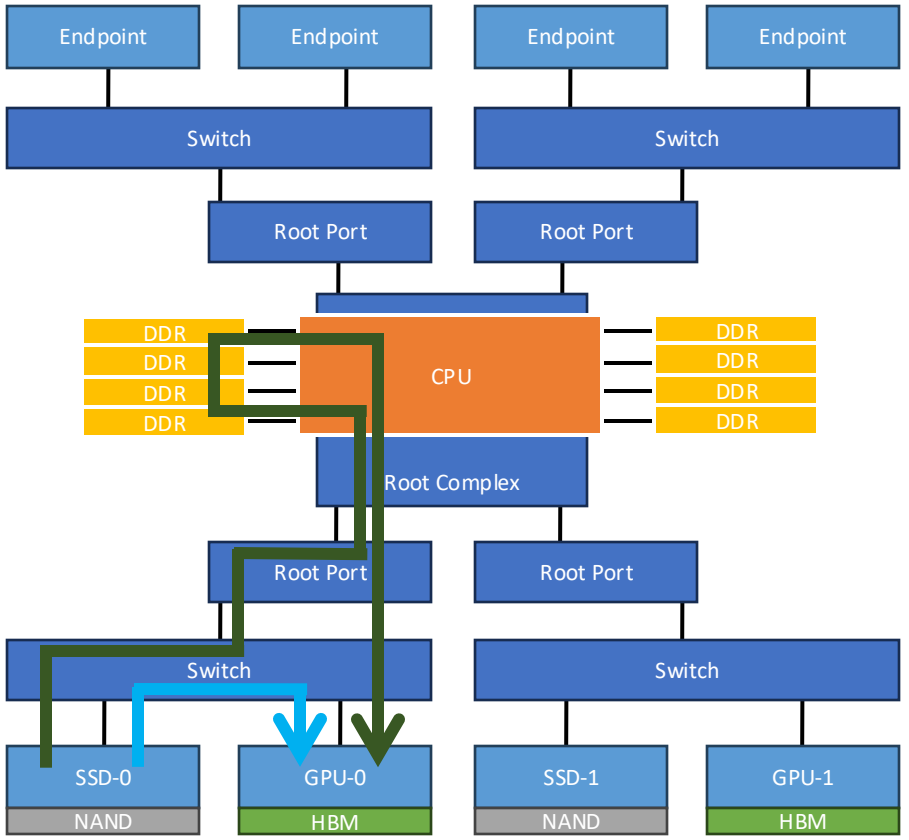


Challenges: data access latency **SW**

- Databases
  - SQL, KV, Vector, Object, … , Time Series

  ➔Data query interfaces

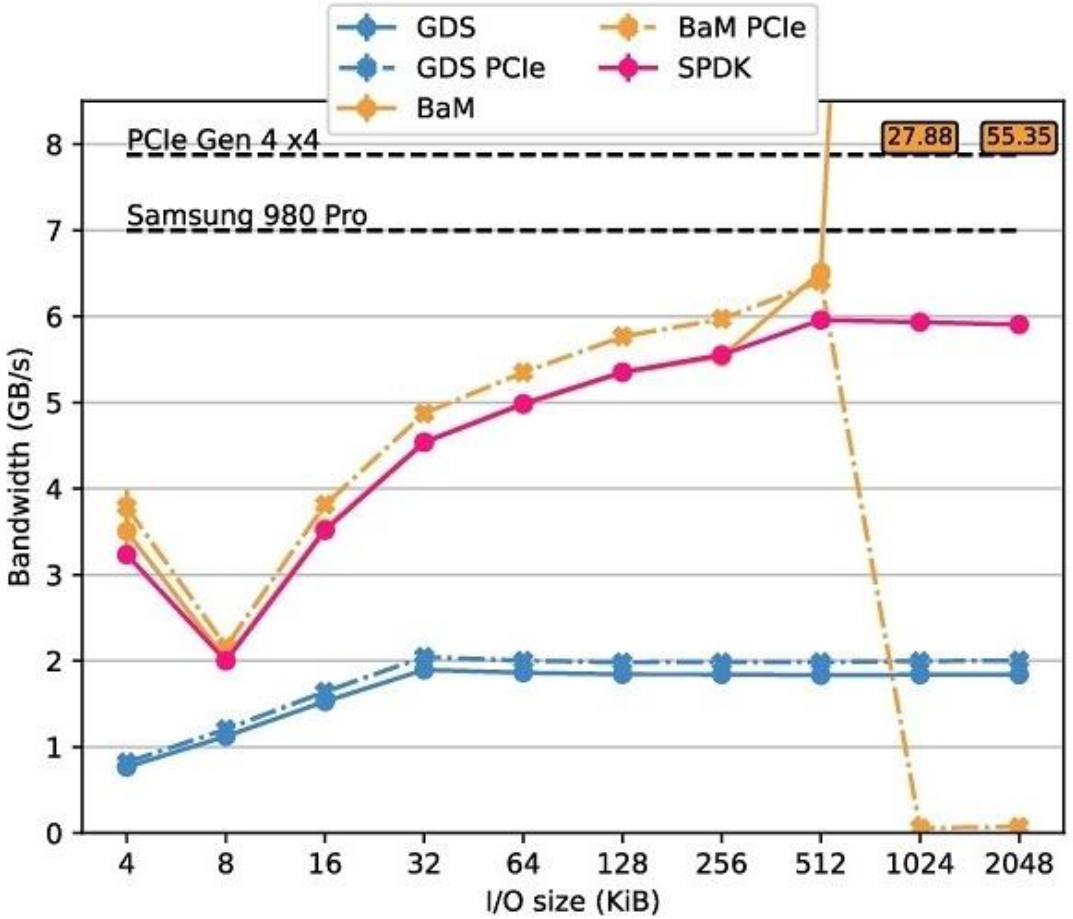
- Files & File-systems
  - File-formats and layout
  - Block allocation and management
- Blocks and Block Devices
  - Storage device and media abstraction



Challenges: state of the art

Benchmark tools:
gdsio, bdevperf, nvm-block-bench

- SPDK and BaM
  - + Optimal performance
  - **- Files / file-system**
- GDS
  - - Subpar performance
  - **+ Files / file-system**



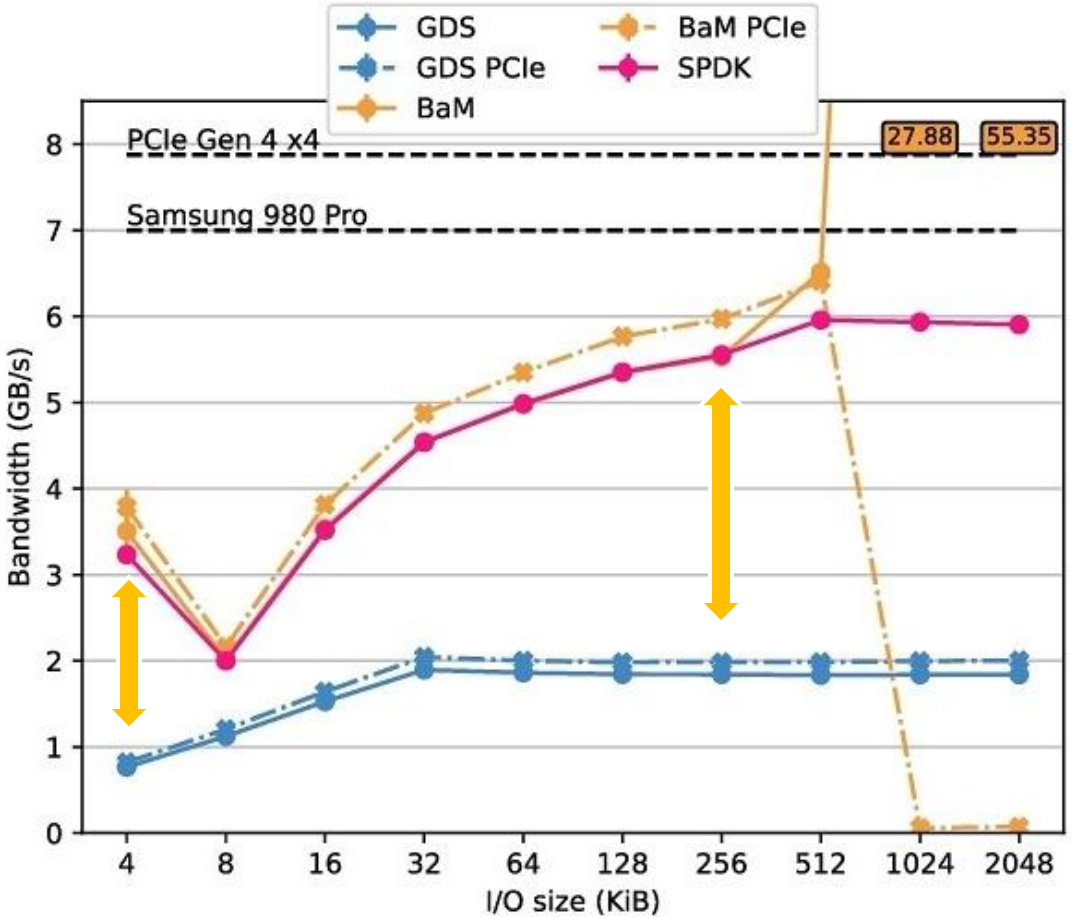
Path to GPU-Initiated I/O for Data-Intensive Systems
https://dl.acm.org/doi/10.1145/3736227.3736232
Paper at SIGMOD / DaMoN25

Challenges: state of the art

Benchmark tools:
gdsio, bdevperf, nvm-block-bench

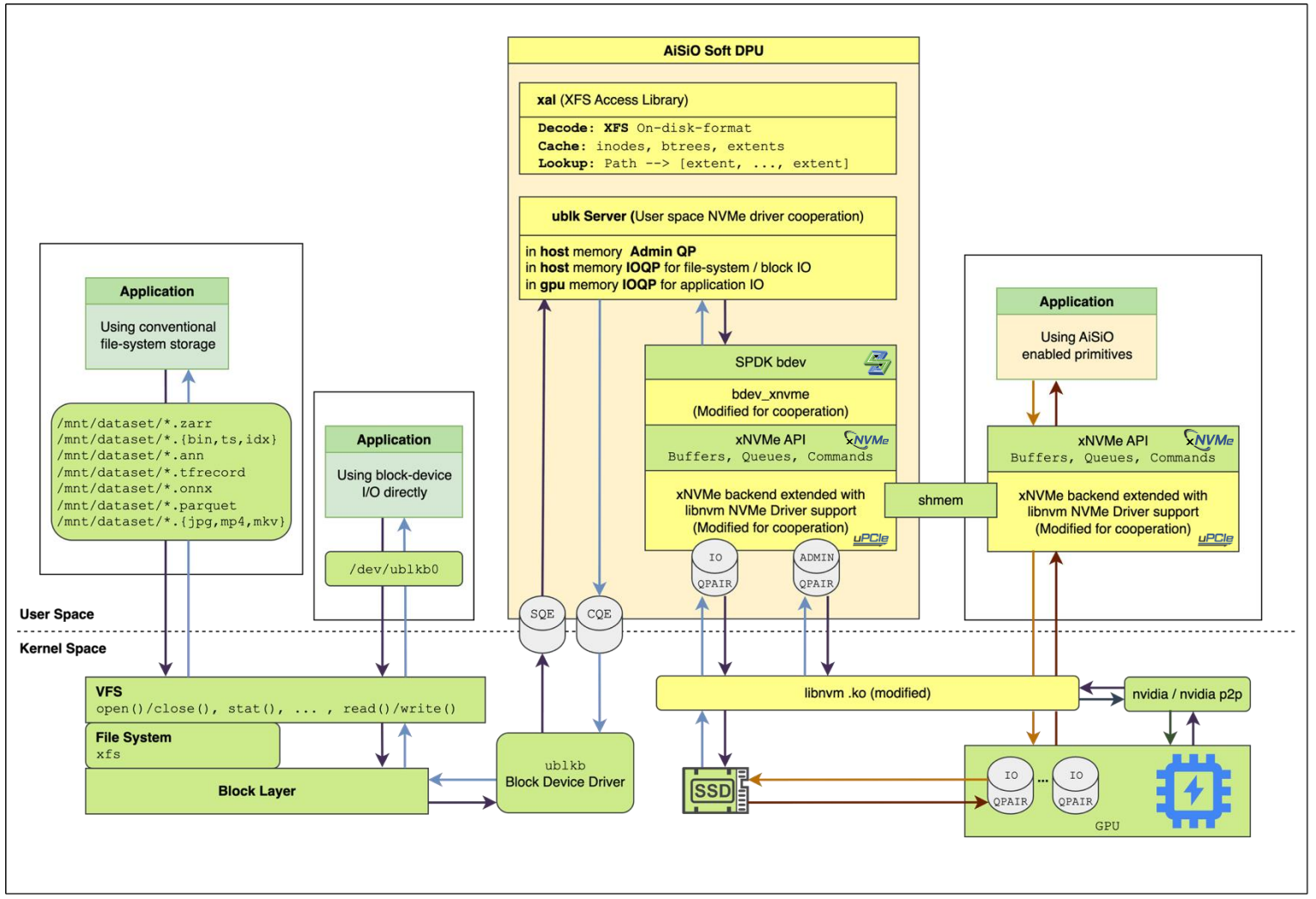
- SPDK and BaM
  - + Optimal performance
  - **- Files / file-system**
- GDS
  - - Subpar performance
  - **+ Files / file-system**

Goal ➔ Close the perf. gap!



Path to GPU-Initiated I/O for Data-Intensive Systems
https://dl.acm.org/doi/10.1145/3736227.3736232
Paper at SIGMOD / DaMoN25

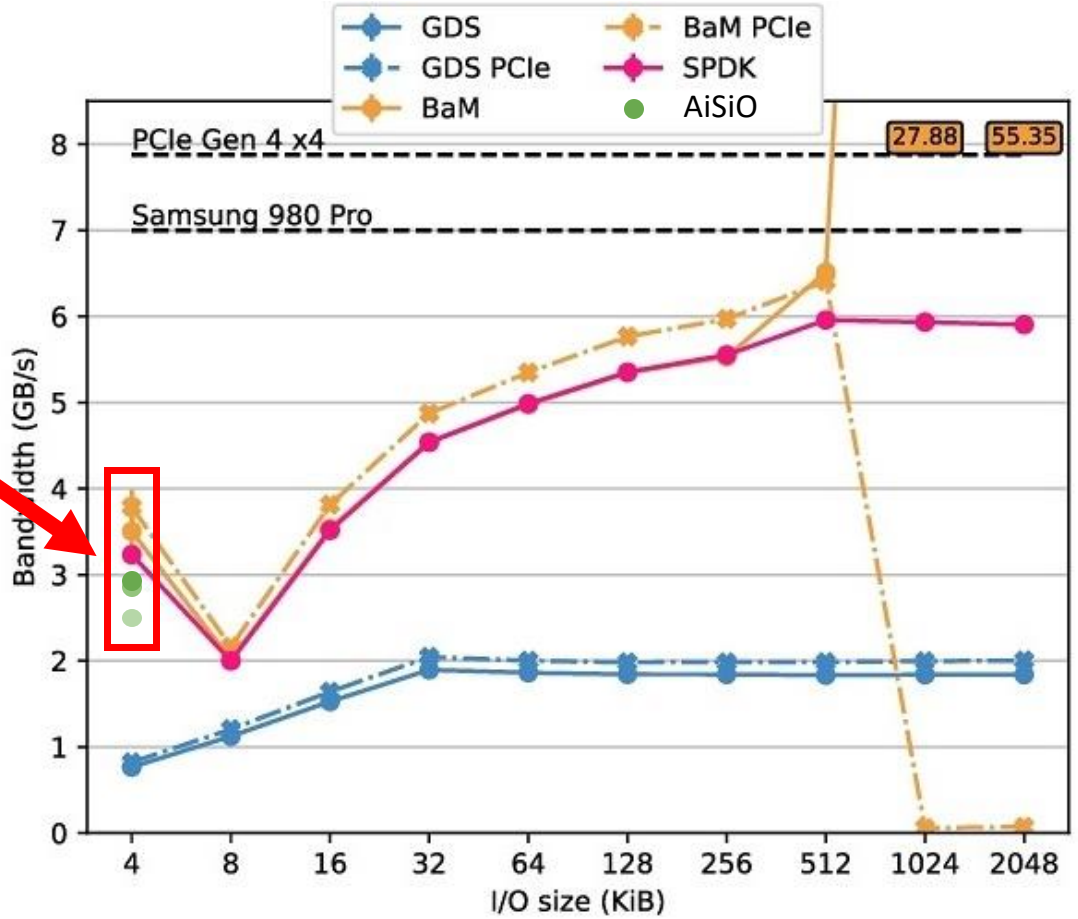
AiSiO: Accelerator Initiated Storage IO



AiSiO: Status

Benchmark tools:
gdsio, bdevperf, nvm-block-bench
xNVMe/SIL **file-loader** benchmark

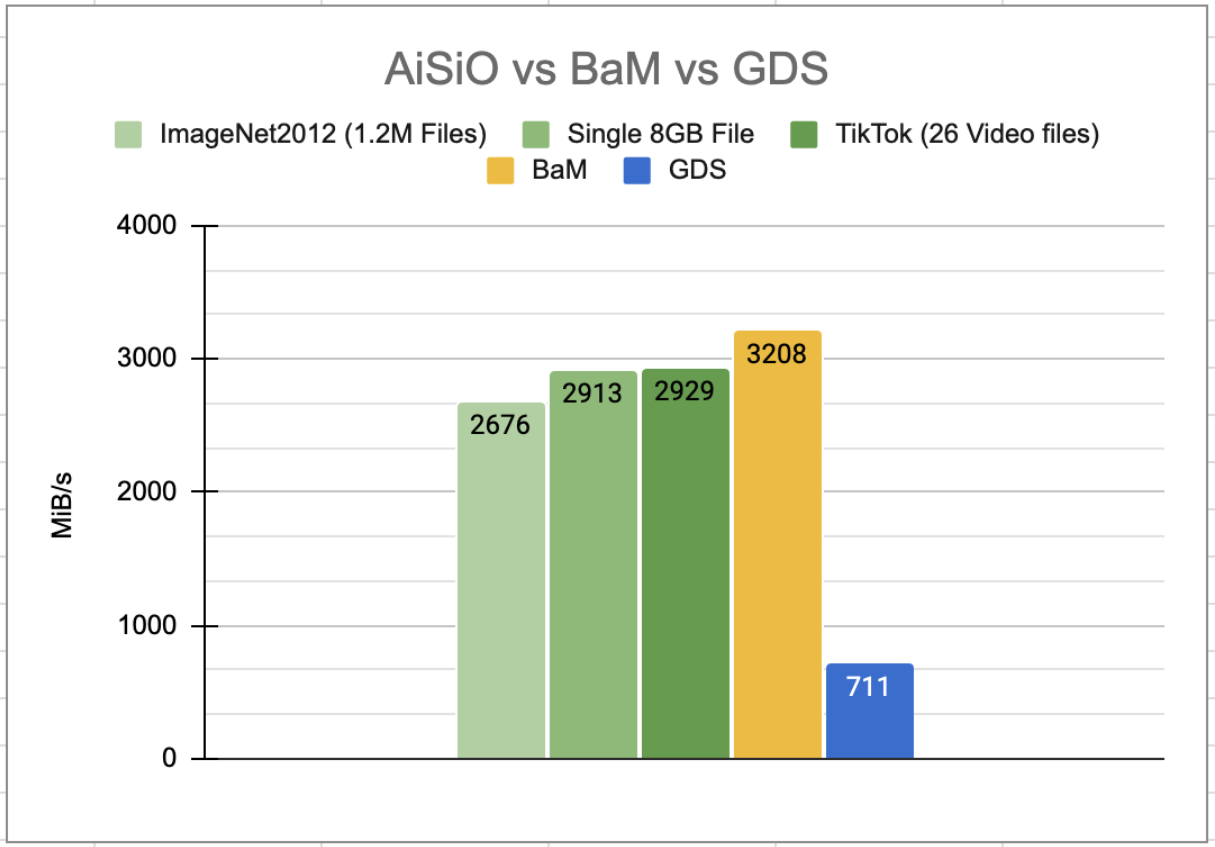
- **Gap** reduced
  - Proximity to performance of **BaM**
  - Achieve **22.1%** of **BaM** with **GDS**
  - Achieve **88.5%** of **BaM** with **AiSiO**
  - A **3.9x** improvement!
- Supports file operations
- Interoperable
- Open Source
- **Work in progress**
  - Aiming to **close** the gap!
  - Compute-kernel implementation.
  - Compute-kernel NVMe Driver



Path to GPU-Initiated I/O for Data-Intensive Systems
https://dl.acm.org/doi/10.1145/3736227.3736232
Paper at SIGMOD / DaMoN25

AiSiO: Status

- **Gap** reduced
  - Proximity to performance of **BaM**
  - Achieve **22.1%** of **BaM** with **GDS**
  - Achieve **88.5%** of **BaM** with **AiSiO**
  - A **3.9x** improvement!
- Supports file operations
- Interoperable
- Open Source
- **Work in progress**
  - Aiming to **close** the gap!
  - Compute-kernel implementation.
  - Compute-kernel NVMe Driver



Community and Industry Requests

- Linux interoperability for safety
  - File-locking and leases
- Linux DMA buffer
  - Common API for accelerator memory allocation for P2P and Zero-Copy
- Accelerator support

Feeding the Beast

Bridging **NVMe** Storage and **GPUs** while Preserving **File** Semantics

Simon A. F. Lund | Principal Engineer Samsung Semiconductor