



OPEN
Compute Project®

What is NVMe?

Why is Collaboration important?

July 2023 | Hanju Lee



Speaker



Hanju Lee

Principal Engineer

Product Planning Team, Memory Division

SAMSUNG

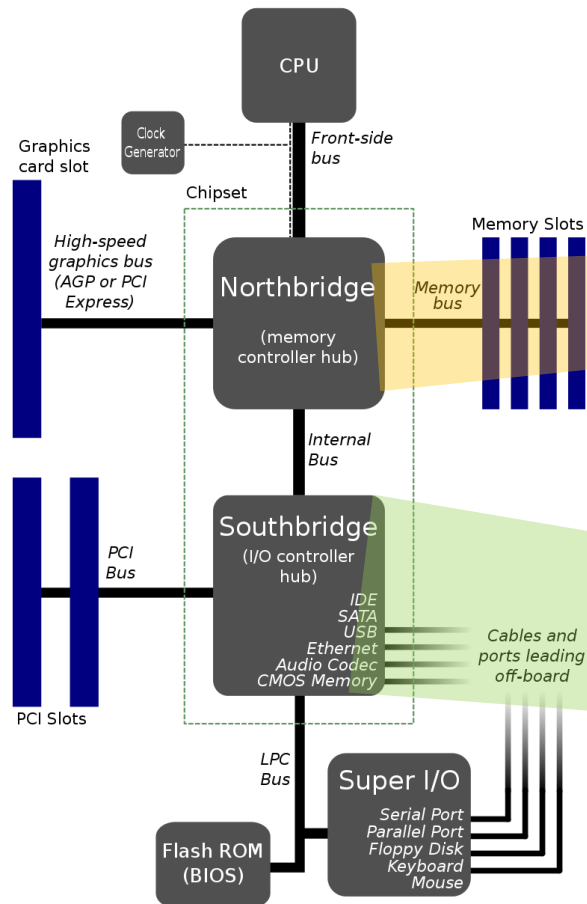
Agenda

1. Technical Trend

2. NVMe Specification

3. Future works

Computer System Diagram



[Typical north/southbridge layout]

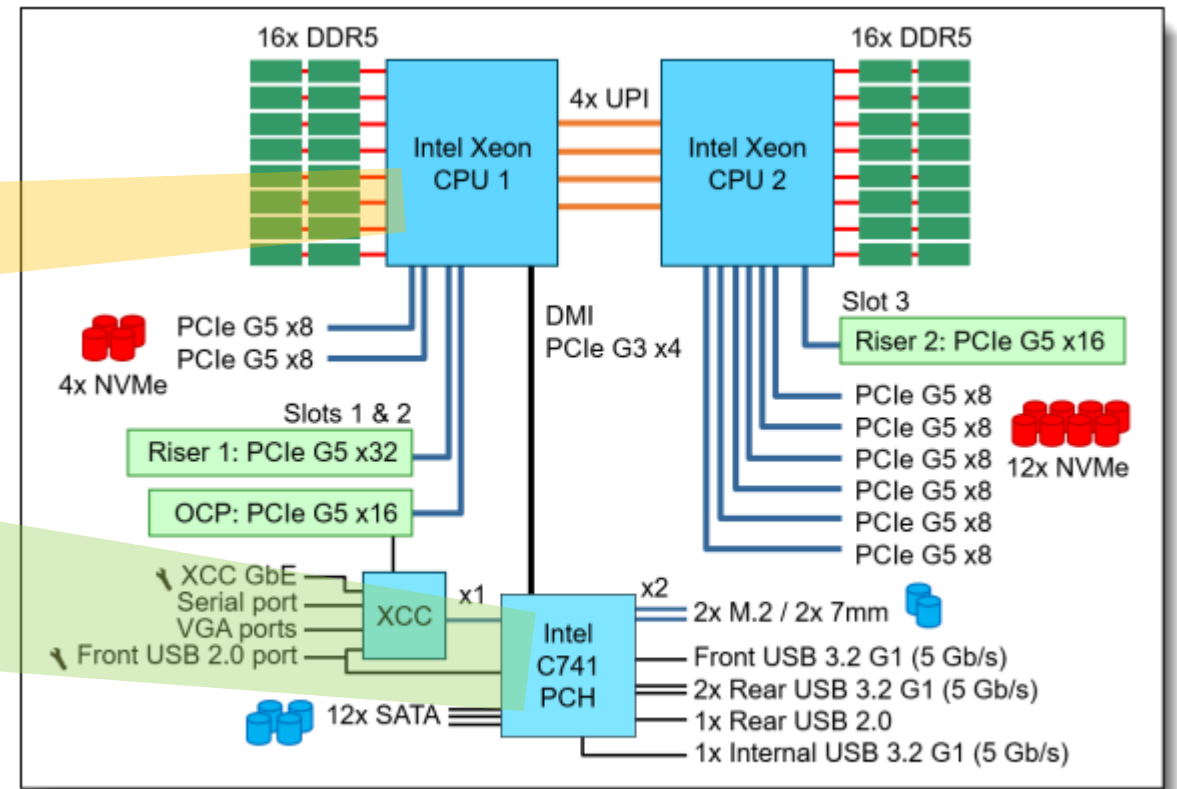
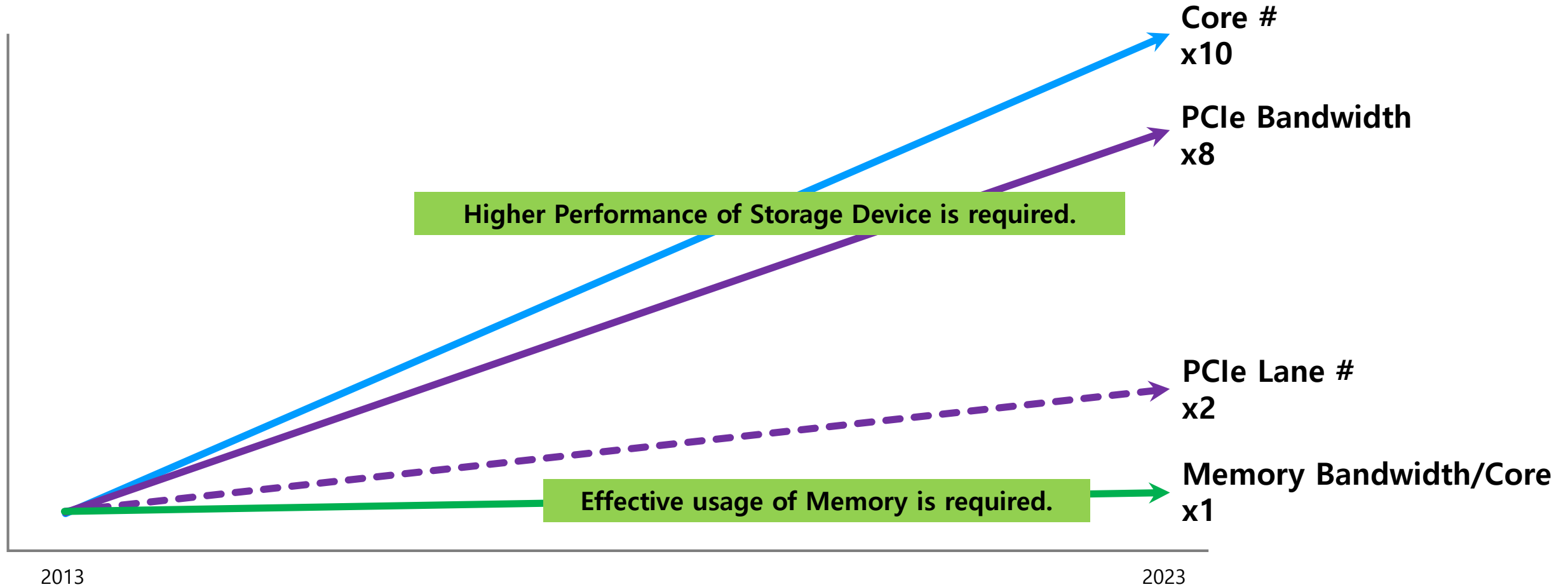


Figure 6. SR630 V3 system architectural block diagram

※ DMI: Direct Media Interface
 ※ PCH: Platform Controller Hub

System Trend

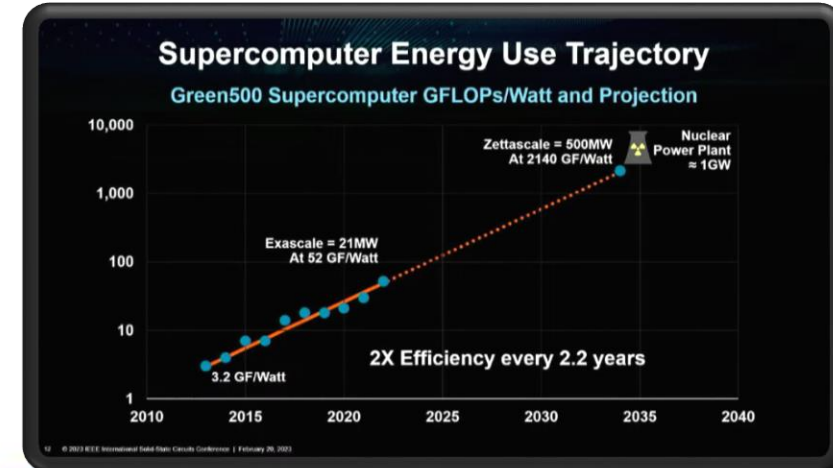


PCIe Trend

PCI-SIG® Roadmap

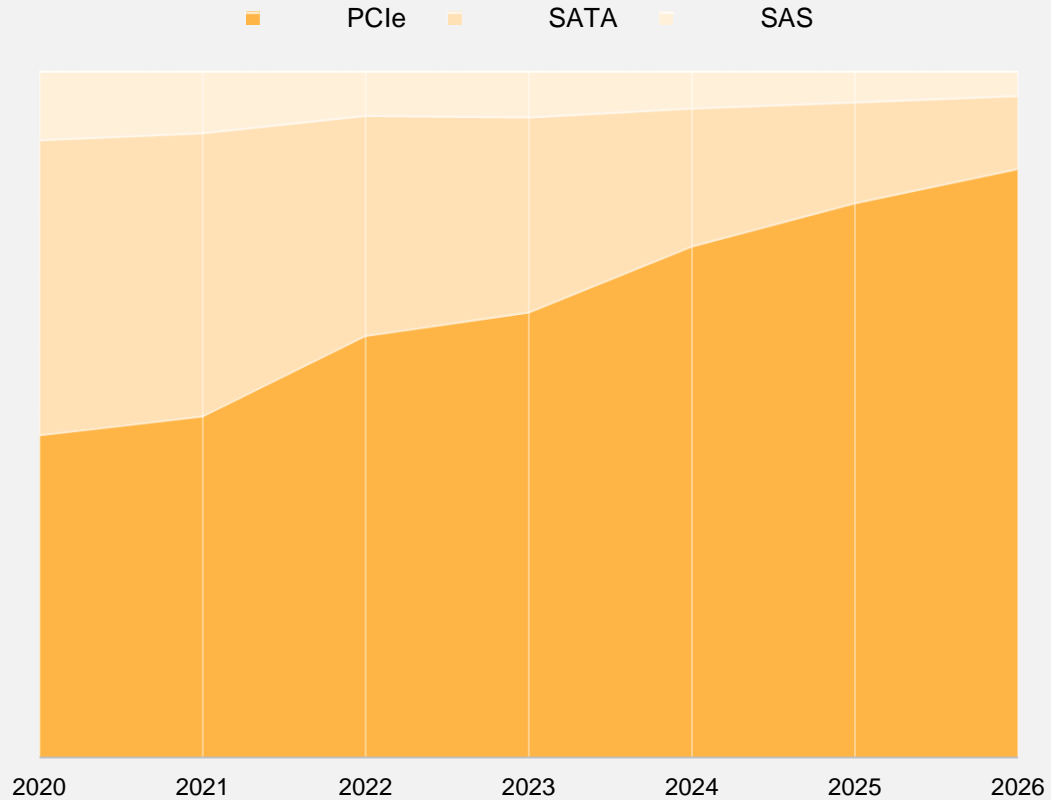
I/O BANDWIDTH DOUBLES EVERY 3 YEARS

Higher Bandwidth
Higher Power Consumption

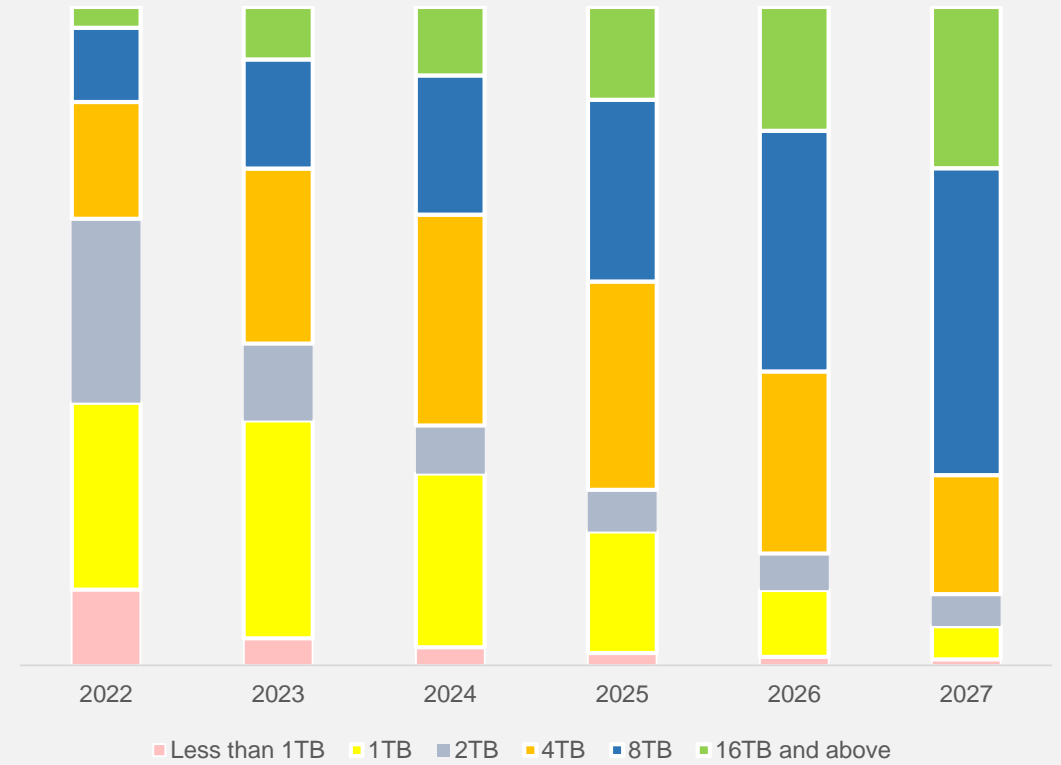


Enterprise SSD Trend

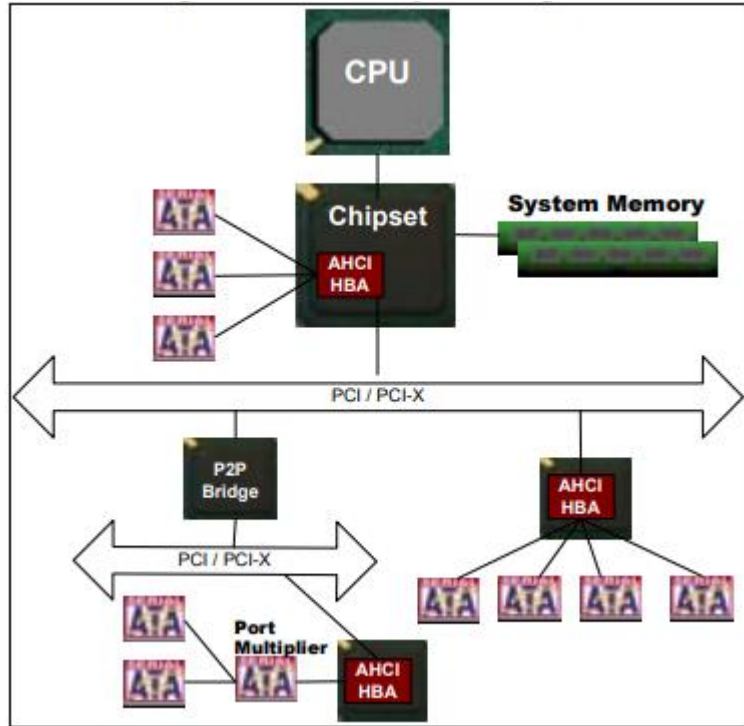
Enterprise SSD Interface



PCIe Server SSD Density



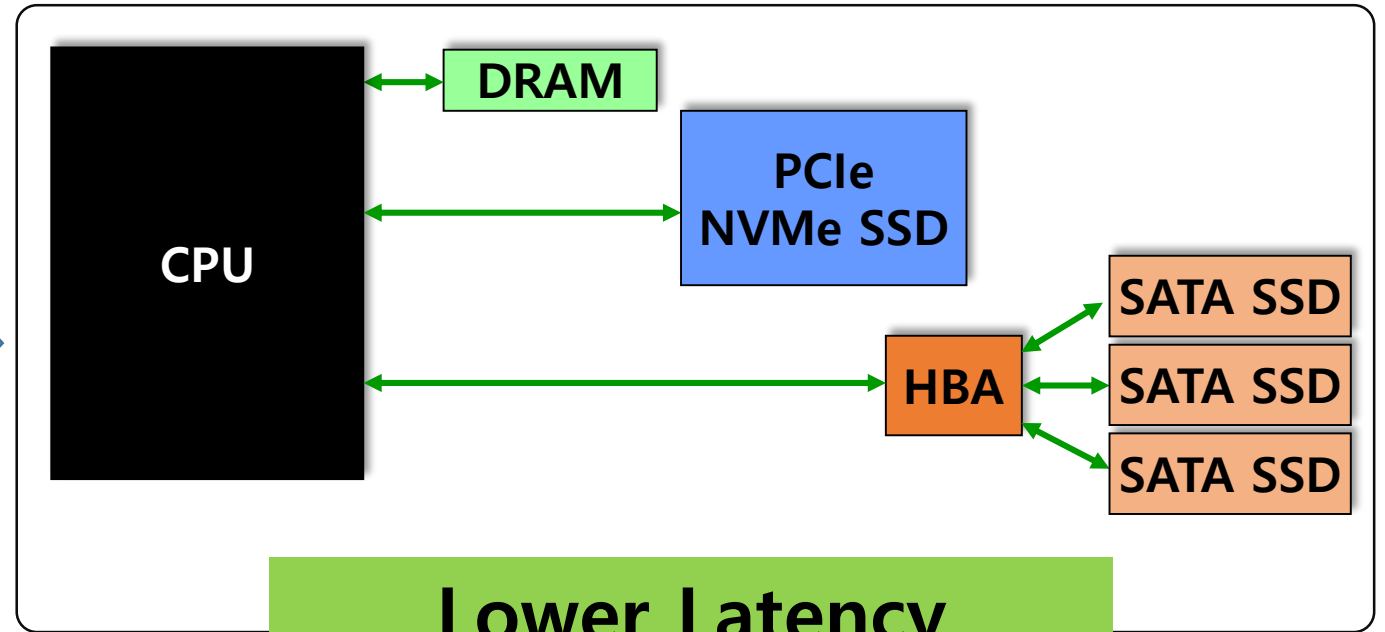
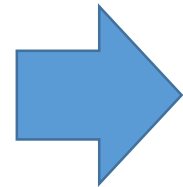
SATA to PCIe



Serial ATA ACHI Specification

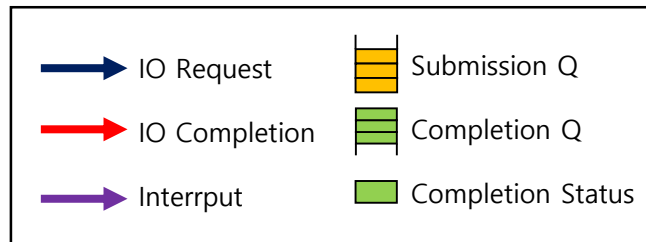
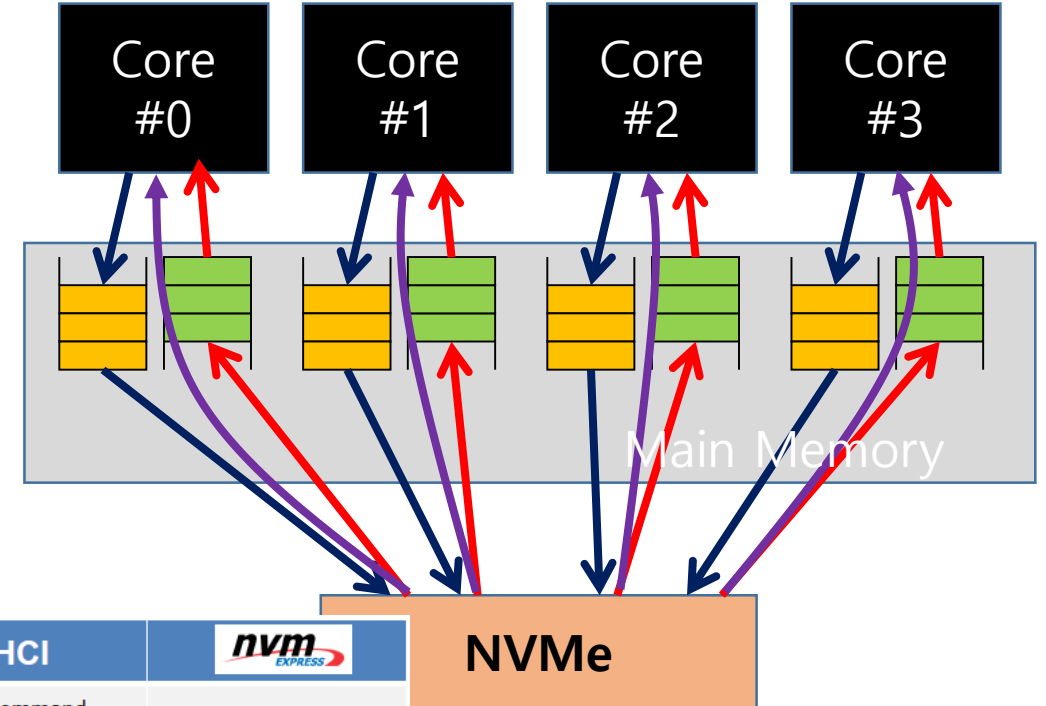
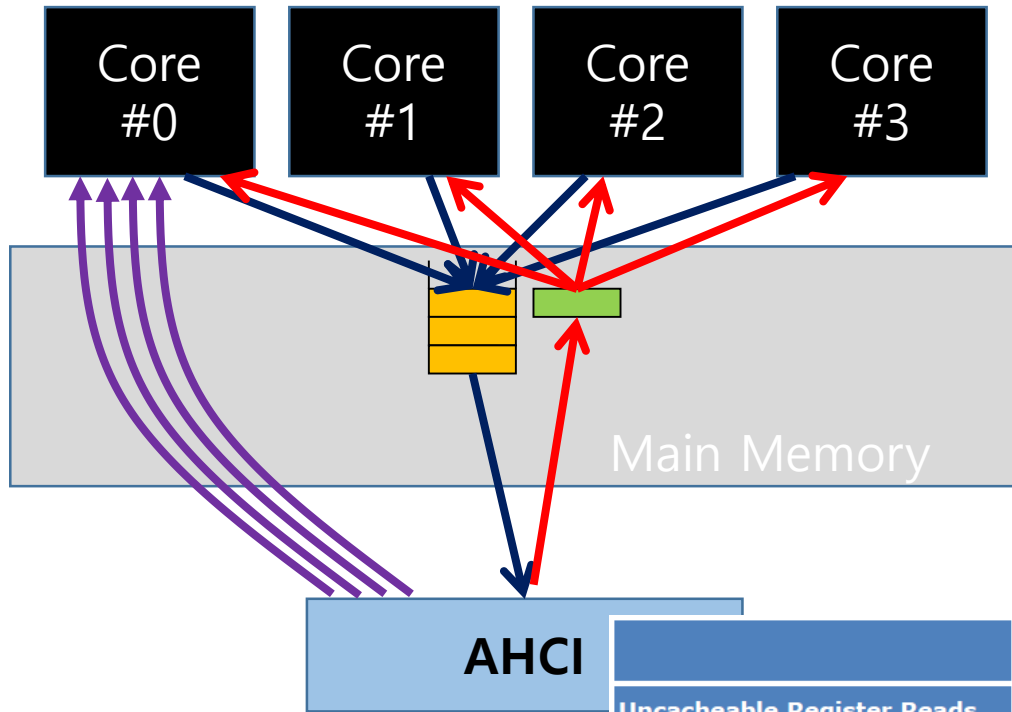
※ HBA: Host Bus Adapter


※ AHCI: Advanced Host Controller Interface



**Lower Latency
Higher Performance**

AHCI vs NVMe



	AHCI	 NVMe
Uncacheable Register Reads Each consumes 2000 CPU cycles	4 per command 8000 cycles, ~ 2.5 μ s	0 per command
MSI-X and Interrupt Steering Ensures one core not IOPs bottleneck	No	Yes
Parallelism & Multiple Threads Ensures one core not IOPs bottleneck	Requires synchronization lock to issue command	No locking, doorbell register per Queue
Maximum Queue Depth Ensures one core not IOPs bottleneck	1 Queue 32 Commands per Q	64K Queues 64K Commands per Q
Efficiency for 4KB Commands 4KB critical in Client and Enterprise	Command parameters require two serialized host DRAM fetches	Command parameters in one 64B fetch

Agenda

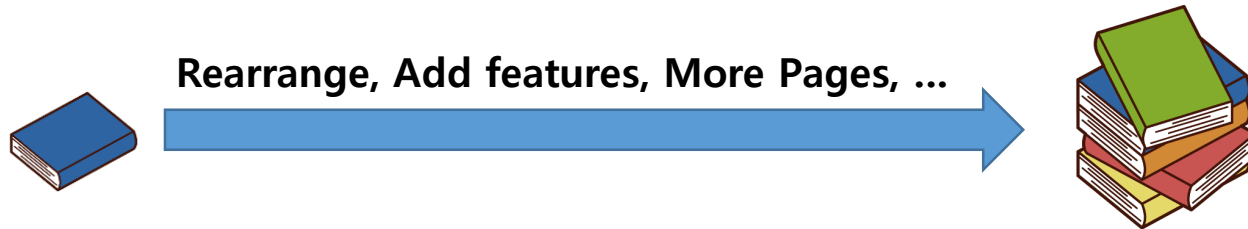
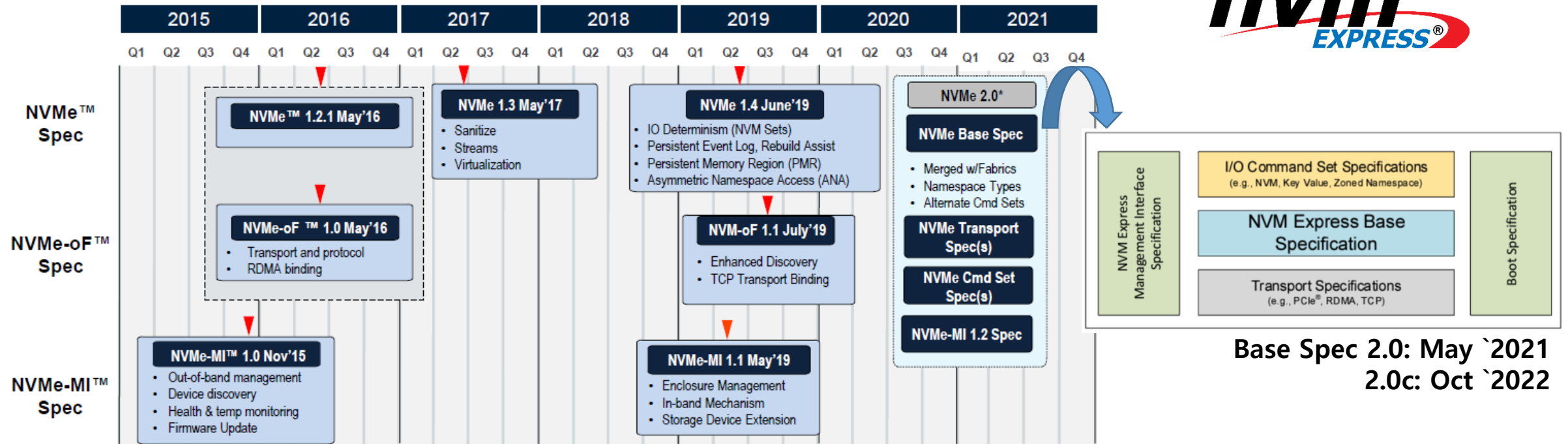
1. Technical Trend

2. NVMe Specification

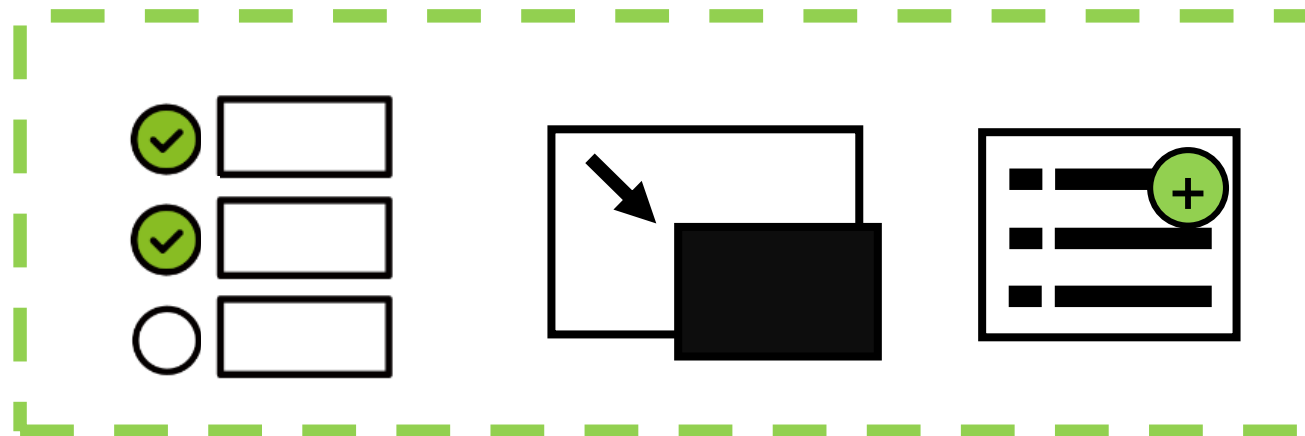
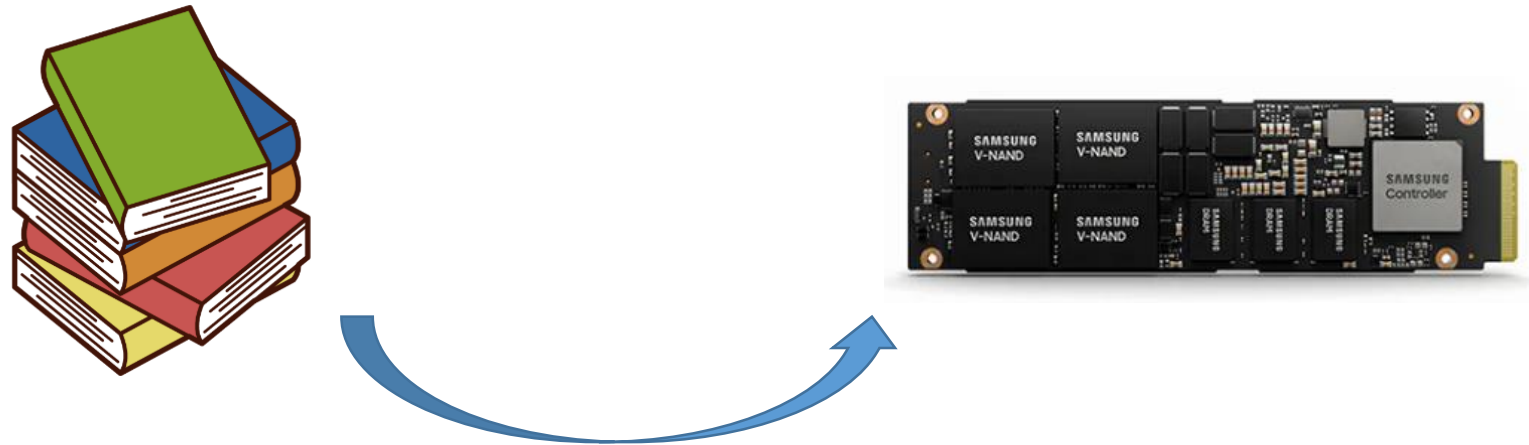
3. Future works

NVMe Specification

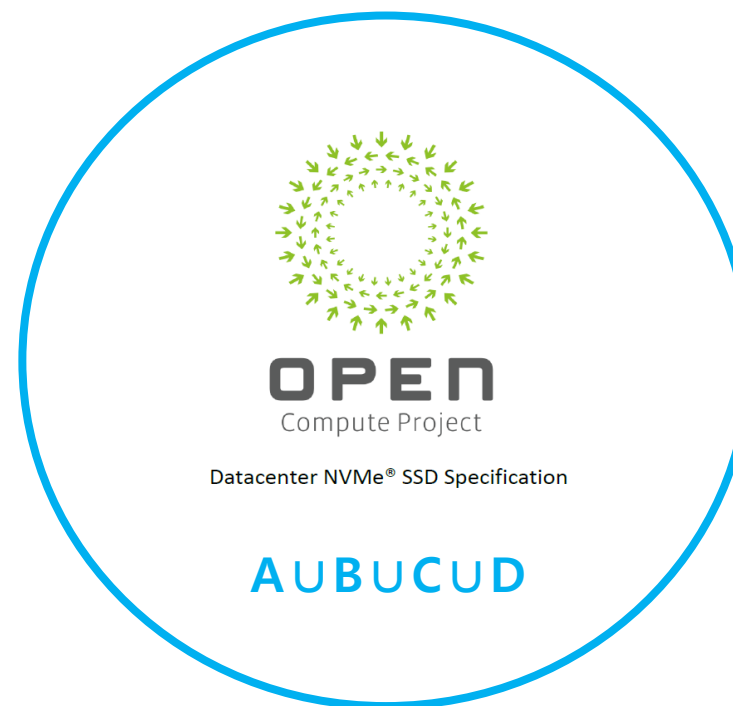
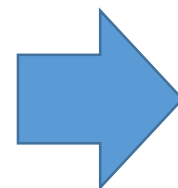
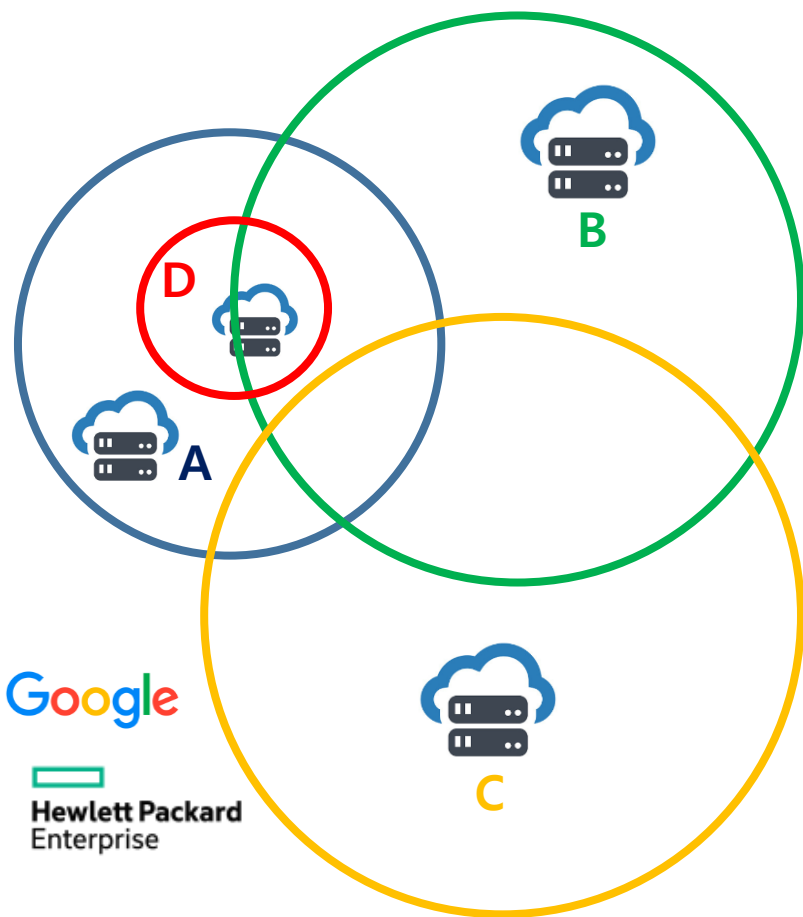
NVM Express Technology Specification Roadmap



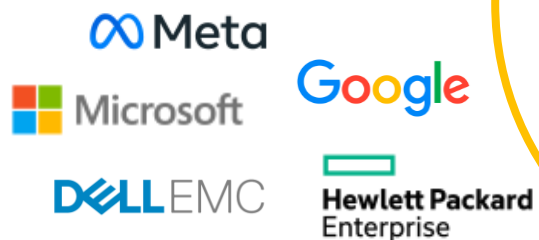
OCP Datacenter NVMe Specification



OCP Datacenter NVMe Specification

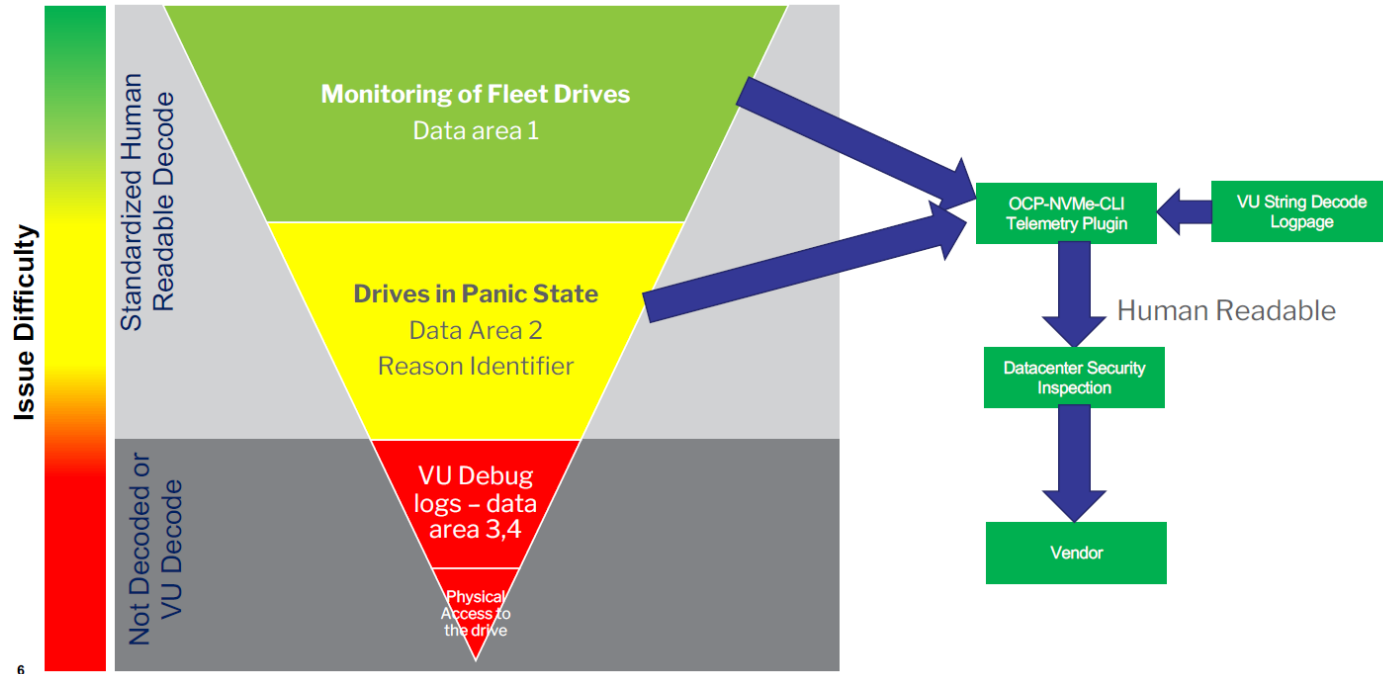


- Telemetry/Debug
- Latency Monitoring
- Media Friendly Solution(FDP/ZNS/...)
- Multi-Tenant
- ...

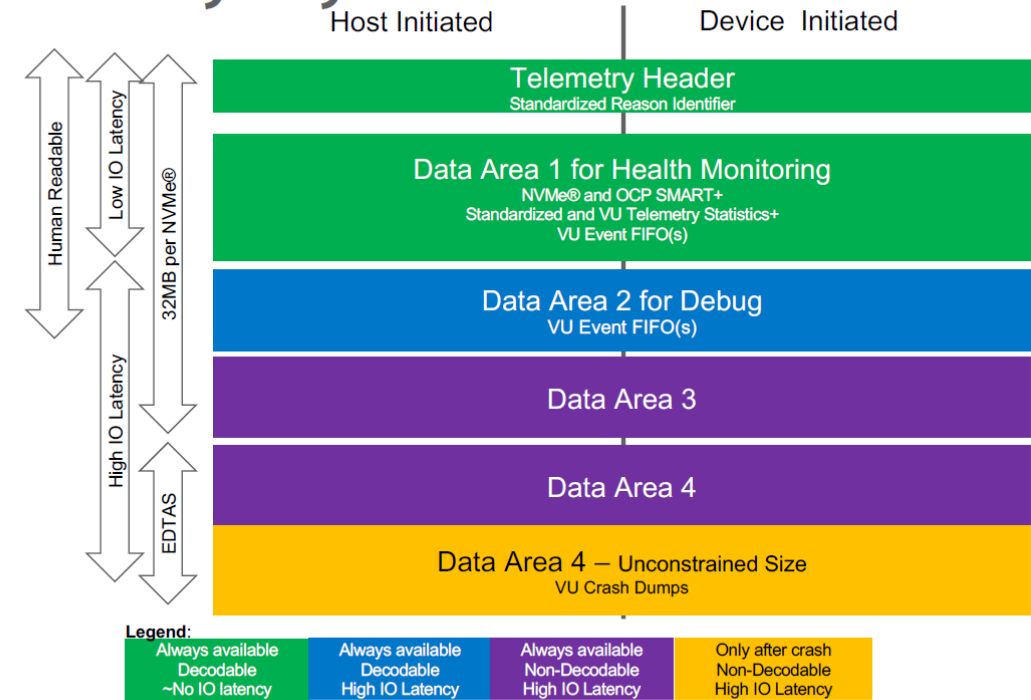


OCP Telemetry Debug

Standardized Telemetry Debug



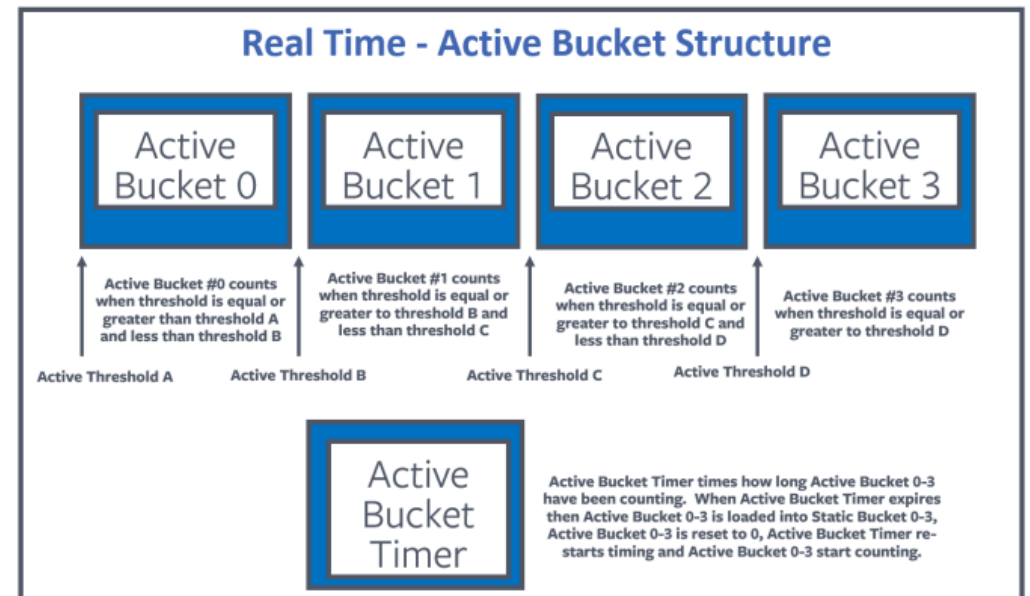
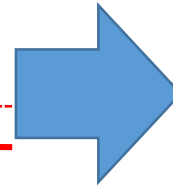
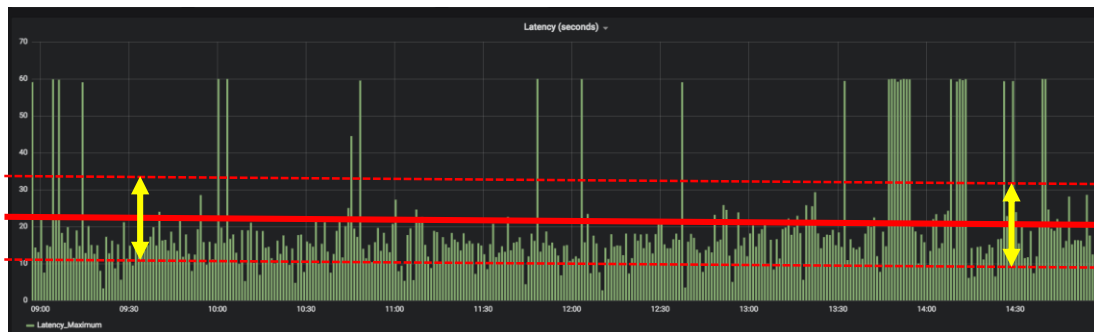
Telemetry Layout Details



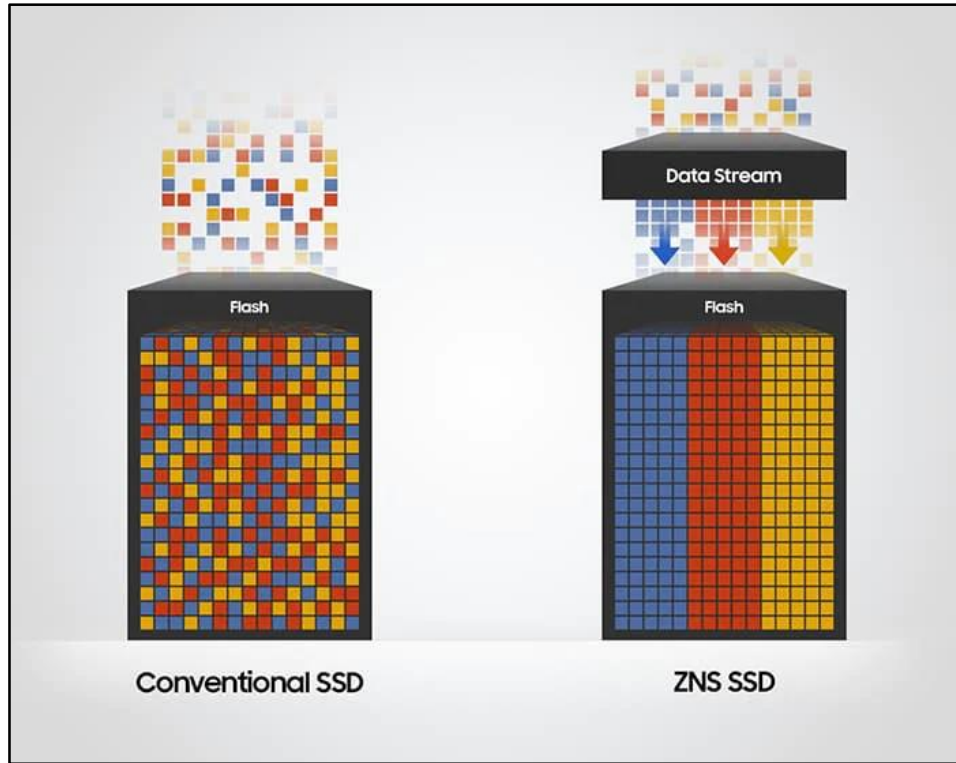
Community-driven hyperscale innovation for all.

https://drive.google.com/file/d/1jLVMQMvDIBbjctvz-QD9tEKix4G0cyn7/view?usp=share_link
 Future Telemetry Standardization in OCP Datacenter NVMe Specification, Steven Wells

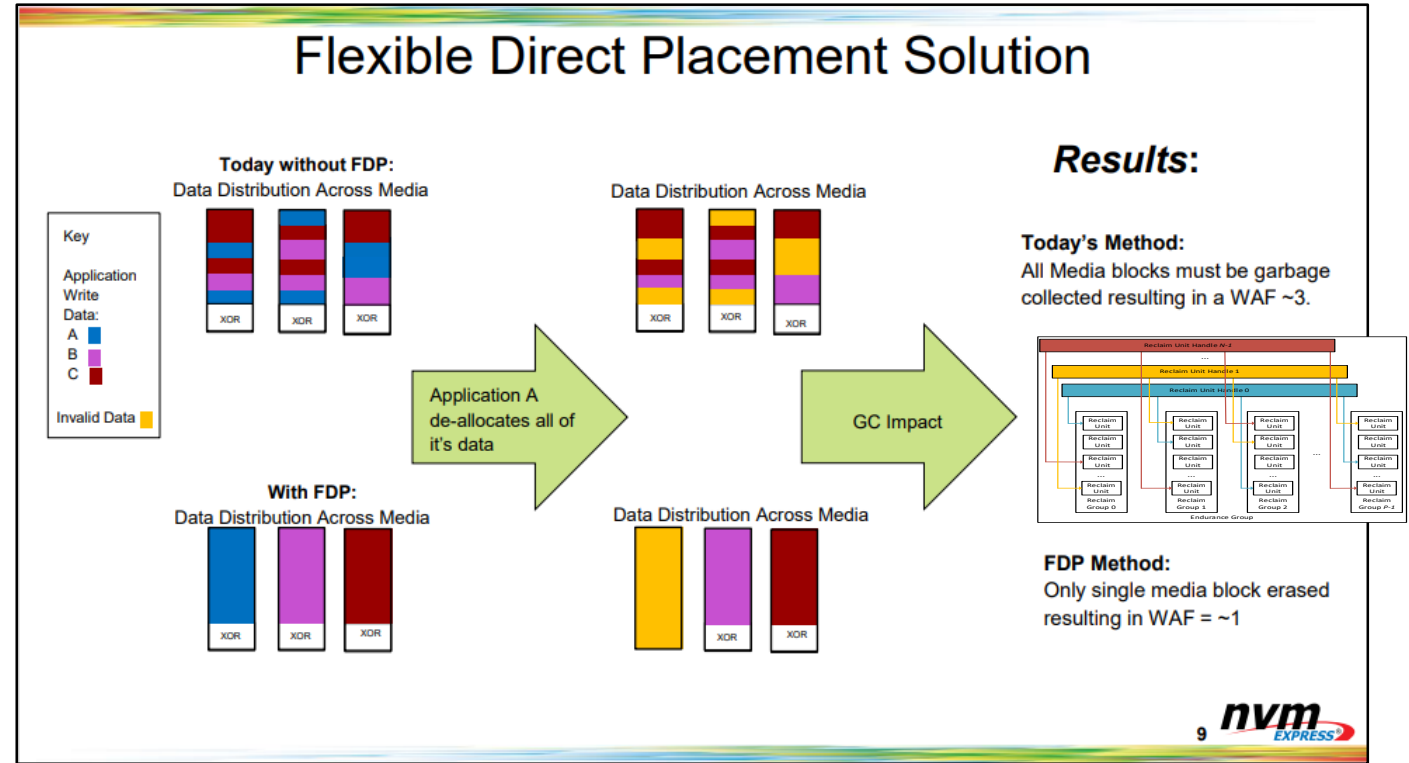
Latency Monitoring



Media Friendly Solution



<https://semiconductor.samsung.com/>

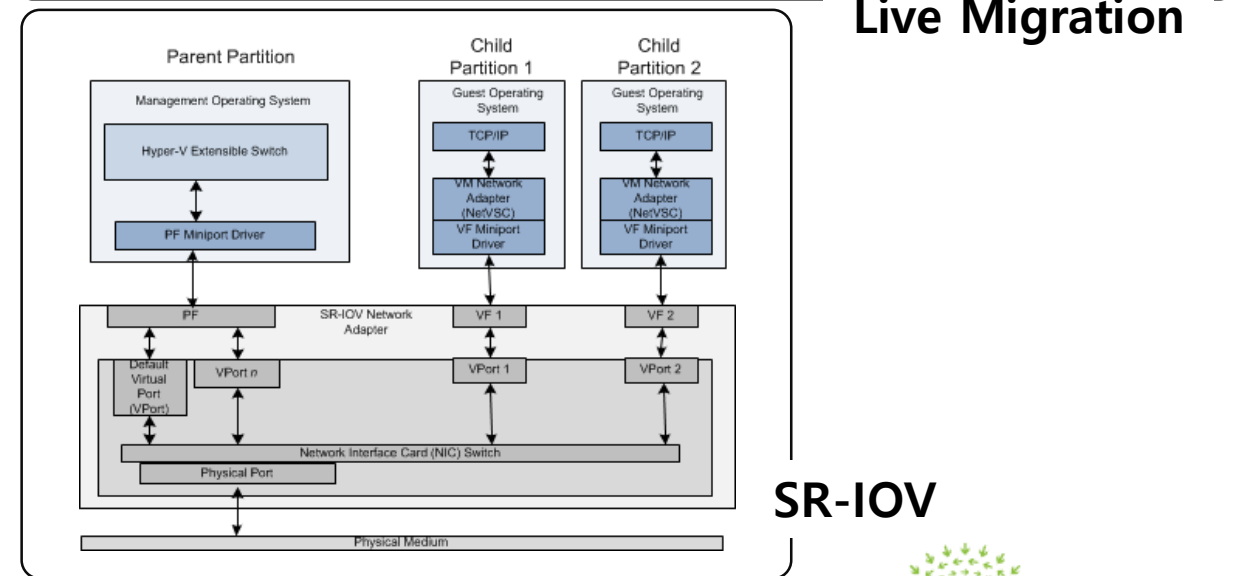
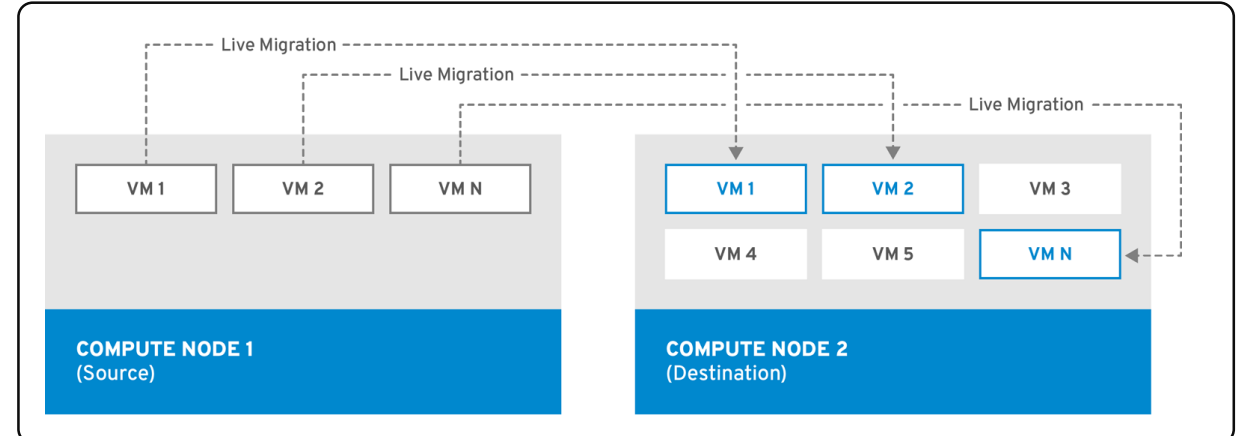
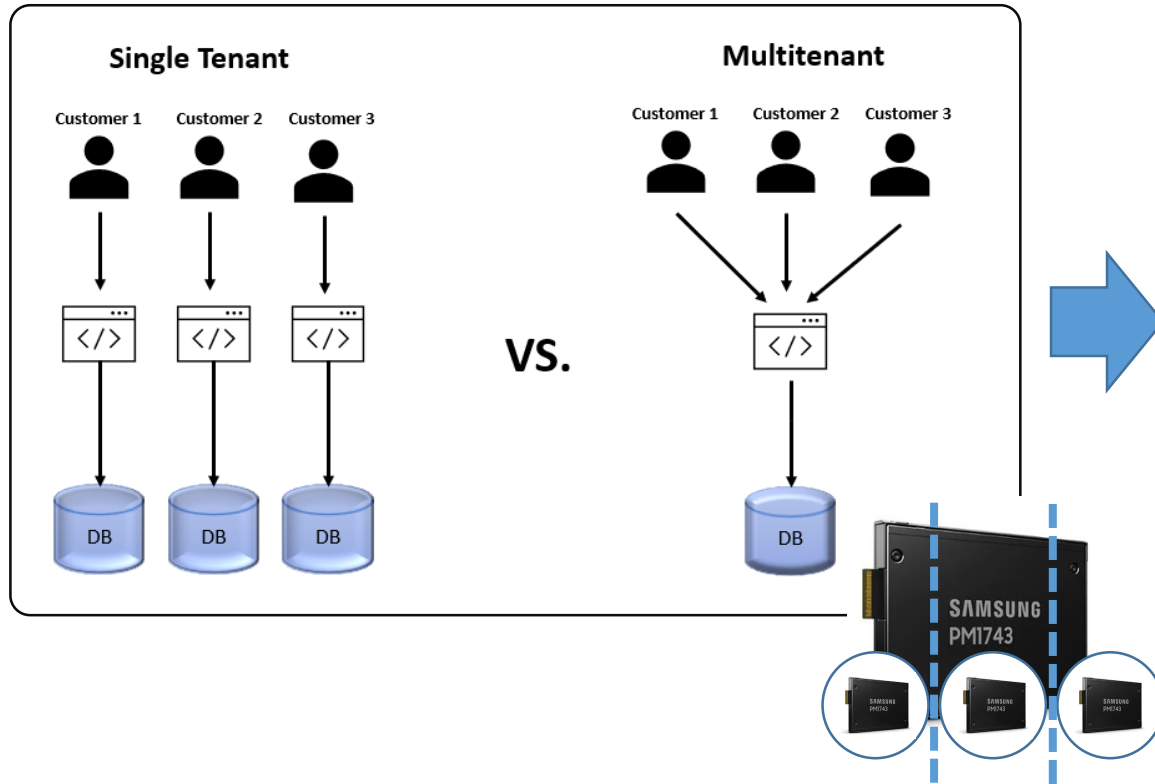


<https://nvmexpress.org/wp-content/uploads/Hyperscale-Innovation-Flexible-Data-Placement-Mode-FDP.pdf>

Ross Stenfort, Meta

WAF (Write Amplification Factor) = ~1

Multi-Tenant Solution



Agenda

1. Technical Trend

2. NVMe Specification

3. Future works

When new feature is added...

Full Stack Optimization



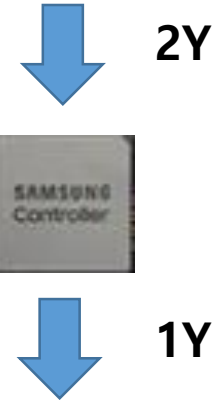
Efficiency ↑
TCO ↓

System
Architect



Storage
Provider

Request



Collaboration is important!



OCP Datacenter NVMe SSD Spec

- NVM Express Specification Features
- Vendor unique log pages
- PCI Express® Specification Features
- SMART Log Requirements
- Reliability, Endurance, Thermal, Power Requirements
- Management
- Security Requirements
- Labeling, Compliance
- Form Factor Requirements
- Open source tool access requirements

14:50~ [SSD virtualization technology trend at multi-tenancy application](#)

15:30~ [FDP and Eco System](#)

16:10~ [Introduce Telemetry of SSD](#)

A journey shared takes us beyond