NVMe® Computational Storage: From Addressing Ransomware to Improving Bandwidth

April 11, 2024

Jason Molgaard (Solidigm), Scott Shadley (Solidigm), Tim Fisher (IBM)

Meet the Presenters

Jason Molgaard
Principal Storage Solutions Architect

Scott Shadley
Director of Strategic Planning

Tim Fisher
STSM, FlashCore Module Chief Architect

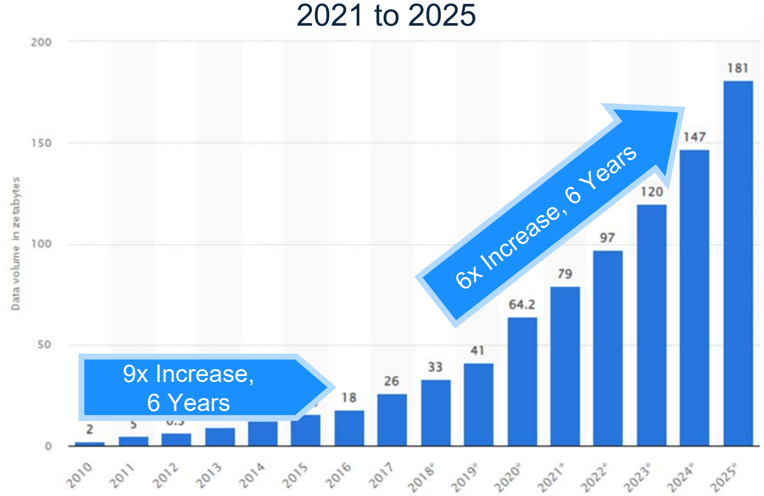
Agenda

- Computational Storage Market
- Protecting Data at Rest with NVMe® Computational Storage
- Enterprise Data Resilience with FlashCore Module and NVMe® Computational Storage
- Q&A

Computational Storage Market

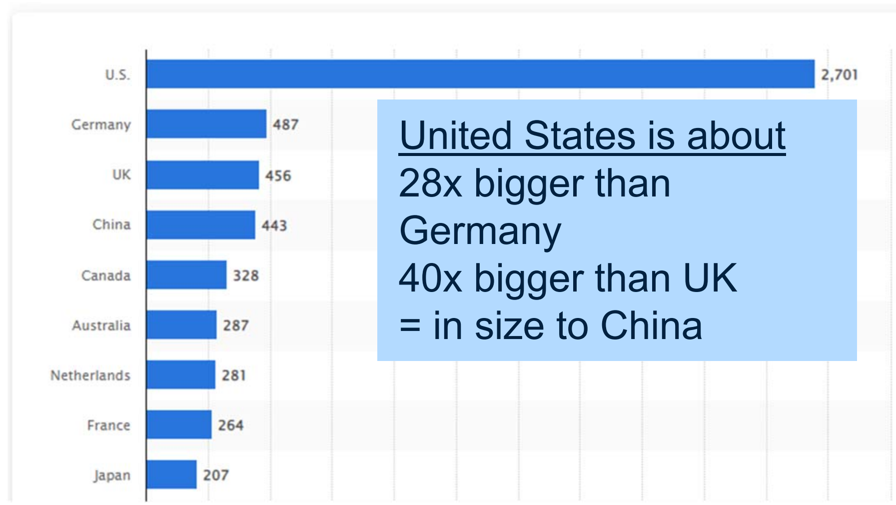
The Forefront of Change

Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025



Source: Total data volume worldwide 2010-2025

Number of data centers worldwide in 2022, by country



Source: Data centers worldwide by country 2022

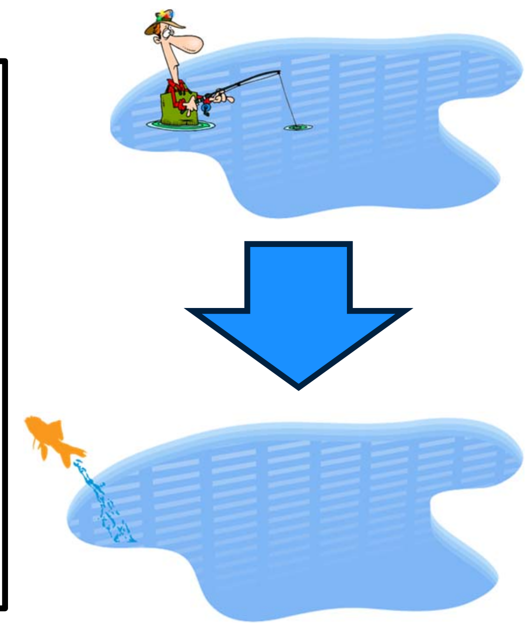
What has Changed?

Why Now?

- Storage is no longer ‘SLOW’
- Memory is no longer ‘Gated’
- Data Gravity, Data Size, Data Locality
- Edge Data Explosion, Transport issue!
- SNIA, NVM Express, CXL, OCP, Others are providing new guidance in new areas of implementation

Key Benefits?

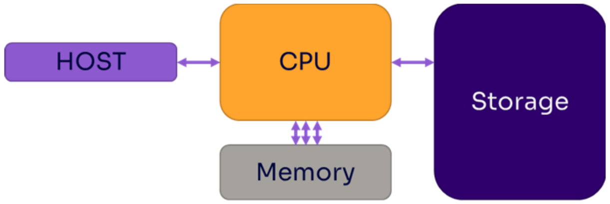
- Faster, Fewer, Easier I/O transfers
- Reducing DRAM/Network tax with new transports, solutions, products
- Redeploy Primary CPU to High Value Work, offer up new services to help
- Improved performance due to parallelism for certain workloads
- Better scheduling of data management and device functionality



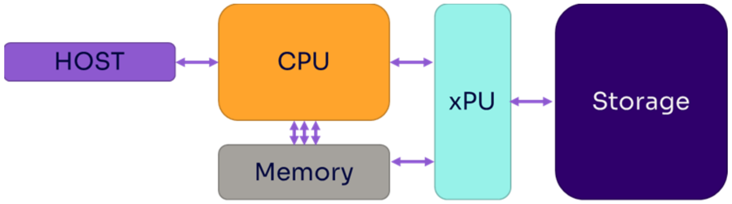
New NVMe® Command Sets, Paired with SNIA Architecture

John von Neumann – The Princeton Architecture

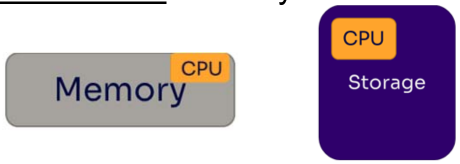
The Ecosystem today, we have our friend J. v N. – CPU/Memory/Storage



The world is evolving and there is a need for compute in more available locations, enter the world of "Accelerators" – SmartNIC, xPU, DPU, GPU, IPU



These are great, but there is room for more! History tells us this much…

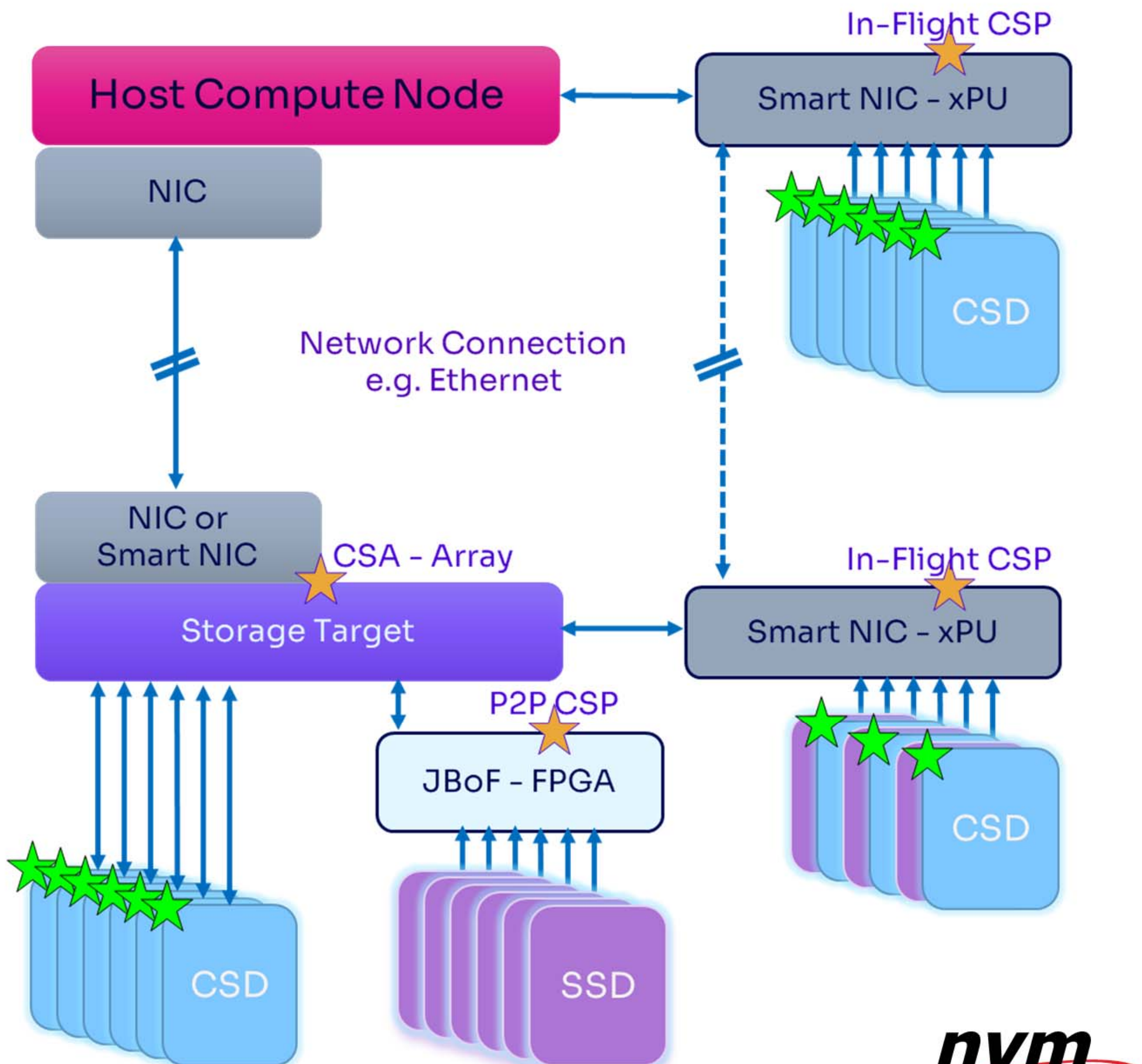


Amdahl's Law

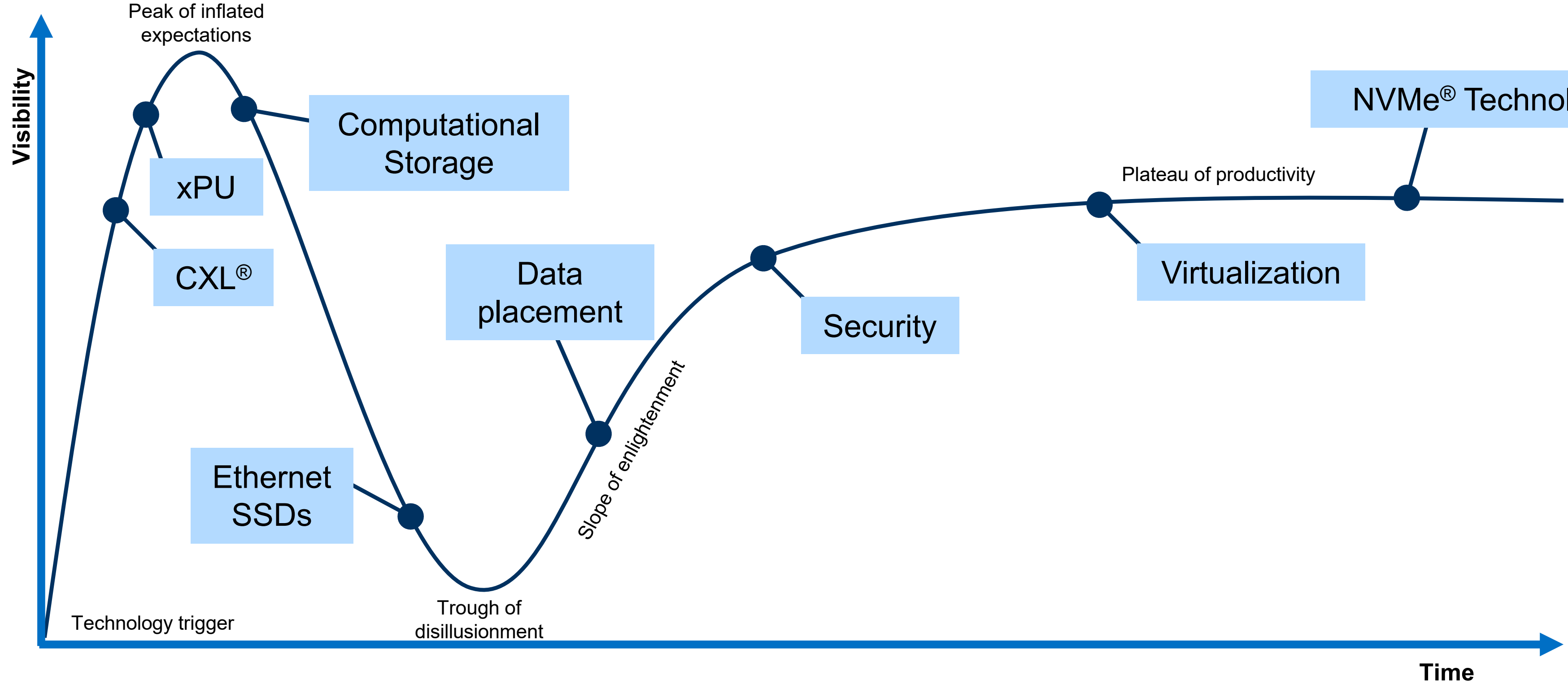
Amdahl's Law
Measures the theoretical speedup of a program's execution latency as a function of the number of processors executing it…

The limiting factor is 'Serial' so move it!!

- In Memory? Sure…
  - New areas of "Computational Memory"
- In Storage? Why Not…
  - Computational Storage Drives, Arrays
- In the Middle? Of Course…
  - Computational Storage Processors
  - SmartNIC, xPU…



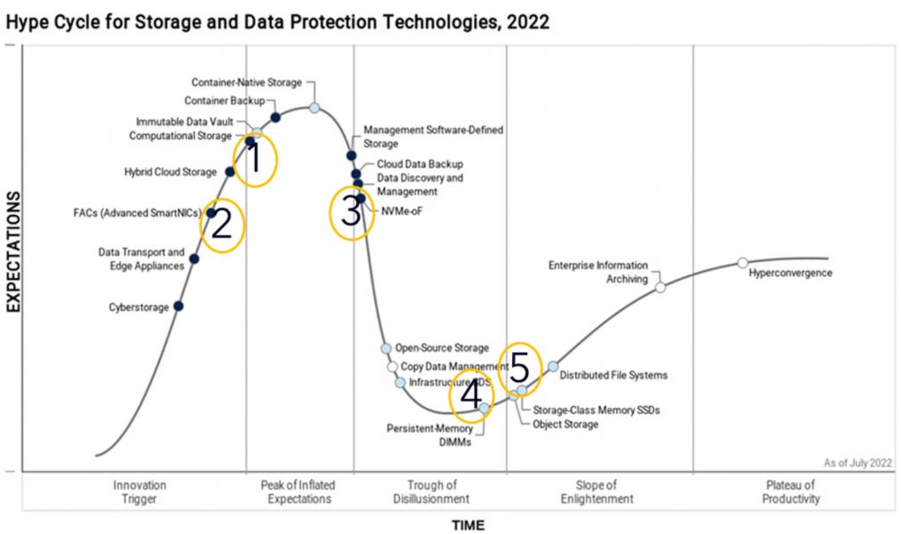
A ‘Spin’ on the Hype Cycle – Not ‘official’

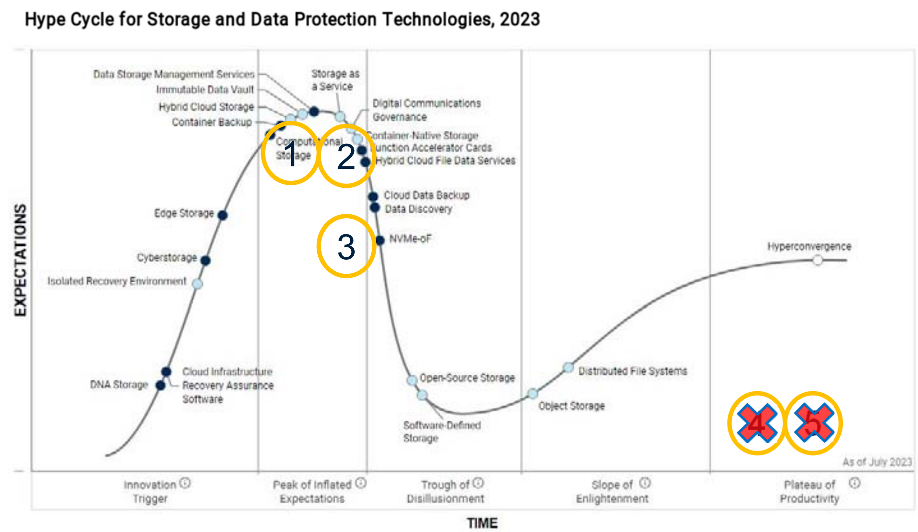


* *Technology hype cycle is a product of the Gartner Research* *https://www.gartner.com/en/research/methodologies/gartner-hype-cycle*

The "Official" Storage Hype Cycle – One Year later

Plateau will be reached: ○ <2 yrs. ○ 2–5 yrs. ● 5–10 yrs. ▲ >10 yrs. ⊗ Obsolete before plateau





1. Computational Storage
2. FACs
3. NVMe-oF™ Technology
4. PM DIMMs
5. Storage-Class Memory SSDs

How to Drive Technology Forward

Pick an Organization, Join the Work, Drive the Market

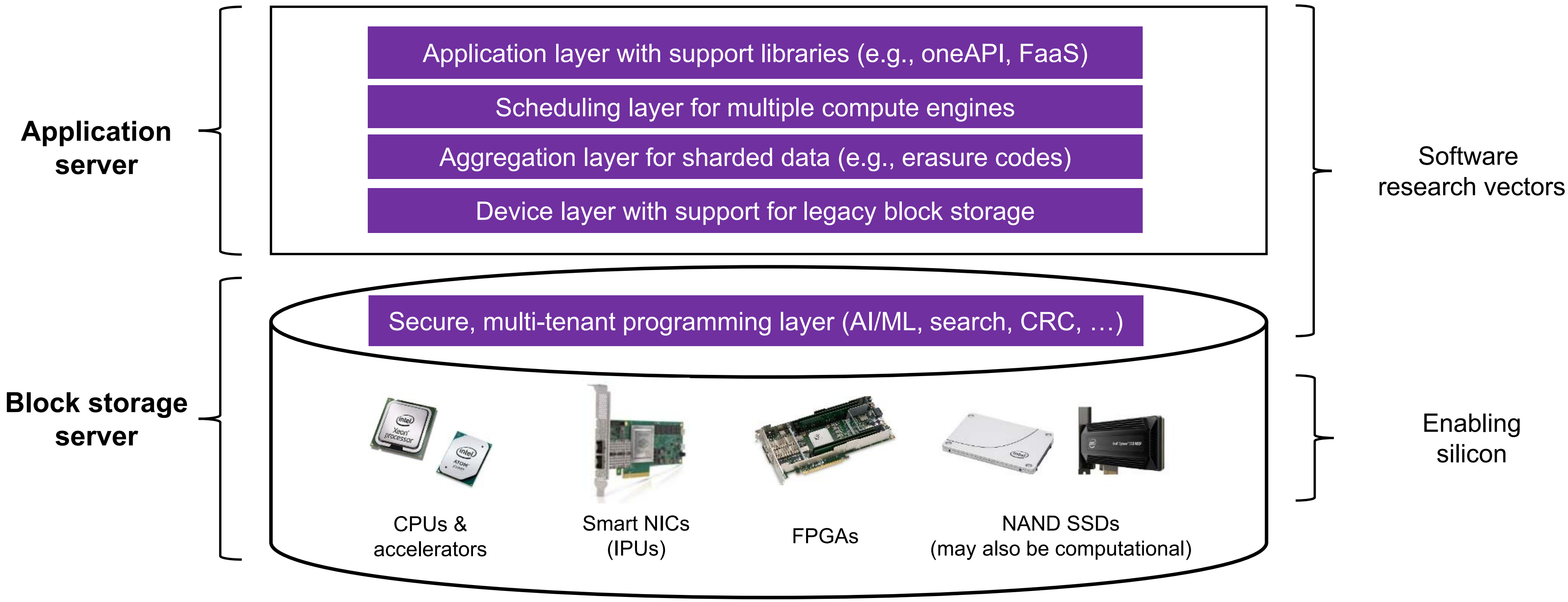
Areas of Work that drive change in Compute Environment:

- SNIA – Computational Storage, SDXI, Security, Zone Storage, Object Storage, DNA Data Storage, DPU, NVMPM…
- Transports – NVM Express, STA, PCI-SIG, CXL®, UCIe…
- Global Work Groups – OCP, SODA…

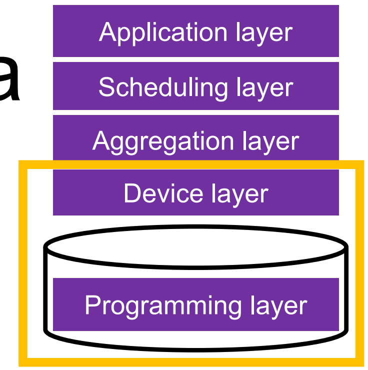
I think you get the picture…

Protecting Data at Rest with NVMe® Computational Storage

NVMe® Computational Storage SW Layers

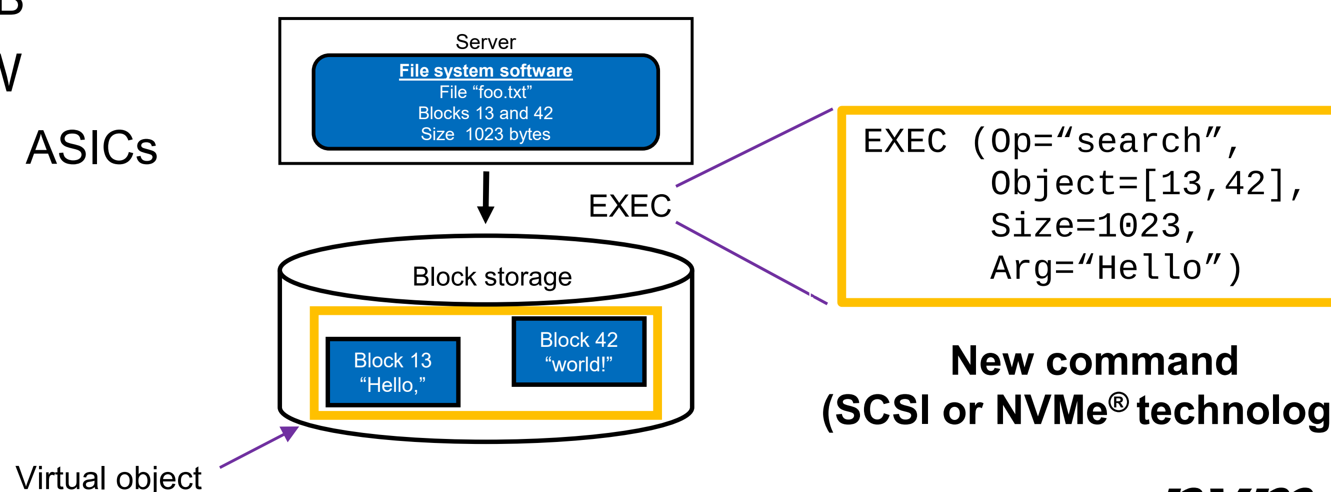


Device Layer – Teaching Storage About Data



1. Teach block storage about data *objects*
   - Files, directories, tables, records, …
2. Specify operations on objects
   - Search in text object A
   - Classify image object B
3. Execute on diverse HW
   - CPUs, GPUs, FPGAs, ASICs

Our research and POCs use virtual objects
USENIX Hot Storage '19 (Adams, Keys, Mesnier)



New command
(SCSI or NVMe® technology)

Use Cases We're Exploring

AI/Deep learning/Machine Learning/Speech/Language

- Video transcoding, image cropping, …
- Training & inference, …
- Large scale data processing

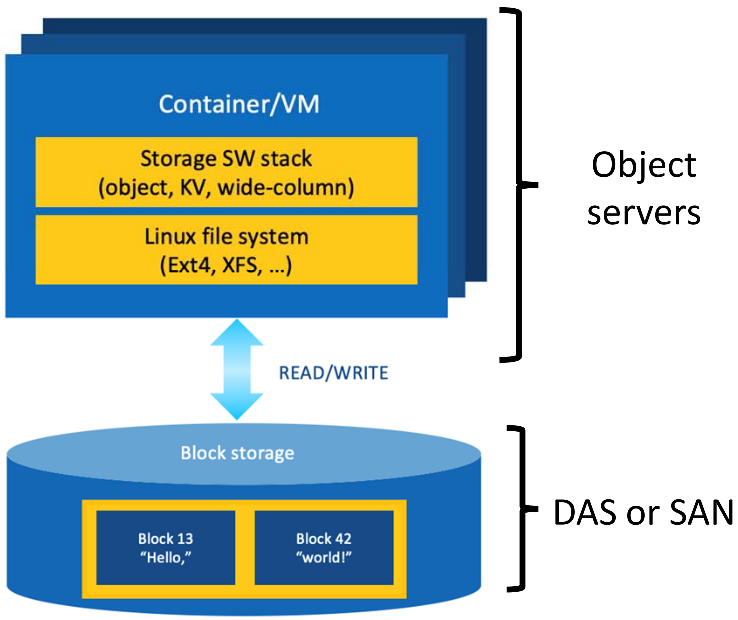
Sorting, searching, filtering

Low-level storage services

- Compression, encryption
- Deduplication (e.g., MD5)
- Data scrubbing (e.g., CRC-32C)

Data Integrity Check as Practical Use Case

Using computational storage for expensive data integrity checks



MINIO | ceph | openstack | cassandra | …

Storage servers regularly scrub all data –

1. Read all data from local FS
2. Hash data (CRC-32C, MD5, Highway Hash)
3. Compare with previously stored hashes

Significant READ traffic generated

Intel® Data Streaming Accelerator (DSA)
can be used for computational storage

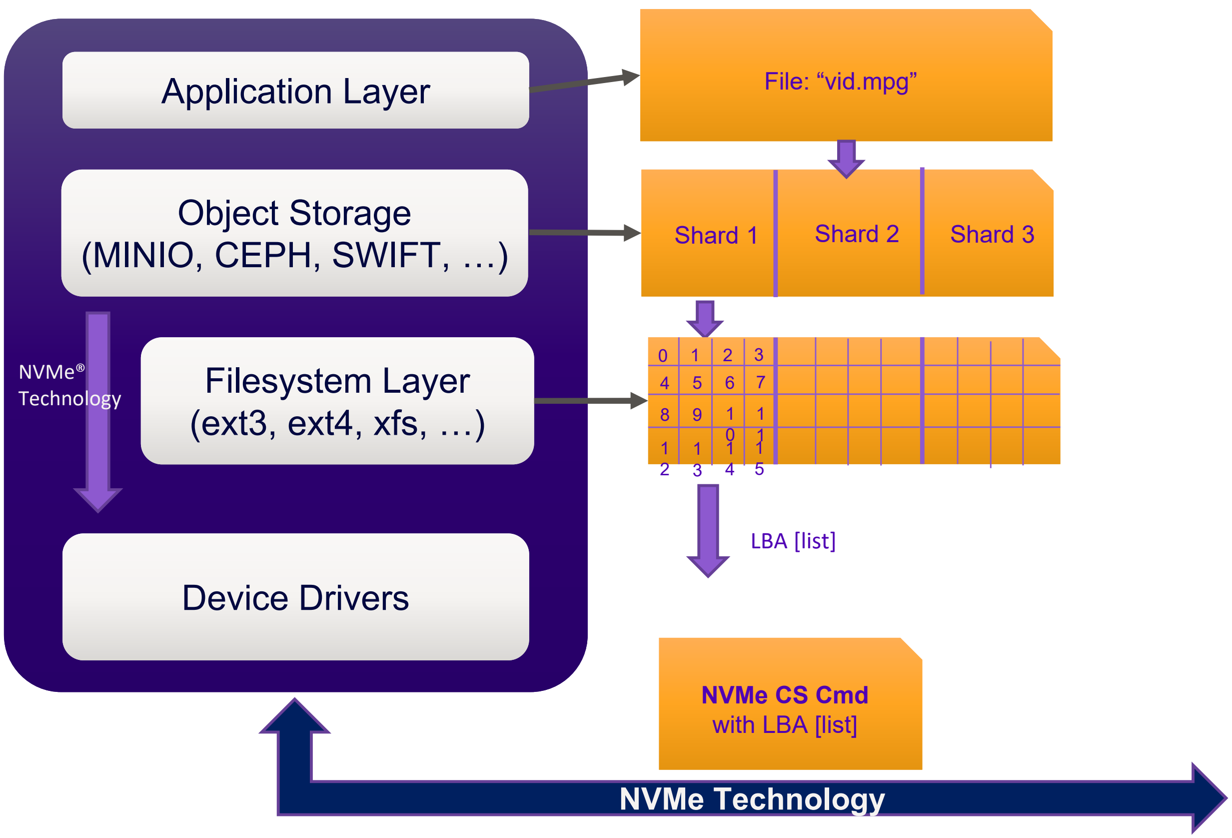
Data integrity check is:

- X Compute intensive
- X Read intensive (SSD & I/O)
- X Memory intensive (host)
- X Not scalable

Why Use a CSD For The Data Integrity Use Case?

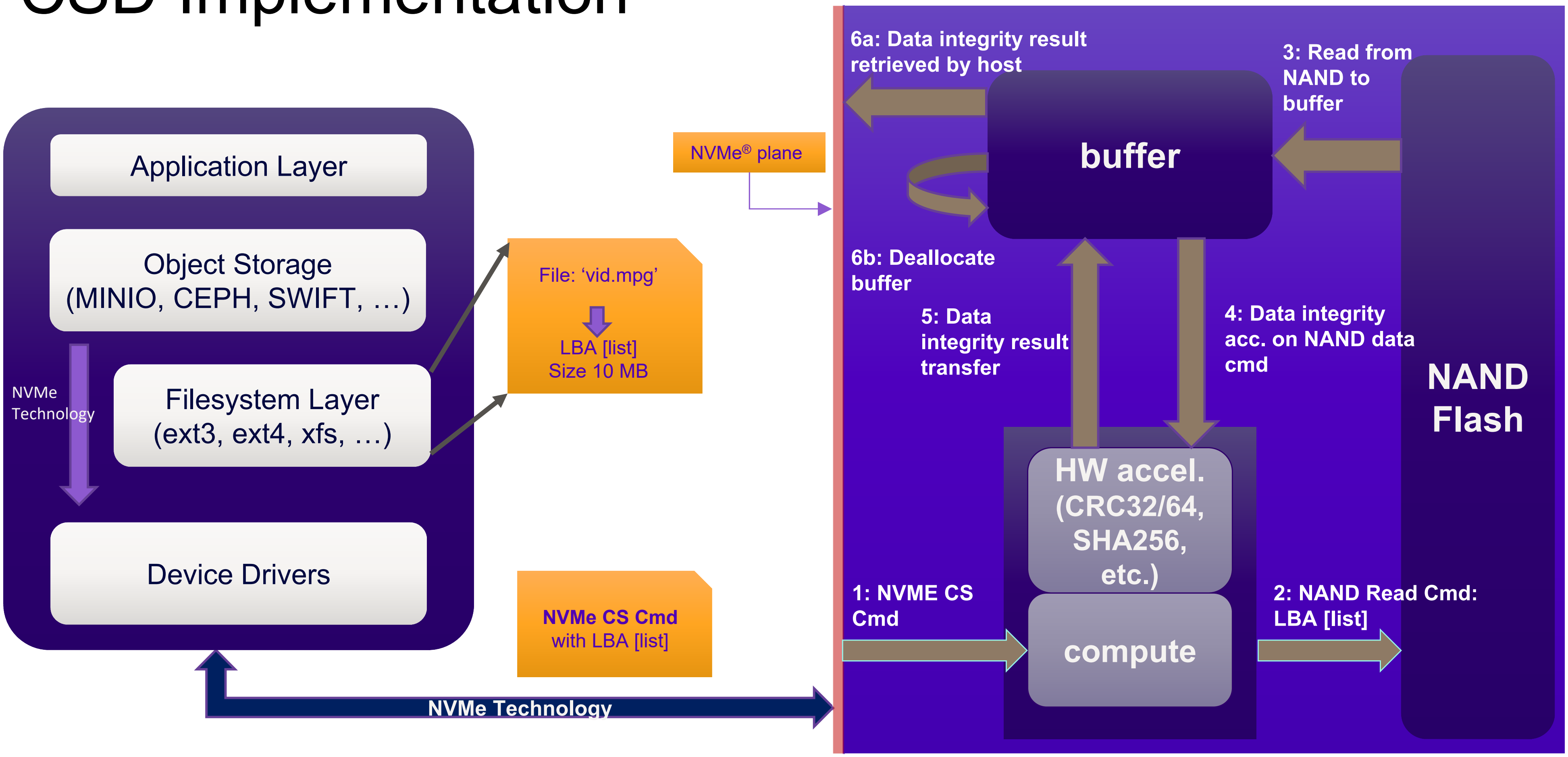
- ✓ Off-load the host
  - ✓ The host is only interested in the data integrity check results
- ✓ Reduces PCIe® traffic
  - ✓ No need to consume bandwidth and power to move the raw data to the host
- ✓ Reduces host memory footprint
  - ✓ All data required for processing is contained in the drive
- ✓ Scalable with storage
  - ✓ Performance increases as drives are added

Software Stack

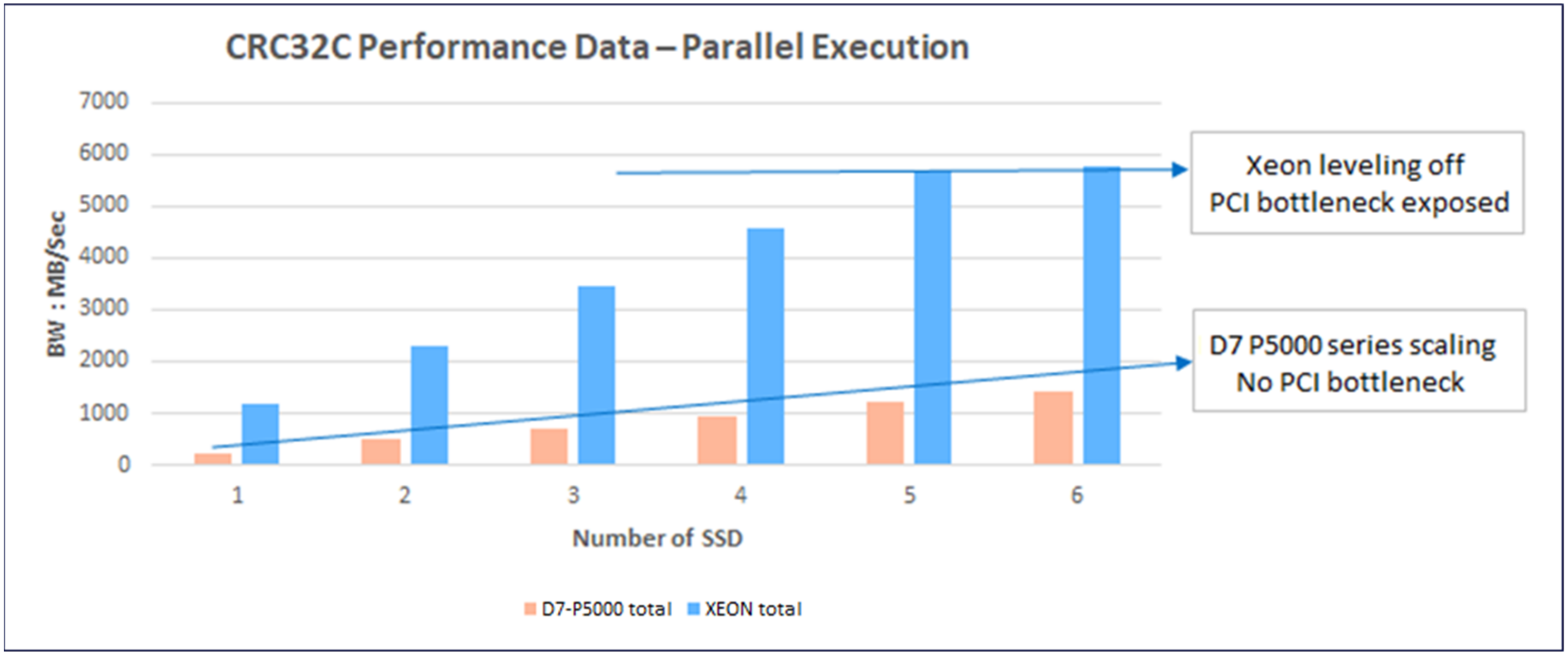


- Data integrity validation of the object shard is decoupled from data transfer.
- Data integrity hash calculation is done by the CSD.
- The object storage node validates the result.

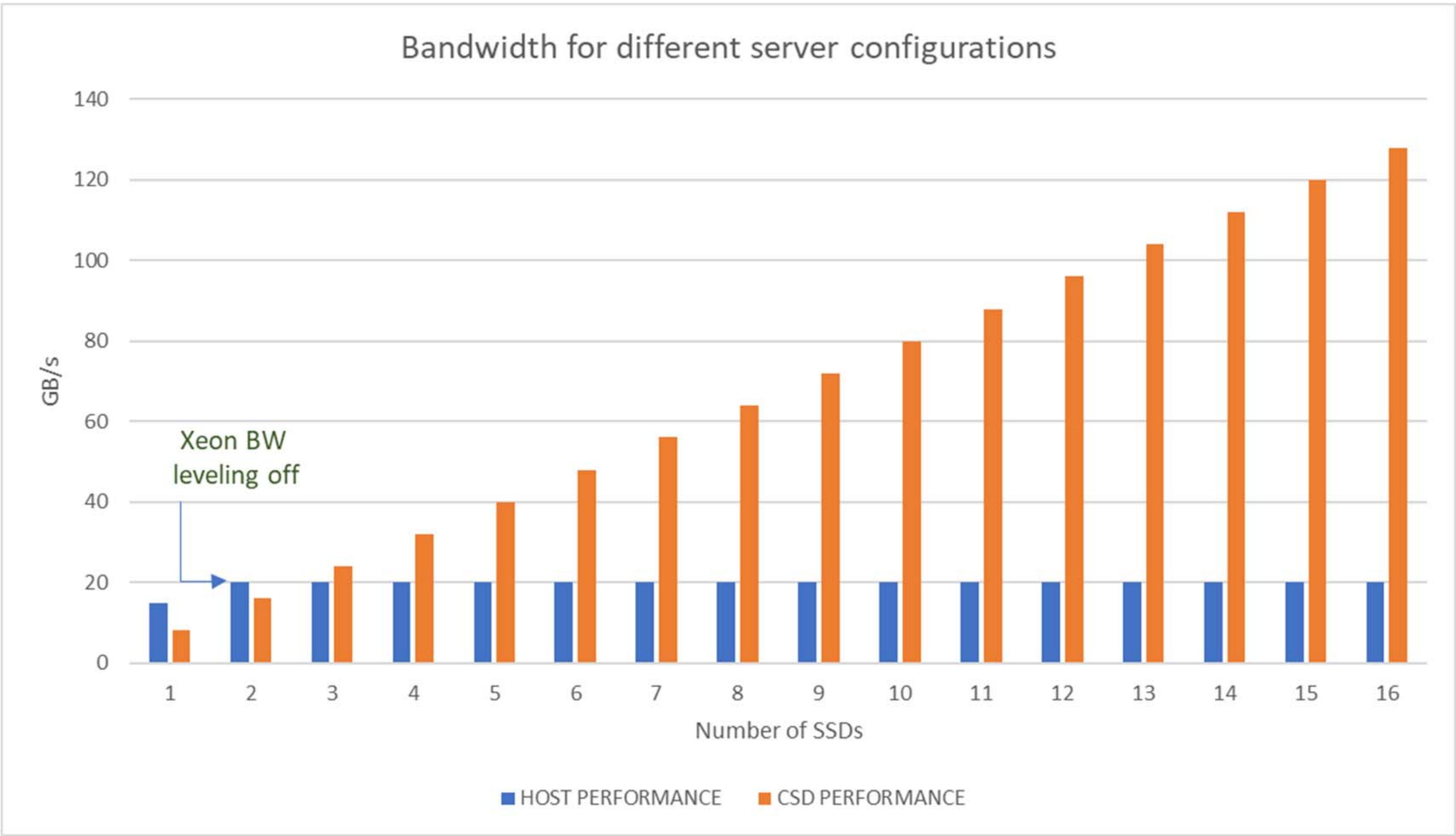
CSD Implementation



Performance and Scalability (1)



Performance and Scalability (2)



Modeling results indicate a high degree of scalability ideal for CPU offload*

*Based on Solidigm's internal analysis. CSD theoretical performance based upon modeling.

Conclusion

- NVMe® Computational Storage is ideal for processing low-level tasks on-drive
  - Utilizes existing HW accelerators and SW solutions, no ‘new’ work required
  - Shared data not an issue for the data integrity use case
    - **Major value add to customer’s concerns of data locality**
  - Scales across multiple CSDs
    - **Works independently, but brings overall increased performance to system**
    - Linear scaling performance with additional drives

- This use-case does not restrict the host from using resources for other computational work on the drive

Enterprise Data Resilience with FlashCore Module and NVMe® Computational Storage

FlashCore Module – A Computational Storage Device

FCM Offering →

Enterprise QLC
2.5" Dual-port
NVMe® SSD
U.2 Form Factor

Hardware Compression and Decompression

Hardware AES-256 Encryption with TCG Opal

FIPS 140-3 L2 Certification in progress

Used exclusively In IBM Storage Appliances
(for now…)

Small

4.8TBu → 22TBe
PCIe Gen3

Medium

9.6TBu → 28.8TBe
PCIe Gen3

Large

19.2TBu → 57.6TBe
PCIe Gen4

XLarge

38.4TBu → 115.2TBe
PCIe Gen4

Market Leading Density

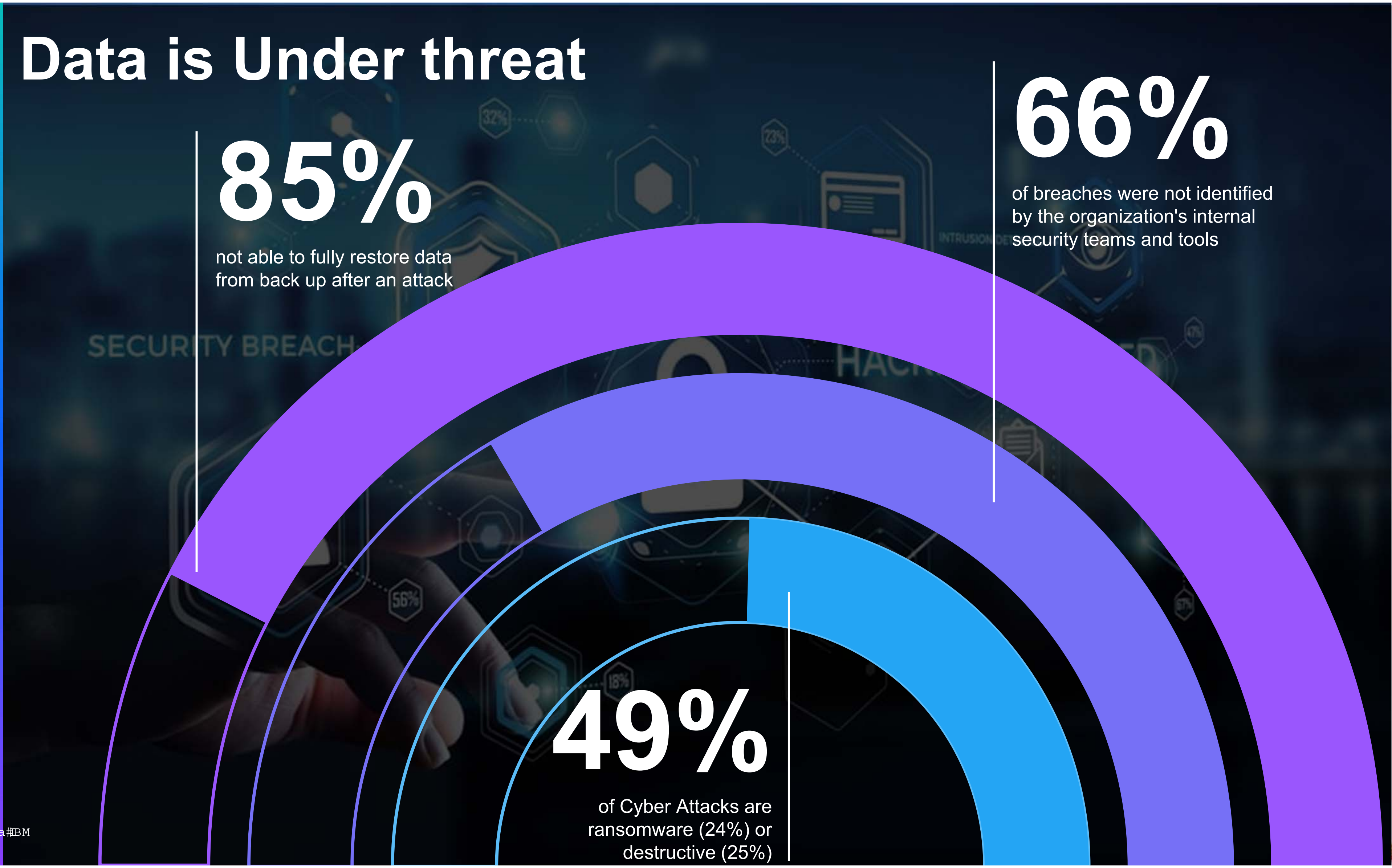
Best Price Point

Best Value

Enterprise Performance and Endurance with *ONLY* QLC since 2020
7 Year Endurance Guaranteed

Secure & Resilient

Data is Under threat



85%

not able to fully restore data from back up after an attack

66%

of breaches were not identified by the organization's internal security teams and tools

49%

of Cyber Attacks are ransomware (24%) or destructive (25%)

What does this have to do with Computational Storage?

An SSD has visibility to the data flowing in the system…

- FCM can Compress and Encrypt - with little to no performance penalty
- Various transforms can be made on the data received by the FCM - with little to no performance penalty
- FCM has intimate knowledge of access patterns, types of data transfer, and IOPs
- Additional Information can be passed to the FCMs: volume, file, etc.
- Trends and predictions can be intelligently made

This makes one consider:

Block Storage is missing some context
other parts of the system have…

BUT: It can compute data needed for identifying
Intrusion attacks with less performance impact
then any other part of the system

Which makes for a very interesting ***Computational Storage*** application!

SSDs can help in early Threat Detection !!!

- Ransomware
- Wiperware
- Exfiltration (stealing data but not hurting it)

A Breakthrough in Cyber Security

FCM-4
FLASHCORE MODULE 4

FCM Computational Storage (CS) Infrastructure

FCM has implemented a standardized process for computational storage, with an infrastructure built on proposed NVMe® technology concepts:

NVM Express® Computational Programs
NVM Express® Subsystem Local Memory

While our use case is still primarily within our own appliances, where we control the software stack, this will allow the FCM to efficiently adapt and modify our computational programs for future use cases and platforms.

Intrusion Detection is the next BIG thing in Computational Storage

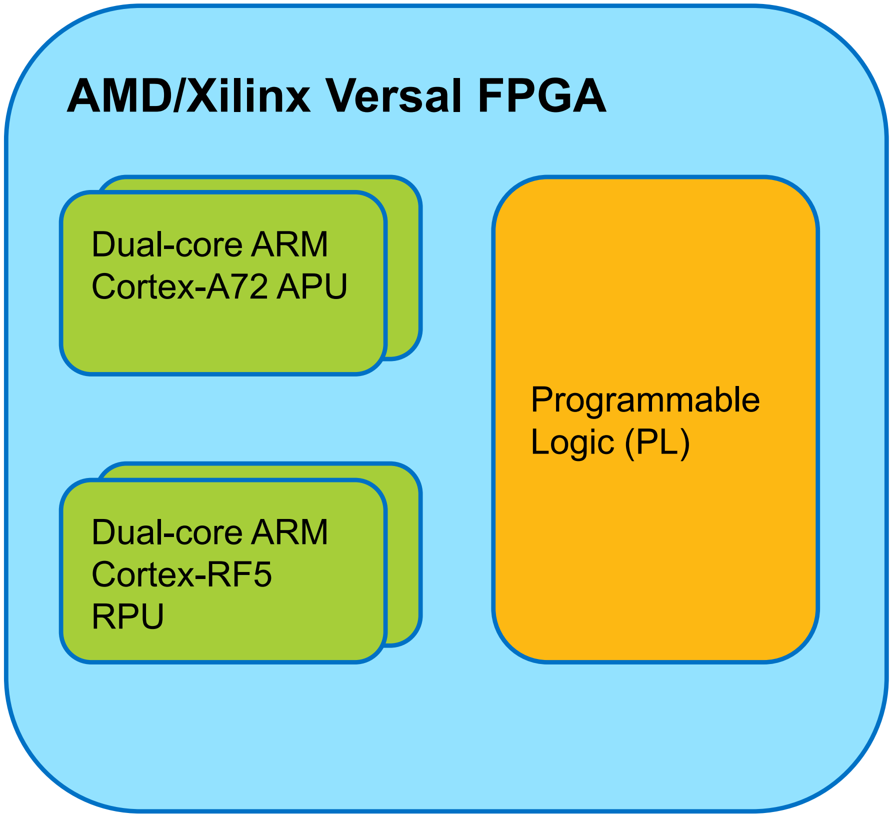
- FCM will analyze different aspects of the data and I/O and provide analysis to FlashSystem Spectrum Virtualize software
- FlashSystem will provide alerts of possible attacks and use Safeguarded Copies to recover data

FCM Controller Overview

- FCM is an FPGA based controller design utilizing AMD/Xilinx Versal technology
- The FCM's primary function is an Enterprise QLC SSD that includes compression, encryption, high-performing FTL, and other features
- This Computational Storage use case is designed to not interfere with the storage performance or function

Computational Components

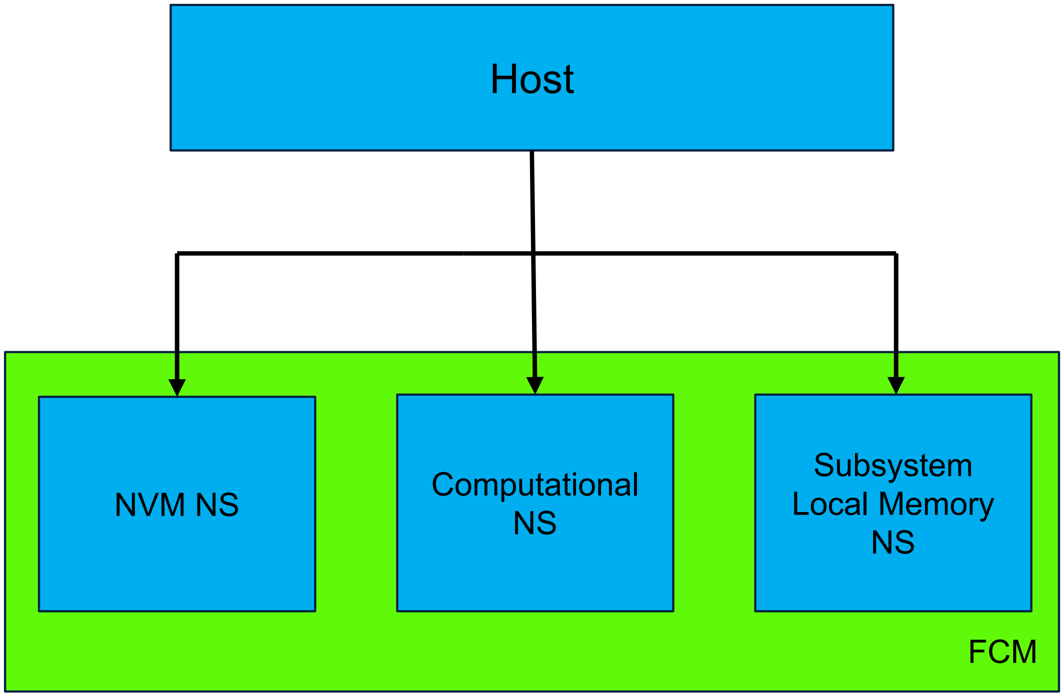
- APU(s), RPU(s), and programmable logic all play a role in the computational storage infrastructure



Computational Memory Host Interface and Namespaces (NS)

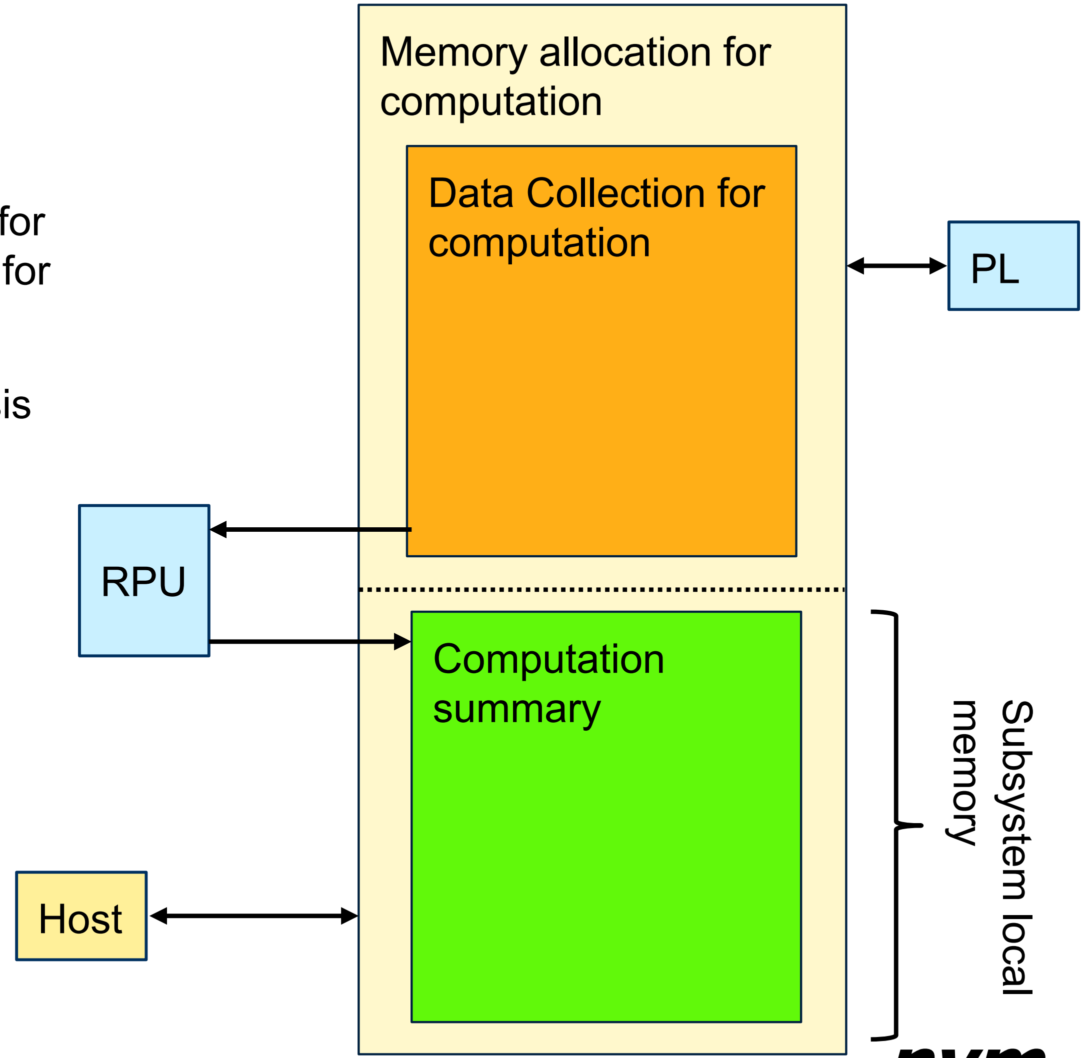
- FCM NVMe® layer will implement a subset of the commands detailed in Computational Programs and Subsystem Local Memory as well as required commands present in the NVMe 2.0 base specification

- FCM will create two additional namespaces on boot. One will conform to the Computational Programs Namespace definitions and the second will conform to the Subsystem Local Memory Namespace definitions



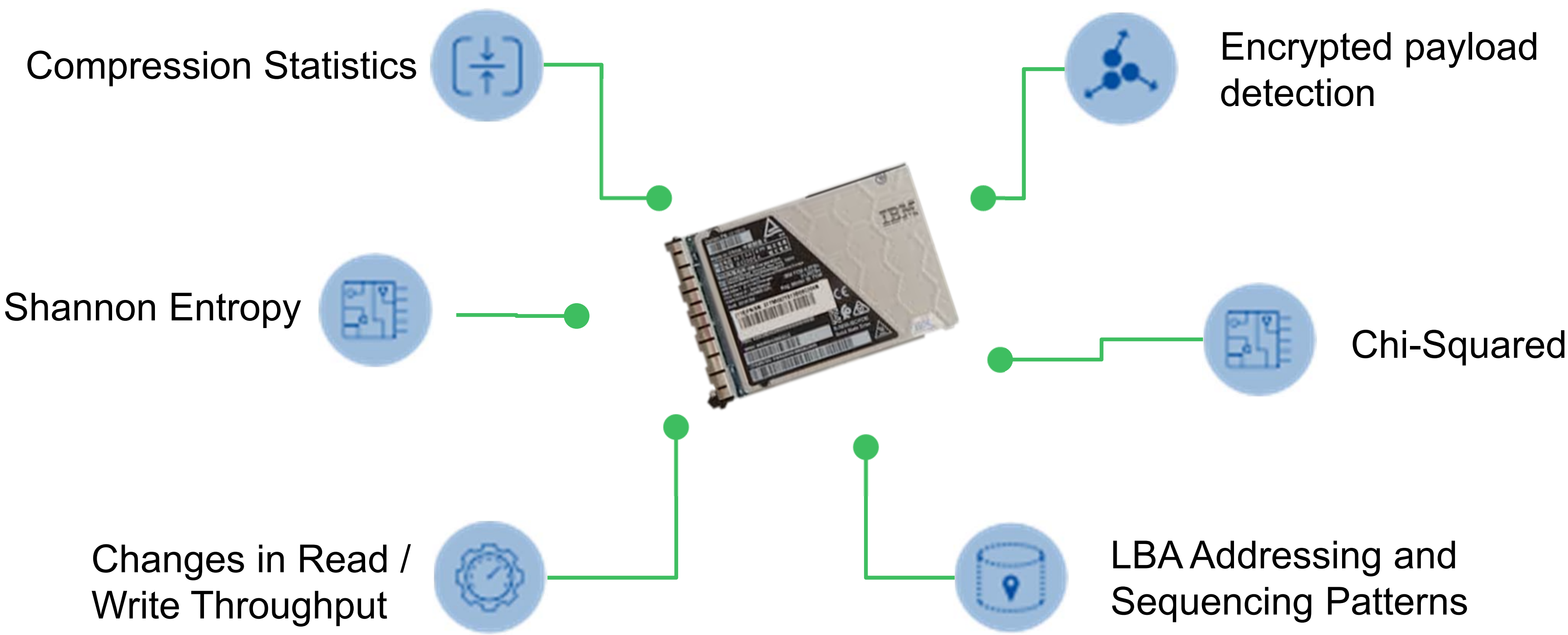
Memory Allocation

- FCM CS infrastructure will include a statically allocated memory region, with memory barriers, for each RPU as well as a dedicated register space for each RPU
- This region will be used to gather data for analysis and summarize the data so that it can be consumed by host or system API
- The PL will be responsible for gathering the data for analysis
- The RPU(s) will be used to analyze, summarize, and format the data for consumption
- The APU(s) will provide the setup, configuration, and management of the CS Infrastructure



Ransomware Threat Detection With FlashCore Module

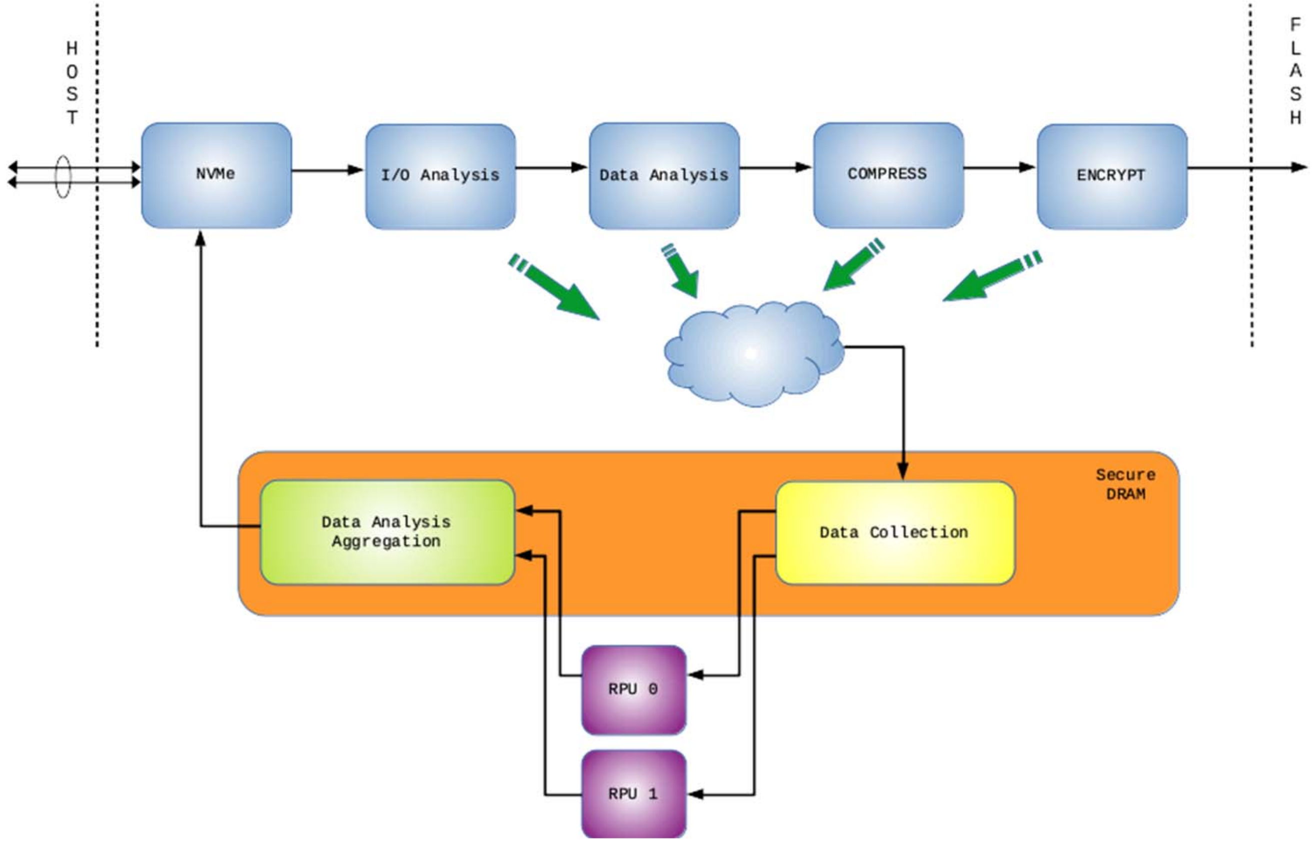
40+ data statistics analyzed in detection engine



Processed on EVERY I/O with ZERO performance impact!

FCM Intrusion Analysis Summary

- The FCM is the IDEAL place to track and analyze data for intrusion detection
- FPGA based controller allows for easy addition of new modules to deal with new threats
- The amount and type of data being collected will lead to an unprecedented amount of information about the work your SSD is doing in your environment!!!



AFA Computational Storage Platform

IBM is Committed to Cyber Resilience throughout the entire storage infrastructure

- IBM Safeguarded Copy
- IBM Storage Defender
- IBM Storage Sentinel
- FCM Intrusion detection hints

Reduce/Eliminate False Positives!!!

- Top reason for Storage Admins to ignore or turn off Threat Detection Software
- FCM hints provides increased accuracy
- In our FlashSystem products, hints from multiple FCMs will be combined for even more confidence!

This work will revolutionize the detection, accuracy, and recovery speed of a malicious attack!

The Future of Storage-Based Resiliency and Recovery Using Computational Storage

Questions?

nvm EXPRESS®

Architected for Performance