FMS
the Future of Memory and Storage

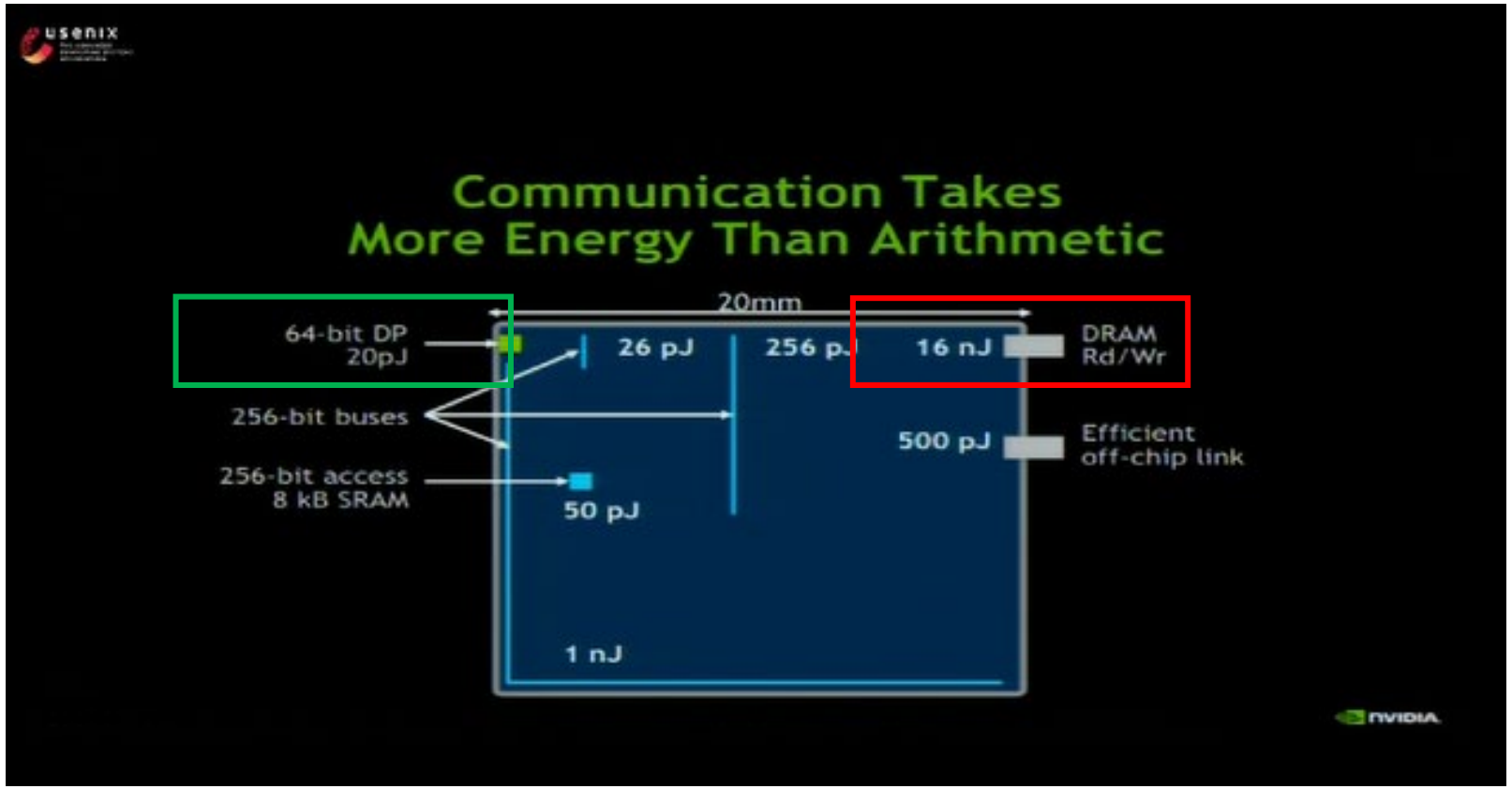
Achieve Significant Reduction in Data Movement by Offloading Data Scrubbing

Presenter: Satvik Vyas

Strategic Marketing, Data Center SSD, KIOXIA

KIOXIA

Data Movement is More Expensive Than Compute



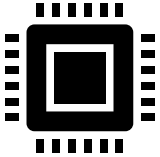
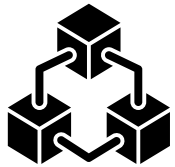
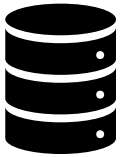
Moving data to compute could consume >100X energy than the compute itself.

Source: William J. Dally, Stanford and NVIDIA through USENIX open access media
https://www.usenix.org/conference/hotpar13/workshop-program/presentation/dally

Data Growth Is Exploding

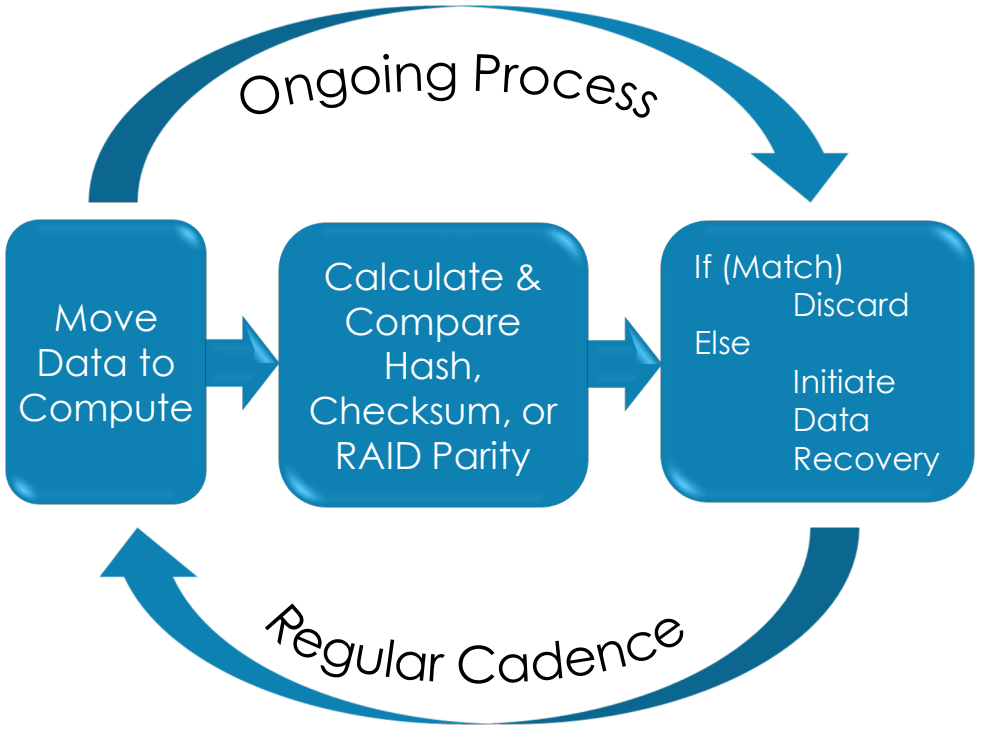
In the time it takes you to read this sentence, the amount of data generated will have grown to be equivalent to the amount needed to hold the entire written works of mankind from the beginning of recorded history until now, in every language, several times over!

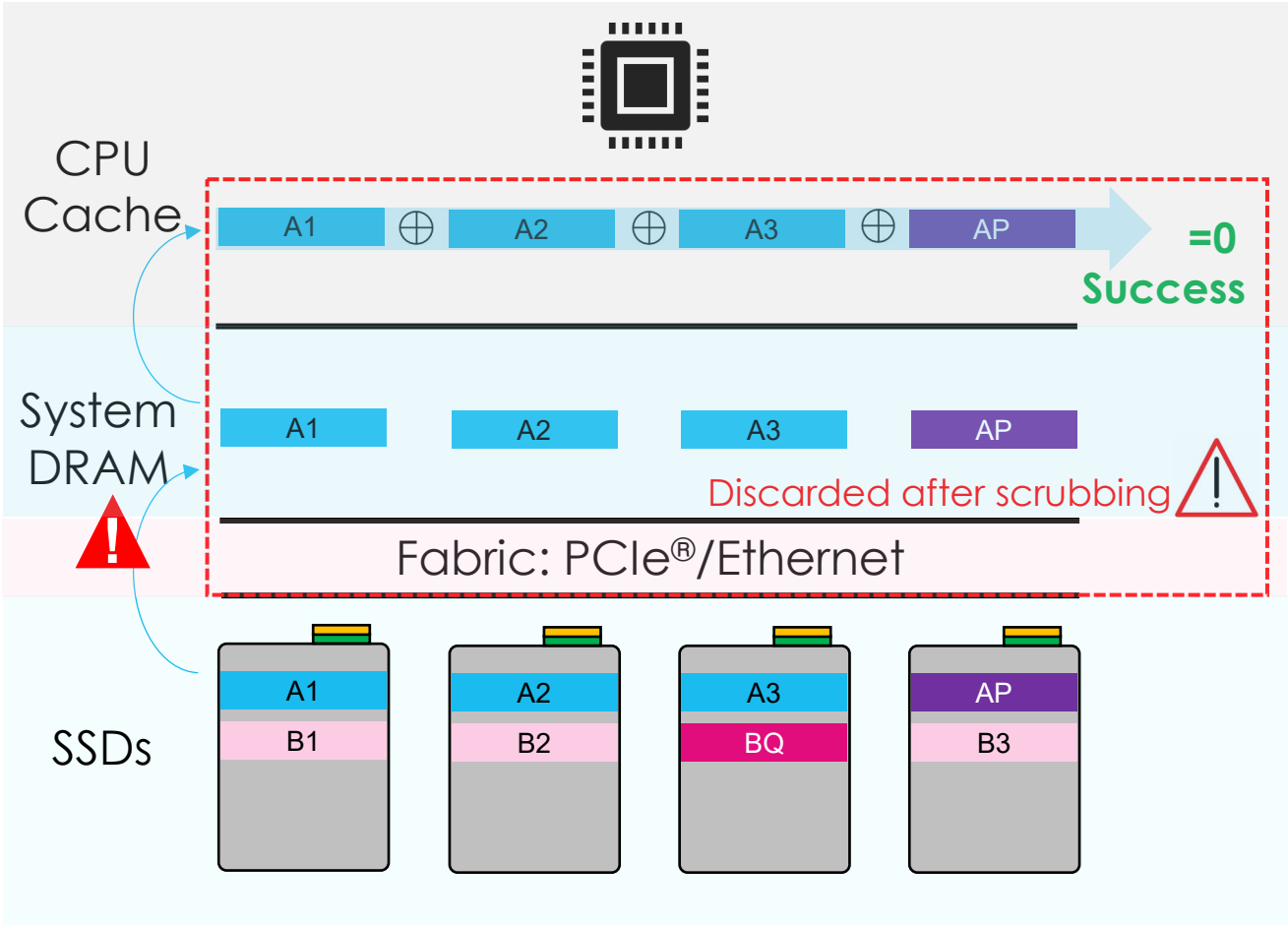
Exploding Growth of Data is Creating New Challenges

	Process	Access	Store
Trend & Enabling Tech	 • More Cores per xPU • Heterogeneous Computing • Edge Computing • FPGA/ Accelerators • CXL®, CM, PIM, HBM • Computational Storage	 • 5G, 400/800 Gigabit Ethernet • SmartNIC, Data Processing Unit • Ultra Accelerator Link • Ultra Ethernet	 • Higher Bit Density (QLC) • Storage Disaggregation • PCIe® Gen5, NVMe-oF™ • Compression, Dedup, EC/RAID • Computational Storage
Challenges	• Feeding Data to Compute While Maintaining Memory Bandwidth per Core	• Power, Security, Latency and Bandwidth to Transfer Data to and from Compute	• Performance/Gigabyte (GB), Endurance, and Reliability
Sustainability	How to efficiently scale-up before scale-out?	How to reduce data movement?	How can storage play a role in improving sustainability?

Ensuring Data Integrity at Scale

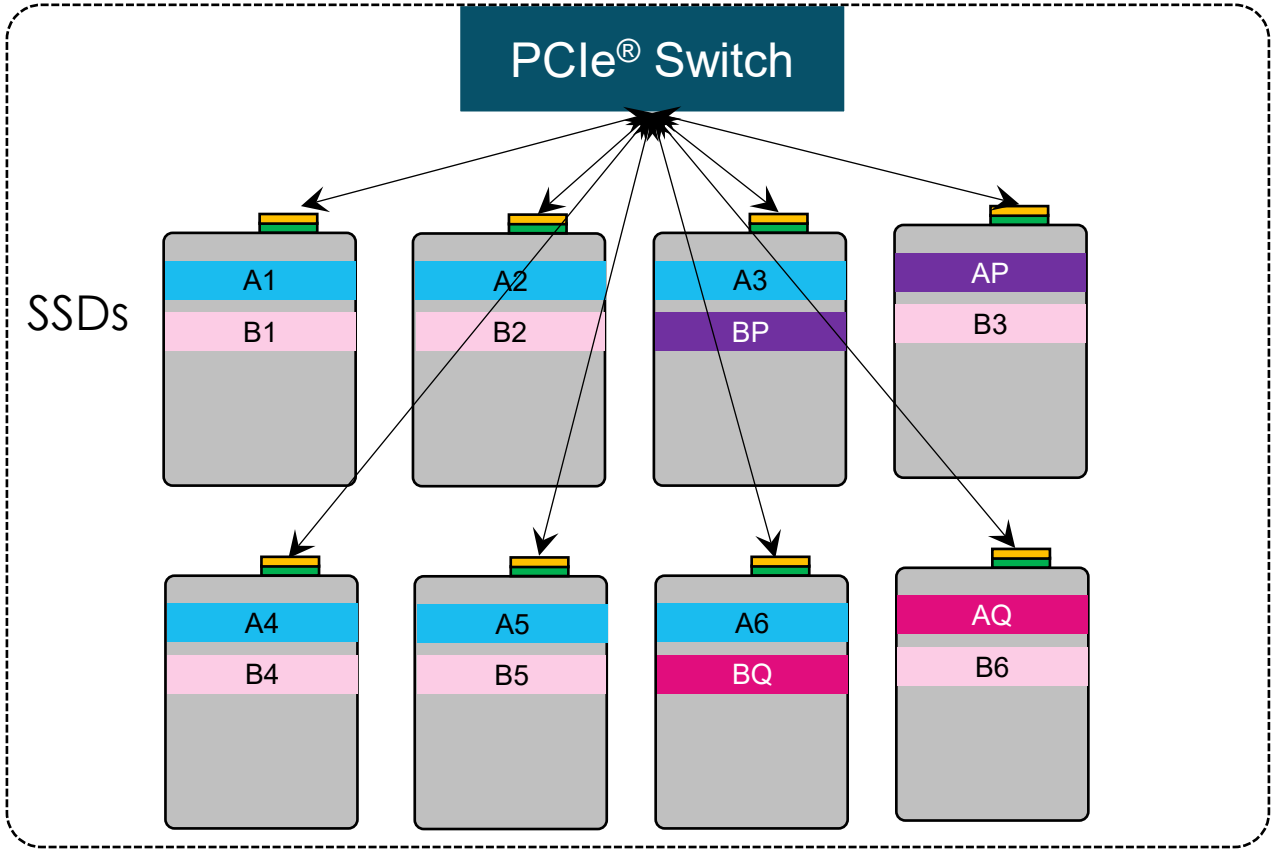
Data Scrubbing: Early detection and correction of errors



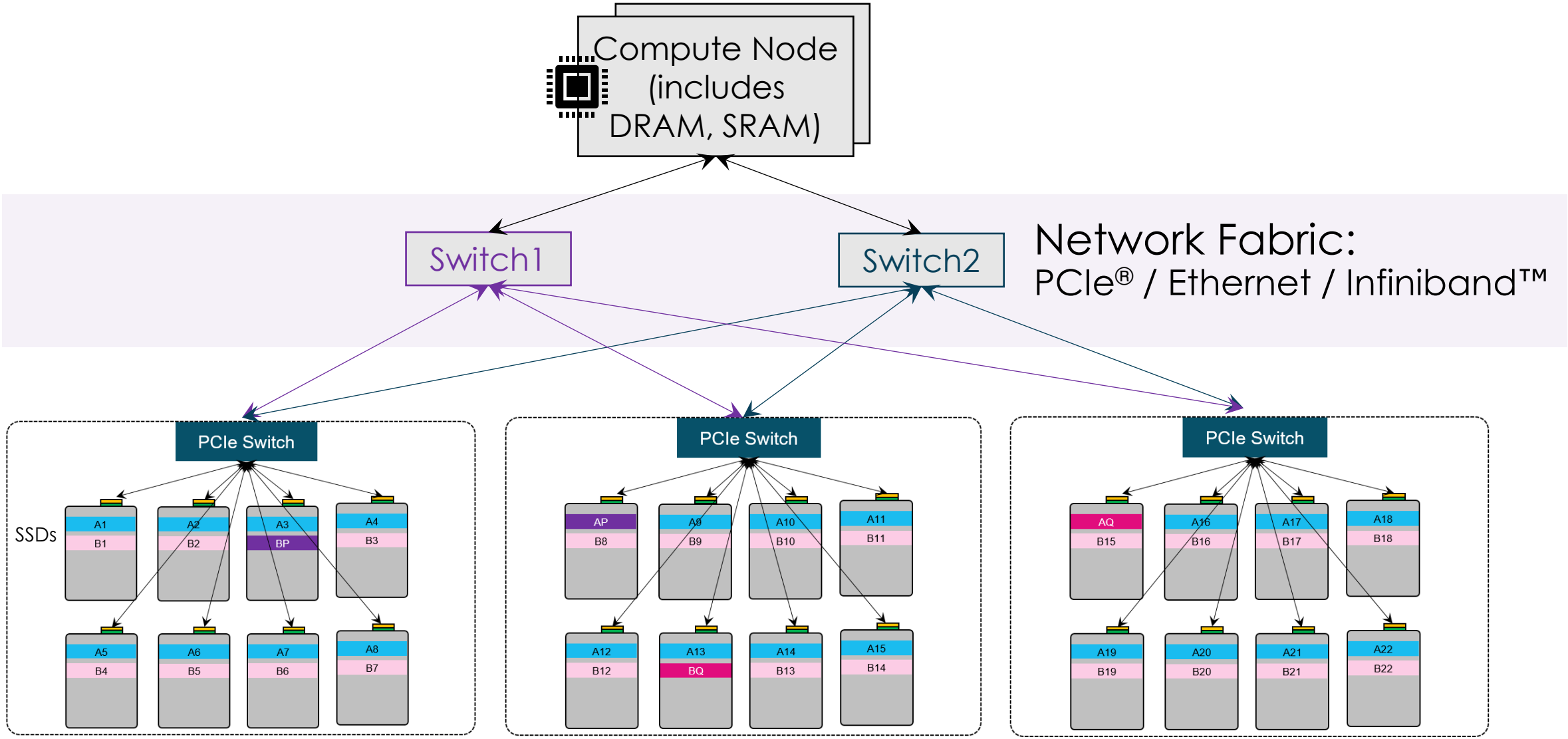


Data Scrubbing is an overhead penalty paid to ensure data integrity.

More Complex in Real Practice



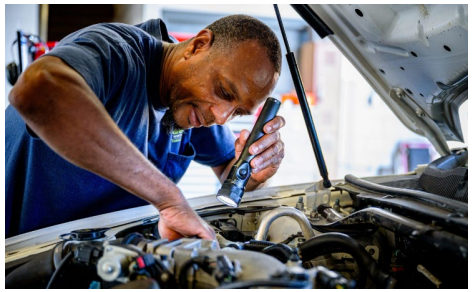
More Complex in Real Practice



Ensuring Data Integrity at Scale

Exabytes of data are scrubbed regularly where the desired outcome constitutes 100% throw-away work.

Data Scrubbing Analogy



Visiting a mechanic every week for car inspection even if you're not aware of any problems.

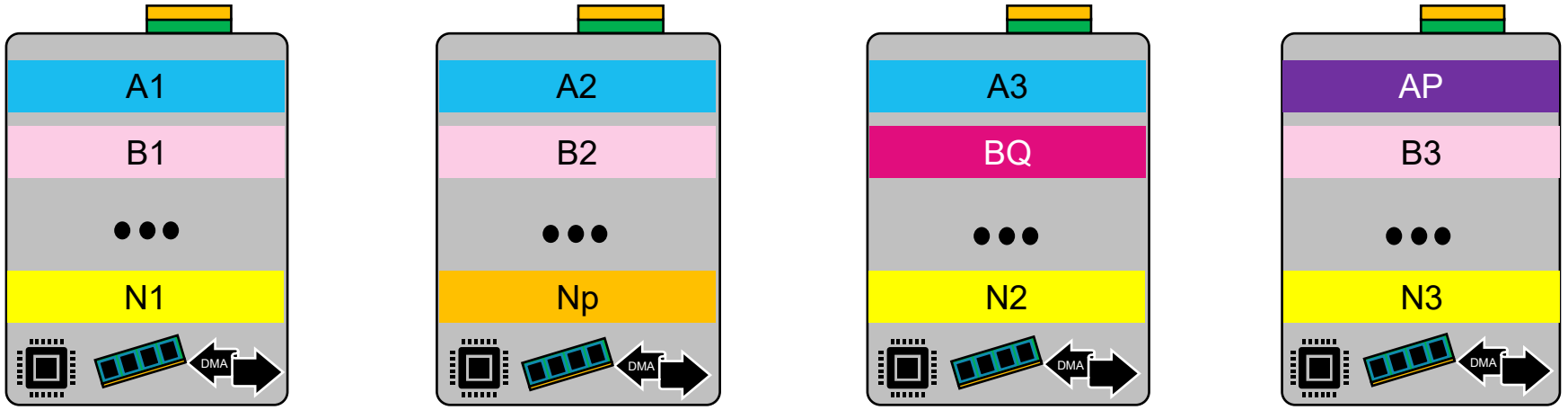
Better approach is that you take a car to a mechanic when built-in diagnosis warns you of a problem.

Image sources:
Engine mechanic, Public domain image, no copyright required, available from creator under creative commons license
Check engine, Wikiuser100000, permission to use under creative commons license
https://commons.wikimedia.org/wiki/File:Vw_engine_check.jpg

Data Scrubbing: Offload Approach with RAID

Host orchestrates data scrubbing
leveraging a standards-based approach.

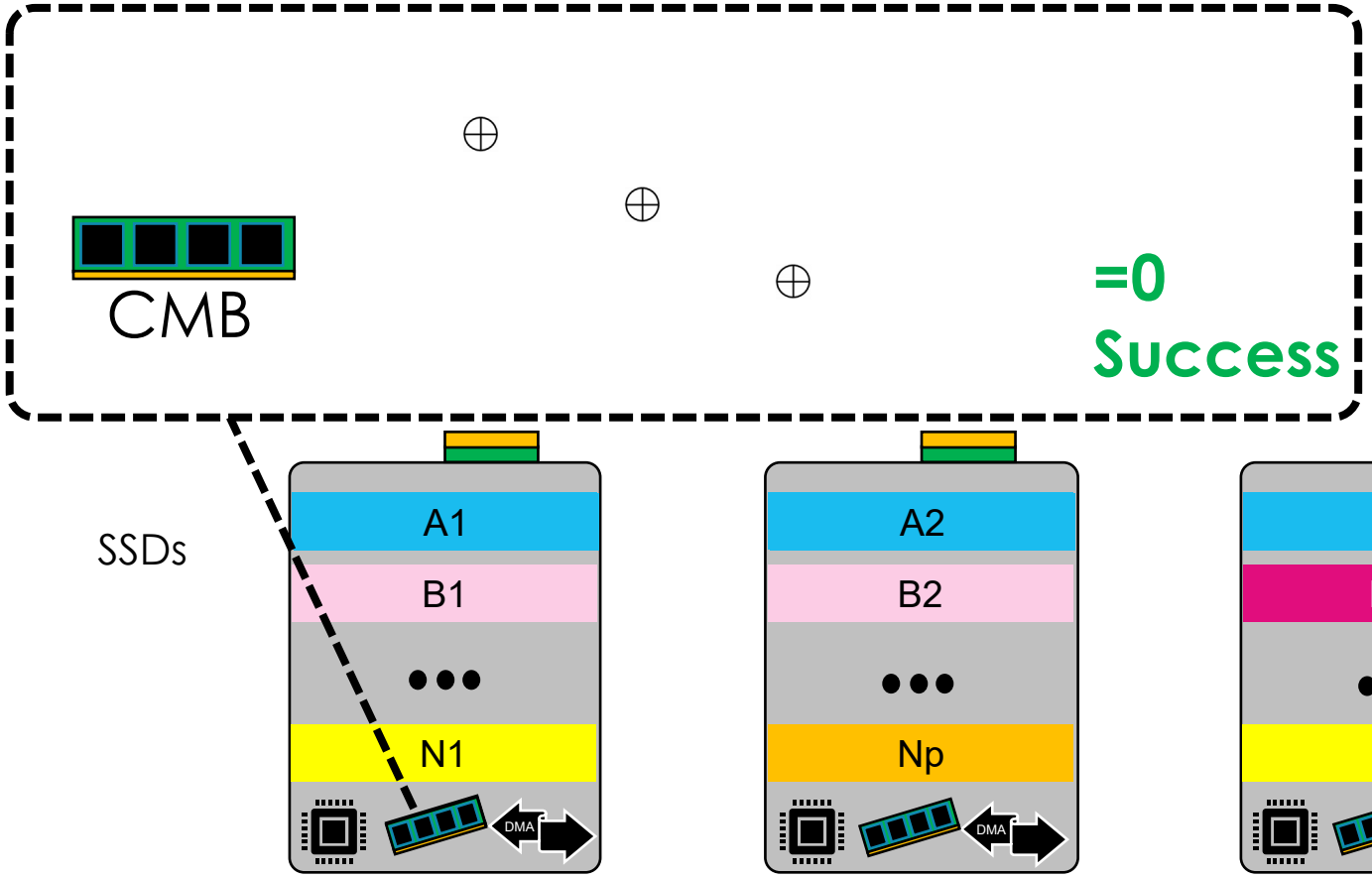
SSDs



Parity Compute Engine, Controller Memory Buffer (CMB), and
Direct Memory Access Controller (DMAC)

Royalty-free icon through Microsoft® 365 subscription.
Representation of DRAM module created by KIOXIA.

Data Scrubbing: Offload Approach #1 with RAID



- Move data into one CMB
- XOR[1]
- Communicate result back to host

Benefit:

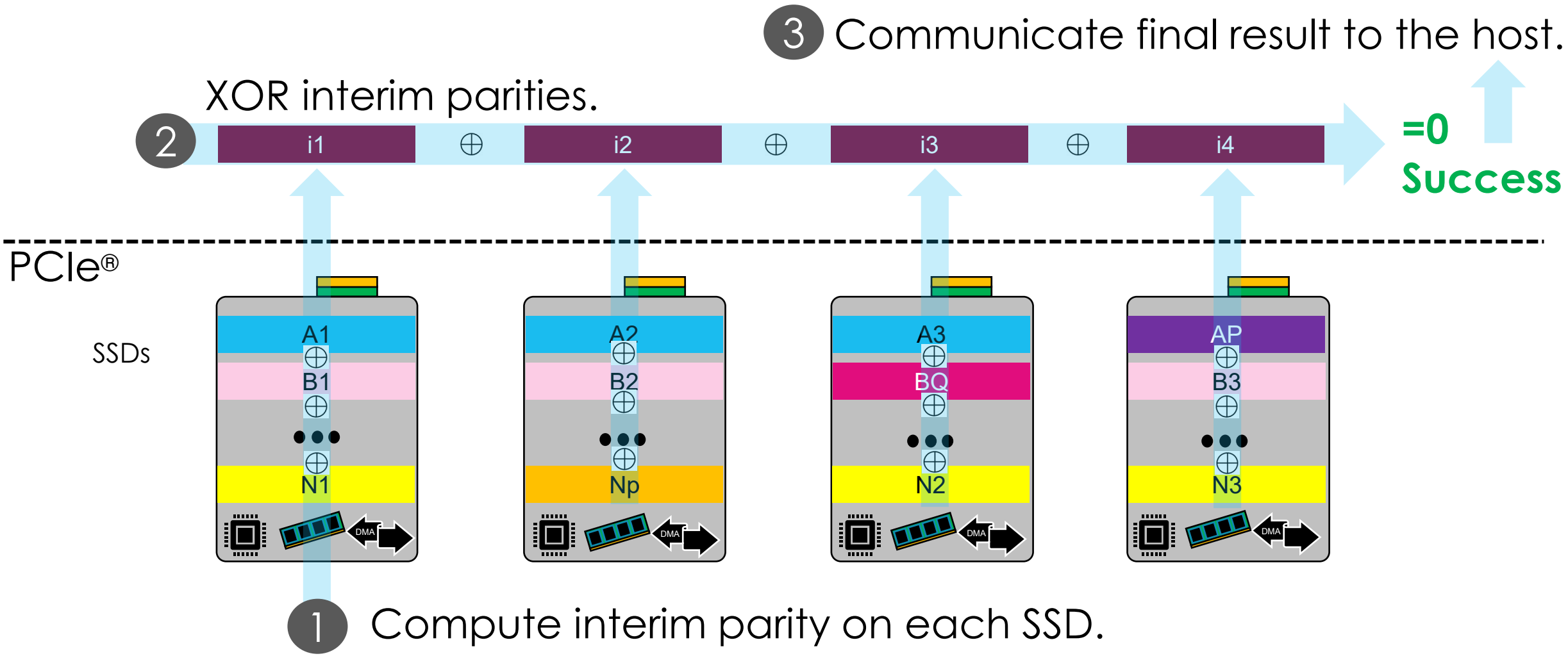
Significant reduction in data movement

Parity Compute Engine, CMB, and DMA Controller

Only the Results of Scrubbing Operation Cross PCIe®.

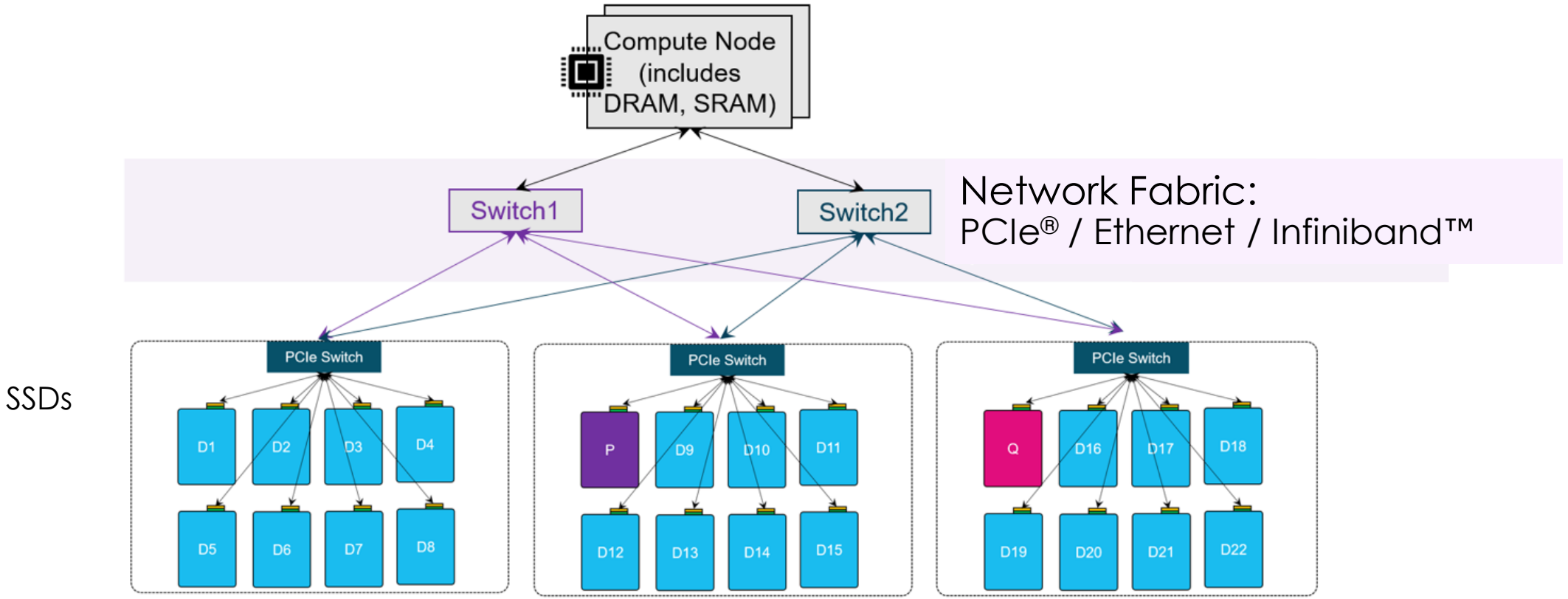
1 A binary operation, stands for Exclusive OR.

Data Scrubbing: Offload Approach #2 with RAID



Further Reduce Data Movement by Scrubbing Multiple Stripes in Parallel.

Example Showing Benefits of Data Scrubbing Offload



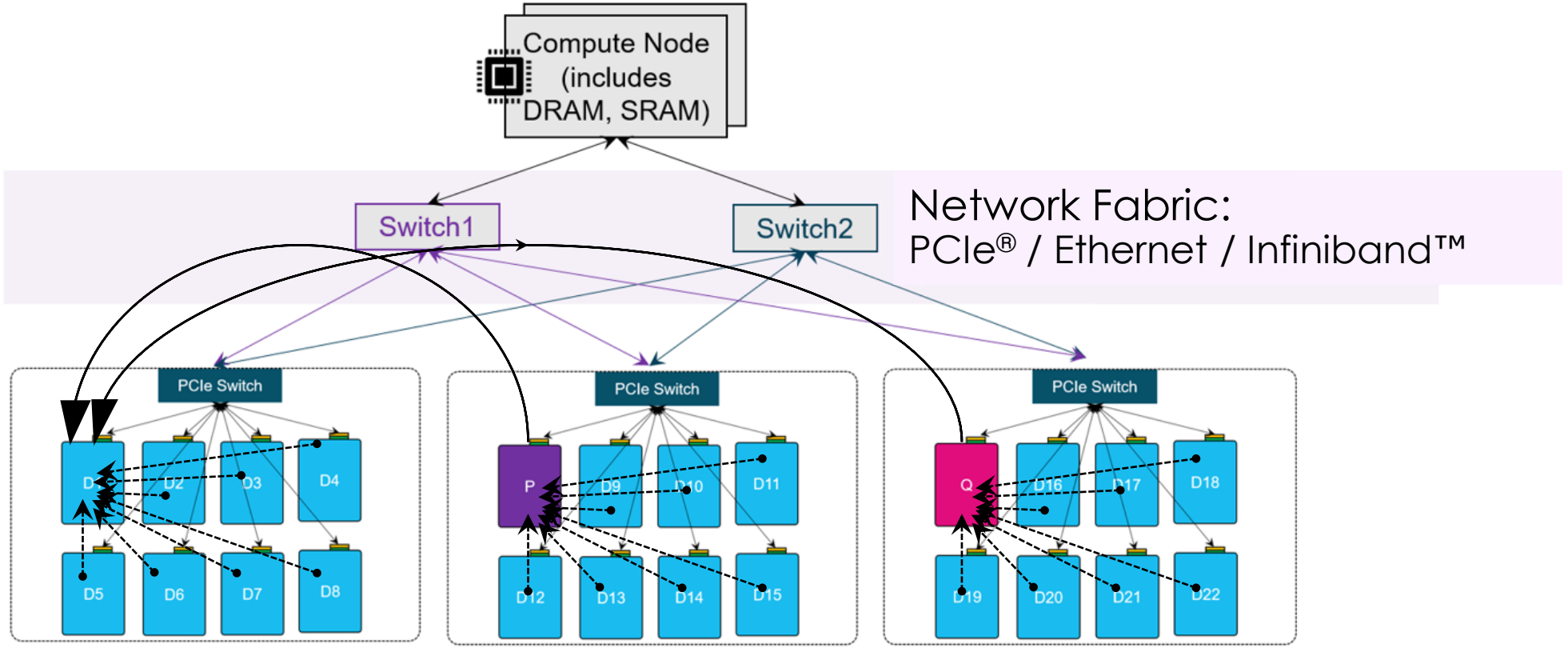
RAID Setup:

- 22 Data & 2 Parity Segments

- Parity_P: D1⊕D2⊕ …⊕D22
- Parity_Q: g1.D1⊕g2.D2⊕ …⊕ g22.D22

Example Showing Benefits of Data Scrubbing Offload #1

Approach#1



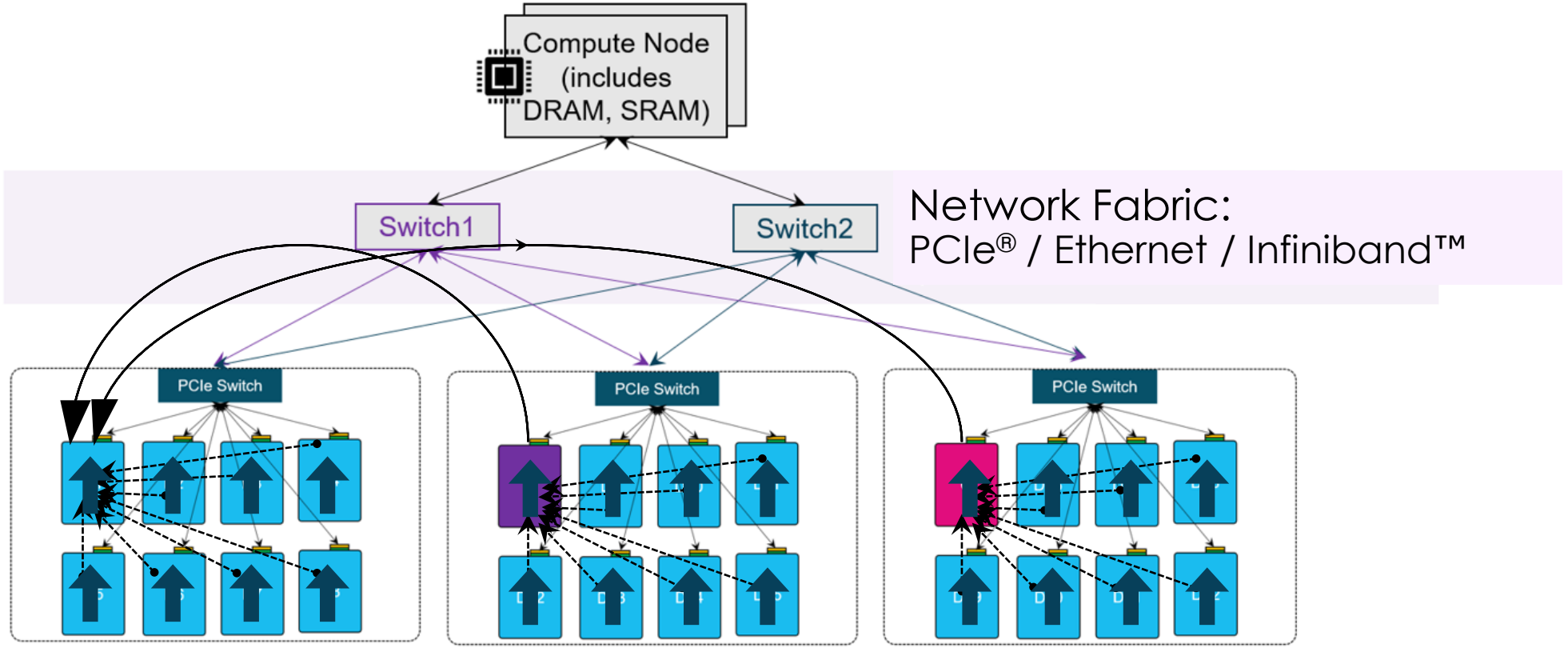
RAID Setup:

- 22 Data & 2 Parity Segments

- Parity_P: D1⊕D2⊕ …⊕D22
- Parity_Q: g1.D1⊕g2.D2⊕ …⊕ g22.D22

Example Showing Benefits of Data Scrubbing Offload #2

Approach#2

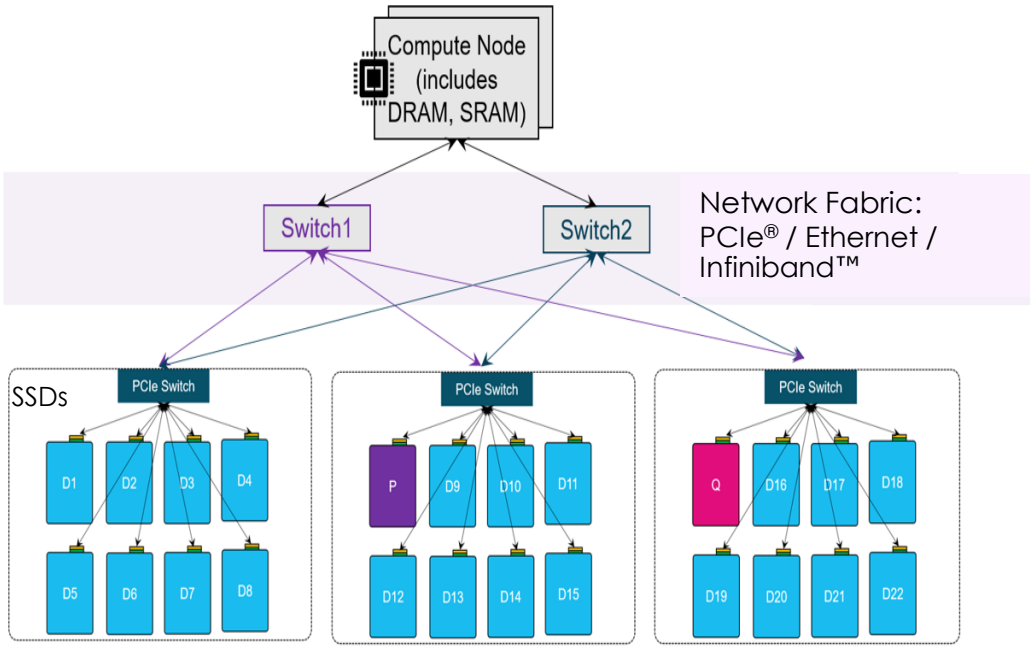


RAID Setup:

- 22 Data & 2 Parity Segments

- Parity_P: D1⊕D2⊕ …⊕D22
- Parity_Q: g1.D1⊕g2.D2⊕ …⊕ g22.D22

Example Showing Benefits of Data Scrubbing Offload



22 Data segments, 2 Parity segments

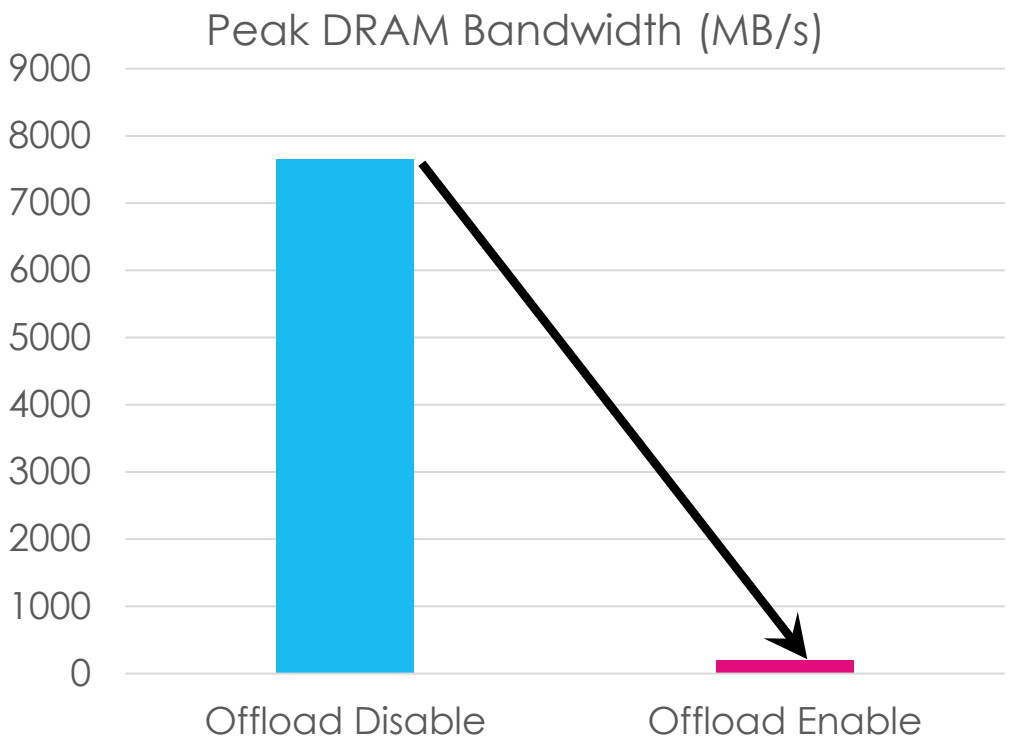
RAID Segments Passed Through For Parity Compute & Check…	Conventional Approach	Offload Approach #1 (1 Stripe per Scrub)	Offload Approach #2 (N Stripes per Scrub)
CPU	100%	0%	0%
DRAM	200%	0%	0%
Network	100%	12.5%	12.5%/N
PCIe	100%	100%	100%/N

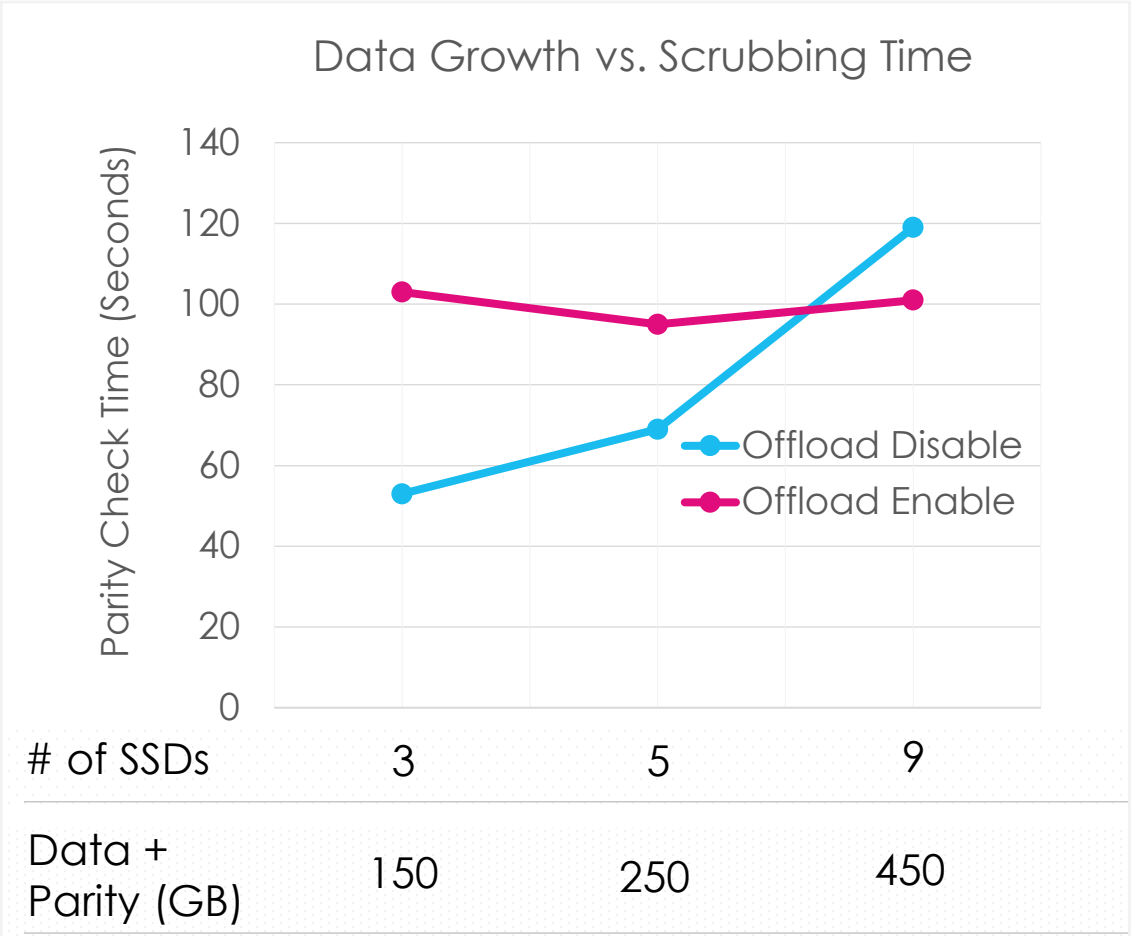
N = number of RAID stripes per Scrub

Significant Reduction in Data Movement. Alleviate Network and Memory Bandwidth Bottlenecks.

Proof of Concept (PoC) Results

150 GB Data+Parity scrubbed on 3 KIOXIA CM7 SSDs





# of SSDs	3	5	9
Data + Parity (GB)	150	250	450

~38X Reduction in Host DRAM Bandwidth Utilization

Data Scrubbing Capability (Compute & Memory) Scales with Each SSD

KIOXIA

Tests and results were conducted in-house by KIOXIA engineering team. Source of graphs: KIOXIA. KIOXIA Corporation defines a megabyte (MB) as 1,000,000 bytes, a gigabyte (GB) as 1,000,000,000 bytes and a terabyte (TB) as 1,000,000,000,000 bytes. A computer operating system, however, reports storage capacity using powers of 2 for the definition of 1GB = 230 = 1,073,741,824 bytes and therefore shows less storage capacity. Available storage capacity will vary based on file size, formatting, settings, software and operating system, and/or pre-installed software applications, or media content.

Future Possibilities: A Call to Action

- Offloading data scrubbing on SSDs can significantly alleviate memory and network bandwidth bottlenecks and reduce data movement.
- Make additional accelerators available to host through a standards-based approach.

Standards Based

Host Controlled

Hardware Accelerators
(Memory, Compute, DMAC)

Let's Collaborate! Visit Booth# 307