SNIA DEVELOPER CONFERENCE

SDC 25

By Developers FOR Developers

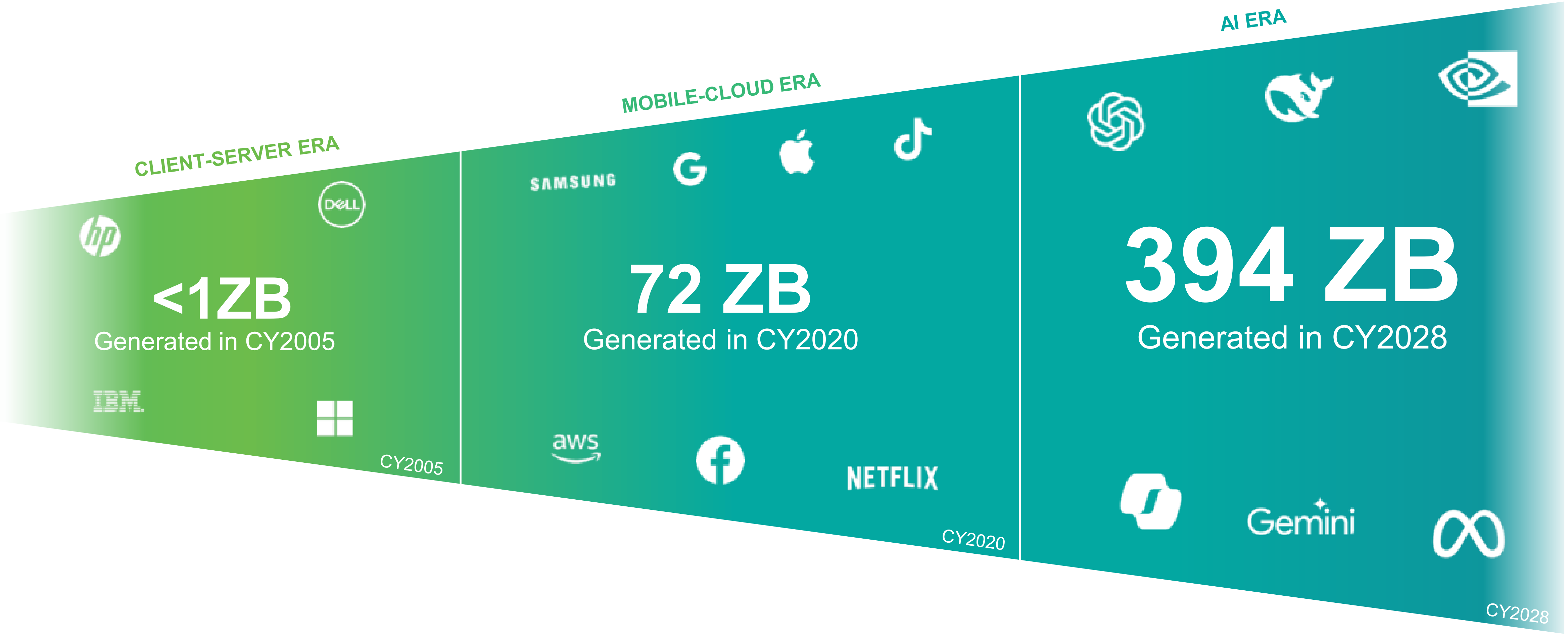
Hyatt Regency Santa Clara, CA
September 15-17, 2025

AI Driven Mass-Storage Evolution

Exploring the evolutionary technologies of the future of mass data-storage in the AI datacenters.

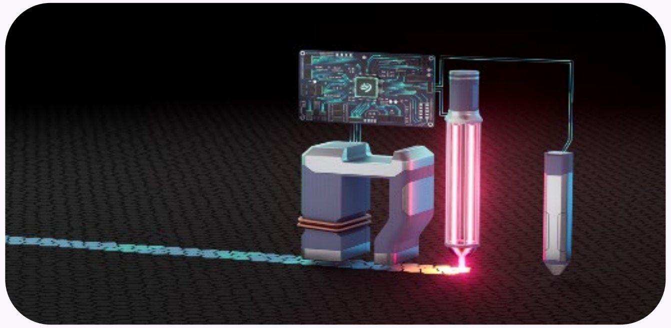
www.sniadeveloper.org

Technology Innovation is Generating Unprecedented Volumes of Data



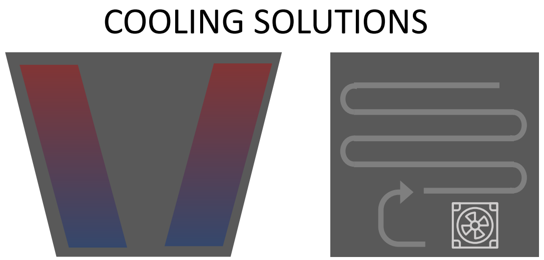
Source: IDC Worldwide Global Datasphere Forecast 2024-2028 doc #US52076424, May 2024.
Zettabytes represent the estimated annual volume of data created and replicated in the years: CY2005, CY2020, CY2028.

AI-Driven Mass-Storage Evolution



Nanoscale Complexity Deployment Simplicity

Efficiency through
Foundational Storage Technology
Superior Media Density Innovation



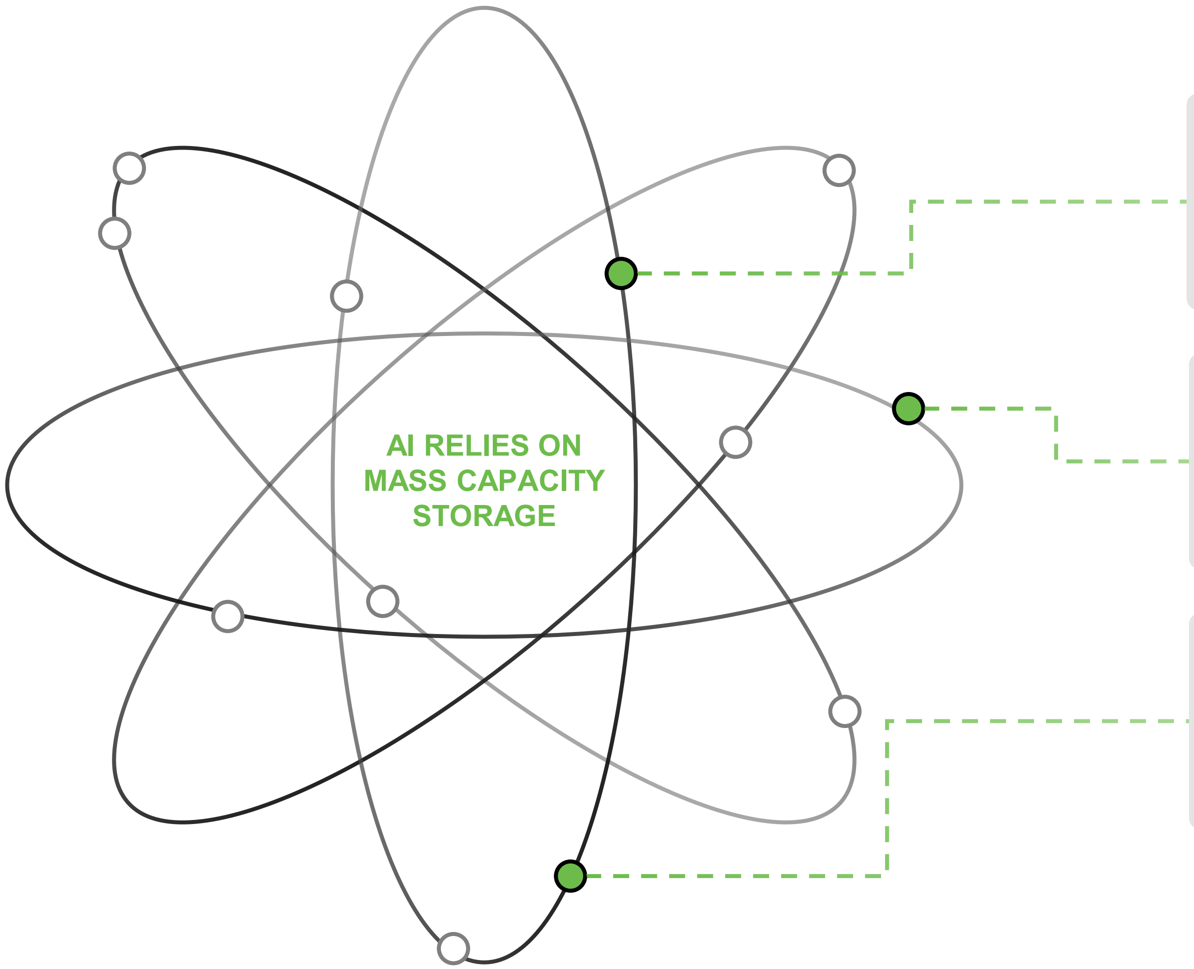
Mass Capacity Systems Innovation

Efficiency through
Sustainability Optimized
High Density Storage Systems

Software Driven TCO Optimization

Efficiency through
Memory & Storage Stack
OS & Application Optimization

AI Driving Next Wave of Mass Capacity Storage Demand

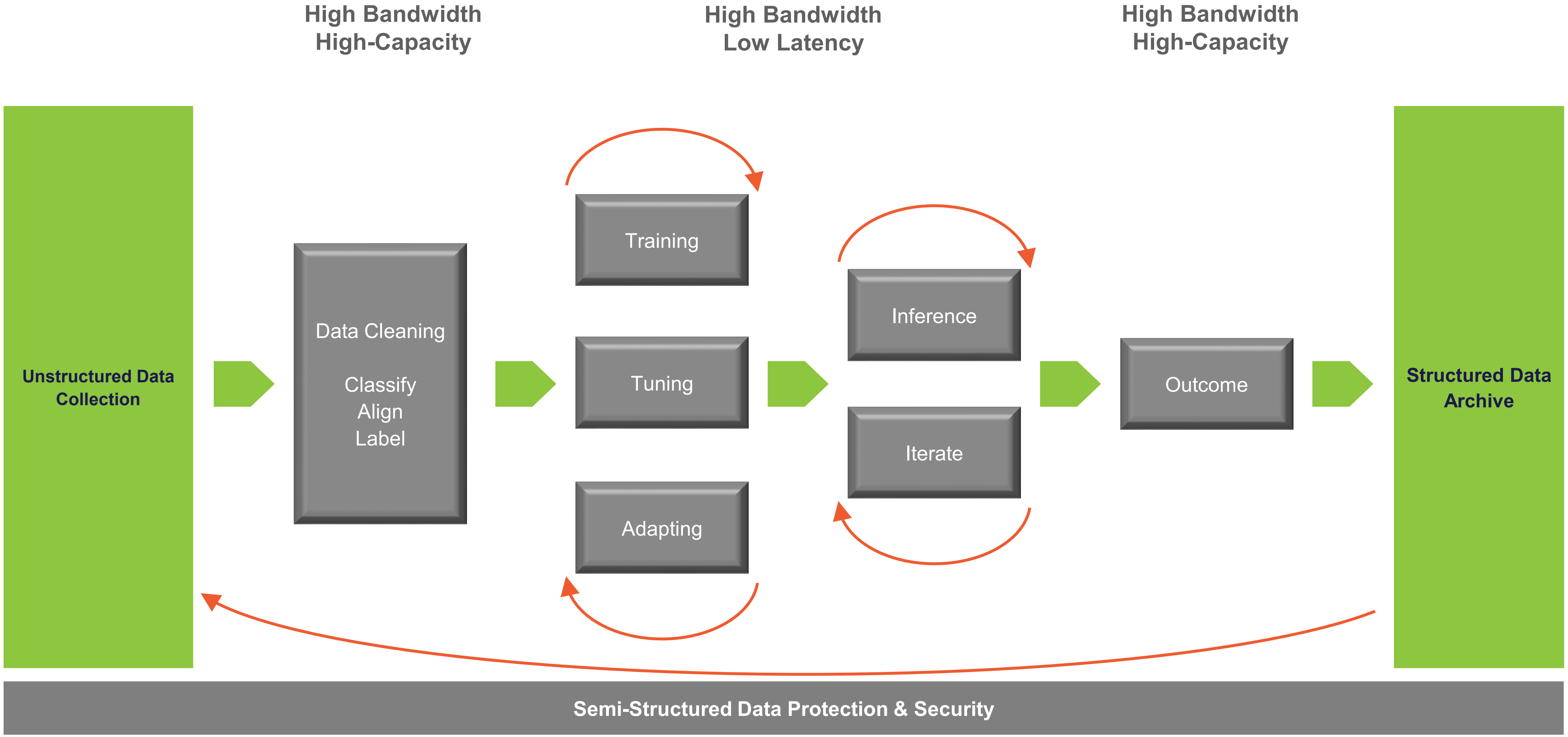


More data improves the quality of AI model output
Hard drives feed data to AI models

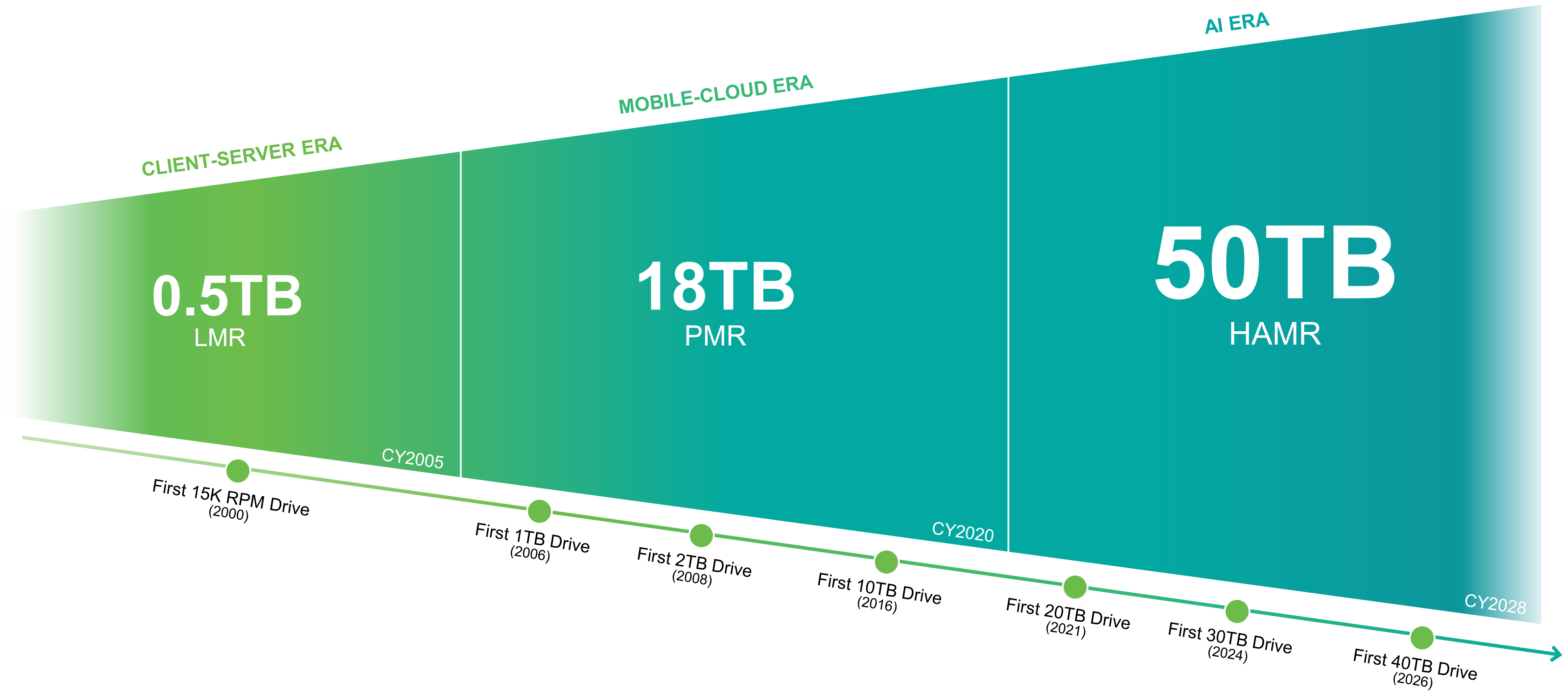
With AI, people and machines will generate data at a pace unlike any before
Hard drives preserve valuable content created from AI models

AI gets more intelligent in an infinite loop
Hard drives maintain AI model checkpoints and query data

AI Relies On Mass Capacity Storage



Seagate's Storage Innovation Enabling Every New Data-Driven Era

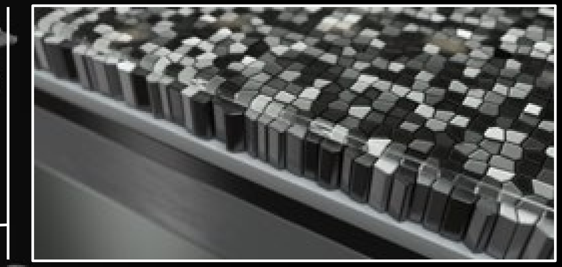


Represents Seagate's highest capacity drive shipped in CY2005, CY2020, and CY2028 (projected).

MOZAIC™ (HAMR)

The Value of Areal Density

Areal density is the **amount of data** that can be stored on the surface area of a hard drive disk.

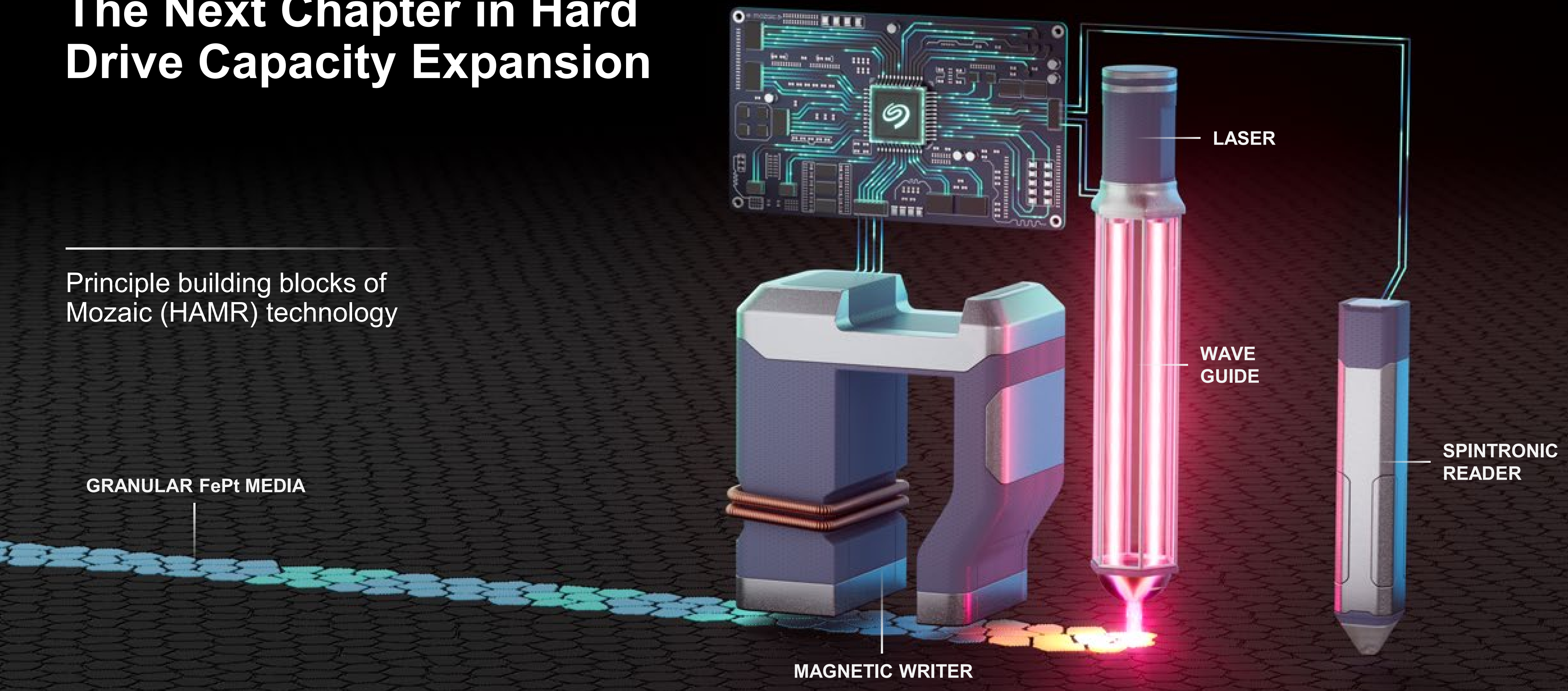


Cost-efficient Capacity Scaling

We are scaling drive capacity by storing more data on every disk. Not by adding more disks.

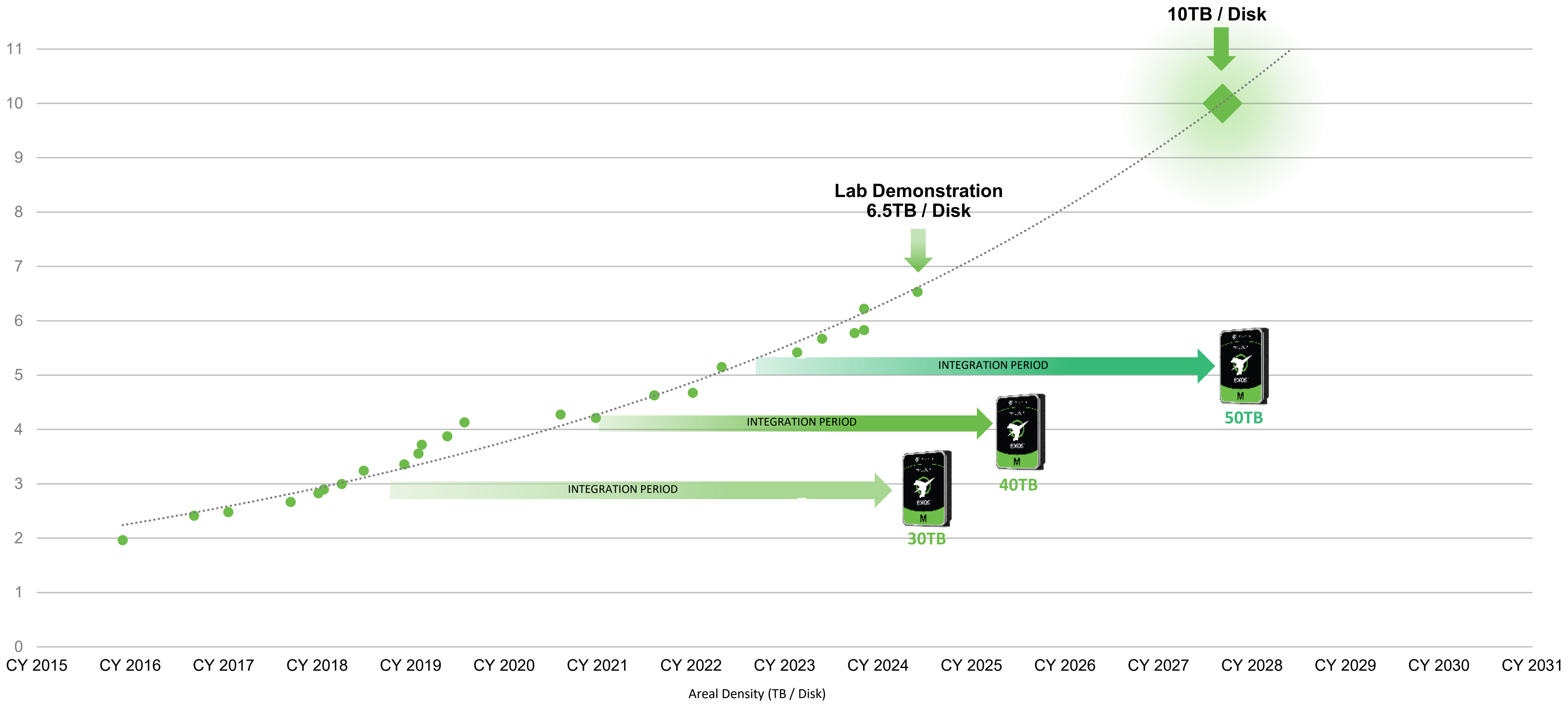
The Next Chapter in Hard Drive Capacity Expansion

Principle building blocks of Mozaic (HAMR) technology



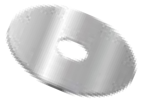
Product Maturity Cycle Within 5 Years

From Lab Demonstration to Productization



40% Improvement with Seagate Mozaic 4+ (HAMR)

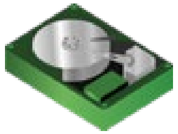
With a 1 exabyte deployment



AT THE DISK LEVEL

4TB / Disk
10 Disks

2.4TB / Disk
10 Disks



AT THE DRIVE LEVEL

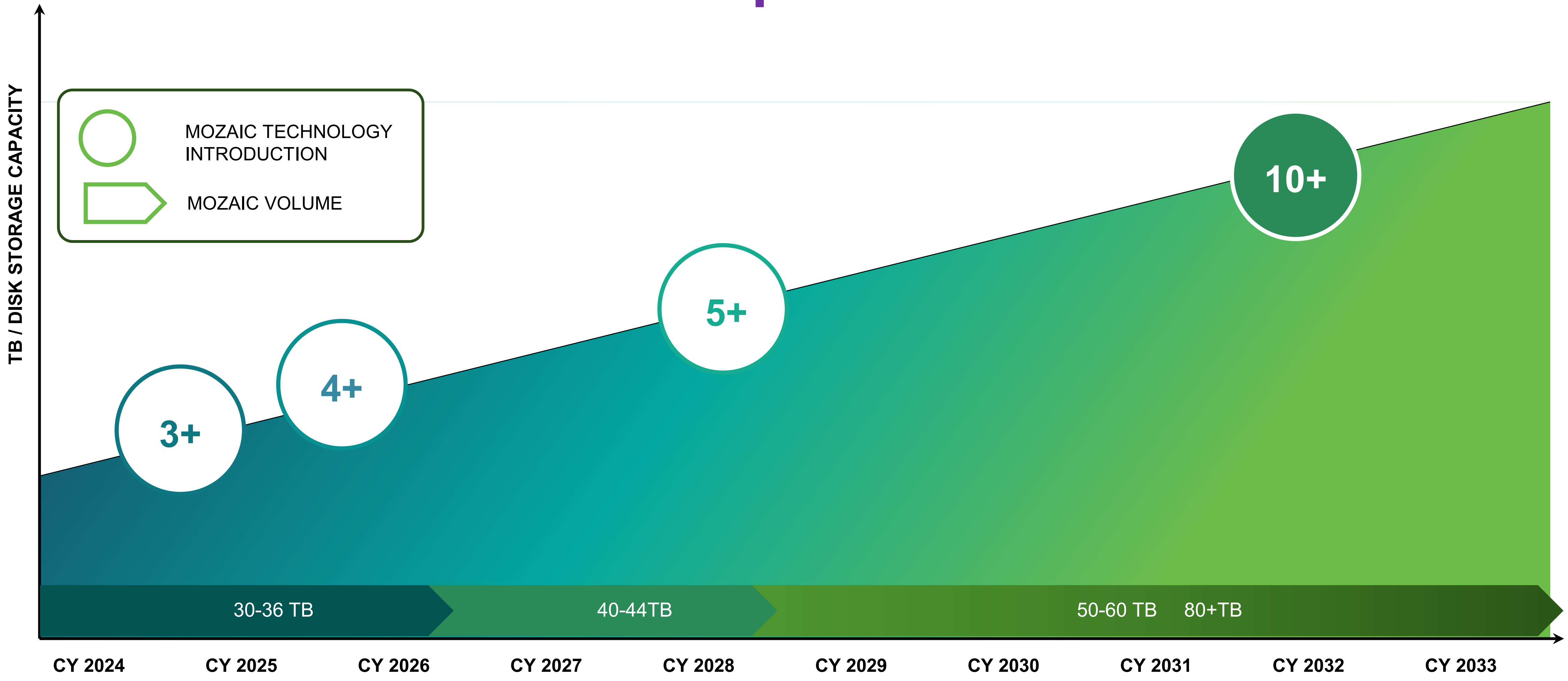
40TB / Drive

24TB / Drive

AT THE DATA CENTER LEVEL

25,000 Units	**240** Sq. Ft.	**~2.0M** kWh / Year
42,000 Units	**400** Sq. Ft.	**~3.3M** kWh / Year

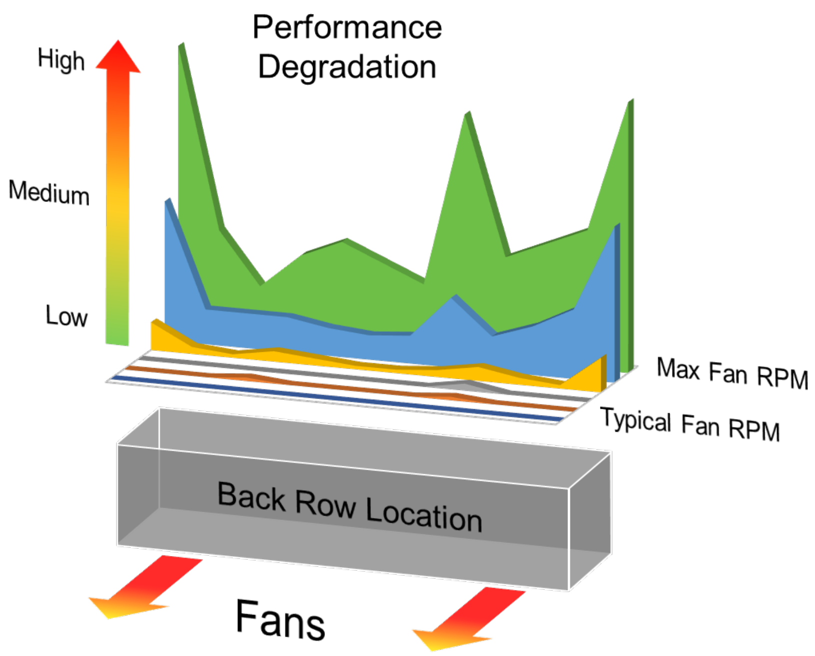
...And Accelerated Transition From Technology Introduction to Volume Shipments



High Density HDD Storage Main Pain Points

- Acoustic fan noise contamination in <u>high-density</u> storage enclosures can cause more Drive Errors

- Reducing the fan's rotation speed to minimize acoustics noise results in higher drive temperatures, which results in higher Annual Failure Rate(AFR)

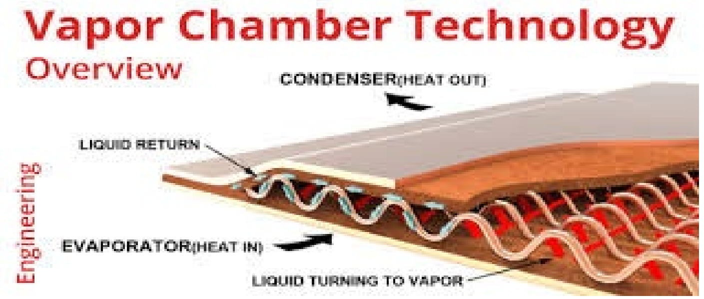


 |

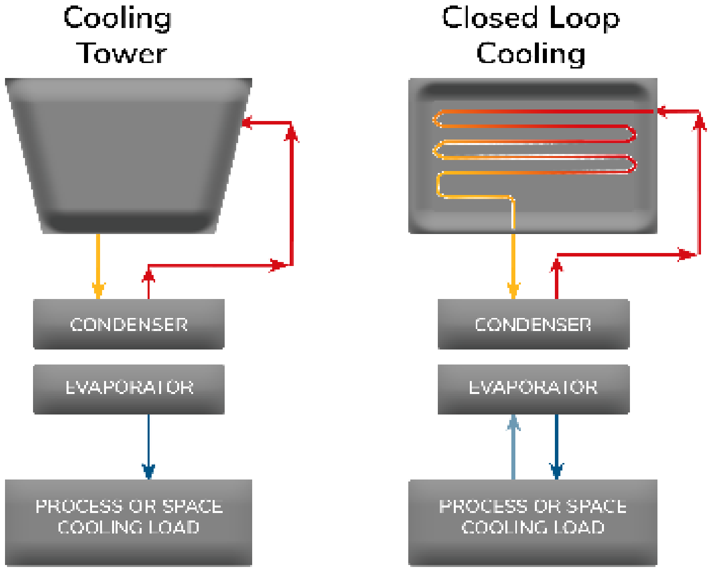
Fan-Less High-Density HDD Storage is the Future

- Replacing traditional air-flow cooling in high-density HDD storage enclosures leverages ASHRAE W27 Open/Closed loop water-cooling infrastructures

- Fan-less Cold-plate & Vapor-chamber open/closed-loop water-cooling results in lower and more uniformed HDD temperatures, thus reduces Power OPEX, improves AFR & RAS, and enables higher ar density, which enhances overall storage TCO



Vapor-chamber plates have 10x better thermal conductivity than copper alone



Cold-Plate & Thin Vapor-Chambers is the Solution

ASHRE W27 Standards Based
Leveraging Modern Datacenter Infrastructure

Assumptions

- ASHRAE W27 class water cooling
- 8 Rows & 14 Columns → 112 HDD Density
- Higher Device Performance demand

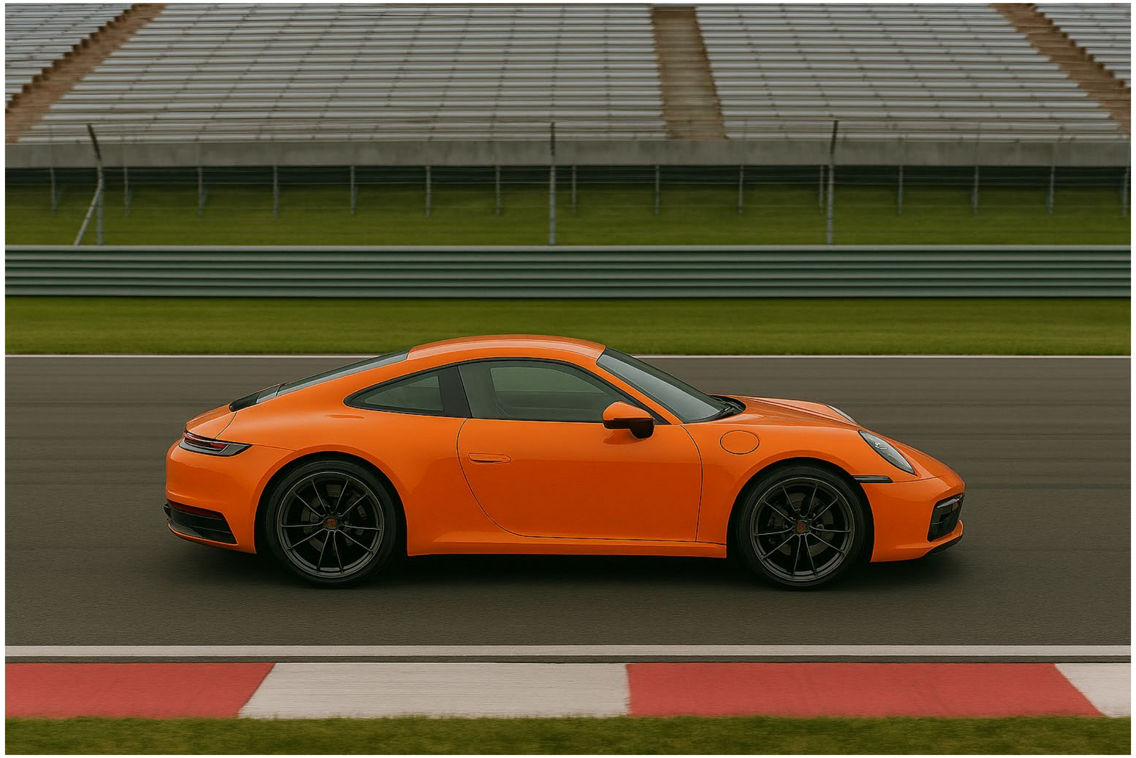
Results

- Lower Max Temp & Avg Deviation
  - → Improved AFR & RAS
  - → Reduced Power & Rebuild Traffic
  - → Denser Media & Storage enclosures
  - → **Optimized Sustainability & TCO**

Conf Options	A	B	C	D
Box CAPEX	~1.1x	~1.3x	~1.2x	~1.2x
Box OPEX	~0.4x	~0.6x	~0.5x	~0.5x
Max Temp	53°C	53°C	54°C	45°C
Temp Diff	6°C	3°C	4°C	3°C

Software Driven Storage TCO Optimization

Managing cost by dynamic provisioning of workloads to optimal storage class



Source: Copilot Aug 2025

Source: Copilot Aug 2025

NVMe HDD Standardization is Needed for Complementary Storage Technologies

Solid-state drives

- Lower latency ($1/100^{th}$ to $1/1000^{th}$)
- Higher throughput/TB (> 10×)
- Higher density/Unit (> 4×)

Best for caches close to processors

Hard drives

- Lower cost per Terabyte (1/6th to 1/9th)
- Higher power efficiency per terabyte (> 4×)
- Lower embodied carbon emission (<1/10th)

Best for mass-capacity storage and longer-term data retention

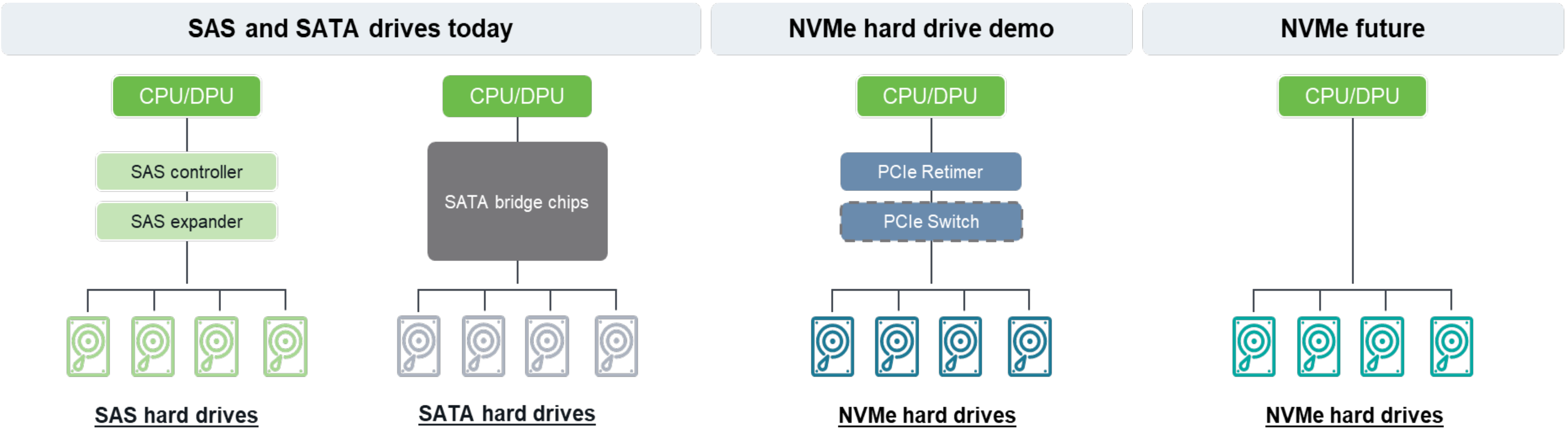
NVMe / SATA Differences

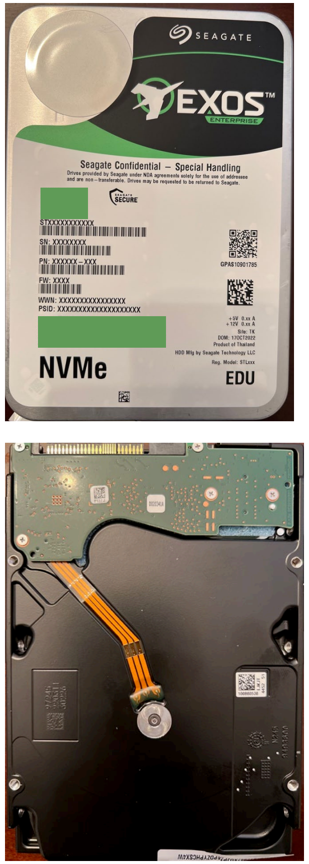
Aspect	NVMe	SATA
Interface Bandwidth	Up to ~1.6 GB/s (PCIe 3.0 x1 Full-Duplex)	Limited to ~500 MB/s Half-Duplex
Fabric Protocol Efficiency	Streamlined protocol with minimal OS overhead	Higher ATA Legacy OS Translation Overhead
CPU Efficiency	Reduced CPU overhead; optimized command handling	Higher CPU usage per operation
Scalability	Ideal for modern datacenters, AI/ML NVMeoF, and virtualization environments	Limited scalability due to legacy design
Command Queues	64K queues with 64K commands each for optimized data priority placement	Single queue with 32 commands
Interface Power Efficiency	Optimized power states (32 states)	Less power efficiency options (4 states)
Future-Proofing	Designed for current and next-gen storage silicon technologies	Legacy hard to integrate in modern silicon process
Driver Differences	Build on Native OS Kernel driver support with advanced IO scheduler	Legacy Standard SAS IOC for STP, or Internal SATA Drivers
General Purpose Ports	Native direct-connect capable over standard silicon PCIe transport layer	Requires specific SerDes and Phys
RDMA Capability	RDMA capable protocol from PCIe to Ethernet transports reduces CPU overhead	Translation and store/forward buffering for Ethernet
Queue Management	Virtual function IO queues direct mapping to CPU cores reducing lock contention	Simple queuing model not supporting virtual functions
Multi-Actuator	Dynamic Namespace Management with one namespace per actuator	Partition based support is less optimized
Fan-Out Support	Up to 5 different PCIe Switch vendors with standards-based implementations	Uniquely implemented PCIe-to-SATA Bridge Solution
Multi-Initiator	Multi-Tenant	Single-Tenant

NVMe-HDD Simplifies Mass Storage Topologies

OCP NVMe HDD Specification was Contributed by 129 Active Members

SAS/SATA to NVMe Device Replacement Reduces System-Level TCO

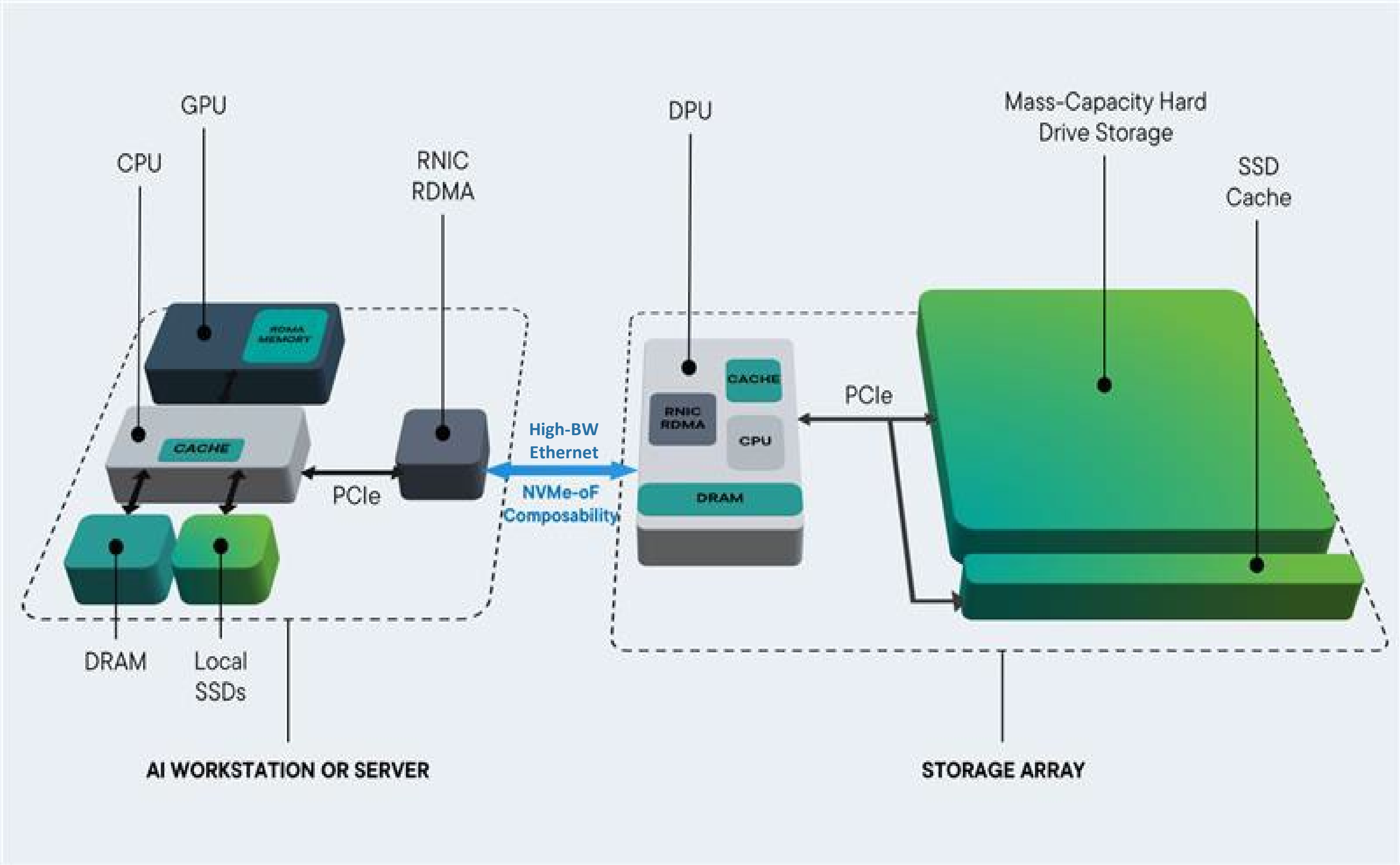




Direct GPU Memory Placement NVMe-oF Driver Support

Optimized TCO with Native PCIe Lane Connectivity

DPU offload the CPU workload at the intersection of network and storage.

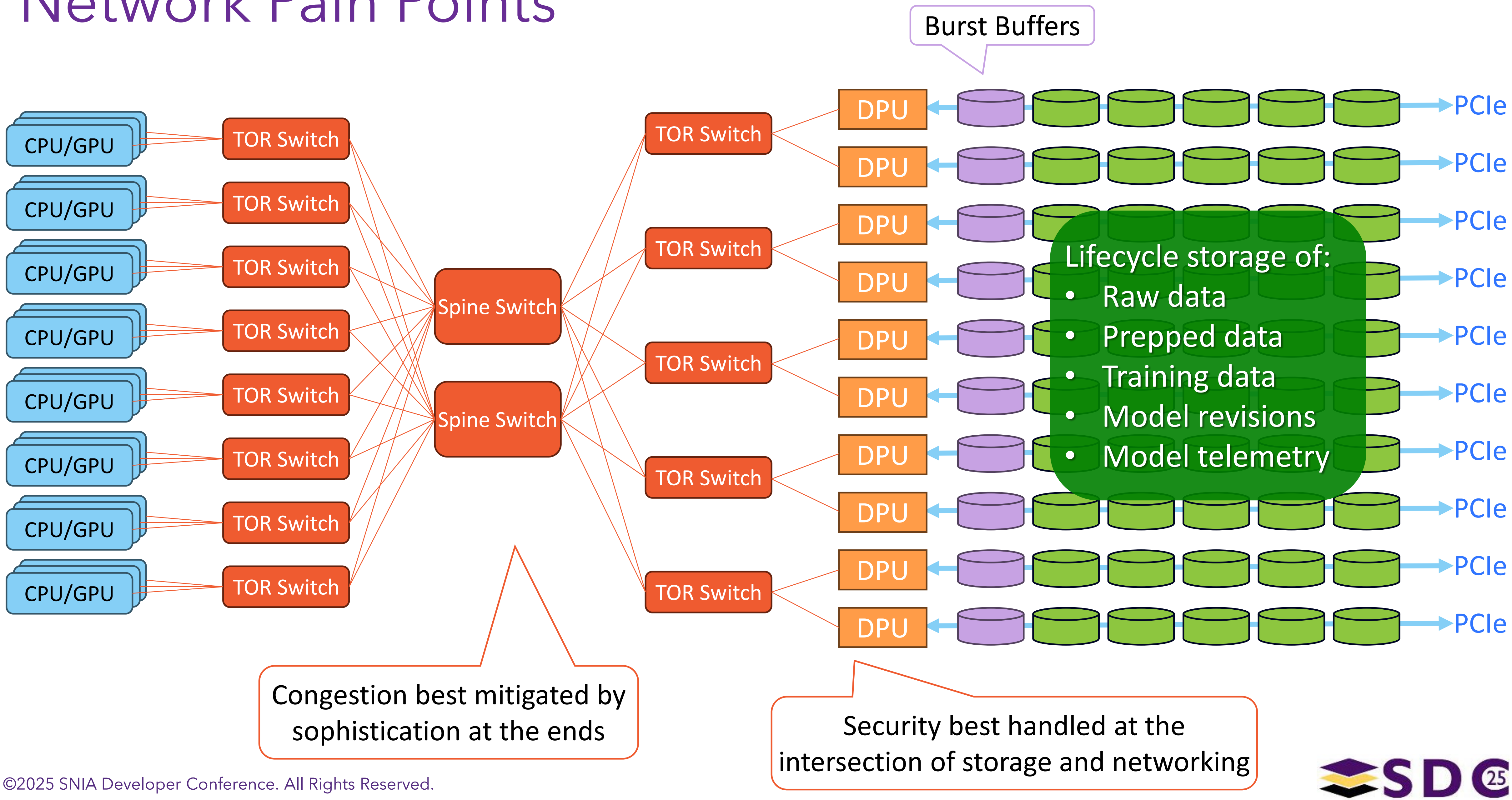


GPU meets PCIe-based hard drives: Seagate and Nvidia demo NVMe HDDs | Tom's Hardware
Seagate begins work on NVMe HDDs for cost-effective AI workloads | TechSpot

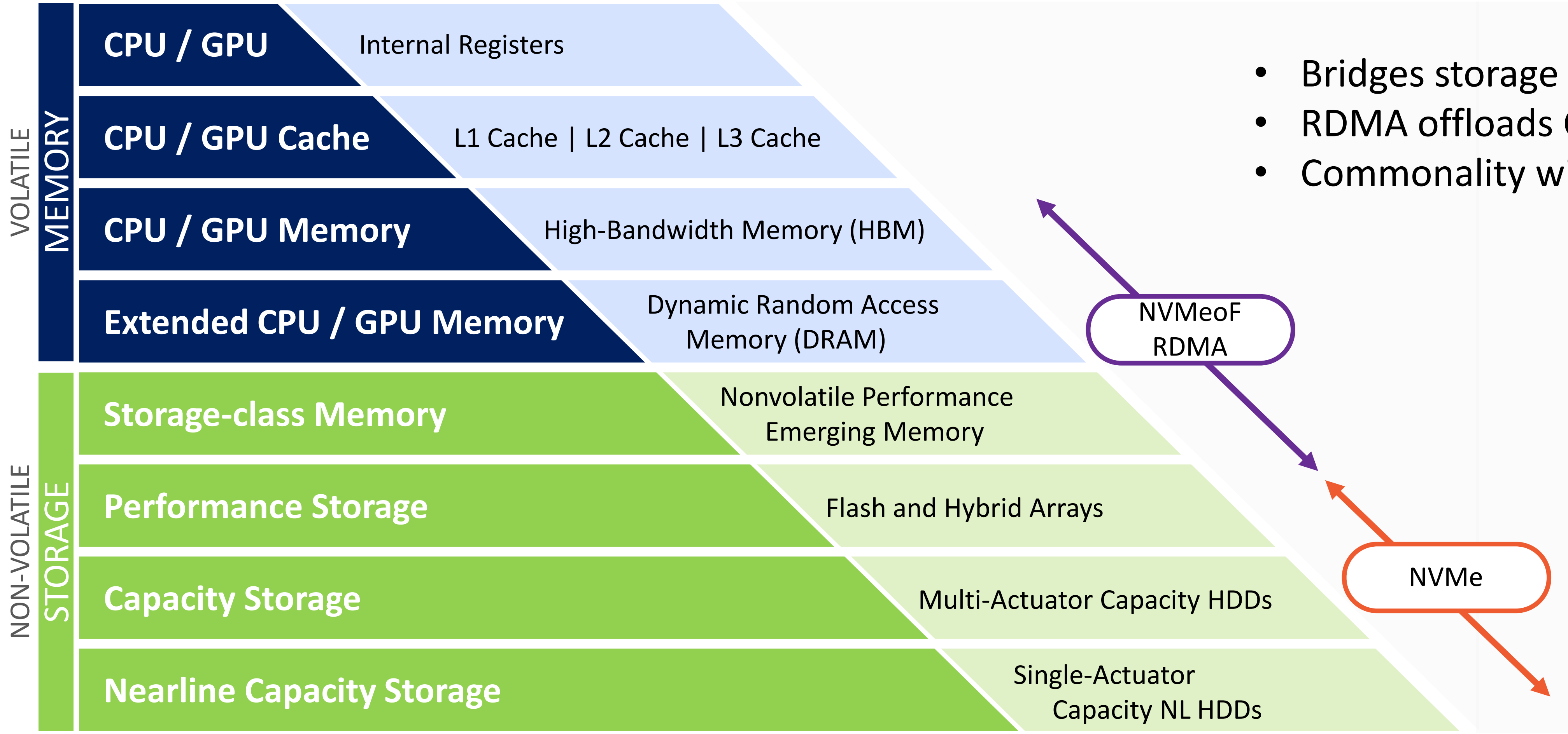
DPUs offers:

- HW automation of RDMA of NVMeoF
- Sophisticated network congestion management
- Enables other protocols besides NVMeoF block I/O
- A separate root of trust for security schemes not effected by a compromised OS.
- Localized hot/cold data migration

Network Pain Points

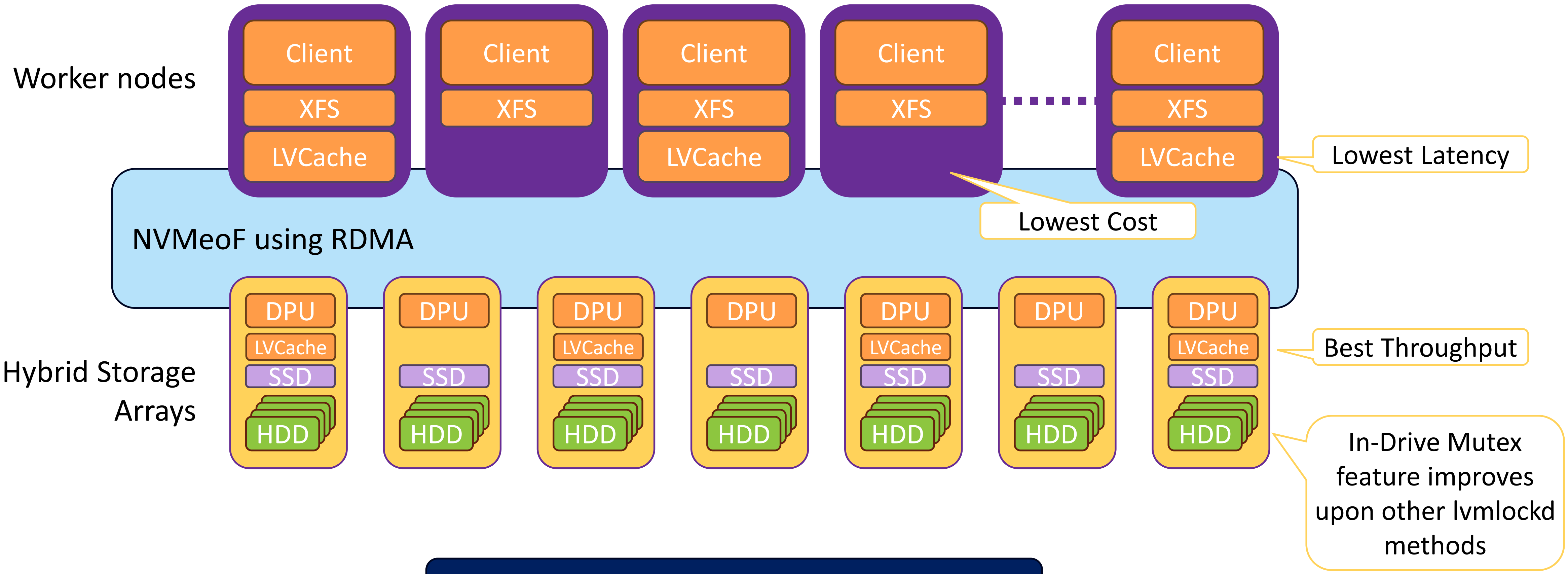


NVMoeF



- Bridges storage and memory
- RDMA offloads CPU
- Commonality with drives

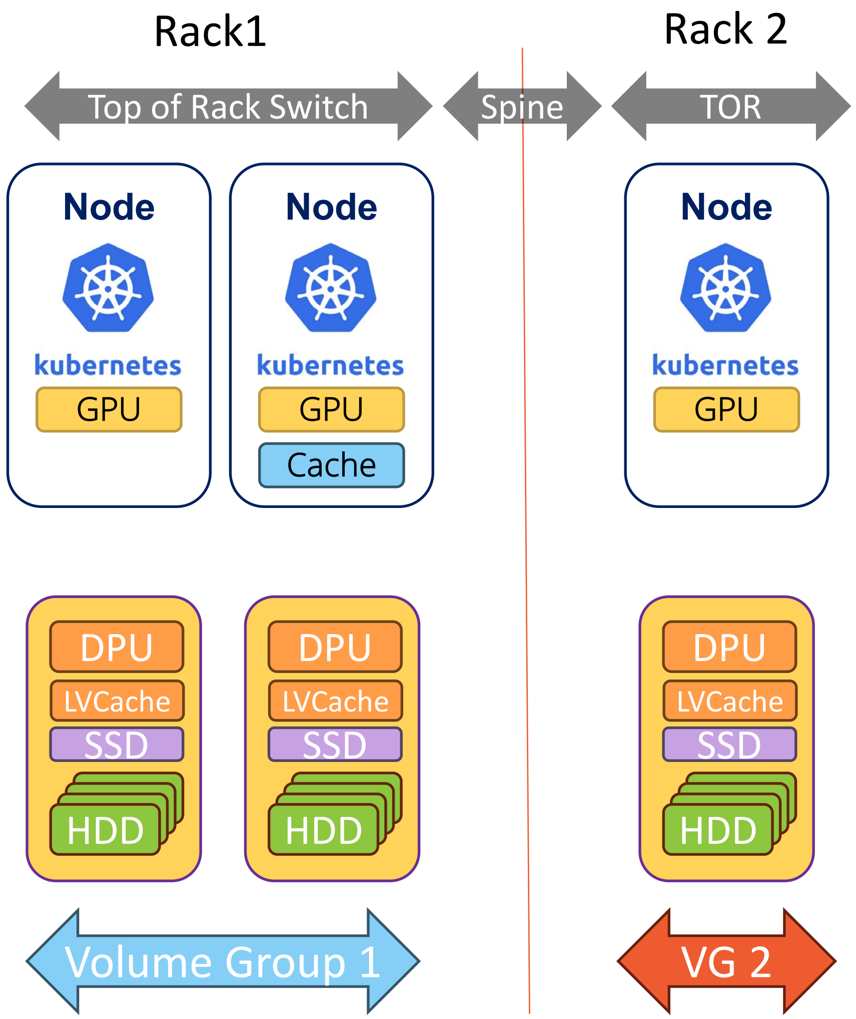
Small scale Linux-only cluster architecture with Linux Volume Manager Shared Volume Groups



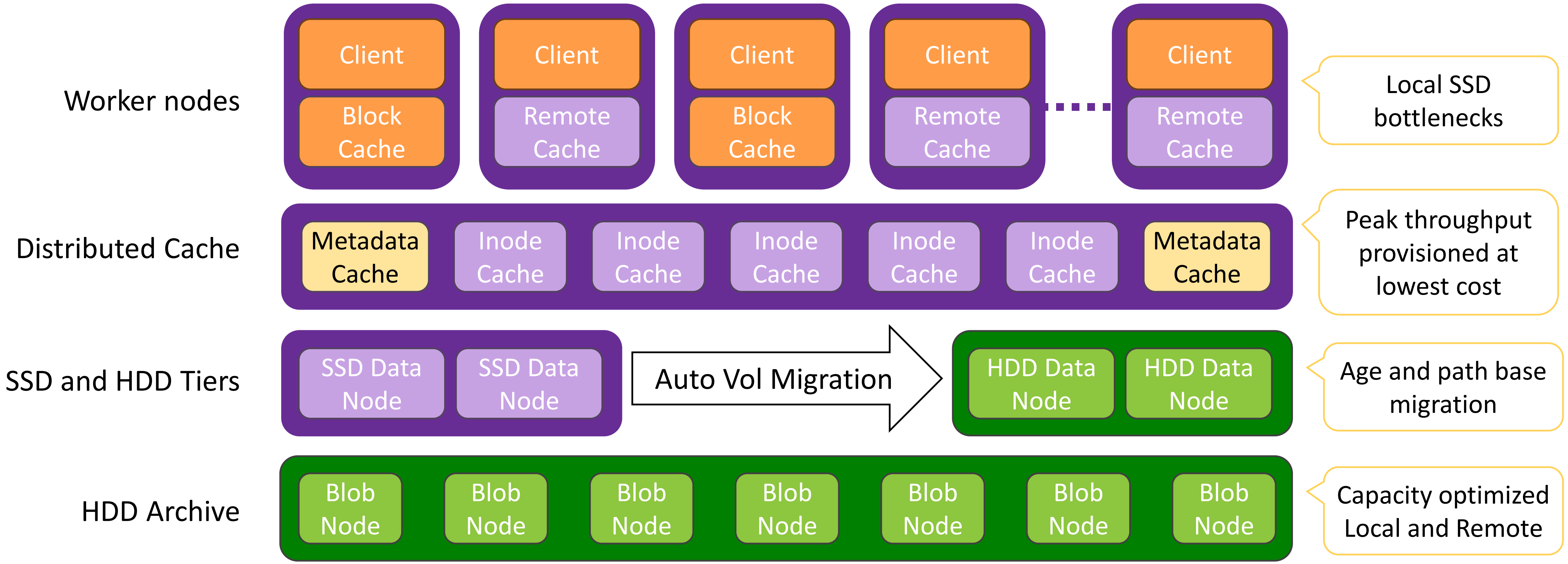
Multi-Node cluster with lockless data path

Infrastructure Dynamic Provisioning

- Kubernetes includes GPU resources in pod requirement
- TopoLVM CSI driver manages storage affinity with available resources.
- CSI Storage Classes enable declarative provisioning of deployment defined quality of service levels.



Large scale deployment offering agility required to achieve TCO benefits from all tiers.

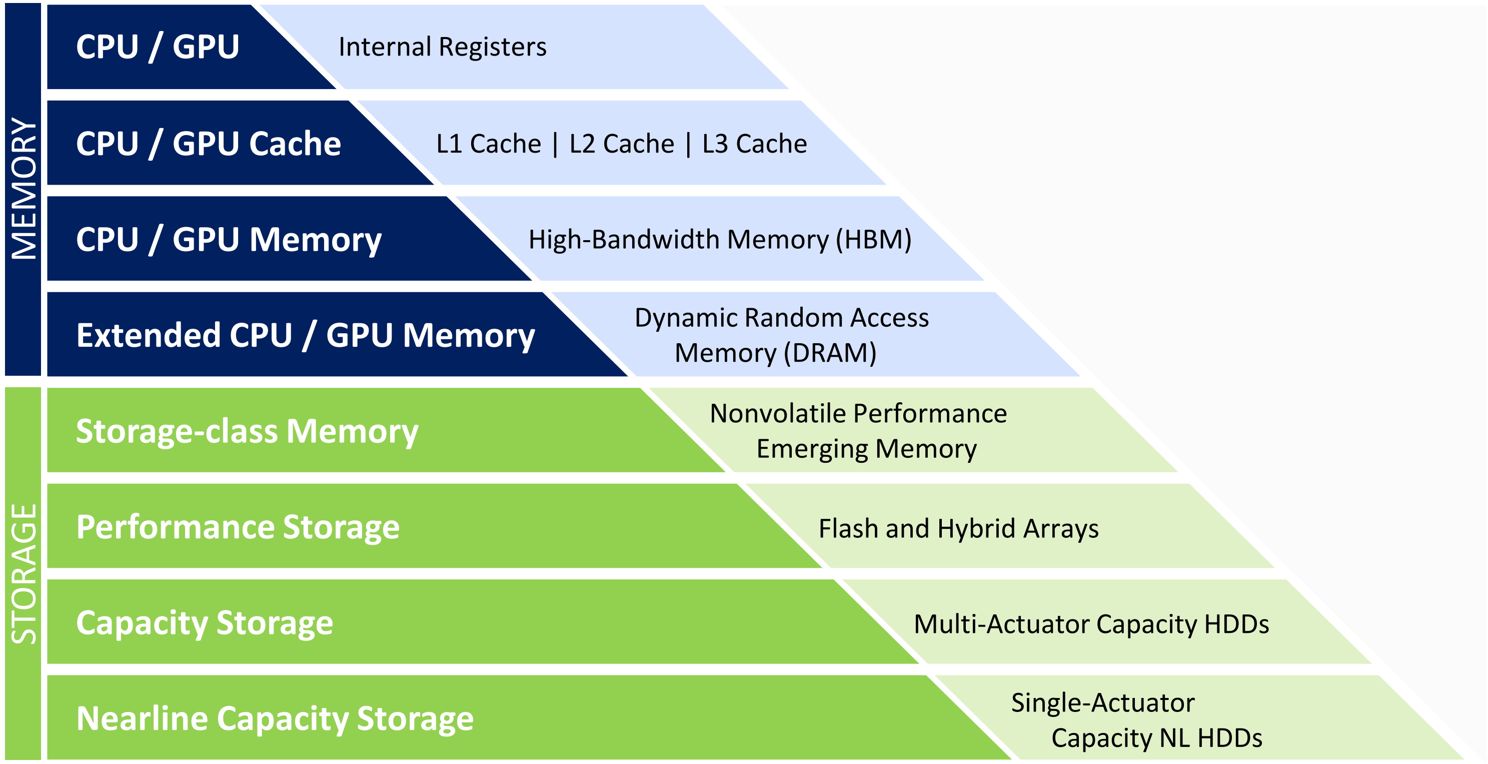


Developer's Considerations

- How will model development be tracked and measured?
- How will datasets be managed over the lifecycle of the project?
  - Will time or naming conventions be used to flag cooling of data?
  - Will active archives be remote or local or both to avoid extraction fees
- Do data prep operations need Read/Write/Many?
- Does the application deployment need to span availability zones?
- What is the duration of:
  - Raw data
  - Model revisions and historical performance
  - Scratch space
- Telemetry gathering and analysis

Storage efficiency comes through utilizing the full data stack

- Dynamically provision for mix of AI usage and workloads over the model's lifecycle.
- Automate caching, tiering and migration
  - Intelligent burst-buffer
  - Direct Data Memory Placement
  - Optimized-data-placement
  - Age out of premium space
- Streamlined sequential data benefits Hard Drives and SSDs
  - Write Amplification Reduction
  - Mechanical Actuator Seeks



Thank you for attending!

Please remember to rate this session. You get access the presentations at
http://sniadeveloper.org/conference

