Storage Architecture Optimized for AI Workloads

Presented by Paul McLeod

Product Director, Storage

Supermicro

SNIA COMPUTE, MEMORY, AND STORAGE SUMMIT

Solutions, Architectures, and Community

VIRTUAL EVENT, MAY 21-22, 2024

Storage Architecture Optimized for AI Workloads

- About Supermicro
- Storage challenges with AIOps and MLOps
- Moving beyond legacy storage
- Solution approach
- Hardware innovation with EDSFF
- Summary

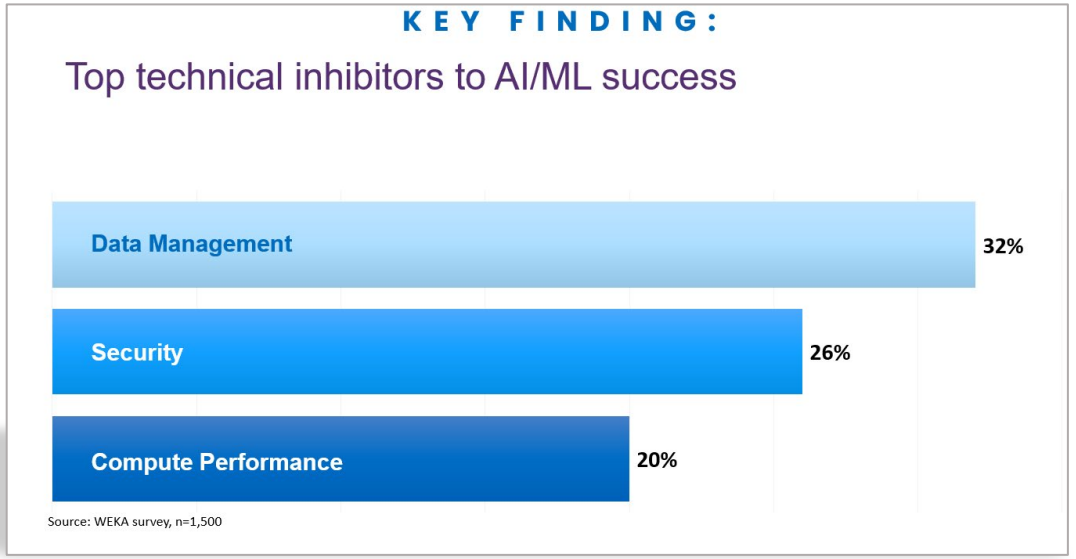
ABOUT SUPERMICRO

Revenue	**$14B+** (FY2024 guidance) $7.1B (FY2023) $5.2B (FY2022)
Worldwide Presence	6M+ Sq ft. Facilities Worldwide 1. Silicon Valley (HQ), 2. Taiwan, 3. The Netherlands, 4. Malaysia and others
Production	$18B/yr Production Capacity (CY23) Top 5 Largest Server System Provider Worldwide (IDC & Gartner 2022), ~1.3M units annually
Human Resource in 4 Campuses	6000+ headcount Worldwide, ~50% Technical / R&D
Key Growth Matrix	**#1** in Generative AI and LLM Platforms 500%+ YoY Growth in Accel. Computing

AI/ML Implementation

Challenges for AI/ML Storage projects

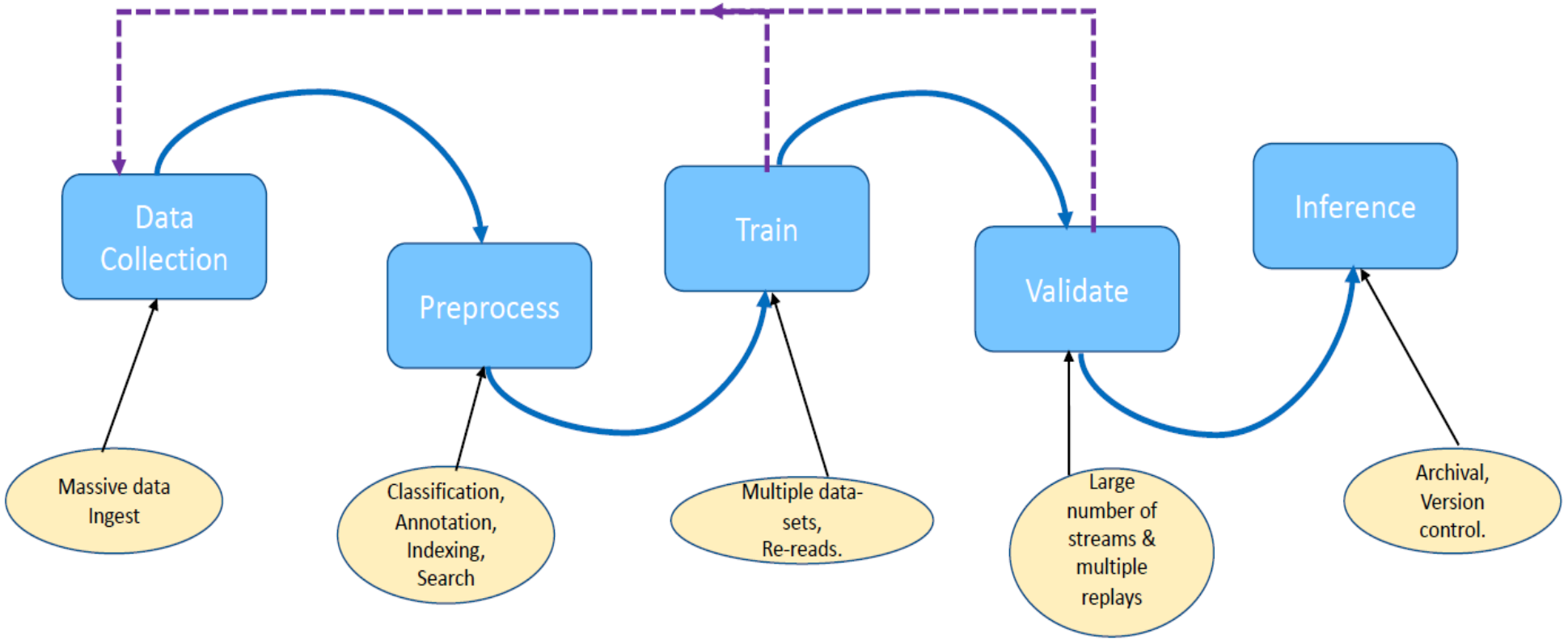
- ✓Large scale, rapid growth
- ✓Mixed data sizes
- ✓High concurrency of I/O Pipelines
- ✓Centralized management
- ✓Integration of emerging technologies



Source: WEKA

AI / ML Workflow

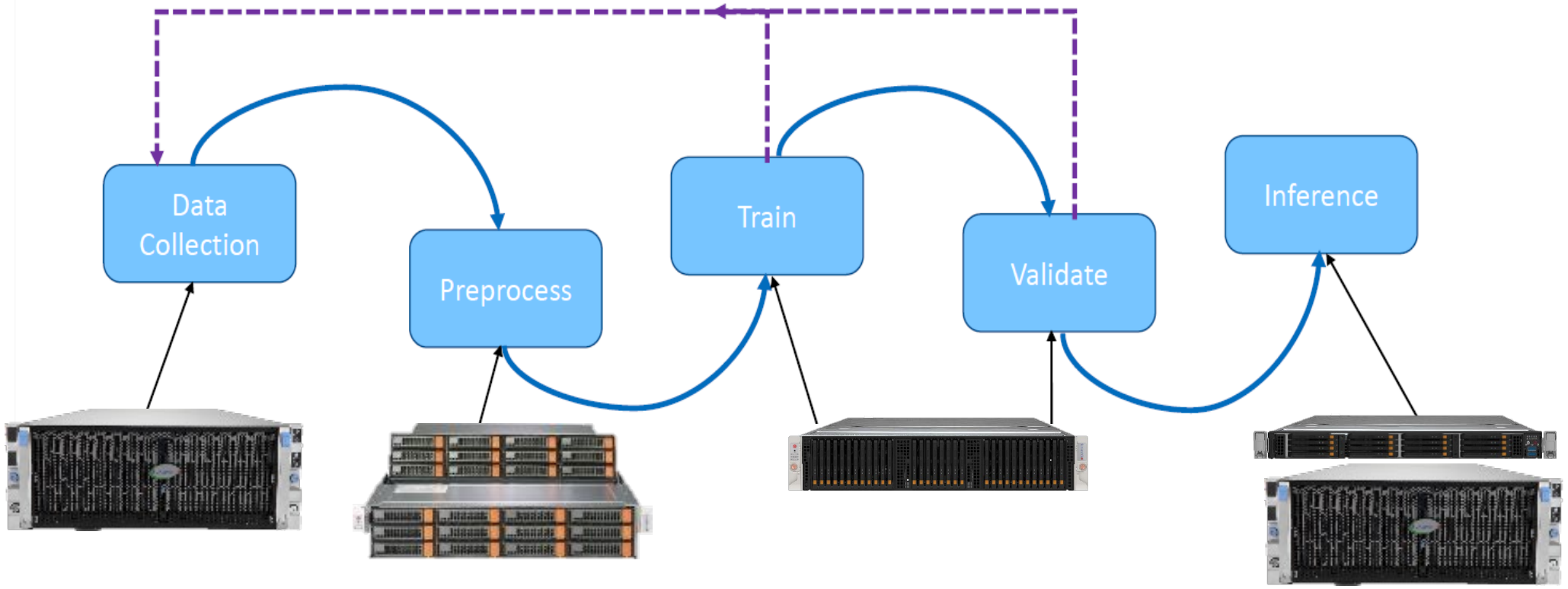
- AI/ ML application different phases



Source: SNIA

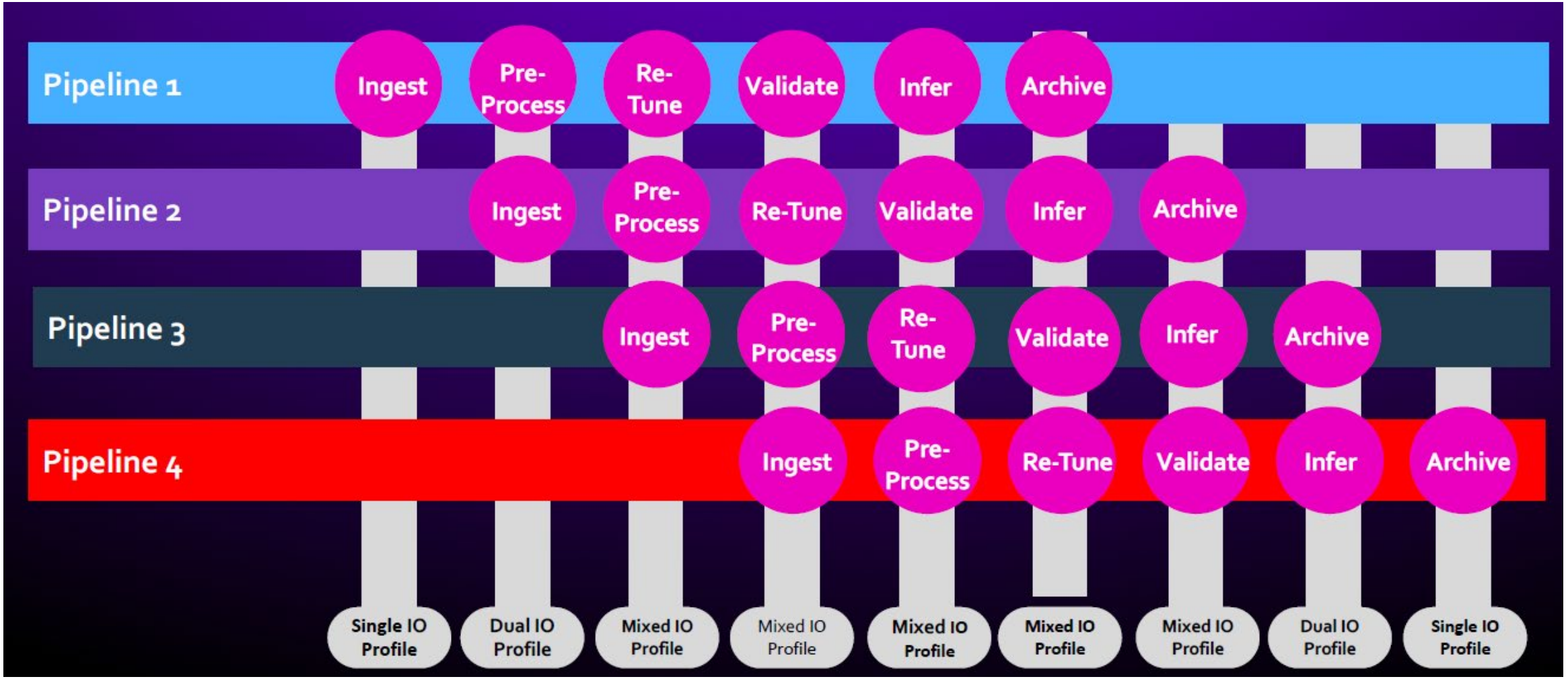
AI / ML Workflow

- AI/ ML application different phases



Source: SNIA

AI data pipeline: Multiple pipelines heating storage

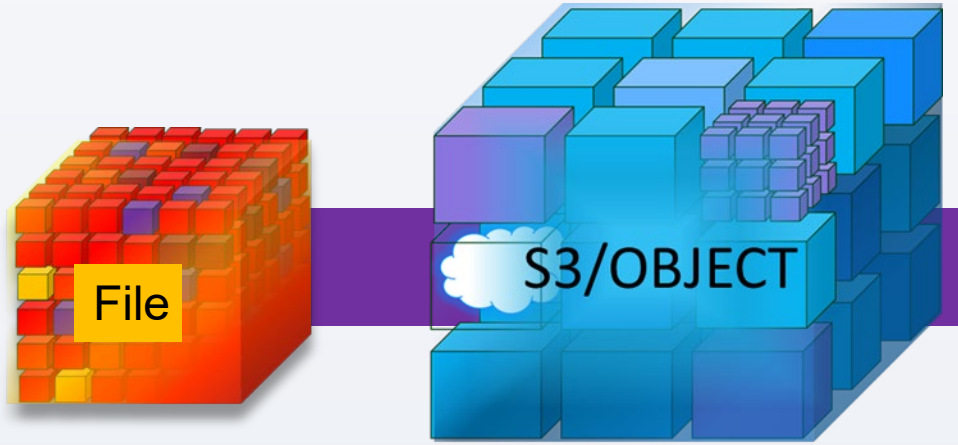


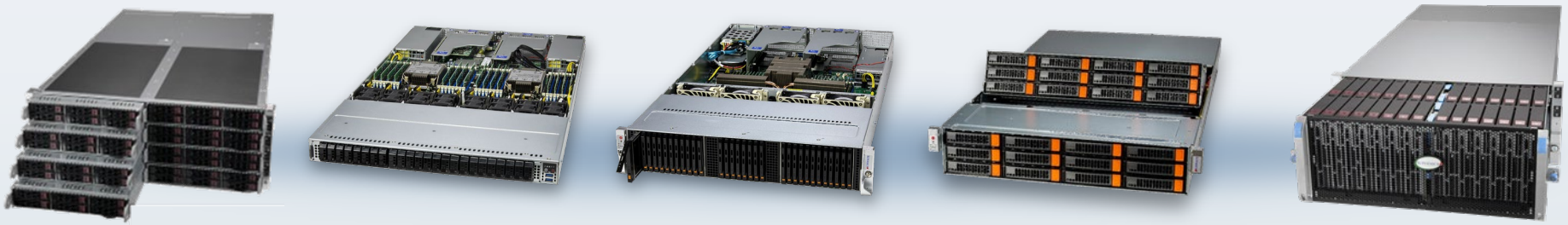
Source: WEKA

GPU Direct Storage (GDS) with WEKA

Supermicro + WEKA

- ✓ GDS provides RDMA with the GPU Memory
- ✓ Lowest latency for the AI Pipeline
- ✓ File-based single namespace for Flash and HDD
- ✓ Transparent file level access to S3 objects
- ✓ Scale-up from 138GB/s with an entry cluster

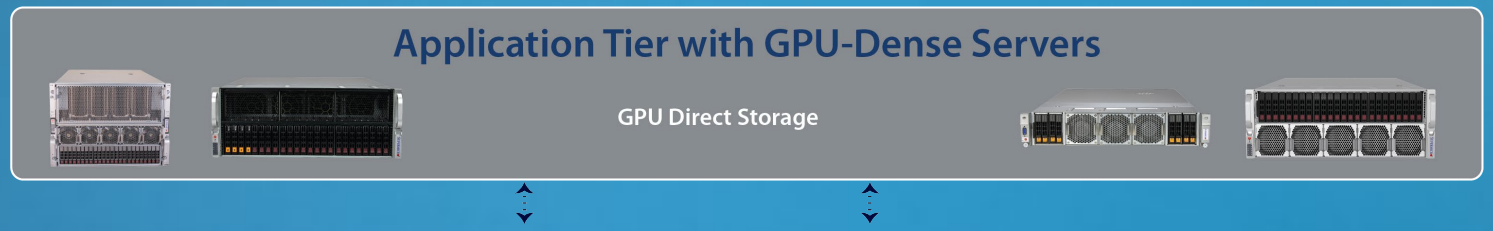




Supermicro Offers Tiered Storage Building Blocks for WEKA

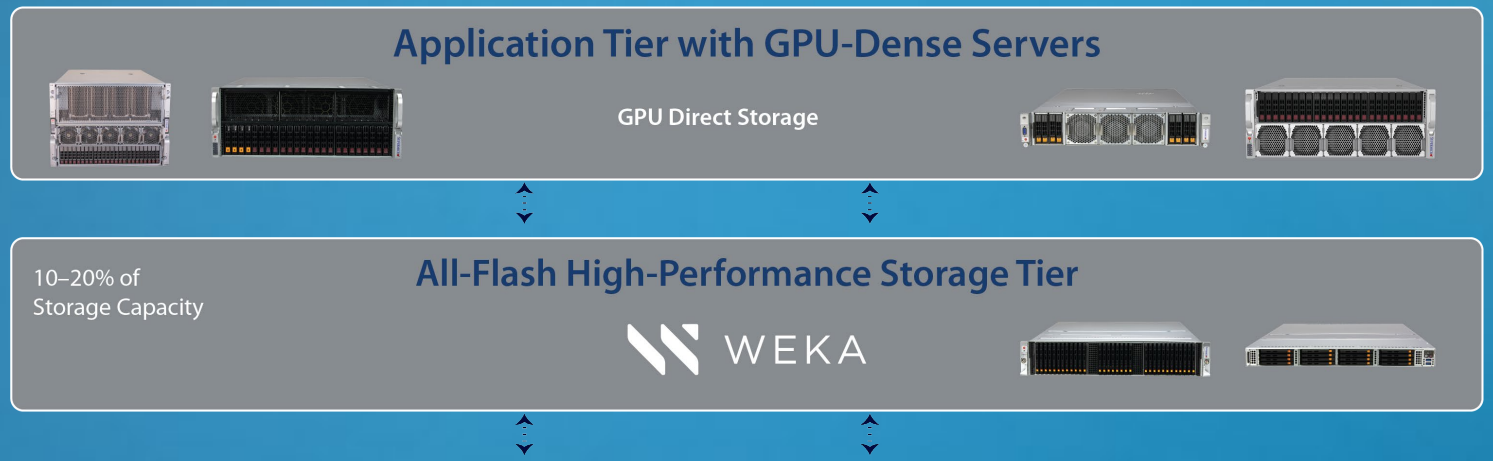
* The performance number is based on six node PCIe4 WEKA storage cluster

Supermicro Powered Application Tier



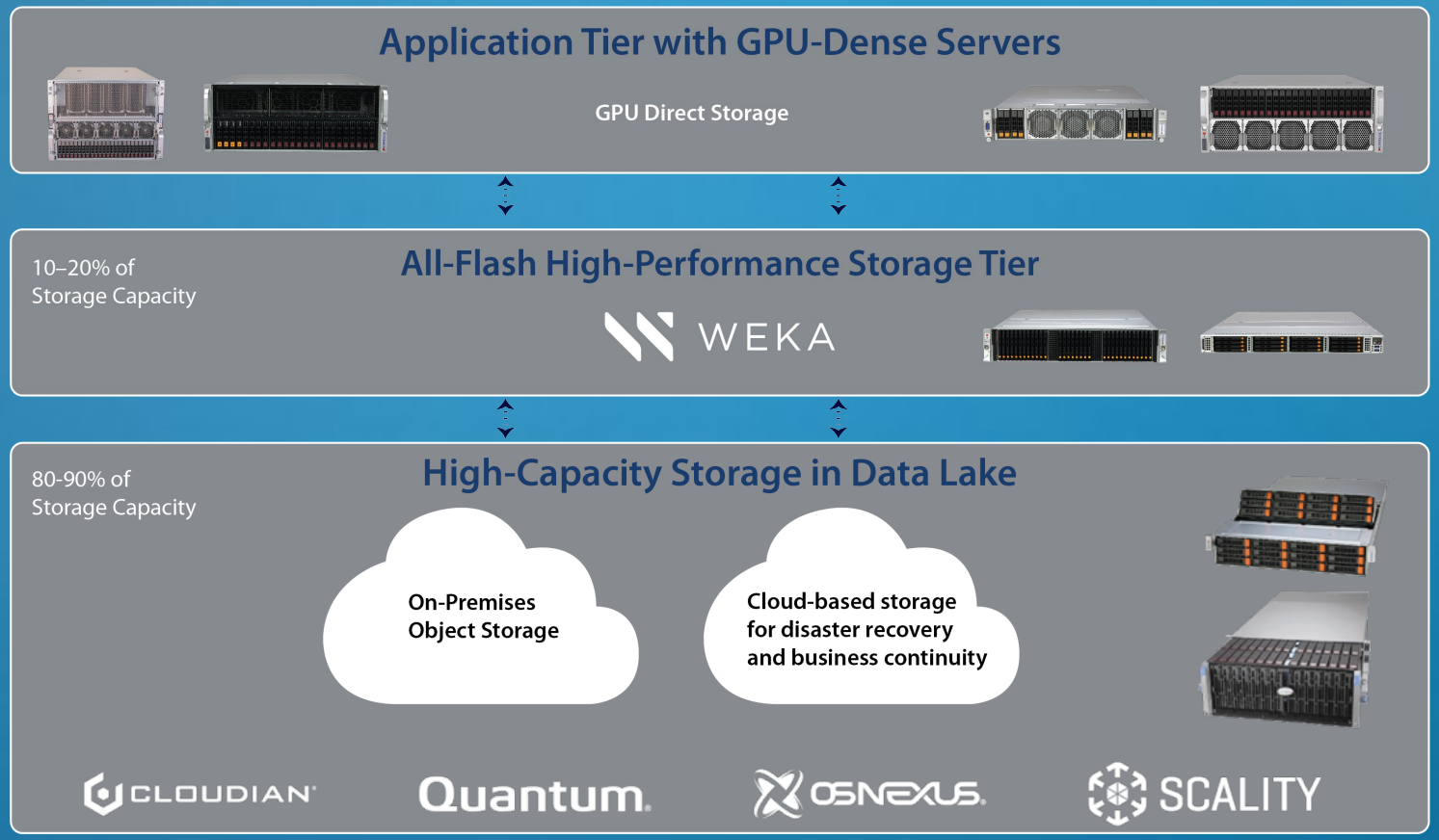
No provider offers more choices for GPU-accelerated computing

High-Performance All-Flash Tier



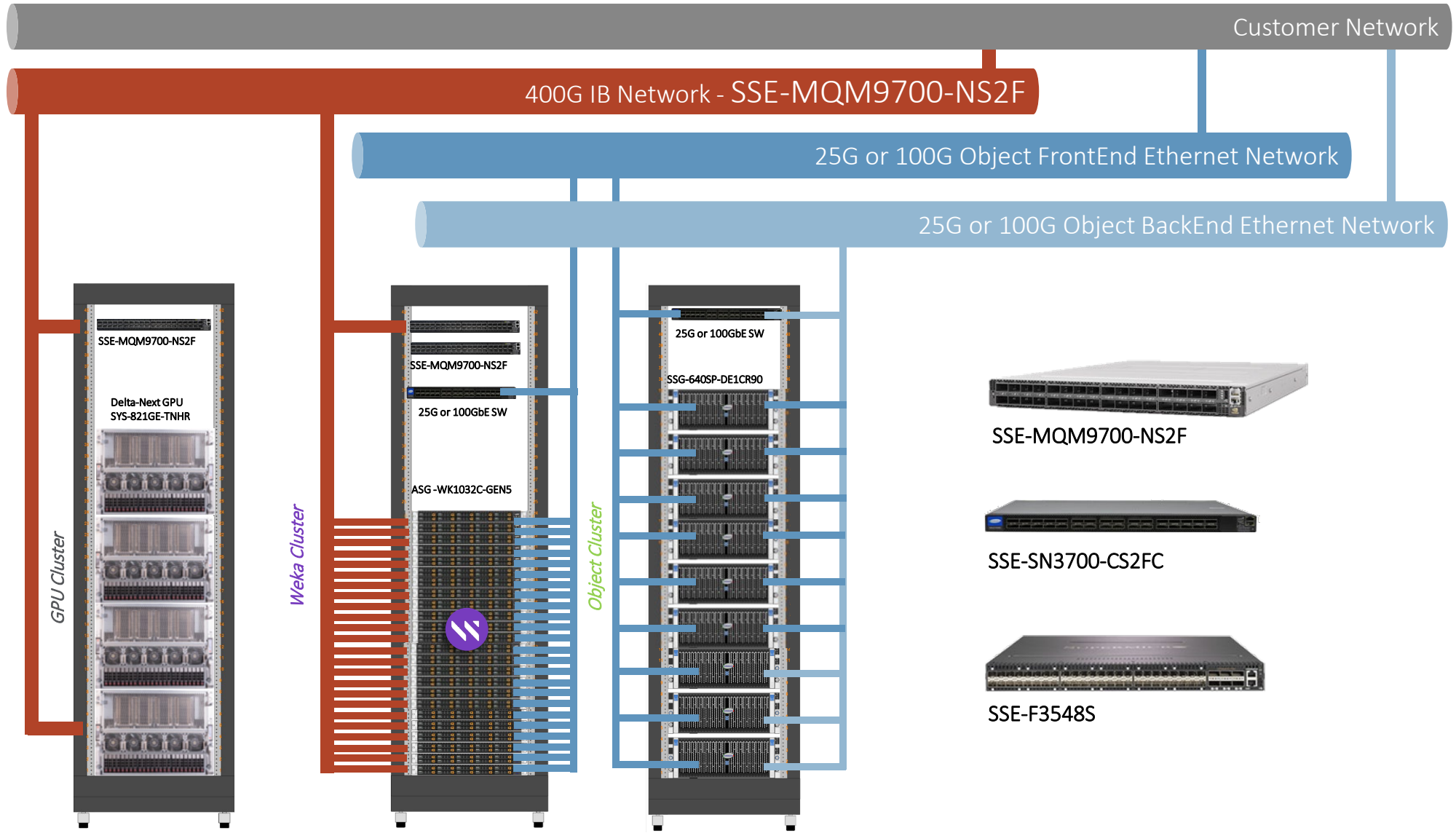
We worked with Weka engineers to optimize for Supermicro storage servers

Multi-Tier Storage Architecture for AI and ML Workloads



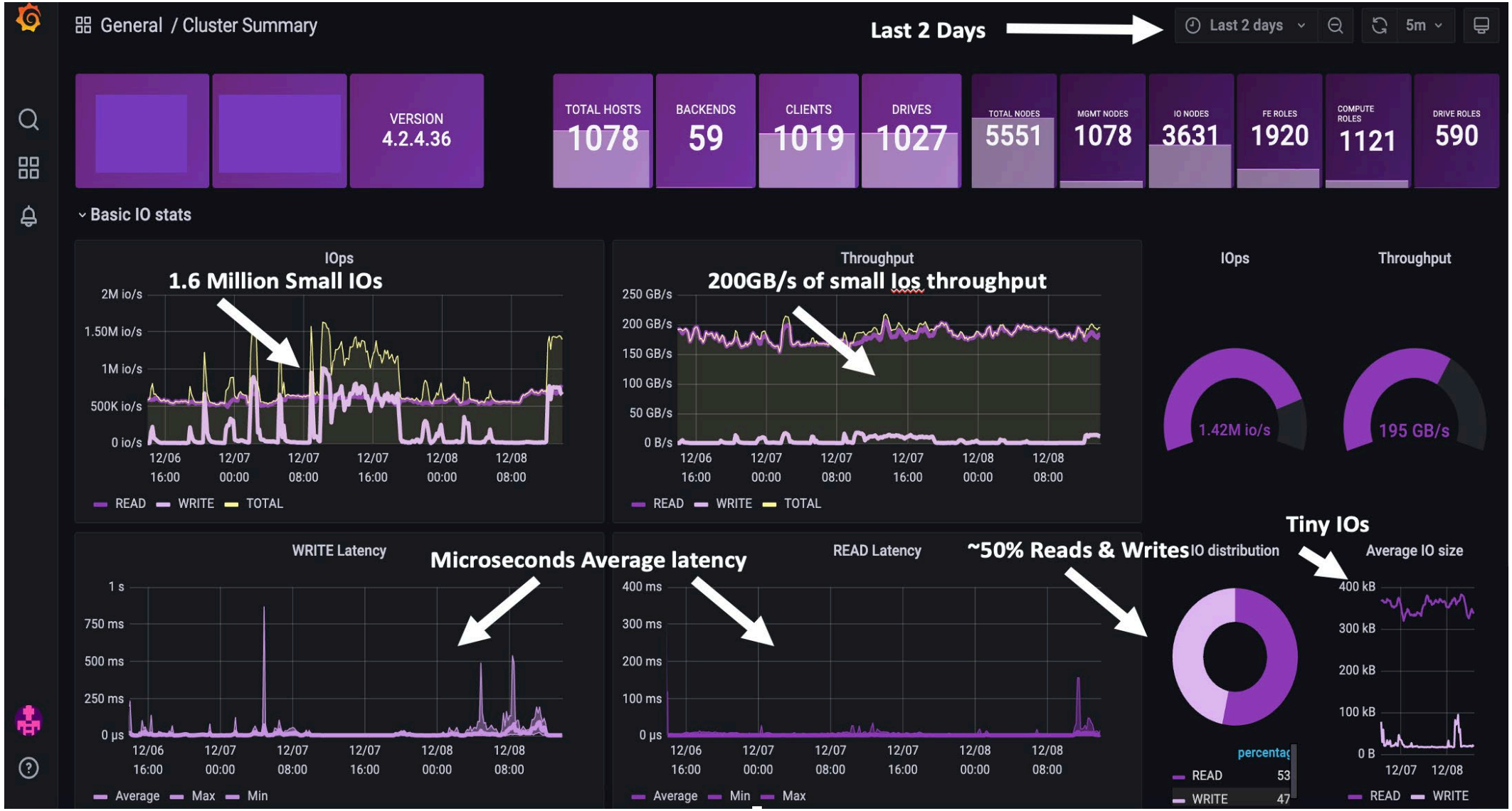
The key to cost-effectively storing all your data, safely, on premises

AI Storage Reference Architecture



For reference only, layout varies depending on actual system quantity

AI customer #1 IO Pattern – Millions of Tiny IOs Reads / Writes



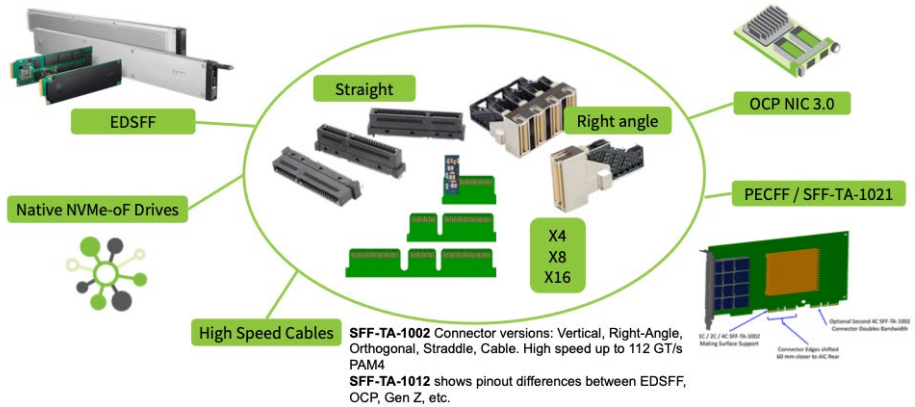
Source: WEKA

Hardware Innovation

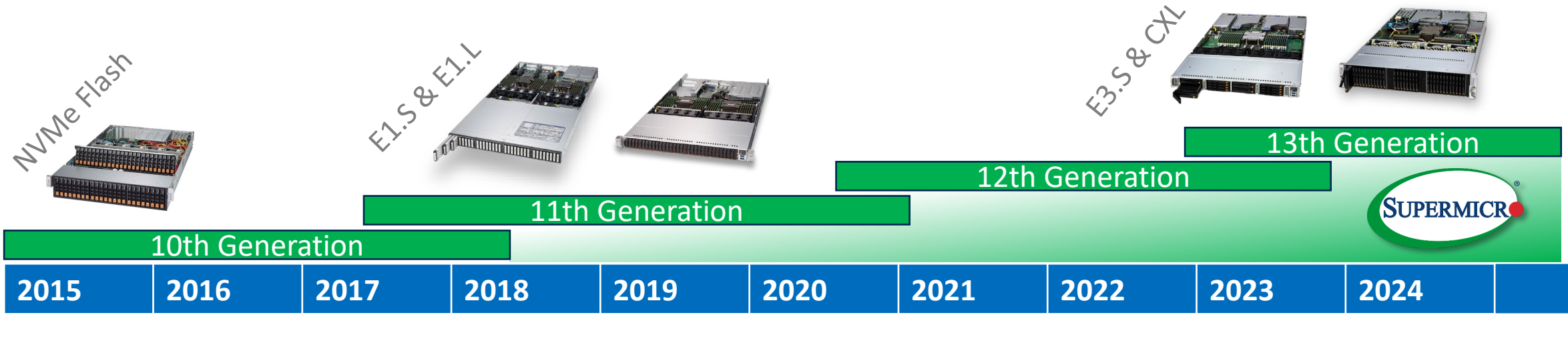
Embracing Emerging Standards

EDSFF and CXL

EDSFF E1.S, E3.S, and E3.L form factors, as well as AICs, have been integrated into the Compute Express Link® (CXL®) ecosystem, underscoring their utility in high-performance, high-capacity server environments promoting robust, scalable, and efficient designs.



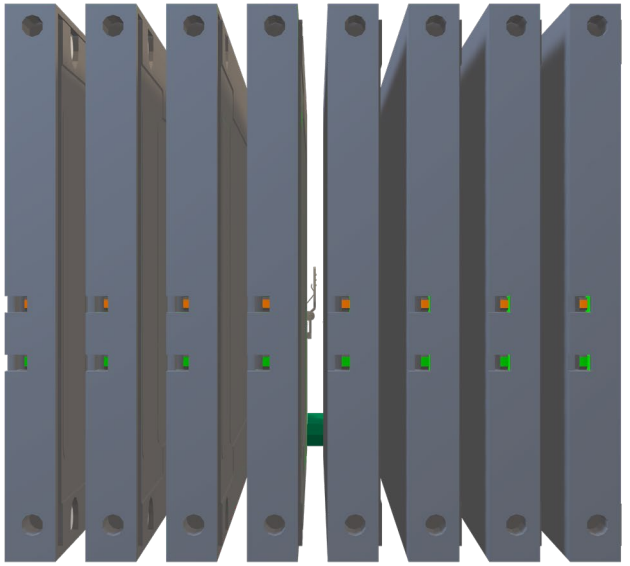
Source: SNIA



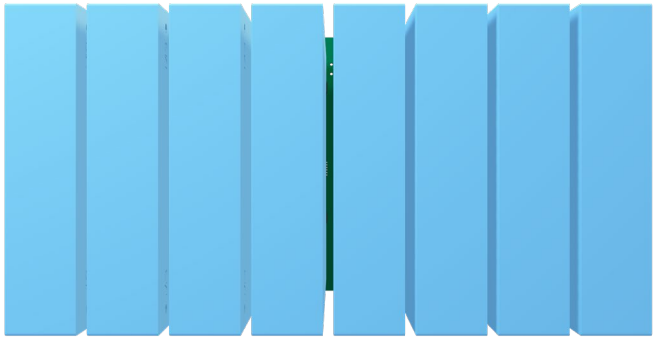
Gen5 EDSFF Petascale Platform Innovation

Reduce Backplane Trace Layout Signal Loss & Improve Air Flow

- **Superior Signal Integrity**
  - Mainboard direct connection to SFF-TA-1002 1C connectors/SSDs and reduce the backplane routing signal loss
  - Reduce ~40% of the signal loss
- **Better Air Flow**
  - No vertical backplane blocking the air flow
  - ~ 75% increase in front opening
  - ~ 20% improvement system CFM



E3.S SSD + EDSFF BPN



U.2 SSD + BPN

Gen5 EDSFF Petascale Platform Innovation

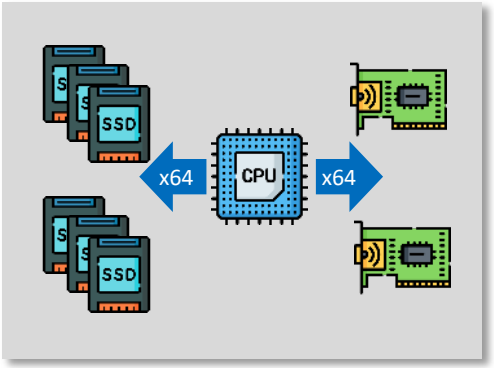
Purposed Built for New All-NVMe and Software-Defined Data Center

- **Unified Chassis**
  - Support
    - Intel DP and AMD UP
    - 1U up to 24 E1 SSD
    - 1U up to 16 E3 SSD and CXL
    - 2U up to 32 E3 SSD
  - Less than 31” chassis depth

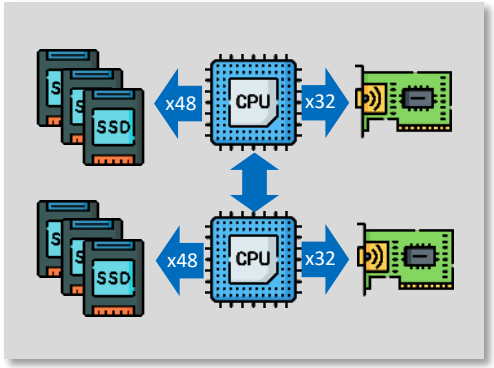
- **Balanced Architecture**
  - Front storage IO and rear networking
  - Eliminate the processor NUMA complexity

< 31” chassis depth

1U16, 1U24, 2U32 and CXL



AMD Single Processor

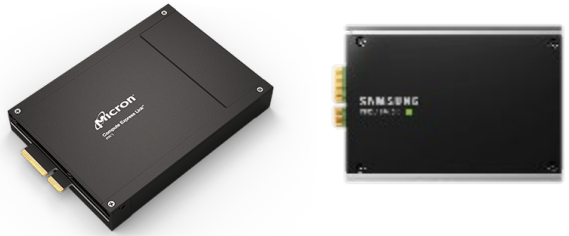


Intel Dual Processor

Petascale + CXL Memory Expansion Solution

Next Gen Memory Expansion Solution for Next Gen Data Center

- CXL (Compute Express Link)
  - A high-speed interconnect, industry-standard interface for communications between processors, accelerators, memory, storage, and other IO devices.
- CXL Memory Expansion
  - Enabling memory cache coherency between CPU memory and attached memory devices
- 1st system supports 4x E3 CXL 2T device.
  - Petascale 1U system (both AMD and Intel)



- 1st industry E3 CXL PoC
  - Partner with Micron CXL team and demonstrate great performance improvement when use Micron CXL CMM and SMC Petascale system

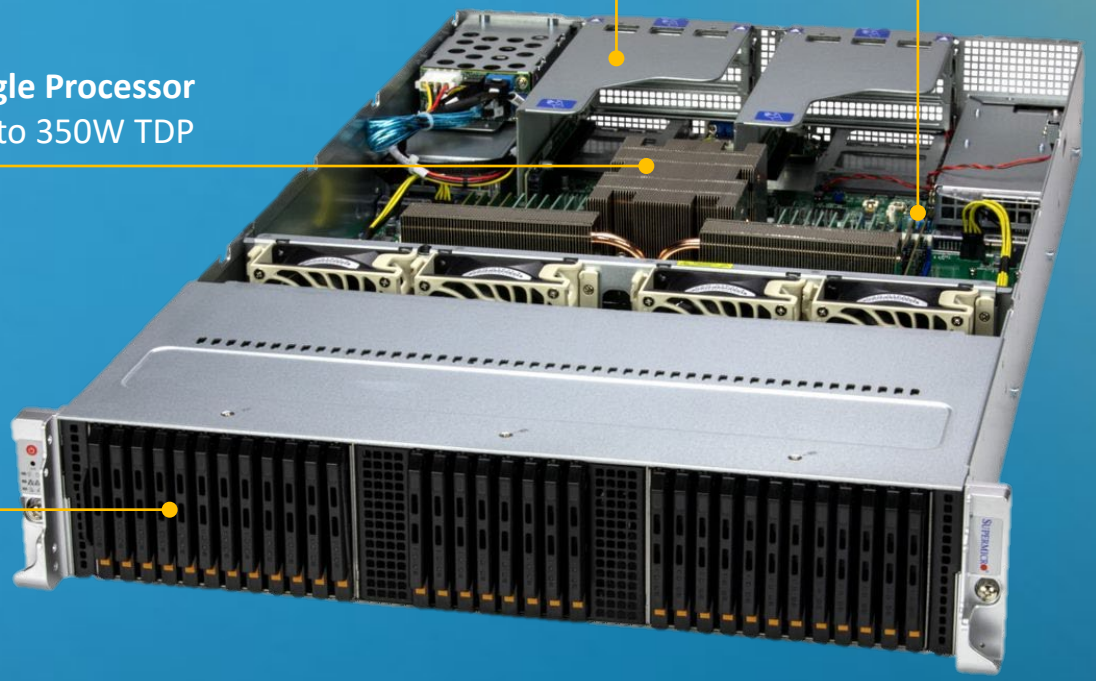
SK hynix Micron SAMSUNG

H13 2U E3.S Petascale All-Flash

PCIe 5.0 Slots
2 x16 slots & 2 AIOMs

DDR5 Slots
Up to 24 DIMMs

Single Processor
Up to 350W TDP



E3.S 1T SSD
Up to 32 E3 1T slots

Redundant Power Supply
2000W (Titanium level)

Hot swap U.2

PCIe 5.0
x16 AIOMs

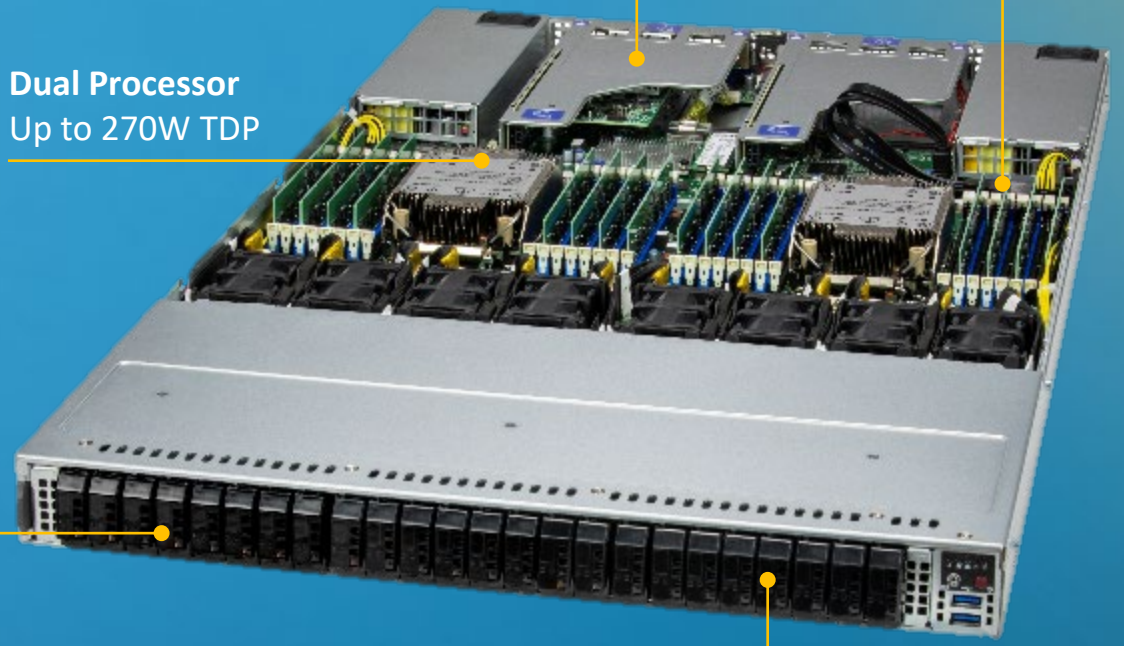
PCIe 5.0
x8 or x16 FHHL slots

X13 1U E1.S Petascale All-Flash

PCIe 5.0 Slots
2 x16 slots & 2x AIOMs

DDR5 Slots
Up to 32 DIMMs

Dual Processor
Up to 270W TDP



E1.S SSD
Up to 24 E1 slots

Support 9.5mm or
15mm E1.S

Redundant Power Supply
2000W (Titanium level)

PCIe 5.0
x16 AIOMs

PCIe 5.0
x16 FHHL slots

H13 1U Petascale CXL & E3 SSD Server

PCIe 5.0 Slots
Up to 2 x16 slots + 2 AIOMs

DDR5 DRAM
Up to 24 DIMMs

Single AMD Genoa Processor

E3.S 1T (x4) SSD

E3.S 2T (x8)
CXL Type 3 Module

Redundant Power Supply
1600W (Titanium level)

PCIe 5.0 Slots
x16 AIOMs

PCIe 5.0 Slots
x16 slots

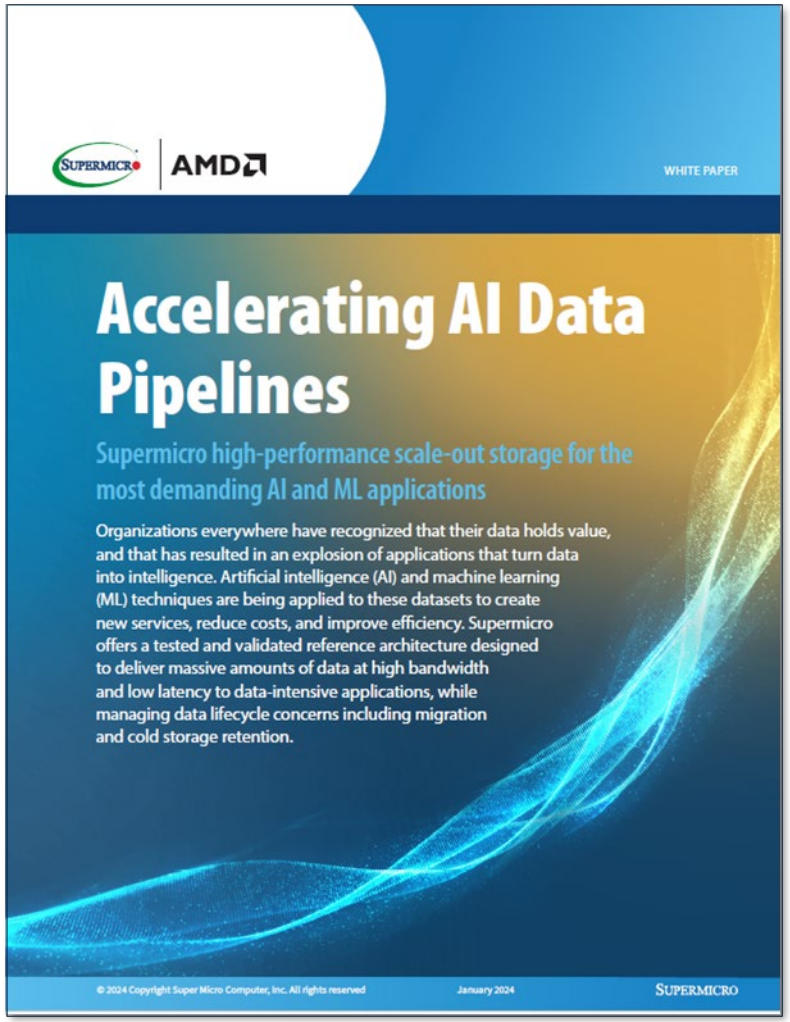
Summary

- Conventional storage approaches aren't well suited to AI and ML workloads
- The "I/O Blender" effect in the data pipeline mixes read/write on small files and multiple simultaneous pipelines produce mix of I/O patterns
- A two-tier storage architecture with a Parallel File System on Supermicro's Petascale All-Flash storage server enables high performance E3.S flash from multiple partners
- An object tier using Supermicro's high-density disk-based SuperStorage storage server provides 90 drives and over 2PB* raw capacity in 4U
- This solution has been deployed with a multinational high tech manufacturing customer with 25PB

.

*Raw value is based on vendor raw base capacity of 24TB. TB is base-10 decimal.

AI Storage White Paper



This paper is available at
www.supermicro.com/en/solutions/ai-storage

For More Information

Supermicro:

- www.supermicro.com/en/solutions/ai-storage
- www.supermicro.com/en/products/storage

Contact Info:

www.supermicro.com/en/contact

Thank You!


SNIA
COMPUTE, MEMORY,
AND STORAGE SUMMIT
Solutions, Architectures, and Community
VIRTUAL EVENT, MAY 21-22, 2024

Please take a moment to rate this session.

Your feedback is important to us.

DISCLAIMER

Super Micro Computer, Inc. may make changes to specifications and product descriptions at any time, without notice. The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors. Any performance tests and ratings are measured using systems that reflect the approximate performance of Super Micro Computer, Inc. products as measured by those tests. Any differences in software or hardware configuration may affect actual performance, and Super Micro Computer, Inc. does not control the design or implementation of third party benchmarks or websites referenced in this document. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to any changes in product and/or roadmap, component and hardware revision changes, new model and/or product releases, software changes, firmware changes, or the like. Super Micro Computer, Inc. assumes no obligation to update or otherwise correct or revise this information.

SUPER MICRO COMPUTER, INC. MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

SUPER MICRO COMPUTER, INC. SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL SUPER MICRO COMPUTER, INC. BE LIABLE TO ANY PERSON FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF SUPER MICRO COMPUTER, Inc. IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ATTRIBUTION