AI Storage Reference Design

Harness the potential of GPU-accelerated AI

WHITE PAPER

June 2025

Contents

Revisions

Date	Description
Date	Description
May 2025	Initial release
May 2025 0.1	AMD additions
May 28, 2025	Updated with new VDURA content and adjust networking
June 6, 2025	Formatted Index
June 12, 2025	Updated and reformatted

06/2025

Introduction

The rapid evolution of Artificial Intelligence (AI) has posed unique challenges to systems infrastructure deployed across a broad spectrum of industrial and academic use cases. In contrast to traditional applications, GenAI applications are designed primarily with reasoning and generalization capabilities. This is achieved via training massive models ranging from billions to trillions of parameters across trillions of tokens of data (text, images, videos, etc.). Because very large models and datasets cannot fit in a single GPU's HBM or even a node, they are distributed with advanced parallelism techniques across nodes and racks. As the models, datasets, and context length grow in complexity and size, scaling infrastructure for training is a critical challenge and poses new opportunities.

Infrastructure requirements are highly dynamic for different phases of the AI pipeline. For instance, prior to AI training, the data is ingested, prepared, and transformed into GPU-amenable structures through Extract-Transform-Load (ETL) pipelines, which requires moving data across the storage network through the front-end network to compute nodes. During AI training, models and data are loaded from remote networked storage to the GPUs (mostly via the front-end network), and the KV caches and model checkpoints are written back across the network to the storage nodes for persistence. This results in large amounts (GBs – TBs) of traffic between GPUs across back-end networks for collectives and front-end traffic especially for persisting and loading checkpoints, requiring high throughput. This may saturate the front-end NICs while also triggering bursts of background traffic across storage systems to enable IO fulfillment services. While emerging, AI inferencing workloads such as AI Vector DBs, Retrieval-Augmented Generation (RAG), and agentic workloads have latency sensitivity. **Infrastructure deployed for AI workloads (training or inference), especially for Large Language Models (LLMs), requires high-performance compute, scalable storage, and robust networking coupled with efficient resource orchestration and data pipeline optimizations.** Reliable blueprints from tested and proven reference architecture are invaluable, offering clear guidance on effectively designing, deploying, and scaling AI infrastructure to enable organizations to harness the potential of GPU-accelerated AI.

The VDURA AI Storage Reference Design on AMD Instinct GPUs Architecture, referred to in this document as the "AI Storage Reference Design", was developed to address and mitigate the challenges that arise when integrating components to build AI systems of any scale, anywhere. This document includes a prescriptive full-rack reference design based on an integrated, tested, and supported configuration optimized for AI workloads. This document contains information about the following design aspects: AI reference architecture, reference design components, compute system design, storage system design, management, and configuration options, and network connectivity.

AI Reference Architecture

A scalable unit (SU) consists of a certain number of compute nodes with NICs in back-end network, local storage, and networked storage interconnected by an Ethernet fabric. In this reference design, the compute nodes consist of AMD Instinct™ MI300 series GPUs, 400 GbE NICs for back-end network, VDURA V5000 network storage, and Ethernet switches. Compute nodes can access each other for full bandwidth, "all-all" communication patterns, as well as the VDURA V5000 storage solution via the VDURA DirectFlow® client software. VDURA PanFS®, the storage operating system and parallel file system that runs in the V5000 storage solution, and the other major components of the AI Reference Architecture are detailed later in the document.

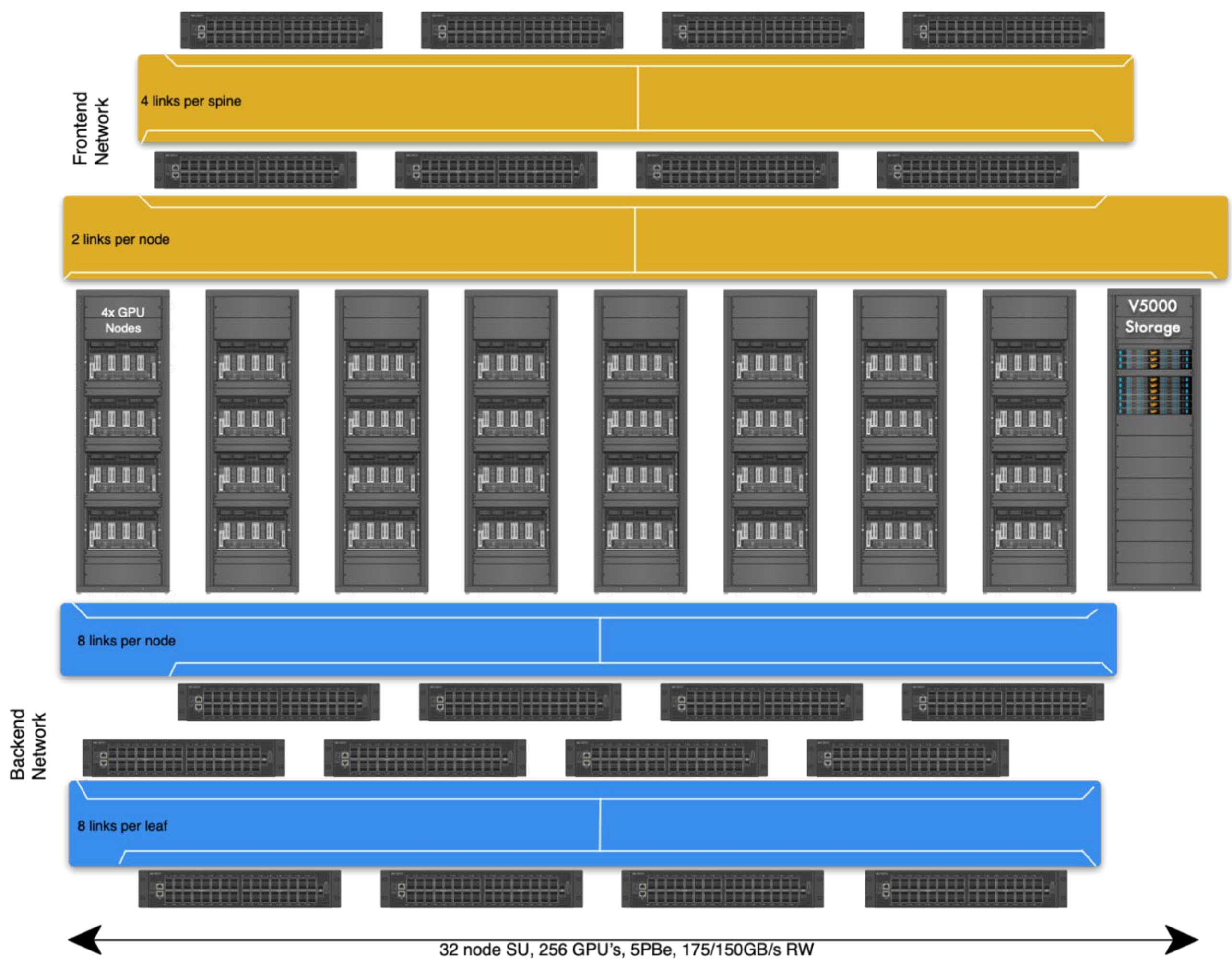


Figure 1. AI Storage Reference Design scalable unit.

The AI Reference Architecture shown in Figure 1 is for a 32-node cluster scalable unit. The basic components of this architecture include VDURA V5000 Storage Nodes, AMD Instinct GPU nodes, VDURA DirectFlow Client on Linux OS, and 400 Gb Ethernet switches.

Reference Design Components

The components used in a 32-GPU node scalable unit design using VDURA V5000 storage and AMD Instinct-based compute nodes with separate front-end and back-end 400 Gb Ethernet-based network fabrics are described in Table 1.

Quantity	Component
6	VDURA V5000 F-Node
3	VDURA V5000 Director Node
20	Ethernet switches 64-port 400 GbE for high-speed networking
32	AMD Instinct-based GPU server
256	AMD Instinct MI300 Series GPU

Table 1. Full-rack reference design components (Ethernet-based compute cluster).

As seen in Figure 2 below, a rack of V5000 enclosures provides 5 PB of usable data storage capacity, and a rack of AMD Instinct MI300 Series compute nodes contains four nodes each with eight GPU nodes.

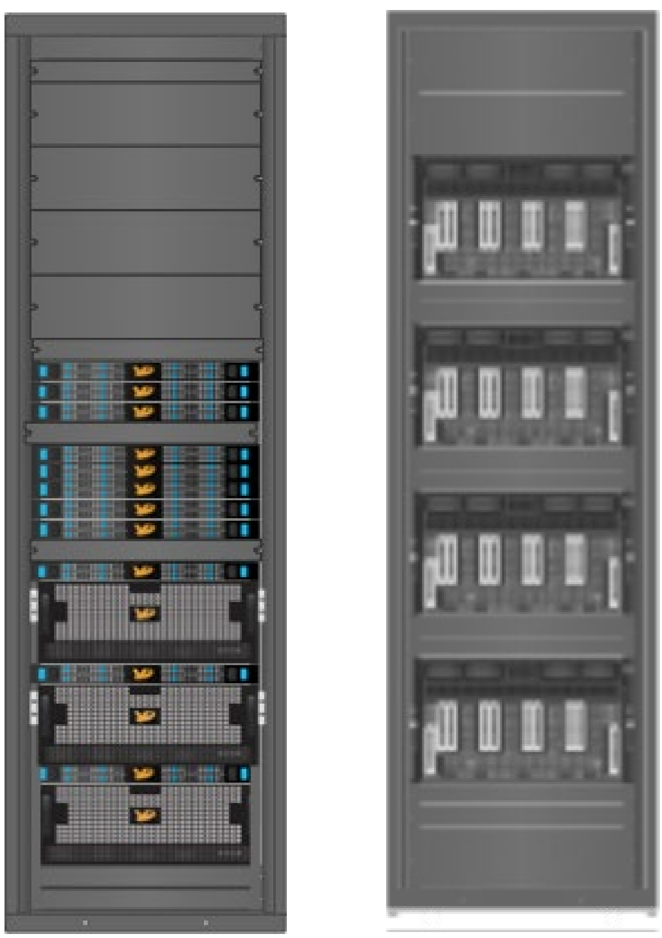


Figure 2. VDURA V5000 storage system and AMD Instinct MI300 series GPU nodes.

Reference Design Network Topology

The reference design network topology shows a no-single-point-of-failure networking topology for the VDURA V5000 reference design installation (see Figure 3).

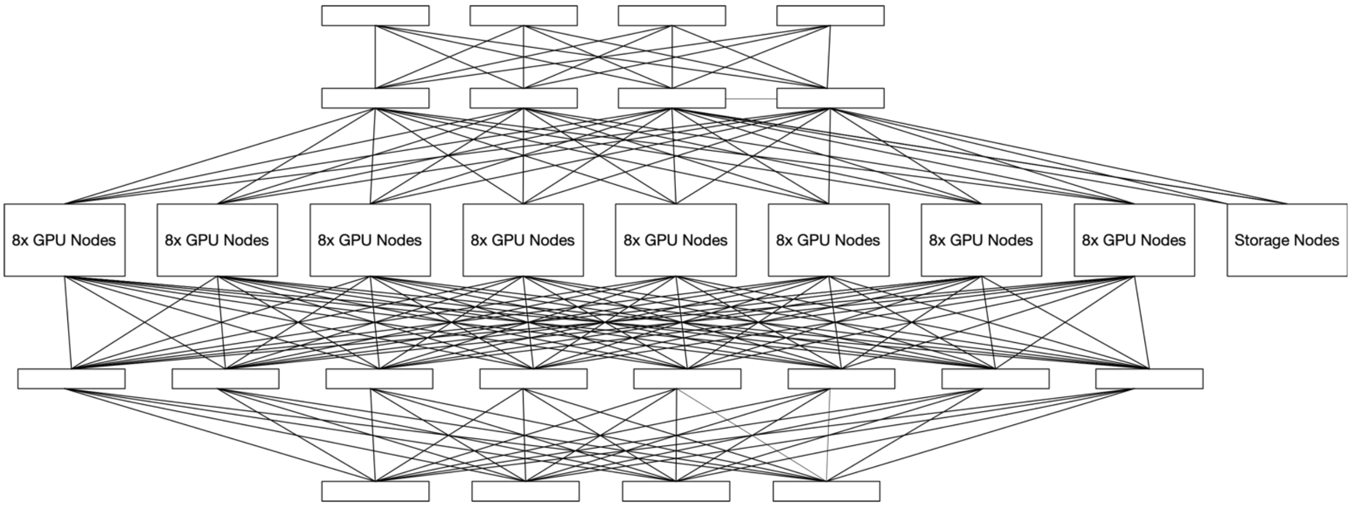


Figure 3. AMD Instinct MI300 Series + VDURA V5000 reference design network topology.

Storage Node Network Connectivity

The VDURA V5000 Storage Nodes defined in the AI Reference Design support 400 GbE networks via two network cards in the rear of each node (see Figure 4). The default configuration upon initial installation is link aggregation across two ports – a 2 x 400 GbE configuration using two 400 GbE QSFP112 cables, with one attached to each port. The VDURA V5000 Storage Nodes support Link Aggregation Control Protocol (LACP) by default; static Link Aggregation Group (LAG), single link, and failover modes are also available.

VDURA V5000 Storage Nodes also contain a single 1 GbE port that may be used as an out-of-band network port or for troubleshooting.

Figure 4. V5000 Storage Node rear view.

Compute System

On the AMD Instinct MI300 series (MI300X, MI325X) the compute node consists of eight OAM form factor GPUs in a Universal Baseboard (UBB) 2.0 design, interconnected by fourth-generation AMD Infinity Fabric™ links. Compute nodes and servers have typical dual-socket CPUs, memory, SSDs, and NICs for network connectivity.

An AMD Instinct MI300 series compute node consists of the following components:

Component	Specification
CPU	2x fourth-generation AMD EPYC™ processors
GPU	8x AMD Instinct MI300 Series Accelerators with AMD Universal Base Board (UBB 2.0)
Memory	Configurable; typical designs use 6 TB (24 x 256 GB DRAM) DDR5
Drives	NVMe SSDs; typical designs use 8-16 2.5-inch drives, 1-2 OS drives, and high performance scratch drives
Networking	8x PCIe 5.0 high-performance networking cards, 400 GbE 2x PCIe cards/DPUs for front-end/storage, 200 GbE Additional network for OOB management, 1-25GbE
Accelerator Interconnect	Incorporating AMD Infinity Architecture platform with Infinity Fabric interconnect
Cooling	Air cooling or liquid cooling

Table 2. AMD Instinct MI300 series compute node components.

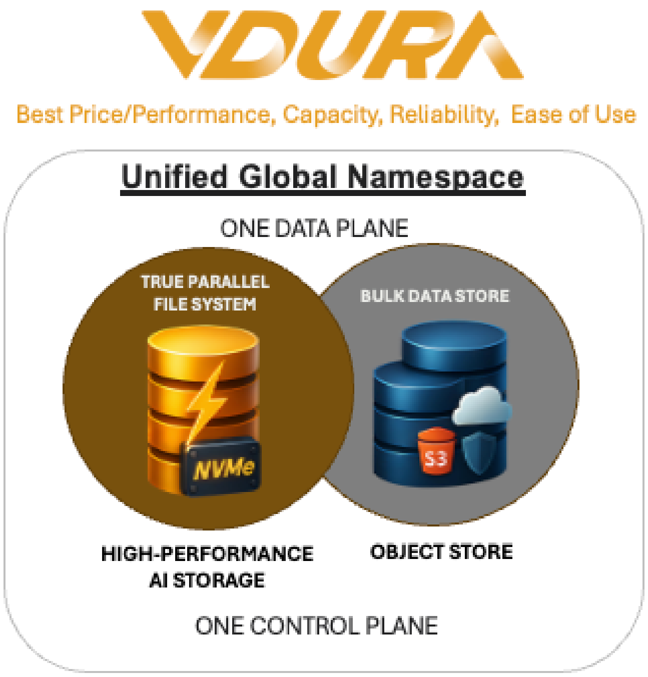
Compute nodes with AMD Instinct MI300 Series platforms are available from select vendors. See the Solution Catalog of AMD Instinct-powered servers.

With MI300 series GPUs, a cluster consists of a group of racks, each containing a group of nodes. The nodes/servers are placed in a rack with back-end switches for the back-end network and front-end/storage switches for front-end/storage network. The rack layouts are scalable and adjustable to meet the data center requirements.

A scalable unit consists of 32 nodes or 256 GPUs with a 64 x 400G back-end switch; a cluster is built from multiple scalable units. The AMD Cluster Reference Architecture Guide outlines the components required to build a back-end network cluster.

The VDURA Data Platform

Software-Defined Architecture Engineered for Performance and Scalability



The VDURA Data Platform V11 is built on a fully software-defined, microservices architecture that combines the speed and efficiency of a true parallel file system with the durability and cost-effectiveness of resilient object storage.

This unified design ensures high performance and simplicity for active and bulk data storage and is designed specifically to address the complexities and requirements of the AI pipeline.

The VDURA Data Platform explicitly separates the control plane handling metadata operations from the data plane, which is dedicated exclusively to user data storage.

With relentless focus on performance, simplicity, and scalability, VDURA empowers organizations to push the boundaries of AI. The platform delivers data at the speed of innovation, turning infrastructure into a competitive advantage for those shaping the future.

Three key components work together to power the VDURA Data Platform:

- Director Nodes are the core of the control plane.
- Storage Nodes are the foundation of the data plane.
- The DirectFlow Client is our high-performance parallel file system driver.

Director Nodes are the core of the control plane. They orchestrate and manage all metadata operations, coordinate the actions of Storage Nodes and DirectFlow Client drivers for file access, maintain the health and membership status within the storage cluster, and oversee all recovery and reliability functions. These nodes are simple, powerful compute servers featuring high-speed networking, substantial DRAM, and NVMe flash optimized for metadata transaction logs.

The proprietary **VDURA VeLO metadata engine** runs on each Director Node. VeLO is a distributed, flash-optimized key-value store designed specifically for high-speed, parallel metadata operations.

This integration ensures ultra-low latency, efficient handling of billions of file operations, and consistent metadata performance at scale.

Storage Nodes form the foundation of the data plane, dedicated exclusively to storing and managing user data. Available in configurations of either all-NVMe flash for peak performance or NVMe flash with HDD capacity expansion for high-performance, economical bulk storage, Storage Nodes deliver versatile and optimized infrastructure.

Each node hosts multiple **Virtualized Protected Object Device (VPOD™)** instances, enabling granular, scalable data management and enhanced reliability through Multi-Level Erasure Coding. VPOD architecture ensures linear scalability and consistent parallel performance, accommodating thousands of nodes seamlessly within a single cluster. This flexible yet robust structure guarantees data integrity, high throughput, and balanced resource utilization, aligning performance and economics precisely with workload requirements.

The VDURA DirectFlow Client is a high-performance parallel file system driver specifically engineered for Linux-based compute environments. Deployed directly on compute servers, DirectFlow seamlessly integrates with existing Linux applications, presenting itself like any conventional file system. It provides fully POSIX-compliant, cache-coherent file operations across a unified global namespace, tightly collaborating with Director and Storage Nodes. By enabling direct, parallel I/O paths from compute servers to Storage Nodes, DirectFlow eliminates traditional bottlenecks and intermediary processing overhead found in NFS or legacy storage solutions. Its support spans all major Linux distributions and versions, ensuring frictionless adoption and integration within existing infrastructure.

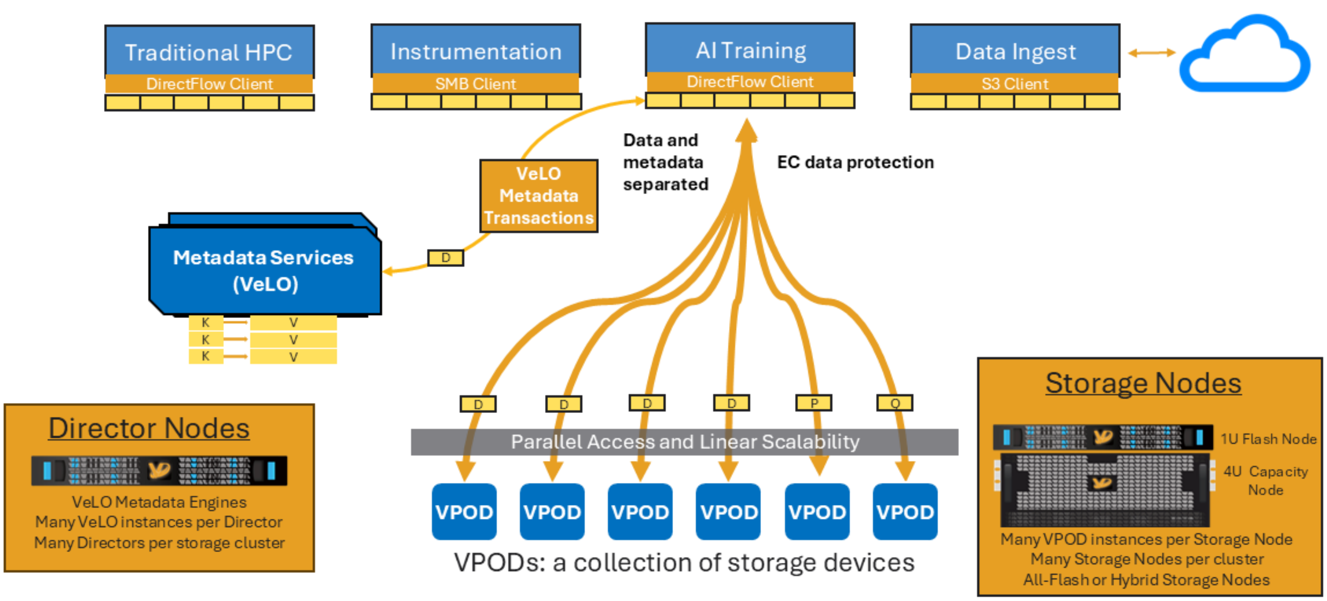


Figure 6. The VDURA Data Platform architecture.

Direct, Parallel Data Access for the AI Pipeline

The VDURA Data Platform is built as a true parallel file system, engineered to handle the intense I/O demands of modern AI and HPC workloads. Each file stored by the VDURA Data Platform is individually striped across many Storage Nodes, allowing each component piece of a file to be read and written in parallel, increasing the performance of accessing every file.

VDURA's parallel architecture dramatically accelerates data access, significantly boosting performance and throughput.

Unlike other enterprise systems which route data through limited head nodes, causing potential bottlenecks and requiring additional back-end network infrastructure, VDURA's DirectFlow Client communicates directly with all relevant Storage Nodes. Each compute server directly accesses the nodes holding the data, bypassing intermediary bottlenecks. Director Nodes manage metadata and coordinate system activity out-of-band, ensuring efficient data flow without interference or congestion.

This direct and parallel design eliminates traditional NAS hotspots, ensures predictable and scalable performance, and simplifies infrastructure by removing the need for a separate, costly back-end network. VDURA architecture delivers seamless scalability, consistently high performance, and exceptional efficiency across every stage of the AI pipeline, from ingest and training to inference and long-term data retention.

Linear Scalability, Seamless Expansion

The VDURA Data Platform delivers true linear scalability across both metadata and data services without compromise or complexity. AI workloads evolve fast, from early experimentation to scaled production across global clusters. Add Director Nodes to boost throughput for metadata-heavy tasks like model versioning and checkpoint tracking. Add Storage Nodes to scale bandwidth and capacity to support more training data, inference logs, or multi-tenant pipelines.

VDURA enables linear scalability, and growth is seamless and predictable. A 50 percent increase in Storage Nodes delivers 50 percent more throughput and capacity—no bottlenecks, no architectural redesigns.

This flexibility is powered by VDURA's fully virtualized, distributed software stack.

- **VeLO Metadata Engine** instances run in-memory across Director Nodes, scaling to billions of parallel metadata operations, perfect for high-frequency file creation, access pattern analysis, and rapid AI job cycles.
- **VPODs** manage user data in independently scalable units, each with its own erasure-coded stripe and logic, ideal for bursty checkpoint writes, long-term data lake retention, or active model training sets.

Together, these software-defined services form an intelligent, self-balancing system that grows effortlessly with the pipeline. Whether scaling GPU clusters, expanding training data, or retaining more checkpoints, VDURA adapts instantly.

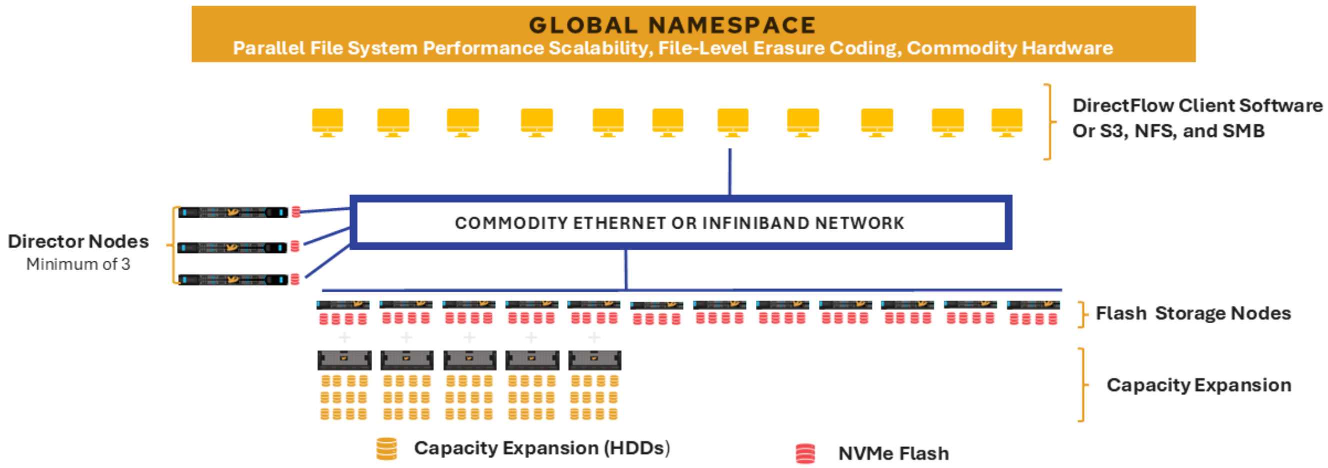


Figure 7. Linear scalability of Director and Storage Nodes.

Director Nodes: The Brain Orchestrating Control in a Parallel World

VDURA separates the **control plane**, which handles metadata, orchestration, and policy, from the **data plane**, which handles user I/O.

Director Nodes serve as the brain in the VDURA architecture. As the control plane's core, they command every stage of the AI pipeline, from ingestion and training to checkpointing and inference. Director Nodes continuously adapt to workload changes, ensuring optimal throughput and seamless orchestration across the system.

Each Director runs **VeLO**, a distributed, flash-optimized key-value metadata engine built to handle billions of operations per second. For modern AI, where performance is dictated as much by metadata velocity as data throughput, VeLO is essential. Tiny files, checkpoint indices, model versions—VeLO accelerates them all.

Director Nodes form the authoritative layer of VDURA's control structure. Every deployment requires a minimum of three. Administrators configure either three or five of the total Director Nodes as a *repset*, a voting quorum that maintains a synchronized, fully replicated configuration database. One node from the repset is elected *realm president* and is tasked with managing configuration, status monitoring, and leading failure recovery. If the current president fails, a new one is elected instantly and automatically.

Beyond coordination, Director Nodes also perform essential tasks at the president's request. These include managing volumes, serving as protocol gateways (NFS, SMB, S3), performing background data scrubbing, recovering failed Storage Nodes, and executing Active Capacity Balancing across

VPODs. All changes are non-disruptive to clients; gateways and volumes can migrate transparently across nodes when necessary.

From petabyte-scale model staging to tiering outputs across flash and hybrid storage, Director Nodes orchestrate every aspect of the control plane. This intelligent coordination empowers the data plane to operate with high efficiency across every stage of the AI pipeline.

Storage Nodes: AI Pipeline Performance from Every Layer

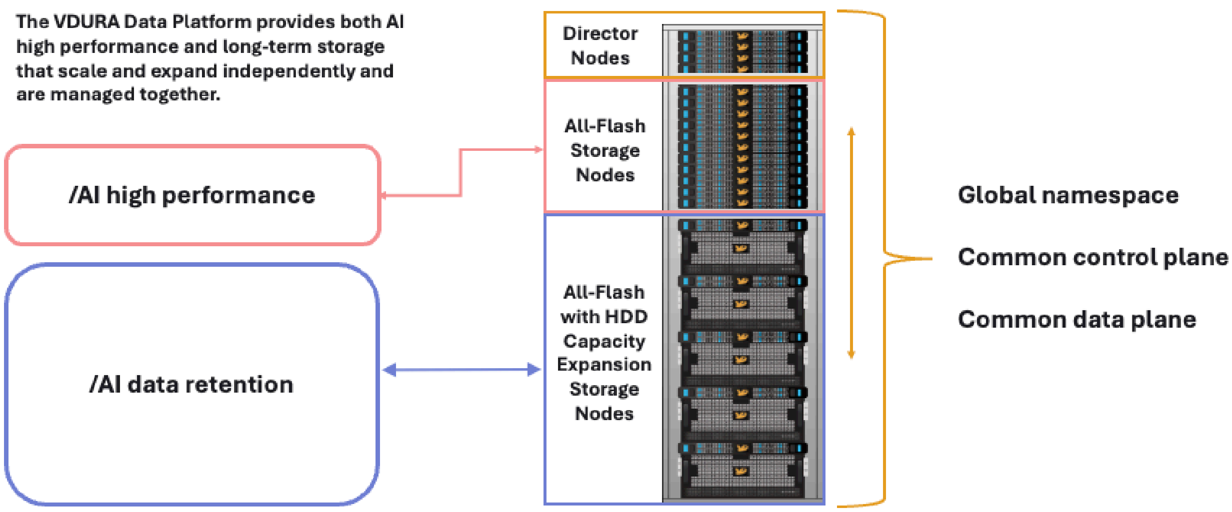


Figure 8. One control plane, one data plane, one single global namespace.

Storage Nodes are the backbone of VDURA's data plane, enabling seamless scale and sustained performance throughout every stage of the AI pipeline. Designed with flexibility and resilience, these nodes combine the best of both all-NVMe flash and flash with HDD capacity expansion storage, orchestrated under a unified control plane and single global namespace.

Optimized for Every Phase of the AI Pipeline

From high-frequency ingest and bursty checkpointing to real-time inference and long-term retraining, each phase of AI benefits from storage tiers purpose-built for performance and durability:

- **All-NVMe flash nodes** deliver ultra-low latency and high IOPs for AI high-performance data and latency-sensitive phases like model loading, active training, and checkpointing.
- **NVMe flash nodes with HDD capacity expansion** combine flash for metadata and active datasets with high-capacity HDDs for scalable, cost-efficient AI data retention. This is ideal for archived model weights, retraining inputs, inference logs, and data lakes that need fast access but are less frequently touched.

All Storage Nodes operate under a **common data plane and global namespace**, allowing applications to seamlessly span performance and capacity tiers without complexity or disruption.

VPOD Architecture: Virtualization for Resilience and Efficiency

Each VDURA Storage Node hosts multiple **VPODs**, rather than treating the entire server as a single failure domain. This architecture introduces a finer unit of failure isolation:

- **The unit of failure is now one VPOD**, not the physical server.
- More VPODs per node increases operational granularity and cluster flexibility.
- **Device-level failures are isolated to a single VPOD**, eliminating the risk of full node failure.
- Storage reconstruction only affects the failed VPOD's component objects, not the entire node.

Files are striped across component objects in multiple VPODs using **N+2 erasure coding**, ensuring high fault tolerance with efficient space utilization. Large POSIX files benefit from this distributed protection model, while small POSIX files are triple replicated across VPODs, delivering optimal performance and storage efficiency.

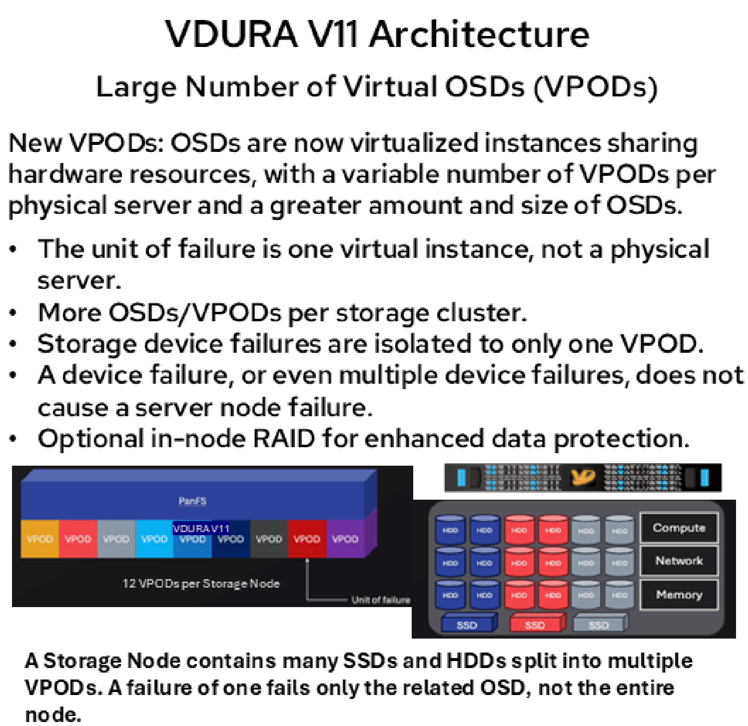


Figure 9. VDURA V11 architecture – Virtual Protected Object Devices (VPODs).

Dynamic Data Acceleration: The Smart Storage Fabric

VDURA **Dynamic Data Acceleration (DDA)** intelligently aligns I/O patterns with the most suitable media layer in real time:

- **Intent-log protection** powered by SSDs replaces legacy NVDIMMs for inflight data integrity.
- **Low-latency NVMe SSDs** store metadata databases for rapid namespace access.
- **High-IOPs SSDs** handle small file workloads.
- **High-bandwidth HDDs** manage large file sequential reads/writes.
- **System DRAM** provides caching for unmodified data and metadata.

Together, these layers form a high-performance, self-optimizing data fabric that minimizes latency and maximizes cost-efficiency.

Resilient Data Reconstruction and Integrity

In the event of a Storage Node failure, the VDURA Data Platform reconstructs only the affected component objects, not the full node's data. Files are rebuilt by pulling erasure-coded data fragments from other nodes. Continuous background scrubbing verifies data consistency across the system by validating erasure codes against stored data.

AI-Aware Placement and Automation

VDURA's intelligent orchestration engine continuously analyzes file size, access pattern, and data temperature to automate data placement across flash and hybrid tiers. Key features include:

- **Flash prioritization** for small and recently accessed files.
- **HDD capacity expansions** for large, sequential, or infrequently accessed data.
- **Continuous Active Capacity Balancing** to eliminate hotspots and evenly distribute load.
- **Real-time system adaptation** to shifting model training cycles, inference loads, and checkpoint bursts, requiring **zero manual tuning**.

The result is a storage system that evolves with the volatility of the AI pipeline, scaling performance and capacity without trade-offs.

VDURA V11 Data Reduction

VDURA V11 performs data reduction at the Storage Node level, ensuring zero impact on client-side CPU or memory resources. Unlike architectures that shift compression or deduplication tasks to the client, consuming valuable compute and memory, VDURA handles all reduction operations within the storage layer itself. This design keeps GPU and application nodes fully dedicated to AI and HPC workloads, maximizing performance and system efficiency.

The data reduction feature can be toggled on or off at any time via the graphical user interface (GUI) or command-line interface (CLI).

VDURA Data Platform V5000 System Overview

The VDURA Data Platform V5000 is an all-NVMe flash appliance engineered for AI and HPC pipelines that demand relentless GPU feed rates. Built on industry-standard servers, V5000 pairs flash performance with optional HDD capacity expansion, giving organizations a cost-balanced path from pilot to petabytes.

V5000 runs VDURA V11, VDURA's flash-tuned parallel file system, streaming multiple terabytes per second from a single global namespace. Working with the DirectFlow Client, VDURA offers parallel redundant data paths that scale linearly, safeguard data with enterprise-class durability, and keep day-to-day management simple.

Each system begins with a minimum of **three Director Nodes** and **three Storage Nodes**, which can be either **all-flash** or **flash with HDD capacity expansion.** Additional nodes can be added seamlessly to expand performance, capacity, or metadata throughput independently.

V5000 has ultimate configurability from 100 percent NVMe Flash to 98 percent HDD.

Add all-flash Storage Nodes to boost throughput and IOPs or attach capacity expansion enclosures for cost-efficient

Figure 10. V5000 configuration options, left to right: all flash, 50 percent flash, 98 percent HDD.

Key Features

- 1U chassis for Director Nodes and Storage Nodes.
- 4U chassis for HDD capacity expansion.
- Configurable as either **all-flash NVMe** or **all-flash NVMe with HDD capacity expansion.**
- Up to **1.4 TB/s throughput per rack** in all-flash configurations.
- Up to **280 GB/s throughput per rack** in all-flash with HDD capacity expansion configurations.

- Up to **1.2 M IOPs per Storage Node**, supporting low-latency, high-frequency workloads. Up to 45 M IOPs per rack with all-flash configuration.
- **Multi-Level Erasure Coding** with up to **12 nines durability (all-flash)** or **11 nines durability (all-flash with HDD capacity expansion).**
- Supports InfiniBand™ (NDR/NDR200) and Ethernet (400/200/100 GbE) networking.
- Hardware-agnostic deployment with full support for commodity components.

V5000 Details

The VDURA V5000 system represents the culmination of decades of engineering expertise in parallel file systems and distributed storage technology. Built for AI/ML and HPC workloads, the V5000 combines enterprise-grade reliability with maximum throughput and flexibility. Its modular architecture allows organizations to independently scale performance, capacity, and metadata operations to create the ideal balance for their specific workload requirements without overprovisioning or underutilization. Each component, from Director Nodes to Storage Nodes, is engineered for maximum efficiency and resilience, with built-in redundancy and intelligent self-healing capabilities that ensure continuous operation even during component failures.

Director Nodes

- Host the **VeLO Metadata Engine**, a distributed key-value store optimized for flash.
- **Contains 12 x NVMe SSDs slots.**
- Up to 333 M inodes and up to 225 K creates/deletes per Director.
- Requires a **minimum of three nodes** for metadata triplication and fault tolerance.
- Scales out to hundreds of instances for greater metadata throughput.

All-Flash Storage Nodes

- Configurable with **12 x NVMe SSDs** in 7.68 TB, 15.36 TB, or 30.72 TB capacities.
- Delivers >40 **GB/s per node and up to 1.4 TB/s throughput per rack**.
- Supports up to **1.2 M IOPs per node**.
- Ideal for AI training pipelines, inference, and checkpointing workloads.
- Fully compatible with Multi-Level Erasure Coding and data reduction features.

All-Flash with HDD Capacity Expansion Storage Nodes

- Combined SSD tier up to 12 3.84 TB, 7.68 TB, 15.36 TB, or 30.72 TB.
- Supports JBOD expansion using **78 or 108 HDDs per node** with 16 TB, 24 TB, 30 TB drive options.
- Up to **280 GB/s** flash accelerated throughput.
- Scalable to **26 PBe per rack** (effective) with inline compression.
- Supports up to **1.2 M IOPs per node**.
- Optimized for high-capacity/data lake storage with flash-accelerated performance.
- Supports dynamic tiering and intelligent placement via the **VDURA Orchestration Engine**.

Connectivity

- **Supported Network Options**
  - InfiniBand: NDR/NDR200
  - Ethernet: 400/200/100 GbE ports
- **Connectivity per Node**
  - Up to 2x NDR200 InfiniBand
  - Up to 4x 100 GbE Ethernet

Expansion Options

Each VDURA V5000 cluster can expand incrementally and non-disruptively.

- Add **Director Nodes** to increase metadata and protocol performance.
- Add **All-Flash Nodes** for higher throughput and IOPs.
- Add **All-Flash with HDD Capacity Expansion Storage Nodes** for cost-efficient capacity growth.
- Mix and match node types within the same realm with **no architectural redesign.**
- Tiered performance and capacity levels managed via **Storage Sets**, ensuring isolation and quality-of-service (QoS).

Physical Connectivity

The VDURA V5000 Director and Storage Nodes support 400/200/100 Gigabit Ethernet (GbE) networks via two network ports in the rear of each node. The default configuration upon initial installation is link aggregation across two ports—a 2 x 200/100 GbE configuration using two 200/100

GbE SFP28 cables, with one attached to each port. The VDURA V5000 nodes support Link Aggregation Control Protocol (LACP) by default; static Link Aggregation Group (LAG), single link, and failover modes are also available.

VDURA V5000 Director and Storage Nodes contain two 25/10 GbE ports for corporate network connectivity.

All nodes also contain a single 1 GbE port that may be used as a general administrative network port or for troubleshooting.

Network Configuration Options

There are four network configuration options:

- Dynamic LACP
- Static LAG
- Single link
- Failover network

The default network configuration for V5000 nodes is LACP across the dual 100 GbE ports.

Generally, protocols other than LACP and static LAG operate in active/passive mode.

Active/Active Link Aggregation Mode

When load balancing is required to optimize performance, V5000 systems can be configured to use either dynamic LACP or static LAG. LACP is preferred, as it is significantly more robust than static LAG.

In static LAG mode, the physical ports are bonded with the IEEE 802.3ad static LAG link-layer protocol. This provides both load balancing and fault tolerance if a port loses its physical carrier status. Static LAG may fail to detect when the port stops functioning properly, but its carrier state will remain active.

In LACP mode, the physical ports are bonded with the IEEE 802.3ad LACP link-layer protocol. This provides load balancing, better fault tolerance, and protection against misconfiguration than static LAG.

Single Link Mode

While single link mode is supported on V5000 systems, it is not optimal since it is a single point of failure and suffers reduced bandwidth. Thus, single link mode should be used with caution, as loss of the one single link will make the node inaccessible.

Network Failover Mode

Network failover is used on V5000 systems when active/passive redundancy is required.

Storage Configuration Options

VDURA provides two mechanisms to manage namespace and capacity: Storage Sets and volumes.

Storage Sets

The **Storage Set** is a physical mechanism that groups Storage Nodes into a uniform storage pool. It is a collection of three or more Storage Nodes grouped together to store data. You can grow a Storage Set by adding more hardware, and you can move data within a Storage Set.

Volumes

A **volume** is a logical mechanism, a sub-tree of the overall system directory structure. A read-only top-level root volume ("/"), under which all other volumes are mounted, and a /home volume are created during setup. All other volumes are created by the user on a particular Storage Set, with up to 1,200 per realm.

A volume does not have a fixed space allocation but instead has an optional quota which can set a maximum size for the volume. Effectively, the volume is a flexible container for files and directories, and the Storage Set is a fixed size container for multiple volumes.

When planning volume configuration, keep the following points in mind:

- Volumes can be used to manage capacity.
- Volumes can be created to complement backup strategies.
- Performance can be enhanced by assigning volumes to different Directors.

RAID and Erasure Coding

Storage nodes in the VDURA Data Platform are highly sophisticated VPODs, and we gain the same scale-out and shared-nothing architectural benefits from our VPODs as any object store would.

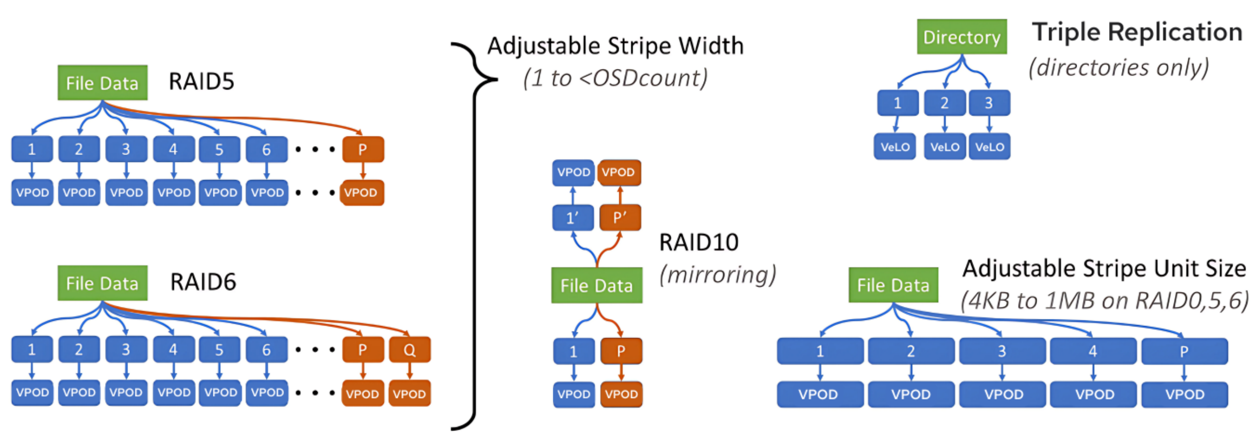


Figure 11. Per-file erasure coding layouts.

VDURA defines Objects used in our VPODs per the Small Computer System Interface (SCSI) standard definition of Objects rather than the Amazon S3 Object definition. The VDURA Data Platform uses SCSI Objects to store POSIX files, but it does so differently than how S3 Objects are typically used to store files. Instead of storing each file in an Object like S3 does, VDURA stripes a large POSIX file across a set of VPODs and adds additional VPODs into that stripe that store the P and Q data protection values of an N+2 erasure coding scheme. Using multiple VPODs per POSIX file enables the striping that is one of the sources of a parallel file system's performance.

A RAID array reconstructs the contents of drives, while VDURA reconstructs the contents of files.

While large POSIX files are stored using erasure coding across multiple VPODs, small POSIX files use triple-replication across three VPODs. This approach delivers higher performance than can be achieved by using erasure coding on such small files, while being more space efficient. Unless the first write to a file is a large one, it will start as a small file. If a small file grows into a large file, the Director Node will transparently transition the file to the erasure coded format at the point that the erasure coded format becomes more efficient.

When a file is created, and as it grows into a large file, the Director Node that manages those operations will randomly assign each of the individual VPODs that make up that file to different Storage Nodes. No two VPODs for any file will be in the same failure domain.

Reliability That Increases with Scale

Any system can experience failures, and as systems grow larger, their increasing complexity typically leads to lower overall reliability. For example, in an old-school RAID system, since the odds of any given HDD failing are roughly the same during the current hour as they were during the prior hour, more time in degraded mode equals higher odds of another HDD failing while the RAID system is still degraded. If enough HDDs were to be in a failed state at the same time, there would be data loss, so recovering back to full data protection levels as quickly as possible becomes the key aspect of any resiliency plan.

If a VDURA Storage Node fails, the system must reconstruct only those VPODs that were on the failed Storage Node, not the entire raw capacity of the Storage Node like a RAID array would. The system would read the VPODs for each affected file from all the other Storage Nodes and use each file's erasure code to reconstruct the VPODs that were on the failed node.

The VDURA Data Platform has linear scale-out reconstruction performance that dramatically reduces recovery time in the event of a storage node failure, so reliability increases with scale.

When a Storage Set is first set up, it sets aside a configurable amount of spare space on all the Storage Nodes in that Storage Set to hold the output from file reconstructions. When the system reconstructs a missing VPOD, it writes it to the spare space on a randomly chosen storage node in the same Storage Set. As a result, during a reconstruction, the system uses the combined write bandwidth of all the storage nodes in that Storage Set. The increased reconstruction bandwidth results in reducing the total time to reconstruct affected files, which reduces the odds of an additional failure during that time and increases the overall reliability of the realm.

VDURA also continuously scrubs the data integrity of the system in the background by slowly reading through all files in the system, validating that the erasure codes for each file match the data in that file. Data scrubbing is a hallmark of enterprise-class storage systems and is only found in one HPC-class storage system, the VDURA Data Platform.

An Architecture of High Reliability

Based on system configuration, the N+2 erasure coding that VDURA implements protects against either one or two simultaneous failures within any given Storage Set without any data loss. The realm can automatically and transparently recover from more than two failures, as long as there are no more than two failed Storage Nodes at any one time in a Storage Set.

If, in extreme circumstances, three Storage Nodes in a single Storage Set were to fail at the same time, VDURA has one additional line of defense that would limit the effects of that failure. All directories are independently stored triplicated—three complete copies of each directory, with no two copies on the same Director Node.

If a third Storage Node were to fail in a Storage Set while two others were being reconstructed, that Storage Set would immediately transition to read-only state. Only the files in the Storage Set that had VPODs on all three of the failed storage nodes would have lost data. All other files in the Storage Set would be unaffected or recoverable using their erasure coding. The number of affected files in these situations becomes smaller as the size of the Storage Set increases.

Since the system will have independent metadata storage that can survive against two simultaneous failures, it can identify the full pathnames of precisely which files need to be restored from a backup or reacquired from their original source and can therefore also recognize which files were either unaffected or recovered using their erasure coding.

VDURA is unique in the way it provides clear knowledge of the impact of a given event, as opposed to other architectures which leave you with significant uncertainty about the extent of the data loss.

Per-File RAID

Instead of relying on hardware RAID controllers that protect data at a drive level and computing RAID on the disks themselves, VDURA architecture uses per-file distributed RAID in software using erasure codes, i.e., per-file erasure coding rather than hardware RAID. Files in the same Storage Set, volume, and even directory can have different RAID/erasure code levels. A file can be seen as a single virtual object that is sliced into multiple component objects.

Users have three RAID levels available: RAID 6, RAID 5, and RAID 10.

- RAID 6 is the default RAID level for all volumes, as it provides a balance between performance, capacity overhead, and RAID protection.
- RAID 5 volumes are optimized for performance and capacity overhead.
- RAID 10 volumes combine striping and mirroring to provide high performance for applications that perform small random writes, while at the same time providing resiliency against a single disk failure.

You can mix RAID levels and any volume layout together in the same Storage Set, with each volume evaluated independently for availability status.

Comprehensive Data Services

VDURA's commitment to data management extends beyond performance and scalability. The platform offers a full suite of data services designed to ensure business continuity, protect data, and streamline management operations.

Snapshots, quotas and data migration are all standard features of VDURA, providing enterprises with the tools they need to manage their data effectively.

Tiered storage is another key feature of VDURA, with the platform's Dynamic Data Acceleration (DDA) technology supporting up to four different performance tiers. This ensures that data is placed on the most appropriate storage media based on its characteristics, optimizing both cost and performance.

Security is woven into every aspect of VDURA, with industry-leading AES-256 encryption for data in flight and at rest, comprehensive access controls, and detailed logging and auditing capabilities. End-to-end encryption provides stronger data confidentiality, integrity, and compliance alignment, surpassing traditional TLS + SED approaches.

PanFS is also available from select OEMs.

System Limits and Requirements

General limits and requirements of the PanFS filesystem and V5000 storage system are provided in Table 3.

Limit	Description
Minimum V5000 system	3 D-Node enclosures with 3 F-Node or S-Node enclosures
Largest V5000 system	No enforced limit VDURA has tested up to 1300 OSDs
Maximum number of files or directories	No enforced limit
Maximum file size	CIFS: 32 TB NFS: 32 TB (FreeBSD buffer cache limitation) 32-bit Linux DirectFlow: 16 TB 64-bit Linux DirectFlow: 8192 PB
Maximum number of volumes	Per Director: no limit Per Storage Set: no limit Per realm: 3600
Maximum number of Realm Managers (RMs) in repset	5
Maximum Storage Set size	No enforced limit
Maximum number of snapshots	32 (enforced) per volume
Maximum time between snapshots	10 minutes (enforced) 30 minutes (recommended) Many volumes within the same Storage Set can be scheduled to have a snapshot taken at the same time; these will be batched and optimized. The system does snapshot work in the background and, in some cases, will reject a snapshot if activity associated with a previous snapshot is still in progress. Snapshots may also be taken more frequently using the manual **Take a Snapshot Now** button.
Maximum number of entries per directory	1,000,000 (enforced) 250,000 (recommended)
Maximum number of clients	30,000
Maximum CIFS clients per Director	No limit
Maximum path length	4096 bytes

Table 3. PanFS and V5000 system limits and requirements.

Networks and Interconnects

Networking for AI places a huge demand on building a robust and resilient framework due to the high bandwidth needs, variants of traffic flows between the type of workloads, and the need to separate the computational traffic from the storage and user data communications.

AMD has deep investment in developing a framework for networking containing both hardware and software elements that are composed of at least three network fabrics. These three fabrics are defined as:

- **Front-End and Storage Network**: The front-end network consists of NICs and Ethernet switches. It handles data ingestion, communication to storage devices, in-band management functions, and secure user access.
- **Back-End "Scale-Out" Network**: Built using PCIe 5.0 NICs, RDMA over Converged Ethernet (RoCEv2, or UEC Ready RDMA) can scale from 16 GPUs to 1M+ GPUs.
- **Accelerator "Scale-Up" Network High BW Low Latency Network**: A network of GPUs connected by AMD Infinity Fabric, enabling high-speed, low-latency inter-GPU communication.

As datasets for AI workloads continue to expand in size, it is becoming increasingly critical that GPUs are not constrained by the I/O network and storage systems. The storage fabric provides the communication path between GPUs and the storage systems.

Ethernet continues to be the leading choice to build the back-end (scale-out) and front-end network. Ethernet is built around open standards and ecosystems, is highly scalable, cost effective, and has widespread adoption with multiple vendors. AMD is participating in standardization efforts related to scale-out Ethernet networking by being a founding member of UEC (Ultra Ethernet Consortium) and having partnerships in ethernet ecosystem.

An Ethernet deployment consists of Ethernet NICs and switches.

- The AMD Pensando™ Pollara 400G NIC is UEC ready. It can run on any Ethernet fabric and offers key innovations for networking in the AI era.
- Ethernet switches for AI networks must meet demanding requirements due to the massive scale, high throughput, and low latency needed for distributed AI and machine learning workloads. Features like end-to-end congestion control, dynamic load balancing, and priority-based flow control are essential to improve network efficiency and maximize job completion times.

Ethernet switches can form either two-tier (leaf-spine) or three-tier (leaf-spine-core) switching fabric. In the *AMD Cluster Reference Architecture Guide*, AMD demonstrates how a design is achieved by connecting storage systems through the storage leaf/spine switches using the storage network fabric on AMD Instinct MI300 series based compute nodes:

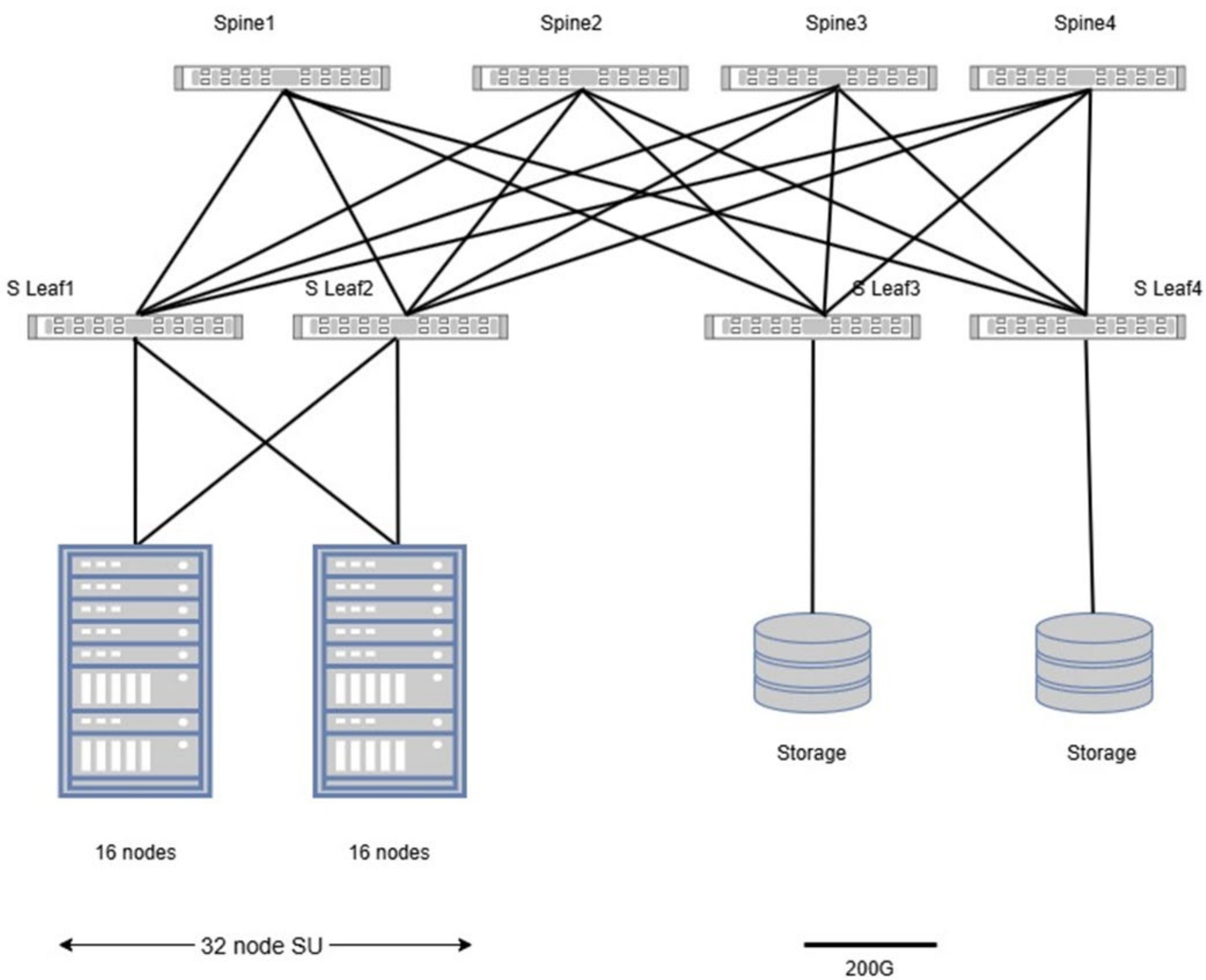


Figure 11. Leaf and spine switching fabric design.

Designing the storage fabric:

- Compute nodes with AMD Instinct MI300 series-powered servers. A catalog of AMD Instinct Accelerator-powered servers is available from [AMD Instinct Solutions](#).
- Ethernet network adapters
    - AMD Pensando™ Pollara 400 (back-end NIC)
    - AMD Pensando DSC3-400 (front-end DPU) or
    - AMD Pensando Giglio DPU 200 (front-end DPU)
- Leaf and spine storage Ethernet switches. Select storage switch vendors are listed in the [AMD MI300 Series Cluster Reference Architecture Guide](#).
- VDURA storage systems

Management Network

VDURA uses a 1 GbE management network for IPMI to V5000 nodes in the AI Storage Reference Design. While IPMI is optional, it makes remote administration much easier.

Support

Outstanding support is an essential ingredient in minimizing downtime in technical computing environments. VDURA provides worldwide enterprise-class solutions, along with OEM support for compute nodes. In addition, customized offerings to meet specific support requirements are available.

VDURA support services include the MyVDURA online self-service portal, where users can access support resources anytime, anywhere. MyVDURA also offers custom software downloads and a robust knowledge base that can help quickly resolve technical issues.

For companies with sensitive data requiring a high level of confidentiality and security, as well as the ability to meet classified data center requirements, VDURA offers support plan options for secure media and/or hardware disposal. The VDURA Support Services team brings many years of operational experience in highly secure environments and understands how to optimize secure installations.

Summary

Navigating the open waters of AI computing is a difficult task, made more so by the deluge of new technologies emerging every day. For many business organizations and academic and government research centers, the effective use of HPC and AI environments provides substantial improvements, from processing and analyzing data, to discoveries and product development, and even to developing and enhancing finance and business strategies.

Design, implementation, and support of HPC infrastructures are all often complicated and confusing tasks and can benefit from a turnkey approach. To that end, the AI Storage Reference Design in this document provides solutions using AMD Instinct GPU -based compute clusters paired with a full-rack configuration of VDURA V5000 storage systems, networking, interconnects, software, and management components that are integrated, tested, and supported.

The reference designs in this guide are both prescriptive and descriptive, as well as instructive in providing future paths. Included are component details and configuration options to guide HPC solution architects and system administrators in their planning of high-performance systems that can accelerate and support the converged application workloads of today's modern HPC users, including concurrent and integrated workflow of modeling and simulation with HPDA and ML. The reference designs are tested to support workloads using both CPUs and accelerators such as GPUs, high-speed Ethernet switches, and high-performance VDURA storage, itself using an architecture and engine designed to meet the rigors of many different users running many different applications at the same time.

To learn more about VDURA and AMD technologies and products, visit us at www.VDURA.com and www.amd.com.