WHITE PAPER

AI STORAGE
REFERENCE DESIGN

VDURA Data Platform V12 + AMD Instinct™ MI350 Series and MI300 Series

Harness the potential of GPU-accelerated AI

WHERE VELOCITY MEETS DURABILITY

A prescriptive, full-rack reference design pairing the VDURA Data Platform V12 with AMD Instinct™ GPUs. Built on the HYDRA architecture, this scalable unit delivers 5 PB of usable storage and 256 GPUs in a single rack with no single point of failure. Purpose-built for AI factories, Neoclouds, and the world's most demanding HPC environments.

Contents

Revisions

Date	Description
May 2025	Initial release with VDURA V11
July 2026	V12 update for AMD AI Reference Architecture

Executive Summary

AI factories and Neocloud operators are deploying GPU infrastructure at unprecedented scale, but storage remains the overlooked bottleneck preventing optimal performance. If storage cannot feed GPUs fast enough, training stalls, checkpoints burn compute dollars, and inference latency spikes. Storage is also the single largest blast radius in any GPU cluster: a 5,000-GPU cluster running at just 98% storage availability bleeds roughly 876,000 GPU-hours per year, approximately $2.6 million in idle compute.

This reference design pairs the VDURA Data Platform V12 with AMD Instinct™ MI3XX Series accelerators to deliver a prescriptive, tested, no-single-point-of-failure blueprint for AI infrastructure at any scale. A 32-node scalable unit delivers **256 GPUs and 5 PB of usable storage** in a single rack, with linear scaling to thousands of nodes. Multiple scalable units combine to build clusters of any size.

Built on the HYDRA architecture (High-Performance, Yield-Optimized, Distributed, Resilient), VDURA V12 combines a true parallel file system with a resilient object store under one data plane, one control plane, and one global namespace. By pairing NVMe flash for performance-critical workloads with high-capacity HDD for data retention, VDURA delivers the same mixed-fleet storage model that hyperscalers like Google, Meta, and Microsoft already deploy in production, now available to AI factories, Neoclouds, and enterprises building AMD-powered GPU clusters.

V12 introduces the Elastic Metadata Engine with up to 20× acceleration in metadata operations, native snapshot support for AI pipeline checkpoints, SMR HDD optimization unlocking 25–30% more capacity per rack, RDMA support for GPU-native data paths that bypass the CPU entirely, Context-Aware Tiering for long-context LLM serving, end-to-end AES-256 encryption, and a native CSI plug-in for Kubernetes. All of this is backed by at least six nines of availability and up to 12 nines of durability, proven across 1,000s of production deployments.

This document outlines the architectural components, performance capabilities, network design, and deployment options of the combined VDURA V12 and AMD Instinct platform, providing the prescriptive guidance needed to design, deploy, and scale GPU-accelerated AI infrastructure.

Introduction

The rapid evolution of Artificial Intelligence has posed unique challenges to systems infrastructure deployed across a broad spectrum of industrial and academic use cases. In contrast to traditional applications, generative AI 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, sensor streams). Because very large models and datasets cannot fit in a single GPU's HBM or even a

single node, they are distributed with advanced parallelism techniques across nodes and racks. As models, datasets, and context length grow in complexity and size, scaling infrastructure for training is a critical challenge and presents new opportunities.

Infrastructure requirements are highly dynamic for different phases of the AI pipeline. Prior to AI training, 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 to TBs) of traffic between GPUs across back-end networks for collectives and front-end traffic, especially for persisting and loading checkpoints, requiring high throughput. AI inferencing workloads such as AI Vector DBs, Retrieval-Augmented Generation (RAG), and agentic workloads have additional latency sensitivity.

> ***Infrastructure deployed for AI workloads (training or inference), especially for Large Language Models, 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 architectures 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, 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, covering: 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 defined number of compute nodes with NICs in the front-end network, local storage, and networked storage interconnected by an Ethernet fabric. In this reference design, the compute nodes consist of AMD Instinct™ MI3XX Series GPUs, 400 GbE NICs for the back-end network, VDURA V5000 networked storage, and Ethernet switches. Compute nodes can access each other for full-bandwidth, all-to-all communication patterns, as well as the VDURA V5000 storage solution via the VDURA DirectFlow® client software. The VDURA Data Platform V12, the storage operating system and parallel file system that runs on the V5000, and the other major components of the AI Reference Architecture are detailed later in this document.

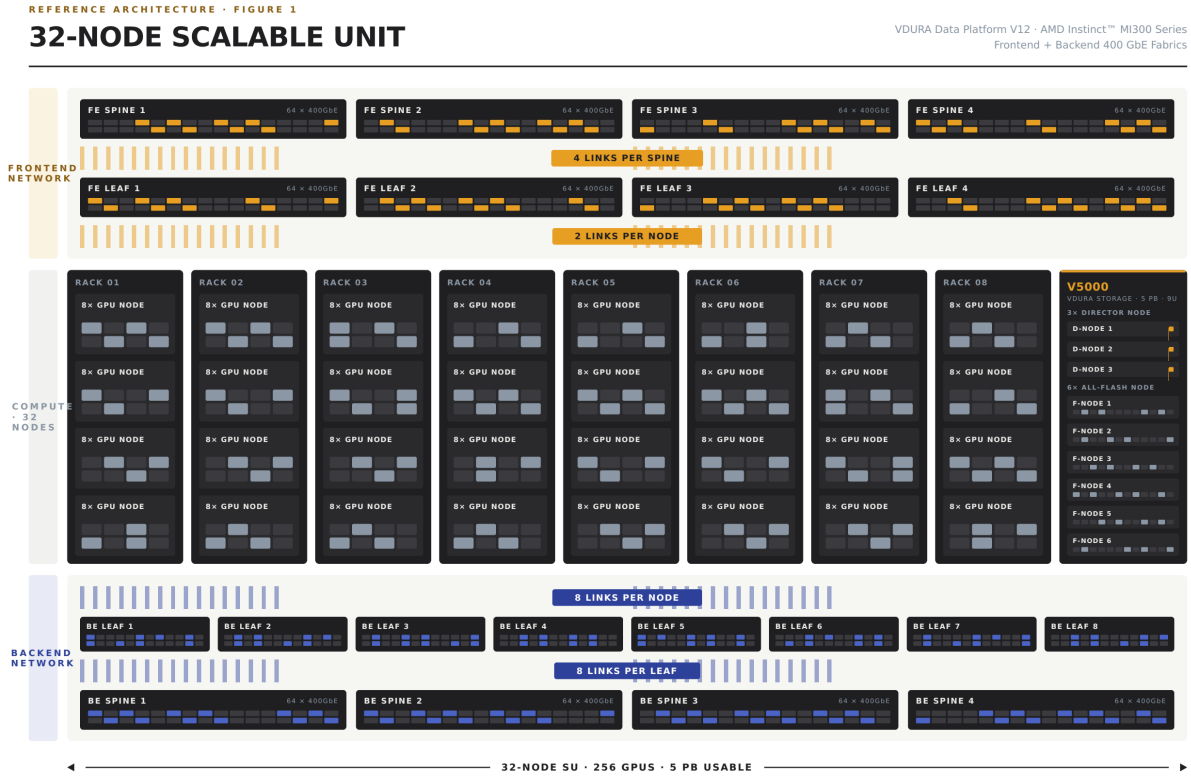

Figure 1. AI Storage Reference Design scalable unit.

The AI Reference Architecture shown in Figure 1 is a 32-node cluster scalable unit. The basic components of this architecture include VDURA V5000 Storage Nodes, AMD Instinct GPU nodes, the VDURA DirectFlow Client on Linux, and 400 GbE 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 GbE network fabrics are described in Table 1.

Quantity	Component
3	VDURA V5000 Director Node (D-Node)
6	VDURA V5000 All-Flash Storage Node (F-Node)
20	Ethernet switches: 64-port 400 GbE for high-speed networking
32	AMD Instinct™-based GPU server
256	AMD Instinct™ MI355X, MI350X, MI325X, or MI300X Series GPU

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

A rack of V5000 enclosures provides 5 PB of usable data storage capacity, and a rack of AMD Instinct MI3XX Series compute nodes contains four nodes each with eight GPUs.

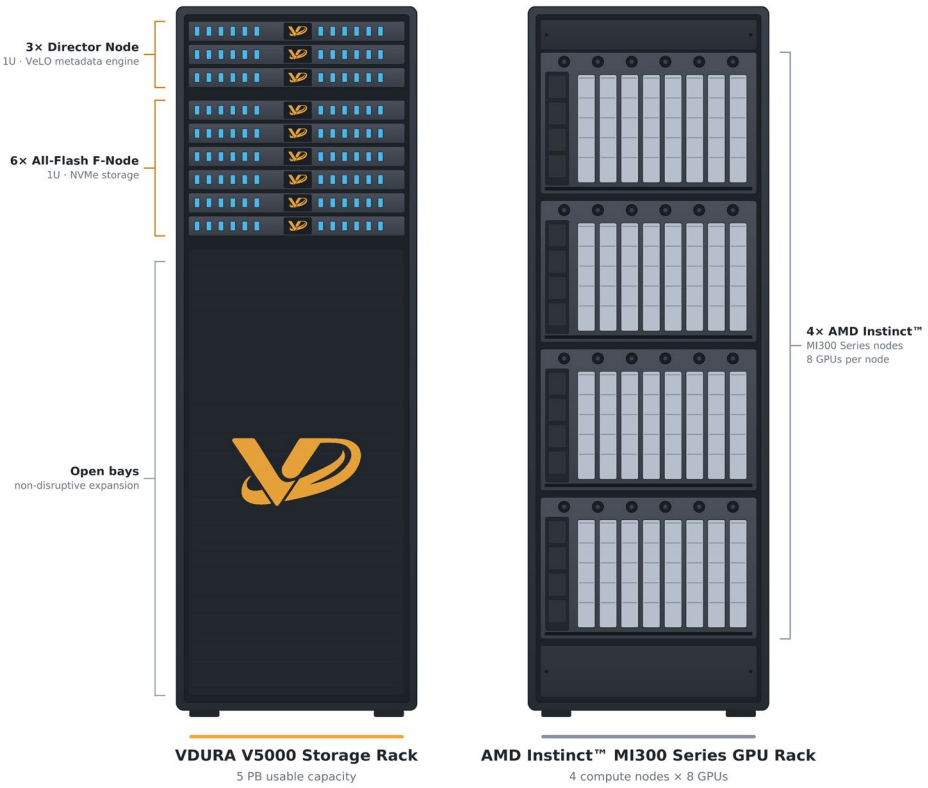

Figure 2. VDURA V5000 storage rack (left) and AMD Instinct MI3XX Series GPU rack (right).

Compute System

On the AMD Instinct MI3XX series 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 MI3XX series compute node consists of the following components:

Component	Specification
CPU	2× fourth-generation AMD EPYC™ processors
GPU	8× AMD Instinct™ MI3XX Series Accelerators with AMD Universal Base Board (UBB 2.0)
Memory	Configurable; typical designs use 6 TB (24 × 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	8× PCIe 5.0 high-performance networking cards, 400 GbE 2× PCIe cards/DPUs for front-end/storage, 200 GbE Additional network for OOB management, 1–25 GbE
Accelerator Interconnect	AMD Infinity Architecture platform with Infinity Fabric™ interconnect
Cooling	Air cooling or liquid cooling

Table 2. AMD Instinct MI3XX Series compute node components.

Compute nodes with AMD Instinct MI3XX Series platforms are available from select vendors. See the AMD Instinct Solution Catalog of AMD Instinct-powered servers.

With MI3XX Series GPUs, a cluster consists of a group of racks, each containing a group of nodes. The nodes and servers are placed in a rack with back-end switches for the back-end network and front-end/storage switches for the front-end/storage network. The rack layouts are scalable and adjustable to meet data center requirements.

A scalable unit consists of 32 nodes or 256 GPUs with a 64 × 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.

Networks and Interconnects

Networking for AI places a heavy demand on building a robust and resilient framework due to the high bandwidth needs, varied traffic patterns between workload types, and the need to separate computational traffic from storage and user data communications. AMD has invested deeply in developing a networking framework, in both hardware and software, that composes the modern AI cluster.

The Three Network Fabrics

An AMD Instinct-powered AI cluster is built on three distinct network fabrics, each optimized for a different traffic pattern:

- **Front-End and Storage Network.** Built from NICs and Ethernet switches. Handles data ingestion, communication between compute and storage, in-band management, and secure user access. This is where VDURA V5000 Storage Nodes attach to the cluster.
- **Back-End Scale-Out Network.** Built from PCIe 5.0 NICs running RDMA over Converged Ethernet (RoCEv2 or UEC-ready RDMA). Carries GPU-to-GPU collective traffic across the cluster, scaling from 16 GPUs to 1M+ GPUs.
- **Accelerator Scale-Up Network.** A high-bandwidth, low-latency network of GPUs connected by AMD Infinity Fabric™, enabling fourth-generation Infinity Fabric inter-GPU communication within and between MI3XX Series UBB 2.0 baseboards.

As datasets for AI workloads continue to expand, it is increasingly critical that GPUs are not constrained by the I/O network and storage systems. The front-end and storage fabric is the communication path between GPUs and the storage systems and is the primary focus of this reference design.

Ethernet continues to be the leading choice for the back-end (scale-out) and front-end networks. Ethernet is built around open standards, is highly scalable, cost-effective, and broadly adopted across multiple vendors. AMD is a founding member of the Ultra Ethernet Consortium (UEC) and participates in standardization efforts and ecosystem partnerships.

AMD Pensando NICs for AI Networking

An Ethernet deployment for this reference design consists of AMD Pensando NICs and supporting Ethernet switches.

- **AMD Pensando™ Pollara 400G NIC** is UEC-ready. It runs on any Ethernet fabric and delivers key innovations for AI-era networking. The Pollara 400G is the back-end NIC for AMD Instinct GPU compute nodes in this reference design.
- **AMD Pensando DSC3-400 or Giglio DPU 200** serve as front-end DPUs, providing storage and front-end network connectivity with hardware offload for stateful services.
- **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. End-to-end congestion control, dynamic load balancing, and priority-based flow control are essential to improve network efficiency and maximize job completion times.

V12 adds RDMA support to all V5000 systems, enabling GPU-to-storage data transfers that bypass the CPU entirely. Built on AMD Pensando™ Pollara 400G NICs and AMD EPYC™ Turin processors, RDMA delivers direct memory access between AMD Instinct GPU server nodes and the VDURA Data Platform, eliminating CPU bottlenecks for the low-latency, high-throughput data paths critical to AI training and inference.

Reference Design Network Topology

The reference design network topology is a no-single-point-of-failure (NSPF) topology for the VDURA V5000 reference design installation. Front-end and back-end networks are physically separated to isolate storage and user-data traffic from GPU collective traffic.

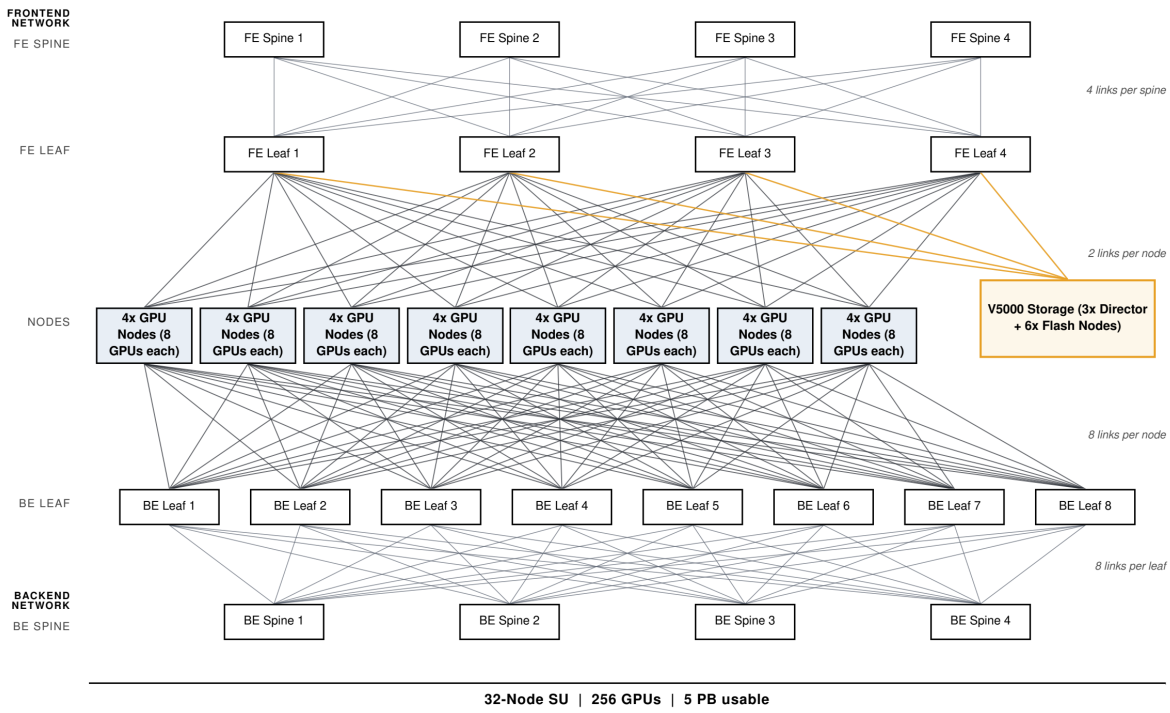

Figure 3. AMD Instinct MI3XX Series + VDURA V5000 reference design network topology.

In this design, eight groups of eight GPU nodes connect through dedicated 400 GbE leaf switches to a 400 GbE spine fabric. VDURA V5000 storage nodes connect to the same fabric through their own leaf switches, providing parallel redundant data paths that scale linearly with the cluster.

Leaf-Spine Fabric Design

Ethernet switches can form either two-tier (leaf-spine) or three-tier (leaf-spine-core) switching fabrics. In the AMD Cluster Reference Architecture Guide, AMD demonstrates a design that connects storage systems through dedicated storage leaf and spine switches using the storage network fabric on AMD Instinct MI3XX Series-based compute nodes:

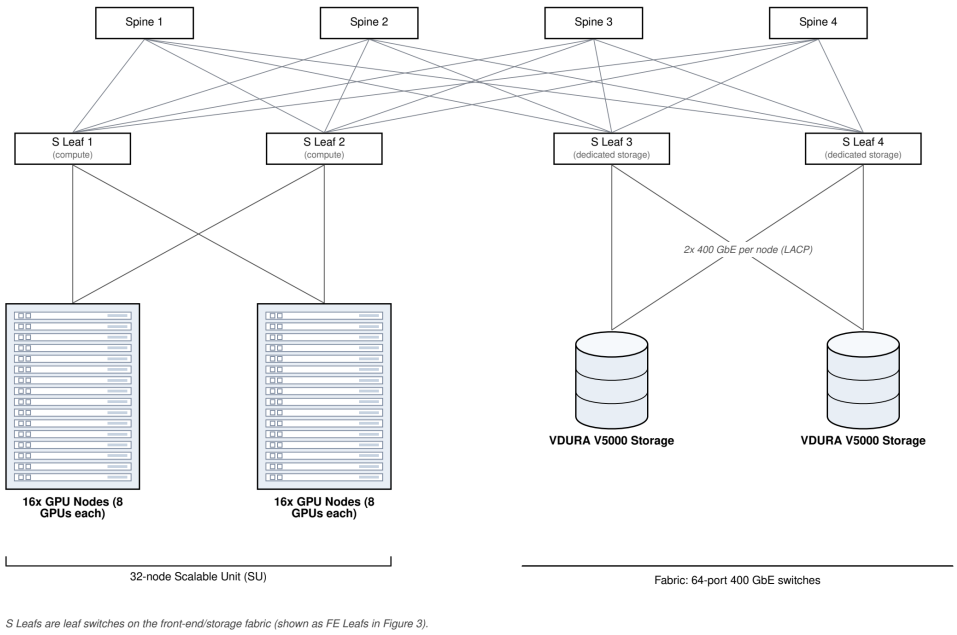

Figure 4. Leaf-and-spine switching fabric design with 32-node scalable units and dedicated storage leaves.

Designing the storage fabric:

- **Compute nodes** with AMD Instinct MI3XX Series-powered servers. A catalog of AMD Instinct Accelerator-powered servers is available from AMD Instinct Solutions.
- **Ethernet network adapters** including AMD Pensando™ Pollara 400 (back-end NIC), AMD Pensando DSC3-400 or Giglio DPU 200 (front-end DPU).
- **Leaf and spine storage Ethernet switches.** Select storage switch vendors are listed in the AMD MI3XX Series Cluster Reference Architecture Guide.
- **VDURA V5000 storage systems** running the VDURA Data Platform V12 with HYDRA architecture.

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. The default configuration upon initial installation is link aggregation across two ports, a 2 × 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.

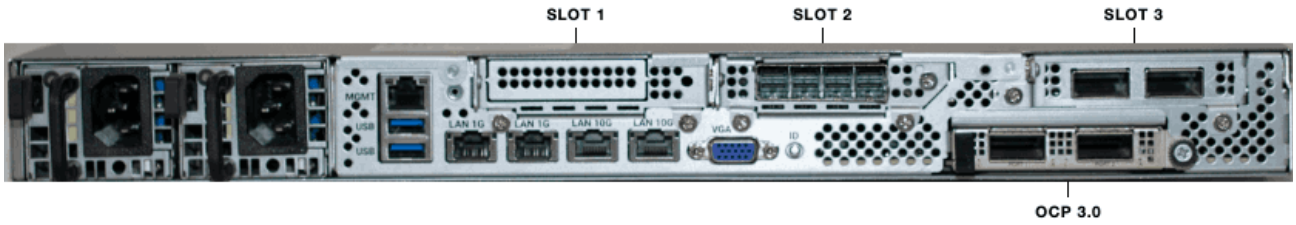

Figure 5. V5000 Storage Node rear view showing dual high-speed network slots and OCP 3.0 expansion.

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. The management network is physically separate from the front-end, back-end, and accelerator fabrics described above.

The VDURA Data Platform V12

HYDRA: High-Performance, Yield-Optimized, Distributed, Resilient

The VDURA Data Platform V12 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 is the HYDRA architecture: High-Performance, Yield-Optimized, Distributed, Resilient.

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.

> ***Three key components work together to power the VDURA Data Platform: Director Nodes (the core of the control plane), Storage Nodes (the foundation of the data plane), and the DirectFlow™ Client (the high-performance parallel file system driver).***

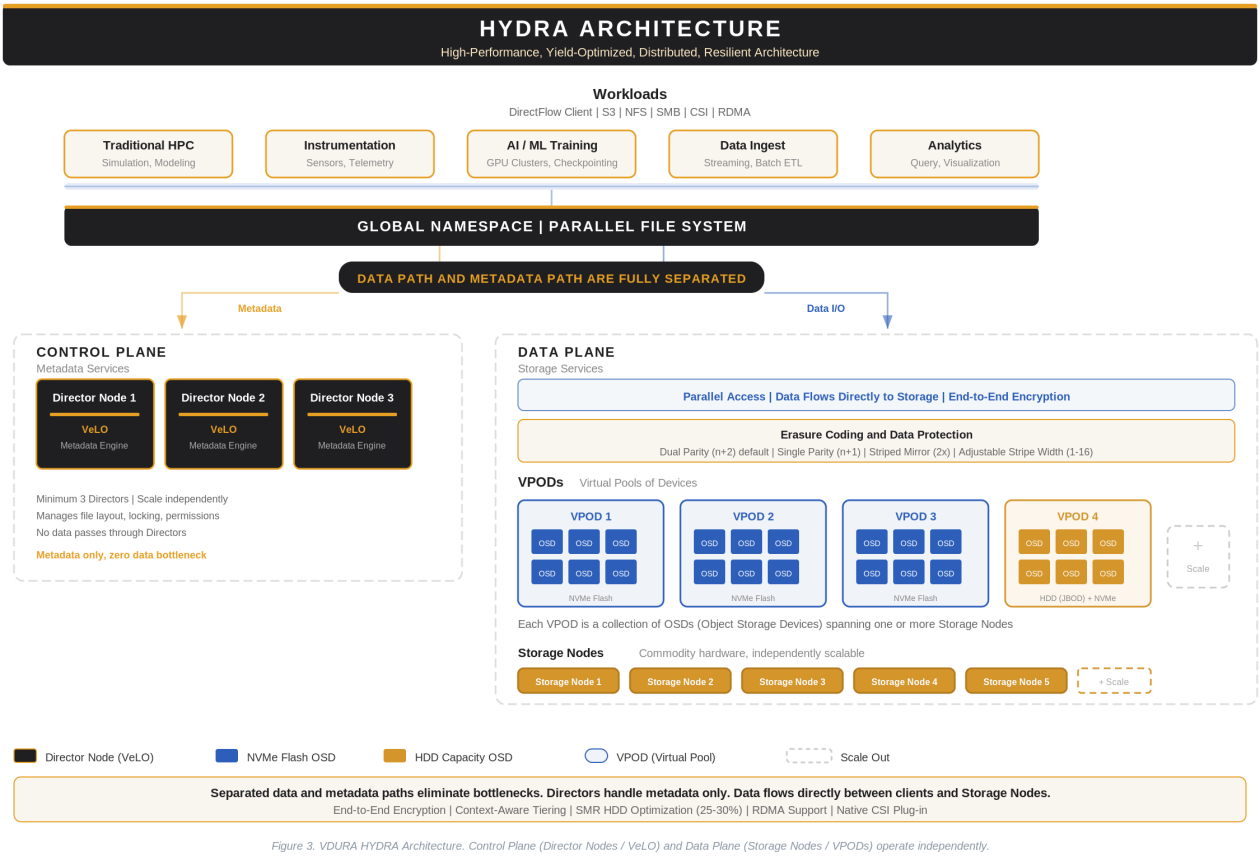

Figure 3. VDURA HYDRA Architecture. Control Plane (Director Nodes / VeLO) and Data Plane (Storage Nodes / VPODs) operate independently.

Figure 6. Control plane (Director Nodes / VeLO) and data plane (Storage Nodes / VPODs) operate independently in VDURA HYDRA.

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 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 VDURA VeLO™ metadata engine runs on each Director Node. VeLO is distributed and flash-optimized, designed specifically for high-speed parallel metadata operations. V12's Elastic Metadata Engine dynamically scales across nodes, delivering up to 20× improvement in metadata operations and supporting billions of files and objects under active use. 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. 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.

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. V12 adds RDMA support for GPU-native data paths that bypass the CPU entirely.

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.

The DirectFlow Client is lightweight, consuming approximately 191 MB of DRAM per compute node, requires zero dedicated CPU cores, and uses the standard Linux page cache rather than the pinned HugePages required by kernel-bypass data paths. CPU cycles are borrowed opportunistically during active I/O and returned immediately to the application. This efficiency matters at fleet scale: a 500-node GPU cluster running VDURA commits roughly 93 GB of DRAM to storage clients, while architectures that require kernel-bypass modes for peak performance reserve 2.5 TB of DRAM and permanently lock 500 to 2,000 CPU cores across the same fleet. Every core and every gigabyte VDURA does not claim stays available to AI applications running on AMD EPYC CPUs and AMD Instinct GPUs. V12 takes this further with RDMA and GPUDirect Storage support, enabling direct DMA transfers between storage and GPU HBM that bypass host DRAM and the CPU entirely.

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 for storage. 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 seamless, predictable growth. A 50% increase in Storage Nodes delivers 50% more throughput and capacity, with no bottlenecks and no architectural redesigns.

- **VeLO metadata engine:** instances run in-memory across Director Nodes, scaling to billions of parallel metadata operations, making them perfect for high-frequency file creation, access pattern analysis, and rapid AI job cycles. V12's Elastic Metadata Engine extends this further with dynamic scaling across nodes.
- **VPODs:** 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.

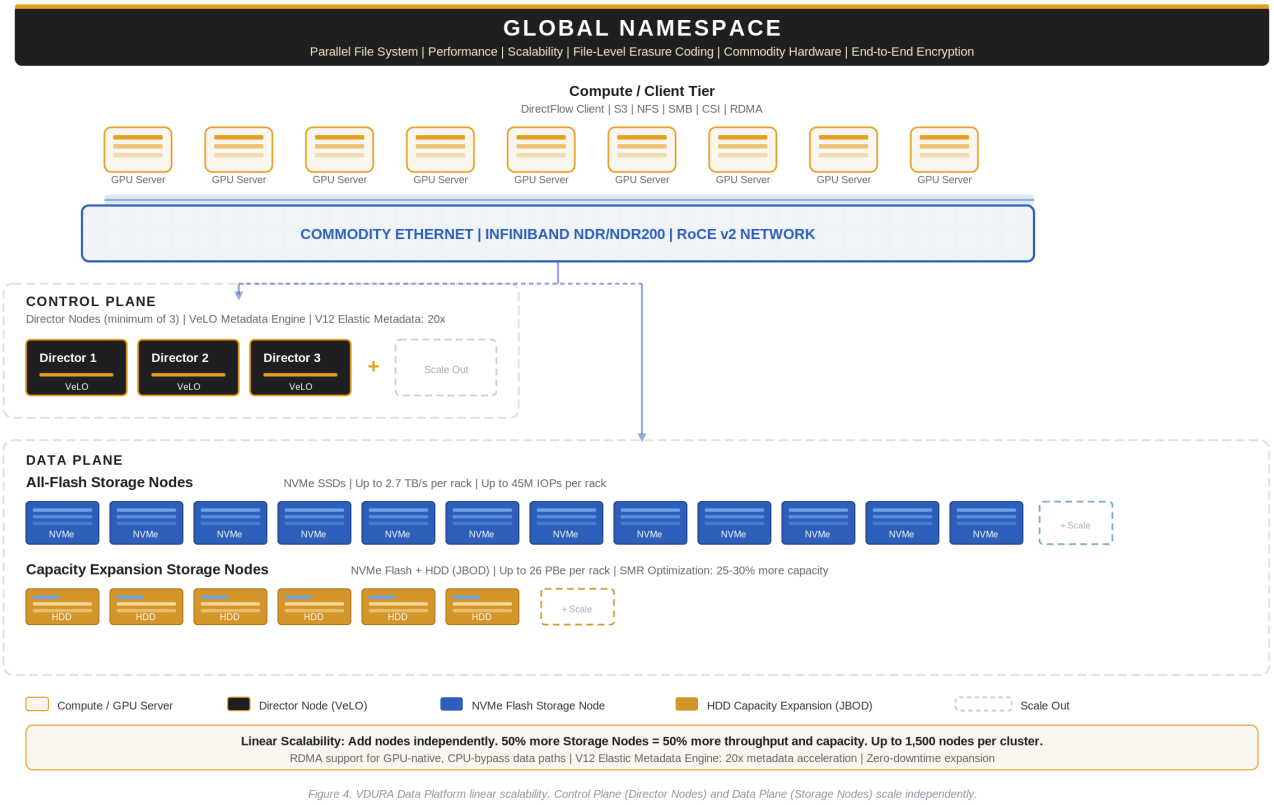

Figure 4. VDURA Data Platform linear scalability. Control Plane (Director Nodes) and Data Plane (Storage Nodes) scale independently.

Figure 7. Linear scalability of Director and Storage Nodes.

Director Nodes: Orchestrating Control in a Parallel World

Director Nodes serve as the brain in the VDURA architecture. VDURA separates the control plane, which handles metadata, orchestration, and policy, from the data plane, which handles user I/O. As the control plane's core, Director Nodes 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 flash-optimized 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. VeLO accelerates everything from tiny files to checkpoint indices to model versions. V12's Elastic Metadata Engine dynamically scales metadata capacity across nodes, delivering up to 20× improvement in operations.

Director Nodes form the authoritative layer of VDURA's control structure and every deployment requires a minimum of three. Administrators configure either three or five of the total Director Nodes as a replication set, or “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.

Storage Nodes: AI Pipeline Performance from Every Layer

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.

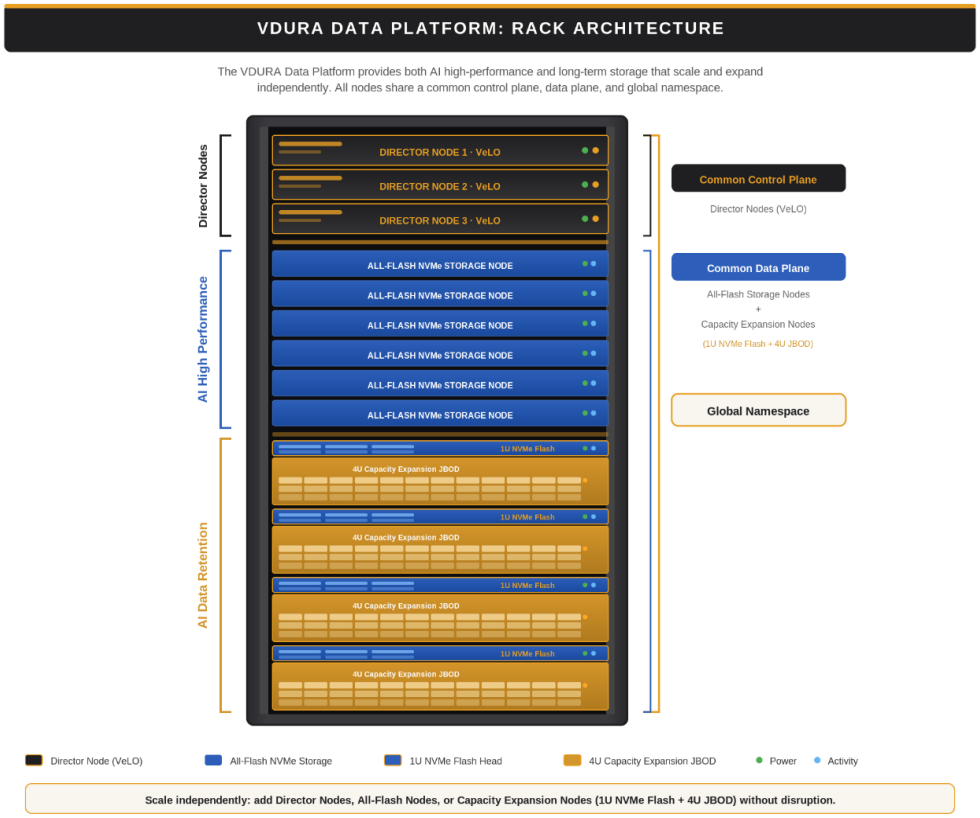

Figure 5. VDURA Data Platform rack architecture. Capacity Expansion Nodes pair a 1U NVMe Flash server with a 4U Capacity Expansion JBOD enclosure.

Figure 8. One control plane, one data plane, one 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. Ideal for archived model weights, retraining inputs, inference logs, and data lakes that need fast access but are less frequently touched. V12's SMR HDD Optimization unlocks 25–30% more capacity per rack without compromising throughput.

VPOD Architecture: Virtualization for Resilience and Efficiency

Rather than treating the entire server as a single failure domain, each VDURA Storage Node hosts multiple Virtualized Object Storage Devices, or VPODs. 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.

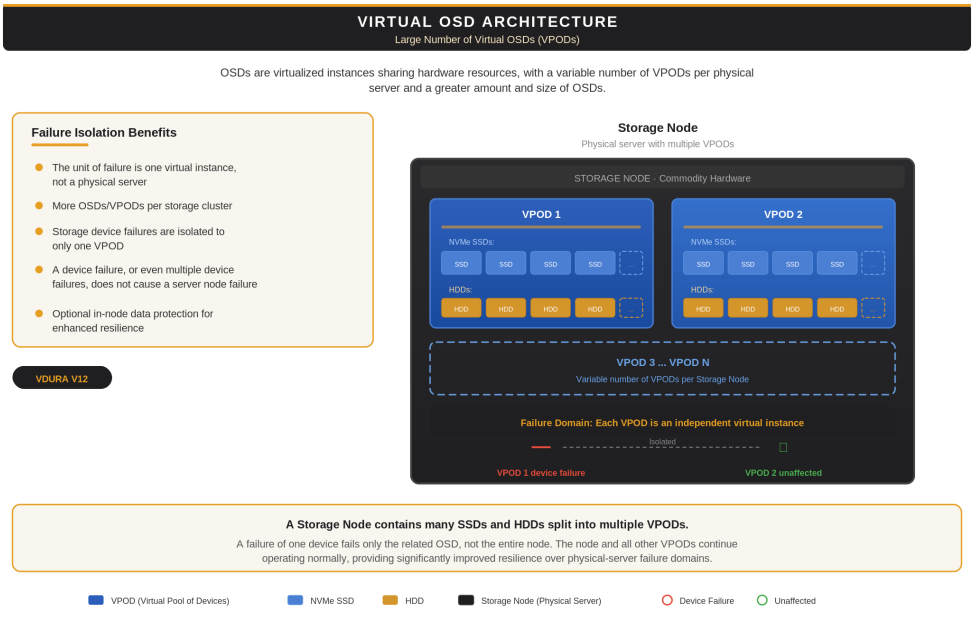

Figure 6. Virtual OSD architecture. VPODs provide failure isolation at the virtual instance level, not the physical server level.

Figure 9. VPOD architecture provides failure isolation at the virtual instance level, not the physical server level.

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.
- **Metadata SSDs:** 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 and 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.
- **Active Capacity Balancing** to eliminate hotspots and evenly distribute load.
- **Real-time 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.

Multi-Tenancy, API Control, and Observability

Neoclouds and AI factories rarely serve a single workload. Storage must operate as a service: many tenants, hard isolation guarantees, API-driven provisioning, and clear operational visibility. V12 treats these as architectural requirements rather than add-ons.

- **Multi-tenant isolation by design.** Per-tenant QoS, namespaces, end-to-end encryption keys, and VLAN isolation are first-class capabilities, exposed through native REST APIs and a Kubernetes CSI plug-in, so providers can onboard tenants on day one without bolt-on tooling.
- **Next-generation control plane.** Announced at ISC 2026 and planned for general availability in the second half of 2026 for all V5000-class systems, VDURA’s next-generation control plane introduces a streamlined tenant administration model: platform and tenant operators manage complex multi-tenant environments from an intuitive dashboard, with an in-place online upgrade path for existing customers.
- **API-first automation.** A REST API for key platform operations — provisioning, quotas, snapshots, and tenant lifecycle — provides the foundation for automation and tooling workflows, with Terraform, Ansible, and Crossplane providers for infrastructure-as-code deployment.

- **Observability and diagnostics.** VDURA provides a high level of observability into performance and resiliency: real-time system health, capacity balancing telemetry, background data-scrubbing status, and the precise scope of any event, complemented by the MyVDURA portal for fleet-level insight.
- **S3 for cloud-native pipelines.** Targeted S3 performance improvements sustain throughput for model checkpointing, inference serving, and large-scale dataset ingestion, while native S3 object tagging (planned for H2 2026) unlocks policy-based lifecycle management, automated tiering, and fine-grained access control across training datasets and model artifacts.

What's New in V12

VDURA Data Platform V12 represents a major release with significant advancements across metadata performance, data management, storage economics, and GPU-native connectivity. V12 delivers more than 20% increase in throughput, 20× metadata acceleration, and over 20% cost-per-TB reductions, all available as a zero-downtime in-place upgrade for V11 customers.

20%+	20×	20%+	25–30%
Throughput Increase	Metadata Acceleration	Cost-per-TB Reduction	More Capacity (SMR)

Elastic Metadata Engine

The V12 Elastic Metadata Engine dynamically scales metadata capacity across Director Nodes, delivering up to 20× improvement in metadata operations. It supports billions of files and objects under active use, eliminating metadata bottlenecks that have traditionally constrained AI pipelines at scale. The engine automatically rebalances metadata distribution as clusters grow, ensuring consistent performance regardless of namespace size.

RDMA Support for GPU-Native I/O

Available now for all V5000 systems, RDMA support enables GPU-to-storage data transfers that bypass the CPU entirely. Built on AMD Pensando™ Pollara 400G NICs and AMD EPYC™ Turin processors, RDMA delivers direct memory access between AMD Instinct GPU server nodes and the VDURA Data Platform, eliminating CPU bottlenecks for the low-latency, high-throughput data paths critical to AI training and inference.

Snapshot Support

V12 introduces native snapshot support with instantaneous, space-efficient, point-in-time copies. Designed for AI pipeline checkpoints, model snapshots, and operational recovery, snapshots can be created manually or via policy-based retention. This capability is essential for protecting training progress, enabling rapid rollback during model development, and maintaining data integrity across complex AI workflows.

SMR HDD Optimization

A new write-placement engine in V12 organizes sequential zones intelligently for Shingled Magnetic Recording (SMR) drives, unlocking 25–30% more capacity per rack without compromising throughput.

Context-Aware Tiering (Phase 1)

Phase 1 of Context-Aware Tiering introduces three capabilities: Extended DirectFlow Buffer to Local SSD, reducing dependency on network storage for hot data; KVCache Writeback for Persistence SLA, minimizing unnecessary I/O while maintaining inference SLA compliance; and Context Cache Tiering Framework for high-speed read/write at LMCache speed, supporting long-context LLM serving and RAG workloads. The roadmap includes deeper application-directed data placement, cross-node cache coherence, and DPU support.

Native CSI Plug-in

V12 includes a native Container Storage Interface (CSI) plug-in that simplifies multitenant, Kubernetes-based deployments with zero-script persistent-volume provisioning and management. Organizations running containerized AI pipelines on Kubernetes can now provision VDURA storage volumes directly through standard Kubernetes APIs, eliminating custom integration work and accelerating time-to-production for cloud-native AI workloads.

End-to-End Encryption

VDURA provides industry-leading end-to-end encryption. V12 delivers comprehensive security with transparent, tenant-per-volume AES-256 encryption that protects data from the moment it leaves the client, through transit, and at rest. This unified encryption architecture replaces the patchwork of TLS for in-flight and SED for at-rest that competitors rely on, providing stronger confidentiality, integrity, and compliance alignment in a single, zero-performance-compromise implementation.

VDURACare Premier

One simple contract covering hardware, software, and support. VDURACare Premier™ includes 10-year, no-cost replacement of drives and 24×7 expert response, delivering comprehensive, risk-free coverage that protects your investment and keeps your AI factory running without interruption.

V12 is available as a zero-downtime in-place upgrade for all V11 customers and reaches general availability in Q2 2026 for all V5000 systems.

VDURA V5000 Certified Platform Hardware

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

V5000 hardware runs the VDURA Data Platform V12, 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.

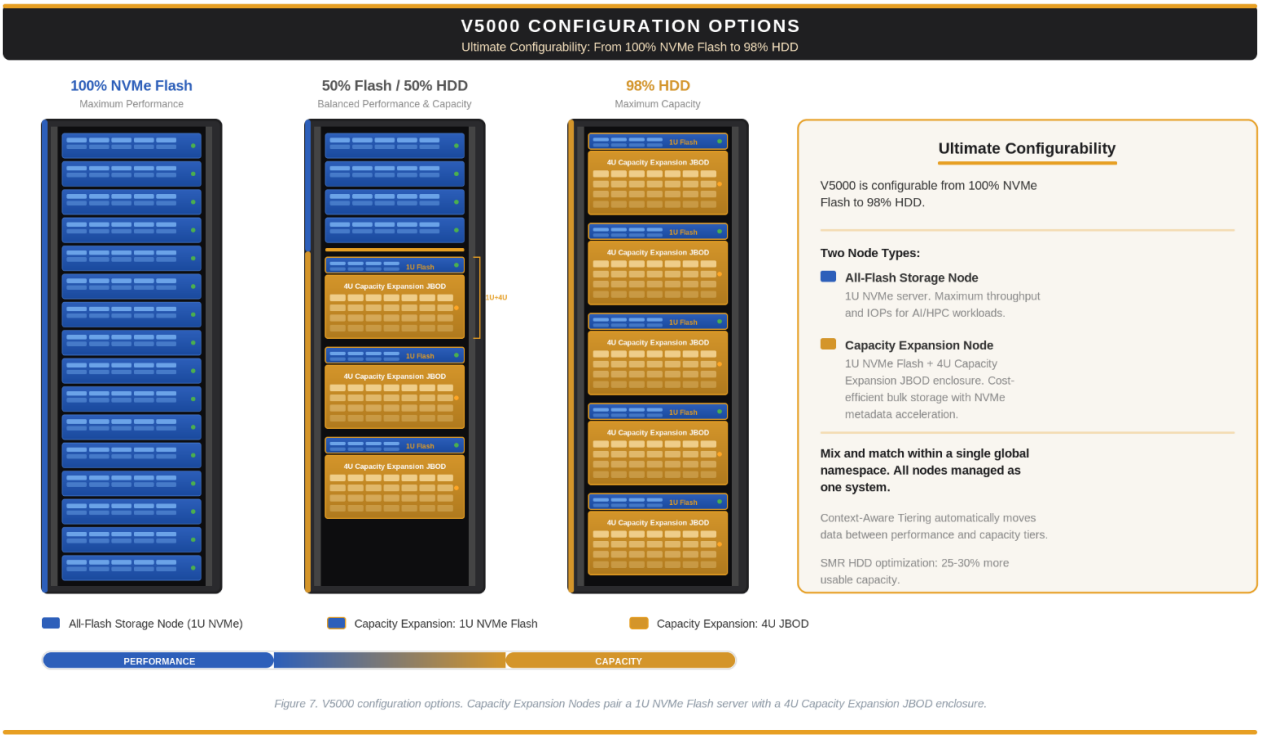

Figure 7. V5000 configuration options. Capacity Expansion Nodes pair a 1U NVMe Flash server with a 4U Capacity Expansion JBOD enclosure.

Figure 10. V5000 configuration options, left to right: 100% NVMe flash, 50% flash / 50% HDD, 98% HDD.

Key Features

- 1U chassis for Director Nodes and All-Flash Storage Nodes.
- 4U chassis for HDD capacity expansion.
- Configurable as either all-flash NVMe or mixed-fleet nodes using flash NVMe with HDD capacity expansion.
- **Up to 2.7 TB/s throughput per rack** in all-flash configurations.
- **Up to 200 GB/s throughput per rack** in flash with HDD capacity expansion configurations.
- **Up to 1.2M IOPS per Storage Node**, supporting low-latency, high-frequency workloads. Up to 45M IOPS per rack with all-flash configuration.
- **Multi-Level Erasure Coding** with up to 12 nines durability (all-flash) or 11 nines durability (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.
- **RDMA support** for GPU-native, CPU-bypass data paths (new in V12).
- **Snapshot support** for AI pipeline checkpoints and operational recovery (new in V12).
- **SMR HDD optimization** unlocking 25–30% additional capacity per rack (new in V12).
- **Native CSI plug-in** for zero-script Kubernetes persistent-volume provisioning (new in V12).
- **Native multi-tenancy** with per-tenant QoS, namespaces, encryption keys, and VLAN isolation (new in V12).

- **REST API control plane** with Terraform, Ansible, and Crossplane providers (new in V12).
- **End-to-end AES-256 encryption** with tenant-per-volume granularity.
- **VDURACare Premier:** 10-year drive replacement, 24×7 expert support, one contract.

V5000 Details

The VDURA V5000 Certified Platform hardware system represents the culmination of decades of engineering expertise in parallel file systems and distributed storage technology. Built for AI/ML and HPC workloads, V5000 hardware 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.

Director Node

- Hosts the VeLO Metadata Engine optimized for flash. V12 Elastic Metadata Engine delivers up to 20× acceleration.
- Contains 12 × NVMe SSD slots.
- Up to 333M inodes and up to 225K 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 Node

- Configurable with 12× NVMe SSDs in 7.68 TB, 15.36 TB, 30.72 TB, 61.44 TB, or 122.88 TB capacities.
- Delivers up to 60 GB/s per node and up to 2.7 TB/s throughput per rack.
- Supports up to 1.2M 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 Node

- 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, or 32 TB drive options.
- Up to 200 GB/s flash-accelerated throughput per rack.
- Scalable to 26 PBe per rack (effective) with inline compression.
- Optimized for high-capacity / data lake storage with flash-accelerated performance.
- Supports dynamic tiering and intelligent placement via the VDURA Orchestration Engine™.
- V12 SMR HDD Optimization unlocks 25–30% additional capacity per rack.

Connectivity

- Supports InfiniBand NDR/NDR200 and Ethernet 400/200/100 GbE ports.
- Nodes connect at up to 2× NDR200 InfiniBand and up to 4× 100 GbE Ethernet.

V5000 Expansion Options

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

- Add Director Nodes to increase metadata and protocol performance.
- Add all-flash Storage Nodes for higher throughput and IOPS.
- Add 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 StorageSets, ensuring isolation and quality-of-service (QoS).

Physical Connectivity

VDURA V5000 Director and Storage Nodes support 400/200/100 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× 200/100 GbE configuration using two 200/100 GbE SFP28 cables, with one attached to each port. 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 LACP mode, the physical ports are bonded with the IEEE 802.3ad LACP link-layer protocol, providing load balancing, better fault tolerance, and protection against misconfiguration.

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. Single link mode should be used with caution.

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: StorageSets and volumes.

StorageSets. The StorageSet 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 StorageSet by adding more hardware, and you can move data within a StorageSet.

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 StorageSet, with up to 1,200 per realm. V12 adds native snapshot support for volumes, enabling instantaneous point-in-time copies for AI pipeline checkpoints and operational recovery.

Erasure Coding and Data Protection

Storage Nodes in the VDURA Data Platform are highly sophisticated Virtualized Protected Object Devices (VPODs); you gain the same scale-out and shared-nothing architectural benefits from our VPODs as any object store would.

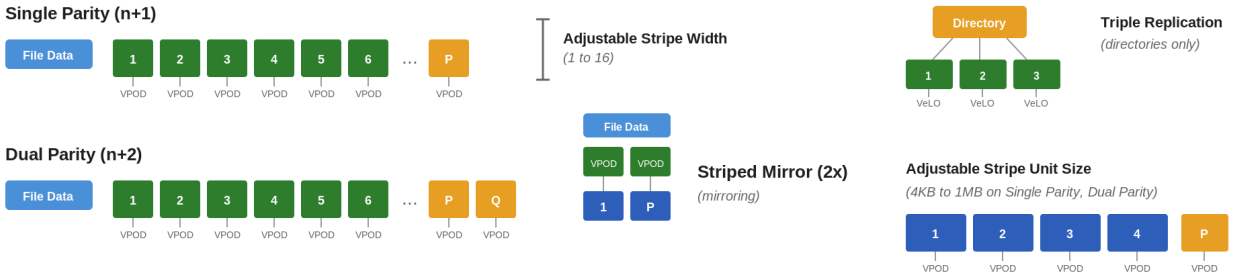

Figure 8. Per-file erasure coding protection layouts.

Figure 11. Per-file erasure coding layouts.

The VDURA Data Platform stripes each large POSIX file across a set of VPODs and adds parity to that stripe using an N+2 erasure coding scheme. Striping every file across multiple VPODs delivers the parallelism behind VDURA's performance, while per-file erasure coding lets administrators match protection to the workload. Rather than a one-size-fits-all approach, VDURA offers a choice of per-volume protection levels — Dual Parity (n+2), Single Parity (n+1), and Striped Mirror (2x) — described in Per-File Erasure Coding below.

> ***A traditional storage 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.

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 a traditional storage system, since the odds of any given drive 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 drive failing while the system is still degraded. If enough drives 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.

> ***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.***

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

When a StorageSet is first set up, it sets aside a configurable amount of spare space on all the Storage Nodes in that StorageSet 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 StorageSet. As a result, during a reconstruction, the system uses the combined write bandwidth of all the Storage Nodes in that StorageSet. 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 StorageSet 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 StorageSet.

If, in extreme circumstances, three Storage Nodes in a single StorageSet were to fail at the same time, VDURA has additional lines of defense that limit the effects of that failure. All directories are independently stored triplicated, with three complete copies of each directory and no two copies on the same Director Node. Should a third Storage Node fail while two others were being reconstructed, the StorageSet would transition to a read-only state to protect data integrity, and access would be affected only for the small subset of files with VPODs on all three failed nodes — all other files remain available or fully recoverable through their erasure coding, and the proportion of affected files shrinks as the StorageSet grows. For environments with stringent SLAs, additional resiliency can be layered in — higher protection levels, added spare capacity, and expanded Director redundancy — to meet the most demanding availability targets.

VDURA provides a high level of observability and diagnostics for performance and resiliency, giving operators clear, real-time insight into system health and the precise scope of any event — enabling confident operation at scale.

Per-File Erasure Coding

Instead of relying on hardware controllers that protect data at a drive level, VDURA architecture uses per-file distributed erasure coding in software. Files in the same StorageSet, volume, and even directory can have different erasure coding protection levels. In this way, a file can be seen as a single virtual object that is sliced into multiple component objects.

Users have three Erasure Coding Protection Levels available: Dual Parity (n+2), Single Parity (n+1), and Striped Mirror (2x).

- **Dual Parity (n+2)** is the default protection level for all volumes, as it provides a balance between capacity overhead and performance. This is the recommended setting for most workloads.
- **Single Parity (n+1)** is optimized for performance with less capacity overhead, ideal for workloads where throughput is prioritized over maximum fault tolerance.
- **Striped Mirror (2x)** is optimized for small write performance, combining striping and mirroring to provide high performance for applications that perform small random writes while providing resiliency against a single disk failure.

The Erasure Coding Protection Level is selected at volume creation time and cannot be changed after a volume is created. You can mix protection levels and any volume layout together in the same StorageSet, with each volume evaluated independently for availability status.

System Limits and Requirements

General limits and requirements of the VDURA Data Platform V12 and V5000 storage system are provided in Table 4.

Limit	Description
Minimum V5000 system	3 Director Nodes with 3 All-Flash or Capacity Expansion Storage Nodes
Largest V5000 system	No enforced architectural limit; production deployment to date +1,500 nodes per cluster
Maximum number of files or directories	No enforced limit
Maximum file size	CIFS: 32 TB | NFS: 32 TB | 64-bit Linux DirectFlow: 8,192 PB
Maximum number of volumes	Per Director: no limit Per StorageSet: no limit Per realm: 3,600

Limit	Description
Maximum number of Realm Managers in repset	5
Maximum StorageSet 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 StorageSet can be scheduled to have a snapshot taken at the same time; these will be batched and optimized.
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	4,096 bytes

Table 4. VDURA Data Platform V12 and V5000 system limits and requirements.

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.

V12 introduces VDURACare Premier, one simple contract covering hardware, software, and support. VDURACare Premier includes 10-year, no-cost replacement of drives and 24×7 expert response, delivering comprehensive, risk-free coverage that protects your investment and keeps your AI factory running without interruption. No surprise costs, no separate maintenance contracts, no finger-pointing between vendors.

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 hardware disposal. The VDURA Support Services team brings many years of operational experience in highly secure environments and understands how to optimize secure installations.

Where Velocity Meets Durability

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

Design, implementation, and support of AI and HPC infrastructure are all often complicated and confusing tasks, and they benefit enormously from a turnkey approach. To that end, the AI Storage Reference Design in this document provides a tested solution 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.

Built on the HYDRA architecture, the VDURA Data Platform V12 unifies a true parallel file system and a resilient object store under one data plane, one control plane, and one global namespace. It is the same hyperscaler-grade model that powers Google, Meta, and Microsoft, productized for every AI factory, Neocloud, and enterprise running AMD Instinct GPUs.

Beyond performance, V12 ships with the data services AI infrastructure teams expect: native snapshots for AI pipeline checkpoints and policy-based retention, quotas and multi-tenant isolation, Dynamic Data Acceleration across performance tiers, RDMA for GPU-native data paths, and end-to-end AES-256 encryption with KMIP key management. One vendor, one stack, one upgrade path, all backed by VDURACare Premier with 10-year drive replacement and 24×7 expert support.

From ingest and training to checkpointing and inference, the VDURA Data Platform V12 and AMD Instinct MI3XX Series reference architecture powers every stage of the AI pipeline. Simple to deploy. Easy to operate. Effortless to scale.

> ***VDURA: Where Velocity Meets Durability.***

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

LEARN MORE · **www.vdura.com**
Worldwide +1-888-726-2727 · info@vdura.com