White Paper

Shared-Nothing Architecture for AI

How Hitachi Virtual Storage Platform One SDS Transforms Hybrid Cloud and Block Workloads

Hitachi Vantara

White Paper

Table of Contents

The New Demands on Enterprise Storage

AI is transforming how data is stored and accessed, making scale, contention and failure impact more critical than ever. Enterprises want to avoid costly hardware and complex reconfiguration, yet AI-driven workloads and hybrid cloud environments push storage to its limits.

Data-intensive applications such as model training, real-time analytics and distributed databases require massive parallel access and elastic scaling across on-premises and cloud environments. Traditional architectures struggle to deliver the scale, predictability and resilience these environments demand.

The shift from sequential OLTP to parallel data processing exposes fundamental weaknesses in shared-everything designs. As AI investments grow, storage becomes the critical path to unlocking business value.

80%

"Data growth will accelerate to 181 zettabytes by 2025, with 80% unstructured."

IDC's Data Age 2025 white paper

Why Current Approaches Fall Short

AI and DevOps workloads overwhelm shared-everything architectures with contention, latency spikes and scale limits — even as costs rise. Adding capacity means bigger arrays and higher risk, not the elastic, node-by-node scaling customers expect.

Centralized controllers create bottlenecks, leading to unpredictable performance, SLA risks, large failure domains and costly redundancy. Forklift upgrades and overprovisioning inflate CapEx and OpEx, clashing with rapid iteration.

Modern applications demand parallelism and resource efficiency. Shared-nothing architecture delivers this by eliminating centralized contention and enabling each node to operate independently. Every node owns its compute and storage stack, serving data locally while coordination happens over the network — removing choke points and metadata bottlenecks.

The result: linear scale, predictable performance and fault isolation. Adding nodes increases capacity and throughput proportionally, while failures remain contained. Cloud-native workloads — microservices, analytics engines, AI pipelines — benefit from rapid scale-out, low latency and graceful degradation.

Aligned with cloud economics, shared-nothing architecture runs on commodity hardware, improving price-performance and avoiding vendor lock-in. It's not just an optimization; it's essential for environments where scale, resilience and cost efficiency are non-negotiable.

30-50%

"Enterprises waste 30-50% of storage capacity due to overprovisioning for unpredictable AI workloads."

Forrester®. Predictions 2026:
Private AI on Private Clouds, and the Rise of the Neoclouds

Shared Nothing Versus Shared Everything

Modern applications are built to use all available resources through parallelism. To be cost-effective at scale, they must avoid bottlenecks that serialize operations or centralize state. Shared-nothing architecture meets this need by eliminating centralized contention points and letting each node operate independently.

In a shared-nothing design, every node owns its full stack — CPU, memory, cache, flash and disk — and serves data locally. Coordination, replication and rebalancing occur over the network without forcing I/O through a single controller or metadata tier. This removes choke points common in shared disks or centralized metadata systems.

The result is **linear, horizontal scale** and **predictable performance.** As you add nodes, both compute and capacity grow proportionally, and throughput stays consistent under concurrency. Failures are fault-isolated to affected nodes or partitions rather than impacting the whole cluster.

Cloud-native workloads — API-driven systems, microservices, serverless functions, analytics engines and AI pipelines — demand rapid scale-out, low-latency access and graceful degradation. Shared-nothing design supports these expectations by enabling node additions or removals with minimal rebalancing and without pushing all operations through a shared layer.

Aligned with cloud economics, shared-nothing architecture runs on commodity x86 hardware instead of specialized arrays, improving price-performance while avoiding vendor-specific bottlenecks. ***It is not just an optimization; it is a prerequisite where scale, resilience and cost efficiency are first-class requirements.***

Architecture Comparison

Aspect	Shared-Everything Architecture	Shared-Nothing Architecture
Scaling	Vertical, controller-limited	Horizontal, linear-node addition
Failure Impact	Cluster-wide disruption	Single-node isolation
Cost Model	Expensive shared arrays	Commodity hardware + software-defined storage (SDS) (Validated via Hitachi Self-Certification App)
Best For	OLTP, small clusters	AI, databases, hybrid cloud

Note: For AI/DevOps, shared-nothing architecture avoids the controller contention that drives tail latency, enabling consistent throughput under heavy parallelism.

When paired with cloud-native deployment models, shared-nothing design removes the dependency on long infrastructure lead times: It allows applications to scale horizontally on demand, rather than waiting for centralized storage or controller upgrades. This unlocks a faster innovation cycle. Teams can iterate, test and deploy at cloud speed without being bottlenecked by monolithic storage layers that can't keep pace with the elasticity of modern pipelines.

> *"Shared-nothing scales linearly and has good price/performance compared to shared memory or shared disk systems."*
>
> **Michael Stonebraker (Turing Award winner, Shared-Nothing originator):**
> The Case for Shared Nothing

Use Cases for Shared-Nothing Architecture

Hybrid Cloud Autonomy

Hybrid cloud environments need flexibility without complexity. Shared-nothing architecture makes this possible by keeping data operations local for speed while synchronizing across sites and clouds for resilience. Policies ensure data stays where it should — meeting compliance requirements like GDPR or HIPAA — while security and retention settings move with the data.

This approach allows workloads to shift easily between on-premises and cloud platforms without being tied to specific hardware, delivering consistent performance and visibility everywhere. The result is scalable growth and reliable operations without the bottlenecks of traditional architectures.

Common SLAs

- **99.95% monthly uptime** for each storage region (allows 22 minutes downtime/month).
- **Data integrity** ≥99.999% annual durability, ensuring highly resilient storage with minimal risk of data loss.

Distributed Data Platforms

Modern data architectures — NoSQL, NewSQL, streaming platforms and data lakehouses — require shared-aware storage that aligns with distributed processing. Shared-nothing designs eliminate cross-node coordination overhead that slows shared-everything systems.

Platforms like Apache Cassandra® and ScyllaDB™ expect local data ownership and minimal network traffic for faster reads without having the entire storage array supplying a read operation. Shared-nothing architecture meets these needs while adding enterprise features, such as automated rebalancing and unified management.

Equally important, shared-nothing architecture delivers predictable performance under mixed read/write workloads, maintaining consistent latency even as clusters scale. This performance is critical for OLTP-adjacent microservices and real-time applications that can't accept latency.

Common SLAs

- **High availability:** 99.95% or higher uptime for the storage cluster, guaranteeing continuous access for distributed databases and real-time applications.
- **Read/write latency:** remains below agreed thresholds (e.g., <10 ms) under mixed workloads, ensuring consistent performance for OLTP-adjacent microservices and real-time applications.

Predictable Scale and Resilience

Shared-nothing architecture stands out in growth scenarios and cloud-enabled architectures (hybrid architecture). Adding capacity also boosts compute power without controller upgrades or bottleneck planning, matching the elastic scale enterprises expect from cloud platforms.

Resilience comes from isolation: If a node fails, partitions automatically shift to healthy nodes while unaffected workloads keep running. Self-healing maintains service during maintenance or unexpected outages. This design limits the impact of failures, reduces downtime risk, safeguards SLAs and removes the need for costly standby systems — delivering reliable operations and lower costs.

Common SLAs

- **High availability:** 99.95% or higher uptime for the storage cluster, ensuring continuous access even during node failures or maintenance.
- **Resilience guarantee:** Automatic failover and partition reassignment within agreed recovery time objectives (e.g., <60 seconds), minimizing downtime and protecting SLAs. This allows applications to degrade gracefully and maintain service continuity, even when the surrounding cloud environment experiences variability. In a cloud-driven world, that resilience empowers teams to innovate at a rapid pace while confidently navigating occasional, unforeseen disruptions.

Building With Hitachi's VSP One SDS

Hitachi Virtual Storage Platform One SDS (VSP One SDS) turns shared-nothing architecture into enterprise reality. This software-defined platform scales across commodity hardware and hybrid cloud environments, delivering node independence without proprietary arrays.

To streamline deployment on third-party servers, Hitachi provides a self-certification app that validates your existing hardware against our performance and resilience standards. This ensures that the software-defined platform delivers the low latency and linear growth promised, regardless of the underlying x86 vendor.

Reads and writes process locally in DRAM and flash before asynchronous writing to permanent storage, ensuring low latency. Hitachi polyphase erasure coding (HPEC) replaces RAID with distributed redundancy that grows linearly with the cluster. Automated data placement and QoS prevent hotspots and guarantee workload isolation, while API-driven provisioning supports DevOps automation and hybrid deployments without lock-in.

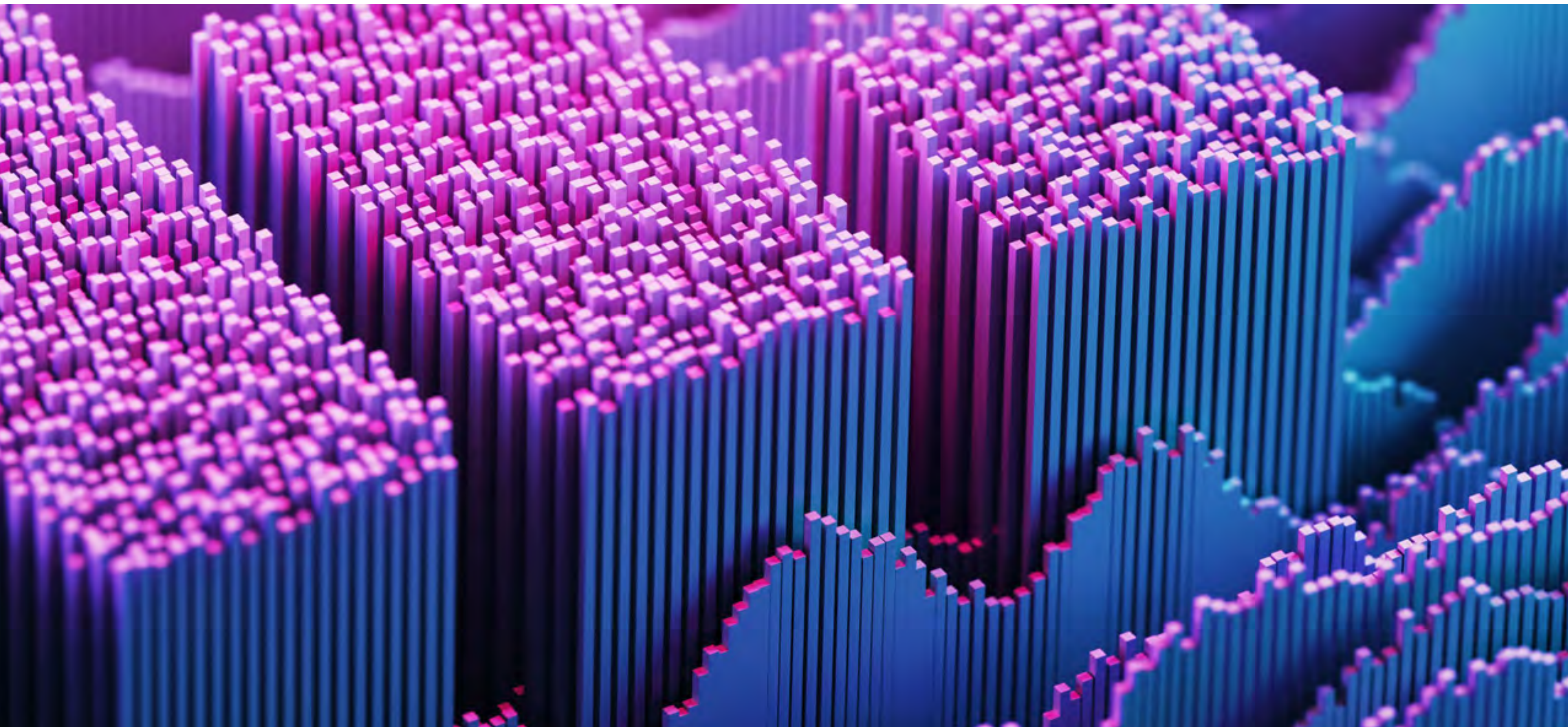
Unlike disaggregated all-flash platforms, VSP One SDS provides true shared-nothing block storage with enterprise-grade resilience across Amazon Web Services® (AWS®), Microsoft® Azure®, Google Cloud Platform® (GCP®), and on-premises environments. It supports mission-critical databases, VMware consolidation and multiple protocol options (protocols are selected per cluster, not mixed within the same cluster). Availability targets are designed for demanding production workloads, delivering cloud-like elasticity and the resiliency enterprises require.

VSP One Block Storage complements SDS deployments by providing ultra-low latency tiers for latency-sensitive databases and AI inference engines. These high-end arrays integrate seamlessly with the SDS control plane, creating tiered storage architectures that optimize both performance and economics.

By leveraging existing arrays, customers can replicate data to the cloud through built-in interoperability, avoiding costly forklift upgrades and reducing total cost of ownership (TCO). This approach ensures SDS acts as a true complement — extending hybrid cloud flexibility while maximizing prior investments.

Consumption models align shared-nothing economics with enterprise financial realities. OpEx-based scaling eliminates CapEx surprises while pay-per-use pricing matches unpredictable AI workload patterns. Forrester TEI analysis confirms 285% ROI with seven-month payback periods.

Hitachi Virtual Storage Platform 360 (VSP 360) management software provides unified observability across SDS, block and hybrid deployments. Its real-time analytics enable capacity right-sizing while AIOps predict maintenance needs before they impact SLAs. And its compliance reporting and security dashboards complete the enterprise management requirements.



Final Thoughts: A Pragmatic Path to Modern Infrastructure

Shared-nothing architecture resolves the fundamental tension between storage scalability, resilience and cost that blocks AI and hybrid cloud success. With VSP One SDS, Hitachi delivers this proven pattern through software-defined innovation backed by decades of enterprise storage leadership. VSP One SDS gives our customers a true hybrid cloud advantage to bridge the traditional applications running their business and provide paths to modern application adoption and innovation.

The integrated portfolio of VSP One SDS Block storage, Hitachi EverFlex Consumption service level option and VSP 360 management software provides enterprises the pragmatic migration path to a shared-nothing architecture. The solution requires no operational disruption or forklift upgrades as it reduces business risk and ensures predictable performance under parallel workloads.

Learn more about VSP One and VSP One SDS

Learn more →

Competitive Advantages

- True shared-nothing block architecture for AI and databases (not just parallel file); modern workload support for Kubernetes® and cloud applications.
- Linear, node-by-node scaling of performance and capacity with deterministic latency.
- Hitachi polyphase erasure coding (HPEC) for efficient, predictable resilience at scale.
- Strong VMware integration for consolidation and hybrid expansion.
- Disaster recovery (DR)/business continuity planning (BCP) with automated failover and policy-driven replication across clouds.
- OpEx and VSP 360 observability for financial and operational control.

Hitachi Vantara

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
hitachivantara.com | community.hitachivantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
hitachivantara.com/contact

HV-BTD-WP-Shared-Nothing-Storage-for-AI-21Jan26-A