HITACHI

Hitachi Vantara

# NVIDIA Cloud Partner High-Performance Storage with Hitachi Content Software for File for NVIDIA HGX Servers with H200, B200, and B300 GPUs

Reference Design

MK-SL-407-00
December 2025

# Feedback

Hitachi Vantara welcomes your feedback. Please share your thoughts by sending an email message to Docs-Feedback@hitachivantara.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

**Revision history**

| Changes | Date |
| --- | --- |
| Initial release | December 2025 |

# Reference Design

The NVIDIA® Cloud Partner Reference Design (NCP RD) for AI training and inference, featuring NVIDIA HGX servers with H200, B200, and B300 GPUs, is the next generation of data center architecture for artificial intelligence (AI) for cloud service providers (CSPs). This document covers the high-performance storage (HPS) Reference Design for Hitachi Content Software for File (HCSF) from Hitachi Vantara.

HCSF powered by WekaIO with NVIDIA HGX servers, delivers a validated, turnkey AI infrastructure that aligns perfectly with the goals of customers pursuing an NCP model. This solution is designed to support flexible consumption, scalable performance, and industry-specific AI outcomes, making it ideal for organizations seeking to modernize their infrastructure without overcommitting capital upfront.

## Solution overview

Unlike conventional AI offerings, the Hitachi Vantara AI portfolio, Hitachi iQ, transcends basic integration and storage capabilities by layering enterprise and industry-specific AI outcomes within the AI solution. This approach ensures that outcomes are finely tuned to the unique needs and objectives of each organization.

At the heart of Hitachi iQ is the data platform, HCSF, a highly scalable, modern approach to the parallel file system. Designed for flexibility, HCSF fully integrates with Hitachi Vantara's Virtual Storage Platform One Object (VSP One Object) and supports on-premises, cloud-based, or hybrid environments. For more information on how VSP One Object extends Hitachi iQ with advanced data management, protection, and governance capabilities, see Appendix: HCSF integration with VSP One Object (on page 14).

HCSF supports a wide range of high-performance workloads, including Artificial Intelligence (AI), machine learning (ML), and advanced analytics. With the ability to scale beyond 10 exabytes in a single cluster, Hitachi Content Software for File is engineered to meet the massive and evolving data demands of modern AI-driven enterprises.

This solution guide covers the HPS Reference Design for NCP integration of HCSF with NVIDIA HGX servers.

# Storage systems

### Hitachi Content Software for File

The unique architecture of Hitachi Content Software for File is different from legacy storage systems, storage appliances, and hypervisor-based software-defined storage solutions as it not only overcomes traditional storage scaling and file sharing limitations but also allows unified file access via POSIX, NFS, SMB, S3 and NVIDIA GPUDirect® Storage. Hitachi Content Software for File provides a rich enterprise feature set, including local snapshot and remote snapshot offload, automated tiering, dynamic rebalancing, filesystem encryption, authentication integration, quotas, and much more.

Hitachi Content Software for File includes the following benefits:

- Highest performance across all IO profiles – ideal for mixed workloads, including heavy metadata operations
- Scalable capacity – start as small as 600 TB and scale to hundreds of petabytes in a single namespace
- Strong security – keep data safe from threats or rogue actors with both software and hardware encryption
- Hybrid Cloud – burst to all the major cloud providers for compute agility or run natively in the cloud
- Backup – offload backups straight to VSP One Object or any public cloud for long-term retention
- Best economics – combine flash and disk for the best cost at scale

What sets Hitachi iQ apart is Hitachi Content Software for File, a fully shared parallel file system that delivers the highest performance file services by leveraging NVMe flash. Also included is integrated tiering that seamlessly expands the filesystem to and from disk-based object storage, without the need for special data management software or complex scripts; all data resides in a single filesystem for easy access and management.

### Hardware specifications and capabilities

The following table lists the hardware specifications.

| Components | Version |
|---|---|
| Server Model | HCSF 31116 |
| Processor | AMD EPYC 9534 64-Core 2.45GHz Processor 256M 280W SP5 |
| Memory | 768 GB<br>12 x 64 GB DDR5 RDIMM 4800MTs Dual Rank |

| Components | Version |
|---|---|
| Storage | ▪ M.2 Boot<br>• 2x 960 GB<br>▪ Internal Storage<br>• 14x-16x NVME per node<br>• 7.68 TB, 15.36 TB, 30.72 TB |
| Network | ▪ 2 × NVIDIA ConnectX-7 1P 400 Gb/s OSFP/QSF112<br>▪ AIOM Dual Port 10G<br>▪ AIOM Dual Port 25G |

A single file system can support billions of directories and trillions of files, delivering a scalability model more akin to object stores than NAS systems, and directories scale with no loss in performance. Hitachi Content Software for File supports up to 1024 file systems and up to 24,000 snapshots in a single cluster.

- 6.4 trillion files or directories
- 14 Exabytes managed capacity with object store
- 6.4 billion files in a directory
- 4 petabytes for a single file

Supported networking technologies support NVIDIA Spectrum-X SN5600/5610 with 400 GbE speed.

**Software capabilities**

The following protocols are supported by the solution:

- POSIX Compliant Client
- NVIDIA GPUDirect® Storage (GDS)
- NFS (Network File System) v3 and v4.1
- SMB (Server Message Block) v2 and v3

**Multi-tenancy capabilities**

HCSF supports two distinct multi-tenancy deployment models that align with CSP operational requirements for isolation, performance, and governance. Within GPU-accelerated environments, multi-tenancy defines how multiple customers, departments, or workloads securely share common infrastructure while maintaining data separation and predictable performance. HCSF enables this through either logical cluster isolation (Organizations) or physical resource partitioning (Composable Clusters), allowing operators to select the model that best fits their deployment objectives.

- Organizations

  In the Organizations model, multi-tenancy is achieved through software-enforced logical boundaries within a single HCSF cluster. Each organization maintains independent users, filesystems, and administrative roles, managed through RBAC and tokenized access control. Capacity quotas can be applied per organization to enable chargeback and usage governance. This architecture is suitable for internal enterprise or departmental environments where governance, security, and management simplicity are prioritized, but strict performance isolation is not required.

- Composable Cluster

  The Composable Cluster model extends multi-tenancy to the infrastructure layer, deploying fully independent clusters across a shared infrastructure substrate via the Kubernetes Operator. Each tenant receives a dedicated allocation of compute, memory, and storage, creating a physically isolated data path with predictable throughput and latency – eliminating noisy-neighbor contention. This design meets the complete NCP requirements for resource isolation, deterministic performance, and independent lifecycle management, making it the recommended architecture for GPUaaS, AI Factory, and sovereign AI deployments requiring certifiable multi-tenant performance and security.

- Comparison table: Multi-tenancy models

| Feature | Organizations (Logical Model) | Composable Clusters (Physical Model) |
|---|---|---|
| Isolation Type | Logical/user/admin/data separation within one cluster | Full resource isolation via independent clusters on shared hardware |
| Hardware Sharing | All tenants share same hardware, drives, CPU, memory | Each tenant gets dedicated hosts/drives/cores/memory slices |
| Performance Predictability | Shared resources: performance may vary under load from other tenants | Dedicated resources<br>more deterministic performance and no “noisy neighbor” interference |
| Use Case Fit | Enterprise departmental or internal multi-user scenarios with governance focus | Service provider, GPUaaS, AI/HPC environments where resource isolation & QoS are critical |
| Administrative Domains | Single cluster, multiple organizations inside it; one Cluster Admin + Org Admins | Multiple independent clusters (one per tenant) each with its own management boundary |
| Limits and quotas | Up to 256 organizations per cluster | Per-cluster limits apply |
| Complexity and Management | Lower complexity: single infrastructure, simpler lifecycle | Higher complexity: multiple clusters, lifecycle automation (Kubernetes Operator) required |
| Certification Alignment | Meets logical isolation and governance requirements | Meets both isolation and performance/QoS requirements for high-end multi-tenant deployments |

**Quality of Service (QoS)**

HCSF supports mount-level Quality of Service (QoS) controls by leveraging client mount options such as `qos_preferred_throughput_mbps` and `qos_max_throughput_mbps`. The `qos_preferred_throughput_mbps` parameter specifies a soft target throughput (in MB/s) that the client will attempt to sustain under normal conditions, while the qos_max_throughput_mbps defines a hard average-based limit that the system enforces over time (bursting above the limit may be allowed briefly). By default, these values are unset that is, unlimited, but cluster administrators may define defaults from `cluster mount-default` commands from the storage management CLI. This will enforce all new mounts to inherit consistent QoS policies. Importantly, QoS limits apply at the frontend process level – all mounts sharing the same frontend process follow those QoS constraints, and each frontend connected to multiple clusters enforces its own limits independently.

# HPS Sizing

The storage performance target for training or inference can vary depending on the type of model and dataset. The guidelines in the following table provide standard throughput for the various GPU system sizes and HPS sizing. The final HPS requirements for throughput and capacity will be specified for each NCP opportunity.

The following table provides standard HPS aggregate storage performance details.

| Description | Number of GPUs | | | | | | |
|---|---|---|---|---|---|---|---|
| | **1,024** | **2,304** | **4096** | **8,192** | **16,384** | **29,952** | **41,472** |
| Read throughput (GB/s) | 160 | 360 | 640 | 1,280 | 2,560 | 4,680 | 6,480 |
| Write throughput (GB/s) | 80 | 180 | 320 | 640 | 1,280 | 2,340 | 3,240 |
| Storage Configuration | | | | | | | |
| Number of appliances | 8 | 8 | 8 | 16 | 32 | 59 | 81 |
| Number of 400G storage ports | 16 | 16 | 16 | 32 | 64 | 118 | 162 |
| Number of out-of-band management connections | 16 | 16 | 16 | 32 | 64 | 118 | 162 |
| Number of rack units | 8 | 8 | 8 | 16 | 32 | 59 | 81 |
| Power (KW) | 9.6 | 9.6 | 9.6 | 19.2 | 38.4 | 70.8 | 97.2 |
| Cooling (BTU/hr) | 35056 | 35056 | 35056 | 70112 | 140224 | 258538 | 354942 |

# Solution design

The solution with NVIDIA HGX servers and HCSF high performance storage is built entirely on the NCP Common networking based on NVIDIA Spectrum-X Ethernet-based architecture.

- Each HCSF storage node connects to two different leaf switches, part of Converged CPU/ Storage (N-S) network, using 2x single-port 400 GbE ConnectX-7 NICs in an active-active configuration for redundancy and high availability for high-throughput storage connectivity.
- For OOB management, 2x 1 GbE connections are required from each HCSF node for storage and BMC management ports, respectively.

The following table lists required transceivers and cables to connect HCSF nodes with SN5600/5610 switches. Use appropriate host side and switch side transceivers based on multi-mode or single-mode option and CX-7 cards (OSFP or QSFP112) used on HCSF nodes.

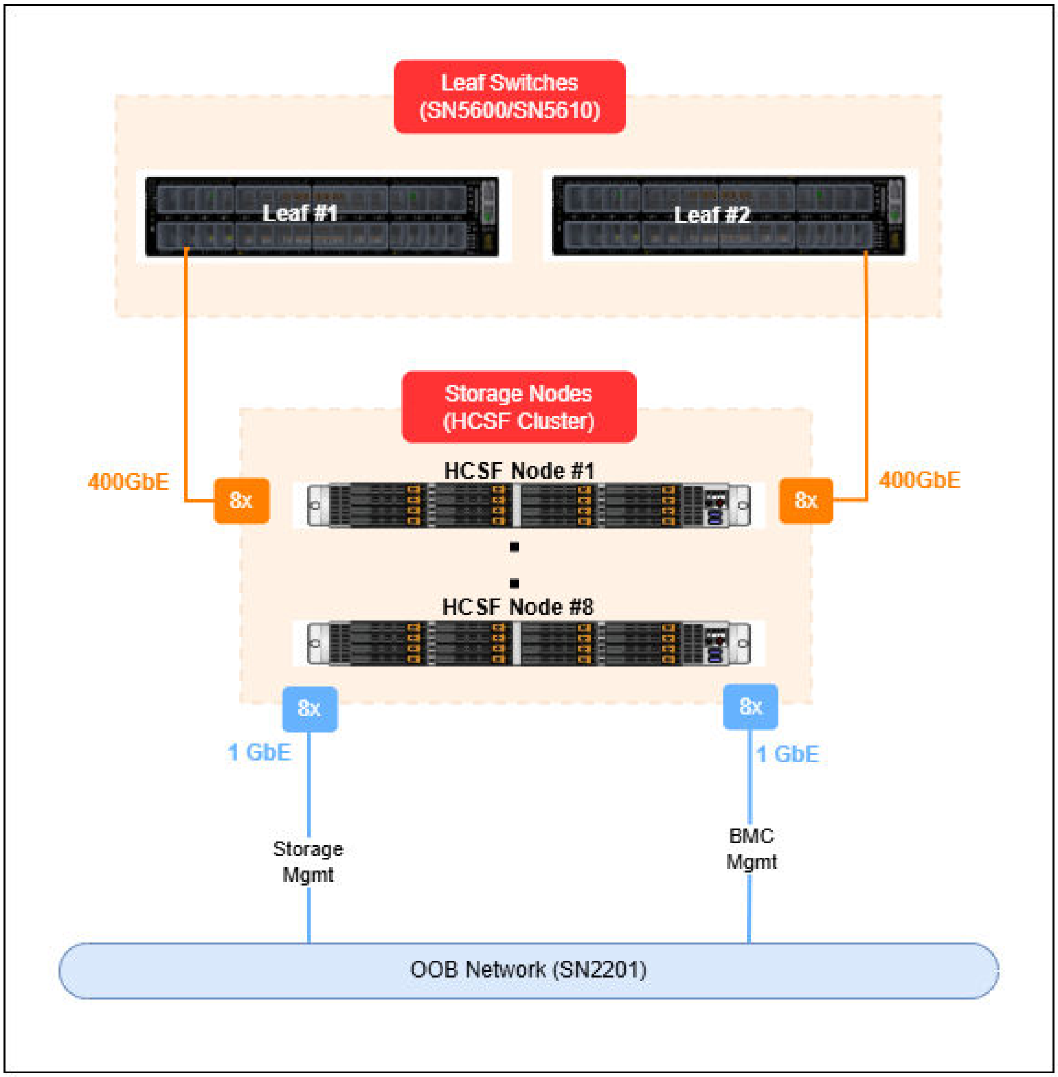
| Component | NVIDIA Part # | Description |
|---|---|---|
| Multi-Mode Connection option (up to 50m) | | |
| HCSF Host Transceiver (OSFP) | MA4Z00-NS400 | 400Gb/s, Single-port, OSFP, SR4 multimode parallel transceiver |
| HCSF Host Transceiver (QSFP112) | MMA1Z00-NS400 | 400Gb/s, Single-port, QSFP112, SR4 multimode parallel transceiver |
| SN5600/5610 Switch Transceiver | MMA4Z00-NS | Twin-port 800G 2xSR4 multimode transceiver |
| Cable | MFP7E10-N0XX | MPO-12/APC-to-MPO12/APC (8 fibers) passive optical Multimode cable<br>XX - 03 to 50 meters |
| Single-Mode Connection option (up to 100m) | | |
| HCSF Host Transceiver (OSFP) | MMS4X00-NS400 | 400Gb/s, Single-port, OSFP, DR4 single mode |
| HCSF Host Transceiver (QSFP112) | MMS1X00-NS400 | 400Gb/s, Single-port, QSFP112, DR4 single mode parallel transceiver |
| SN5600/5610 Switch Transceiver | MMS4X00-NS | Twin-port 800G 2xDR4 single mode transceiver |
| Cable | MFP7E30-N0xx | MPO-12/APC-to-MPO12/APC (8 fibers) passive optical single-mode cable xx - 01 to 100 meters |

See https://docs.nvidia.com/networking/display/400g100gpam4listpub for more information about the cables and transceivers highlighted in the previous table.

## NCP Deployments with 128 HGX systems (1,024 GPUs)

The following illustration is a Reference Design for NCP deployments with 128 HGX systems that need a minimum of HCSF cluster nodes for standard HPS aggregate storage performance requirements.

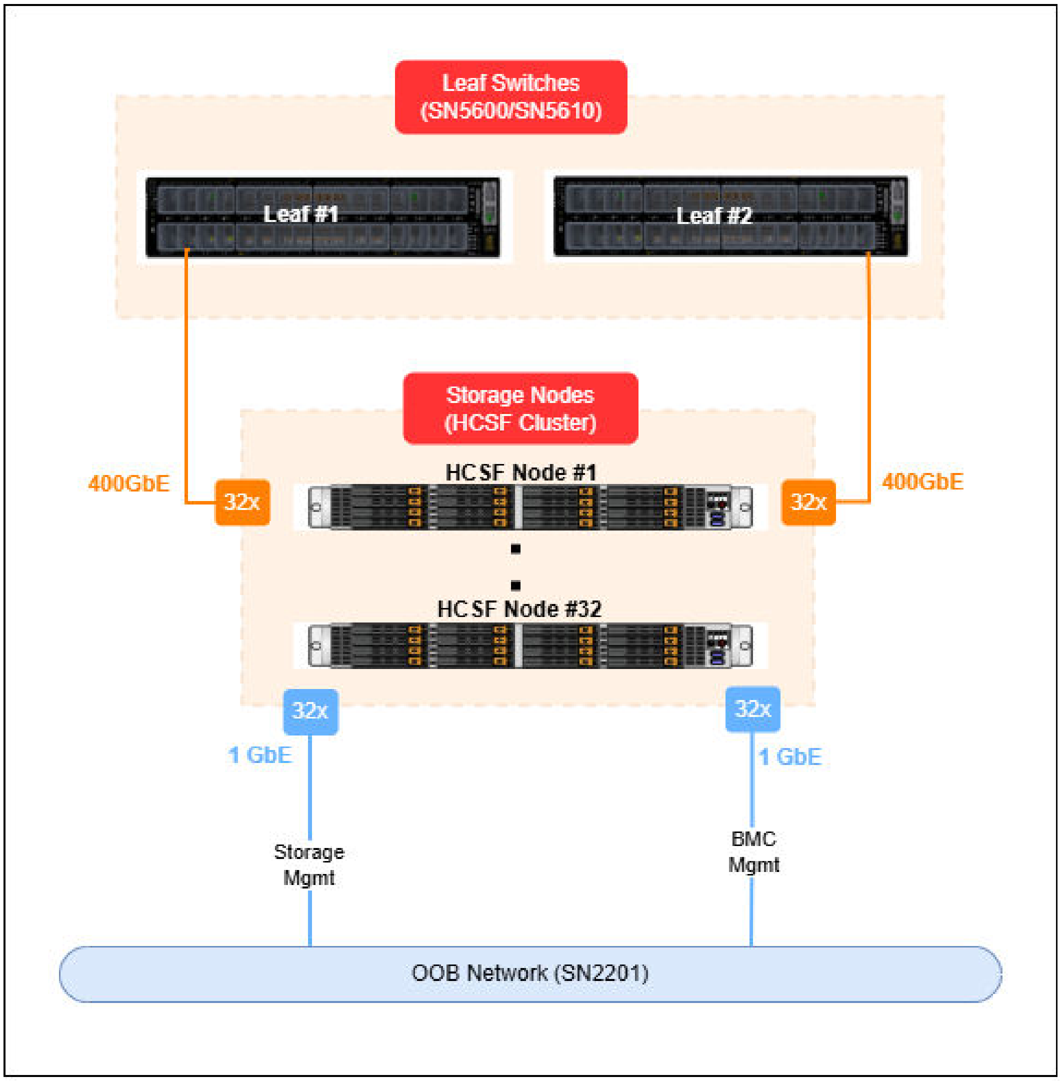
Two leaf switches are required to connect 16x 400 GbE connections from 8x HCSF nodes.




## NCP Deployments with 2048 HGX systems (16,384 GPUs)

The following illustration is a Reference Design for NCP deployments with 2048 HGX systems that need a minimum of 32 HCSF cluster nodes for standard HPS aggregate storage performance requirements.

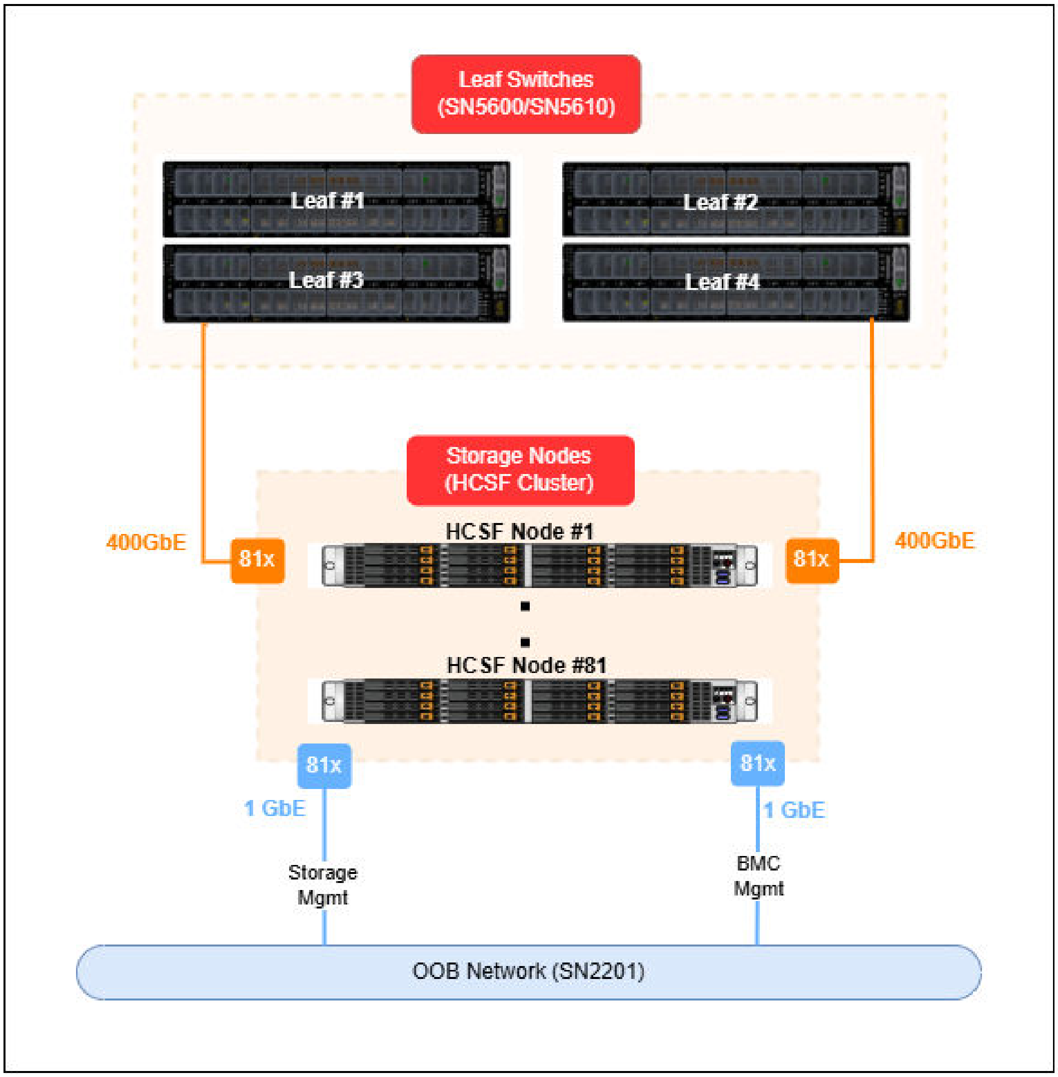
Two leaf switches are required to connect 64x 400 GbE connections from 32x HCSF nodes.




## NCP Deployments with 5184 HGX systems (41,472 GPUs)

The following illustration is a Reference Design for NCP deployments with 5184 HGX systems that need a minimum of 81 HCSF cluster nodes for standard HPS aggregate storage performance requirements.

Four leaf switches are required to connect 162x 400 GbE connections from 81x HCSF nodes.




# Storage partner solution performance validation

In close collaboration with NVIDIA, testing was performed using RoCE configuration on an 8-node HCSF cluster to validate storage performance and design, aligning with the NVIDIA NCP Reference Design and performance requirements, at scale.

As part of the NVIDIA-Certified Storage program, Hitachi Vantara HCSF 31116 storage achieved NCP certification for High Performance Storage.

# Conclusion

This Reference Design defines a new standard in AI infrastructure for cloud service providers. The Hitachi Vantara solution is meticulously engineered to deliver scalable, high-performance storage that meets the rigorous demands of modern AI, machine learning, and advanced analytics workloads. By integrating HCSF and its modern approach to high-performance parallel file system capabilities, with the power of NVIDIA HGX servers, CSPs can provide unprecedented flexibility, performance, and efficiency to their next-generation AI customer base.

For cloud service providers seeking to provide the latest in hosted AI capabilities, the NCP Reference Design offers turnkey deployment, point-and-click simplicity, and validated performance at scale. Whether supporting hundreds or tens of thousands of GPUs, the solution’s flexible consumption model empowers organizations to innovate without overcommitting capital. With built-in data protection, ransomware resilience, and governance capabilities, Hitachi Vantara and NVIDIA deliver peace of mind and operational continuity.

This Reference Design positions Hitachi as a leader in next-generation AI infrastructure. It enables customers to unlock the full potential of their data, accelerate time-to-insight, and achieve competitive advantage in a rapidly evolving digital landscape. For cloud service providers who want to modernize their offerings supporting enterprises who are ready to embrace the future of AI-driven outcomes, this solution is the foundation for scalable, secure, and high-performance innovation.

# Appendix: HCSF integration with VSP One Object

VSP One Object is a Hitachi Vantara enterprise object storage solution built for modern data environments. It delivers scalable, cost-efficient protection and long-term retention and integrates with analytics, AI and ML, and data lakehouse workloads through native S3 Table and Apache Iceberg support. Its software-defined, microservices-based architecture enables flexible deployment and independent scaling of compute and storage.

Within the NCP Reference Design, VSP One Object serves as the object storage layer of the Hitachi iQ architecture. It provides the capacity tier that complements HCSF, extending access from high-performance NVMe flash to durable object storage. Together, HCSF and VSP One Object create a unified data foundation that supports high-throughput AI workloads and long-term retention within a single managed environment.

The NCP Reference Design accelerates the AI data lifecycle, from ingestion to training to inference, through tightly integrated, high-performance infrastructure. Within this framework, Hitachi iQ extends the architecture beyond GPU-accelerated compute to deliver unified data lifecycle management across the flash and object tiers. Integrating HCSF with VSP One Object enhances the NCP Reference Design with advanced capabilities for data management, protection, and governance, allowing NVIDIA HGX workloads to securely scale while maintaining integrity and consistency across the entire storage environment.

Together, HCSF and VSP One Object deliver an appliance-like experience with a single namespace spanning NVMe flash and object storage for always-hot access to data. This architecture simplifies management at scale, enabling administrators to configure, monitor, and optimize system behavior through one intuitive interface.

### Management and control

HCSF delivers point-and-click simplicity for configuration and day-to-day management. Administrators can rapidly provision new storage, create and expand file systems within the global namespace, and set policies for tiering, protection, encryption, authentication, and permissions. Protocols such as NFS, SMB, and S3 are supported natively and managed from the same interface. Read-only or read-write snapshots, snapshot-to-object policies, and quality-of-service controls can be easily defined. Detailed event logging and integrated time-series graphing provide precise visibility into system health, event history, and performance trends. Beyond day-to-day management, HCSF automates key data placement and protection functions through policy-based controls.

### Automated data management

HCSF includes a built-in, policy-based automation engine that intelligently moves data between performance and capacity tiers according to data temperature. Active datasets remain on the NVMe flash tier for maximum throughput, while colder or less active data is transparently tiered to VSP One Object or other cloud object storage for long-term retention or backup. The same policy framework governs snapshot handling, allowing snapshots to be offloaded to VSP One Object or compatible cloud tiers for cost-efficient protection and recovery without breaking the unified namespace.

### Intelligent tiering and scalability

Automated tiering extends the HCSF namespace from high-performance flash to economical object-based storage such as VSP One Object or hybrid-cloud object tiers. This design optimizes scalability and total cost of ownership while eliminating bottlenecks found in legacy protection architectures. The result is a unified, high-throughput data fabric that balances speed, scale, and efficiency for modern workloads. This distributed design also reinforces data protection strategies, ensuring consistency and availability even as datasets scale.

### Data protection and business continuity

Protection and resiliency are designed into every layer of the solution. HCSF supports user-defined snapshots for local protection, backup, or disaster recovery across both flash and object tiers. Logical snapshots, multiple protection levels, erasure coding, and remote replication ensure data consistency and availability. These protection mechanisms also strengthen the foundation for governance and compliance, ensuring that resilience extends from system recovery to data integrity and control.

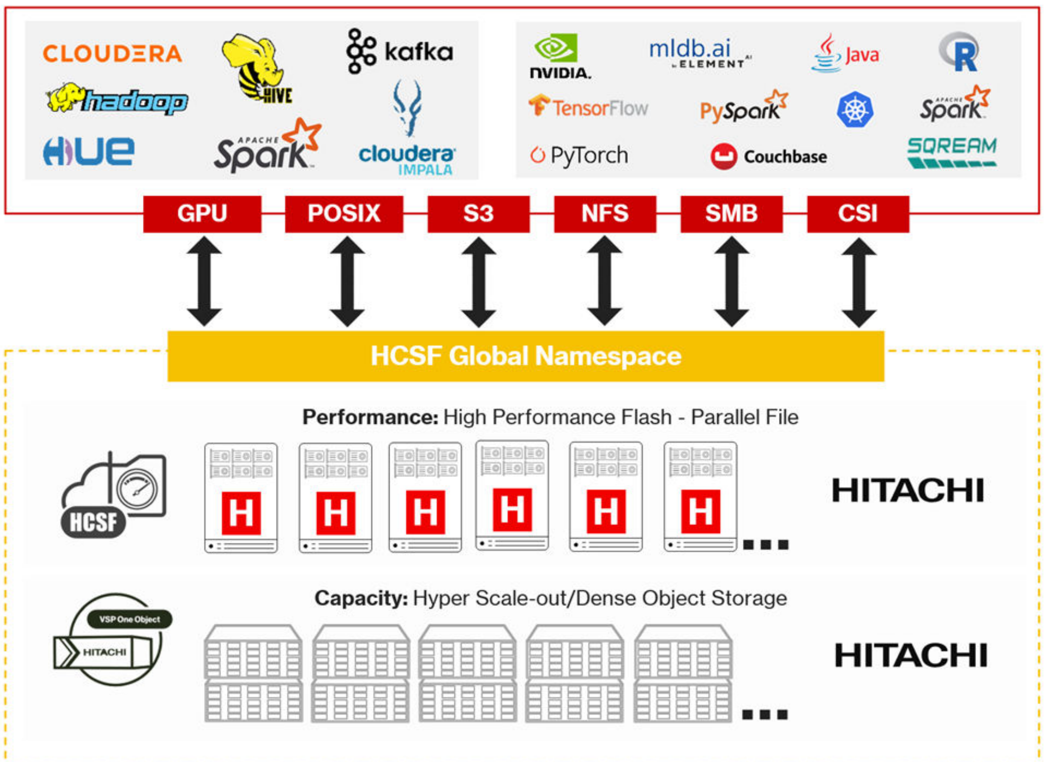
### Governance and security

When integrated with VSP One Object, HCSF inherits additional safeguards including immutability, versioning, and encryption for ransomware resilience. VSP One Object also extends governance with retention enforcement, secure destruction, authenticity validation, access control, and audit capabilities that help organizations meet stringent data governance and compliance requirements. Together, these capabilities create a consistent, compliant, and secure data foundation. With governance enforced at the storage layer and accessibility unified across file and object tiers, organizations can confidently extend this architecture to analytics and AI workloads. This trusted foundation supports modern data lakehouse environments that depend on governed, high-performance data pipelines.

**Data lakehouse architecture**

Built on a unified namespace that spans NVMe flash and VSP One Object, HCSF provides the foundation for data lakehouse architectures. Hot data is ingested and processed directly on the flash tier for GPU-accelerated analytics, while colder or derived datasets are automatically tiered to VSP One Object for durable, cost-optimized retention. On the object tier, S3-compatible access, native S3 Table support, and Apache Iceberg catalog integration enable open, in-place queries and analytics without unnecessary data movement or external metastores. This unified structure allows enterprises to consolidate pipelines, reduce duplication, and simplify access for both operational and analytical workloads within a single managed namespace.

The following illustration shows HCSF integration with VSP One Object.



**Hitachi Vantara**

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA

HitachiVantara.com/contact