Hitachi iQ M Series with HCSF Storage Based on NVIDIA

Reference Configuration

MK-SL-340-00
July 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.

Thank you!

Revision history

Changes	Date
Initial release	July 2025

Reference Configuration

Purpose

As organizations accelerate their digital transformation, artificial intelligence (AI) has emerged as a key driver of innovation, operational efficiency, and competitive advantage. Building AI-ready infrastructure that balances performance, scalability, and cost-efficiency remains a significant challenge. Hitachi Vantara, in collaboration with NVIDIA, addresses this by delivering integrated solutions grounded in NVIDIA Enterprise Reference Architectures. These architectures define a balanced, high-performance GPU cluster architecture built upon NVIDIA-certified servers. Designed to harness NVIDIA NVLink™ technology within each individual server, optimize capacity and performance, and provide guidance for scaling to meet workload demands with additional server resources.

The PCIe-Optimized 2-8-5-200 Reference Configuration defines a balanced design with 2 CPUs, up to 8 GPUs, and 5 NICs, delivering 200 GbE or greater bandwidth. This model scales from 4 to 32 nodes and is ideal for high-performance, distributed AI workloads.

The Hitachi iQ solution empowers enterprises to deploy intelligent, flexible, and high-performing infrastructure tailored to their unique AI workloads and business needs. Whether deploying general-purpose AI environments or industry-specific use cases, Hitachi iQ streamlines the path from concept to implementation, enhancing automation, improving customer experience, and accelerating time-to-value.

Unlike conventional solutions focusing solely on infrastructure, Hitachi iQ solution incorporates domain-specific AI capabilities across finance, manufacturing, energy, and transportation sectors. This strategic layering of AI services enables faster deployment and adoption, simplifying the AI journey and delivering outcomes with real-world relevance.

Hitachi iQ solutions scale from foundation-level deployments to enterprise-class environments. This document focuses on the Hitachi iQ M series solution and outlines the design and reference configuration of the 2-8-5-200 foundation storage solution based on NVIDIA Enterprise RA guidelines. Leveraging Hitachi Content Software for File, Hitachi iQ M Series has been tailored to support the NVIDIA 2-8-5-200 reference configuration, suitable for medium scale AI/ML and generative AI workloads.

The intended audience of this document is:

- Data scientists and engineers
- AI developers and architects
- Security operations teams
- Storage and systems administrators
- IT infrastructure professionals

This technical paper assumes that you are already familiar with the following:

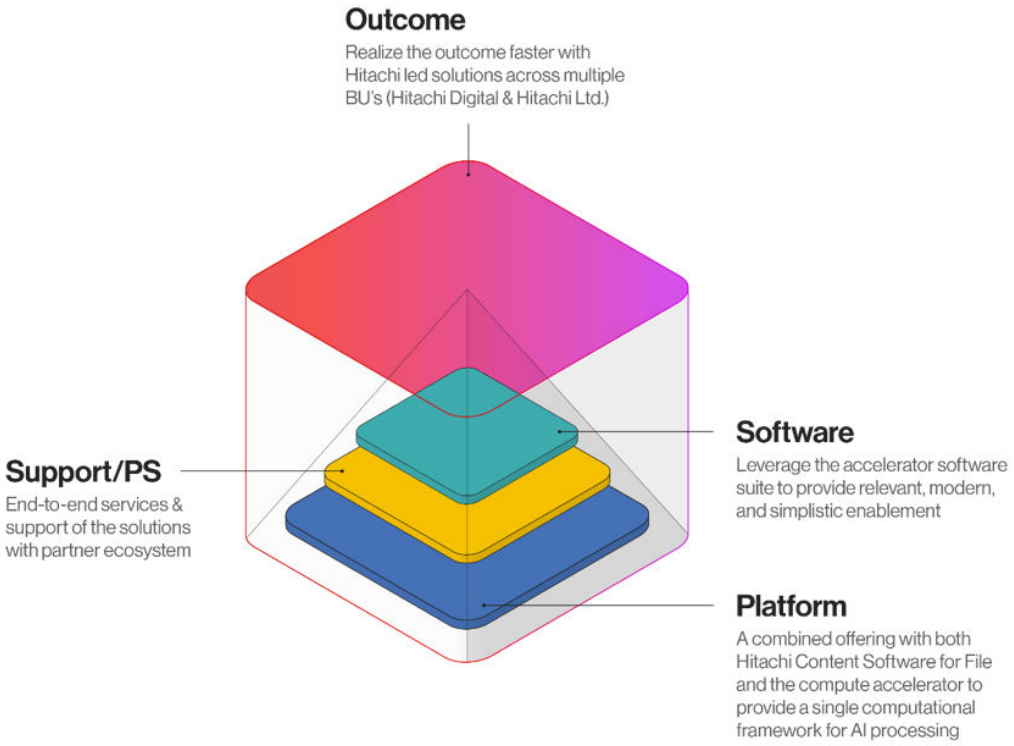
- GPU-based accelerated compute platforms
- General IT storage concepts and best practices
- General networking concepts
- Kubernetes orchestration

Solution overview

Unlike traditional AI infrastructure offerings, the Hitachi Vantara AI portfolio, Hitachi iQ, goes beyond basic integration and storage functionality. By incorporating industry-specific AI outcomes, Hitachi iQ offers full turn-key AI systems that align closely with the distinct requirements and strategic goals of organizations and their AI Factories.

Hitachi iQ M Series stands out as the only accelerated solution that offers unified data access across environments, while maintaining robust capabilities around data explainability, lineage, accuracy, security, and traceability. These characteristics are critical for mission-critical applications, particularly as organizations transition from proof-of-concept to full-scale, production-grade AI deployments.

At the core of the Hitachi iQ platform is Hitachi Content Software for File, a highly scalable, software-defined, parallel file system. Designed for flexibility, Hitachi Content Software for File fully integrates with Hitachi Vantara VSP One Object while supporting on-premises, cloud-based, and hybrid environments. It supports a wide range of high-performance workloads, including AI, machine learning (ML), and advanced analytics. With the ability to scale beyond 14 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 describes a Foundation Storage certified system integrating Hitachi Content Software for File with Hitachi Vantara accelerated GPU servers and NVIDIA Spectrum™-X switching. In this solution, compute, storage, and networking are optimized to provide a balanced, high-throughput infrastructure for AI, GenAI and other data-intensive workloads.

Key components and technologies

Hitachi Content Software for File (HCSF) powered by WEKA

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 Magnum IO™ 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, 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
- Data Protection – offload snapshots straight to VSP One Object or any public cloud for long-term retention
- Best economics – combine flash and object 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 object storage, without the need for special data management software or complex scripts; all data maintains native filesystem access for ease of use and management.

Hardware Capabilities

A single file system can support billions of directories and trillions of files, delivering a scalability model more akin to object storage 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 in the global namespace
- 6.4 billion files in a directory
- 4 petabytes for a single file

Supported networking technologies:

- NVIDIA Quantum InfiniBand Platform - NDR400 (400 Gb), NDR200 (200 Gb), HDR (200 Gb), and EDR (100 Gb)
- Ethernet – 100 Gb, 200 Gb, 400 Gb, 800 Gb

Software Capabilities

Point-and-click simplicity allows users to rapidly provision new storage; create and expand file systems within a global filesystem, create multi-tenant isolation, establish tiering policy, data protection, encryption, authentication, permissions, NFS, SMB and S3 configuration, read-only or read-write snapshots, snapshot-to-objects, and quality of service policies, as well as monitor overall system health. Detailed event logging provides users the ability to view system events and status over time or drill down into event details with point-in-time precision via the time-series graphing function.

Hitachi Content Software for File has a built-in, policy-based automated data management feature that transparently moves data across storage types according to the data temperature. Hitachi Content Software for File supports moving data from the NVMe flash storage tier to Hitachi Vantara VSP One Object. Snapshots can be saved to lower-cost cold storage, such VSP One Object storage.

Hitachi Content Software for File supports user-definable snapshots for routine data protection, including backup. For example, snapshots can be used to back up files locally on the flash tier as well as to make copies to cloud storage tiers for backup or disaster recovery.

To provide multi-tenant isolation, Hitachi Content Software for File now supports composable clusters, which are flexible groupings of resources that can be dynamically configured. This allows for the creation of logically isolated micro-clusters, or smaller, independent clusters within the deployed infrastructure. This functionality ensures multiple discreet environments can be provided within the same deployed infrastructure. For example, a company could use composable clusters to separate development and production environments, ensuring that data and workloads remain isolated while sharing the same physical infrastructure. Each composable cluster can scale to meet the unique demand of each tenant, providing maximum flexibility through modern capabilities.

The following protocols are supported by the solution:

- POSIX Compliant Client
- NVIDIA Magnum IO GPUDirect Storage (GDS)
- NFS (Network File System) v3 and v4.1
- SMB (Server Message Block) v2 and v3
- S3 (Simple Storage Service)

Hitachi iQ GPU Compute Node

The Hitachi iQ accelerated compute offering, including the latest generation Intel dual processor platform combined with up to 8 NVIDIA GPUs, provides unparalleled performance for data-intensive applications such as AI, machine learning, and high-performance computing.

Key features:

- Dual Intel® Xeon® 6900 series processors with P-cores, up to 500W TDP
- Support for up to 8 double-width NVIDIA RTX Pro 6000 Blackwell Server Edition, H100, H200, or L40S PCIe GPUs
- 24 DIMM slots, ECC RDIMM / MRDIMM DDR5
- Up to 13 PCIe 5.0 x16 FHFL slots
- 24 × U.2/U.3 NVMe and 2 × M.2 NVMe
- MCX75310AAS-NEAT 400-Gigabit single port

- NVIDIA BlueField®-3 BF3220 DPU 200 GbE Dual Port
- Flexible networking with AIOM slots
- 6 Redundant (N+N) 2700W Titanium level power supplies

Hitachi iQ networking

The networking design leverages NVIDIA's expertise in AI data centers to optimize traffic flow, delivering maximum AI performance, scalability, and enterprise-grade manageability and security. The solution architecture is built on the NVIDIA Spectrum-X Ethernet platform, integrating Spectrum-4 Ethernet switches with NVIDIA ConnectX®-7 NICs and BlueField-3 DPUs and SuperNICs to ensure high throughput and low latency. It also includes guidance for selecting the appropriate network topology based on deployment, scale, and workload requirements.

NVIDIA Spectrum SN5600 Switch

The NVIDIA SN5600 switch offers 64 total ports of 800 Gbps connectivity for compute, in-band management, and the storage fabric, supporting the complete solution.

The NVIDIA SN5600 switch ports can provide speeds between 10 and 800 Gbps, supporting up to 128 × 400 Gbps or 256 × 200 Gbps break-out ports best suited for use by compute, storage resources, or any other connectivity needs.

NVIDIA Spectrum SN2201 Switch

The NVIDIA SN2201 switch offers 48 × 1/10 Gbps ports providing connectivity to support Out-of-Band management. Out-of-Band management provides consolidated management connectivity for all components in the solution.

NVIDIA ConnectX-7 NIC

NVIDIA ConnectX-7 NICs support a wide range of Ethernet speeds, including 25, 50, 100, 200, and 400 Gbps. In this Ethernet-based architecture, Hitachi AC520 GPU compute nodes are equipped with single-port ConnectX-7 NICs operating at 400 Gbps for high-speed compute fabric connectivity. The HCSF 31116 storage nodes utilize two single-port ConnectX-7 NICs at 400 Gbps each to deliver a combined high-throughput storage network.

NVIDIA BlueField-3 DPUs and SuperNICs

The NVIDIA BlueField-3 networking platform is a high-performance network accelerator engineered to enhance hyperscale AI workloads. It delivers up to 400 Gb/s of remote direct memory access (RDMA) over Converged Ethernet (RoCE), facilitating rapid and efficient data transfer between GPU servers. This capability is crucial for optimizing the performance of network-intensive, massively parallel computing tasks typical in AI applications.

Kubernetes

Kubernetes is an open-source container orchestration platform for deployment automation, scaling, and management of containerized applications.

NVIDIA AI Enterprise

NVIDIA AI Enterprise is a comprehensive, cloud-native software platform designed to accelerate data science pipelines and streamline the development and deployment of production-grade co-pilots and generative AI applications. It offers easy-to-use microservices that deliver optimized model performance, coupled with enterprise-grade security, support, and stability.

This ensures a seamless transition from prototype to production for businesses that rely on AI to operate.

NVIDIA GPUDirect RDMA

GPUDirect RDMA provides a direct P2P (Peer-to-Peer) data path between the GPU memory directly to and from NVIDIA NIC and SuperNIC devices. This reduces GPU-to-GPU communication latency and completely offloads the CPU, removing it from all GPU-to-GPU communications across the network.

NVIDIA GPU Direct Storage

GPUDirect Storage (GDS) is like GPUDirect peer-to-peer (see Enhancing Data Movement and Access for GPUs) that enables a direct memory access (DMA) path between the memory of two graphics processing units (GPUs) and GPUDirect RDMA that enables a DMA path to a network interface card (NIC), GDS enables a DMA data path between GPU memory and storage, thus avoiding a bounce buffer through the CPU. This direct path can increase system bandwidth while decreasing latency and utilization load on the CPU.

Solution components

This reference configuration integrates a comprehensive set of hardware and software components designed to deliver a scalable, high-performance AI infrastructure.

Hardware Components

This reference configuration integrates validated compute, storage, networking, and management hardware to deliver scalable and high-performance AI infrastructure.

Component	Description / Specification
Hitachi AC520 GPU Server	5U dual CPU socket GPU server, supports up to 8 double-width PCIe GPU accelerator cards. Features dual Intel® Xeon® 6900 series processors, 24 DIMM slots for DDR5 memory, and up to 24 NVMe drive bays.
NVIDIA RTXPro 6000/ H100/H200/L40S PCIe GPUs	PCIe-based GPUs with NVLink connectivity for AI/ML workloads are used in the AC520 compute nodes
NVIDIA Bluefield-3 SuperNIC	Dual-port 200 GbE NIC used for in-band management and storage traffic (AC520 and HA810 G3 nodes).
NVIDIA ConnectX-7 NIC	Single-port 400 GbE NIC used in both AC520 GPU compute nodes and HCSF 31116 storage nodes for high-speed compute fabric and data.
HCSF 31116 Storage Node (8 total nodes)	1U NVMe-based storage appliance with AMD EPYC Genoa 9534 (64-core) CPU, 768 GB RAM, and 16 × NVMe SSDs (7.6 TB or 15.3 TB each).
Hitachi Advanced Server HA810 G3 (Optional)	1U dual-socket server used for control plane and system management services.
NVIDIA Spectrum-4 SN5600 Switch	64-port 800 GbE or 128-port 400 GbE or 256-port 200/100 GbE (using breakout cables) switch providing converged compute, storage, and management networking.
NVIDIA Spectrum SN2201 Switch	1U 48 ×1 GbE + 4 ×100 GbE OOB switch used for hardware-level management (BMC/iLO).
Hitachi Universal V3 Rack	42U standard rack with 80% ventilated doors, cable management trays, and PDU mounting brackets.
Raritan PX3-5541R-E2 iPDU	Intelligent PDU for compute racks, 208V 3-phase 48A, 17.3kVA capacity.
Raritan PX4-532A-E7V2 iPDU	Intelligent PDU for storage and infrastructure racks, 208V 3-phase 48A, 17.3kVA capacity.

Software Components

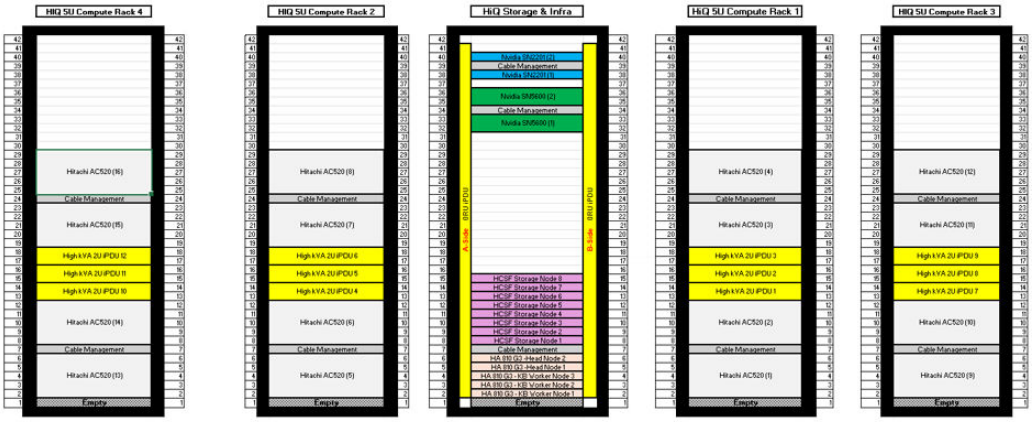
This reference architecture is built on a curated set of software components to enable high-performance AI workloads, efficient orchestration, and simplified infrastructure management.

Component	Description / Purpose
RHEL 9.4	Linux OS deployed on all compute and management nodes.
Kubernetes	Container orchestration platform for deploying and scaling AI/ML workloads.
OFED	Enables RDMA support and low-latency network communication over high-speed fabrics.
NVIDIA Drivers and Tools (CUDA, GDS)	Includes GPU drivers, CUDA toolkit, and GPUDirect Storage (GDS) for accelerated compute.
NVIDIA GPU and Network Operators	Automates lifecycle management of GPUs and NICs within the Kubernetes environment.
NVIDIA AI Enterprise and NVIDIA Base Command™ Manager	Provides a validated AI software suite and centralized GPU infrastructure management.
NVIDIA Cumulus Linux	Network OS used on NVIDIA Spectrum switches.

Solution design

This is a detailed solution example of how the Hitachi iQ M Series with Hitachi AC520 Compute Nodes and Hitachi Content Software for File Storage is configured.

The following is a reference rack elevation design for 16 Hitachi AC520 compute nodes. This reference design offers a scalable and modular architecture to start as small as 4 compute nodes and grows up to 32 compute nodes across additional racks. The rack layout can be adjusted to meet local datacenter requirements, such as maximum power per rack and rack layout between system, storage, and network components to meet local needs for power and cooling distribution.

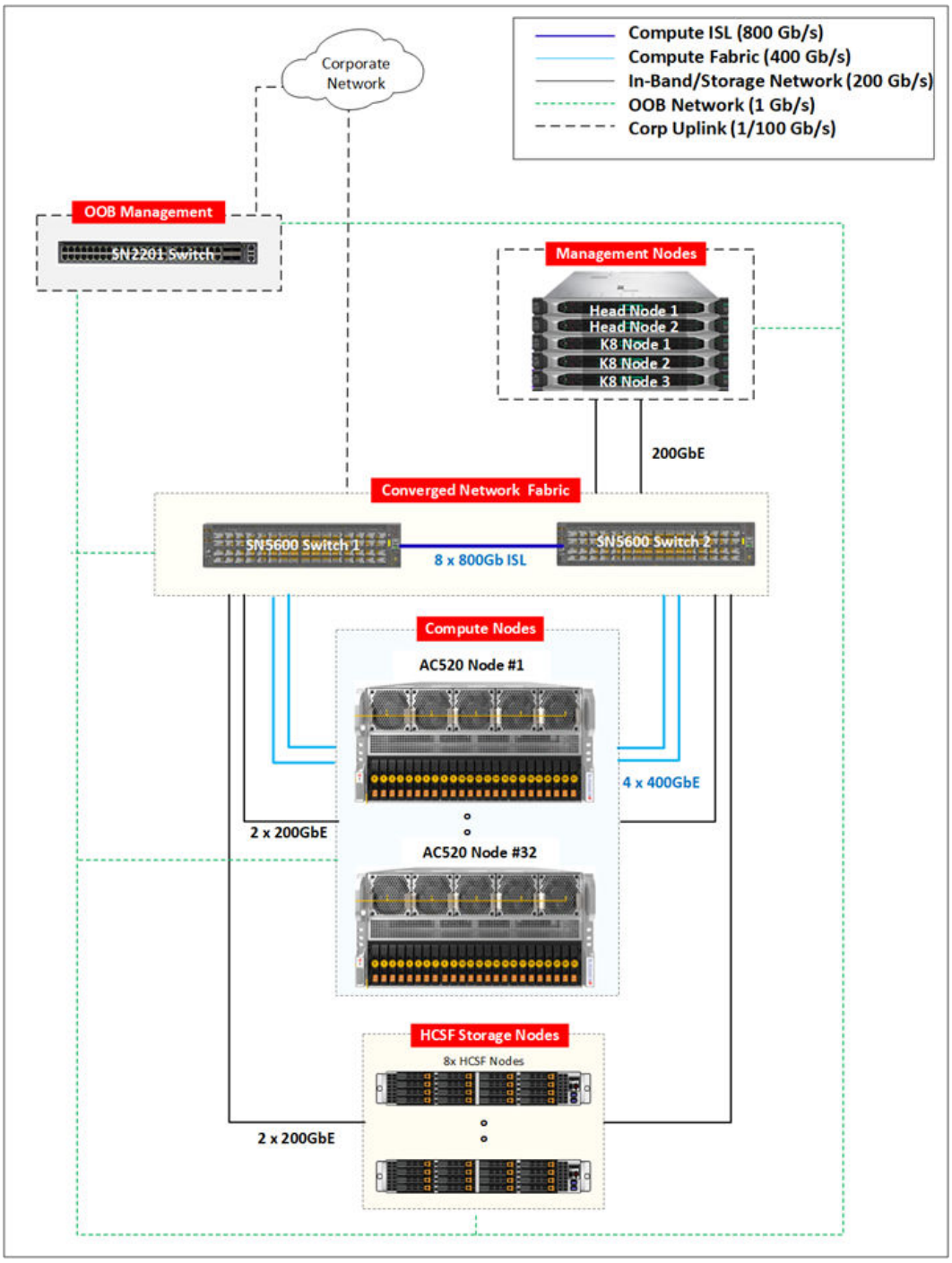


Physical Architecture

The following are Hitachi iQ M Series with AC520 Compute nodes and HCSF storage architecture and components:

- Hitachi Advanced Server HA810 G3 are used as management nodes.
  - 2 × NVIDIA Base Command Manager head nodes
  - 3 × K8s control plane nodes
- Hitachi AC520 servers with up to 8 GPUs are used as compute nodes to form K8S cluster worker nodes, which scale from 4 nodes to 32 nodes.
- 8 × HCSF 31116 servers as storage nodes.
- 2 × NVIDIA Spectrum-4 SN5600 switches for converged network fabric, these switches will be used for compute (East-West), storage, in-band management, and North-South traffic.
- 2 × NVIDIA Spectrum SN2201 switches for out-of-band management network.

The following illustration shows the system architecture:



Network Overview

The solution is built entirely on an Ethernet-based architecture and is categorized into four main logical networks (compute, storage, in-band, and out-of-band) to support the high-performance compute, storage, and management operations.

Converged Network Fabric:

- The core of the network is built on two SN5600 Ethernet switches operating across multiple 800 Gbps ISLs.
- This converged fabric handles both compute traffic via 400 GbE and storage/in-band management traffic via 200 GbE.
- All Hitachi AC520 GPU compute nodes connect to the fabric using:
  - 4 × single-port 400 GbE ConnectX-7 NIC for compute fabric.
  - 1 × dual-port 200 GbE BlueField-3 DPU for storage and in-band management.
- HCSF storage nodes connect using:
  - 2 × single-port 400 GbE ConnectX-7 NICs for high-throughput storage communication.
- Management nodes also participate in this converged network using:
  - 1 × dual-port 200 GbE BlueField-3 SuperNIC for shared in-band management and storage access.
- VLANs can be configured for to segregate networks.

Out-of-Band (OOB) Management Network:

- Managed through SN2201 switch providing 1 GbE connectivity.
- Provides 1 GbE access for BMC/iLO management of compute, storage, management nodes and other devices.
- This network is isolated from the data and control planes for secure hardware-level access.

Corporate Uplink Network:

- Provides external connectivity for management nodes through multiple 1/100 GbE uplinks to the corporate network.
- These connections are used for external cluster control, orchestration, enterprise data and internet access when needed.

Scalability Matrix: Network and Infrastructure Requirements

The following table outlines the infrastructure scaling requirements based on the number of compute nodes in the deployment. It provides a comprehensive view of the necessary network ports, node counts, rack space, and switch configurations required to scale from a small to a large deployment using the Hitachi AC520 compute nodes and HCSF 31116 storage nodes.

Compute Nodes	Compute Rack	Compute Ports	In-band / Storage Ports	HCSF Nodes	HCSF In-band / Storage Ports	Management Nodes	Management In-band Ports	Storage & Infra Rack	Total required Racks	Total Storage Ports	Converged Network Switches	Total Out of Band Ports	OOB Switch
4	1	**16**	8	8	16	5	10	1	**2**	**34**	2	**21**	**1**
8	2	**32**	16	8	16	5	10	1	**3**	**42**	2	**26**	**1**
16	4	**64**	32	8	16	5	10	1	**5**	**58**	2	**36**	**1**
32	8	**128**	64	8	16	5	10	1	**9**	**90**	2	**56**	**2**

Cabling and Transceiver Matrix

The following table provides recommended switch-side and host-side transceiver types, along with corresponding cable types and intended use cases across the architecture. While indicative cable length ranges are listed, actual lengths should be selected based on customer-specific deployment environments, rack layout, and distance requirements.

Switches	Hosts	Switch Side Transceivers	Host Side Transceivers	Cables	Purpose	Length Options
SN5600	N/A	OSFP	N/A	DAC Cable	ISL connectivity between SN5600 switches	1.5m or as needed
	AC520	OSFP	Single Port 400 GbE OSFP CX-7	400 GbE MPO/APC Cable	GPU network connectivity	Up to 30m or as needed
	AC520	OSFP	Dual port 200GbE QSFP112 BlueField-3	800 Gb to 200 Gb breakout AOC Cable	Compute node in-band/storage connectivity	Up to 30m or as needed
	HCSF Node	OSFP	Single port 400GbE CX-7	800 Gb to 400 Gb breakout AOC Cable	Storage node in-band/storage connectivity	Up to 3m or as needed
	Management Node	OSFP	Dual port 200GbE QSFP112 BlueField-3	800 Gb to 200 Gb breakout AOC Cable	Management node connectivity	Up to 3m or as needed
SN2201	All hosts	RJ45	RJ45	CAT-6 Cable	Out of Band Management Connectivity	Up to 30m or as needed

Conclusion

Hitachi iQ M Series with AC520 and HCSF storage solution offers the unparalleled computational power of Hitachi AC520 accelerated compute systems with the scalable and efficient storage capabilities of Hitachi Content Software for File offering a comprehensive solution that addresses the complexities of modern high-performance computing and AI-driven workloads.

This reference configuration serves as a foundational guide, providing the necessary tools and best practices to leverage the combined strengths of Hitachi AC520 accelerated compute systems and Hitachi Content Software for File Storage, ultimately driving success in the era of data-driven decision-making.

Key benefits of this reference architecture include:

- Enhanced performance and scalability for demanding AI and HPC applications.
- Reliable and secure data management, ensuring data integrity and compliance.
- Simplified deployment and management, reducing operational complexities.
- Comprehensive support for advanced AI workflows, driving innovation and insights.

By adopting this solution, organizations can effectively harness the power of AI and HPC, transforming data into actionable insights and achieving strategic objectives with greater efficiency.

Hitachi Vantara

Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA

HitachiVantara.com/contact