Datasheet

Hitachi iQ M Series

Affordable, Adaptable AI Infrastructure Solutions

Hitachi iQ M Series brings a modular approach to AI infrastructure that simplifies deployments with turnkey solutions that can start small, go big, and evolve as needs change.

Organizations are constantly looking for ways to automate, accelerate time to market and develop new insights, products or innovations to propel their business forward. At the same time, AI and generative AI technologies are revolutionizing industries by enhancing existing capabilities and transforming how quickly and creatively business problems can be solved. Whether building specific AI solutions or just starting to identify general-purpose capabilities, Hitachi iQ has the power to automate your business processes and improve your AI experience.

Hitachi iQ allows you to automate, expedite and streamline your business through intelligent, performant, scalable and flexible AI and genAI solutions. Unlike conventional AI offerings, Hitachi iQ transcends basic integration and storage capabilities by also layering industry-specific AI outcomes within the AI solution. This approach ensures outcomes are finely tuned to your organization's unique needs and objectives.

Hitachi iQ provides unified access to data — irrespective of where it resides — while ensuring explainability, lineage, data accuracy, security and traceability at any given point for mission-critical solutions. By optimizing AI deployments with end-to-end software-defined AI and data analytics software, Hitachi iQ streamlines development and deployment of production-grade AI applications, from pilot to production.

Hitachi iQ M Series provides right-sized, modular solutions for AI processes and application modernization initiatives. Start with a small configuration to experiment with your genAI projects and scale up as your initiatives grow.

Incremental Scalability

Adopt a modular approach to AI with appropriately sized initial deployments, safeguarding infrastructure investments and adapting as AI requirements evolve.

Affordable Solutions

Slash entry costs and operational expenses for in-house AI with modular accelerated compute, integrated file system and storage options, and IP-based networking.
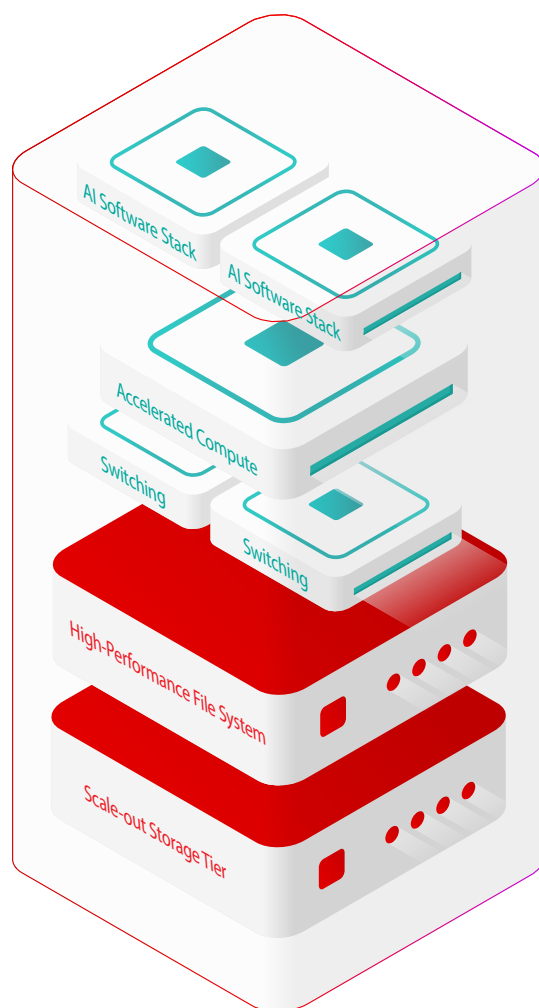
Industry Relevant, Meaningful Outcomes

Recognizing that not all organizations possess the AI skills for their industries, we bring our expertise to help bridge the gap, creating cutting-edge AI applications and improving your outcomes.

Hitachi iQ: Flagship AI Portfolio From Hitachi, Powered by NVIDIA® Technologies

When organizations try to implement their own AI solutions, they face the challenges of creating complicated systems and performing necessary integrations on their own. Hitachi iQ M Series eases the process by offering prebuilt platforms with bespoke customization and delivery — all from a single vendor.

Engineered with Hitachi and NVIDIA solutions, Hitachi iQ accelerates integration based on a proven foundation. For AI-accelerated compute, Hitachi iQ M Series platforms combine the best software, infrastructure and expertise in unified and scalable AI development solutions. M Series storage consolidates storage systems into a unified, high-performance global data environment, eliminating the delays and inefficiencies associated with traditional data silos. The global namespace ensures that massive datasets, such as petabytes of data that feed AI workloads, are readily available to GPU clusters, regardless of the location of the data or compute resources.

Hitachi iQ solutions also include NVIDIA-specified compute and storage infrastructure and interconnect components. These can be implemented with NVIDIA Base Command Manager software suite for provisioning, managing and monitoring and NVIDIA AI Enterprise (NVAIE) for AI deployment and enablement. Hitachi Vantara provides end-to-end services and support of the solutions with our partner ecosystem.

Workloads and Use Cases

Whether you are looking for industry-specific AI solutions or just starting to identify general-purpose capabilities, AI has the power to automate your business processes and improve your customer experience. The industry is already using AI for things like creating art and graphic design elements, writing code and generating marketing taglines. And new use cases are being identified every day that can provide immediate benefits to any organization. For example:

Data Analysis and Transformation

- Advanced data analytics
- Automated reporting/visualization
- Predictive analytics
- Data mining/business intelligence
- Data cleaning/error correction/integration

Processes Automation and Optimization

- Processess automation
- Resource allocation
- Asset management
- Scheduling
- Route planning
- Document analysis

Conversational AI and Virtual Assistants

- Natural language processing
- Chatbots
- Voice assistants
- Speech recognition/synthesis
- Sentiment analysis
- Intent recognition

Information Exploration and decision Support
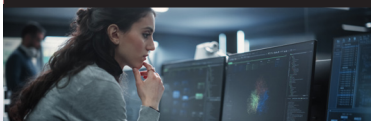
- Information retrieval
- Knowledge discovery/extraction
- Summarization
- Risk assessment
- Classification/trend analysis
- Strategic planning

Content Generation and Personalization

- Automated writing
- Personalized content
- Personalized marketing/learning
- Image, audio, video and graphic creation
- Synthetic data generation

With Hitachi iQ, Hitachi leverages our deep domain experience with 110+ years of experience in the operational technology (OT) market in industries such as finance, transportation, energy, media and entertainment, manufacturing and healthcare. With the Hitachi OT and IT legacy, Hitachi iQ portfolio integrates infrastructure and capabilities to deliver industry-specific AI solutions tailored to your needs.
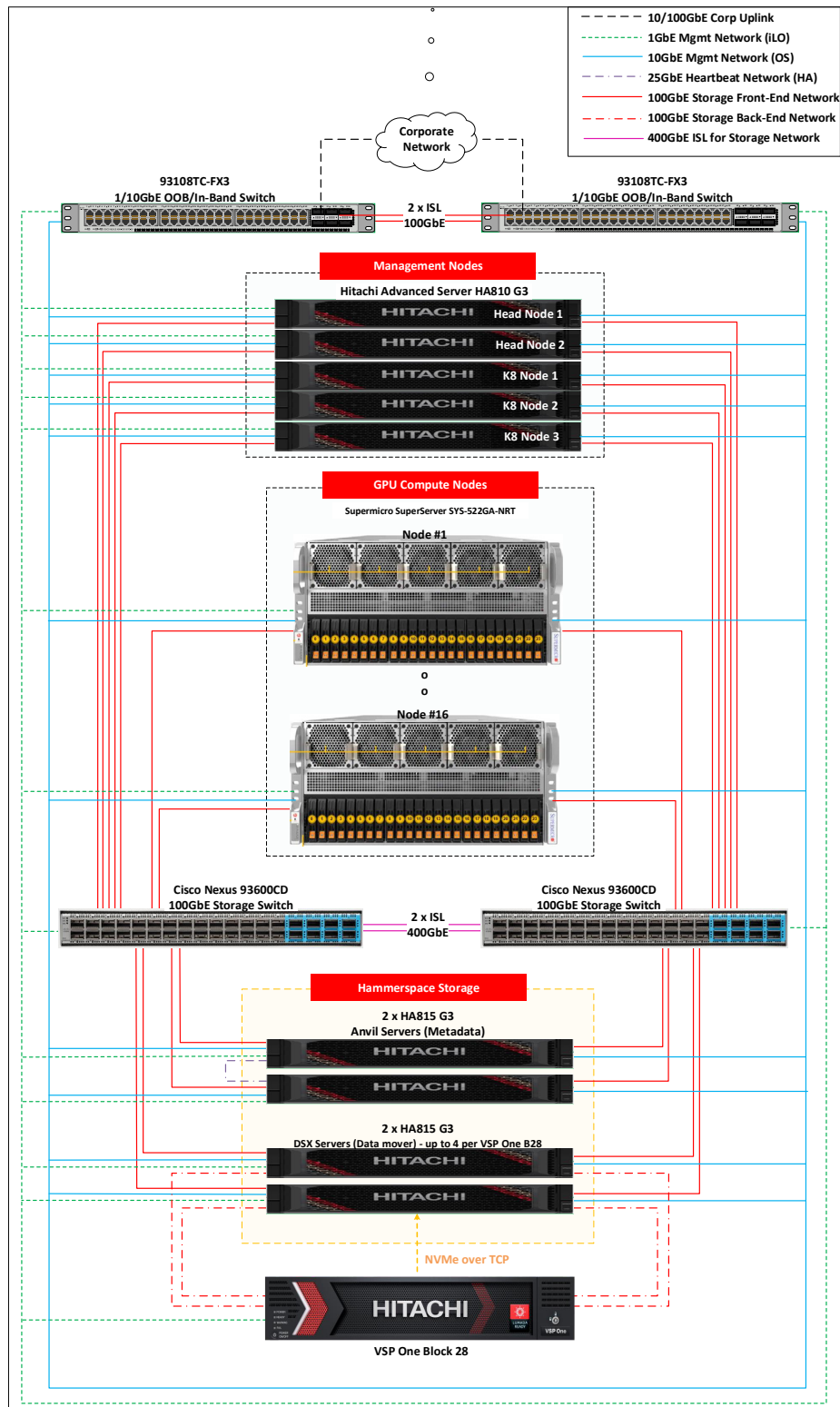
Hitachi iQ Solution Architecture

Figure 1. Hitachi iQ M Series Reference Architecture

Management nodes	**Chassis:** 3x HA G3 nodes for management. If BCMe software is used, 2 additional HA G3 nodes will be needed as with the HGX program. These servers will need NFS access over IP to the Hammerspace storage. **Notes:** Optional/customer supplied. In-band management to manage the entire cluster – OS, container orchestration and applications.
Compute nodes	**Chassis:** Starting with 1, minimum of 16x SMCI X-14 AC520 GPU scalable nodes. The solution will be 6 racks. Maximum number SMCI X-14 AC520 units should be fit in these racks. **Notes:** Server needs to be NVIDAI qualified and be on the NVIDIA qualified server list. **Note:** SMCI should take responsibility for this certification and posting.
	CPU: Dual Intel 6th gen GNR AP CPUs (6900-AP). Need a few CPU options. **Note:** Target 6 cores for each GPU. Need to confirm this metric with NVIDIA and SMCI.
	GPU: NVIDIA H200 NVL, H100 NVL and L40S. **Notes:** Support GPU count — up to max possible in chassis; target 4 DW GPUs for average configu-rations. A single GU type is supported in a single chassis.
	RAM: 12x 32/64/128GB DIMM per CPU with 6400MT/s or higher, 12 channels per CPU and 1 DIMM per channel. **Notes:** Need to understand minimum RAM for the range of GPUs in the system. Expecting about 256GB per GPU. Need to confirm with SMCI.
	Local storage: ● NVMe read-intensive SFF, 1 DWDP. This can be U.2/U.3. ● Internal boot option. NICs: Ethernet, minimum two 100Gb per port. Preferred using 2P@100Gb CX-6 VPI.
Storage	**Support the Hammerspace storage HW with parallel NFS connectivity:** ● 2x Anvil servers – for min configuration refer to Hammerspace SRD. ● 2x DSX servers – for min configuration refer to Hammerspace SRD. ● 1x VSP One Block 28 – single tray with 24 NVMe drives.
Connectivity	**Data networking (minimum):** 1. AC520 nodes and switch – 100GbE 2. Control nodes and switch – 100GbE 3. Head nodes and switch – 100GbE 4. Anvil nodes and switch – 100GbE 5. DSX nodes and switch – 100GbE 6. DSX nodes and VSP One Block 28 – 8x 100GbE (direct con-nect) 7. Corporate and switch – 400GbE **Management networking:** ● Operating system – 10GbE (can use OCP) **High-speed high-availability (HS HA) networking** ● HA for HS Anvil nodes – 10GbE (reuse the data networking) **Management [out of band (OOB)]:** ● IP networking @ 1GbE or 10GbE, single or redundant switch
Host Operating Environment	● K8s-based running on RHEL 9.X ● Support NVIDIA AI Enterprise (NVAIE) Essentials, starting at v5.1 or later
Management Software	● SMCI server management software ● Base Command Manager Essentials v10 or later included in NVAIE Essentials ● Switch management software (optional)

Table 1: Hitachi iQ With NVIDIA HGX Components

About Hitachi Vantara

Hitachi Vantara is transforming the way data fuels innovation. A wholly owned subsidiary of Hitachi, Ltd., we're the data foundation the world's leading innovators rely on. Through data storage, infrastructure systems, cloud management and digital expertise, we build the foundation for sustainable business growth.

Hitachi Vantara

HV-BTD-DS-Hitachi iQ-M-Series-7May25-A