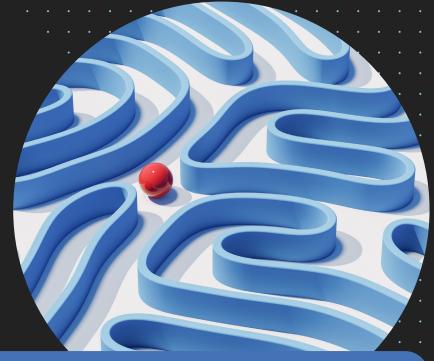
Solution Profile

# Hitachi iQ M Series

The Shortest Distance from AI to ROI



# Simplify AI

Use validated and seamlessly integrated GenAl technology solutions to accelerate your transformation journey and quickly gain insights.

# AI That Works for You

Confidently run your own GenAl, for your industry, with your own data.

## Start Smart, Be Nimble

Choose from a range of speeds, sizes and tools to fit your immediate needs and stay adaptable to rapidly evolving demands and technologies.

# Take an Affordable, Adaptable Approach to Al

The pace of AI developments can make long-term architecture investment decisions challenging. What is deployed today may not fit tomorrow's needs, and tomorrow's needs may shift dramatically. Rather than over-investing on day-one to meet expectations of growth down the line, a modular approach to AI infrastructure is a better choice for many organizations. Hitachi iQ M Series provides right-sized, adaptable infrastructure for AI outcomes, that can scale as usage evolves.

# Fully integrated solutions for Al no matter where data resides.

#### Hitachi iQ

In today's fast-evolving digital landscape, businesses are racing to harness the power of artificial intelligence (AI) to drive innovation, streamline operations, and stay ahead of the competition. However, building an AI-ready infrastructure that delivers optimal performance at the right cost can be a complex puzzle. By seamlessly integrating cutting-edge hardware, scalable software, and tailored services, Hitachi Vantara empowers customers to tackle their AI projects with confidence, efficiency, and precision. Hitachi iQ bridges the gap between ambition and execution, delivering AI-ready solutions that fit both the job and the budget. Hitachi iQ allows organizations to automate, expedite, and streamline their business through intelligent, performant, scalable and flexible AI infrastructure and solutions.

## Turn AI Into ROI With Trusted AI Solutions Tailored For Business Growth



Finance



Energy



Transportation



Media & Entertainment



Manufacturing



Healthcare

Hitachi iQ leverages its deep domain experience with 110+ years of experience in the Operational Technology (OT) industries such as finance, transportation, energy, media & entertainment, manufacturing, and healthcare.

With Hitachi's OT and IT legacy, the Hitachi iQ portfolio integrates infrastructure and capabilities to deliver industry-specific Al solutions.

## **Mass Market Al Adoption**

As Al adoption accelerates across organizations of all sizes, the need for cost-optimized deployment and scalable, high-performance Al infrastructure solutions has increased. High-end solutions, while powerful, are not practical for all Al solutions. Most Al deployments will not require massive scale model development and training. More often, organizations will download open-source models, train and tune the models against their proprietary datasets, and utilize retrieval augmented generation (RAG) to enhance the context supporting the specific outcome for a specific populations of users.

#### **RAG & Fine Tuning**

- Retrieval augmented generation (RAG) allows the optimization of AI models by connecting them with internal knowledge bases to deliver more relevant responses at a higher quality.
- Fine-tuning models requires synchronized access to both the retriever (search/indexing) and the generator (LLM).
- Data pipeline inefficiencies can lead to mismatches, latency, and poor performance.

#### Inferencing

- Al inferencing requires immediate access to relevant data, especially in latency-sensitive applications like chatbots, recommendation systems, and autonomous systems.
- Delays in retrieving data can degrade user experience and model effectiveness.

#### **Process Optimization**

- Automating processes requires training and running models on proprietary datasets demanding high-speed data access and processing capabilities.
- Bottlenecks or gaps in data access can reduce the effectiveness of automation.

## Hitachi iQ M Series

Hitachi iQ M Series is ideal for cost-sensitive, moderate scale AI use cases. It provides right-sized, modular solutions with a flexible, enterprise-grade AI infrastructure that offers a balanced approach to performance and cost. M Series makes it faster, easier, and less expensive to experiment with in-house AI projects, support initial deployments and scale up for mass production as needs evolve.

M Series integrates modular, accelerated compute servers with robust networking and enterprise or high-performance file system and storage options into a scalable, adaptable, cost-effective GenAl infrastructure solution.

With flexible configurations using NVIDIA GPUs, M Series can handle various data types, sizes, and workload demands, making it ideal for businesses operating in dynamic and data-intensive environments. And as AI models evolve and data demands scale up or down, the system can be upgraded or reconfigured without requiring a complete infrastructure overhaul, ensuring long-term viability.

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 Compute, Networking, Storage

Slash entry costs and operational expenses for inhouse AI with modular accelerated compute, integrated file system and storage options, and IP-based networking.



#### **Consolidated Infrastructure**

Simplify management and maximize utilization with an integrated, flexible, and scalable AI infrastructure.



#### **Mixed Workloads**

Run mixed, multi-modal workloads, such as language and graphics models, or model customization and inferencing projects on the same infrastructure.

### Hitachi iQ M Series

## **Hitachi iQ M Series Components**

The Hitachi iQ solution architecture includes all the components necessary to build a fully functioning and optimized AI infrastructure from entry all the way to high-end AI. This includes the GPU servers, high-speed reliable storage, high-speed networking components, and a full management stack. M Series provides flexible options for both entry and midrange offerings with tailored pricing and performance for the desired outcome.

#### Hitachi iQ M Series hardware includes:

- Up to 8 NVIDIA Data Center GPUs
- High performance data storage, protection, resiliency and 100% data availability
- Robust, high-performance networking for GPU and storage connectivity

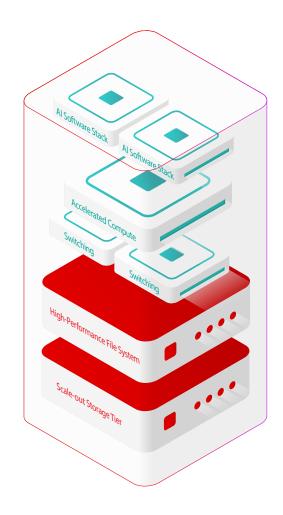
#### With options for:

- High-performance data storage systems
  - VSP One Block QLC with Hammerspace
  - 。 VSP One Object with Hitachi Content Software for File
- AI/ML software, models, and tools

The Hitachi iQ M Series offers a robust, scalable, and adaptable Al infrastructure solution tailored to meet the diverse needs of modern businesses. By leveraging Hitachi's extensive experience in both operational and information technology, the M Series provides a cost-effective and high-performance platform that can evolve with the rapidly changing demands of Al and GenAl applications. Whether you're looking to start small with initial deployments or scale up for mass production, the M Series ensures that your Al initiatives are supported by a flexible and integrated infrastructure, enabling you to turn Al into ROI efficiently and effectively.

## 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.



Learn more about Hitachi iQ solutions





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