Solution Brief

NetApp EF600 All-Flash Array with NVIDIA DGX SuperPOD

Turbocharge your AI with high-performance NVMe storage

Key Benefits

Accelerate Time to Insight
- Enable fast, continuous feeding of data to your AI application with high-performance, low-latency storage.
- Reduce downtime with 99.9999% availability.

Future-Proof Your Investment
- Quickly respond to changing workload demands and exponential data growth with a building-block architecture that enables you to seamlessly scale performance and capacity as needed.

Maximize Cost Efficiency
- Reduce operating costs with high-density drives and price/performance–optimized storage building blocks.

Reduce Risk and Enable Success
- Rely on a fully integrated, validated AI infrastructure from industry leaders to help you gain a competitive edge.

The Challenge
Although AI enhances consumers' lives and helps organizations in all industries worldwide to innovate and to grow their businesses, it is a huge disrupter for IT. To support the business, IT departments are scrambling to deploy high-performance computing (HPC) solutions that can meet the extreme demands of AI workloads. As the race to win with AI intensifies, the need for an easy-to-deploy, easy-to-manage solution becomes increasingly urgent.

The Solution
In the race to AI, you can accelerate your competitive advantage by turbocharging your NVIDIA DGX SuperPOD with the high-performance NetApp® EF600 all-flash array.

The NVIDIA DGX SuperPOD makes supercomputing infrastructure easily accessible for your organization and delivers the extreme computational power that you need to solve even the most complex AI problems. To help you deploy at scale today, this NVIDIA and NetApp turnkey solution removes the complexity and guesswork from infrastructure design and delivers a complete, validated solution (including best-in-class compute, networking, storage, and software).

Accelerate Time to Insight
In the highly competitive world of business, speed is everything. However, even the fastest supercomputer can't meet expectations if it doesn't have equally fast storage to support it.

The NetApp EF600 all-flash array gives you consistent, near-real-time access to your data while supporting any number of workloads simultaneously. To enable fast, continuous feeding of data to your AI application, EF600 all-flash arrays deliver up to 2 million cached read IOPS, response times of under 100 microseconds, and 44GBps sequential read bandwidth in one enclosure. With 99.9999% reliability from EF600 all-flash arrays, data for your AI operations is available whenever and wherever you need it.

■ NetApp®

Figure 1) DGX SuperPOD with NetApp EF600 all-flash arrays.
Visit m.kaon.com/c/na to view a 3D model of NVIDIA DGX SuperPOD with NetApp EF600 all-flash NVMe storage.

Future-Proof Your Investment

From ingestion to processing, AI operations generate a lot of data. To manage and process all that data, you need a solution that can quickly respond to data growth.

With NetApp EF600 all-flash arrays at the foundation of your NVIDIA DGX SuperPOD, you get an agile AI solution that scales easily and seamlessly. The flexibility and scalability of the solution enable it to support and adapt to evolving workloads, making it a strong foundation to meet your future storage requirements. Modular storage building blocks give you a granular approach to growth. You can scale seamlessly from terabytes to petabytes by adding capacity in any increment—one drive or multiple drives at a time. By increasing the number of storage building blocks, you can scale up the performance and capacity of the file system, enabling your solution to handle the most extreme workloads with ease.

Maximize Cost Efficiency

When you're considering the cost of an AI solution, your initial investment in hardware and software is just the beginning. To accurately assess cost, you must look at the TCO for the solution. So, in addition to the deployment costs, you must also consider the costs of operating the solution after you have deployed it.

With the NetApp EF600 all-flash array, you get enterprise storage in price/performance–optimized building blocks that make small to large configurations cost-efficient. The EF600 all-flash array is purpose-built for capacity-intensive environments that require efficient space, power, and cooling utilization.
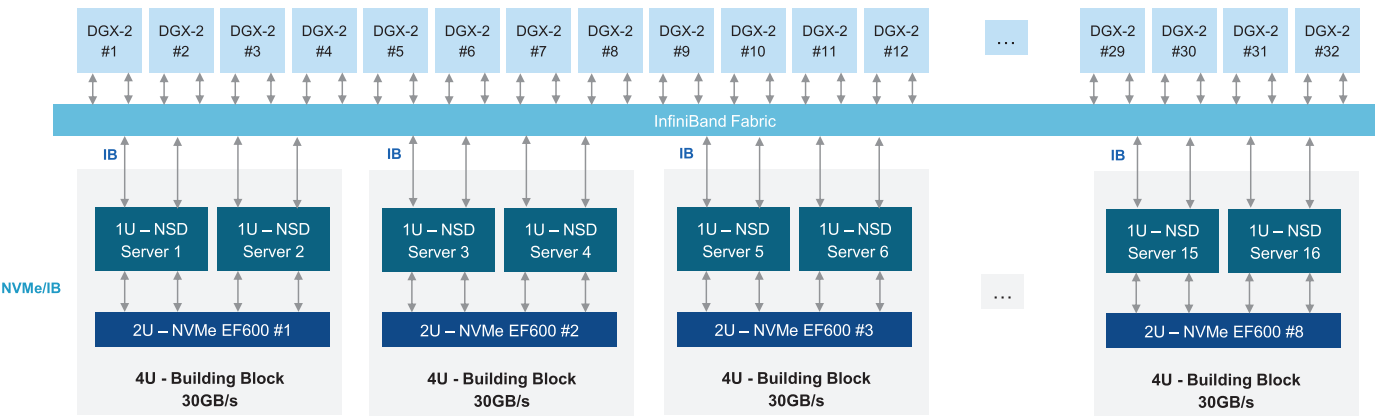

Figure 2) NetApp EF600 offers extreme scalability to support your most demanding AI workloads.

FEATURE	EF600 SPECIFICATIONS (2U, 24-DRIVE NE224 SHELF)
Controller processor	Intel Skylake 12-core
Controller memory	32GB or 128GB DDR4 per controller (64GB or 256GB per array)
Base host ports per controller	No base host ports
Optional host connectivity ports	Supports dual host interface cards: quad 32Gb FC or 100Gb InfiniBand
Maximum drives	24 SSDs, no expansion
Drive options	NVMe SSD: 1.9TB, 3.8TB, 7.6TB, or 15.3TB
I/O interface options	• 16 x 32Gb NVMe over Fibre Channel (NVMe/FC) • 16 x 32Gb SCSI FC • 8 x 100Gb NVMe over InfiniBand (NVMe/IB) • 8 x 100Gb NVMe over RoCE (NVMe/RoCE)
System performance maximum (dual controllers)	• 2 million random reads IOPS • 44GBps sequential read bandwidth

Table 1) NetApp EF600 specifications.

Reduce Risk and Enable Success

NetApp and NVIDIA have a long history of collaboration to deliver a portfolio of AI solutions to market. DGX SuperPOD with the NetApp EF600 all-flash array is a proven, validated solution that you can deploy with confidence. This fully integrated, turnkey architecture takes the risk out of deployment and puts you on the path to winning the race to AI leadership.

About NetApp

NetApp is the leader in cloud data services, empowering global organizations to change their world with data. Together with our partners, we are the only ones who can help you build your unique data fabric. Simplify hybrid multicloud and securely deliver the right data, services, and applications to the right people at the right time. Learn more at **www.netapp.com**.