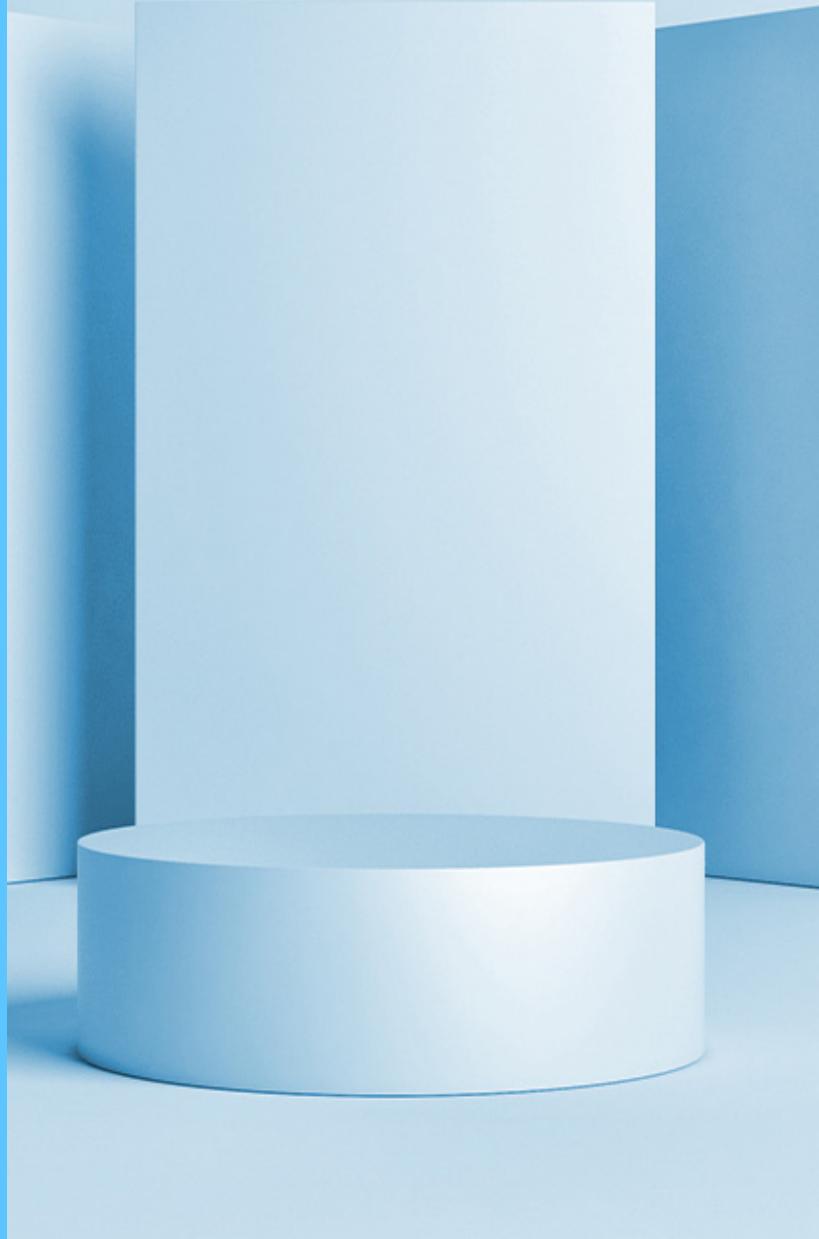


SOLUTION BRIEF

NVIDIA DGX SuperPOD with NetApp

Turbocharge your AI and HPC
workloads with high-performance
NVMe storage and BeeGFS
parallel file system



The challenge

Although AI enhances consumers' lives and helps organizations in all industries worldwide to innovate and to grow their businesses, it is a 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-scale, and easy-to-manage solution becomes increasingly urgent.

The solution

NVIDIA DGX SuperPOD with NVIDIA Base Command Manager brings together a design-optimized combination of AI computing, network fabric, storage, software, and support. Every implementation is validated on a dedicated acceptance cluster at NVIDIA. The customer's design is replicated beforehand, and a suite of performance results is produced. So, when DGX SuperPOD is deployed on site, it runs exactly as it was intended.

AI supercomputers also require extremely high-speed storage to run at peak capacity. In a well-architected system, storage solutions need to handle various data types—such as text, tabular data, audio, and video—in parallel and with unwavering performance. NetApp® EF600 all-flash NVMe storage combined with the BeeGFS parallel file system is certified for NVIDIA DGX SuperPOD. To promote success, the EF600 array was carefully selected and tested for the unique demands of AI workloads.

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.

All DGX SuperPOD certified storage is guaranteed to meet or to exceed the baseline NVIDIA performance threshold. The NetApp EF600 all-flash array, combined with the BeeGFS parallel file system, has been tested and certified by NVIDIA. Each EF600 and BeeGFS-based scalable building block adds up to 76GBps/23GBps of sequential read/write performance and 431TB of capacity. Capacity and performance can easily be sized and optimized for metadata operations, data storage, or any mix of the two. With proven 99.9999% availability, the EF600 array also significantly reduces your system's downtime.

Key benefits

Accelerate time to insight

- Get going faster by eliminating design complexity and guesswork with a certified solution.
- Simplify deployment with full integration into NVIDIA Base Command Manager.

Future-proof your investment

- Integration with BeeGFS enables next-generation workloads at scale.
- Quickly respond to changing workload demands and exponential data growth with a building-block architecture that seamlessly scales performance and capacity as needed.
- Get proven reliability, evidenced by over 20 years of and 1 million NetApp E-Series and EF-Series installations.

Maximize cost efficiency

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

The glue between the NetApp EF600 and the NVIDIA DGX A100™ Systems is the BeeGFS parallel file system. BeeGFS was originally created by the Fraunhofer Center for High-Performance Computing in Germany to solve the pain points of legacy parallel file systems. The result is a file system with a modern, user space architecture that is now developed and delivered by ThinkParQ, and used by many supercomputing environments. No more hacking away at kernels to get your parallel file system up and running. No more hardware vendor lock-in. No more paying for premium features that you don't need for AI workloads—and no more complicated pricing. You get just a blazing-fast HPC file system that's automated and integrated into the overall DGX SuperPOD experience.

The same software that's used to manage thousands of NVIDIA systems, Base Command Manager, is the best-in-class infrastructure solution for provisioning and lifecycle management, monitoring, telemetry, logging, alerting, and scheduling. NetApp EF600 and BeeGFS integration with Base Command Manager makes deploying and managing your DGX SuperPOD storage much simpler.

Future-proof your investment

From ingestion to processing, AI operations generate a lot of data. To manage and to 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 DGX SuperPOD, you get an agile AI solution that easily scales nondisruptively. BeeGFS enables your entire storage capacity to be served up in a single namespace—significantly reducing data management headaches.

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 and beyond by adding capacity, one building block 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. Reduce operating costs with high-density drives and price/performance-optimized storage building blocks. NetApp's world-class enterprise support organization supports the entire solution stack, so you never have to worry about outages or downtime.

Engage the specialists

NVIDIA and NetApp experts work together to help you every step along the way. Assistance with planning, design, performance projections, installation, post-installation testing, and ongoing

support are all part of the DGX SuperPOD experience. 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. And the NetApp support team is certified and trained to help with supporting any issues that might arise with your BeeGFS deployment. This fully integrated, turnkey architecture takes the risk out of deployment and puts you on the path to winning the race to AI leadership.

Solution components

- NVIDIA DGX SuperPOD (starting with 20 NVIDIA DGX A100 systems)
- NetApp EF600 all-flash storage
- NVIDIA Quantum QM8700 switches
- NVIDIA DGX software stack
- NVIDIA Base Command Manager
- ThinkParQ BeeGFS parallel file system

About NVIDIA

The invention of the GPU in 1999 by NVIDIA sparked the growth of the PC gaming market, redefined modern computer graphics and revolutionized parallel computing. More recently, GPU deep learning ignited modern AI—the next era of computing—with the GPU acting as the brain of computers, robots and self-driving cars that can perceive and understand the world. More information at www.nvidia.com.

About NetApp

In a world full of generalists, NetApp is a specialist. We're focused on one thing, helping your business get the most out of your data. NetApp brings the enterprise-grade data services you rely on into the cloud, and the simple flexibility of cloud into the data center. Our industry-leading solutions work across diverse customer environments and the world's biggest public clouds.

As a cloud-led, data-centric software company, only NetApp can help build your unique data fabric, simplify and connect your cloud, and securely deliver the right data, services and applications to the right people—anytime, anywhere. www.netapp.com

