# NetApp AIPod with NVIDIA

## NetApp Solutions

NetApp
August 14, 2024

# Table of Contents


NetApp AI with NVIDIA . . . 1
- NetApp AIPod with NVIDIA DGX Systems . . . 1
- NetApp ONTAP AI with NVIDIA DGX A100 Systems . . . 1
- NetApp ONTAP AI with NVIDIA DGX A100 Systems and Mellanox Spectrum Ethernet Switches . . . 1
- NetApp AIPod with NVIDIA DGX Systems - Introduction . . . 1
- NVA-1151-DESIGN: NetApp ONTAP AI with NVIDIA DGX A100 systems design guide . . . 11
- NVA-1151-DEPLOY: NetApp ONTAP AI with NVIDIA DGX A100 systems . . . 12
- NVA-1153-DESIGN: NetApp ONTAP AI with NVIDIA DGX A100 systems and Mellanox Spectrum Ethernet switches . . . 12
- NVA-1153-DEPLOY: NetApp ONTAP AI with NVIDIA DGX A100 systems and Mellanox Spectrum Ethernet switches . . . 12

# NetApp AI with NVIDIA

Overview of ONTAP AI converged infrastructure solutions from NetApp and NVIDIA.

## NetApp AIPod with NVIDIA DGX Systems

- NetApp AI Pod with NVIDIA DGX Systems

## NetApp ONTAP AI with NVIDIA DGX A100 Systems

- Design Guide
- Deployment Guide

## NetApp ONTAP AI with NVIDIA DGX A100 Systems and Mellanox Spectrum Ethernet Switches

- Design Guide
- Deployment Guide

## NetApp AIPod with NVIDIA DGX Systems - Introduction

This section provides an introduction to the NetApp AIPod with NVIDIA DGX systems.

NetApp Solution Engineering

The NetApp™ AIPod with NVIDIA DGX™ systems and NetApp cloud-connected storage systems, simplifies infrastructure deployments for machine learning (ML) and artificial intelligence (AI) workloads by eliminating design complexity and guesswork. Building on the NVIDIA DGX BasePOD design to deliver exceptional compute performance for next-generation workloads, AIPod with NVIDIA DGX systems adds NetApp AFF storage systems that allow customers to start small and grow non-disruptively while intelligently managing data from the edge to the core to the cloud and back. NetApp AIPod is part of the larger portfolio of NetApp AI solutions, show in the figure below-

*NetApp AI Solutions Portfolio*
image::aipod_nv_portfolio.png[]

This document describes the key components of the AIPod reference architecture, system connectivity information and solution sizing guidance. This document is intended for NetApp and partner solutions engineers and customer strategic decision makers interested in deploying a high-performance infrastructure for ML/DL and analytics workloads.

### NetApp AIPod with NVIDIA DGX Systems - Introduction

This section provides an introduction to the NetApp AIPod with NVIDIA DGX systems.

NetApp Solution Engineering

The NetApp™ AIPod with NVIDIA DGX™ systems and NetApp cloud-connected storage systems, simplifies infrastructure deployments for machine learning (ML) and artificial intelligence (AI) workloads by eliminating

design complexity and guesswork. Building on the NVIDIA DGX BasePOD design to deliver exceptional compute performance for next-generation workloads, AIPod with NVIDIA DGX systems adds NetApp AFF storage systems that allow customers to start small and grow non-disruptively while intelligently managing data from the edge to the core to the cloud and back. NetApp AIPod is part of the larger portfolio of NetApp AI solutions, show in the figure below-

*NetApp AI Solutions Portfolio*
image::aipod_nv_portfolio.png[]

This document describes the key components of the AIPod reference architecture, system connectivity information and solution sizing guidance. This document is intended for NetApp and partner solutions engineers and customer strategic decision makers interested in deploying a high-performance infrastructure for ML/DL and analytics workloads.

## NetApp AIPod with NVIDIA DGX Systems - Hardware Components

This section focuses on the hardware components for the NetApp AIPod with NVIDIA DGX systems.

### NetApp AFF Storage Systems

NetApp AFF state-of-the-art storage systems enable IT departments to meet enterprise storage requirements with industry-leading performance, superior flexibility, cloud integration, and best-in-class data management. Designed specifically for flash, AFF systems help accelerate, manage, and protect business-critical data.

#### AFF A900 storage systems

The NetApp AFF A900 powered by NetApp ONTAP data management software provides built-in data protection, optional anti-ransomware capabilities, and the high performance and resiliency required to support the most critical business workloads. It eliminates disruptions to mission-critical operations, minimizes performance tuning, and safeguards your data from ransomware attacks. It delivers:
• Industry-leading performance
• Uncompromised data security
• Simplified non-disruptive upgrades

*NetApp AFF A900 storage system*
image::aipod_nv_A900.png[]

### Industry-leading Performance

The AFF A900 easily manages next-generation workloads like deep learning, AI, and high-speed analytics as well as traditional enterprise databases like Oracle, SAP HANA, Microsoft SQL Server, and virtualized applications. It keeps business-critical applications running at top speed with up to 2.4M IOPS per HA pair and latency as low as 100µs—and increases performance by up to 50% over previous NetApp models. With NFS over RDMA, pNFS and Session Trunking, customers can achieve the high level of network performance required for next-generation applications using existing data center networking infrastructure.
Customers can also scale and grow with unified multi-protocol support for SAN, NAS, and Object storage and deliver maximum flexibility with unified and single ONTAP data management software, for data on-premises or in the cloud. In addition, system health can be optimized with AI-based predictive analytics delivered by Active IQ Digital Advisor (also known as Digital Advisor) and Cloud Insights.

### Uncompromised Data Security

AFF A900 systems contain a full suite of NetApp integrated and application-consistent data protection software. It provides built-in data protection and cutting-edge anti-ransomware solutions for pre-emption and post-attack recovery. Malicious files can be blocked from ever being written to disk, and storage abnormalities are easily monitored to gain insights.

### Simplified Non-Disruptive Upgrades

The AFF A900 is available as a non-disruptive in-chassis upgrade to existing A700 customers. NetApp makes it simple to refresh and eliminate disruptions to mission-critical operations through our advanced reliability, availability, serviceability, and manageability (RASM) capabilities. In addition, NetApp further increases operational efficiency and simplifies day-to-day activities for IT teams because ONTAP software automatically applies firmware updates for all system components.

For the largest deployments, AFF A900 systems offer the highest performance and capacity options while other NetApp storage systems, such as the AFF A800, AFF C800, AFF A400, AFF C400 and AFF A250 offer options for smaller deployments at lower cost points.

## NVIDIA DGX BasePOD

NVIDIA DGX BasePOD is an integrated solution consisting of NVIDIA hardware and software components, MLOps solutions, and third-party storage. Leveraging best practices of scale-out system design with NVIDIA products and validated partner solutions, customers can implement an efficient and manageable platform for AI development. Figure 1 highlights the various components of NVIDIA DGX BasePOD.

*NVIDIA DGX BasePOD solution*
image::aipod_nv_basepod_layers.png[]

#### NVIDIA DGX H100 Systems

The NVIDIA DGX H100™ system is the AI powerhouse that is accelerated by the groundbreaking performance of the NVIDIA H100 Tensor Core GPU.

*NVIDIA DGX H100 system*
image::aipod_nv_H100_3D.png[]

Key specifications of the DGX H100 system are:
• Eight NVIDIA H100 GPUs.
• 80 GB GPU memory per GPU, for a total of 640GB.
• Four NVIDIA NVSwitch™ chips.
• Dual 56-core Intel® Xeon® Platinum 8480 processors with PCIe 5.0 support.
• 2 TB of DDR5 system memory.
• Four OSFP ports serving eight single-port NVIDIA ConnectX-7 (InfiniBand/Ethernet) adapters, and two dual-port NVIDIA ConnectX-7 (InfiniBand/Ethernet) adapters.
• Two 1.92 TB M.2 NVMe drives for DGX OS, eight 3.84 TB U.2 NVMe drives for storage/cache.
• 10.2 kW max power.
The rear ports of the DGX H100 CPU tray are shown below. Four of the OSFP ports serve eight ConnectX-7 adapters for the InfiniBand compute fabric. Each pair of dual-port ConnectX-7 adapters provide parallel pathways to the storage and management fabrics. The out-of-band port is used for BMC access.

*NVIDIA DGX H100 rear panel*
image::aipod_nv_H100_rear.png[]

#### NVIDIA Networking

### NVIDIA Quantum-2 QM9700 Switch

*NVIDIA Quantum-2 QM9700 InfiniBand switch*
image::aipod_nv_QM9700.png[]

NVIDIA Quantum-2 QM9700 switches with 400Gb/s InfiniBand connectivity power the compute fabric in NVIDIA Quantum-2 InfiniBand BasePOD configurations. ConnectX-7 single-port adapters are used for the InfiniBand compute fabric. Each NVIDIA DGX system has dual connections to each QM9700 switch, providing multiple high-bandwidth, low-latency paths between the systems.

### NVIDIA Spectrum-3 SN4600 Switch

*NVIDIA Spectrum-3 SN4600 switch*
image::aipod_nv_SN4600_hires_smallest.png[]

NVIDIA Spectrum-3 SN4600 switches offer 128 total ports (64 per switch) to provide redundant connectivity for in-band management of the DGX BasePOD. The NVIDIA SN4600 switch can provide for speeds between 1 GbE and 200 GbE. For storage appliances connected over Ethernet, the NVIDIA SN4600 switches are also used. The ports on the NVIDIA DGX dual-port ConnectX-7 adapters are used for both in-band management and storage connectivity.

### NVIDIA Spectrum SN2201 Switch

*NVIDIA Spectrum SN2201 switch*
image::aipod_nv_SN2201.png[]

NVIDIA Spectrum SN2201 switches offer 48 ports to provide connectivity for out-of-band management. Out-of-band management provides consolidated management connectivity for all components in DGX BasePOD.

### NVIDIA ConnectX-7 Adapter

*NVIDIA ConnectX-7 adapter*
image::aipod_nv_CX7.png[]

The NVIDIA ConnectX-7 adapter can provide 25/50/100/200/400G of throughput. NVIDIA DGX systems use both the single and dual-port ConnectX-7 adapters to provide flexibility in DGX BasePOD deployments with 400Gb/s InfiniBand and 100/200Gb Ethernet.

## NetApp AIPod with NVIDIA DGX Systems - Software Components

This section focuses on the software components of the NetApp AIPod with NVIDIA DGX systems.

### NVIDIA Software

#### NVIDIA Base Command

NVIDIA Base Command™ powers every DGX BasePOD, enabling organizations to leverage the best of NVIDIA software innovation. Enterprises can unleash the full potential of their investment with a proven platform that includes enterprise-grade orchestration and cluster management, libraries that accelerate compute, storage and network infrastructure, and an operating system (OS) optimized for AI workloads.

*NVIDIA BaseCommand solution*
image::aipod_nv_BaseCommand_new.png[]

#### NVIDIA GPU Cloud (NGC)

NVIDIA NGC™ provides software to meet the needs of data scientists, developers, and researchers with various levels of AI expertise. Software hosted on NGC undergoes scans against an aggregated set of common vulnerabilities and exposures (CVEs), crypto, and private keys. It is tested and designed to scale to multiple GPUs and in many cases, to multi-node, ensuring users maximize their investment in DGX systems.

*NVIDIA GPU Cloud*
image::aipod_nv_ngc.png[]

#### NVIDIA AI Enterprise

NVIDIA AI Enterprise is the end-to-end software platform that brings generative AI into reach for every enterprise, providing the fastest and most efficient runtime for generative AI foundation models optimized to run on the NVIDIA DGX platform. With production-grade security, stability, and manageability, it streamlines the development of generative AI solutions. NVIDIA AI Enterprise is included with DGX BasePOD for enterprise developers to access pretrained models, optimized frameworks, microservices, accelerated libraries, and enterprise support.

### NetApp Software

#### NetApp ONTAP

ONTAP 9, the latest generation of storage management software from NetApp, enables businesses to modernize infrastructure and transition to a cloud-ready data center. Leveraging industry-leading data management capabilities, ONTAP enables the management and protection of data with a single set of tools, regardless of where that data resides. You can also move data freely to wherever it is needed: the edge, the core, or the cloud. ONTAP 9 includes numerous features that simplify data management, accelerate, and protect critical data, and enable next generation infrastructure capabilities across hybrid cloud architectures.

### Accelerate and protect data

ONTAP delivers superior levels of performance and data protection and extends these capabilities in the following ways:

- Performance and lower latency. ONTAP offers the highest possible throughput at the lowest possible latency, including support for NVIDIA GPUDirect Storage (GDS) using NFS over RDMA, parallel NFS (pNFS), and NFS session trunking.
- Data protection. ONTAP provides built-in data protection capabilities and the industry’s strongest anti-ransomware guarantee with common management across all platforms.
- NetApp Volume Encryption (NVE). ONTAP offers native volume-level encryption with both onboard and External Key Management support.
- Storage multitenancy and multifactor authentication. ONTAP enables sharing of infrastructure resources with the highest levels of security.

### Simplify data management

Data management is crucial to enterprise IT operations and data scientists so that appropriate resources are used for AI applications and training AI/ML datasets. The following additional information about NetApp technologies is out of scope for this validation but might be relevant depending on your deployment.

ONTAP data management software includes the following features to streamline and simplify operations and reduce your total cost of operation:

- Snapshots and clones enable collaboration, parallel experimentation and enhanced data governance for ML/DL workflows.
- SnapMirror enables seamless data movement in hybrid cloud and multi-site environments, delivering data where and when it's needed.
- Inline data compaction and expanded deduplication. Data compaction reduces wasted space inside storage blocks, and deduplication significantly increases effective capacity. This applies to data stored locally and data tiered to the cloud.
- Minimum, maximum, and adaptive quality of service (AQoS). Granular quality of service (QoS) controls help maintain performance levels for critical applications in highly shared environments.
- NetApp FlexGroups enable distribution of data across all nodes in the storage cluster providing massive capacity and higher performance for extremely large datasets.
- NetApp FabricPool. Provides automatic tiering of cold data to public and private cloud storage options, including Amazon Web Services (AWS), Azure, and NetApp StorageGRID storage solution. For more information about FabricPool, see TR-4598: FabricPool best practices.
- NetApp FlexCache. Provides remote volume caching capabilities that simplify file distribution, reduces WAN latency, and lowers WAN bandwidth costs. FlexCache enables distributed product development across multiple sites, as well as accelerated access to corporate datasets from remote locations.

### Future-proof infrastructure

ONTAP helps meet demanding and constantly changing business needs with the following features:

- Seamless scaling and non disruptive operations. ONTAP supports the online addition of capacity to existing controllers and to scale-out clusters. Customers can upgrade to the latest technologies, such as NVMe and 32Gb FC, without costly data migrations or outages.
- Cloud connection. ONTAP is the most cloud-connected storage management software, with options for software-defined storage (ONTAP Select) and cloud-native instances (NetApp Cloud Volumes Service) in all public clouds.
- Integration with emerging applications. ONTAP offers enterprise-grade data services for next generation platforms and applications, such as autonomous vehicles, smart cities, and Industry 4.0, by using the same infrastructure that supports existing enterprise apps.

#### NetApp DataOps Toolkit

The NetApp DataOps Toolkit is a Python-based tool that simplifies the management of development/training workspaces and inference servers that are backed by high-performance, scale-out NetApp storage. The DataOps Toolkit can operate as a stand-alone utility, and is even more effective in Kubernetes environments leveraging NetApp Astra Trident to automate storage operations. Key capabilities include:

- Rapidly provision new high-capacity JupyterLab workspaces that are backed by high-performance, scale-out NetApp storage.
- Rapidly provision new NVIDIA Triton Inference Server instances that are backed by enterprise-class NetApp storage.
- Near-instantaneous cloning of high-capacity JupyterLab workspaces in order to enable experimentation or rapid iteration.
- Near-instantaneous snapshots of high-capacity JupyterLab workspaces for backup and/or traceability/baselining.

- Near-instantaneous provisioning, cloning, and snapshots of high-capacity, high-performance data volumes.

#### NetApp Astra Trident

Astra Trident is a fully supported, open-source storage orchestrator for containers and Kubernetes distributions, including Anthos. Trident works with the entire NetApp storage portfolio, including NetApp ONTAP, and it also supports NFS, NVMe/TCP, and iSCSI connections. Trident accelerates the DevOps workflow by allowing end users to provision and manage storage from their NetApp storage systems without requiring intervention from a storage administrator.

## NetApp AIPod with NVIDIA DGX Systems - Solution Architecture

This section focuses on the architecture for the NetApp AIPod with NVIDIA DGX systems.

### NetApp AI Pod with DGX H100 systems

This reference architecture leverages separate fabrics for compute cluster interconnect and storage access, with 400Gb/s InfiniBand (IB)connectivity between compute nodes. The drawing below shows the overall solution topology of NetApp AIPod with DGX H100 systems.

*NetApp AIpod solution topology*
image::aipod_nv_a900topo.png[]

### Network configuration

In this configuration the compute cluster fabric uses a pair of QM9700 400Gb/s IB switches, which are connected together for high availability. Each DGX H100 system is connected to the switches using eight connections, with even-numbered ports connected to one switch and odd-numbered ports connected to the other switch.

For storage system access, in-band management and client access, a pair of SN4600 Ethernet switches is used. The switches are connected with inter-switch links and configured with multiple VLANs to isolate the various traffic types. For larger deployments the Ethernet network can be expanded to a leaf-spine configuration by adding additional switch pairs for spine switches and additional leaves as needed.

In addition to the compute interconnect and high-speed Ethernet networks, all of the physical devices are also connected to one or more SN2201 Ethernet switches for out of band management. For more details on DGX H100 system connectivity please refer to the NVIDIA BasePOD documentation.

### Client configuration for storage access

Each DGX H100 system is provisioned with two dual-ported ConnectX-7 adapters for management and storage traffic, and for this solution both ports on each card are connected to the same switch. One port from each card is then configured into a LACP MLAG bond with one port connected to each switch, and VLANs for in-band management, client access, and user-level storage access are hosted on this bond.

The other port on each card is used for connectivity to the AFF A900 storage systems, and can be used in several configurations depending on workload requirements. For configurations using NFS over RDMA to support NVIDIA Magnum IO GPUDirect Storage, the ports are configured in an active/passive bond, as RDMA is not supported on any other type of bond. For deployments that do not require RDMA, the storage interfaces can also be configured with LACP bonding to deliver high availability and additional bandwidth. With or without RDMA, clients can mount the storage system using NFS v4.1 pNFS and Session trunking to enable parallel access to all storage nodes in the cluster.

### Storage system configuration

Each AFF A900 storage system is connected using four 100 GbE ports from each controller. Two ports from each controller are used for workload data access from the DGX systems, and two ports from each controller are configured as an LACP interface group to support access from the management plane servers for cluster management artifacts and user home directories. All data access from the storage system is provided through NFS, with a storage virtual machine (SVM) dedicated to AI workload access and a separate SVM dedicated to cluster management uses.

The workload SVM is configured with a total of eight logical interfaces (LIFs), with two LIFs on each physical port. This configuration provides maximum bandwidth as well as the means for each LIF to fail over to another port on the same controller, so that both controllers stay active in the event of a network failure. This configuration also supports NFS over RDMA to enable GPUDirect Storage access. Storage capacity is provisioned in the form of a single large FlexGroup volume that spans all storage controllers in the cluster, with 16 constituent volumes on each controller. This FlexGroup is accessible from any of the LIFs on the SVM, and by using NFSv4.1 with pNFS and session trunking, clients establish connections to every LIF in the SVM, enabling data local to each storage node to be accessed in parallel for significantly improved performance. The workload SVM and each data LIF are also configured for RDMA protocol access. For more details on RDMA configuration for ONTAP please refer to the ONTAP documentation.

The management SVM only requires a single LIF, which is hosted on the 2-port interface groups configured on each controller. Other FlexGroup volumes are provisioned on the management SVM to house cluster management artifacts like cluster node images, system monitoring historical data, and end-user home directories. The drawing below shows the logical configuration of the storage system.

*NetApp A900 storage cluster logical configuration*
image::aipod_nv_A900logical.png[]

### Management plane servers

This reference architecture also includes five CPU-based servers for management plane uses. Two of these systems are used as the head nodes for NVIDIA Base Command Manager for cluster deployment and management. The other three systems are used to provide additional cluster services such as Kubernetes master nodes or login nodes for deployments utilizing Slurm for job scheduling. Deployments utilizing Kubernetes can leverage the NetApp Astra Trident CSI driver to provide automated provisioning and data services with persistent storage for both management and AI workloads on the AFF A900 storage system.

Each server is physically connected to both the IB switches and Ethernet switches to enable cluster deployment and management, and configured with NFS mounts to the storage system via the management SVM for storage of cluster management artifacts as described earlier.

## NetApp AIPod with NVIDIA DGX Systems - Solution Validation and Sizing Guidance

This section focuses on the solution validation and sizing guidance for the NetApp AIPod with NVIDIA DGX systems.

### Solution Validation

The storage configuration in this solution was validated using a series of synthetic workloads using the open-source tool FIO. These tests include read and write I/O patterns intended to simulate the storage workload generated by DGX systems performing deep learning training jobs. The storage configuration was validated using a cluster of 2-socket CPU servers running the FIO workloads concurrently to simulate a cluster of DGX systems. Each client was configured with the same network configuration described previously, with the addition of the following details.

The following mount options were used for this validation:

| | |
|---|---|
| vers=4.1 | enables pNFS for parallel access to multiple storage nodes |
| proto=rdma | sets the transfer protocol to RDMA instead of the default TCP |
| port=20049 | specify the correct port for the RDMA NFS service |
| max_connect=16 | enables NFS session trunking to aggregate storage port bandwidth |
| write=eager | improves write performance of buffered writes |
| rsize=262144,wsize=262144 | sets the I/O transfer size to 256k |

In addition the clients were configured with an NFS max_session_slots value of 1024. As the solution was tested using NFS over RDMA, the storage networks ports were configured with an active/passive bond. The following bond parameters were used for this validation:

| | |
|---|---|
| mode=active-backup | sets the bond to active/passive mode |
| primary=<interface name> | primary interfaces for all clients were distributed across the switches |
| mii-monitor-interval=100 | specifies monitoring interval of 100ms |
| fail-over-mac-policy=active | specifies that the MAC address of the active link is the MAC of the bond. This is required for proper operation of RDMA over the bonded interface. |

The storage system was configured as described with two A900 HA pairs (4 controllers) with two NS224 disk shelves of 24 1.9TB NVMe disk drives attached to each HA pair. As noted in the architecture section, storage capacity from all controllers was combined using a FlexGroup volume, and data from all clients was distributed across all the controllers in the cluster.

**Storage System Sizing Guidance**

NetApp has successfully completed the DGX BasePOD certification, and the two A900 HA pairs as tested can easily support a cluster of eight DGX H100 systems. For larger deployments with higher storage performance requirements, additional AFF systems can be added to the NetApp ONTAP cluster up to 12 HA pairs (24 nodes) in a single cluster. Using the FlexGroup technology described in this solution, a 24-node cluster can provide over 40 PB and up to 300 GBps throughput in a single namespace. Other NetApp storage systems such as the AFF A400, A250 and C800 offer lower performance and/or higher capacity options for smaller deployments at lower cost points. Because ONTAP 9 supports mixed-model clusters, customers can start with a smaller initial footprint and add more or larger storage systems to the cluster as capacity and performance requirements grow. The table below shows a rough estimate of the number of A100 and H100 GPUs supported on each AFF model.

*NetApp storage system sizing guidance*

| | | Throughput[2] | Raw capacity *(typical / max)* | Connectivity | # NVIDIA A100 GPUs supported[3] | # NVIDIA H100 GPUs supported[4] |
|---|---|---|---|---|---|---|
| **NetApp® AFF A900** | **1 HA pair[1]** | 28GB/s | 182TB / 14.7PB | **100 GbE** | 1 - 64 | 1-32 |
| | **12 HA pairs** | 336GB/s | 2.1PB / **176.4PB** | | **768** | **384** |
| **AFF A800** | **1 HA pair** | 25GB/s | 368TB / 3.6PB | **100 GbE** | 1 - 64 | 1-32 |
| | **12 HA pairs** | 300GB/s | 4.4PB / **43.2PB** | | **768** | **384** |
| **AFF C800** | **1 HA pair** | 21GB/s | 368TB / 3.6PB | **100 GbE** | 1-48 | 1-24 |
| | **12 HA pairs** | 252GB/s | 4.4PB / **43.2PB** | | **576** | **288** |
| **AFF A400** | **1 HA pair** | 11GB/s | 182TB / 14.7PB | **40/100 GbE** | 1 - 32 | 1-16 |
| | **12 HA pairs** | 132GB/s | 2.1PB / **176.4PB** | | **384** | **192** |
| **AFF C400** | **1 HA pair** | 8GB/s | 182TB / 14.7PB | **40/100 GbE** | 1 - 16 | 1-8 |
| | **12 HA pairs** | 128GB/s | 2.1PB / **176.4PB** | | **192** | **96** |
| **AFF A250** | **1 HA pair** | 7.4GB/s | 91.2TB / 4.4PB | **25 GbE 40/100GbE** | 1 - 16 | 1-8 |
| | **4 HA pairs** | 29.6GB/s | 364.8TB / 17.6PB | | 64 | 32 |
| **AFF C250** | **1 HA pair** | 5 GB/s | 91.2TB / 4.4PB | **25 GbE 40/100GbE** | 1-8 | 1-4 |
| | **4 HA pairs** | 20 GB/s | 364.8TB / 17.6PB | | 32 | 8 |

1 – 1 AFF = 1 HA pair = 2 Nodes. 12 HA pairs = 24 nodes
2 – 100% sequential read
3 – Based on workload testing in NVA-1153
4 – Based on BasePOD validation test results

## NetApp AIPod with NVIDIA DGX Systems - Conclusion and Additional Information

This section includes references for additional information for the NetApp AIPod with NVIDIA DGX systems.

### Conclusion

The DGX BasePOD architecture is a next-generation deep learning platform that requires equally advanced storage and data management capabilities. By combining DGX BasePOD with NetApp AFF systems, the NetApp AIPod with DGX systems architecture can be implemented at almost any scale up to 48 DGX H100 systems on a 24-node AFF A900 cluster. Combined with the superior cloud integration and software-defined capabilities of NetApp ONTAP, AFF enables a full range of data pipelines that spans the edge, the core, and the cloud for successful DL projects.

### Additional Information

To learn more about the information described in this document, please refer to the following documents and/or websites:

- NetApp ONTAP data management software — ONTAP information library

  https://docs.netapp.com/us-en/ontap-family/

- NetApp AFF A900 storage systems-

  https://www.netapp.com/data-storage/aff-a-series/aff-a900/

- NetApp ONTAP RDMA information-

https://docs.netapp.com/us-en/ontap/nfs-rdma/index.html

- NetApp DataOps Toolkit

  https://github.com/NetApp/netapp-dataops-toolkit

- NetApp Astra Trident

  Overview

- NetApp GPUDirect Storage Blog-

  https://www.netapp.com/blog/ontap-reaches-171-gpudirect-storage/

- NVIDIA DGX BasePOD

  https://www.nvidia.com/en-us/data-center/dgx-basepod/

- NVIDIA DGX H100 systems

  https://www.nvidia.com/en-us/data-center/dgx-h100/

- NVIDIA Networking

  https://www.nvidia.com/en-us/networking/

- NVIDIA Magnum IO GPUDirect Storage

  https://docs.nvidia.com/gpudirect-storage

- NVIDIA Base Command

  https://www.nvidia.com/en-us/data-center/base-command/

- NVIDIA Base Command Manager

  https://www.nvidia.com/en-us/data-center/base-command/manager

- NVIDIA AI Enterprise

  https://www.nvidia.com/en-us/data-center/products/ai-enterprise/

**Acknowledgements**

This document is the work of the NetApp Solutions and ONTAP Engineering teams- David Arnette, Olga Kornievskaia, Dustin Fischer, Srikanth Kaligotla, Mohit Kumar and Rajeev Badrinath. The authors would also like to thank NVIDIA and the NVIDIA DGX BasePOD engineering team for their continued support.

# NVA-1151-DESIGN: NetApp ONTAP AI with NVIDIA DGX A100 systems design guide

David Arnette and Sung-Han Lin, NetApp

NVA-1151-DESIGN describes a NetApp Verified Architecture for machine learning and

artificial intelligence workloads using NetApp AFF A800 storage systems, NVIDIA DGX A100 systems, and NVIDIA Mellanox network switches. It also includes benchmark test results for the architecture as implemented.

NVA-1151-DESIGN: NetApp ONTAP AI with NVIDIA DGX A100 systems design guide

## NVA-1151-DEPLOY: NetApp ONTAP AI with NVIDIA DGX A100 systems

David Arnette, NetApp

NVA-1151-DEPLOY includes storage system deployment instructions for a NetApp Verified Architecture (NVA) for machine learning (ML) and artificial intelligence (AI) workloads using NetApp AFF A800 storage systems, NVIDIA DGX A100 systems, and NVIDIA Mellanox network switches. It also includes instructions for running validation benchmark tests after deployment is complete.

NVA-1151-DEPLOY: NetApp ONTAP AI with NVIDIA DGX A100 systems

## NVA-1153-DESIGN: NetApp ONTAP AI with NVIDIA DGX A100 systems and Mellanox Spectrum Ethernet switches

David Arnette and Sung-Han Lin, NetApp

NVA-1153-DESIGN describes a NetApp Verified Architecture for machine learning (ML) and artificial intelligence (AI) workloads using NetApp AFF A800 storage systems, NVIDIA DGX A100 systems, and NVIDIA Mellanox Spectrum SN3700V 200Gb Ethernet switches. This design features RDMA over Converged Ethernet (RoCE) for the compute cluster interconnect fabric to provide customers with a completely ethernet-based architecture for high-performance workloads. This document also includes benchmark test results for the architecture as implemented.

NVA-1153-DESIGN: NetApp ONTAP AI with NVIDIA DGX A100 systems and Mellanox Spectrum Ethernet switches

## NVA-1153-DEPLOY: NetApp ONTAP AI with NVIDIA DGX A100 systems and Mellanox Spectrum Ethernet switches

David Arnette, NetApp

NVA-1153-DEPLOY includes storage-system deployment instructions for a NetApp Verified Architecture for machine learning (ML) and artificial intelligence (AI) workloads using NetApp AFF A800 storage systems, NVIDIA DGX A100 systems, and NVIDIA Mellanox Spectrum SN3700V 200Gb Ethernet switches. It also includes instructions for executing validation benchmark tests after deployment is complete.

NVA-1153-DEPLOY: NetApp ONTAP AI with NVIDIA DGX A100 systems and Mellanox Spectrum Ethernet

switches

## Copyright information



## Trademark information