AI Converged Infrastructures

NetApp Solutions

NetApp
July 31, 2024

Table of Contents

AI Converged Infrastructures

NetApp AFF A400 with Lenovo ThinkSystem SR670 V2 for AI and ML Model Training

TR-4810: NetApp AFF A400 with Lenovo ThinkSystem SR670 V2 for AI and ML Model Training

Sathish Thyagarajan, David Arnette, NetApp
Mircea Troaca, Lenovo

This solution presents a mid-range cluster architecture using NetApp storage and Lenovo servers optimized for artificial intelligence (AI) workloads. It is meant for small- to medium-sized enterprises for which most compute jobs are single node (single or multi-GPU) or distributed over a few computational nodes. This solution aligns with most day-to-day AI training jobs for many businesses.

This document covers testing and validation of a compute and storage configuration consisting of eight-GPU Lenovo SR670V2 servers, a mid-range NetApp AFF A400 storage system and 100GbE interconnect switch. To measure the performance, we used ResNet50 with the ImageNet dataset, a batch size of 408, half precision, CUDA, and cuDNN. This architecture provides an efficient and cost-effective solution for small and medium-sized organizations just starting out with AI initiatives that require the enterprise-grade capabilities of NetApp ONTAP cloud-connected data storage.

Target audience

This document is intended for the following audiences:

- Data scientists, data engineers, data administrators, and developers of AI systems
- Enterprise architects who design solutions for the development of AI models
- Data scientists and data engineers who are looking for efficient ways to achieve deep learning (DL) and machine learning (ML) development goals
- Business leaders and OT/IT decision makers who want to achieve the fastest possible time to market for AI initiatives

Solution architecture

This solution with Lenovo ThinkSystem servers and NetApp ONTAP with AFF storage is designed to handle AI training on large datasets using the processing power of GPUs alongside traditional CPUs. This validation demonstrates high performance and optimal data management with a scale-out architecture that uses either one, two, or four Lenovo SR670 V2 servers alongside a single NetApp AFF A400 storage system. The following figure provides an architectural overview.

This NetApp and Lenovo solution offers the following key benefits:

- Highly efficient and cost-effective performance when executing multiple training jobs in parallel
- Scalable performance based on different numbers of Lenovo servers and different models of NetApp storage controllers

- Robust data protection to meet low recovery point objectives (RPOs) and recovery time objectives (RTOs) with no data loss
- Optimized data management with snapshots and clones to streamline development workflows

Technology overview

This section introduces the major components of this solution in greater detail.

NetApp AFF systems

NetApp AFF storage systems enable businesses 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.
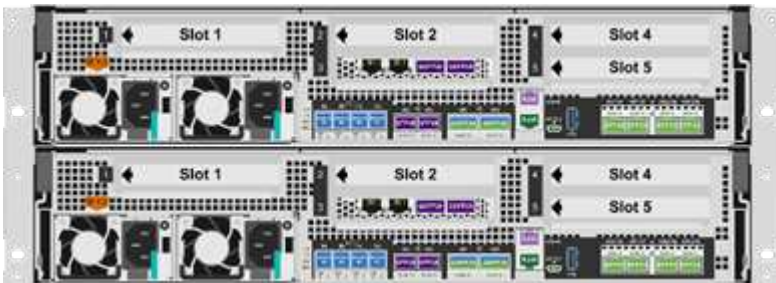
NetApp AFF A400 is a mid-range NVMe flash storage system that includes the following features:

- Maximum effective capacity: ~20PB
- Maximum scale-out: 2-24 nodes (12 HA pairs)
- 25GbE and 16Gb FC host support
- 100GbE RDMA over Converged Ethernet (RoCE) connectivity to NVMe expansion storage shelves
- 100GbE RoCE ports can be used for host network attachment if NVMe shelves aren't attached
- Full 12Gbps SAS connectivity expansion storage shelves
- Available in two configurations:
  - Ethernet: 4x 25Gb Ethernet (SFP28) ports
  - Fiber Channel: 4x 16Gb FC (SFP+) ports
- 100% 8KB random read @.4 ms 400k IOPS

NetApp AFF A250 features for entry level AI/ML deployments include the following:

- Maximum effective capacity: 35PB
- Maximum scale out: 2-24 nodes (12 HA pairs)

- 440k IOPS random reads @1ms
- Built on the latest NetApp ONTAP release ONTAP 9.8 or later
- Two 25Gb Ethernet ports for HA and cluster interconnect

NetApp also offers other storage systems, such as the AFF A800 and AFF A700 that provide higher performance and scalability for larger-scale AI/ML deployments.

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. Data can also be moved freely to wherever it's needed: the edge, the core, or the cloud. ONTAP 9 includes numerous features that simplify data management, accelerate and protect critical data, and future-proof infrastructure across hybrid cloud architectures.

Simplify data management

Data management is crucial to enterprise IT operations so that appropriate resources are used for applications and datasets. ONTAP includes the following features to streamline and simplify operations and reduce the total cost of operation:

- **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 (QoS).** Granular QoS controls help maintain performance levels for critical applications in highly shared environments.
- **ONTAP FabricPool.** This feature automatically tiers cold data to public and private cloud storage options, including Amazon Web Services (AWS), Azure, and NetApp StorageGRID object storage.

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.
- **Data protection.** ONTAP provides built-in data protection capabilities with common management across all platforms.
- **NetApp Volume Encryption.** ONTAP offers native volume-level encryption with both onboard and external key management support.

Future-proof infrastructure

ONTAP 9 helps meet demanding and constantly changing business needs:

- **Seamless scaling and nondisruptive operations.** ONTAP supports the nondisruptive addition of capacity to existing controllers as well as 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 OpenStack, Hadoop, and MongoDB by using the same infrastructure that supports existing enterprise apps.

NetApp FlexGroup volumes

Training datasets are typically a collection of potentially billions of files. Files can include text, audio, video, and other forms of unstructured data that must be stored and processed to be read in parallel. The storage system must store many small files and must read those files in parallel for sequential and random I/O.

A FlexGroup volume (the following figure) is a single namespace made up of multiple constituent member volumes that is managed and acts like a NetApp FlexVol volume to storage administrators. Files in a FlexGroup volume are allocated to individual member volumes and are not striped across volumes or nodes. They enable the following capabilities:

- Up to 20 petabytes of capacity and predictable low latency for high-metadata workloads

- Up to 400 billion files in the same namespace

- Parallelized operations in NAS workloads across CPUs, nodes, aggregates, and constituent FlexVol volumes

image::a400-thinksystem-image5.png["This image depicts an HA-pair of storage controllers containing many volumes with main files within a FlexGroup.]"

Lenovo ThinkSystem portfolio

Lenovo ThinkSystem servers feature innovative hardware, software, and services that solve customers' challenges today and deliver an evolutionary, fit-for-purpose, modular design approach to address tomorrow's challenges. These servers capitalize on best-in-class, industry-standard technologies coupled with differentiated Lenovo innovations to provide the greatest possible flexibility in x86 servers.

Key advantages of deploying Lenovo ThinkSystem servers include the following:

- Highly scalable, modular designs that grow with your business

- Industry-leading resilience to save hours of costly unscheduled downtime

- Fast flash technologies for lower latencies, quicker response times, and smarter data management in real time

In the AI area, Lenovo is taking a practical approach to helping enterprises understand and adopt the benefits of ML and AI for their workloads. Lenovo customers can explore and evaluate Lenovo AI offerings in Lenovo AI Innovation Centers to fully understand the value for their particular use case. To improve time to value, this customer-centric approach gives customers proofs of concept for solution development platforms that are ready to use and optimized for AI.

Lenovo SR670 V2

The Lenovo ThinkSystem SR670 V2 rack server delivers optimal performance for accelerated AI and high-performance computing (HPC). Supporting up to eight GPUs, the SR670 V2 is suited for the computationally intensive workload requirements of ML, DL, and inference.

4x SXM GPUs with 8x 2.5-inch HS drives and 2x PCIe I/O slots

4x double-wide or 8x single-wide GPU slots and 2x PCIe I/O slots
with 8x 2.5-inch or 4x 3.5-inch HS drives
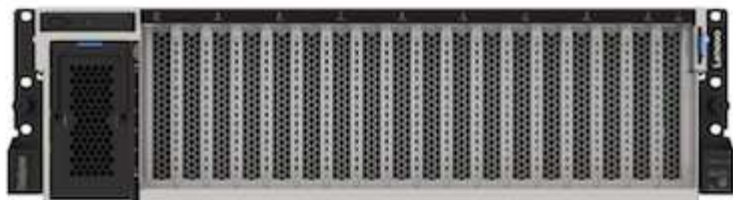
8x double-wide GPU slots with 6x EDSFF HS drives and 2x PCIe I/O slots

With the latest scalable Intel Xeon CPUs that support high-end GPUs (including the NVIDIA A100 80GB PCIe 8x GPU), the ThinkSystem SR670 V2 delivers optimized, accelerated performance for AI and HPC workloads.

Because more workloads use the performance of accelerators, the demand for GPU density has increased. Industries such as retail, financial services, energy, and healthcare are using GPUs to extract greater insights and drive innovation with ML, DL, and inference techniques.

The ThinkSystem SR670 V2 is an optimized, enterprise-grade solution for deploying accelerated HPC and AI workloads in production, maximizing system performance while maintaining data center density for supercomputing clusters with next-generation platforms.

Other features include:

- Support for GPU direct RDMA I/O in which high-speed network adapters are directly connected to the GPUs to maximize I/O performance.
- Support for GPU direct storage in which NVMe drives are directly connected to the GPUs to maximize storage performance.

MLPerf

MLPerf is the industry-leading benchmark suite for evaluating AI performance. In this validation, we used its image-classification benchmark with MXNet, one of the most popular AI frameworks. The MXNet_benchmarks training script was used to drive AI training. The script contains implementations of several popular conventional models and is designed to be as fast as possible. It can be run on a single machine or run in distributed mode across multiple hosts.

Test plan

In this validation, we performed image recognition training as specified by MLPerf v2.0.

Specifically, we trained the ResNet v2.0 model with the ImageNet dataset until we reached an accuracy of 76.1%. The main metric is the time to reach the desired accuracy. We also report training bandwidth in images per second to better judge scale-out efficiency.

The primary test case evaluated multiple independent training processes (one per node) running concurrently. This simulates the main use case, a shared system used by multiple data scientists. The second test case evaluated scale-out efficiency.

Test results

The following table summarizes the results for all tests performed for this solution.

Test description	Results summary
Image recognition training: multiple concurrent jobs	Highly efficient performance. All jobs ran at full speed even when the cluster was fully used. The NetApp storage systems delivered training performance comparable to local SSD storage while enabling easy sharing of data between servers.
Image recognition training: scale out	Highly efficient for up to four nodes. At that point, scale out was less efficient but still feasible. Using a higher-speed computational network improves scalability. The NetApp storage system delivered training performance comparable to local SSD storage while enabling easy sharing of data between servers.

Test configuration

This section describes the tested configurations, the network infrastructure, the SR670 V2 server, and the NetApp storage provisioning details.

Solution architecture

We used the solution components listed in the following table for this validation.

Solution components	Details
Lenovo ThinkSystem servers	• Two SR670 V2 servers each with eight NVIDIA A100 80GB GPU cards • Each server contains 2 Intel Xeon Platinum 8360Y CPUs (28 physical cores) and 1TB RAM
Linux (Ubuntu – 20.04 with CUDA 11.8)	

Solution components	Details
NetApp AFF storage system (HA pair)	• NetApp ONTAP 9.10.1 software • 24x 960GB SSDs • NFS protocol • 1 interface group (ifgrp) per controller, with four logical IP addresses for mount points

In this validation, we used ResNet v2.0 with the ImageNet basis set as specified by MLPerf v2.0. The dataset is stored in a NetApp AFF storage system with the NFS protocol. The SR670s were connected to the NetApp AFF A400 storage system over a 100GbE switch.

ImageNet is a frequently used image dataset. It contains almost 1.3 million images for a total size of 144GB. The average image size is 108KB.

The following figure depicts the network topology of the tested configuration.
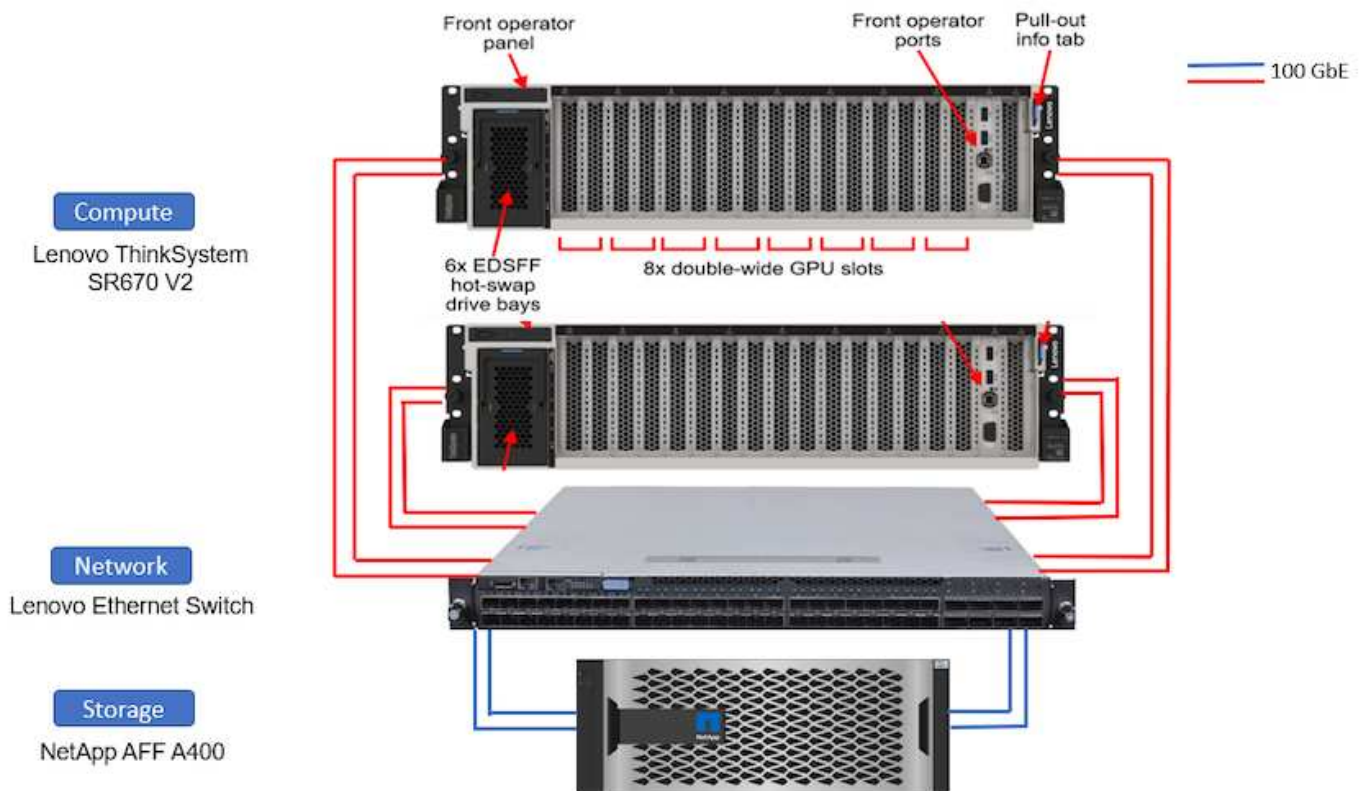
Storage controller

The following table lists the storage configuration.

Controller	Aggregate	FlexGroup volume	Aggregate size	Volume size	Operating system mount point
Controller1	Aggr1	/a400-100g	9.9TB	19TB	/a400-100g
Controller2	Aggr2	/a400-100g	9.9TB		/a400-100g

ⓘ   The /a400-100g folder contains the dataset used for ResNet validation.

Test procedure and detailed results

This section describes the detailed test procedure results.

Image recognition training using ResNet in ONTAP

We ran the ResNet50 benchmark with one and two SR670 V2 servers. This test used the MXNet 22.04-py3 NGC container to run the training.

We used the following test procedure in this validation:

1. We cleared the host cache before running the script to make sure that data was not already cached:

```
sync ; sudo /sbin/sysctl vm.drop_caches=3
```

2. We ran the benchmark script with the ImageNet dataset in server storage (local SSD storage) as well as on the NetApp AFF storage system.

3. We validated network and local storage performance using the dd command.

4. For the single-node run, we used the following command:

```
python train_imagenet.py --gpus 0,1,2,3,4,5,6,7 --batch-size 408 --kv
-store horovod --lr 10.5 --mom 0.9 --lr-step-epochs pow2 --lars-eta
0.001 --label-smoothing 0.1 --wd 5.0e-05 --warmup-epochs 2 --eval-period
4 --eval-offset 2 --optimizer sgdwfastlars --network resnet-v1b-stats-fl
--num-layers 50 --num-epochs 37 --accuracy-threshold 0.759 --seed 27081
--dtype float16 --disp-batches 20 --image-shape 4,224,224 --fuse-bn-relu
1 --fuse-bn-add-relu 1 --bn-group 1 --min-random-area 0.05 --max-random
-area 1.0 --conv-algo 1 --force-tensor-core 1 --input-layout NHWC --conv
-layout NHWC --batchnorm-layout NHWC --pooling-layout NHWC --batchnorm
-mom 0.9 --batchnorm-eps 1e-5 --data-train /data/train.rec --data-train
-idx /data/train.idx --data-val /data/val.rec --data-val-idx
/data/val.idx --dali-dont-use-mmap 0 --dali-hw-decoder-load 0 --dali
-prefetch-queue 5 --dali-nvjpeg-memory-padding 256 --input-batch
-multiplier 1 --dali- threads 6 --dali-cache-size 0 --dali-roi-decode 1
--dali-preallocate-width 5980 --dali-preallocate-height 6430 --dali-tmp
-buffer-hint 355568328 --dali-decoder-buffer-hint 1315942 --dali-crop
-buffer-hint 165581 --dali-normalize-buffer-hint 441549 --profile 0
--e2e-cuda-graphs 0 --use-dali
```

5. For the distributed runs, we used the parameter server's parallelization model. We used two parameter servers per node, and we set the number of epochs to be the same as for the single-node run. We did this because distributed training often takes more epochs due to imperfect synchronization between processes. The different number of epochs can skew comparisons between single-node and distributed cases.

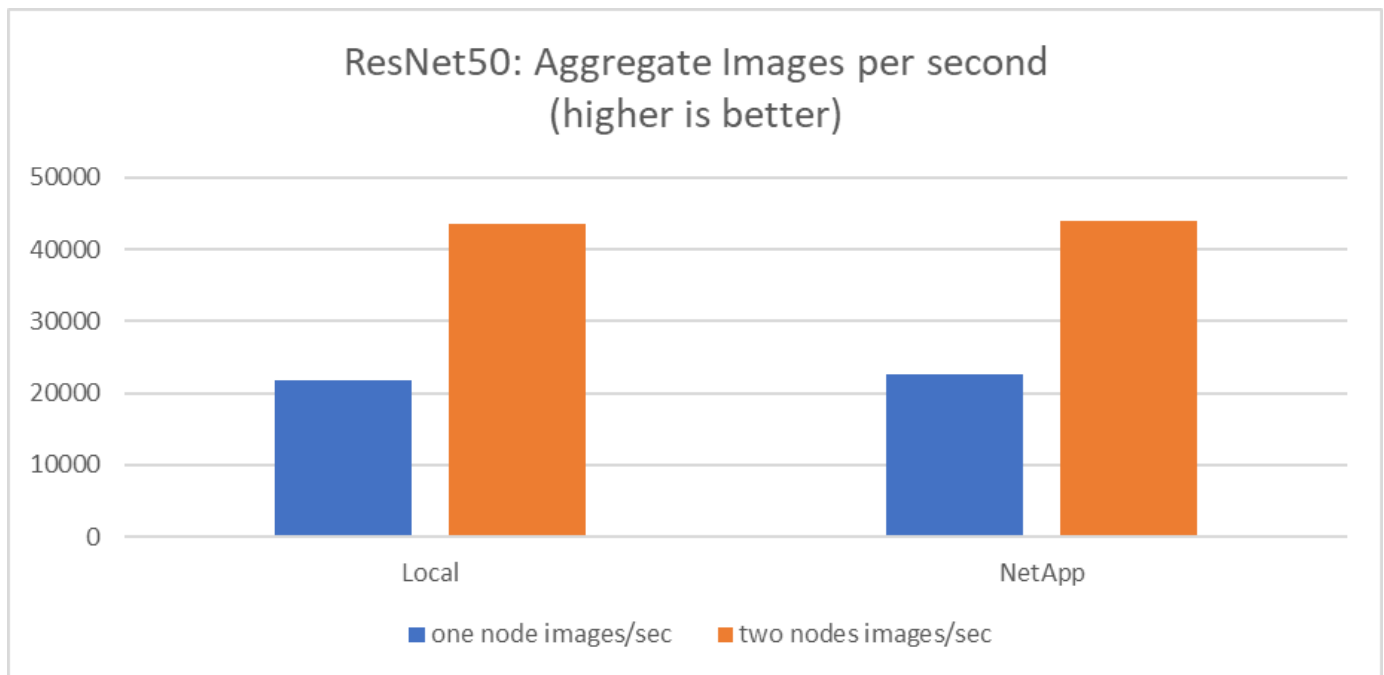**Data read speed: Local versus network storage**

The read speed was tested by using the `dd` command on one of the files for the ImageNet dataset. Specifically, we ran the following commands for both local and network data:
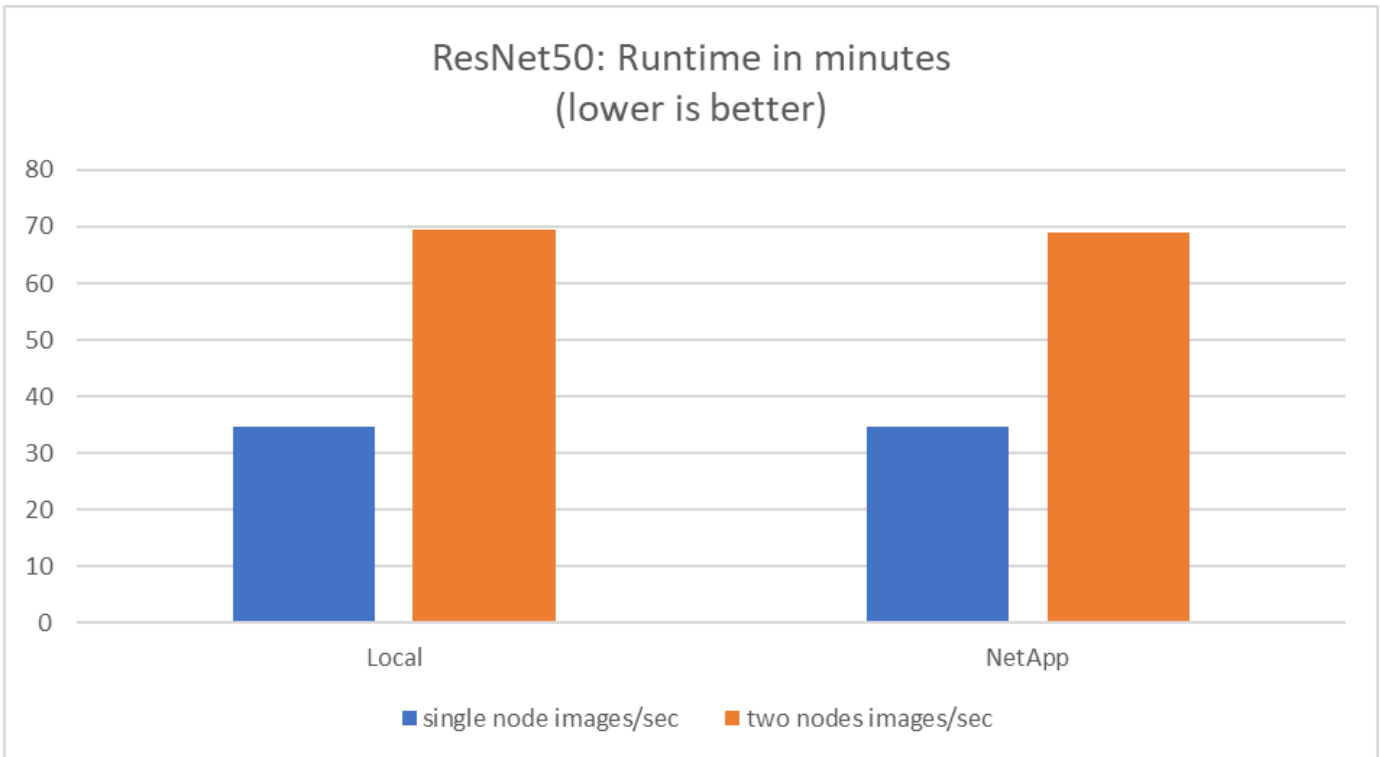
```
sync ; sudo /sbin/sysctl vm.drop_caches=3dd if=/a400-100g/netapp-
ra/resnet/data/preprocessed_data/train.rec of=/dev/null bs=512k
count=2048Results (average of 5 runs):
Local storage: 1.7 GB/s Network storage: 1.5 GB/s.
```

Both values are similar, demonstrating that the network storage can deliver data at a rate similar to local storage.

Shared use case: Multiple, independent, simultaneous jobs

This test simulated the expected use case for this solution: multi-job, multi-user AI training. Each node ran its own training while using the shared network storage. The results are displayed in the following figure, which shows that the solution case provided excellent performance with all jobs running at essentially the same speed as individual jobs. The total throughput scaled linearly with the number of nodes.

ResNet50: Runtime in minutes
(lower is better)

Legend: ■ single node images/sec ■ two nodes images/sec

These graphs present the runtime in minutes and the aggregate images per second for compute nodes that used eight GPUs from each server on 100 GbE client networking, combining both the concurrent training model and the single training model. The average runtime for the training model was 35 minutes and 9 seconds. The individual runtimes were 34 minutes and 32 seconds, 36 minutes and 21 seconds, 34 minutes and 37 seconds, 35 minutes and 25 seconds, and 34 minutes and 31 seconds. The average images per second for the training model were 22,573, and the individual images per second were 21,764; 23,438; 22,556; 22,564; and 22,547.

Based on our validation, one independent training model with a NetApp data runtime was 34 minutes and 54 seconds with 22,231 images/sec. One independent training model with a local data (DAS) runtime was 34 minutes and 21 seconds with 22,102 images/sec. During those runs the average GPU utilization was 96%, as observed on nvidia-smi. Note that this average includes the testing phase, during which GPUs were not used, while CPU utilization was 40% as measured by mpstat. This demonstrates that the data delivery rate is sufficient in each case.

Architecture adjustments

The setup used for this validation can be adjusted to fit other use cases.

CPU Adjustments

We used a Skylake Intel Xeon Platinum 8360Y processor for this validation, as recommended by Lenovo. We expect that the equivalent Cascade Lake CPU, an Intel Xeon Gold 6330 processor, would deliver similar performance because this workload is not CPU bound.

Storage Capacity Increase

Based on your storage capacity needs, you can increase the share storage (NFS volume) on demand, provided that you have the additional disk shelves and controller models. You can do this from the CLI or from the NetApp web interface of the storage controller as the admin user.

Conclusion

The NetApp and Lenovo solution validated here is a flexible scale-out architecture that is ideal for entry into mid-level enterprise AI.

NetApp storage delivers the same or better performance as local SSD storage and offers the following benefits to data scientists, data engineers, and IT decision makers:

- Effortless sharing of data between AI systems, analytics, and other critical business systems. This data sharing reduces infrastructure overhead, improves performance, and streamlines data management across the enterprise.
- Independently scalable compute and storage to minimize costs and improve resource utilization.
- Streamlined development and deployment workflows using integrated snapshots and clones for instantaneous and space-efficient user workspaces, integrated version control, and automated deployment.
- Enterprise-grade data protection for disaster recovery and business continuance.

Acknowledgments

- Karthikeyan Nagalingam, Technical Marketing Engineer, NetApp
- Jarrett Upton, Admin, AI Lab Systems, Lenovo

Where to find additional information

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

- NetApp All Flash Arrays product page

  https://www.netapp.com/us/products/storage-systems/all-flash-array/aff-a-series.aspx

- NetApp AFF A400 page

  https://docs.netapp.com/us-en/ontap-systems/a400/index.html

- NetApp ONTAP data management software product page

  http://www.netapp.com/us/products/data-management-software/ontap.aspx

- MLPerf

  https://mlperf.org

- TensorFlow benchmark

  https://github.com/tensorflow/benchmarks

- NVIDIA SMI (nvidia-smi)

  https://developer.nvidia.com/nvidia-system-management-interface

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
image::aipod_nv_sizing_new.png[]

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/

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.

NetApp EF-Series AI with NVIDIA

Overview of EF-Series AI converged infrastructure solutions from NetApp and NVIDIA.

EF-Series AI with NVIDIA DGX A100 Systems and BeeGFS

- Design Guide
- Deployment Guide
- BeeGFS Deployment Guide

NVIDIA DGX SuperPOD with NetApp - Design Guide

Amine Bennani, David Arnette and Sathish Thyagarajan, NetApp

Executive summary

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 NVIDIA DGX SuperPOD makes supercomputing infrastructure easily accessible for any organization and delivers the extreme computational power needed to solve even the most complex AI problems. To help customers 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.

Program summary

NVIDIA DGX SuperPOD with NVIDIA DGX H100 systems and NVIDIA Base Command brings together a design-optimized combination of AI computing, network fabric, storage, software, and support. The BeeGFS on NetApp architecture has been previously validated on a dedicated acceptance cluster at NVIDIA. The latest architecture extends that validation by maintaining the proven design while incorporating support for the latest hardware from NVIDIA.

Solution overview

NVIDIA DGX SuperPOD is an AI data center infrastructure platform delivered as a turnkey solution for IT to support the most complex AI workloads facing today's enterprises. It simplifies deployment and management while delivering virtually limitless scalability for performance and capacity. In other words, DGX SuperPOD lets you focus on insights instead of infrastructure.

With NetApp EF600 all-flash arrays at the foundation of an NVIDIA DGX SuperPOD, customers 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 current and future storage requirements. Modular storage building blocks allow a granular approach to growth and scale seamlessly from terabytes to petabytes. By increasing the number of storage building blocks, customers can scale up the performance and capacity of the file system, enabling the solution to manage the most extreme workloads with ease.

Solution technology

- NVIDIA DGX SuperPOD with NVIDIA DGX H100 systems leverates DGX H100 systems with validated externally attached shared storage:

  ◦ Each DGX SuperPOD scalable unit (SU) consists of 32 DGX H100 systems and is capable of 640 petaFLOPS of AI performance at FP8 precision. It usually contains at least two NetApp BeeGFS building blocks depending on the performance and capacity requirements for a particular installation.

A high-level view of the solution
image::EF_SuperPOD_HighLevel.png[]

- NetApp BeeGFS building blocks consists of two NetApp EF600 arrays and two x86 servers:

  ◦ With NetApp EF600 all-flash arrays at the foundation of NVIDIA DGX SuperPOD, customers get a reliable storage foundation backed by six 9s of uptime.

  ◦ The file system layer between the NetApp EF600 and the NVIDIA DGX H100 systems is the BeeGFS parallel file system. BeeGFS was 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.

  ◦ NetApp support for BeeGFS aligns NetApp's excellent support organization with customer requirements for performance and uptime. Customers get access to superior support resources, early access to BeeGFS releases, and access to select BeeGFS enterprise features such as quota enforcement and high availability (HA).

- The combination of NVIDIA SuperPOD SUs and NetApp BeeGFS building blocks provides an agile AI solution in which compute or storage scales easily and seamlessly.

NetApp BeeGFS building block
image::EF_SuperPOD_buildingblock.png[]

Use Case Summary

This solution applies to the following use cases:

- Artificial Intelligence (AI) including machine learning (ML), deep learning (DL), natural language processing (NLP), natural language understanding (NLU) and g
  generative AI (GenAI).

- Medium to large scale AI training

- Computer vision, speech, audio, and language models

- HPC including applications accelerated by message passing interface (MPI) and other distributed computing techniques

- Application workloads characterized by the following:

  ◦ Reading or writing to files larger than 1GB

- Reading or writing to the same file by multiple clients (10s, 100s, and 1000s)
- Multiterabyte or multipetabyte datasets
- Environments that need a single storage namespace optimizable for a mix of large and small files

Technology Requirements

This section covers the technology requirements for the NVIDIA DGX SuperPOD with NetApp solution.

Hardware requirements

Table 1 below lists the hardware components that are required to implement the solution for a single SU. The solution sizing starts with 32 NVIDIA DGX H100 systems and two or three NetApp BeeGFS building blocks. A single NetApp BeeGFS building block consists of two NetApp EF600 arrays and two x86 servers. Customers can add additional building blocks as the deployment size increases. For more information, see the NVIDIA DGX H100 SuperPOD reference architecture and NVA-1164-DESIGN: BeeGFS on NetApp NVA Design.

Table 1. Hardware requirements

Hardware	Quantity
NVIDIA DGX H100	32
NVIDIA Quantum QM9700 switches	8 leaf, 4 spine
NetApp BeeGFS building blocks	3

Software requirements

Table 2 below lists the software components required to implement the solution. The software components that are used in any particular implementation of the solution might vary based on customer requirements.

Table 2. Software requirements

Software
NVIDIA DGX software stack
NVIDIA Base Command Manager
ThinkParQ BeeGFS parallel file system

Solution verification

NVIDIA DGX SuperPOD with NetApp was validated on a dedicated acceptance cluster at NVIDIA by using NetApp BeeGFS building blocks. Acceptance criteria was based on a series of application, performance, and stress tests performed by NVIDIA. For more information, see the NVIDIA DGX SuperPOD: NetApp EF600 and BeeGFS Reference Architecture.

Conclusion

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

Where to find additional information

To learn more about the information that is described in this document, review the following documents and/or websites:
NVA-1164-DESIGN: BeeGFS on NetApp NVA Design
https://www.netapp.com/media/71123-nva-1164-design.pdf
NVA-1164-DEPLOY: BeeGFS on NetApp NVA Deployment
https://www.netapp.com/media/71124-nva-1164-deploy.pdf
NVIDIA DGX SuperPOD Reference Architecture
https://docs.nvidia.com/dgx-superpod/reference-architecture-scalable-infrastructure-h100/latest/index.html#
NVIDIA DGX SuperPOD Data Center Design Reference Guide
https://docs.nvidia.com/nvidia-dgx-superpod-data-center-design-dgx-h100.pdf
NVIDIA DGX SuperPOD: NetApp EF600 and BeeGFS
https://nvidiagpugenius.highspot.com/viewer/62915e2ef093f1a97b2d1fe6?iid=62913b14052a903cff46d054&source=email.62915e2ef093f1a97b2d1fe7.4

NVA-1156-DESIGN: NetApp EF-Series AI with NVIDIA DGX A100 systems and BeeGFS

Abdel Sadek, Tim Chau, Joe McCormick and David Arnette, NetApp

NVA-1156-DESIGN describes a NetApp Verified Architecture for machine learning (ML) and artificial intelligence (AI) workloads using NetApp EF600 NVMe storage systems, the BeeGFS parallel file system, NVIDIA DGX A100 systems, and NVIDIA Mellanox Quantum QM8700 200Gbps IB switches. This design features 200Gbps InfiniBand (IB) for the storage and compute cluster interconnect fabric to provide customers with a completely IB-based architecture for high-performance workloads. This document also includes benchmark test results for the architecture as implemented.

NVA-1156-DESIGN: NetApp EF-Series AI with NVIDIA DGX A100 systems and BeeGFS

NVA-1156-DEPLOY: NetApp EF-Series AI with NVIDIA DGX A100 systems and BeeGFS

Abdel Sadek, Tim Chau, Joe McCormick, and David Arnette, NetApp

This document describes a NetApp Verified Architecture for machine learning (ML) and artificial intelligence (AI) workloads using NetApp EF600 NVMe storage systems, the ThinkParQ BeeGFS parallel file system, NVIDIA DGX A100 systems, and NVIDIA Mellanox Quantum QM8700 200Gbps InfiniBand (IB) switches. This document also includes instructions for executing validation benchmark tests after the deployment is complete.

NVA-1156-DEPLOY: NetApp EF-Series AI with NVIDIA DGX A100 systems and BeeGFS

TR-4859: Deploying IBM spectrum scale with NetApp E-Series storage - Installation and validation

Chris Seirer, NetApp

TR-4859 describes the process of deploying a full parallel file system solution based on IBM's Spectrum Scale software stack. TR-4859 is designed to provide details on how to install Spectrum Scale, validate the infrastructure, and manage the configuration.

TR-4859: Deploying IBM spectrum scale with NetApp E-Series storage - Installation and validation

TR-4815: NetApp AFF A800 and Fujitsu Server PRIMERGY GX2570 M5 for AI and ML model training workloads

David Arnette, NetApp
Takashi Oishi, Fujitsu

This solution focuses on a scale-out architecture to deploy artificial intelligence systems with NetApp storage systems and Fujitsu servers. The solution was validated with MLperf v0.6 model-training benchmarks using Fujitsu GX2570 servers and a NetApp AFF A800 storage system.

TR-4815: NetApp AFF A800 and Fujitsu Server PRIMERGY GX2570 M5 for AI and ML model training workloads