Technical Report

NetApp AFF A800 and Fujitsu Server PRIMERGY GX2570 M5 for AI and ML Model Training Workloads

David Arnette, NetApp
Takashi Oishi, Fujitsu
February 2020 | TR-4815

Abstract

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.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

This solution focuses on a clustered architecture using NetApp® storage systems and Fujitsu PRIMERGY servers optimized for artificial intelligence (AI) workflows. It covers testing and validation for PRIMERGY GX2570 M5 servers and a NetApp AFF A800 storage system. In this validation, we demonstrate an efficient and cost-effective solution for high-performance distributed training with NVIDIA Tesla V100 GPUs and the enterprise-grade data management capabilities of NetApp ONTAP® cloud-connected data storage.

1.1 Target Audience

This document is intended for the following audiences:

- Enterprise architects who design solutions for the development of AI models and software
- Data scientists and data engineers who are looking for efficient ways to achieve deep learning (DL) and machine learning (ML) development goals
- IT decision makers, and business leaders who are interested in achieving the fastest time to market from AI initiatives

1.2 Use Case Summary

Computer vision capabilities are having a significant impact in a wide variety of industry settings, from autonomous vehicles to AI-assisted medical diagnosis. Training the DL and ML algorithms used for computer vision applications requires massive quantities of data and significant computing power. This technical report demonstrates the performance of the NetApp and Fujitsu solution using industry-standard computer vision models and datasets.

2 Solution Technology

This NetApp and Fujitsu solution is designed to handle large datasets by using the processing power of GPUs alongside traditional CPUs. This validation demonstrates high performance and optimal data management and protection with a scale-out architecture using Fujitsu PRIMERGY GX2570 servers.

This solution offers the following key benefits:

- 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
- Full scalability of data storage and computation power with multiple Fujitsu servers

2.1 NetApp AFF Systems

State-of-the-art NetApp AFF 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.

The NetApp AFF A800 system (Figure 1) is the industry's first end-to-end NVMe solution. For NAS workloads, a single AFF A800 system supports throughput of 25GBps for sequential reads and 1 million IOPS for small random reads at sub-500µs latency. AFF A800 systems support the following features:

- Massive throughput of up to 300GBps and 11.4 million IOPS in a 24-node cluster
- 100 Gigabit Ethernet (GbE) and 32Gb FC connectivity
- Up to 30TB SSDs with multistream write
- High density with 2PB in a 2U drive shelf

- Scaling from 200TB (two controllers) up to 9.6PB (24 controllers)
- NetApp ONTAP 9.5, with a complete suite of data protection and replication features for industry-leading data management

Other NetApp storage systems, such as the AFF A700, AFF A320, and AFF A220, offer lower performance and capacity options for smaller deployments at lower cost points.

Figure 1) NetApp AFF A800.

NetApp ONTAP 9

ONTAP 9 is the latest generation of storage management software from NetApp that 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 to ensure 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. These benefits apply 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 provides automatic tiering of 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 cloud-connected storage management software offers 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.

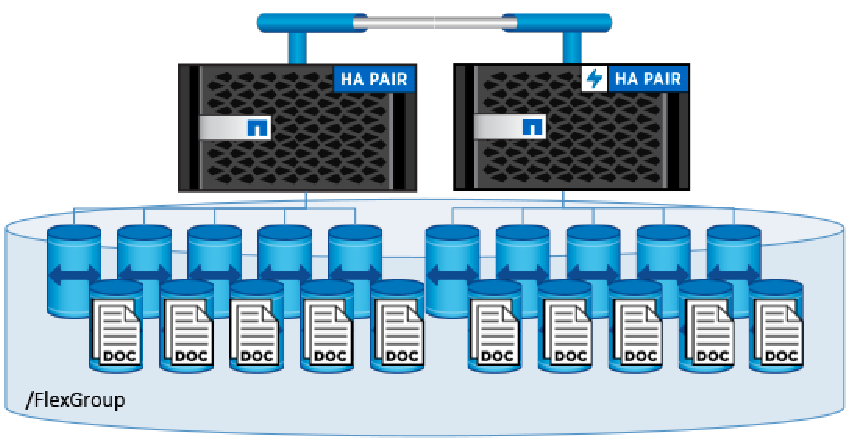
NetApp FlexGroup Volumes

A training dataset is 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 (Figure 2) 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

Figure 2) NetApp FlexGroup volumes.



2.2 Fujitsu Server PRIMERGY GX2570 M5

The PRIMERGY GX2570 M5 (Figure 3) is a 4U, dual-socket, PFLOPS-class high-performance server with the latest processor from the Intel Xeon scalable processor family. It features eight 125 TFLOPS NVIDIA Tesla V100 SXM2 GPUs (HBM 32GB) for AI and high-performance computing (HPC) workloads.

The PRIMERGY GX2570 M5 uses a balanced PCIe topology with GPUs evenly distributed on each CPU to achieve efficient transfer of data between system memory and GPU memory. The GPU-GPU connection with NVLink provides up to 50GBps of bandwidth. The PRIMERGY GX2570 M5 also has a 16-lane PCIe slot for every two GPUs that supports up to four low-profile interconnect cards. Using Mellanox ConnectX-5 host channel adapters (HCA) enables remote direct memory access (RDMA) data transfer between nodes at 100Gbps for both Ethernet and InfiniBand. These HCAs also allow larger workloads to use multiple nodes with linear performance scalability.

Internal storage includes six SATA 3.0 HDDs or SSDs through PCH and four faster PCIe Gen3 4-lane NVMe SSDs. Internal storage can be used for caching frequently used data or temporary space for deep learning, while the external storage is used for training with larger datasets.

The applications and the frameworks for maximizing the parallel processing power of NVIDIA Tesla V100 SXM2 GPUs are available on NVIDIA GPU Cloud (NGC) with validated containers.

Figure 3) Fujitsu PRIMERGY GX2570 M5 server.

2.3 Cisco Nexus 3232C

The Cisco Nexus 3232C switch (Figure 4) is a dense, power-efficient, low-latency, high-performance, 100Gbps switch designed for the data center. This compact, 1RU model offers wire-rate layer-2 and layer-3 switching on all ports with a latency of 450ns. This switch is a member of the Cisco Nexus 3200 platform and runs the industry-leading Cisco NX-OS software operating system, providing comprehensive features and functions that are widely deployed. The Cisco Nexus 3232C is a Quad Small Form-Factor Pluggable (QSFP) switch with 32 QSFP28 ports. Each QSFP28 port can operate at 10, 25, 40, 50, and 100Gbps, up to a maximum of 128 ports of 25Gbps.

Figure 4) Cisco Nexus switch with NX-OS support for CEE standards and RoCE v1 and v2.

2.4 RDMA over Converged Ethernet (RoCE)

Direct memory access (DMA) technology enables hardware subsystems such as network adapters and GPUs to access system memory and to read and write data without CPU processing.

RDMA allows network adapters to transfer data between application memory across server nodes without using the OS or device driver. This feature greatly reduces CPU overhead and latency and has traditionally been used in infrastructures that handle large workloads such as HPC. RDMA over Converged Ethernet (RoCE) is a technology that enables RDMA using Ethernet.

Ethernet is commonly used in enterprise data centers rather than the InfiniBand interconnects that have traditionally been used for RDMA. In the past, Ethernet was the best-effort data transfer method; however, it was not suitable for the low-latency and high-bandwidth data transfers required for communication between GPU nodes. RoCE leverages newer Converged Enhanced Ethernet (CEE) capabilities and allows priority flow control for lossless forwarding of Ethernet packets by allocating bandwidth to specific traffic on physical layer media. This enables simultaneous NFS storage access using the same 100GbE link, while guaranteeing bandwidth for RDMA between GPU nodes.

2.5 Automatically Build and Sustain System Infrastructure with Ansible

Ansible is a DevOps-style configuration management tool developed by Red Hat. The desired configuration of a system is defined in easily readable YAML files, including the software and hardware configuration of servers, storage, and networking. Ansible enables the automated initial deployment of infrastructure as well as ongoing configuration management and version control of the whole system. Ansible includes native modules for NetApp ONTAP storage systems, Fujitsu PRIMERGY servers, and Cisco Nexus network switches to streamline system management and reduce overhead during hardware and software configuration changes.

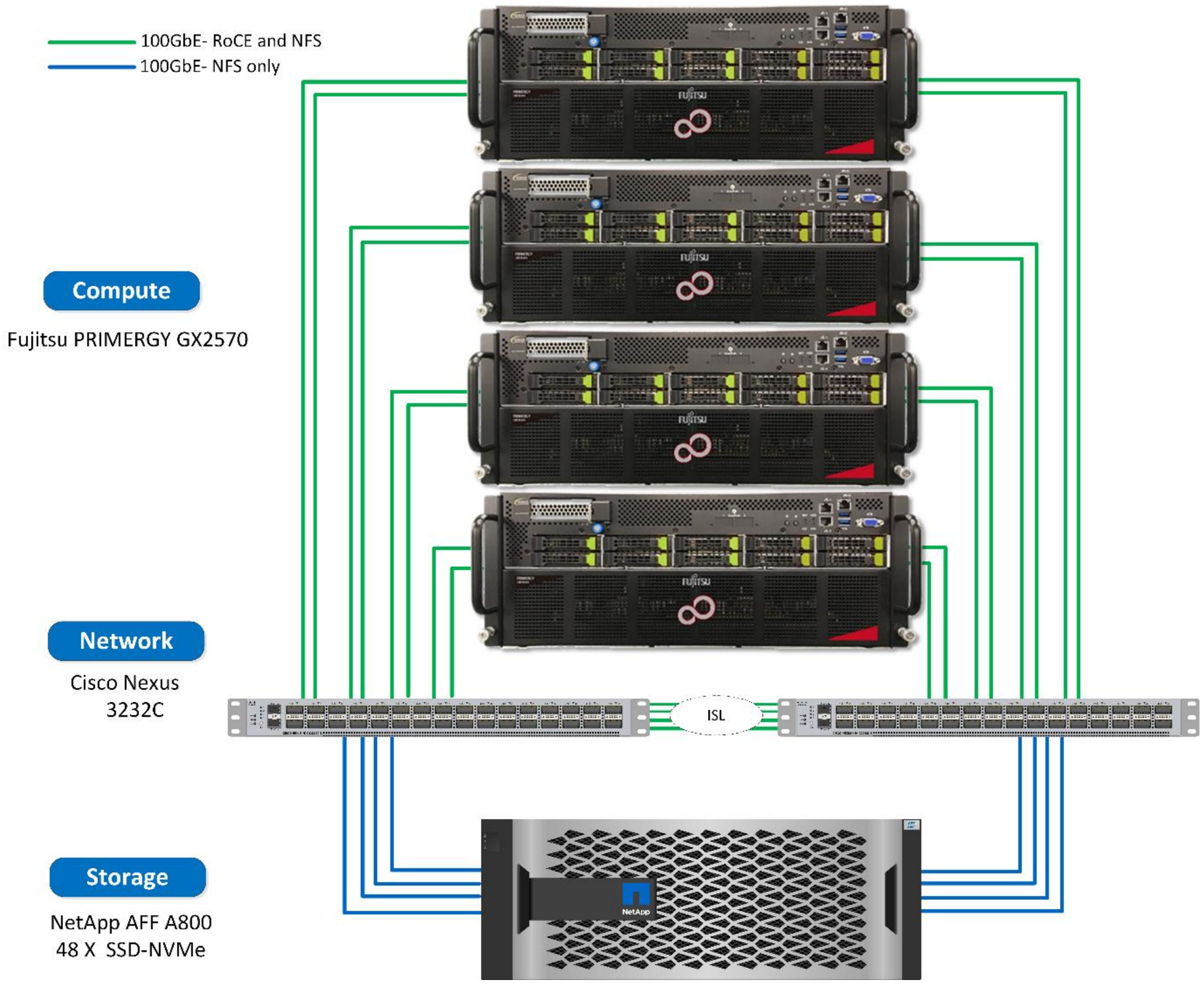
3 Test Configuration Details

This section describes the tested configurations, the network infrastructure, the GX2570 server, and the storage provisioning details.

3.1 Solution Architecture

This solution was validated with one NetApp AFF A800 storage system, four Fujitsu PRIMERGY GX2570 servers, and two Cisco Nexus 3232C 100GbE switches. As illustrated in Figure 5, each PRIMERGY system is connected to the Nexus switches with four 100GbE connections that are used for inter-GPU communications with RoCE. Traditional IP communications for NFS storage access also occur on these links. Each storage controller is connected to the network switches with four 100GbE links.

Figure 5) Network topology of tested configuration.



3.2 Hardware Requirements

Table 1 lists the hardware components required to implement the solution as tested.

Table 1) Hardware requirements.

Hardware	Quantity	Note
Fujitsu PRIMERGY GX2570 M5 servers	4	**CPU:** Dual, 24-core Intel Xeon 8280. **System memory:** 12 DDR4 32GB 2933MHz. **GPUs:** 8 NVIDIA Tesla V100 SXM2 32GB. **Storage:** 2 SATA SSDs, 2 NVMe SSDs. **Network:** 2 Mellanox ConnectX-5 2ports HCA. **Power consumption:** Max 3,656W.
NetApp AFF A800 system	1	High-availability (HA) pair, including two controllers and 48 NVMe SSDs.
Cisco Nexus 3232C network switches	2	

3.3 Software Requirements

Table 2 lists the software components required to implement the solution as tested.

Table 2) Software requirements.

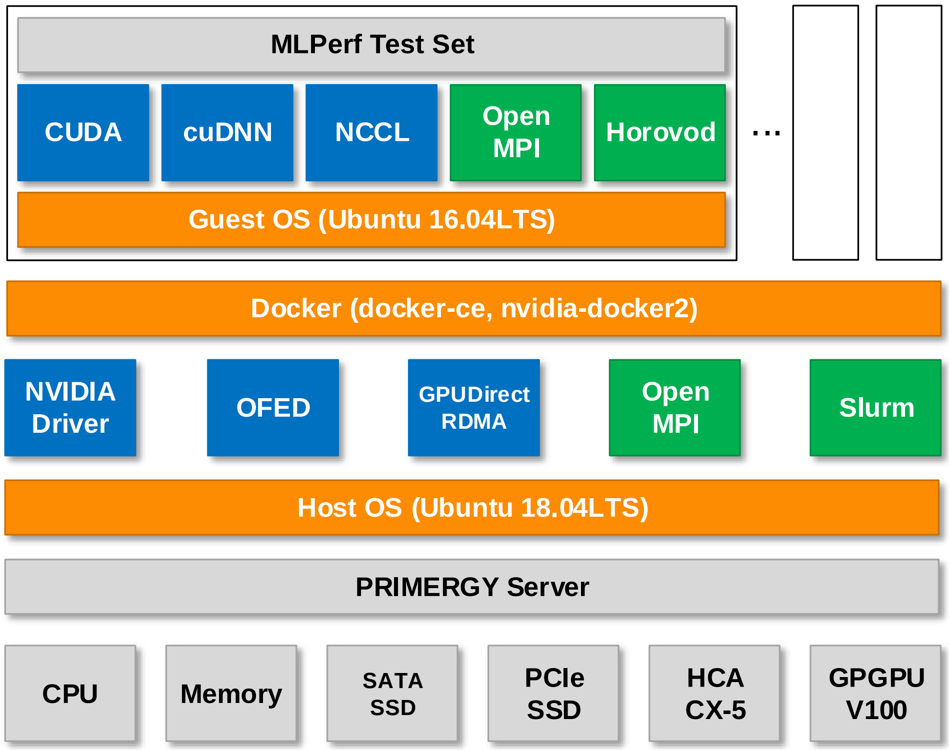
Software	Version or Other Information
OS	Ubuntu 18.04.3 LTS
Docker container platform	18.09.4-ce
NVIDIA-Docker container	nvcr.io/NVIDIA/pytorch:19.05-py3 nvcr.io/NVIDIA/mxnet:19.05-py3 nvcr.io/NVIDIA/tensorflow:19.05-py3
Cuda	10.1
Slurm	19.05.4
Benchmark software	SSD300 v1.1 for PyTorch
	Mask R-CNN for PyTorch
	ResNet-50 v1.5 for Mxnet
	Minigo for TensorFlow
Mellanox HCA driver	OFED 4.6
NVIDIA Peer Memory Client	Nv-peer-memory 1.0.8
Munge	
Open-mpi	3.1.5

4 Validation Test Plan and Results

4.1 Validation Test Plan

This solution was validated using the MLPerf v0.6 benchmark models and testing procedure. Each model was trained using one PRIMERGY GX2570 server and one NetApp AFF A800 storage system in the configuration described in Section 3.1. In addition to testing each model with a single server, the ResNet-50 model was trained using two, three, or four PRIMERGY GX2570 servers to demonstrate the scalability of the architecture. Figure 6 shows the software stack used in this validation.

Figure 6) MLPerf benchmark software stack.



MLPerf v0.6 testing procedures were used to produce these results. The following notes apply to the test results that follow:

- Each result was computed by executing five test runs, dropping the fastest and slowest, and then taking the mean of the remaining three runs.
- We tested each model with the recommended dataset, as noted below. Each test used a standardized dataset size and makeup in order to normalize results across various implementations. No modifications were made to the datasets in any of these tests.
- Each model was tested using the parameters recommended in the MLPerf v0.6 testing procedures.
- Training for each model ran until it reached the quality target specified for each model in the MLPerf v0.6 testing procedures.

The following models were used in this validation:

- **SSD.** One of the most popular object detectors for multiple categories. SSD can achieve faster frames-per-second (FPS) rates than Faster R-CNN while producing markedly superior detection accuracy. In addition to stressing GPU resources, SSD requires CPU capabilities to complete complex object detection processing. This test was performed with the CoCo2017 dataset.
- **Mask R-CNN.** Extends Faster R-CNN to perform both object detection and segmentation with minimal processing overhead. This test was also performed with the CoCo2017 dataset.
- **ResNet-50.** Another popular convolutional neural network (CNN) model for image classification. This test was performed with an MXNet RecordIO database containing the ImageNet dataset with default image sizes.
- **Minigo.** A benchmark of reinforcement learning for the 9x9 version of the boardgame Go. Minigo is an independent effort from AlphaGo and is based on Brian Lee's MuGo. Three primary phases are performed in each iteration: self-play, training, and model evaluation.

4.2 Validation Test Results

MLPerf Model Training Results

Each MLPerf training benchmark measured the processing time required to train a model on the specified dataset to achieve the specified quality target. Table 3 shows the clock-time result for each of the models trained. Because the datasets, training parameters, and quality targets are standardized in these benchmarks, these results can be compared to other publicly available MLPerf results.

Table 3) MLPerf benchmark processing time for tested models.

Model	Training Time Result
SSD	19.54 minutes
MaskR-CNN	186.22 minutes
ResNet-50	94.76 minutes
Minigo	24.97 minutes

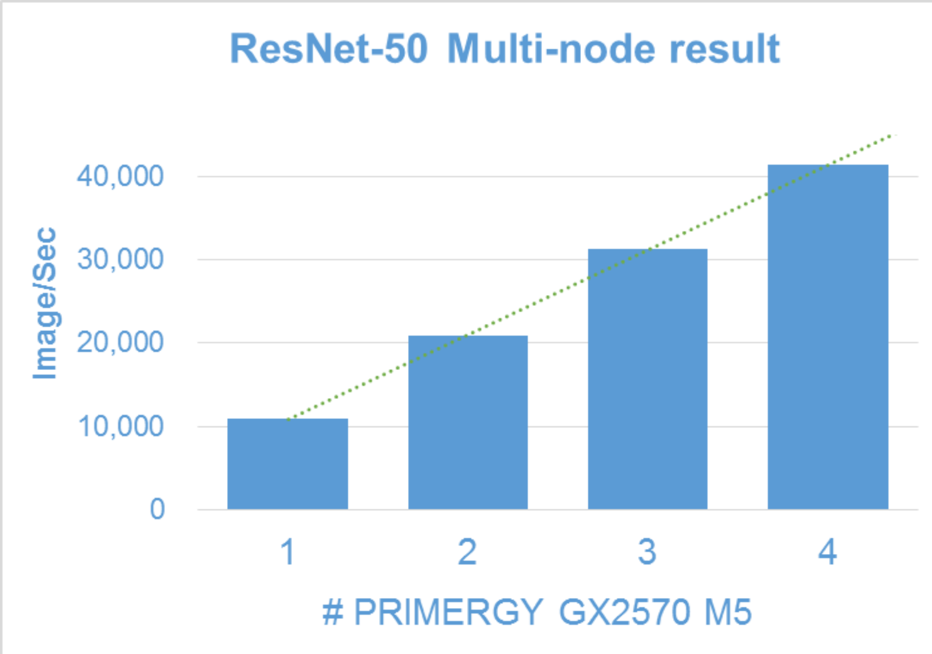
Note that these are unverified scores of v0.6 on the MLPerf image classification benchmark and have not been submitted for review or validation. However, they do compare favorably to official publicly available results[1].

Scale-Out Model Training with ResNet-50

In addition to basic benchmark testing, the solution was validated with ResNet-50 on up to four PRIMERGY GX2570 M5 servers to demonstrate scalability across multiple GPUs with RoCE as the interconnect. Horovod was used to distribute training across multiple nodes. Memory, the GPU core, HCAs, and logical network ports on every node were associated with each Message Passing Interface (MPI) and NVIDIA Collective Communications Library (NCCL) process in detail to provide communication bandwidth between nodes. As shown in Figure 7, this configuration provides scalability of computing across multiple GPUs and nodes. Figure 7 also shows that the training images per second scaled in a linear fashion as PRIMERGY GX2570 M5 servers were added, validating that there were no bottlenecks in the network or storage subsystems.

[1]These results were not verified by MLPerf. The MLPerf name and logo are trademarks. See www.mlperf.org for more information.

Figure 7) Training results for ResNet-50 with up to four PRIMERGY servers.



5 Conclusion

The Fujitsu PRIMERGY GX2570 server is an extremely powerful DL platform that benefits from equally powerful storage and network infrastructure to deliver maximum value. The combined solution of PRIMERGY servers and NetApp storage systems provides an infrastructure that delivers excellent performance and seamless scalability with industry-leading data management.

Acknowledgments

The authors gratefully acknowledge the contributions that were made to this technical report by our esteemed colleagues from Fujitsu, Munenori Maeda, Akihiko Kasagi, Takumi Danjo, Yasushi Hara, and Kai Yutaka.

The authors would also like to acknowledge the contributions of our NetApp team members, Sung-Han Lin and Sathish Thyagarajan.

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 ONTAP data management software product page
  http://www.netapp.com/us/products/data-management-software/ontap.aspx
- MLPerf
  https://mlperf.org/
- MLPerf training rules
  https://github.com/mlperf/training_policies/blob/master/training_rules.adoc

- SSD model
  https://github.com/NVIDIA/DeepLearningExamples/tree/master/PyTorch/Detection/SSD)
- Mask R-CNN
  https://github.com/NVIDIA/DeepLearningExamples/tree/master/PyTorch/Segmentation/MaskRCNN
- ResNet-50
  https://github.com/NVIDIA/DeepLearningExamples/tree/master/MxNet/Classification/RN50v1.5
- Minigo
  https://github.com/mlperf/training_results_v0.6/tree/master/NVIDIA/benchmarks/minigo/implementations/tensorflow
- Horovod
  https://github.com/horovod/horovod

Version History

Version	Date	Document Version History
Version 1.0	February 2020	Initial release

Refer to the [Interoperability Matrix Tool (IMT)] on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.