NVIDIA DGX SuperPOD with WEKApod Data Platform Appliance

Reference Architecture

Featuring NVIDIA DGX Systems

RA-11335-001 V1

May 2024

Abstract

The NVIDIA DGX SuperPOD™ with NVIDIA DGX™ systems is an artificial intelligence (AI) supercomputing infrastructure, providing the computational power necessary to train today's state-of-the-art deep learning (DL) models and to fuel future innovation. The DGX SuperPOD delivers groundbreaking performance, deploys as a fully integrated system, and is designed to solve the world's most challenging computational problems.

This DGX SuperPOD reference architecture (RA) for DGX systems with x86 architecture CPUs is the result of collaboration between DL scientists, application performance engineers, and system architects to build a system capable of supporting the widest range of DL workloads. The groundbreaking performance delivered by the DGX SuperPOD with DGX systems enables the rapid training of DL models at great scale. The integrated approach of provisioning, management, compute, networking, and fast storage enables a diverse, multi-tenant system that can span data analytics, model development, and AI inference.

The WEKA®Data Platform on the WEKApod appliance was evaluated for suitability for supporting DL workloads when connected to the DGX SuperPOD. This fully integrated storage appliance is validated with the DGX SuperPOD, and WEKApod is fully supported by WEKA support services.

DGX SuperPOD with the WEKApod Data Platform is designed to make large-scale AI simpler, faster, and easier to manage for every organization and their IT team. The WEKApod Data Platform has key capabilities that benefit organizations as they elevate their AI initiatives to DGX SuperPOD scale:

> The WEKA file system provides all the performance required for the most demanding HPC and AI data management.

> Easy to deploy, WEKApod Data Platform appliances are pre-configured, and provide a highly capable scale-out platform for capacity and performance.

> Fully optimized for all types of I/O patterns and data layouts, WEKApod Data Platform appliances deliver data to applications, ensuring excellent GPU resource utilization even with distributed applications running on multiple computing servers. The WEKA shared parallel protocol engages the full capabilities of the data infrastructure, and AI applications achieve excellent results, reliably, and consistently.

Learn more about the NVIDIA and WEKA collaboration at:
https://www.weka.io/resources/video/introducing-wekapod/

Contents

Storage Overview

Training performance may be limited by the rate at which data can be read and reread from storage. The key to performance is the ability to write data multiple times, ideally from local storage. The closer the data is cached to the GPU, the faster it can be read. Both persistent and nonpersistent storage needs to be designed to balance the needs of performance, capacity, and cost.

Storage Caching Hierarchy

The storage caching hierarchy for DGX systems is shown in Table 1. Depending on data size and performance needs, each tier of the hierarchy can be leveraged to maximize application performance.

Table 1. DGX system storage and caching hierarchy

Storage Hierarchy Level	Technology	Total Capacity[1]	Performance[1]
RAM	DRAM	2 TB[2]	> 200 GB/s
Internal storage	Flash storage	30 TB[3]	> 50 GB/s
1. Total capacity and performance values are per system. 2. Shared between the operating system, application, and other system processes 3. PCIe Gen 4 NVMe SSD storage			

Caching data in local RAM provides the best performance for reads. This caching is transparent once the data is read from the filesystem. While local storage is fast, it is not practical to manage a dynamic environment with local disk alone in a multi-node environment. Functionally, centralized storage may be as quick as local storage on many workloads.

Storage Performance Requirements

Performance requirements for high-speed storage greatly depend on the types of AI models and data formats used. The DGX SuperPOD has been designed as a fully capable system that can manage any workload both today and in the future. However, if systems are going to focus on a specific workload, such as natural language processing (NLP), it may be possible to better estimate performance needs of the storage system.

To allow customers to characterize their own performance requirements, some general guidance on common workloads and datasets is shown Table 2.

Table 2: Characterizing different I/O workloads

Storage Performance Level Required	Example Workloads	Dataset Size
Good	NLP	Most to all datasets fit in cache
Better	Image processing with compressed images, ImageNet/ResNet-50	Many to most datasets can fit within the local system's cache.
Best	Training with 1080p, 4K, or uncompressed images, offline inference, ETL	Datasets are too large to fit into cache, massive first epoch I/O requirements, workflows that only read the dataset once

Achieving these performance characteristics may require the use of optimized file formats such as TFRecord, RecordIO, or HDF5.

The high-speed storage provides a shared view of an organization's data to all systems. It needs to be optimized for small, random I/O patterns, and provide high peak system performance and high aggregate filesystem performance to meet the variety of training workloads an organization may encounter.

About the WEKApod Data Platform

The WEKApod Data Platform meets performance characteristics with an architecture that delivers not just speed and scale but also operational efficiency and ease of use to quickly deploy and support large-scale AI initiatives with DGX SuperPOD.

Figure 1. WEKApod Data Platform appliance

Some WEKApod Data Platform appliance features include:

• Appliance deployment as a scalable building block that can be easily clustered into a single namespace that scales seamlessly in capacity, performance, and capability. It is fully integrated which streamlines deployment and simplifies management operations.

•Configuration for 500 TB or 1 PB usable capacity with TLC NVMe drives. The base appliance can be expanded seamlessly in a single namespace that provides up to several hundreds of petabytes of useable capacity, with linear performance scaling.

• Native support of the NVIDIA Quantum InfiniBand Platform is a key feature that maximizes performance and minimizes CPU overhead. Since the DGX SuperPOD compute fabric is InfiniBand, designing the storage fabric as InfiniBand means only one high-speed fabric type needs to be managed, simplifying operations.

• Appliance communications with DGX SuperPOD nodes using multiple NVIDIA Quantum InfiniBand connections for performance, load balancing, and resiliency.

• The WEKA shared parallel protocol enables a single DGX node to access data at over 90 GB/s and 3M IOPS from the WEKApod appliance.. Multiple nodes can access data on the appliance simultaneously, with performance available to all systems and distributed dynamically. Nodes can be optimized for ultra-low latency workloads.

• The all-NVMe architecture of the WEKApod Data Platform appliance provides excellent random read performance, often as fast as sequential read patterns.

• The WEKApod Data Platform appliance is space and power efficient with dense performance and capacity to optimize data center utilization and improve operational economics.

Validation Methodology

Three classes of validation tests are used to evaluate a particular storage technology and its configuration for use with the DGX SuperPOD: microbenchmark performance, real application performance, and functional testing. The microbenchmarks measure key I/O patterns for DL training and can be run on CPU-only nodes. Real DL training applications are then run on a DGX SuperPOD to confirm that the applications meet expected performance. Beyond performance, storage solutions are evaluated for robustness and resiliency as part of functional testing.

The NVIDIA DGX SuperPOD storage validation process leverages a “Pass or Fail” methodology. Specific targets are set for the microbenchmark test. Each benchmark result is graded as good, fair, or poor. A passing grade is one where at least 80% of the tests are good, and none are poor. In addition, there must be no catastrophic issues created during testing. For application testing, a passing grade is one where all cases are completed within 5% of the performance set by running the same tests with data staged on the DGX local RAID. For functional testing, a passing grade is one where all functional tests meet their expected outcomes.

Microbenchmarks

In the Storage Performance Requirements section, there are several high-level performance metrics that storage systems must meet to qualify as a DGX SuperPOD solution. Current testing recommends that the solutions meet the “Best” criteria discussed in Table 2. In addition to these high-level metrics, several groups of tests are run to validate the overall capabilities of the proposed solutions. These include single node tests where the number of threads is varied and multi-node tests where a single thread count is used and as the number of nodes vary. In addition, each test runs in both Buffered and DirectIO modes and when I/O is performed to separate files.

Five different read patterns are run. The first is sequential write to the shared file system. The second read operation is sequential where no data is in the cache. The second read operation is executed immediately thereafter to evaluate the ability for the filesystem to cache data. The cache is purged and then the data is read again, this time randomly. Lastly, the data is reread again randomly, to evaluate data caching.

Hero Benchmark Performance

The hero benchmark helps establish the peak performance capability of the entire solution. Storage parameters, such as filesystem settings, I/O size, and controlling CPU affinity, were tuned to achieve the best read and write performance. Storage devices were expected to demonstrate that quoted performance was close to measured performance. Other tests are crafted to demonstrate performance of real workloads. The delivered solution for a single Scalable Unit (SU) had to demonstrate over 20 GiB/s for writes and 65 GiB/s for reads, all when using 8 threads per node. Ideally, the write performance should be at least 50% of the read performance. However, some storage architectures have a different balance between read and write performance, so this is only a guideline.

Single-Node, Multi-File Performance

For single-node performance, I/O read and write performance is measured by varying the number of threads in incremental steps. Each thread writes (and reads) to (and from) its own file in the same directory.

For single-node performance tests, the number of threads varies from 1 to the ideal number of threads to maximize performance (typically more than half the cores but no more than the total physical cores). The I/O size varies between 128 KiB and 1 MiB and the tests are run with Buffered I/O and Direct I/O.

The target performance for a single thread for these tests is shown in Table 3.

Table 3. Single-node, multi-file performance targets

Thread Count	Buffered or DirectIO	I/O size (KiB)	Performance (MiB/s)				
			Write	**Read**	**Reread**	**Random Read**	**Random Reread**
1	Buffered	128	512	1,024	1,536	256	1,536
1	Buffered	1024	800	3,072	4,608	768	1,024
1	Direct	128	768	1,024	1,024	768	768
1	Direct	1024	1,024	1,024	1,024	1,024	1,024

When maximizing single-node performance, the thread count may vary; however, it is expected that performance does not drop significantly when additional threads are used beyond the optimal thread count.

Target performance for single-node performance with multiple threads is in Table 4. The optimal number of threads may vary for any storage configuration.

Table 4. Single-node, multi-threaded performance targets

Thread Count	Buffered or DirectIO	I/O size (KiB)	Performance (MiB/s)				
			Write	Read	Reread	Random Read	Random Reread
Varies	Buffered	128	8,000	12,000	18,000	12,000	18,000
	Direct	128	8,000	15,000	15,000	15,000	15,000
	Buffered	1024	10,000	20,000	30,000	20,000	30,000
	Direct	1024	10,000	20,000	20,000	20,000	20,000

For Buffered I/O tests, the reread performance relative to read performance can vary substantially between different storage solutions. The reread performance should be at least 50% of the read performance for both sequential and random reads.

Multi-Node, Multi-File Performance

The next test performed is a multi-node I/O read and write test to make sure that the storage appliance can provide the minimum required buffered read and write per system for the DGX SuperPOD. This benchmark determines the capacity of a filesystem to scale performance of different I/O patterns. Performance should scale linearly from one to a few nodes, reach a maximum performance, and not drop off significantly as more nodes are added to the job.

The target performance for a single test cluster (1 scalable unit (SU)) is 65 GiB/s for reads with an I/O size of 128 KiB or 1,024 KiB, and if the I/O is Direct or Buffered. The write performance should be at least 20 GiB/s, but ideally it would be 50% of the read performance. Results from these tests must be interpreted carefully as it is possible to add more hardware to achieve these levels. Overall performance is the goal, but it is desirable that the performance comes from an efficient architecture that is not over designed for its use.

Application Testing

Microbenchmarks provide indications of the peak performance of key metrics. However, it is application performance that is most important. A subset of the MLPerf Training benchmarks is used to validate storage performance and function. Here, both single-node and multi-node configurations are evaluated to ensure that the filesystem can support different I/O patterns and workloads. Training performance when data is staged on the DGX RAID was used as the baseline for performance. The performance goal is for the total time to train when data is staged on the shared filesystem to be within 5% of those measured when data is staged on the local RAID. This is not just for individual runs, but also when multiple cases are run across the DGX SuperPOD at the same time.

ResNet-50

ResNet-50 is the canonical image classification benchmark. The size of the training dataset is over 100 GiB and it has a requirement for fast data ingestion. On a DGX system, a single node training requires approximately 3 GiB per second and the dataset is small enough that it can fit into GPU memory. Preprocessing can vary, but the typical image size is approximately 128 KiB. One challenge of this benchmark is that at NVIDIA the processed images are stored in the RecordIO format (i.e. one large file for the entire dataset) since this provides the best performance for MLPerf. Since it is a single file, this can stress shared filesystem architectures that do not distribute the data across multiple targets or controllers.

Summary

NVIDIA evaluations show that the WEKApod Data Platform appliance meets the DGX SuperPOD performance and functionality requirements and is a certified choice to pair with a DGX SuperPOD to meet current and future storage needs.

As storage requirements grow, WEKApod Data Platform appliances can be added to seamlessly scale capacity, performance, and capability. The combination of NVME hardware and WekaFS file system architecture provides excellent random read performance, often just as fast as sequential read patterns. With the 1U form factor of the storage appliance, storage pools can be configured to meet or exceed the performance requirements of a DGX SuperPOD.

DGX SuperPOD customers should be assured that the WEKApod Data Platform appliance will meet their most demanding I/O needs for their most challenging AI workloads at any scale.

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation ("NVIDIA") makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete. NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer ("Terms of Sale"). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer's own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer's sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application to avoid a default of the application or the product. Weaknesses in customer's product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices. THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, NVIDIA DGX POD, NVIDIA DGX SuperPOD are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2024 NVIDIA Corporation. All rights reserved.

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, CA 95051

http://www.nvidia.com