White paper

NVIDIA DGX SuperPOD: VAST Data Reference Architecture

Designed and validated for NVIDIA DGX GB200/GB300 Systems

3/26 Version 1.1

Table of Contents

Table of Contents **2**

VAST Data Partnership with NVIDIA **3**

VAST Data Storage with NVIDIA DGX SuperPOD Reference Architecture 3

Executive Overview 3

Storage Architecture Overview **4**

About the VAST Data AI Operating System **5**

VAST DATA with NVIDIA DGX SuperPOD Solution Integration **7**

VAST Data with NVIDIA DGX SuperPOD with NVIDIA Quantum InfiniBand Networking **9**

NVIDIA DGX SuperPOD Scaling **10**

Storage Performance Guidance from NVIDIA **11**

Summary **12**

Appendix A: VAST Data Components for 2SU 13

Appendix B: Deployment Considerations 14

VAST Data Partnership with NVIDIA

The VAST Data® deep partnership engagement with NVIDIA® is a key foundation for building high-performance, AI-optimized infrastructure. By combining VAST Data AI Operating System (AI OS) and Data Platform with NVIDIA DGX SuperPOD™, this solution provides the powerful, scalable data storage needed to support today's data-intensive AI and machine learning demands. Designed to process massive large-scale datasets with ease, VAST Data is perfectly suited for applications like large language models, computer vision, and deep learning.

This integration creates a flexible foundation that can grow alongside an organization's AI needs. VAST Data's high-density, low-latency architecture, combined with NVIDIA DGX systems, supports non-disruptive scaling as AI usage expands, all while improving data center efficiency with lower power consumption. Built for long-term growth, this solution adapts to the fast pace of AI evolution without requiring extensive reconfiguration.

For those interested in the technical details, this guide provides a comprehensive look into deploying and managing DGX SuperPOD systems. Readers can explore architecture components, networking setups, and deployment steps, gaining practical insights on how to get the most out of the VAST Data and NVIDIA solution. With this foundation, users can start at any scale and confidently expand as their AI requirements grow, ensuring their infrastructure keeps pace with the latest AI advancements.

Learn more about the NVIDIA / VAST partnership at vastdata.com/nvidia

VAST Data Storage with NVIDIA DGX SuperPOD Reference Architecture

Executive Overview

NVIDIA DGX SuperPOD with VAST Data Platform is a turn-key data center infrastructure platform for enterprises with scale-out storage, optimized for AI. It is a multi-user system optimized for large scale AI and HP applications. It functions as a unified system, managing simultaneous users with advanced access controls for resource scheduling.

NVIDIA DGX SuperPOD represents the next generation of scale-out compute clusters that demand the most from their storage fabrics. Newer protocols like RoCE have further increased the need for simple and performant solutions that can handle the complexity, and high performance demands of today's AI workloads.

The data ingestion, processing, and storage requirements for large-scale enterprise AI require high performance, highly efficient storage that VAST Data was architected to deliver. Scaling to meet these workloads is an ever-growing part of pushing the boundaries of what is possible to achieve the next AI plateau. Some of the key features of DGX SuperPOD are:

- **Scalability:** It can scale up to tens of thousands of NVIDIA GPUs and multiple PetaBytes on VAST Data making it suitable for large-scale AI model training and inference.
- **Performance:** DGX SuperPOD is optimized for enterprise-scale inference, delivering unique architectural benefits including its NVLink domain, unmatched GPU memory, and leading compute performance. This inference-optimized architecture enables businesses to maximize token throughput and quality of service to maximize AI revenue.
- **Software Integration:** NVIDIA DGX SuperPOD leverages NVIDIA Mission Control for AI infrastructure operations and

workload orchestration. The software simplifies AI operations for enterprises by delivering the tools for agility, resilience, and hyperscale efficiency.

- **Enterprise Support:** Get dedicated expertise and services to help enterprises deploy and manage their AI infrastructure effectively.

Storage Architecture Overview

Integrating VAST Data into NVIDIA DGX SuperPOD storage fabric provides an AI data storage platform with enterprise-grade data services that accelerates AI data pipelines with a parallel architecture and multi-dimensional performance. VAST Data allows capacity and performance to be scaled independently and can be continuously improved over time without disruption to meet the changing and growing needs of an NVIDIA DGX SuperPOD environment.

Ease of Use & Management: VAST AI Operating System (AI OS) is designed to simplify AI infrastructure management by consolidating multiple layers of data processing into a single, highly efficient system. The VAST OS Management Service (VMS) integrates an API-first design, enabling seamless automation through RESTful interfaces while also offering a modern GUI for visual oversight. The system collects and analyzes performance, capacity, and health metrics in real time, providing AI engineers with deep insights into workload behavior and system efficiency.

Multi-tenant Data access and VAST Cluster enables multiple users within the same group or organization. It supports external user authorization providers such as LDAP, Active Directory (AD) and a local provider. Users can simultaneously access data via NFS and object storage protocols along with interchangeability in reading and writing to meet growing demands.

Multi-Protocol Support: The VAST AI Operating System (AI OS) gives customers native support for industry standard protocols like NFS, SMB, S3 and NVMe/TCP enhancing the agility of Enterprises to support demanding AI and HPC workloads seamlessly and efficiently

Enterprise Grade Reliability: Observed production systems realize 99.9999% Availability. With support for enterprise-grade resiliency, including non-disruptive updates, high availability DBoxes, and automated system health monitoring, VAST empowers AI organizations to build and maintain data infrastructures that are both future-proof and operationally efficient.

Non-Disruptive Upgrades: Traditional storage architecture often requires downtime for software upgrades, leading to service disruptions and administrative complexities. VAST Data eliminates this challenge through its fully automated, non-disruptive upgrade process. By leveraging a stateless VAST Server architecture, the platform ensures that system updates do not require taking any part of the infrastructure offline. The VAST Management Service (VMS) orchestrates the upgrade process by transferring workloads across available servers before updating each node sequentially, ensuring continuous operation. Similarly, VAST Enclosures (DBoxes) are upgraded without downtime by dynamically reprogramming PCIe switches to maintain active connectivity to SSDs while updating software on CNodes. This approach minimizes operational risk, eliminates the need for scheduled maintenance windows, and ensures that security patches and performance enhancements are deployed without service interruptions.

Seamless Performance and Capacity Scaling: Meet and exceed the performance and capacity needs of the enterprise with a seamless VAST DASE architecture.

Trusted AI Industry Leaders: VAST Data has earned the trust of AI infrastructure leaders, leading model providers by delivering an architecture that balances performance, scalability, and reliability. The platform's ability to support clusters spanning multiple

generations of hardware—while ensuring automatic load balancing and seamless upgrades—demonstrates its long-term viability. By decoupling computing and storage through its Disaggregated Shared-Everything (DASE) architecture, VAST enables organizations to scale infrastructure based on workload demands without requiring disruptive forklift upgrades. Its integration with high-performance NVMe fabrics, robust data reduction technologies, and strong security controls further solidify its position as a trusted solution in AI-driven enterprises.

Enterprise AI Use Cases: This platform handles a multitude of use cases like Agentic AI, Visual Language models, Large Language Models, Computer vision, deep learning recommendation systems - GANS, high performance computing, and more.

About the VAST Data AI Operating System

The VAST Data AI Operating System (AI OS) was subjected to a series of NVIDIA DGX SuperPOD tests to validate optimal performance and configuration.

The VAST AI Operating System (AI OS) is the first platform built from the ground up to manage the complete AI lifecycle. One unified foundation for storage, database, and compute runtime - eliminating the need to stitch together dozens of tools for training, inference, autonomous agents, and everything in between.

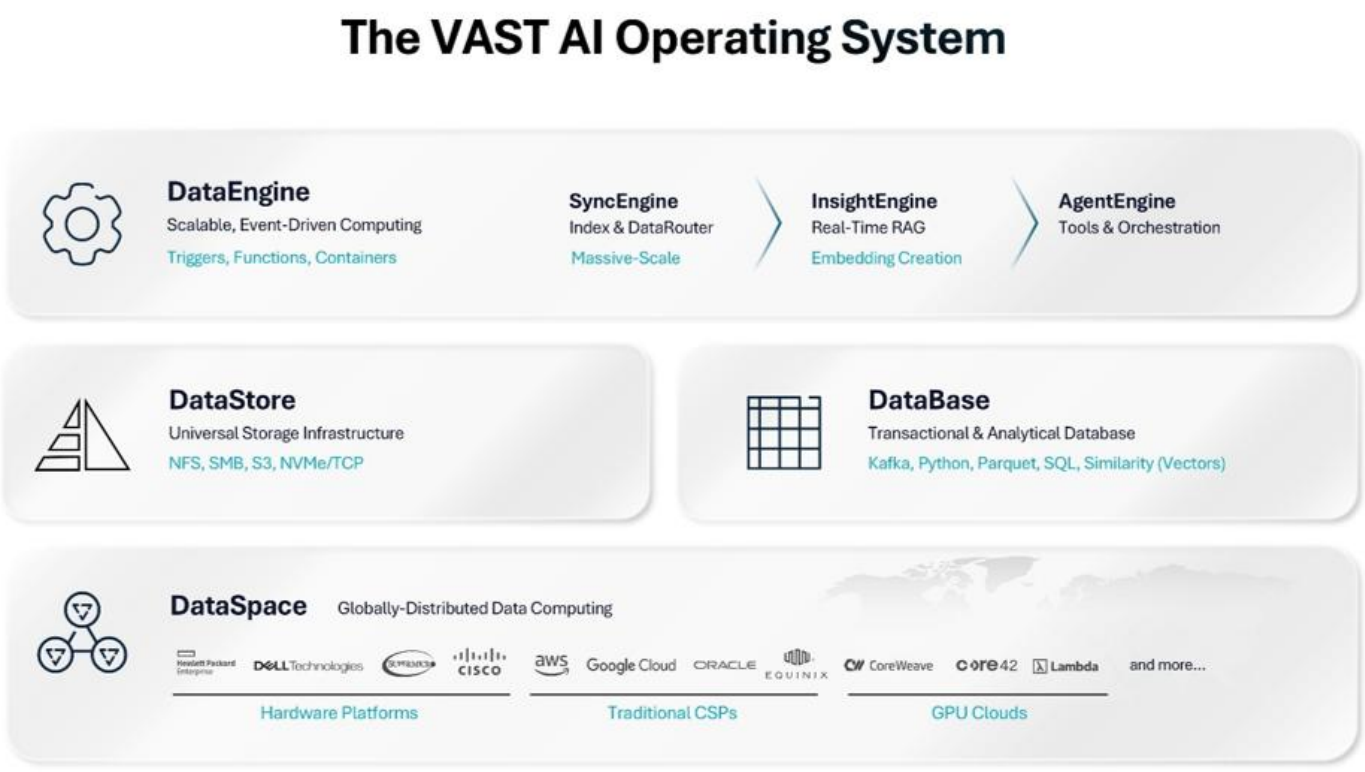


Figure 1. The VAST AI Operating System

The VAST AI OS meets performance requirements with the Disaggregated, Shared Everything (DASE) architecture that delivers not just speed and scale but also operational efficiency and ease of use such that any IT team can deploy and support large-scale AI initiatives with DGX SuperPOD.

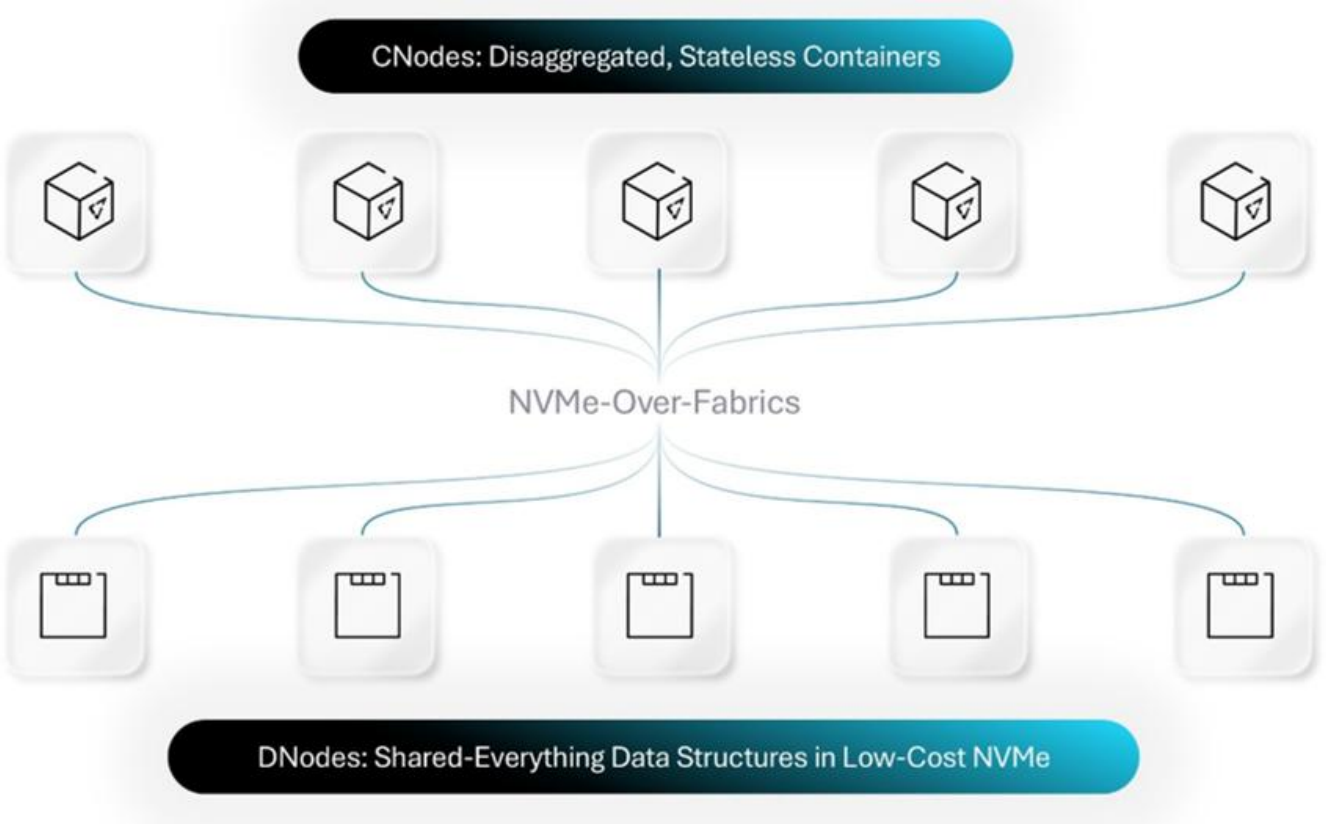


Figure 2. The VAST Disaggregated, Shared Everything (DASE) architecture disaggregates CPUs and connects them to a globally accessible pool of SCM and ultra-dense NVMe flash SSD



Figure 3. Highly available, NVIDIA Bluefield-3 DPU integrated NVMe enclosure

VAST challenges the long-held assumption that NFS performance is inadequate for AI and HPC workloads. The VAST Disaggregated, Shared-Everything architecture consists of two building blocks that are scaled across a common NVMe fabric. First, the state and storage capacity of the system is built from resilient, high-density NVMe-oF storage enclosures. Second, the logic of the system is implemented by stateless containers that each can connect to and manage all the media in the enclosures. By disaggregating compute from storage, it is possible to spread I/O across the system to achieve levels of parallelism for massive performance.

Other benefits of the VAST AI OS for DGX SuperPOD include:

- Independent scaling of performance and capacity with support for mixed generations of hardware in a single exabyte-scale namespace.
- Native support for both InfiniBand and Ethernet in the same namespace.
- Archive tier economics via next-generation global data reduction algorithms, support for hyperscale QLC flash with ten years of endurance, and ultra-efficient locally decodable erasure codes.
- Non-disruptive, online system expansions and software upgrades.
- Encryption, authentication, and external key management.
- VAST Catalog, a built-in metadata index, allows customers to find and manage data via SQL queries.
- Enterprise-grade data protection with support for n-1 and 1-n replication topologies and up to 1 million ransomware-proof snapshots.
- The VAST Data GUI management interface provides thousands of metrics via an API-first architecture unlocking real-time visibility into performance metrics.

VAST DATA with NVIDIA DGX SuperPOD Solution Integration

NVIDIA DGX SuperPOD clusters are transforming how data is processed and analyzed, especially when integrated with advanced Data storage solutions. These Data storage systems play a crucial role in managing the large amounts of data that AI applications require, enabling seamless access and sharing for distributed computing environments. This combination allows for efficient collaboration among nodes within a compute cluster, ensuring that AI models can be trained on large datasets without delays.

As the demand for AI-driven insights continues to grow, the reliance on powerful compute clusters alongside efficient storage solutions will expand, paving the way for innovation across diverse AI fields.

The integration of VAST Data with NVIDIA DGX SuperPOD compute clusters significantly enhances these capabilities. For example, in healthcare, AI can analyze extensive medical records and imaging data, with distributed storage providing quick retrieval and processing. In the financial sector, these systems support the storage of historical market data, allowing AI models to rapidly analyze trends and make predictions. Additionally, industries like media and entertainment benefit from this integration by managing large volumes of video and audio data for AI-driven content creation and editing.

VAST Data connects to an NVIDIA DGX SuperPOD environment via 400GbE or NDR ports. This allows the VAST Data to seamlessly scale from a minimal configuration into multiple racks in a simple, non-disruptive manner. VAST Data supports both NVIDIA Spectrum with RDMA over Converged Ethernet (RoCE) and NVIDIA Quantum InfiniBand connectivity options for customers. Details of this networking are discussed in the next section.

Ethernet networking is a cornerstone of enterprise data centers due to its low total cost of ownership (TCO), extensive interoperability, and proven reliability. When VAST Data is connected via Ethernet to DGX SuperPOD storage fabric, it delivers exceptional performance for large AI clusters. Additionally, the use of RDMA over Converged Ethernet (RoCE) further enhances this setup by minimizing latency and offloading CPU workloads, resulting in faster data transfers and improved overall efficiency for demanding AI applications.

VAST Data with NVIDIA DGX SuperPOD with NVIDIA Spectrum with RoCE Networking

VAST Data AI OS integrates effortlessly with NVIDIA DGX SuperPOD clusters using established Ethernet connectivity. It further enhances performance by supporting RDMA over Converged Ethernet (RoCE) communication between DGX systems and VAST Data, ensuring low-latency data transfer. Connecting CNodes to a pair of NVIDIA Spectrum-X SN5600/5610 switches follows a standard installation process that has been successfully implemented in thousands of installations worldwide.

The VAST Data AI OS, engineered with the industry's first Disaggregated, Shared-Everything (DASE) architecture, comprises three key physical infrastructure components: CNodes (compute), DBox (capacity), and NVMe Fabric (NVMe-oF) switches. This architecture enables a groundbreaking approach to storage systems design by enabling the ability to independently scale performance (via CNodes) and capacity (via DBox). This flexibility for enterprises and service providers—especially those with growing or expanding infrastructure needs—alleviates common capacity forecasting challenges while simultaneously meeting the needs of evolving AI workloads. For AI workloads, clients should be configured with a multi-path NFS driver using NFS-over-RDMA and spreading traffic across both local ports and remote ports. Additionally, Priority Flow Control (PFC) using DSCP, as well as ECN, is implemented in the converged network fabric to enable lossless behavior and, if using VXLAN, propagating DSCP from the inner header to the outer header. When deploying the VAST Data AI OS, each VAST CNode has a NVIDIA ConnectX®-7 (CX-7) network adapter and each VAST Ceres V2 DBox has two NVIDIA Bluefield-3 DPU's that connect up to three 400 GbE ports on the SN5600/5610 leaf switches as shown in figure 4. For 1SU with 512 GPUs sizing requires 14x 400 GbE ports on the SN5600/5610 leaf switches to support the VAST HPS solution with 8 CNodes and 3 DBoxes at a standard performance level.

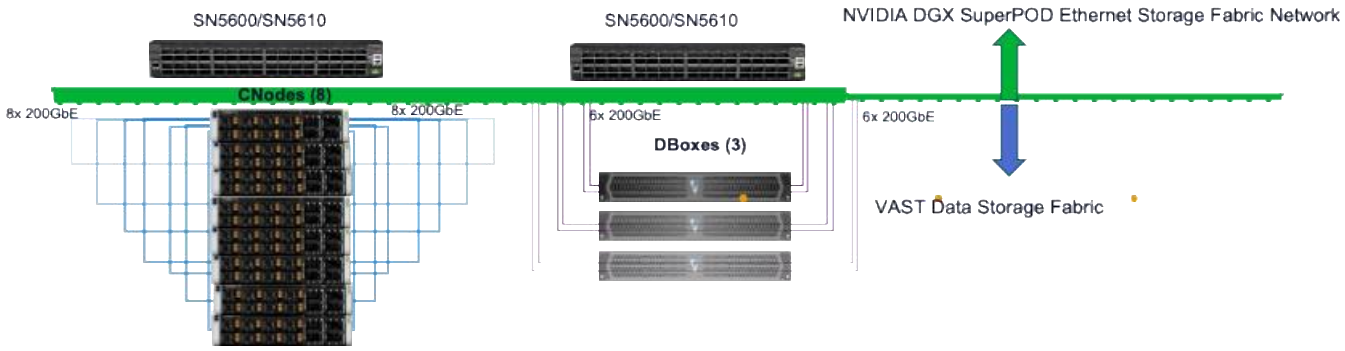


Figure 4. Switch connectivity from SN5600/5610 to eight CNodes and three DBoxes

As a default, for the CNode, VAST recommends the MMA4Z00-NS, OSFP finned-top twin-port 2x400G SR8 multimode transceivers in the SN5600/5610 connecting to a MMA1Z00-NS400 QSFP112 flat-top SR4 transceiver in the CX-7 using 2, split MPO-12/APC to MPO-12/APC cables. For the DBox, VAST recommends MMA4Z00-NS, OSFP finned-top twin-port 2x400G SR8 multimode transceivers in the SN5600/5610 connecting to a MMA1Z00-NS400 QSFP112 flat-top SR4 transceivers in the Bluefield-3 using 2, split MPO-12/APC to MPO-12/APC cables.

The VAST Data NVIDIA DGX SuperPOD Components BOM can be found in Appendix A.

VAST Data with NVIDIA DGX SuperPOD with NVIDIA Quantum InfiniBand Networking

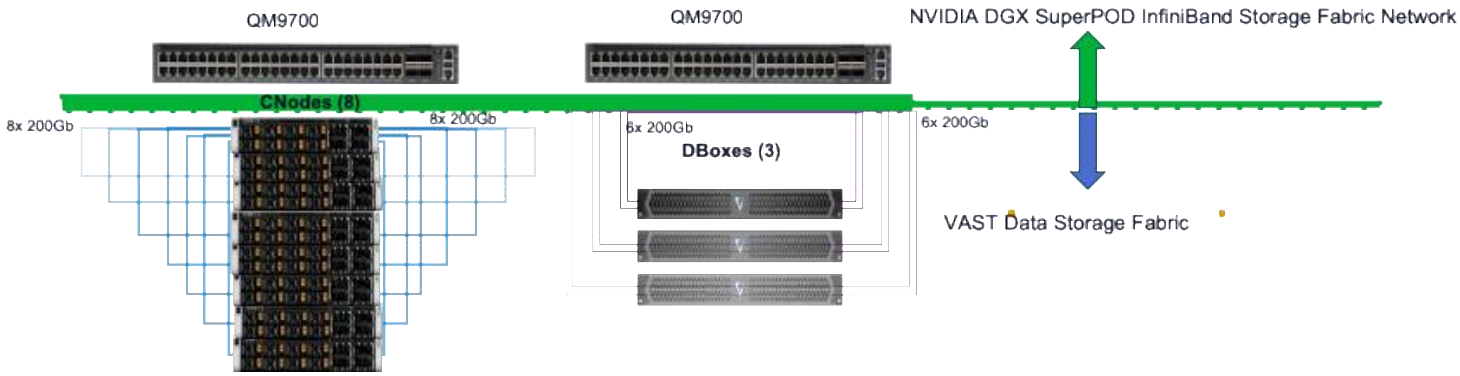


Figure 5. Switch connectivity from QM9700 across eight CNodes and three DBoxes

The VAST Data solution adheres to the NVIDIA Converged Network Design and connects to the NVIDIA Quantum QM9700 leaf switches. As part of the DGX SuperPOD RA with a Infiniband Storage network, QM9700 leaf switches are included in the bill of materials (BOM) articulated by NVIDIA. When deploying the VAST Data AI OS, each VAST CNode has a NVIDIA ConnectX®-7 (CX-7) network adapter and each VAST Ceres V2 DBox has two NVIDIA Bluefield-3 DPU's that connect to NDR Infiniband ports on the QM9700 leaf switches as shown in figure 5.

The VAST Data with NVIDIA DGX SuperPOD Components BOM can be found in Appendix A

NVIDIA DGX SuperPOD Scaling

The basic building block for NVIDIA DGX SuperPOD is the Scalable Unit (SU). A single SU consists of 32 DGX systems with 128 GPUs using NVIDIA Quantum InfiniBand Platform for compute fabric (also called East-West networking). Each SU communicates with the storage through the storage fabric (also called North-South networking).

The table below provides DGX SuperPOD with VAST Data platform scaling is based on a single SU to 64SU. As the scalable unit grows, VAST Data CNodes and DBoxes can be seamlessly expanded to provide additional performance scaling, data storage capacity requirements.

Scalable Unit	GPUsS	CNodes	DBox	Usable Capacity (TB)	Rack Units	Rack	Total kW
1	576	8	3	860	11U	1	6.25
2	1152	13	5	1480	18U	1	10.25
3	1728	21	8	2392	29U	2	16.5
4	2304	26	10	2992	36U	2	20.5
8	4608	51	20	5986	71U	4	40.5

Table 1. Scalable standard Configurations for VAST Data in a DGX SuperPOD

Scalable Unit	GPUs	CNodes	DBox	Usable Capacity (TB)	Rack Units	Rack	Total kW
1	576	21	8	2392	29U	2	16.5
2	1152	41	16	4781	57U	2	32.5
3	1728	61	24	7172	85U	4	48.5
4	2304	81	32	9562	113U	6	64.5
8	4608	159	63	18,832	222U	11	126.75

Table 2. Scalable Enhanced Configurations for VAST Data in a DGX SuperPOD

Storage Performance Guidance from NVIDIA

Performance requirements for high-speed storage greatly depend on the types of AI models and data formats used. NVIDIA DGX SuperPOD has been designed as a capability- class 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.

NVIDIA provides guidelines for DGX SuperPOD workload IO based on the following table. The numbers here are meant to represent the best-case scenario and can be used to help understand the workload that a single SU would generate in terms of IO load. See the table below for more detailed information. To allow for customers to characterize their own performance requirements, some general guidance on common workloads and datasets is shown Table 2.

Level	Work Description	Data Set Size
Standard	Multiple concurrent LLM or fine-tuning training jobs and periodic checkpoints, where the compute requirements dominate the data I/O requirements significantly.	Most datasets can fit within the local compute system's memory cache during training. The datasets are single modality, and models have millions of parameters.
Enhanced	Multiple concurrent multimodal training jobs and periodic checkpoints, where the data I/O performance is an important factor for end-to-end training time.	Datasets are too large to fit into local compute systems' memory cache requiring more I/O during training, not enough to obviate the need for frequent I/O. The datasets have multiple modalities and models have billions (or higher) of parameters.

Table 3: Characterizing different I/O workloads

Performance estimates for the storage system necessary to meet the guidelines in Table 4 are in:

Performance Characteristic	Standard (GBps) (Text/Image)	Enhanced (GBps) (4K Video)
1 SU aggregate system read	90	280
1 SU aggregate system writes	45	140
2 SU aggregate system read	180	560
2 SU aggregate system write	90	280
4 SU aggregate system read	360	1120
4 SU aggregate system write	180	560
8 SU aggregate system read	720	2240
8 SU aggregate system write	360	1120

Summary

The VAST AI OS meets the NVIDIA DGX SuperPOD performance and functionality requirements. As an enterprise NAS solution, the VAST AI OS paired with NVIDIA DGX SuperPOD enables customers to take on demanding AI workloads without the complexity and specialized skills typically associated with HPC storage solutions.

As requirements grow the VAST Data Platform may be seamlessly scaled by adding compute and or storage resources tailored to meet performance and capacity targets. The VAST Data Platform supports multiple generations of infrastructure in a cluster, allowing customers to mitigate supply chain issues and select from a range of VAST certified hardware solutions.

NVIDIA DGX SuperPOD customers can be confident that the VAST Data Platform will meet their most challenging AI workloads on any scale.

Appendix A: VAST Data Components for 2SU

Below is an example configuration for a 2SU NVIDIA DGX SuperPOD. It is recommended to work with a VAST Data services representative to optimally configure the VAST Data platform for NVIDIA DGX SuperPOD clusters.

Compatible Hardware

SKU	Description	Quantity
1U-1N-GEN5-1NIC	Single Server 1U Chassis inc (Single NIC, Dual Port 200GbE QSFP112 or Dual Port HDR200 QSFP112)	13
DF-3015-V2	Ceres Enclosure: 1U HA Enclosure (up to 67GB/s): incl. 338TB NVMe Flash, 6.4TB SCM, 4 x 200Gb Bluefield-3 (ETH or IB)	5

Key specifications of each VAST CNode include:

- Single AMD Genoa 9454P UP (48 Cores / 96 Threads) CPU, 2.75GHz
- Memory: 384GB (12 × 32 GB) DDR5-4800
- I/O Connectivity:
    - Dual-port Connect-X 7 for frontend (client-facing & Internal-facing)
- Power Consumption: 0.5 kW Active

Key specifications of each VAST DBox include

- Each DBox contains two trays
- Each tray supports one NVIDIA Bluefield-3 DPUs
- Up to 2,700 TB maximum raw capacity per DBox
- Power Consumption: 0.75 kW Active

Appendix B: Deployment Considerations

VAST Data recommends two physical network ports per CNode and Four Logical network IP addresses per CNode for VAST Cluster deployments. This requirement is applicable to both NVIDIA Spectrum Ethernet and NVIDIA Quantum InfiniBand storage network fabric environments.

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 the 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 the 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

© 2025 NVIDIA Corporation. All rights reserved.