

High-Performance Storage Reference Architecture for NVIDIA Cloud Partners

June 2024

Table of Contents

About VAST Data and NVIDIA Cloud Partners **3**

VAST Data Platform with NVIDIA HGX H100 Servers **4**

VAST Data Platform: Hardware Components **5**

VAST CNodes 6

VAST DBoxes (High Availability NVMe Enclosures) 7

NVMe Fabric (NVMe-oF) Switches 8

VAST Data Platform: Software Architecture **9**

VAST DataStore 10

VAST DataBase 11

VAST DataSpace 11

NVIDIA Recommended Storage Sizing for NCP **12**

VAST Data Architectures for NCP Deployments **15**

Recommended Storage Networking 15

Recommended Backend Network Sizing 16

NCP Deployments with 127 HGX H100 Servers 17

NCP Deployments with 255 HGX H100 Servers 20

NCP Deployments with 1023 HGX H100 Servers 20

NCP Deployments with 2047 HGX H100 Servers 21

VAST Data Solution Performance Validation **21**

Summary **22**

Appendix A: Consolidate Storage Silos, Accelerate Insight **23**

Appendix B: The Imperative for Secure Multi-Tenant Storage **24**

Benefits of a Zero-Trust Multi-Tenant Data Platform 24

VAST Data's Secure Multi-Tenant Data Platform 25

Advanced Tenancy Capabilities 25

Cost Optimization and Performance Benefits 26

Compliance and Auditability 26

About VAST Data and NVIDIA Cloud Partners

The VAST Data Platform is an integrated set of storage and data services that enable service providers with a comprehensive catalogue of data-centric offerings, including multi-protocol data ingest, accelerated data preprocessing for feature engineering, high-performance storage (HPS) for model training, edge data management for model serving, and end-to-end cataloguing to meet audit and compliance requirements. Compared to traditional AI infrastructure with siloed storage solutions, NVIDIA® Cloud Partners (NCPs) can develop and deploy these data services to market faster while effortlessly meeting availability service level agreements (SLAs), and fortifying security with a true Zero Trust framework that fully supports multi-tenancy, fine-grained quality-of-service (QoS) controls, and tenant-specific data encryption.

While the VAST Data Platform is robust in capabilities, this document—in collaboration with NVIDIA—was created to detail a validated reference design that specifically adheres to the NCP Reference Architecture (NCP RA) specifications for HPS. The remainder of this document provides an overview of the VAST Data solution (both hardware and software) along with exact configurations that have been validated to exceed the performance requirements necessary for NCPs building AI infrastructure at scale.

VAST Data Platform with NVIDIA HGX H100 Servers

The VAST Data Platform is the foundation for some of the world's largest GPU cloud deployments.

VAST Data's support of the NCP RA builds on years of experience supporting large-scale NVIDIA HGX-based server deployments with some of the largest NCPs in the world. This fully integrated, turnkey architecture is validated at scale and is backed directly by the VAST Data global support services team.

The VAST Data Platform uniquely marries the high performance and scalability demanded by AI workloads with the simplicity of enterprise manageability and resiliency needed by CSPs. The result is an ideal platform for service operators thanks to VAST Data's industry leadership in the HPS market both in terms of production uptime as well as total cost of ownership (TCO). The VAST Data Platform gives customers native support for industry standard protocols like NFS, SMB, and S3, enhancing the agility of CSPs to support demanding AI and HPC workloads seamlessly and efficiently.

The VAST Data Platform architecture is engineered to fully exploit the performance and operational characteristics of NVMe flash storage. By pioneering new, low-overhead, erasure codes and advanced data reduction technology, VAST reduces the costs of all-NVMe storage to a level comparable to the combined cost of tiered, multi-cluster solutions. This allows NCPs to realize the performance benefits of flash at archive-tier economics. With the VAST Data Platform, NCPs can leverage the simplicity of standard file and object protocols, while streamlining AI pipelines and eliminating the complexity and security challenges of managing a multi-tenant environment.

Benefits for NVIDIA Cloud Partners:

Service provider-grade reliability
Observed production systems realizing 99.9999% availability

Secure multi-tenancy
Per tenant data encryption, flexible network segmentation, robust audit capabilities

Fine-grained workload isolation
Granular QoS policies prevent multi-tenant IO contention ("noisy neighbour")

Infrastructure cost savings
Consolidate storage silos into a single, affordable all-flash solution and permanently eliminate the overhead of data copies and sprawl

Improved operational efficiency
Online operations (such as upgrades) never require downtime, leading to better customer experience with fewer admins required

VAST Data Platform: Hardware Components

The VAST Data Platform, engineered with the industry's first Disaggregated, Shared-Everything (DASE) architecture, comprises three key physical infrastructure components: CNodes (compute), DNode (capacity), and NVMe Fabric (NVMe-oF) switches. This architecture enables a ground breaking approach to storage systems design by enabling the ability to independently scale performance (via CNodes) and capacity (via DNode). This flexibility for service providers—especially those with growing or expanding infrastructure needs—alleviates common capacity forecasting challenges while simultaneously meeting the needs of evolving AI workloads.



VAST CNodes

Cluster of x86-based compute nodes each running a collection of stateless containers that handle all aspects data access and management.



NVMe Fabric Switches

Low-latency, high-bandwidth connectivity providing RDMA access from every CNode to every SSD within each DNode across the cluster.



VAST DBoxes

High-density NVMe enclosures in a high-availability configuration to store all file and object storage data as well as metadata in a single namespace.

VAST CNodes

VAST CNodes constitute the core processing units of the VAST Data Platform. CNodes are stateless, containerized servers that manage all critical functions, including protecting user data via erasure coding, performing background data reduction, processing complex database queries, and handling file and object storage requests.

CNodes handle every processing request without ever persisting data or metadata locally (in fact, there are no local disks except for boot drives). Instead, every CNode has access to every remote SSD in the cluster as if it were locally attached. The stateless nature of the CNode design is what enables truly linear performance scaling (no east-west communication in the data path), dramatically better reliability (failed CNodes do not trigger disk rebuilds), as well as simplified management (all maintenance operations—including upgrades—are done completely online).

Each CNode today leverages industry-standard x86 processors, equipped with PCIe Gen 4.0 interfaces to maximize data throughput and minimize latency. Up to 4 CNodes are housed within a 2RU (rack units) CBox, ensuring compact and efficient use of data center space. The design is optimized to support intense I/O operations, with each node contributing processing power to the aggregate throughput of an entire cluster.



Key specifications of each VAST CBox include:

- **Servers:**
  4 x Stateless CNodes
- **I/O Connectivity:**
  Dual-port Connect-X 7 for frontend (client-facing)
  Dual-port Connect-X 6 for backend (internal-facing)
- **Physical CPU Cores:**
  32 x 2.9GHz cores
- **Memory:**
  16 × 16 GB 2400MHZ RDIMM
- **Form Factor:**
  2 Rack Units (RU)
- **Power Consumption:**
  2.1 kW Active

VAST DBoxes (High Availability NVMe Enclosures)

VAST DBoxes are high-density, 100% NVMe, highly-redundant systems that store all the persistent state (both data and metadata) of the cluster. VAST DBoxes house a combination of Storage Class Memory (SCM) and hyperscale NVMe flash SSDs and present each drive—via multiple Ethernet or InfiniBand network interfaces—onto an internal, ultra-low-latency NVMe fabric.

One specific DBox model, “Ceres”, leverages NVIDIA BlueField DPUs to enhance networking capabilities and offload essential data processing tasks, enhancing security and efficiency. These DPUs are instrumental in managing the network flow and securing the data paths, ensuring robust HPS operations.



Key specifications of each 1RU Ceres DBox include:

- **Severs:**
  Each DBox contains 4 x DNodes
- **I/O connectivity:**
  2 ports from BlueField DPU (Connect-X 5)
- **Processors:**
  1x NVIDIA BlueField DPUs
- **Total Capacity:**
  Up to 1,375 TB maximum raw capacity per DBox
- **Drives:**
  Up to 22 x E1.L (9.5 mm) Hyperscale QLC drives and 8 x SCM drives
- **Form Factor:**
  1 Rack Unit (RU)
- **Power Consumption:**
  700W Active

NVMe Fabric (NVMe-oF) Switches

NVMe-oF technology is the backbone of the VAST DASE architecture, enabling high-performance, low-latency network connectivity between CNodes and DBoxes. This technology allows VAST CNodes to access storage devices (within DBoxes) across a high-speed network, as if they were locally connected in the same physical server. By leveraging Remote Direct Memory Access (RDMA), NVMe-oF provides the essential high-speed data transfer capabilities required to meet the demands of data-intensive applications.

NVMe-oF is integral to ensuring that each CNode can efficiently interact with every SSD across all DBoxes in the deployment, regardless of physical location within the data center. This setup supports the disaggregated storage architecture of the VAST Data Platform, where data processing and capacity resources are scaled and managed independently.

VAST Data ships and supports the following NVIDIA Networking switches as part of every deployment:

Switch	Model	Datasheet
NVIDIA Spectrum™ SN4600	MSN4600-VS2FC	https://www.nvidia.com/content/dam/en-zz/Solutions/networking/br-sn4000-series.pdf
NVIDIA Spectrum SN5600	SN5600	https://nvdam.widen.net/s/mmvbnpk8qk/networking-ethernet-switches-sn5000-datasheet-us

Table 1. Supported backend fabric switches for VAST deployments

VAST Data Platform: Software Architecture

The VAST Data Platform, renowned for its innovative approach to data management, integrates seamlessly with the robust physical components of the VAST DASE architecture. Utilizing containers on CNodes to encapsulate operational logic, each DBox processes I/O requests efficiently, ensuring high performance and streamlined data handling. The platform's unique use of wide erasure code stripes that span multiple DBoxes and similarity-based data reduction across the cluster significantly reduces storage overhead and total cost of ownership, making it a cost-effective solution for large-scale data environments.

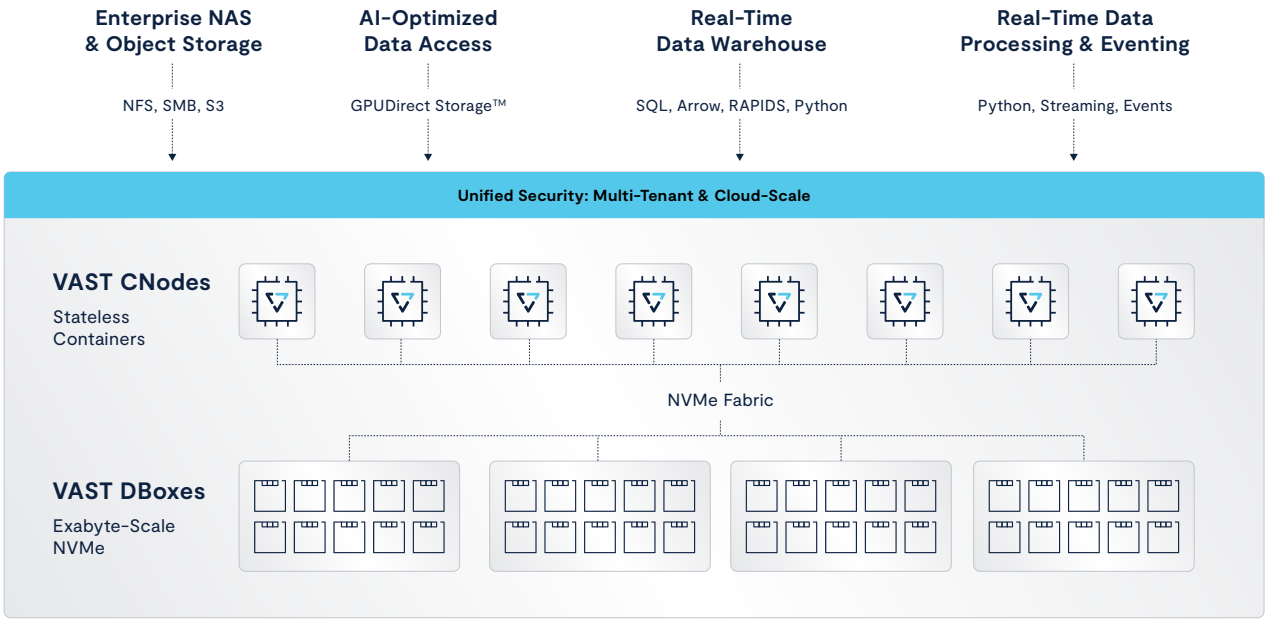

Figure 1. VAST's DASE architecture disaggregates CPUs and connects them to a globally accessible pool of SCM and ultra-dense NVMe flash SSD

Communication within the platform is facilitated through a semantic layer on high-speed SCM drives, allowing CNodes to operate without cross-talk and ensuring data integrity. This setup enhances performance and supports simultaneous access to data via multiple protocols such as NFS, SMB, S3, NFSoRDMA, and NVIDIA Magnum IO GPUDirect™ Storage (GDS). The robust security framework of the platform adheres to zero-trust principles, incorporating per-tenant encryption and comprehensive access controls, including QoS, ABAC, RBAC, and audit logging.

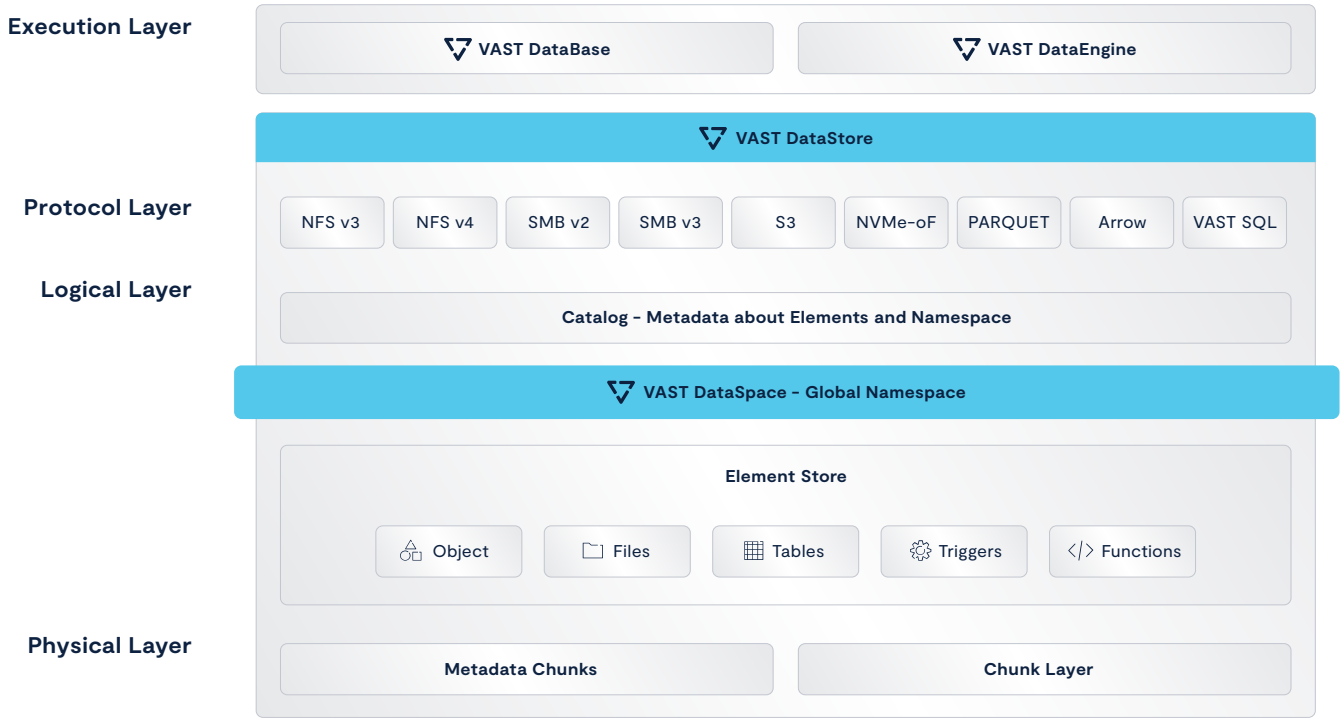

Figure 2. VAST Data Platform is a single namespace solution that provides multi-protocol file and object storage (DataStore), high-performance data warehousing (DataBase), and multi-site global data access (DataSpace)

The VAST Data Platform provides a secure, scalable, and economically efficient foundation for data storage. While there is a multitude of capabilities within the platform worth discussing, there are three critical components that will be discussed in further detail in this section: VAST DataStore, VAST DataBase, and VAST DataSpace. These components collectively expand an NCP's offerings and platform capabilities for managing and analyzing unstructured and structured data sets across AI and data pipelines.

VAST DataStore

The VAST DataStore is a scalable, HPS solution with support for multiple file (NFS, SMB) and object (S3) interfaces. For peak performance during intensive tasks like AI model training, the DataStore facilitates RDMA access over NFS (v3/v4) and employs NFS multipathing, enabling terabyte-per-second throughput from a single NFS mount point. Notably, it was the industry's first system to integrate GDS through standard NFS-over-RDMA, marking a significant technological advancement. With its native implementation of an S3-compatible object store, the VAST DataStore can also simultaneously be leveraged for data collection and ingest as part of a broader data lake offering.

VAST DataBase

The VAST DataBase is an exabyte-scale data warehouse capable of delivering robust performance for both transactional and analytical workloads. It excels in environments with intensive data preprocessing needs by accelerating data access for frameworks such as Apache Spark, Trino, and NVIDIA Data Science (via RAPIDS-enabled applications). Featuring advanced predicate pushdowns (filtering), the VAST DataBase optimizes query processing and retrieval times, thus streamlining ETL/ELT stages and shortening AI pipelines end-to-end. It also powers the VAST Catalog, which automatically indexes metadata for every file and object within the DataStore, enhancing capabilities for audit, reporting, and data provenance essential for AI regulatory compliance.

VAST DataSpace

VAST DataSpace extends the platform's capabilities into a cohesive environment for global data management, offering a unified namespace across on-prem and multiple regions across NCP data centers. It ensures consistent and reliable data access with policies that optimize data placement and synchronization while maintaining high performance. DataSpace supports multiple consistency modes for global read/write access and point-in-time data snapshots, facilitating complex data operations and recovery scenarios. The VAST DataSpace not only enhances data availability and disaster recovery but also simplifies management across diverse and geographically dispersed infrastructure, essential for large-scale, data-intensive applications.

NVIDIA Recommended Storage Sizing for NCP

The storage performance required to maximize training performance can vary depending on the type of model and dataset. The guidelines in Table 2 and Table 3 below provide guidance and targets to determine the I/O levels required for different types of models.

Performance Level	Workload Examples	Dataset Size
Good	Classical NLP workload, such as BERT	Datasets generally fit within local cache
Better	Training compressed images, compressed audio and text data, such as LLM training	Many to most datasets can fit within the local system's cache, or datasets are distributed across different nodes
Best	Training with large video and image files (such as AV replay), offline inference, ETL, generative networks such as stable diffusion, 3 images such as Medical uNet, genomics workload and protein prediction such as AlphaFold	Datasets are too large to fit into cache, massive first epoch I/O requirements, workflows that only read the dataset once, or in-place data processing

Table 2. Storage performance requirements

Performance Characteristic	Good (GBps)	Better (GBps)	Best (GBps)
Peak Single-GPU read	0.5	1	5
Peak Single-GPU write	0.25	0.5	2.5
Peak Single-node read	4	8	40
Peak Single-node write	2	4	20
Single SU aggregate system read	15	40	125
Single SU aggregate system write	7	20	62
4 SU (1k SU group) aggregate system read	60	160	500
4 SU (1k SU group) aggregate system write	30	80	250
8 SU (2k SU group) aggregate system read	120	320	1000
8 SU (2k SU group) aggregate system write	60	160	500
32 SU (8k SU group) aggregate system read	480	1280	4000
32 SU (8k SU group) aggregate system write	240	640	2000
64 SU (16k SU group) aggregate system read	960	2560	8000
64 SU (16k SU group) aggregate system write	480	1280	4000

Table 3. Guidelines for storage performance

While NLP cases often do not require as much read performance for training, peak performance for reads and writes are needed for creating and reading checkpoint files. This is a synchronous operation and training stops during this phase. When looking for best end-to-end training performance, I/O operations for checkpoints are important to the HPS design. Modern LLM workloads have significant requirements for write performance in order to minimize time spent writing checkpoints.

As a reference, Table 4 has hypothetical calculations for required write throughput needed for LLM training of various sizes. To calculate the write throughput requirement, we assume these constants commonly seen in the industry:

- Number of bytes per parameter: **14**
- Total write time percentage of total training time: **1%**
- Seconds per checkpointing interval: **3,600s**

Number of Parameters (in Billions)	Size (TB)	Tensor Parallel Domain Size	Pipeline Parallel Domain Size	Number of Files	Total Write Rate (GB/s)	Per Node Write Rate (GB/s)	Per GPU Write Rate (GB/s)
175	2.4	8	16	128	68.06	4.25	0.53
530	7.2	8	32	280	206.11	5.89	0.74
1,000	13.7	8	64	512	388.89	6.08	0.76

Table 4. Estimates of LLM checkpoint size

As with any other infrastructure project, it is always best to characterize workloads and organizational needs before finalizing performance and capacity requirements.

VAST Data Architectures for NCP Deployments

Recommended Storage Networking

The VAST Data solution adheres to the NVIDIA Converged Network Design and connects to the SN5600 leaf switches. As part of the NCP RA, SN5600 leaf switches are included in the bill of materials (BOM) articulated by NVIDIA.

When deploying the VAST Data Platform, each VAST CNode has a ConnectX-7 (CX-7) network adapter that connects up to two ports on the SN5600 leaf switches. Table 5 describes the number of ports required on the SN5600 leaf switches (and remaining ports available for expansion) to support the VAST HPS solution at a Better performance level.

There are multiple copper and optical breakout cable options from NVIDIA for connectivity between SN5600 leaf switches and the CNode CX-7 network adapters. As a default, VAST recommends MMA4Z00-NS, OSFP finned-top twin-port 2x400G SR8 multimode transceivers in the SN5600 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, which are available from NVIDIA. This results in one CBox consuming one port on the SN5600. In addition, for distances less than five meters, NVIDIA offers passive and active copper cables options.

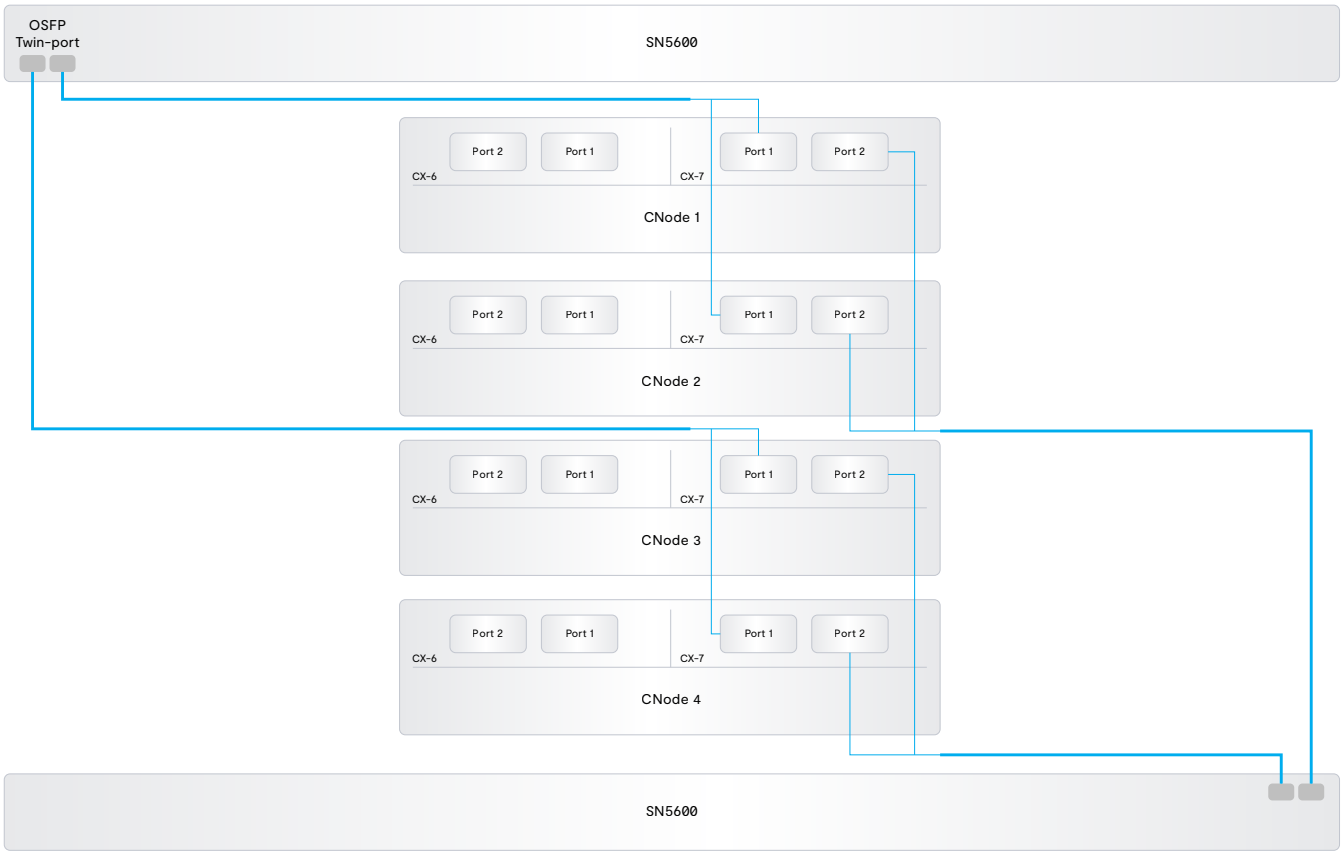

Figure 3. Switch connectivity from SN5600 leafs to CX-7 adapters across four CNodes (one CBox)

GPUs	Scales To	SN5600 200 GbE Leaf Ports Used	SN5600 200 GbE Leaf Ports Available
1K	4K	80	240
2K	4K	160	160
4K	8K	320	0
8K	8K	640	0
16K	16K	1280	0

Table 5. Number of ports consumed from the CNF at various scale points of the VAST solution

Recommended Backend Network Sizing

With the DASE architecture, every CNode also contains a network adapter that connects to a backend (BE) fabric that is internal to the VAST Data Platform. This BE fabric is used exclusively for communication between CNodes and DBoxes within a cluster. Please note that DBoxes are solely connected to this BE fabric and should not be connected to the SN5600 leaf switches discussed in the previous section.

Table 6 details the BE fabric switching requirements across various compute scale points. The network switches required for the BE fabric are included in the BOM provided and supported by VAST Data as a part of the overall HPS solution.

	GPUS			
	1K	2K	8K	16K
Leaf Switch (SN4600)	4	6	22	44
Spine Switch (SN4600)	4	4	8	16

Table 6. Number of switches for VAST's backend fabric across the various compute scale points

VAST does not span leaf to CBox or DBox connectivity in the BE fabrics between racks, therefore the number of leaf switches in the design will vary based on the number CBoxes and DBoxes per rack. The above designs assume a maximum of eight CBoxes and six DBoxes per rack. There should not be fewer than 4 leaf or spine switches in the converged network fabric or BE fabric to ensure that the loss of a single switch does not cause more than a 25% impact to available bandwidth.

Tenant isolation of client connectivity is at the discretion of the NCP. The VAST HPS solution has been validated with VXLAN in alignment to the NCP RA converged network fabric design.

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.

NCP Deployments with 127 HGX H100 Servers

The below table specifies three different VAST Data configurations that meet and exceed the storage requirements outlined in Table 3 for 4 SU (1K SU group) compute clusters.

	Good	Better	Best
VAST CBoxes	4	10	30
VAST DBoxes	3	8	25

Table 7. Recommended VAST configurations for 4 SU compute clusters[1]

While VAST Data can scale linearly to meet the performance needs of any workload, the HPS storage requirements are driven primarily based on the size of the model being trained and the dataset associated.

For example, a 4 SU compute cluster is typically best suited for LLM training with model sizes of 175B parameters or smaller. With this example in mind and based on the checkpointing table detailed earlier in the document, both the Better and Best configurations beat the target of achieving less than 1% of total training time writing checkpoints. **For NCPs targeting similar LLM training workloads, the Better configuration is ideal as it exceeds the necessary performance targets for maximum GPU utilization.**

[1] Upcoming VAST v5.2 (ETA Q3 2024) will further improve write performance on existing clusters by 70% (reducing the overall hardware required by 40%).

The following diagram outlines the networking architecture for the recommended VAST configuration:

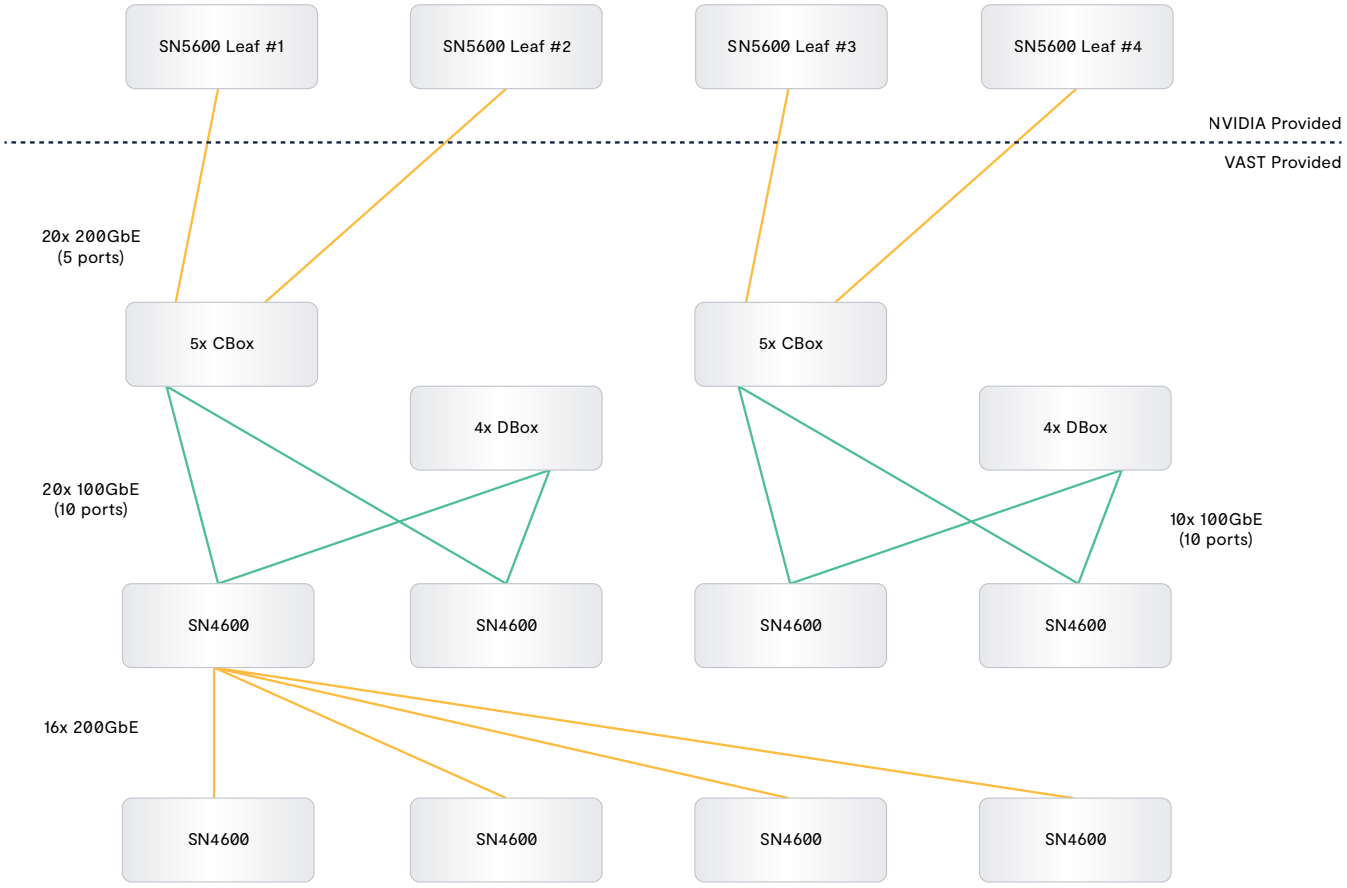

Figure 4. Networking architecture (both frontend and backend) for the recommended VAST configuration



Rack Specifications

Mgmt ports	34
C13 PDU port	14
C19 PDU port	10
Max Power	18.2 kW
Normal Power	15.6 kW
Weight	862 lbs

chassis, not cables or rails



Rack Specifications

Mgmt ports	34
C13 PDU port	14
C19 PDU port	10
Max Power	18.2 kW
Normal Power	15.6 kW
Weight	862 lbs

chassis, not cables or rails

NCP Deployments with 255 HGX H100 Servers

The below table specifies three different VAST Data configurations that meet and exceed the storage requirements outlined in Table 3 for 8 SU (2K SU group) compute clusters.

	Good	Better	Best
VAST CBoxes	8	20	62
VAST DBoxes	6	16	50

Table 8. Recommended VAST configurations for 8 SU compute clusters[2]

The below table While VAST Data can scale linearly to meet the performance needs of any workload, VAST and NVIDIA recommend using the Better configuration for AI Model training with compressed images, compressed audio and text data, such as LLM training.

NCP Deployments with 1023 HGX H100 Servers

The below table specifies three different VAST Data configurations that meet and exceed the storage requirements outlined in Table 3 for 32 SU (8K SU group) compute clusters.

	Good	Better	Best
VAST CBoxes	30	80	250
VAST DBoxes	24	64	200

Table 9. Recommended VAST configurations for 32 SU compute clusters[2]

While VAST Data can scale linearly to meet the performance needs of any workload, VAST and NVIDIA recommend using the Better configuration for AI Model training with compressed images, compressed audio and text data, such as LLM training.

[2] Upcoming VAST v5.2 (ETA Q3 2024) will further improve write performance on existing clusters by 70% (reducing the overall hardware required by 40%).

NCP Deployments with 2047 HGX H100 Servers

The below table specifies three different VAST Data configurations that meet and exceed the storage requirements outlined in Table 3 for 64 SU (16K SU group) compute clusters.

	Good	Better	Best
VAST CBoxes	60	160	500
VAST DBoxes	48	128	400

Table 10. Recommended VAST configurations for 64 SU compute clusters[3]

A 64 SU compute cluster can support LLM training with models of all sizes, including those greater than 1T parameters. New research is pushing the boundary of model sizes beyond several trillion parameters but often employs different types of checkpointing techniques (i.e., async) that are beyond the scope of this document. However, if we were to assume a T parameter-sized model trained on a 64 SU compute cluster, then all three configurations (Good, Better, and Best) beat the target of achieving less than 1% of total training time writing checkpoints.

VAST Data Solution Performance Validation

The solution outlined in this reference architecture document was validated by NVIDIA on a VAST Data v5.1 cluster. A benchmarking test suite from NVIDIA was utilized to characterize overall cluster performance typical of long-running AI model training jobs across various compute scale points. During this testing, VAST validated the write demands of large-scale checkpointing, as well as the read demands during token loading and restoring from a checkpoint. All results of this testing were submitted to and validated by NVIDIA.

For a more in-depth look on checkpointing, please see:
https://www.vastdata.com/blog/a-checkpoint-on-checkpoints-in-llms

In addition to the benchmark results, VAST has provided NVIDIA real-world production examples—hosted at NCPs—showing load profiles of various AI model training jobs running across six thousand NVIDIA GPUs in parallel, training one trillion parameter models backed by VAST Data for HPS.

Lastly, a variety of scaling, resilience, and QoS tests were performed. It was demonstrated that the system could linearly scale between performance levels (i.e., from Good to Better), as well as based on the number of GPUs (i.e., from 1K to 2K) without rebuilding or reinstalling and that cluster performance was not negatively impacted while expanding. The VAST cluster could also tolerate loss of network interfaces, CNodes, and DNodes with only a proportional impact, no greater. Finally, QoS levels could be enforced while read and write operations were on-going. These results were also reviewed and validated by NVIDIA as exceeding expectations.

[3] Upcoming VAST v5.2 (ETA Q3 2024) will improve write performance on existing clusters by 70% (reducing the overall hardware required by 40%).

Summary

The VAST Data Platform pioneered an architecture that delivers not only performance and scale but also operational efficiency and ease of use. With flexible configurations to support over 16K GPUs in a single cluster, NCPs can confidently pair VAST Data solutions with any scale of large AI infrastructure deployments. Based on the NCP RA, featuring NVIDIA HGX H100 platforms, this jointly validated VAST Data reference architecture meets and exceeds the intensive demands of AI, high-performance computing, and other data-intensive applications.

This document details a validated reference design from VAST Data that specifically adheres to the NCP RA specifications for HPS. The VAST Data Platform also provides additional data services—such as multi-protocol data ingest, accelerated ELT for feature engineering—useful in helping CSPs accelerate data pipelines for their customers, all underpinned by a multi-tenant Zero Trust security model. This holistic approach not only streamlines AI pipelines by consolidating storage across various data stages but also significantly minimizes complexity and costs, thereby boosting overall system performance and security for CSPs.



For more information on Universal Storage and how it can help you solve your application problems, reach out to us at hello@vastdata.com.

©2024 VAST Data, Inc. All rights reserved. All trademarks belong to their respective owners.

Appendix A: Consolidate Storage Silos, Accelerate Insight

Because the VAST Data Platform supports file, object, and tabular data formats, organizations can consolidate high-performance storage (training data), data management systems (feature stores, model artifacts), and data lake (data collection and preprocessing) storage requirements onto a single platform. By eliminating data transfers across separate storage systems, AI training, inference, and model serving processes are simplified, enhancing speed and reliability. This convergence of data formats and workflow stages streamlines the entire AI data lifecycle, allowing organizations to focus on innovation rather than managing complex storage architectures and cumbersome data copy management tasks.

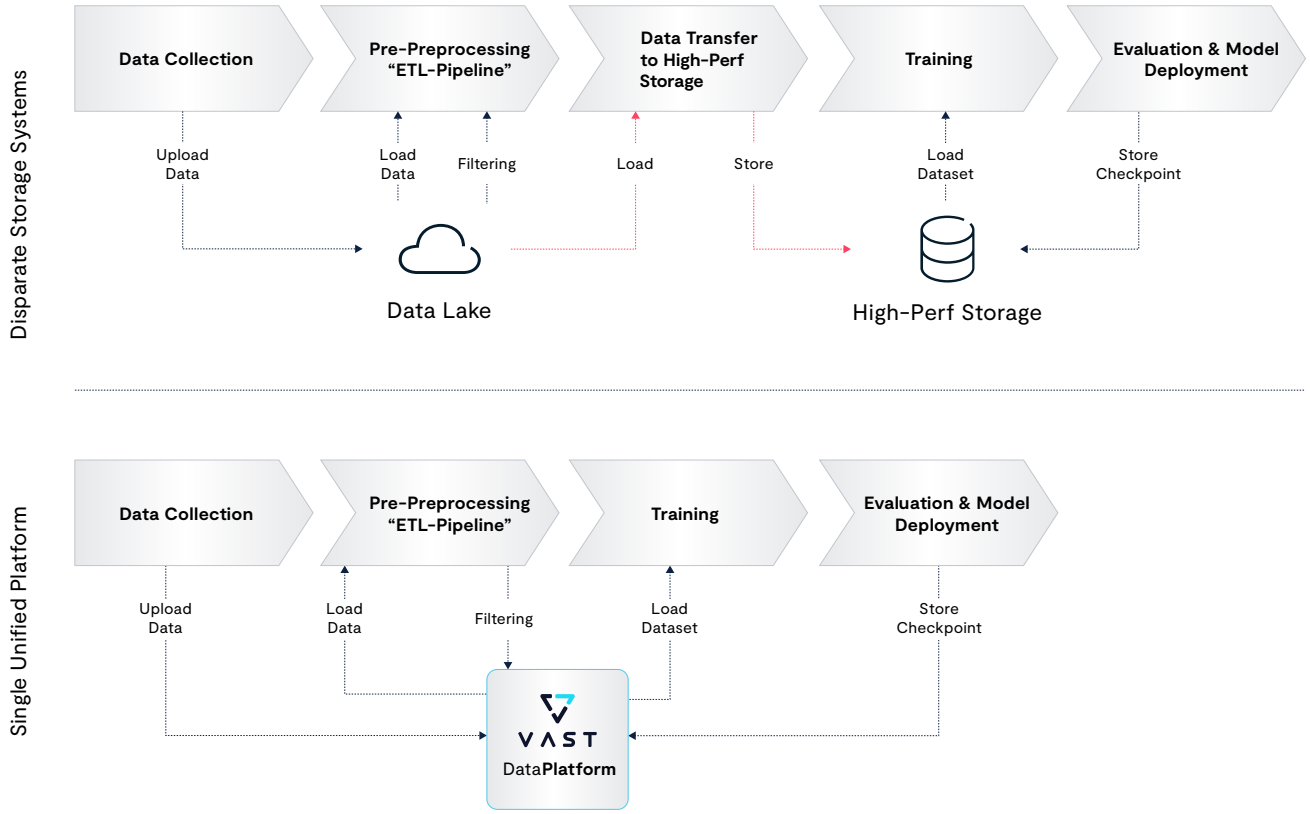

Figure 5. (Top) Typical workflow of an AI pipeline and corresponding storage locations requiring data transfer between storage systems. (Bottom) Simplified AI workflow with the VAST Data Platform whereby data is accessed in place and is available over file and object protocols

By consolidating the various storage systems into a single, unified data platform, VAST enables NVIDIA Cloud Partners to offer a highly differentiated platform that accelerates the entirety of an AI pipeline. This empowers NCPs to create and deliver services to their customers faster and with lower operational overhead, compared to traditional multi-cluster, tiered storage solutions. Additionally, NVIDIA Cloud Partners can improve the security posture of their service offerings by reducing the attack surface and eliminating the need to manage complex permissions across multiple systems.

Appendix B: The Imperative for Secure Multi-Tenant Storage

When deploying the VAST Data Platform in an NCP RA environment, especially if constructing a technology stack around NVIDIA's HGX-based AI compute systems, selecting the right data services layer is essential. This layer must not only provide the high performance needed to power NVIDIA GPUs but also incorporate multi-tenant storage capabilities to enhance the service's efficiency, security, and scalability. By implementing full-stack multi-tenancy, including at the data services level, cloud providers can significantly reduce both capital expenditures (CAPEX) and operational expenditures (OPEX), leading to reduced administrative costs and increased profit margins.

Because CSP customers will be working with sensitive, proprietary and oftentimes personal information of their own customers, these systems must meet regulatory requirements as well to satisfy GDPR, PCI DSS, HIPAA, FedRamp, and Zero Trust Controls. The regulatory requirements are based on a common set of controls that are developed and maintained by NIST. Let's take a look at key controls for a Storage-as-a-Service (STaaS) offering that will support AI/ML workloads.

The key overlapping controls include:

- Access Control Policies (NIST 800-53r5) and Data Authorization (NIST 800-207r2)
- Encryption and Key Management (NIST 800-53r5), Data Protection (NIST 800-207r2), and Data Encryption (NIST 800-207r2)
- Audit and Accountability (NIST 800-53r5) and Audit and Monitoring (NIST 800-207r2) Incident Response (NIST 800-53r5)
- System and Communications Protection (NIST 800-53r5) and Segmentation (NIST 800-207r2)
- Identity and Authentication (NIST 800-53r5)

While there are some differences in the specific terminology used, the underlying principles and objectives of these controls are closely aligned between ZTA and FedRAMP. Both frameworks aim to protect sensitive data, enforce strict access controls, maintain continuous monitoring and auditing, and ensure the confidentiality, integrity, and availability of systems and data.

Benefits of a Zero-Trust Multi-Tenant Data Platform

Enhanced Data Security and Privacy: Ensures data isolation and protection, providing robust security mechanisms that keep each tenant's data private and secure, critical for complying with industry regulations such as GDPR and HIPAA.

Cost Efficiency and Resource Optimization: Allows for the efficient use of storage resources by pooling them and dynamically allocating these resources among multiple tenants based on demand, reducing the need for dedicated storage systems and overprovisioning

Scalability and Flexibility: Offers scalable storage options that can easily expand to meet increasing demands, ensuring the AIaaS offering can support a wide range of applications, from small-scale experiments to large, enterprise-level AI deployments

Simplified Management and Maintenance: Managing a single, multi-tenant storage infrastructure is far simpler and more cost-effective than maintaining separate storage systems for each customer, reducing administrative overhead and the risk of human error

Performance and Availability: Designed for high-performance workloads, these solutions deliver the high throughput and low latency required for AI and ML tasks, with built-in redundancy and failover capabilities to ensure high availability and data protection

Compliance and Auditability: Provides the necessary tools and frameworks to ensure that data handling and processing meet regulatory standards, including features for data encryption, access controls, and audit trails, which are vital for demonstrating compliance with regulations and for conducting audits

VAST Data's Secure Multi-Tenant Data Platform

VAST Data's platform is the industry's only secure multi-tenant data platform built for large-scale AI clouds. It delivers a comprehensive solution for NVIDIA Cloud Partners seeking to implement a zero-trust, multi-tenant infrastructure. The platform integrates multi-category security (MCS) and a hybrid model of Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC), providing granular access controls and ensuring strict data isolation among tenants.

Advanced Tenancy Capabilities

The VAST Data Platform offers a range of advanced tenancy capabilities that enable cloud providers to deliver secure, isolated, and dynamically provisioned AIaaS offerings:

Dynamic Provisioning: Supports dynamic provisioning of storage resources, allowing cloud providers to allocate and deallocate storage capacity on-demand based on tenant requirements

Network Segmentation: Provides robust network segmentation capabilities, creating isolated virtual networks for each tenant to ensure that tenants' data traffic remains separate and secure

Micro-Segmentation: Supports micro-segmentation, enabling fine-grained control over access to specific data sets and resources within a tenant's environment

Data Separation via View and Tenant-Level Encryption: Offers data separation at both the view and tenant levels, with tenant-level encryption ensuring that each tenant's data is encrypted with unique keys, further enhancing data isolation and security

Automated Data Labeling: Administrators can define labels/attributes at the "View" level for data classification and authorization

Granular Access with Hybrid RBAC-ABAC: Combines RBAC and ABAC for finer-grained access control, supported with extensive audit capabilities and unique encryption keys for data isolation

Data Encryption and Key Management: Uses external key management with KMIP for encrypting data at rest and in transit, providing secure key exchange and management

In-Flight Encryption: Based on the TLS protocol 1.3, offering enhanced security for encrypted communications using FIPS validated encryption ciphers

NVIDIA Magnum IO GPUDirect Storage Encryption: Supports GDS with end-to-end encryption, enabling secure and high-performance data transfer between storage and GPUs

Object Storage Encryption: Supports server-side encryption (SSE) and client-side encryption for enhanced data protection

Crypto Erase: The ability to destroy encryption keys and purge data in a way that meets NIST 800.88 controls that ensures that a data set is purged

Cost Optimization and Performance Benefits

The VAST Data Platform is designed to optimize costs and maximize performance for CSPs offering AI-as-a-Service (AIaaS). Key benefits include simplified AI infrastructure, HPC performance with enterprise simplicity, data insights & management, and flexible scalability.

Compliance and Auditability

The VAST Data Platform provides the necessary tools and frameworks to ensure that data handling and processing meet regulatory standards, including features for data encryption, access controls, and audit trails, which are vital for demonstrating compliance with regulations and for conducting audits.