



NVIDIA Spectrum-X Technology for AI Storage Fabrics and the VAST Data Platform

January 2025

Table of Contents

Introduction	3
NVIDIA Spectrum-X Technology for AI Systems	3
The VAST Data Platform	3
Testing Configuration	4
Testing Environment Configurations & Variables	4
Test Types	5
Scaled Client	5
Checkpoint Restore	5
Many-To-Many	6
Noisy Neighbor	6
Results	6
Conclusion	10

Abstract:

The combination of NVIDIA Spectrum-X Technology for AI Storage Fabrics and the VAST Data Platform offers an ideal solution for the modern AI data center. Together, NVIDIA and VAST Data deliver an optimized, consistent, and congestion-reducing solution that empowers real-time agentic enterprises, AI inference and training, and various AI-driven workloads to operate at peak performance. Spectrum-X demonstrates its effectiveness as inter-peer links, client sessions, and infrastructure utilization expand.

Introduction

With the evolution of Artificial Intelligence (AI) workloads in data centers, the need for consistent data ingress and egress performance has become increasingly vital. Variability in I/O patterns can create whipsaw patterns in network performance, creating traffic spikes that disrupt the ability of multiple systems to coherently process, store, and retrieve the data vital to continued AI operations. To improve data pipeline functionality, efforts have been made to ensure that intrasystem connections between AI accelerators are as efficient as possible through the network infrastructure. VAST's Disaggregated Shared-Everything (DASE) architecture also ensures storage system efficiency between control and data planes. However, intersystem connectivity, whether through Ethernet or NVIDIA InfiniBand, remains a linkage point that needs additional attention.

NVIDIA's Spectrum-X Technology for AI Storage Fabrics addresses this challenge by delivering consistent performance to interconnected systems. This capability is a steppingstone for enterprises, hyperscalers, cloud and managed services providers, and other AI-enabled organizations that require performance and consistent operation.

NVIDIA Spectrum-X Technology for AI Systems

Spectrum-X is the network technology that brings to Ethernet what has been standard practice in InfiniBand for decades: Adaptive Routing (AR) and Congestion Control (CC).

AR and CC techniques meaningfully improve the sustained bandwidth of AI workloads through the Ethernet fabric. These capabilities enhance AI workloads' network performance and improve RDMA over Converged Ethernet (RoCE) storage traffic, radically impacting AI storage.

The VAST Data Platform

The VAST Data Platform, supported by the DASE architecture, exemplifies the peak of AI-enhanced data platforms tailored for the rise of the real-time agentic enterprise. By disaggregating control and data planes, leveraging NVMe over Fabrics (NVMeoF) for interconnections, and utilizing NVIDIA BlueField®-3 DPUs for storage, VAST improves data traffic optimization with adaptive routing and congestion management to ensure efficient data access for the AI-era. This proven platform is at the nexus of frontier model providers like CoreWeave and has demonstrated capabilities suited for the ever-evolving AI-centric data center.

Testing Configuration

Testing Environment Configurations & Variables

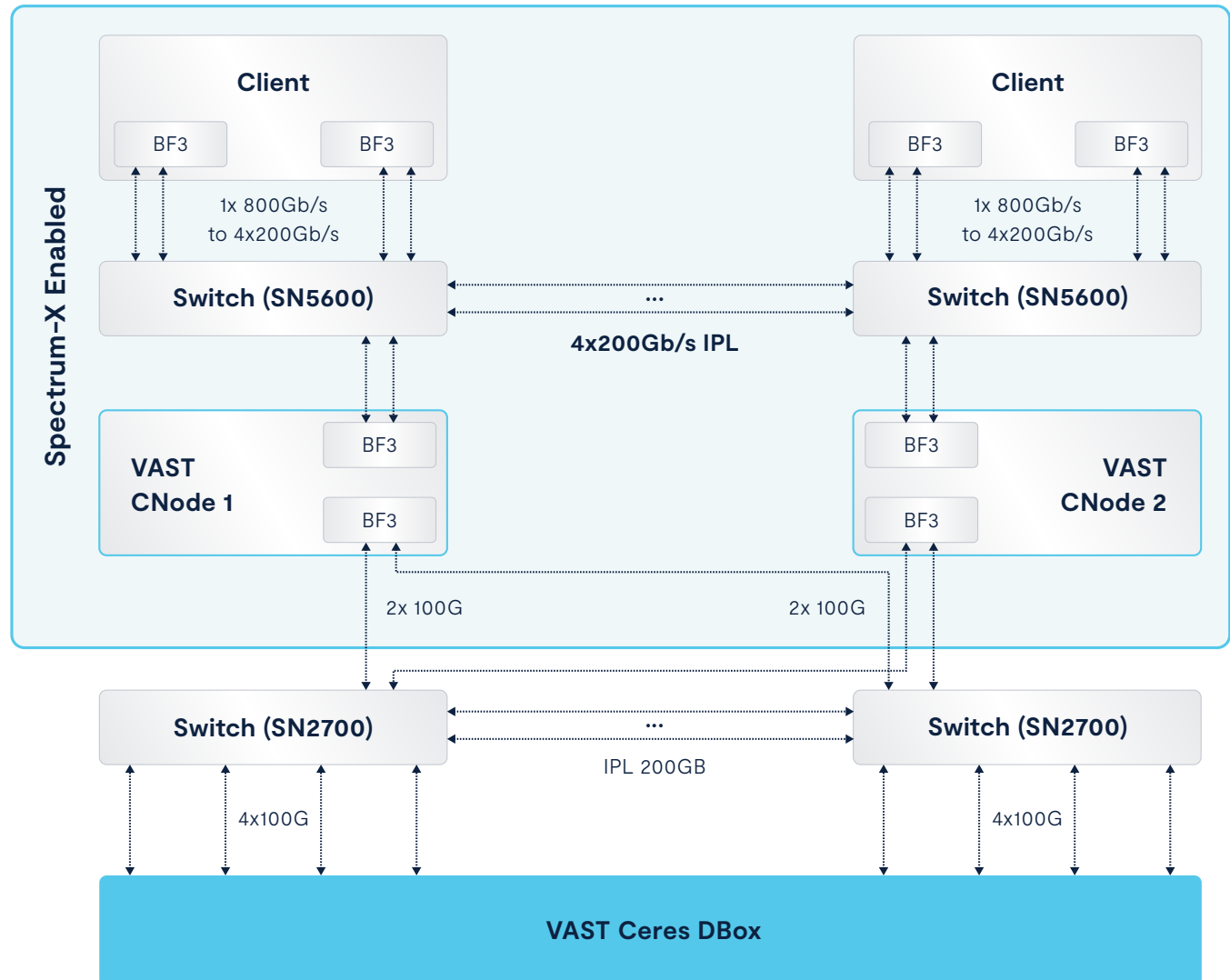


Figure 1 Spectrum-X and VAST Data Platform Testing Environment

The overall architecture utilized for testing is illustrated in Figure 1. It consists of the following components:

- Two clients, each containing four NVIDIA BlueField®-3 DPUs configured as NVIDIA BlueField®-3 SuperNICs with 200Gb links
- Two NVIDIA SN5600 switches with an interlink (IPL) bandwidth of 4x200Gb
- Two VAST CNodes running VAST OS 5.2 with NFS driver version 4.0.30, featuring two BF3 DPUs configured as SuperNICs with 200Gb links.
- One VAST Ceres DBox with 22 x 15Tb NVMe drives and 8 x SCM drives
- Two NVIDIA SN2700 switches with an interlink bandwidth of 2x200Gb¹

All tests were conducted using [Fio](#). Each scenario was initially configured with Spectrum-X AR and CC **enabled** across the clients, CNodes, and switches. Subsequently, the same environment was tested with Spectrum-X **disabled** to compare the performance of both configurations.

Test Types

To accurately assess Spectrum-X's effect on AI platforms, it is essential to profile data activity that closely resembles commonly encountered workloads. The effects of Spectrum-X's AR and CC are most evident when evaluating the resulting data. The following test types were employed to showcase the influence of Spectrum-X innovations on established I/O patterns commonly used for data transfer within AI-ready data centers and to compare them against a control setup without Spectrum-X integration.

The test types include *Scaled Client*, *Checkpoint Restores*, *Many-to-Many*, and *Noisy Neighbor* scenarios, all of which represent typical workloads in contemporary AI-driven infrastructures.

The descriptions and objectives of each testing regimen are outlined as follows:

Scaled Client

The scaled client workload rigorously tests connectivity among clients, CNodes, and DNodes, highlighting improvements in loading and connections. This type of test enhances the understanding of interpeer switch links (IPL), client, and CNode/DNode loading while emphasizing AR and CC's relative effectiveness on the VAST Data Platform.

Checkpoint Restore

The checkpoint restore test simulates a standard AI training and inference function: recovering a checkpoint file due to a GPU calculation error or other errata. Checkpoint files are created from a subset of active GPUs and stored within a training epoch. Restoring from a checkpoint file is a sequential, read-intensive operation that demands maximum bandwidth and minimal latency to resume training or inference.

¹ The NVIDIA SN2700 switches are currently *End-of-Life (EOL)* and have been replaced by the SN3700 series switches. More information on the SN3700 can be found [here](#).

Many-To-Many

Hash-based distribution mechanisms, such as Link Aggregation Control Protocol (LACP) or Equal-Cost Multi-Pathing (ECMP), can create imbalances across links and increase latency when aggregate link utilization exceeds approximately 70% in typical network implementations. This issue may degrade storage systems' ability to acknowledge reads or writes promptly while maintaining acceptable latency. Testing this scenario illustrates how Spectrum-X influences reads while attempting to reach utilization thresholds of 70%, 80%, 90%, and 100% of line rate.

Noisy Neighbor

Imbalances across IPL links can result when elephant flows consume disproportionate amounts of available bandwidth compared to other, less resource-intensive flows (e.g., mice flows). These imbalances can degrade latency and traffic optimization, resulting in congested links that disrupt upstream applications. This testing scenario represents a use case for multimodal content or managed service providers (CSP/MSP), where training, inference, and enterprise applications may coexist within a similar data context.

Results

The results collected from testing, based on the environment and variables noted previously, are summarized below. The evaluation begins with the performance of scaled client connections to the VAST Data Platform. Testing started with low connection and user counts, scaling incrementally to identify the pressure points. The initial testing was done with Spectrum-X AR and CC **enabled** across the infrastructure, followed by the same tests with Spectrum-X features **disabled** for comparison.

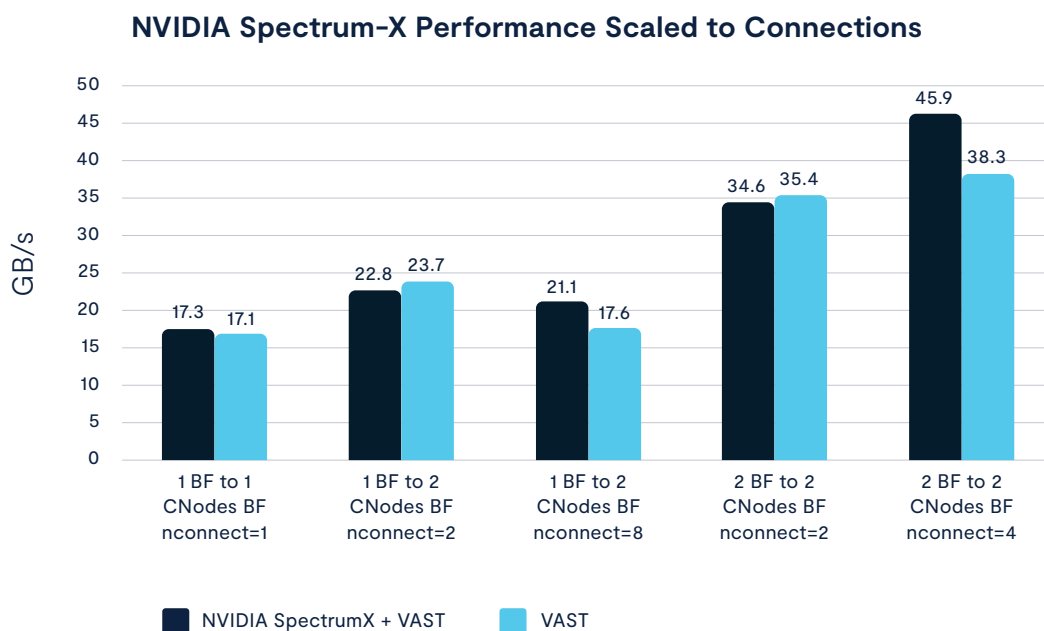


Figure 2 NVIDIA Spectrum-X Performance Scaling per Connection

As illustrated in Figure 2, the impact of Spectrum-X was exceptionally notable in scenarios with high connection counts and asymmetric traffic pressure. For instance, in the configurations *1 BF to 2 CNodes, nconnect = 8* and *2 BF to 2 CNodes nconnect = 4*, which had a performance delta of approximately 20%. These scenarios provide a salient example of the importance of optimized infrastructure, such as the NVIDIA DGX SuperPOD™, where scaled client connections need consistent writes and reads to the VAST Data Platform.

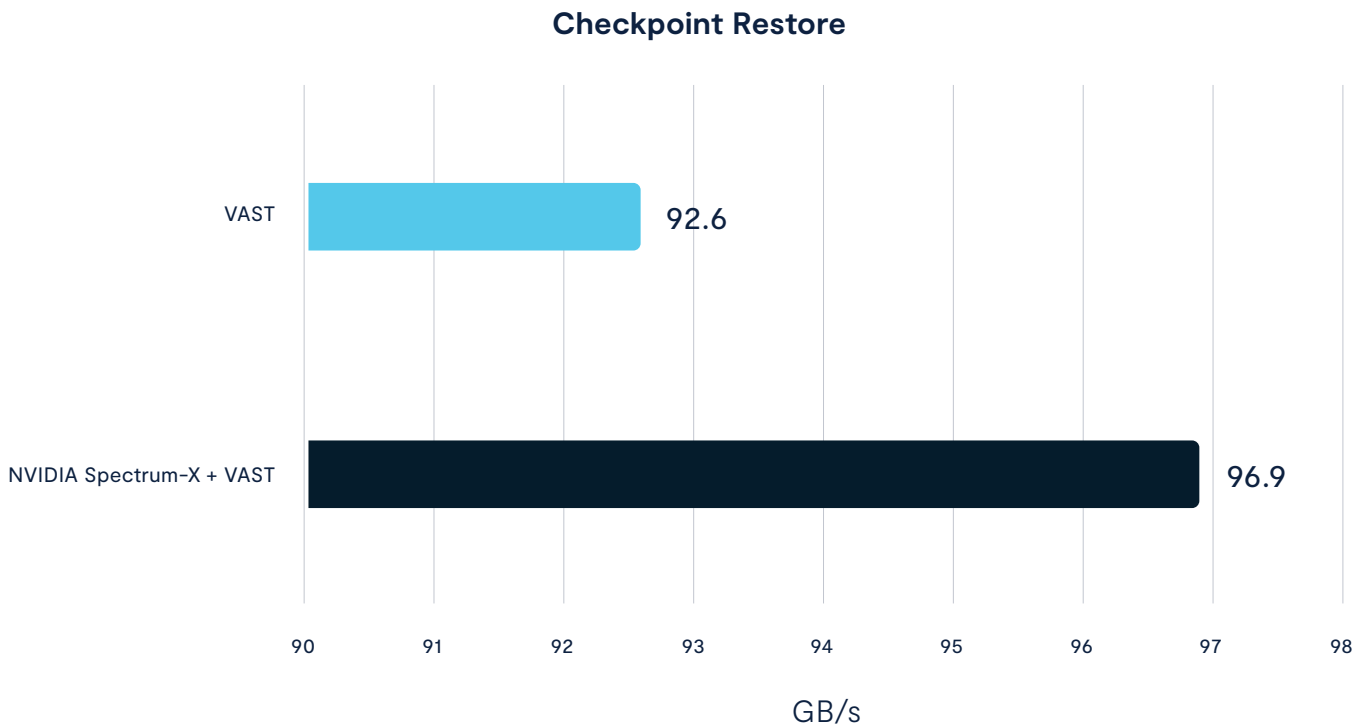


Figure 3 Checkpoint Restore

As shown in Figure 3, Spectrum-X delivered a performance advantage of nearly **5%** compared to a non-AR and CC solution across tested iterations. While this performance delta may appear insignificant at first glance, these differentials add up at scale. For example, utilizing an NVIDIA GB200 NVL72 solution requires that up to 8 GPUs out of 72 write their checkpoint files as part of a training or inference epoch. At 1.5Tb in aggregate file writes (*and reads, in the case of checkpoint restoration noted here*), this process relies heavily on high bandwidth and low latency to and from the VAST Data Platform.

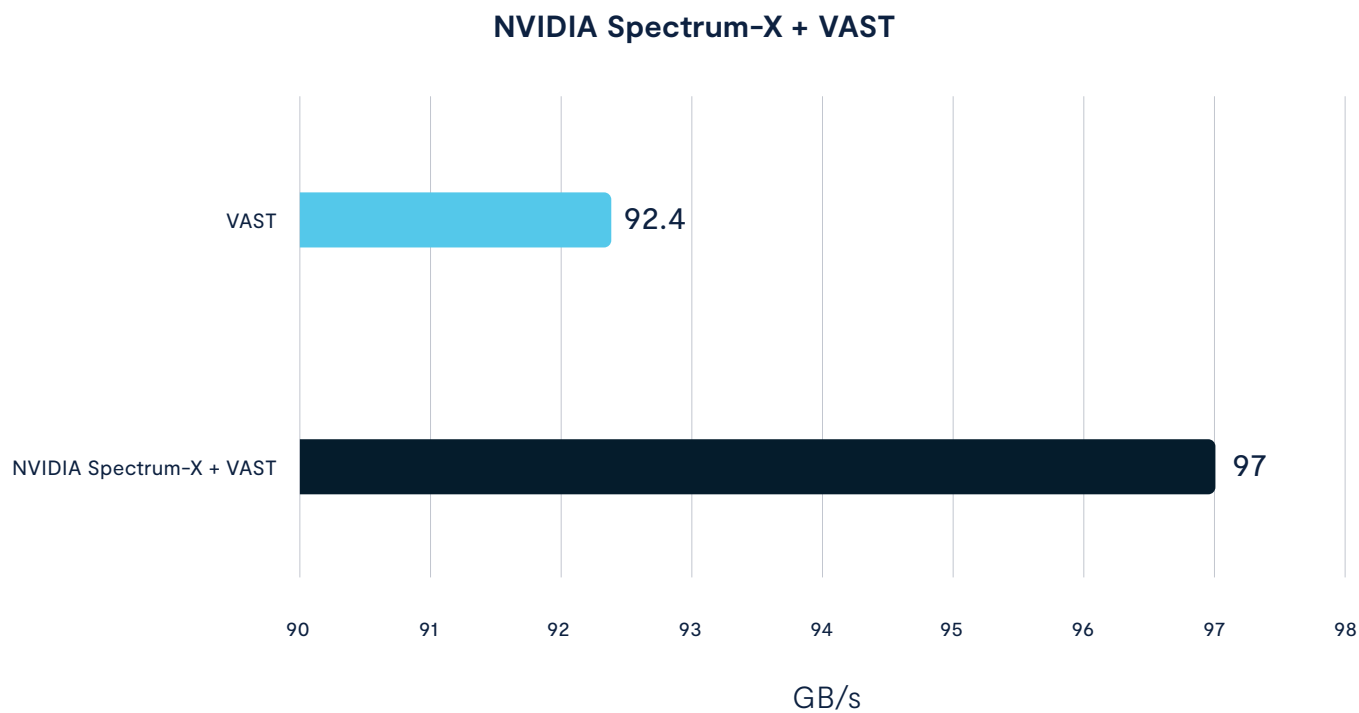


Figure 4 Many to Many Scaling

Per Figure 4, NVIDIA Spectrum-X delivered a performance advantage of nearly **5%** compared to a VAST-only solution. When considered alongside the Checkpoint Restore test, this result underscores the importance of acknowledging system read and write requests while maintaining low latency at scale—a key requirement for the modern AI data center. The impact of Spectrum-X on traffic congestion validates the initial premise outlined in this document: utilizing AR and CC within the context of a holistic storage network infrastructure is paramount for achieving a performant, AI-ready data center.

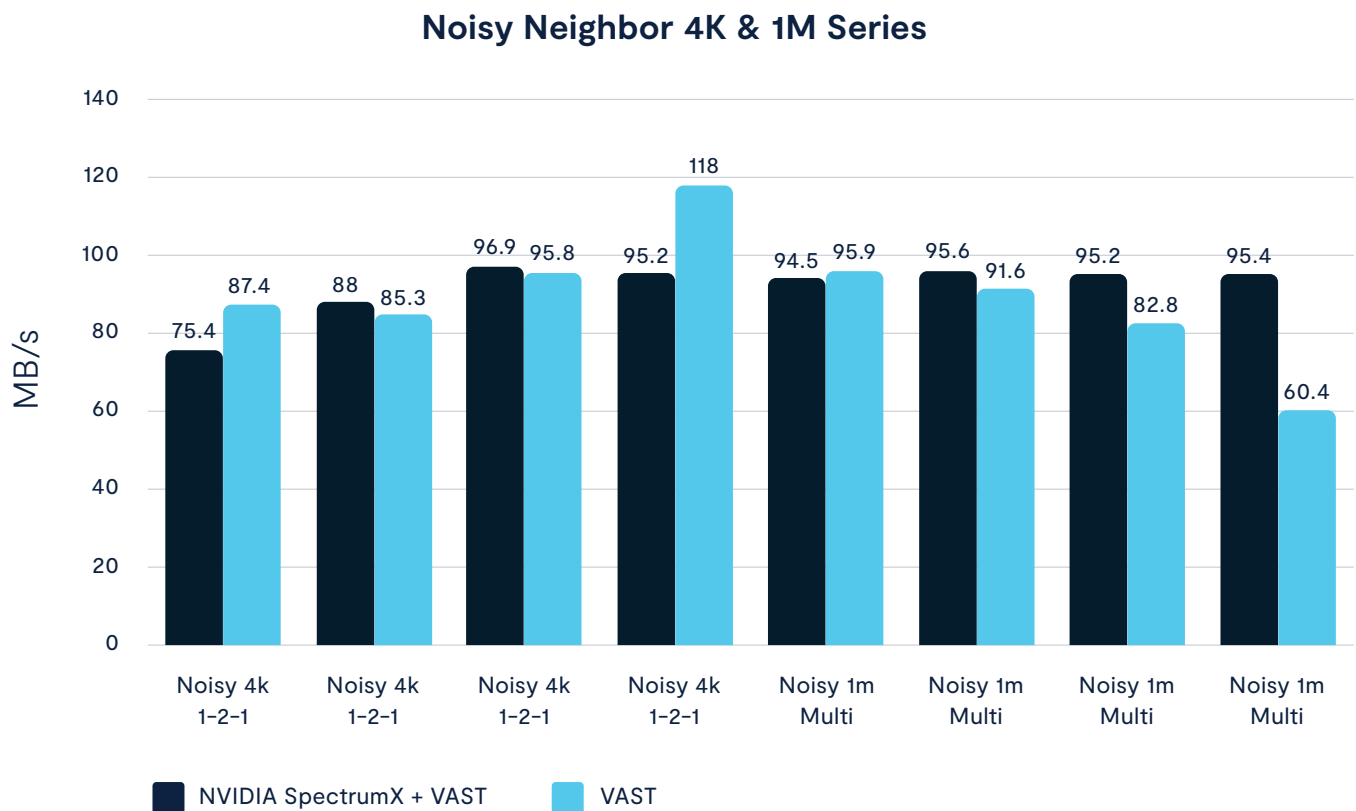


Figure 5 Noisy Neighbor 1m Series

Figure 5 highlights Spectrum-X's role in maintaining balanced traffic on IPL links to achieve a stable performance curve. Across multiple iterations of the Noisy Neighbor scenario, utilizing various block sizes, effective CC and AR demonstrated low variance between executions and minimal interference from the heavy workloads generated by a Noisy Neighbor. This shows that with heavy workloads, Spectrum-X enables better noise tolerance with up to a 37% increase in bandwidth.

Conclusion

Spectrum-X's effect on the architecture and resulting data flows is consistently present in all test scenarios. While some variances occur, the tests validate Spectrum-X's ultimate mission: balancing the storage network's I/O. This achievement directly benefits the output of the data pipelines that rely on it, as seen in the rapid proliferation of foundation models and the emerging agentic enterprises seeking to harness these capabilities.

The greatest benefit is seen when line utilization exceeds 70% and IPL links are saturated. In these scenarios, Spectrum-X's ability to adapt to increasing congestion to enable timely responses and bandwidth utilization is unmatched. In an environment where client-side software such as the VAST nconnect client is installed, the combination of increased session counts, bandwidth utilization, and storage payloads enhances the overall effectiveness of the AR and CC algorithms.

Spectrum-X does not require a hyperscaler or CSP to realize the benefits of Adaptive Routing and Congestion Control. Instead, it can be implemented more granularly, constructed from the building blocks of NVIDIA BlueField-3 DPUs integrated with AI-accelerated servers and the VAST Data Platform. By providing a scalable solution to infrastructure congestion, NVIDIA Spectrum-X Technology for AI Storage Fabrics and the VAST Data Platform are uniquely positioned to support our customers and partners as they build the future.

For more information on NVIDIA Spectrum-X Technology for AI Storage Fabrics, please visit <https://www.nvidia.com/en-us/networking/spectrumx/>

For more information on the VAST Data Platform, please visit <https://www.vastdata.com/whitepaper>



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