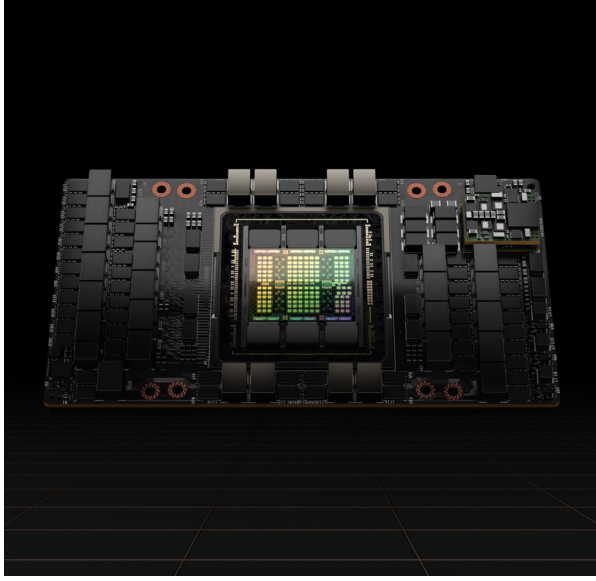
Datasheet

# NVIDIA Ethernet Storage Fabrics

Accelerated Ethernet for AI/ML storage.



## Accelerated Ethernet Networking for AI Storage Workloads at Cloud Scale

Exponential growth in data volume, real-time data processing, virtualized workloads, and AI has driven a demand for scale-out, software-defined storage infrastructure for data centers and cloud, built on faster storage devices and clustered storage servers. As faster storage needs faster networking, a dedicated Ethernet storage fabric (ESF) is adopted for needed performance and cloud scalability, with uncompromised quality of service, reliability, and security. These ESFs provide a unified network for all storage types, including block, file, object, and hyperconverged infrastructure, and for all business applications, from end-user computing, databases, and big data analytics to performance-intensive AI and machine learning.

In today's data centers, business applications are developed, virtualized, and deployed across geographic boundaries, generating and consuming large-volume data in rapidly increasing size. This paradigm shift is calling for a new storage infrastructure that's fast to expand with rapid data growth, agile to accommodate various performance requirements, and efficient to operate at cloud scale. Traditional storage, such as complex, expensive, and proprietary storage area network (SAN)-based systems, cannot meet these requirements. The result is that modern data centers are moving away from these legacy storage systems and migrating to scale-out, software-defined storage (SDS) or hyperconverged infrastructure (HCI) with faster storage devices such as NVMe SSDs.

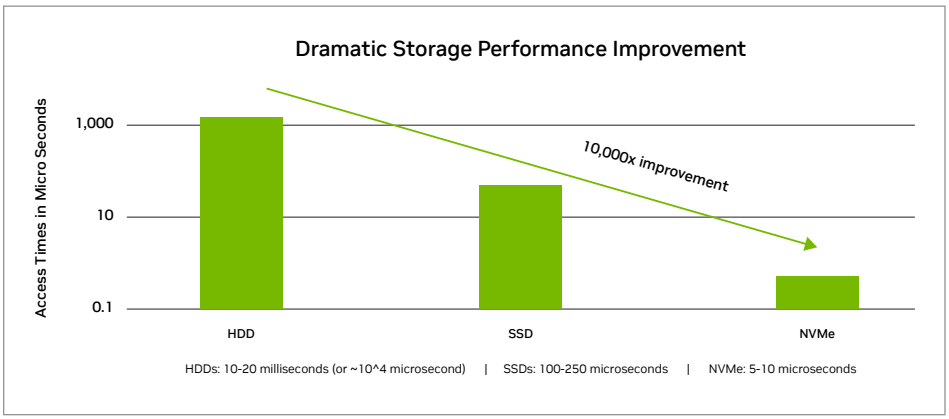



Figure 1. Faster data center storage

## Key Features

- Legacy storage infrastructure is complex and expensive to scale out.
- Traditional storage networks have performance bottlenecks and cannot cope with faster storage.
- Common data center switches aren't built for modern storage network requirements of performance and rack efficiency.

### Benefits

- Purpose-built for data center storage
- Performance acceleration for all storage types and workloads, particularly AI/ML
  - Simplified deployment and automated management
  - Accelerated time to insight and resolution

The NVMe protocol in particular is designed for orders-of-magnitude faster data access. Today's NVMe drives can deliver sustained read-write performance that easily saturates a 25 gigabits per second (Gb/s) link. Clearly, performance-bottlenecked legacy storage networks, such as Fibre Channel, cannot sustain such high-bandwidth traffic without packet drops. To overcome this challenge, dedicated ESFs have become mainstream, providing fast, scalable, reliable, and high-bandwidth networking for performance- intensive storage infrastructure.

## Ethernet Storage Fabric (ESF)

Ethernet is scalable, simple, and used ubiquitously in data centers. An Ethernet storage fabric, converging all data traffic within a scale-out infrastructure, eliminates network silos (in contrast with Fibre Channel used with legacy SAN). ESF supports all storage types—block, file, and object. Combined with remote direct-memory access (RDMA) over converged Ethernet (RoCE), ESF is optimized for performance-intensive storage systems such as NVMe over Fabrics (NVMe-oF). By accelerating storage performance over a single unified network fabric, ESF significantly improves data center efficiency with fast deployment, simplified operation and management, and easy integration with the cloud.

To achieve these goals, ESF needs switches optimized to storage traffic patterns. ESF switches must have superior capabilities in micro-burst absorption, low-latency input/output (I/O), and fair, predictable performance, using any combination of switch ports, port speeds, and packet sizes.

Not every data center switch can meet these requirements. In addition, it's common for storage racks to contain 16 or fewer servers with connectivity at 10/25GbE or higher. Most of the data center switches in the market today are 1U and have 48x 10/25GbE ports. In such cases, many of the switch ports go unused due to mismatches in the rack-server density and switch-port density.

In general, an ESF switch must provide the following capabilities:

- Fast, guaranteed performance
- Highly utilized and flexible storage connectivity
- Simplified deployment and provisioning
- Automated management
- Easy scalability and cloud integration

## NVIDIA Spectrum for ESF

The NVIDIA Spectrum™ platform is an end-to-end ESF solution built for a broad range of applications, including NVMe-oF, hyperconverged infrastructure (HCI), AI, and hybrid cloud. NVIDIA Spectrum consists of **Spectrum switch systems**, NVIDIA® ConnectX® smart network interface cards (SmartNICs), NVIDIA BlueField® data processing units (DPUs), and LinkX® cables and transceiver modules. Inside storage servers, ConnectX SmartNICs and BlueField DPUs provide hardware accelerators for RoCE, **NVIDIA Magnum IO™ GPUDirect®** Storage, NVIDIA Software-Defined Network Accelerated Processing (SNAP), and security. In the fabric, Spectrum switch systems deliver low latency and high bandwidth and are optimized for RoCE and GPUDirect Storage.

## Spectrum Platform Components for ESF

Powered by the Spectrum switch ASIC and the **NVIDIA Cumulus® Linux networking operating system**, Spectrum switches deliver accelerated ESF with cloud-scale efficiency:

- Purpose-built ESF switches: The complete portfolio of NVIDIA ESF switches ranges in speeds from 1Gb/s to 800Gb/s in switch form factors that provide flexibility for various storage systems and workloads. NVIDIA Spectrum switches, featuring a fully shared buffer architecture, provide a fair, predictable, low-latency, and high- throughput data path (**Tolly Report**), and make ESF transparent to scale-out, software-defined storage. Advanced RoCE capabilities, including adaptive routing supported in the Spectrum SN3000/4000/5000 switches, provide data path enhancements to intelligently minimize congestion of storage data flows across a wide variety of workload.

| Full width (19”) Switches<br>1 Switch fits in 1U | Half-width (8.5”) Switches<br>2 Switches fit in 1U |
|---|---|
|  |  |
| **SN3420**<br>Ideal 25GbE/100GbE aggregation for storage<br>48x25/10GbE + 12x100/40GbE | **SN2010**<br>Optimized 25GbE/100GbE TOR for storage<br>18x10/25GbE + 4x40/100GbE |
|  |  |
| **SN3700**<br>Ideal 25GbE/100GbE aggregation for storage<br>48x25/10GbE + 12x100/40GbE | **SN2100**<br>Ideal 25GbE/100GbE aggregation for storage<br>48x25/10GbE + 12x100/40GbE |
|  | |
| **SN4410**<br>Ideal 100/400GbE aggregation for storage<br>48x100/40GbE + 8x400GbE | |

Figure 2. 25/100/200/400G Spectrum ESF switches

The unique half-width, 1U top-of-rack (ToR) switches are specifically designed for ESF, eliminating underutilization of switch ports and enabling ToR high availability (HA) in 1U rack space. Combined with low-power consumption, the half-width Spectrum switches enable compact design of storage racks for use cases such as edge or remote office/branch office (ROBO) use cases.

- Simplified deployment and automated management: The NVIDIA Cumulus Linux operating system is built for operational efficiency. For example, most RoCE configurations include complex steps and manual configuration that can be error prone. Cumulus Linux with the NVIDIA User Experience (NVUE) command-line interface (CLI) enables RoCE configuration in a single command with best practice settings for optimal performance (refer to the **Cumulus Linux Configuration Guide for ESF**). Cumulus Linux offers a DevOps approach to data center

operations, where a complete data center network can be simulated in a digital twin hosted on the **NVIDIA Air platform**.
ESF can be implemented in the digital twin first, where it's verified with required business policies and compliance. When the physical equipment is installed, connected, and powered on, the ESF can be up and running in minutes.
**NVIDIA NetQ™** automates monitoring and validation of the ESF, allowing the continuous integration and continuous delivery (CI/CD) operational model. This approach greatly simplifies storage network configuration, reduces manual errors and downtime, and streamlines storage expansion and integration with other parts of the data center and cloud.

- Actionable telemetry for storage data flows: NVIDIA NetQ provides unparalleled visibility into ESF with hardware-accelerated actionable telemetry—NVIDIA What Just Happened® (WJH). WJH is a hardware-accelerated monitoring solution where the switch ASIC monitors every storage data flow at line rate and will alert on performance problems due to packet drops, congestion events, routing loops, or other data plane problems.
  With detailed and contextual telemetry, Spectrum improves time to resolution with cloud-scale insight and validation.
- End-to-end optimization for NVIDIA AI platforms: AI is data centric. Throughout the AI pipeline, data ingestion, curation and transformation, and model training all require high-performance data transfers. NVIDIA ESF is the only end-to-end solution maximizing the performance for AI-driven storage workflows.

## ESF for AI Storage

The large datasets needed for deep learning and generative AI make training workloads very storage- and I/O-intensive. Traditional Ethernet is unable to provide the high-effective bandwidth and low latency needed for highly parallel data flows in all-flash arrays and scale-out parallel file systems. To that end, NVIDIA's Spectrum-X is designed to accelerate AI workloads and improve the performance and efficiency of AI clouds built with Ethernet. This breakthrough technology offers 1.7X better overall AI performance and 1.7X higher power efficiency, along with consistent, predictable performance in multi-tenant environments. Spectrum-X is built on network innovations powered by the tight coupling of the Spectrum-4 switch and BlueField-3 DPU, reducing AI run times and allowing researchers and scientists to obtain accurate results and make informed decisions faster. For ESF solutions powering AI clouds built on Spectrum-4 switches and BlueField-3 DPUs, Spectrum-X is recommended to provide optimal performance.

## Key Benefits of Spectrum-X

1. **Improved AI cloud performance**: Spectrum-X enhances AI cloud performance by 1.7X and higher, accelerating processing, analysis, and execution of AI workloads and, in turn, the development and deployment of AI solutions.
2. **Standard Ethernet connectivity**: Spectrum-X is powered by NVIDIA innovation, is fully standards-based Ethernet, and is completely interoperable with Ethernet-based stacks.
3. **Increased power efficiency**: By improving performance, Spectrum-X contributes to a more power-efficient AI environment. This leads to reduced power consumption and lower operational costs for AI clusters.
4. **Enhanced multi-tenant performance**: Performance isolation in multi-tenant environments ensures that each tenant's workloads perform optimally and consistently, resulting in higher customer satisfaction and improved service quality.

5. **Better AI fabric visibility**: Visibility into the flows running across the AI cloud makes it possible to identify performance bottlenecks and is a key part of a modern, automated fabric-validation solution.
6. **Higher AI scalability**: Unprecedented scale to 256x 200G ports in a single hop or 16K ports in a two-tier leaf/spine topology supports the expansion of AI clouds while maintaining high levels of performance, making it an ideal solution for organizations with evolving AI infrastructure needs.
7. **Faster network setup**: The automated, end-to-end configuration of advanced networking functionality is fully tuned for AI workloads.
8. **Higher resiliency**: With higher resilience, the cascading performance issues that occur with a lost link are eliminated, limiting the loss in bandwidth to that single link.

| 400/200G Switches | Spectrum-X Switches |
|---|---|
|  |  |
| **SN4700**<br>Ideal 400GbE TOR/aggregation for storage<br>32x400GbE | **SN5400**<br>High-density 400G aggregation for AI storage<br>64x400GbE |
|  |  |
| **SN4600**<br>High-density 200GbE TOR/aggregation for storage \| 64x200GbE | **SN5600**<br>High-density 800GbE aggregation for storage<br>64x800GbE |

Figure 3. Spectrum ESF switches for AI. Spectrum-X is only compatible with SN5400 and SN5600.

## The End-to-End Solution for All Storage Fabric Needs

Ethernet storage fabric (ESF) has become the de facto network fabric for storage in data centers and cloud. The NVIDIA Spectrum platform is an end-to-end ESF solution, purpose-built to achieve accelerated Ethernet at an optimal total cost of ownership. With hardware acceleration for storage data transfers and the software suite for automated configuration, validation, monitoring, and security, NVIDIA Spectrum-based ESF delivers the performance and efficiency needed for all storage use cases, including RoCE-accelerated NVMe-oF and GPUDirect-powered AI and machine learning applications.

## Ready to Get Started?

To learn more about NVIDIA Ethernet, visit **https://www.nvidia.com/en-us/networking/ethernet-switching/**



ASBIS® SUCCESS THROUGH FOCUS

NVIDIA.