



DDN A³I[®] Solutions with Supermicro NVIDIA HGX Supercluster

Fully-integrated and optimized infrastructure solutions for accelerated at-scale AI, Analytics and HPC featuring Supermicro NVIDIA HGX Systems.

EXECUTIVE SUMMARY

DDN A³I Solutions are proven at-scale to deliver optimal data performance for Artificial Intelligence (AI), Data Analytics and High-Performance Computing (HPC) applications running on GPUs in Supermicro HGX H100/H200 systems. DDN data platforms power nearly all large HGX clusters and NVIDIA DGX SuperPOD™ systems currently in operation, including NVIDIA SELENE, Cambridge-1 and many others.

This document describes fully validated reference architectures for scalable HGX Cluster configurations. The solutions integrate DDN AI400X2 appliances and DDN Insight software with HGX Clusters.

1. DDN A³I END-TO-END ENABLEMENT FOR SUPERMICRO HGX CLUSTERS

DDN A³I solutions (Accelerated, Any-Scale AI) are architected to achieve the most from at-scale AI, Data Analytics and HPC applications running on HGX systems and HGX clusters. They provide predictable performance, capacity, and capability through a tight integration between DDN and Supermicro systems. Every layer of hardware and software engaged in delivering and storing data is optimized for fast, responsive, and reliable access.

DDN A³I solutions are designed, developed, and optimized in close collaboration with Supermicro. The deep integration of DDN AI appliances with HGX systems ensures a reliable experience. DDN A³I solutions are configured for flexible deployment in a wide range of environments and scale seamlessly in capacity and capability to match evolving workload needs.

DDN brings the same advanced technologies used to power the world's largest supercomputers in a fully integrated package for HGX systems that's easy to deploy and manage. DDN A³I solutions are proven to maximize benefits for at-scale AI, Analytics and HPC workloads on HGX systems.

This section describes the advanced features of DDN A³I Solutions for HGX clusters.

1.1 DDN A³I SHARED PARALLEL ARCHITECTURE

The DDN A³I shared parallel architecture and client protocol ensures high levels of performance, scalability, security, and reliability for HGX systems. Multiple parallel data paths extend from the drives all the way to containerized applications running on the GPUs in the HGX system. With DDN's true end-to-end parallelism, data is delivered with high-throughput, low-latency, and massive concurrency in transactions. This ensures applications achieve the most from HGX systems with all GPU cycles put to productive use. Optimized parallel data-delivery directly translates to increased application performance and faster completion times. The DDN A³I shared parallel architecture also contains redundancy and automatic failover capability to ensure high reliability, resiliency, and data availability in case a network connection or server becomes unavailable.

1.2 DDN A³I STREAMLINED DEEP LEARNING WORKFLOWS

DDN A³I solutions enable and accelerate end-to-end data pipelines for deep learning (DL) workflows of all scale running on HGX systems. The DDN shared parallel architecture enables concurrent and continuous execution of all phases of DL workflows across multiple HGX systems. This eliminates the management overhead and risks of moving data between storage locations. At the application level, data is accessed through a standard highly interoperable file interface, for a familiar and intuitive user experience.

Significant acceleration can be achieved by executing an application across multiple HGX systems in an HGX cluster simultaneously and engaging parallel training efforts of candidate neural networks variants. These advanced optimizations maximize the potential of DL frameworks. DDN works closely with Supermicro and its customers to develop solutions and technologies that allow widely-used DL frameworks to run reliably on HGX systems.

1.3 DDN A³I MULTIRAIL NETWORKING

DDN A³I solutions integrate a wide range of networking technologies and topologies to ensure streamlined deployment and optimal performance for AI infrastructure. The latest generation NVIDIA Quantum InfiniBand (IB) and Spectrum Ethernet technology provide both high-bandwidth and low-latency data transfers between applications, compute servers and storage appliances.

DDN A³I Multirail enables grouping of multiple network interfaces on a HGX system to achieve faster aggregate data transfer capabilities. The feature balances traffic dynamically across all the interfaces, and actively monitors link health for rapid failure detection and automatic recovery. DDN A³I Multirail makes designing, deploying, and managing high-performance networks very simple, and is proven to deliver complete connectivity for at-scale infrastructure for HGX cluster deployments.

1.4 DDN A³I ADVANCED OPTIMIZATIONS FOR HGX SYSTEM ARCHITECTURE

The DDN A³I client's NUMA-aware capabilities enable strong optimization for HGX systems. It automatically pins threads to ensure I/O activity across the HGX system is optimally localized, reducing latencies and increasing the utilization efficiency of the whole environment. Further enhancements reduce overhead when reclaiming memory pages from page cache to accelerate buffered operations to storage. The DDN A³I client software for HGX systems has been validated at-scale with the largest HGX clusters and DGX SuperPODs currently in operation.

1.5 DDN A³I HOT NODES

DDN Hot Nodes is a powerful software enhancement that enables the use of the NVME devices in a HGX system as a local cache for read-only operations. This method significantly improves the performance of applications if a data set is accessed multiple times during a particular workflow.

This is typical with DL training, where the same input data set or portions of the same input data set are accessed repeatedly over multiple training iterations. Traditionally, the application on the HGX system reads the input data set from shared storage directly, thereby continuously consuming shared storage resources. With Hot Nodes, as the input data is read during the first training iteration, the DDN software automatically writes a copy of the data on the local NVME devices. During subsequent reads, data is delivered to the application from the local cache rather than the shared storage. This entire process is managed by the DDN client software running on the HGX system. Data access is seamless and the cache is fully transparent to users and applications. The use of the local cache eliminates network traffic and reduces the load on the shared storage system. This allows other critical DL training operations like checkpointing to complete faster by engaging the full capabilities of the shared storage system.

DDN Hot Nodes includes extensive data management tools and performance monitoring facilities. These tools enable user-driven local cache management, and make integration simple with task schedulers. For example, training input data can be loaded to the local cache on a HGX system as a pre-flight task before the AI training application is engaged. As well, the metrics expose insightful information about cache utilization and performance, enabling system administrators to further optimize their data loading and maximize application and infrastructure efficiency gains.

1.6 DDN A³I CONTAINER CLIENT

Containers encapsulate applications and their dependencies to provide simple, reliable, and consistent execution. DDN enables a direct high-performance connection between the application containers on the HGX system and the DDN parallel filesystem. This brings significant application performance benefits by enabling low latency, high-throughput parallel data access directly from a container. Additionally, the limitations of sharing a single host-level connection to storage between multiple containers disappear. The DDN in-container filesystem mounting capability is added at runtime through a universal wrapper that does not require any modification to the application or container.

Containerized versions of popular DL frameworks specially optimized for HGX systems are available. They provide a solid foundation that enables data scientists to rapidly develop and deploy applications on HGX systems. In some cases, open-source versions of the containers are available, further enabling access and integration for developers. The DDN A³I container client provides high-performance parallelized data access directly from containerized applications on HGX system. This provides containerized DL frameworks with the most efficient dataset access possible, eliminating all latencies introduced by other layers of the computing stack.

1.7 DDN A³I S3 DATA SERVICES

DDN S3 Data Services provide hybrid file and object data access to the shared namespace. The multi-protocol access to the unified namespace provides tremendous workflow flexibility and simple end-to-end integration. Data can be captured directly to storage through the S3 interface and accessed immediately by containerized applications on a HGX system through a file interface. The shared namespace can also be presented through an S3 interface, for easy collaboration with multisite and multicloud deployments. The DDN S3 Data Services architecture delivers robust performance, scalability, security, and reliability features.

2. DDN A³I SOLUTIONS WITH SUPERMICRO HGX H100 AND H200 SYSTEMS

The DDN A³I scalable architecture integrates HGX systems with DDN AI shared parallel file storage appliances and delivers fully-optimized end-to-end AI, Analytics and HPC workflow acceleration on GPUs. DDN A³I solutions greatly simplify the deployment of HGX systems in HGX cluster configurations, while also delivering performance and efficiency for maximum GPU saturation, and high levels of scalability.

This section describes the components integrated in DDN A³I Solutions for HGX clusters.

2.1 DDN AI400X2 APPLIANCE

The AI400X2 appliance is a fully integrated and optimized shared data platform with predictable capacity, capability, and performance. Every AI400X2 appliance delivers over 90 GB/s and 3M IOPS directly to HGX systems in an HGX cluster. Shared performance scales linearly as additional AI400X2 appliances are integrated to an HGX cluster. The all-NVMe configuration provides optimal performance for a wide variety of workload and data types and ensures that HGX cluster operators achieve the most from at-scale GPU applications, while maintaining a single, shared, centralized data platform.

The AI400X2 appliance integrates the DDN A³I shared parallel architecture and includes a wide range of capabilities described in section 1, including automated data management, digital security, and data protection, as well as extensive monitoring. The AI400X2 appliances enables HGX cluster operators to go beyond basic infrastructure and implement complete data governance pipelines at-scale.

The AI400X2 appliance integrates with HGX clusters over InfiniBand, Ethernet and RoCE. It is available in 250 and 500 TB all-NVMe capacity configurations. Optional hybrid configurations with integrated HDDs are also available for deployments requiring high-density deep capacity storage. Contact DDN Sales for more information.



Figure 1. DDN AI400X2 all-NVME storage appliance.

2.2 DDN INSIGHT SOFTWARE

DDN Insight is a centralized management and monitoring software suite for AI400X2 appliances. It provides extensive performance and health monitoring of all DDN storage systems connected to HGX cluster from a single web-based user interface. DDN Insight greatly simplifies IT operations and enables automated and proactive storage platform management guided by analytics and intelligent software.

Performance monitoring is an important aspect of operating an HGX cluster efficiently. Provided the several variables that affect data I/O performance, the identification of bottlenecks and degradation is crucial while production workloads are engaged. DDN Insight provides deep real-time analysis across the entire HGX cluster, tracking I/O transactions from applications running on HGX nodes all the way through individual drives in the AI400X2 appliances. The embedded analytics engine makes it simple for HGX cluster operators to visualize I/O performance across their entire infrastructure through intuitive user interfaces. These include extensive logging, trending, and comparison tools, for analyzing I/O performance of specific applications and users over time. As well, the open backend database makes it simple to extend the benefits of DDN Insight and integrate other AI infrastructure components within the engine, or export data to third party monitoring systems.

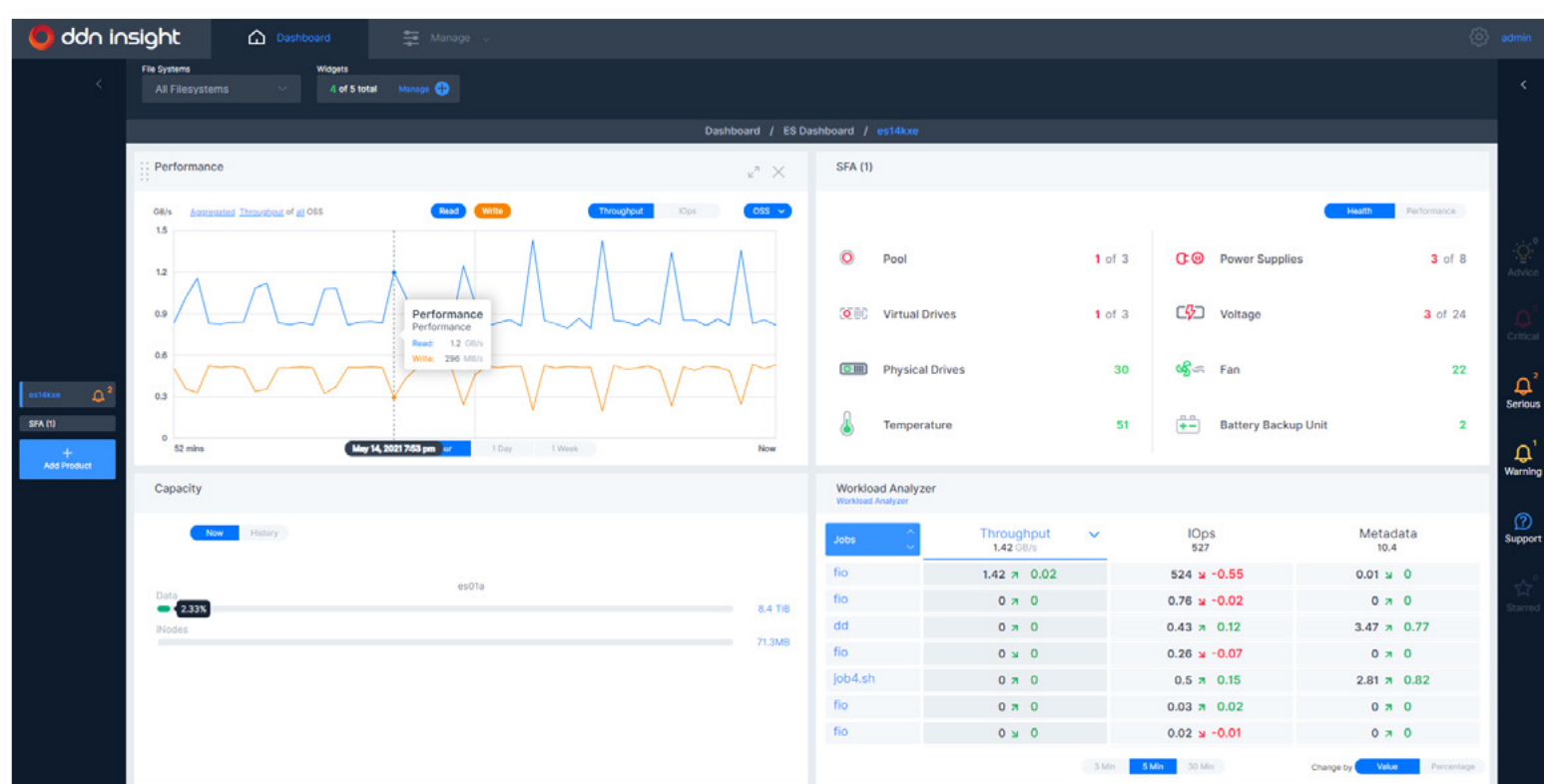


Figure 2. DDN Insight Workload Analyzer Tool.

2.3 SUPERMICRO AS -8125GS-TNHR SYSTEM

8U Supermicro GPU server with 2x Genoa 9654 CPUs (96C @ 2.4GHz, 384MB L3 Cache) 1x Dedicate IPMI, 1,536GB RAM (24x64), 8x 15.3TB NVME, 2x 960GB NVME, 10x 400Gb/s NDR IB NVIDIA ConnectX-7, Single-port OSFP NIC, and 2x 10GbE RJ45

<https://www.supermicro.com/en/products/system/gpu/8u/as-8125gs-tnhr>

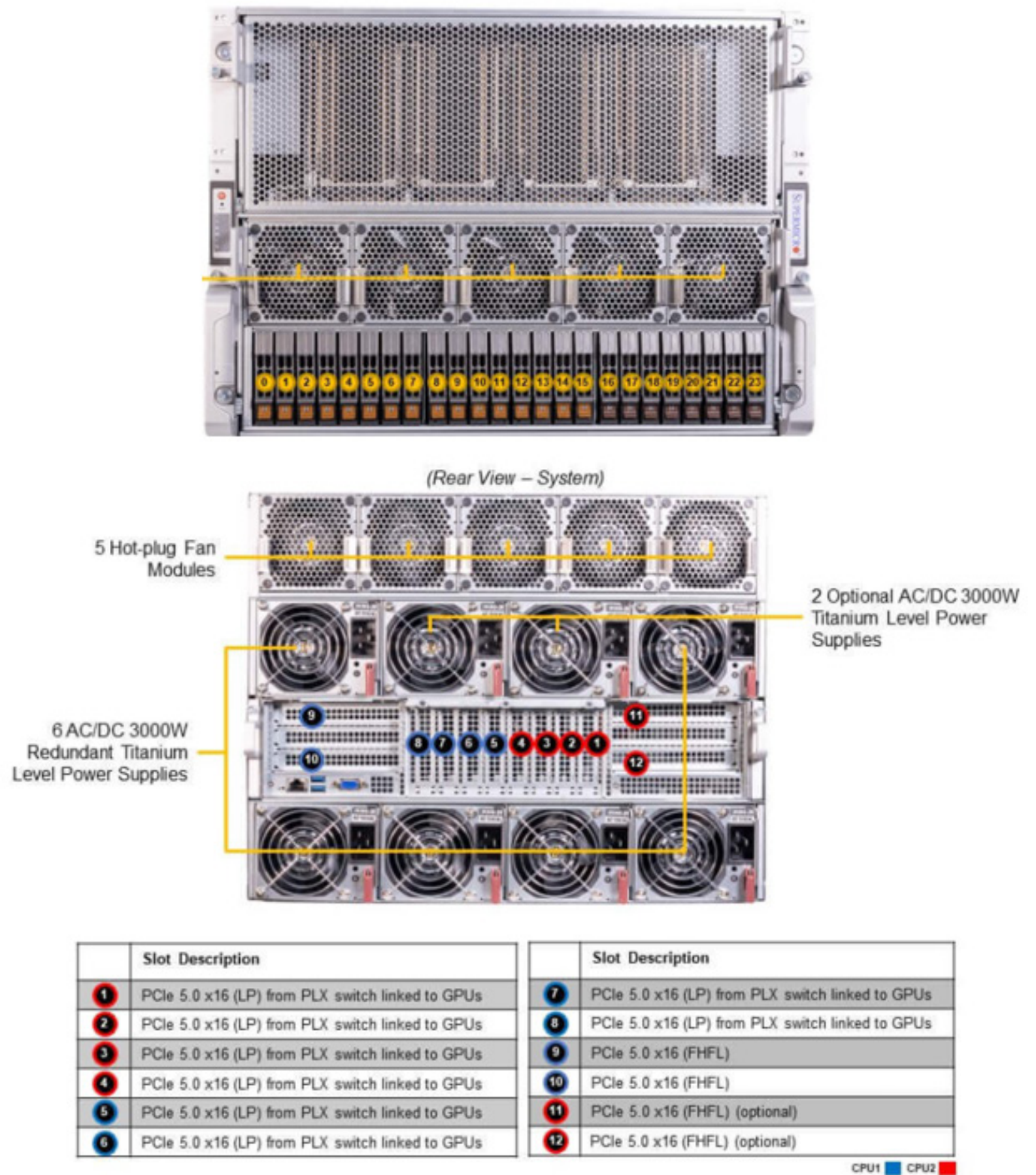


Figure 3. Supermicro AS -8125GS-TNHR HGX H100 system.

KEY APPLICATIONS

- A. High Performance Computing
- B. AI/Deep Learning Training
- C. Industrial Automation, Retail
- D. Climate and Weather Modeling

KEY FEATURES

01. High density 8U system with NVIDIA® HGX™ H100 8-GPU
02. Highest GPU communication using NVIDIA® NVLINK™ + NVIDIA® NVSwitch™
03. NIC for GPU direct RDMA (1:1 GPU Ratio)
04. 24 DIMM Slots; Up to 6TB DRAM; 4800 ECC DDR5 LRDIMM; RDIMM;
05. PCIe Gen 5.0 X16 LP, and up to 4 PCIe Gen 5.0 X16 FHFL Slots
06. Flexible networking options
07. 2x 2.5" Hot-swap SATA drive bays
08. M.2 NVMe for boot drive only
09. Up to 16x 2.5" Hot-swap NVMe drive bays (12 by default + 4 optional)
10. Heavy duty fans with optimal fan speed control
11. 6x 3000W redundant Titanium level power supplies

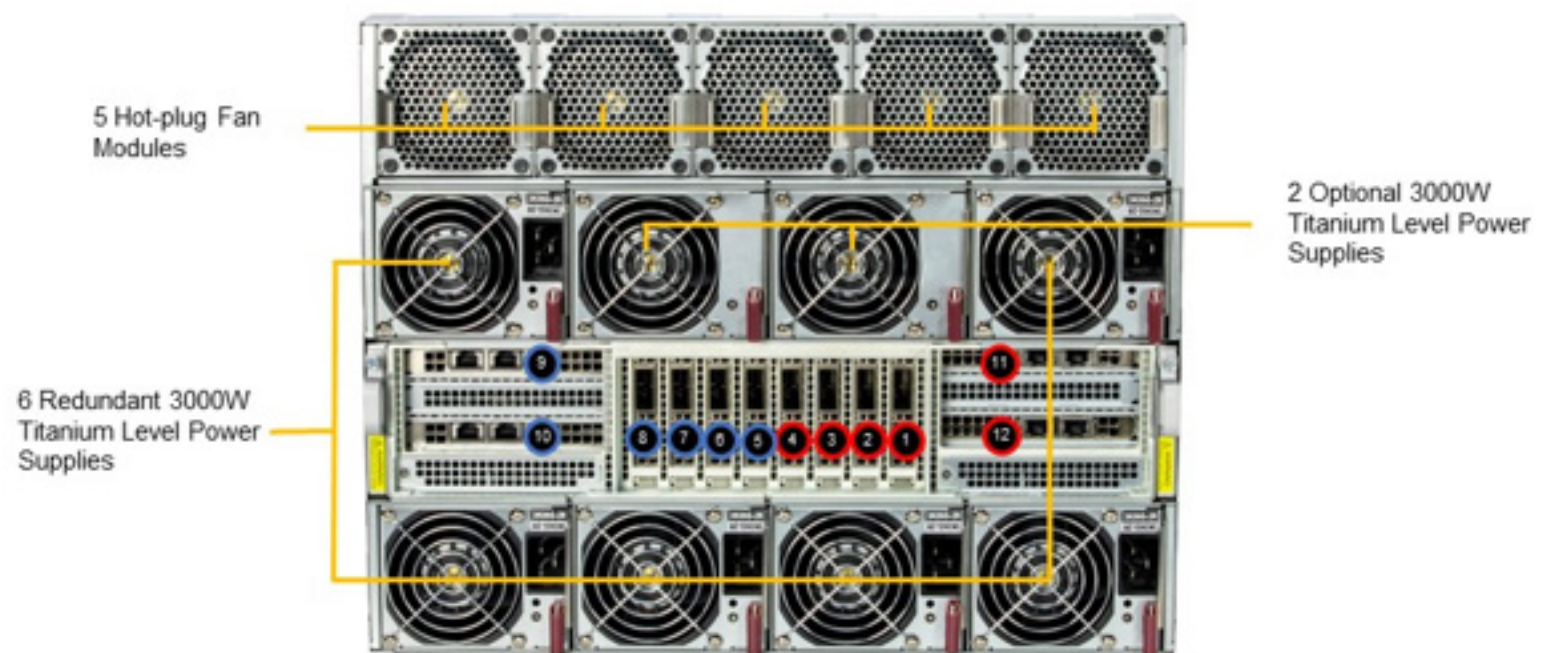
2.4 SUPERMICRO SYS-821GE-TNHR SYSTEM

8U Supermicro GPU server with 2x EMR8592V CPUs (64C @ 2 GHz, 320MB Cache) 1x Dedicate IPMI, 2,048GB RAM (32x64), 8x 15.3TB NVME, 2x 960GB NVME, 10x 400Gb/s NDR IB NVIDIA ConnectX-7, Single-port OSFP NIC, and 2x 10GbE RJ45.

<https://www.supermicro.com/en/products/system/gpu/8u/sys-821ge-tnhr>



(Rear View – System)



Slot	Slot Description
1	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
2	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
3	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
4	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
5	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
6	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs

Slot	Slot Description
7	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
8	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
9	PCIe 5.0 x16 (FHHL) via PLX switch (optional)
10	PCIe 5.0 x16 (FHHL) via PLX switch (optional)
11	PCIe 5.0 x16 (FHHL) via PLX switch
12	PCIe 5.0 x16 (FHHL) via PLX switch

CPU1 CPU2

Figure 4. Supermicro SYS-821GE-TNHR HGX H100 system.

KEY APPLICATIONS

- A. High Performance Computing
- B. Industrial Automation, Retail
- C. Conversational AI
- D. Drug Discovery
- E. Finance & Economics
- F. AI/Deep Learning Training
- G. Healthcare
- H. Business Intelligence & Analytics
- I. Climate and Weather Modeling

KEY FEATURES

- 01. 5th/4th Gen Intel® Xeon® Scalable processor support
- 02. 32 DIMM slots Up to 8TB: 32x 256 GB DRAM
Memory Type: 5600MTs ECC DDR5
- 03. 8 PCIe Gen 5.0 X16 LP
- 04. 2 PCIe Gen 5.0 X16 FHHL Slots, 2 PCIe
Gen 5.0 X16 FHHL Slots (optional)
- 05. Flexible networking options
- 06. 2 M.2 NVMe for boot drive only
- 07. 16x 2.5" Hot-swap NVMe drive bays
(12x by default, 4x optional)
- 08. 3x 2.5" Hot-swap SATA drive bays
- 09. Optional: 8x 2.5" Hot-swap SATA drive bays
- 10. 10 heavy duty fans with optimal fan speed control
- 11. Optional: 8x 3000W (4+4) Redundant
Power Supplies, Titanium Level
- 12. 6x 3000W (4+2) Redundant Power Supplies, Titanium Level

3. DDN A³I RECOMMENDATIONS FOR HGX CLUSTERS

DDN proposes the following recommended architectures for HGX Cluster configurations. DDN A³I solutions are fully validated with Supermicro and already deployed with several HGX cluster customers worldwide.

The DDN AI400X2 appliance is a turnkey appliance fully-validated and proven with HGX Cluster. The AI400X2 appliances delivers optimal GPU performance for every workload and data type in a dense, power efficient 2RU chassis. The AI400X2 appliance simplifies the design, deployment, and management of a HGX cluster and provides predictable performance, capacity, and scaling. The appliance is designed for seamless integration with HGX systems and enables customers to move rapidly from test to production. As well, DDN provides complete expert design, deployment, and support services globally. The DDN field engineering organization has already deployed dozens of solutions for customers based on the A³I reference architectures.

3.1 STORAGE SIZING FOR HGX CLUSTERS

As general guidance, DDN recommends the shared storage be sized to ensure at least 1 GB/s per second of read and write throughput for every H100/H200 GPU in a HGX cluster (Table 1). This ensures minimum performance required to operate the GPU infrastructure. All four configurations are illustrated in this guide.

		Scalable Units (SUs)						
		¼	½	1	2	3	4	8
Compute components	HGX Systems	8	16	32	64	96	128	256
	H100/H200 GPUs	64	128	256	512	768	1024	2048
DDN Storage components	AI400X2 combo appliances	1	2	4	8	12	-	-
	AI400X2 metadata appliances	-	-	-	-	-	3	6
	AI400X2 data appliances	-	-	-	-	-	16	32
DDN Storage specification	Aggregate read throughput	90 GB/s	180 GB/s	360 GB/s	720 GB/s	1 TB/s	1.4 TB/s	2.8 TB/s
	Aggregate write throughput	65 GB/s	130 GB/s	260 GB/s	520 GB/s	780 GB/s	1 TB/s	2 TB/s
	Per GPU read throughput	1.4 GB/s	1.4 GB/s	1.4 GB/s	1.4 GB/s	1.4 GB/s	1.4 GB/s	1.4 GB/s
	Per GPU write throughput	1 GB/s	1 GB/s	1 GB/s	1 GB/s	1 GB/s	1 GB/s	1 GB/s
	Useable capacity 500TB appliance option	500 TB	1 PB	2 PB	4 PB	6 PB	8 PB	16 PB
	Useable capacity 250TB appliance option	250 TB	500 TB	1 PB	2 PB	3 PB	4 PB	8 PB
	NDR/200 GbE ports	8	16	32	64	96	152	304
	1 GbE ports	4	8	16	32	48	76	152
	Physical, rack units	2	4	8	16	24	38	76
	Power, nominal	2.2 KW	4.3 KW	8.6 KW	17.3 KW	25.9 KW	40.2 KW	80.5 KW
	Cooling, nominal	7.4 kBTU/hr	14.7 kBTU/hr	29.5 kBTU/hr	58.9 kBTU/hr	88.4 kBTU/hr	137.2 kBTU/hr	274.5 kBTU/hr

Table 1. DDN recommended storage sizing for HGX clusters.

3.2 STORAGE NETWORK FOR HGX CLUSTERS

The HGX cluster reference design includes several networks, one of which is dedicated to storage traffic. The storage network provides connectivity between the AI400X2 appliances, the compute nodes and management nodes. This network is designed to meet the high-throughput, low-latency, and scalability requirements of HGX cluster.

The NVIDIA QM9700 InfiniBand Switch is recommended for HGX cluster storage connectivity. It provides 64 ports of NDR over 32 OSFP ports in a 1RU form factor. The cables listed in section 3.1.3 are validated with the QM9700 switch.



Figure 5. NVIDIA QM9700 InfiniBand Switch.

An overview of the HGX cluster storage network architecture is shown in Figure 6, recommended network connections for each HGX system on Figure 7, and recommended network connections for each AI400X2 appliance on Figure 8.

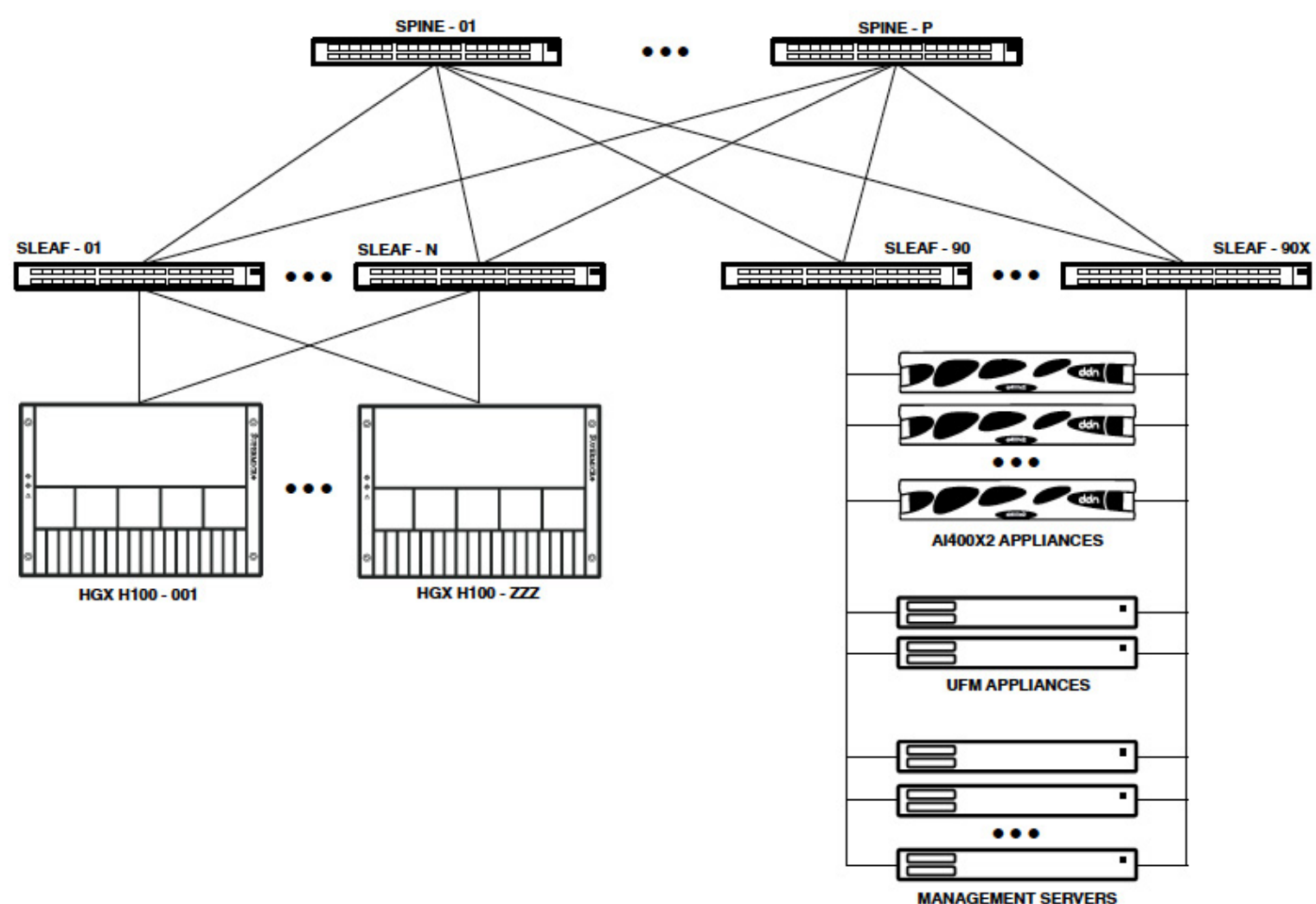


Figure 6. Overview of the HGX cluster storage network architecture.

3.2.1 HGX SYSTEMS NETWORK CONNECTIVITY

For HGX clusters, DDN recommends ports 1 to 4 on the HGX systems be connected to the compute network. Ports 5 and 7 should be connected to the in-band management network, which also serves as the storage network for Ethernet storage configurations. For InfiniBand storage configurations, ports 6 and 8 should be connected to the InfiniBand storage network. As well, the management BMC (“B”) port should be connected to the out-of-band management network.

For HGX clusters, Supermicro recommends 2 dual-ported network cards (Supermicro PN: AOC-SMP-CX7660030-SQ0) for the storage network.

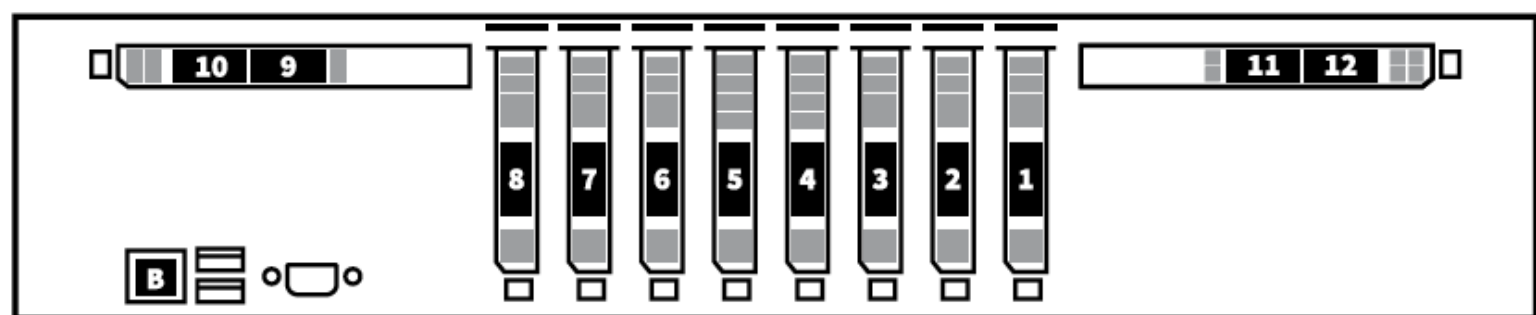


Figure 7. Recommended HGX system network port connections.

3.2.2 AI400X2 APPLIANCE NETWORK CONNECTIVITY

For HGX clusters, DDN recommends ports 1 to 8 on the AI400X2 appliance be connected to the storage network. As well, the management (“M”) and BMC (“B”) ports for both controllers should be connected to the out-of-band management network. Note that each AI400X2 appliance requires one inter-controller network port connection (“I”) using short ethernet cable supplied.

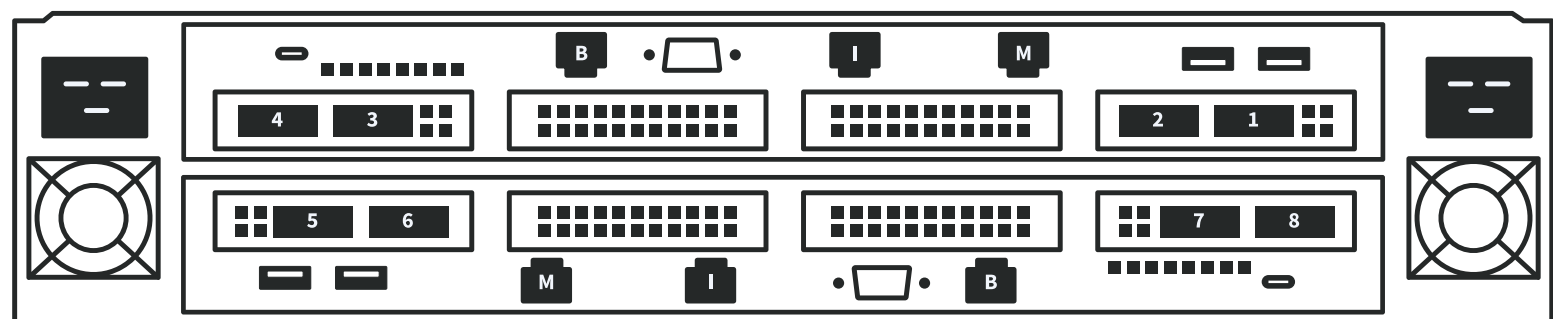


Figure 8. Recommended AI400X2 appliance network port connections.

3.2.3 AI400X2 APPLIANCE CABLING WITH NDR INFINIBAND SWITCHES

The AI400X2 appliance connects to the HGX cluster storage network with 8 200 Gbps InfiniBand interfaces. Particular attention must be given to the cabling selection to ensure compatibility between different InfiniBand connectivity and data rates. DDN has validated the following cables to connect AI400X2 appliances with QM9700 switches. The use of splitter cables ensures most efficient use of switch ports.

Copper Cable

Requires 2 cables per AI400X2 appliance

MCP7Y40-N0xx	NVIDIA DAC 1:4 splitter, InfiniBand 2 x 400Gb/s to 4 x 200Gb/s, OSFP-Finned top to QSP112 xxx indicates length in meters: 001, 01A (1.5m), 002
--------------	---

Optical Cable Set

Requires 2 sets per AI400X2 appliance

Qty	Part Number	Description
1	MMA4Z00-NS	NVIDIA Twin-port OSFP Finned Top Transceiver SKU 980-91510-00NS00
2	MFP7E20-N0xx	NVIDIA Passive Fiber Cable xx indicates length in meters: 03-50
4	MMA1Z00-NS400	NVIDIA Single-port, QSFP112 Multimode SR4 Transceiver

Note: 400Gb/s transceivers used with 1:2 fiber splitters result in 200Gb/s transceivers. 200Gb/s transceivers are not offered.

3.3 HGX CLUSTER WITH 8 HGX H100 SYSTEMS

Figure 9 illustrates the DDN A³I architecture for HGX cluster with 8 HGX H100 systems, 1 DDN AI400X2 appliance and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. The AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliance over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

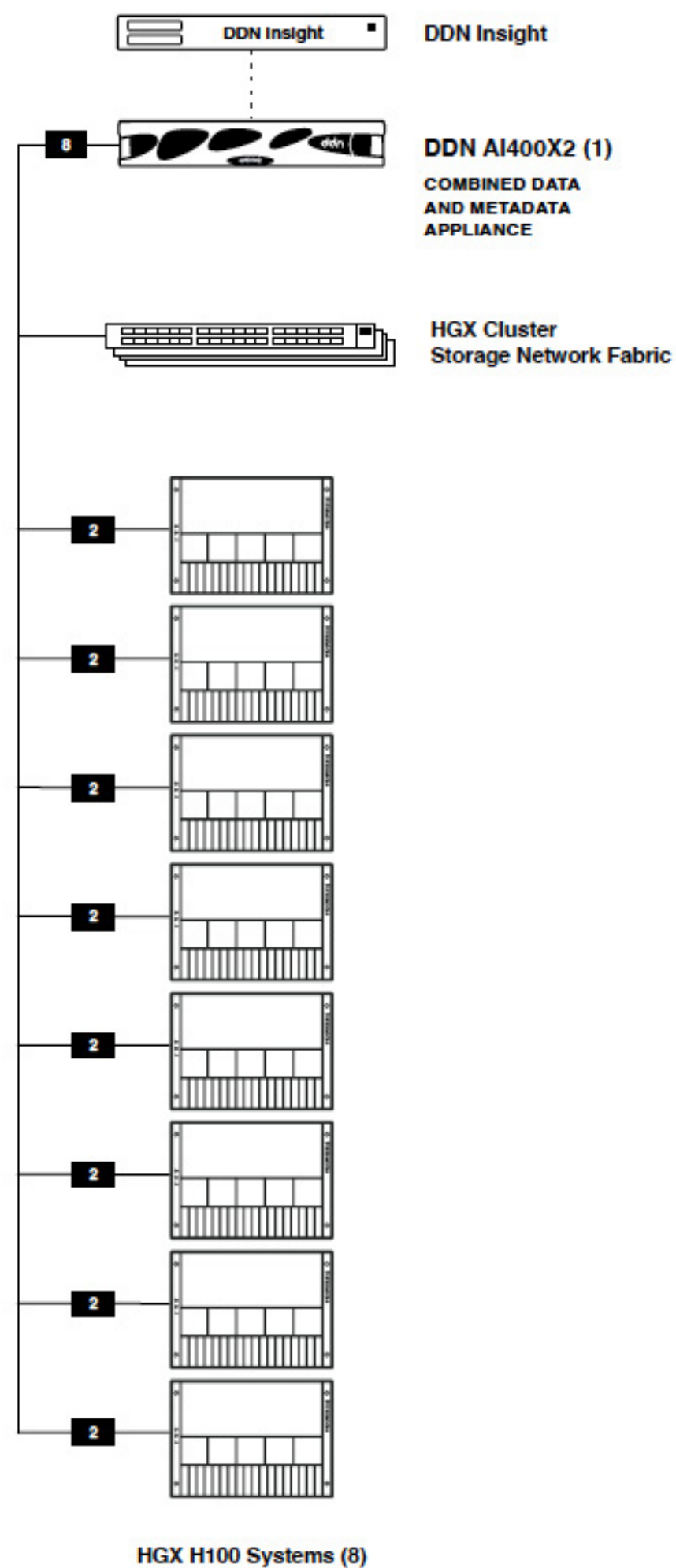


Figure 9. DDN A³I HGX cluster reference architecture with 8 HGX H100 systems.

3.4 HGX CLUSTER WITH 16 HGX H100 SYSTEMS

Figure 10 illustrates the DDN A³I architecture for HGX cluster with 16 HGX H100 systems, 2 DDN AI400X2 appliances and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. Each AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliances over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

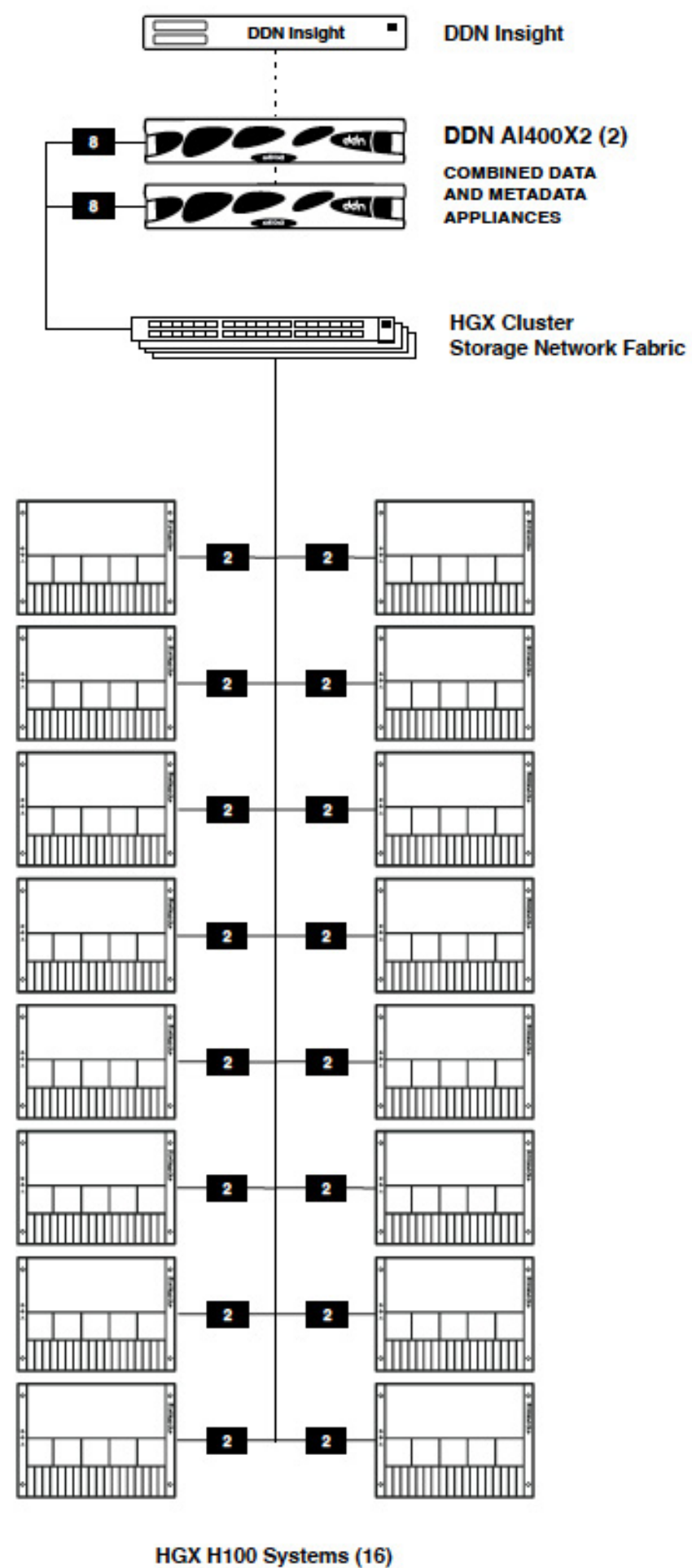


Figure 10. DDN A³I HGX cluster reference architecture with 16 HGX H100 systems.

3.5 HGX CLUSTER WITH 32 HGX H100 SYSTEMS

Figure 11 illustrates the DDN A³I architecture for HGX cluster with 32 HGX H100 systems, 4 DDN AI400X2 appliances and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. Each AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliances over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

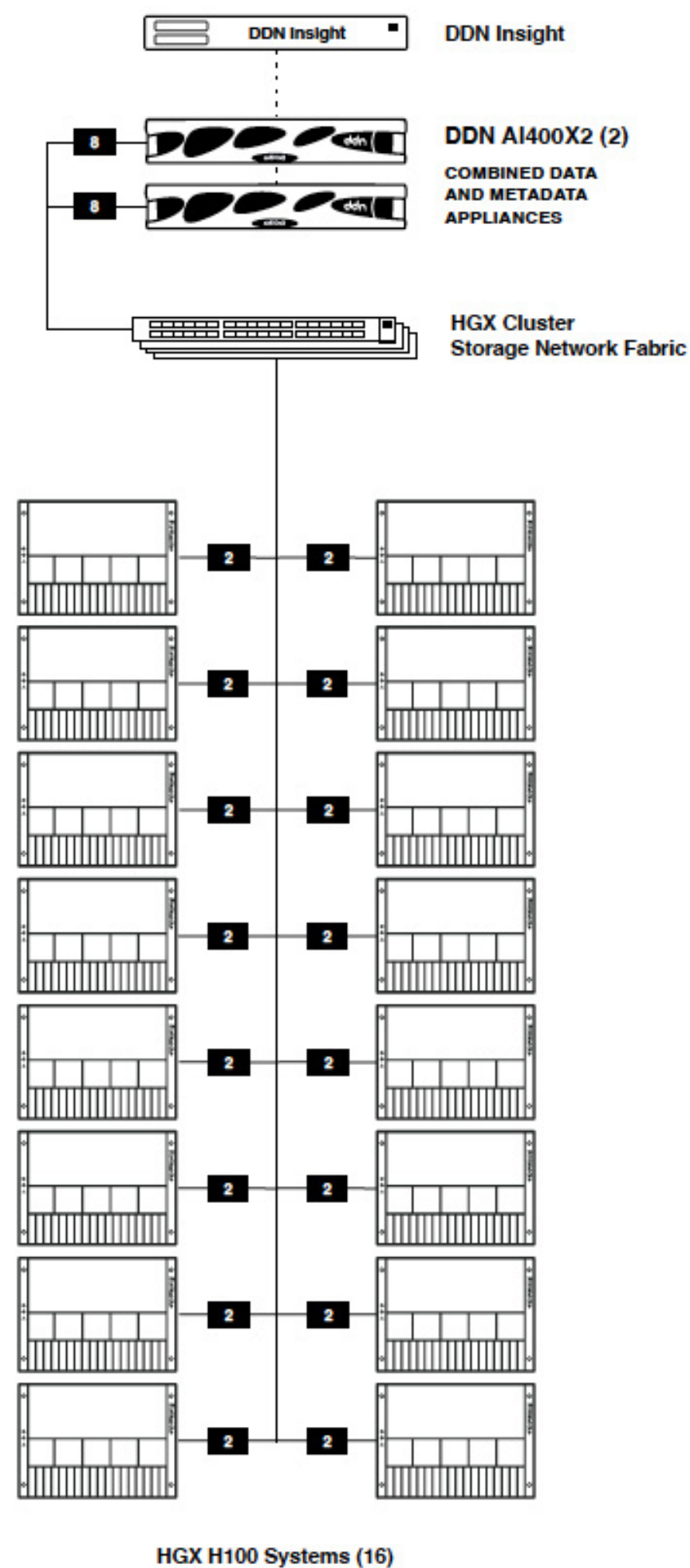


Figure 11. DDN A³I HGX cluster reference architecture with 32 HGX H100 systems.

3.6 HGX CLUSTER WITH 64 HGX H100 SYSTEMS

Figure 12 illustrates the DDN A³I architecture for HGX cluster with 64 HGX H100 systems, 8 DDN AI400X2 appliances and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. Each AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliances over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

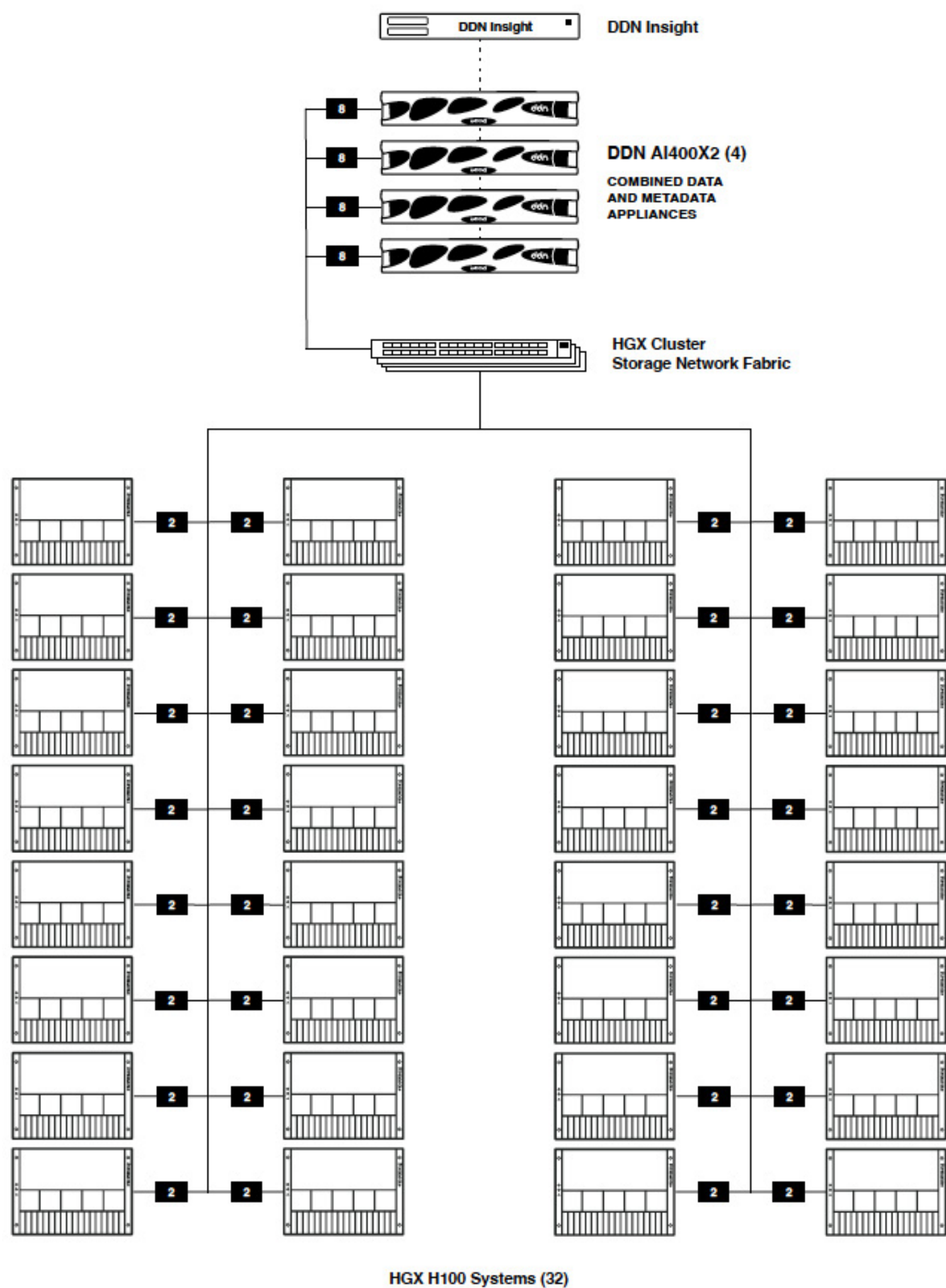


Figure 12. DDN A³I HGX cluster reference architecture with 64 HGX H100 systems.

3.7 HGX CLUSTER WITH 96 HGX H100 SYSTEMS

Figure 13 illustrates the DDN A³I architecture for HGX cluster with 96 HGX H100 systems, 12 DDN AI400X2 appliances and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. Each AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliances over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

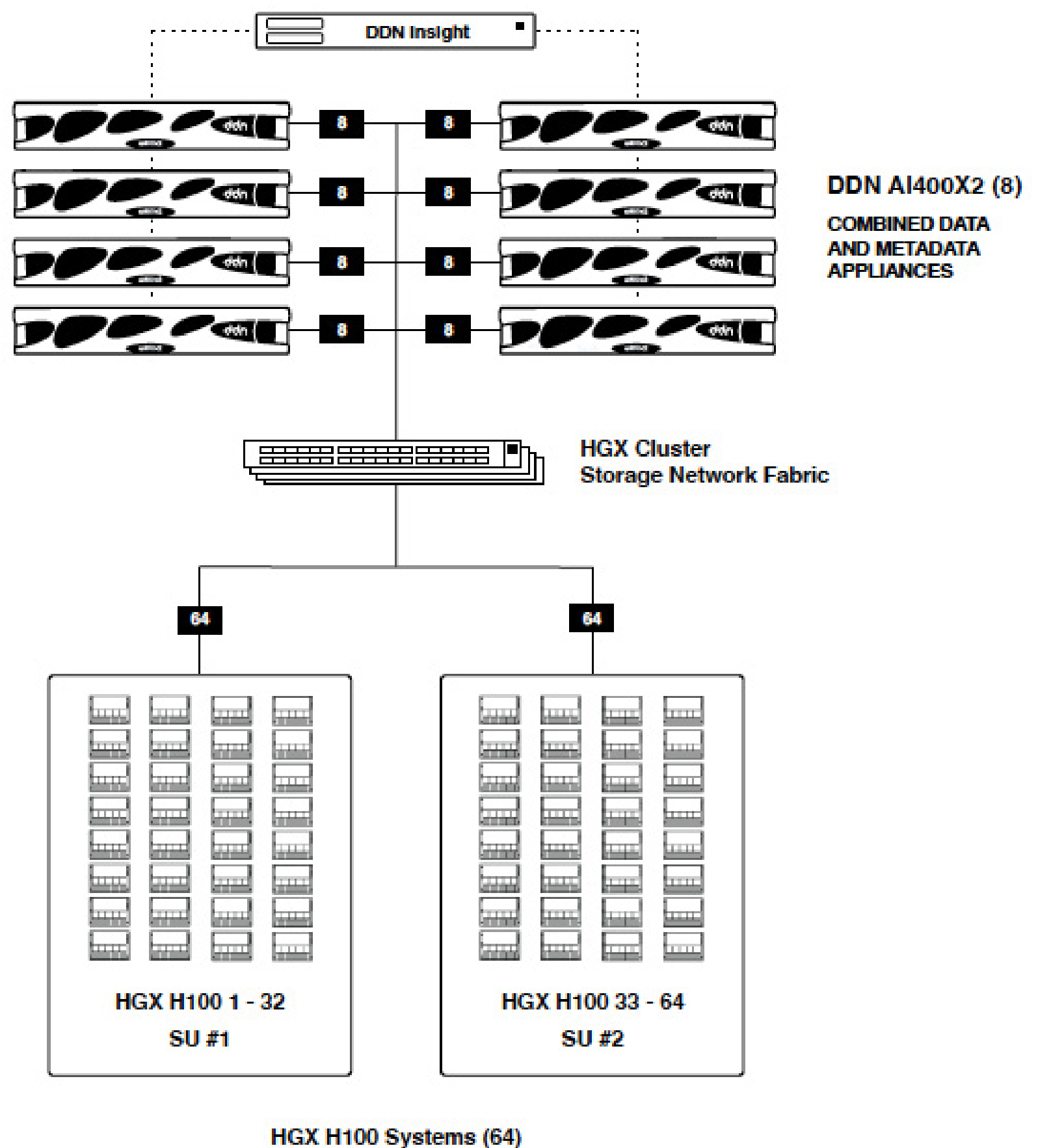


Figure 13. DDN A³I HGX cluster reference architecture with 96 HGX H100 systems.

3.8 HGX CLUSTER WITH 128 HGX H100 SYSTEMS

Figure 14 illustrates the DDN A³I architecture for HGX cluster with 128 HGX H100 systems, 3 DDN AI400X2 metadata appliances, 16 DDN AI400X2 data appliances and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. Each AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliances over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

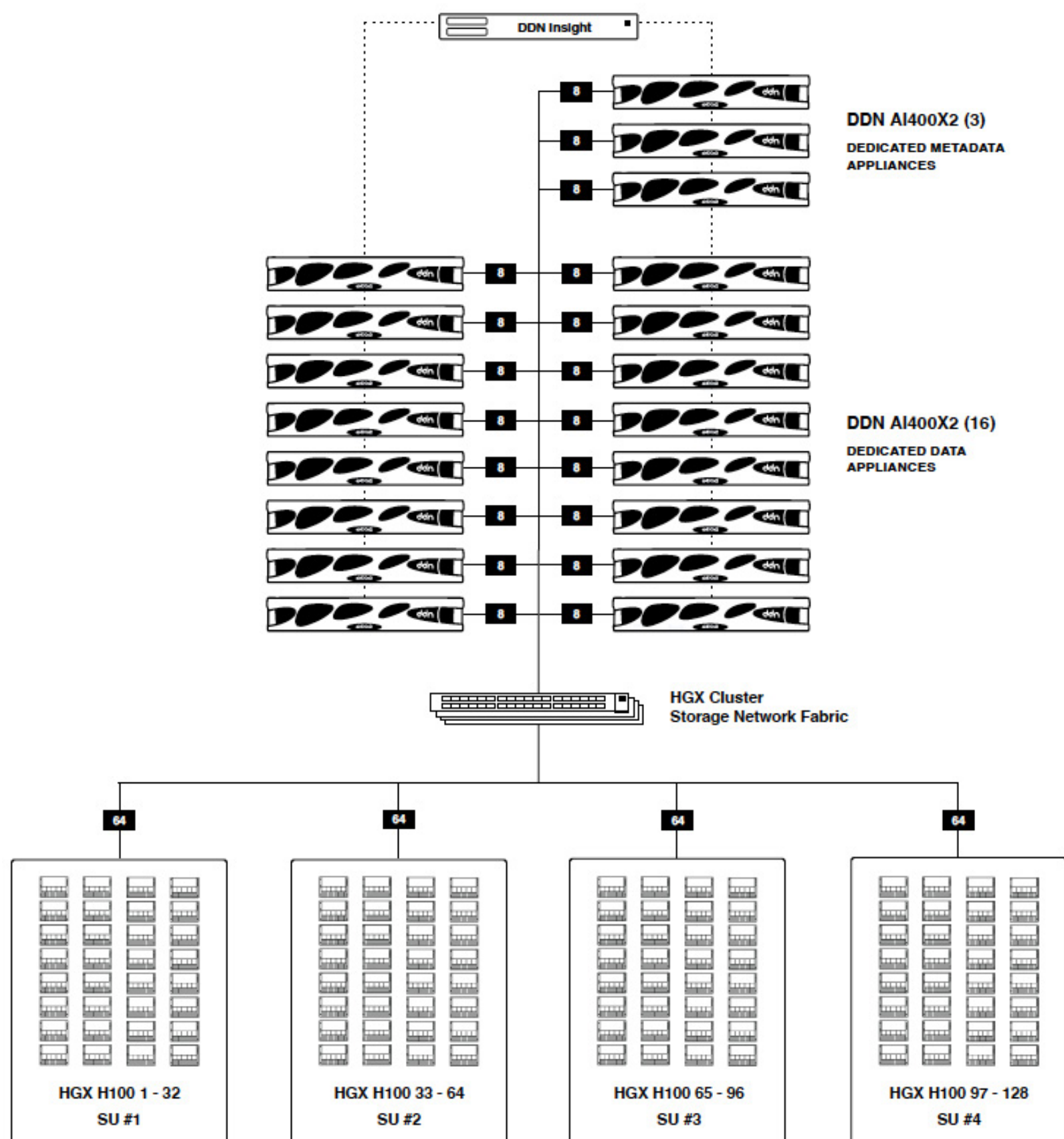


Figure 14. DDN A³I HGX cluster reference architecture with 128 HGX H100 systems.

3.9 HGX CLUSTER WITH 256 HGX H100 SYSTEMS

Figure 15 illustrates the DDN A³I architecture for HGX cluster with 256 HGX H100 systems, 6 DDN AI400X2 metadata appliances, 32 DDN AI400X2 data appliances and a DDN Insight server. Every HGX H100 system connects to the storage network with two NDR 400Gb/s InfiniBand links. Each AI400X2 appliance connects to the storage network with 8 InfiniBand links using the appropriate cable type. The DDN Insight server connects to the AI400X2 appliances over the 1GbE out-of-band management network. It does not require a connection to the InfiniBand storage network.

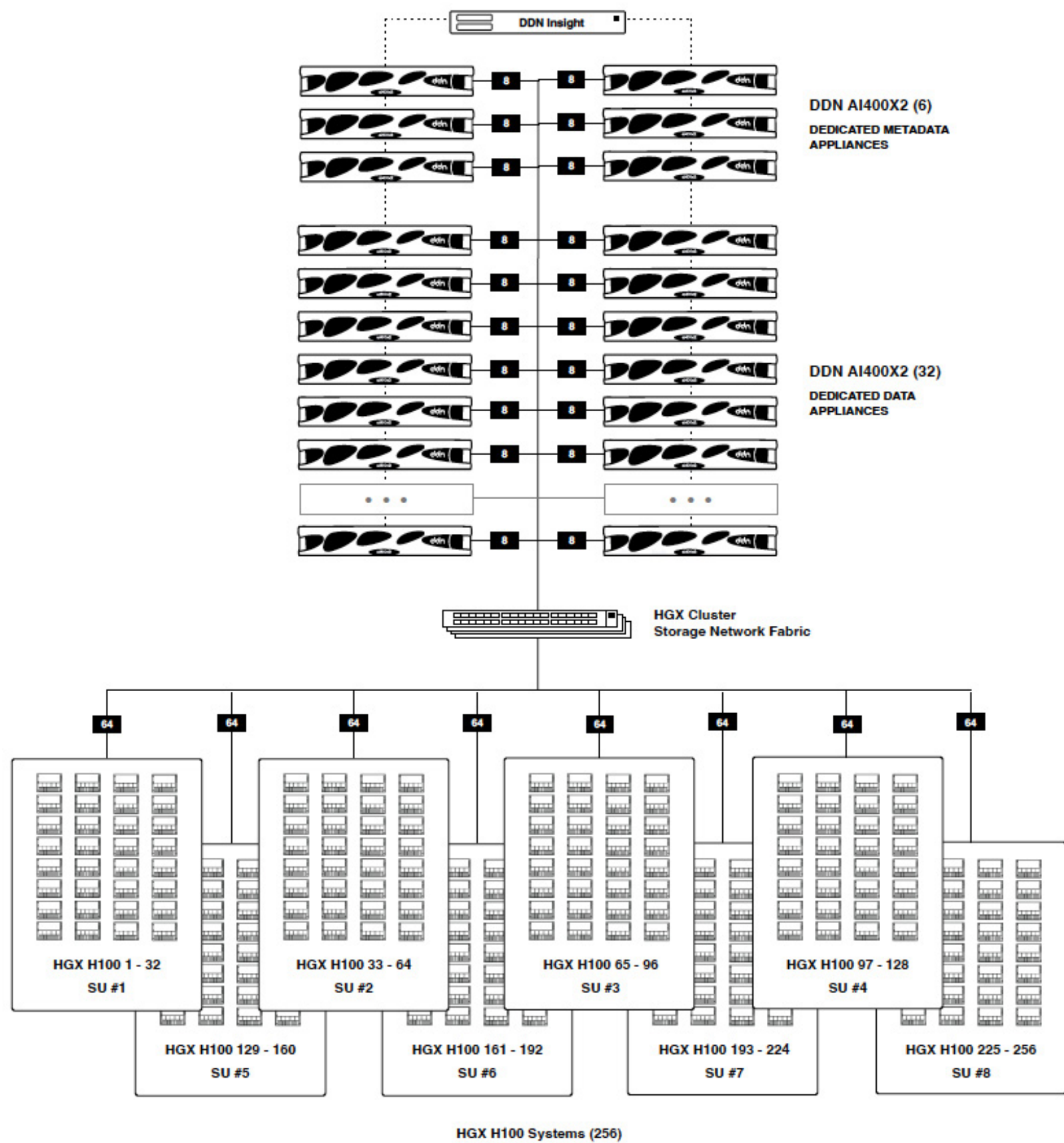


Figure 15. DDN A³I HGX cluster reference architecture with 256 HGX H100 systems.

4. DDN A³I SOLUTIONS WITH SUPERMICRO HGX CLUSTER VALIDATION

DDN conducts extensive engineering integration, optimization, and validation efforts in close collaboration with partners and customers to ensure the best possible end-user experience using the reference designs in this document. The joint validation confirms functional integration, and optimal performance out-of-the-box for HGX cluster configurations.

Performance testing on the DDN A³I architecture is conducted with industry standard synthetic throughput and IOPS applications, as well as widely used DL frameworks and data types. The results demonstrate that with the DDN A³I shared parallel architecture, GPU-accelerated applications can engage the full capabilities of the data infrastructure and the HGX systems. Performance is distributed evenly across all the HGX systems in the HGX cluster, and scales linearly as more HGX systems are engaged.

This section details some of the results from recent at-scale testing integrating AI400X2 appliances with HGX H100 systems in a HGX cluster deployment.

4.1 HGX CLUSTER PERFORMANCE VALIDATION

The tests described in this section were executed on thirty-two HGX H100 systems each equipped with eight H100 GPUs and four AI400X2 appliances running DDN EXAScaler 6.3.1.

For the storage network, all thirty-two HGX H100 systems are connected to NVIDIA Quantum-2 QM9700 switches with two NDR 400Gb/s InfiniBand links, one per dual-ported adapter (see recommendation in section 3.1). The AI400X2 appliances are connected to the same network with eight NDR 200Gb/s InfiniBand links each.

This test environment allows us to demonstrate storage performance with a HGX H100 Cluster scalable unit configuration.

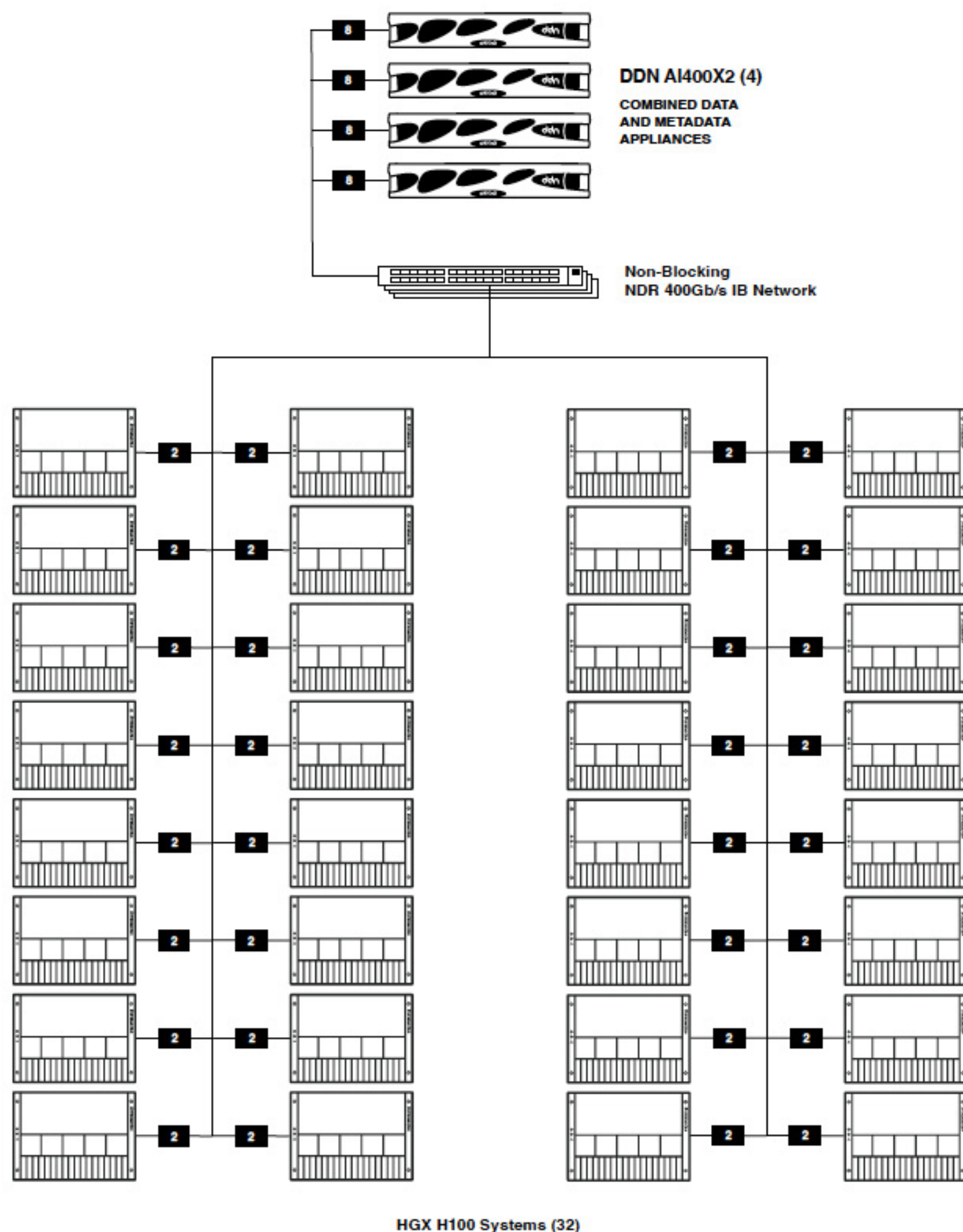


Figure 16. Test environment with 32 HGX H100 systems and 4 AI400X2 appliances.

This test demonstrates the peak performance of the DDN HGX cluster reference architecture using the fio open-source synthetic benchmark tool. The tool is set to simulate a general-purpose workload without any performance-enhancing optimizations. Separate tests were run to measure both 100% read and 100% write workload scenarios.

Below are the specific FIO configuration parameters used for these tests:

- » blocksize = 1024k
- » ioengine = posixaio
- » direct = 1
- » bw-threads = 255
- » iodepth = 128

In Figure 17, test results demonstrate that the DDN solution can deliver over 99 GB/s of read throughput and over 90 GB/s of write throughput to a single HGX H100 system. In this test, data is accessed through a single posix mount provided by the DDN shared parallel filesystem client installed on the HGX H100 system. The test demonstrates that the HGX H100 system can utilize the full network capabilities on the HGX cluster storage network.

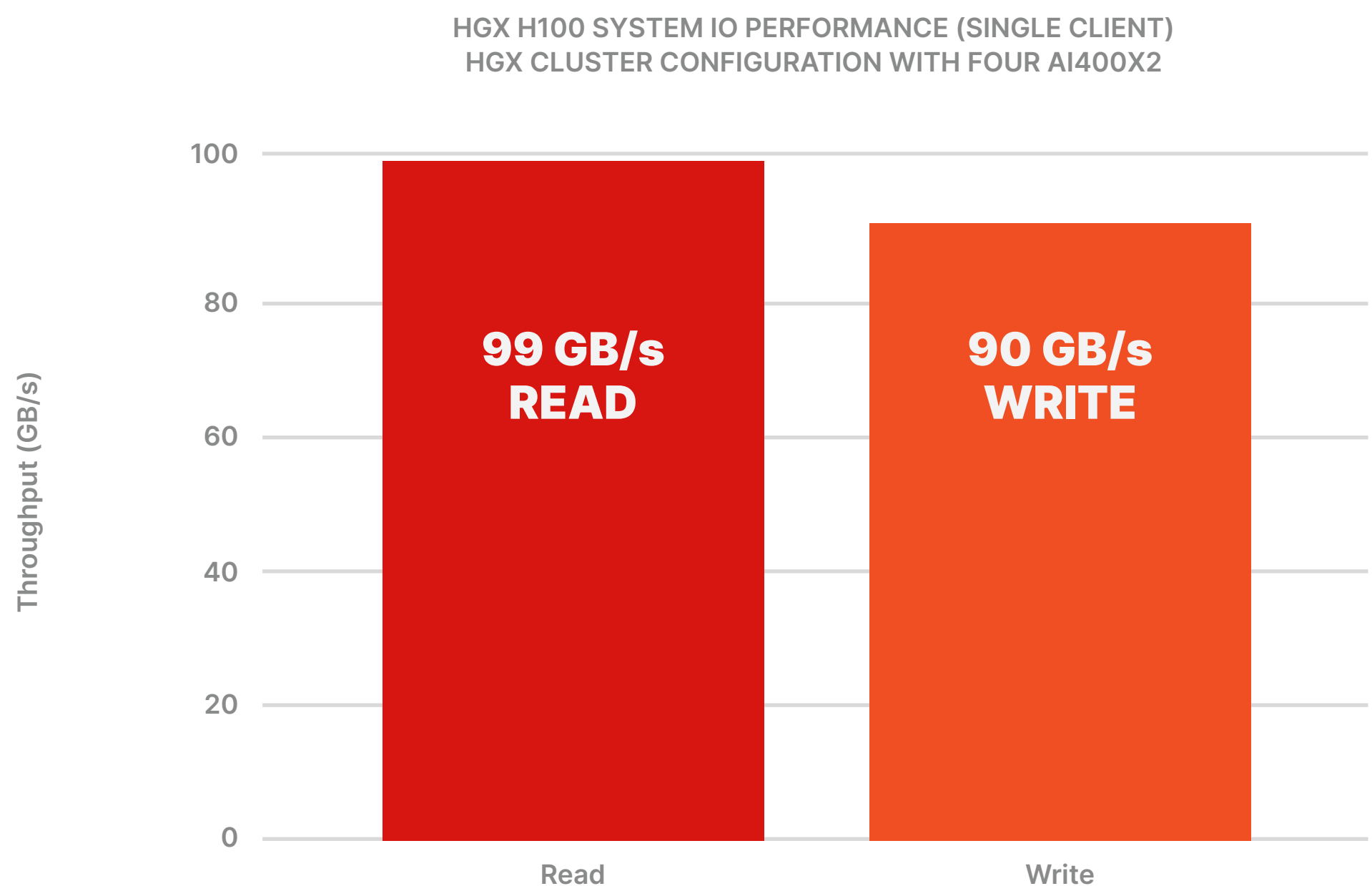


Figure 17. FIO throughput with HGX cluster configuration using a single HGX H100 system.

Figure 18 demonstrates that the DDN software evenly distributes the full read and write performance of the AI400X2 appliances with all thirty-two HGX H100 systems engaged simultaneously. The DDN solution utilizes both network links on every HGX H100 system, ensuring optimal performance for a very wide range of data access patterns and data types.

This test demonstrates that the storage solution is provisioned to deliver at least 1.4 GB/s of read and 1 GB/s of write throughput for every GPU in the HGX cluster.

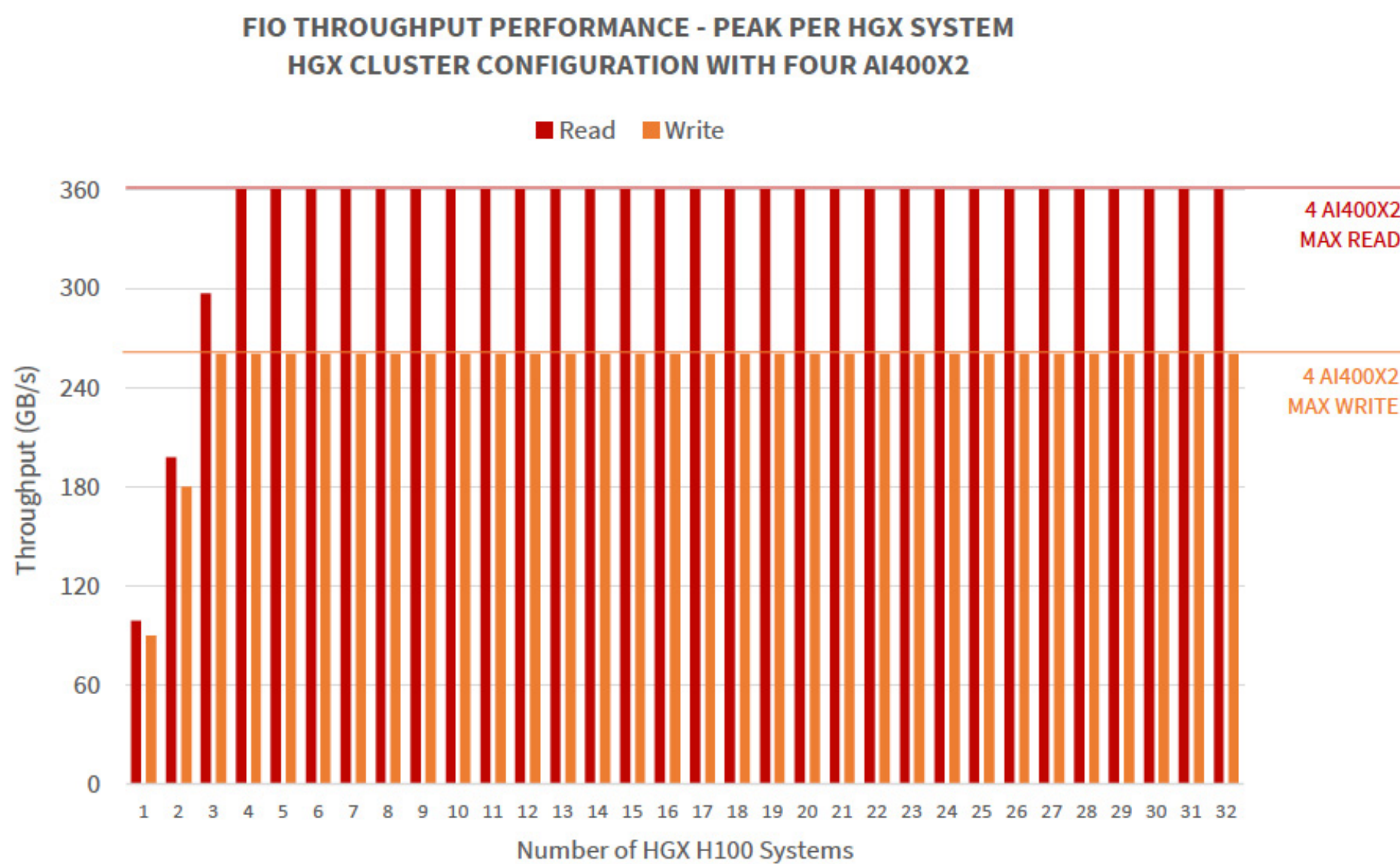


Figure 18. FIO throughput with HGX cluster configuration using 32 HGX H100 systems.

6. CONTACT DDN TO UNLEASH THE POWER OF YOUR HGX CLUSTER

DDN has long been a partner of choice for organizations pursuing at-scale data-driven projects. Beyond technology platforms with proven capability, DDN provides significant technical expertise through its global research and development and field technical organizations.

A worldwide team with hundreds of engineers and technical experts can be called upon to optimize every phase of a customer project: initial inception, solution architecture, systems deployment, customer support and future scaling needs.

Strong customer focus coupled with technical excellence and deep field experience ensures that DDN delivers the best possible solution to any challenge. Taking a consultative approach, DDN experts will perform an in-depth evaluation of requirements and provide application-level optimization of data workflows for a project. They will then design and propose an optimized, highly reliable and easy to use solution that best enables and accelerates the customer effort.

Drawing from the company's rich history in successfully deploying large scale projects, DDN experts will create a structured program to define and execute a testing protocol that reflects the customer environment and meet and exceed project objectives. DDN has equipped its laboratories with leading GPU compute platforms to provide unique benchmarking and testing capabilities for AI and DL applications.

Contact DDN today and engage our team of experts to unleash the power of your AI projects.