FS H100 InfiniBand Solution Delivery Manual

Contents

1. Program Overview

1.1 Overview

In the comprehensive AI industry solution for large-scale model training, several strategies are crucial. These include building an efficient distributed deep learning framework, implementing high-performance network configurations to optimize data transfer efficiency, adopting intelligent data management and optimization strategies to enhance computational resource utilization, and employing model compression and quantization techniques to reduce computational load.
Specifically, the integration of InfiniBand (IB) technology is a key component of this solution. IB technology significantly optimizes data transfer efficiency in distributed systems by providing high throughput and low latency network communication, thus reducing communication delays during the training process and accelerating the training of large models. The combined application of these strategies and technologies not only accelerates model training but also improves training efficiency and accuracy, providing robust support for the innovation and application of AI technology.

1.2 InfiniBand Networking Solutions

1.2.1 Business Background

In the rapidly evolving AI industry, particularly in computationally intensive applications such as large-scale model training, the demand for computational resources is sharply increasing. These complex models, similar to deep learning networks used for natural language processing or image recognition, are undergoing swift iterations. Their parameter counts are growing, which in turn heightens the need for computing power.
For instance, advanced models that process and analyze massive datasets have complexities and computational requirements comparable to cutting-edge language models like GPT-4. Traditional single GPU servers can no longer meet the demands of training these advanced models, as they fall short in terms of memory, storage capacity, and training efficiency. Therefore, building distributed training systems has become essential to support the technological advancements in large model training.

Distributed training involves connecting multiple GPU nodes to build a cluster with powerful computational and memory capabilities. However, the overall computational power of the cluster does not increase linearly with the number of GPU nodes. The efficiency of model training is not only influenced by the computation time of individual GPUs but also by the efficiency of communication between nodes. Therefore, building a high-performance AI/ML network directly impacts the efficiency of communication between GPU nodes, thereby significantly affecting the performance and throughput of the entire AI/ML cluster. This is crucial for the development of large-scale model training techniques.

In addressing the challenges faced by high-performance Ethernet networks in supporting AI and GC applications such as large-scale model training, InfiniBand (IB) network technology offers an efficient solution. IB networks provide significant advantages to distributed training systems in the field of large-scale model training:

- High bandwidth and low latency: IB networks offer transfer rates of up to 400Gb/s and latency in the microsecond range, providing necessary network performance for large-scale distributed training tasks.
- Efficient load balancing and congestion control: With advanced load balancing and congestion control mechanisms, IB networks can effectively improve bandwidth utilization, addressing related issues in traditional Ethernet networks.
- Native RDMA support: IB networks support native RDMA, allowing direct data exchange between nodes' memories, bypassing the CPU, thus reducing latency and CPU load and simplifying the optimization process.
- Ease of management and scalability: IB networks provide efficient and flexible solutions for configuration and management, adapting to the rapidly changing demands of large-scale model training technologies.

These advantages make IB networks an ideal choice for supporting large-scale, efficient AI model training, contributing to further development and innovation in AI technology.

1.2.2 Construction Needs

The core goal of constructing high-performance networks for AI/ML is to improve the computational efficiency of business and further reduce the time required for training.

The construction goals mainly include three aspects:

- Large-scale networking: AIGC businesses require large-scale AI/ML cluster networking and distributed parallel computing. Balancing the scale of the cluster with GPU efficiency is essential. Therefore, AI/ML networks need to support connections of thousands or even tens of thousands of GPUs and facilitate flexible and scalable expansion.
- High-throughput networks: With the increase in the proportion of inter-machine network communication for parallel large models, network transmission efficiency becomes increasingly important for GPU computing efficiency. Upgrading server access bandwidth from 100G to 400G is gradually becoming mainstream. To maximize business throughput efficiency, networks continuously pursue higher bandwidth utilization rates while also requiring extremely low forwarding latency and avoiding packet loss events. Therefore, reducing forwarding hops, lowering latency, and achieving lossless transmission are core objectives.
- Easy deployment and operation: The construction cycle of AI/ML networks is tight, requiring support for rapid business deployment. At the same time, network failures can have a significant impact on business. AI/ML networks need to support network telemetry, enabling rapid fault location within minutes and fast self-recovery to ensure uninterrupted AI/ML training tasks.

1.2.3 Design Ideas

When selecting network technologies suitable for AI/ML applications, especially considering the different requirements of large model training and AI inference, the industry mainly considers two technological paths: InfiniBand and RoCE v2. Here is a summary of the main advantages and applicable scenarios for these two technologies:

1.2.3.1 InfiniBand Technology

- Ultra-low latency: InfiniBand provides extremely low static and dynamic forwarding latency, which is crucial for latency-sensitive large model training scenarios.
- High-performance computing suitability: It is ideal for high-performance computing and processing large-scale AI/ML workloads, ensuring fast and accurate data transfer between nodes.
- Enhanced compatibility and interoperability: Although traditionally seen as relatively closed, InfiniBand is evolving to support a broader range of applications and environments, improving its compatibility and interoperability.
- Applicable scenarios: Particularly well-suited for large model training, where the demand for network performance is exceptionally high.

1.2.3.2 RoCE v2 Technology

- Performance similar to InfiniBand: It provides competitive network performance, with dynamic latency comparable to InfiniBand.
- Standardization and openness: Based on widely adopted Ethernet standards, it is easier to integrate into existing network architectures, enhancing openness and standardization.
- Cost-effectiveness: It offers relatively lower initial investment and operational costs, making it suitable for budget-sensitive projects.
- Configuration and optimization: While it may require complex network configuration and optimization to achieve optimal performance, this provides network professionals the flexibility to fine-tune the network to meet specific needs.
- Applicable scenarios: More suitable for AI inference scenarios, where cost-effectiveness and network compatibility are primary considerations.

Based on the analysis above, InfiniBand, with its outstanding performance and low latency, is more suitable for large model training tasks that typically require extremely high data transfer efficiency and low communication latency. In contrast, RoCE v2, due to its cost-effectiveness, ease of integration, and standardization, is better suited for AI inference scenarios, especially in applications where cost and network compatibility are significant considerations.

1.2.4 Solution Architecture

1.2.4.1 GPU Computing Network

Based on IB technology, FS has introduced an IB network solution:

- GPU server interconnection: This ensures the connection between all GPU servers for efficient data exchange and processing capabilities.
- 400G high-speed network interconnect: Utilizing a 400G InfiniBand network to achieve high-speed interconnectivity between GPU servers, providing extremely high data transfer rates and low-latency communication to support large-scale AI/ML workloads.
- Subnet manager optimized architecture: Employing a multi-plane network architecture optimized with Subnet Manager ensures that GPUs with the same ID on different servers connect to the same Top of Rack (ToR) switch. This design reduces the number of hops for communication between GPUs with the same ID, thereby optimizing AI/ML computational performance.
- InfiniBand's GPU direct RDMA technology: Leveraging InfiniBand's GPU Direct RDMA technology for efficient communication between GPUs. GPU Direct RDMA allows direct data transfer between GPUs, bypassing the CPU, thus reducing latency and improving data transfer efficiency, ensuring high-performance execution of computational tasks.

By using InfiniBand technology, the GPU computing network can achieve more efficient data processing and transmission capabilities, making it particularly suitable for large-scale AI/ML computing environments with stringent latency and throughput requirements. This optimized network architecture further enhances computational performance, ensuring optimal performance in executing complex AI/ML tasks.

1.2.4.2 Storage Network

- All Server San servers and GPU servers are interconnected.
- Use 100G network to interconnect GPU servers and storage servers.
- The BGP protocol is used between Leaf and Spine. BGP's powerful routing control capability ensures the optimal forwarding path of the storage network and ensures that storage does not experience high latency.

1.2.4.3 Storage Management Network

- All Server San servers and GPU servers are interconnected.
- All Server San storage management networks use separate 25G network cards for interconnection. Ensure that the management network is isolated from other networks to avoid mutual influence.
- The BGP protocol is used between Leaf and Spine. BGP's powerful routing control capability ensures the optimal forwarding path of the storage network and ensures that the storage is in a low-latency forwarding state.

1.2.4.4 Device Management Network

- The management ports of all servers and switches in the network are connected to the management network.
- The switch uses 1G network in the downlink to connect various devices, and 10G equipment in the uplink to interconnect the switches.

Figure 1-1 GPU Computing Network Architecture Diagram Example

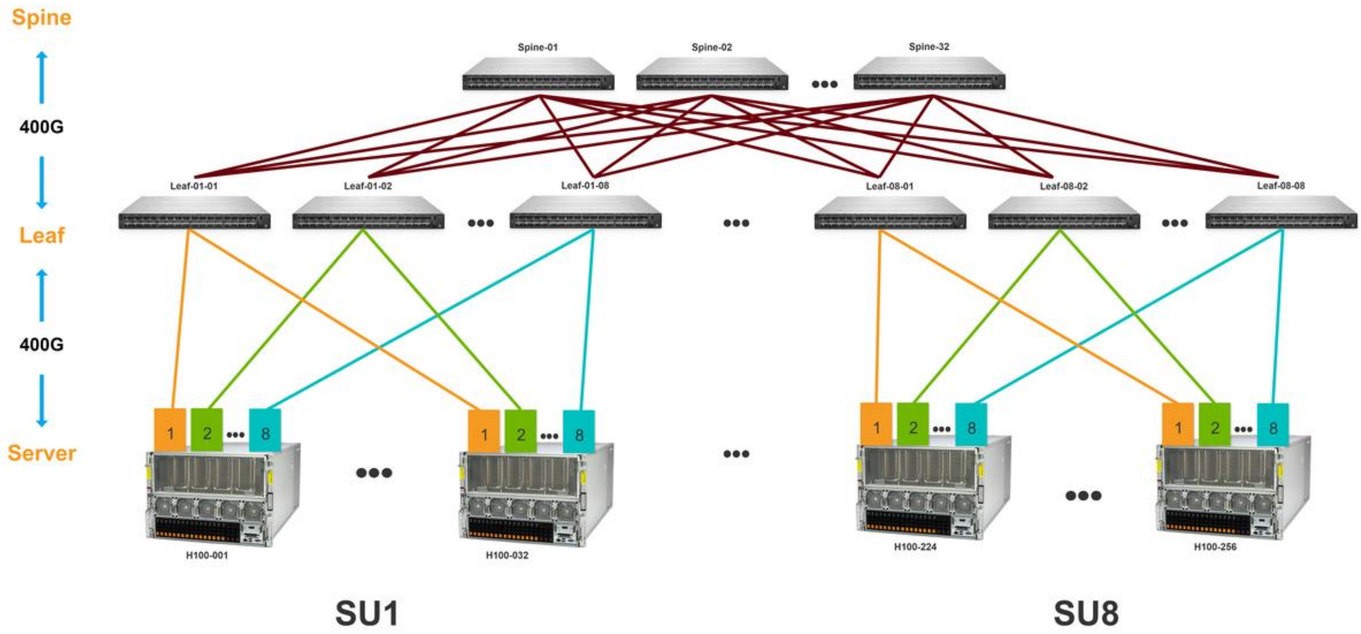


Figure 1-2 Storage Network Architecture Diagram Example

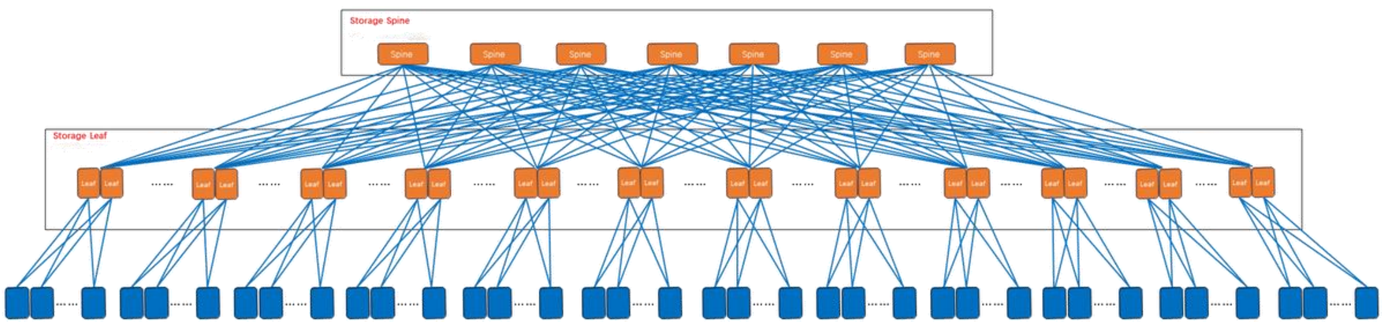
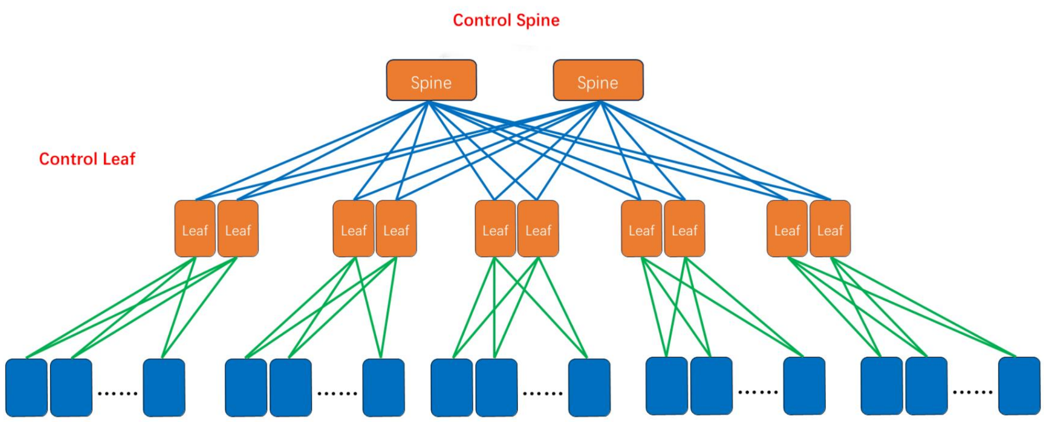


Figure 1-3 Control (Storage + Device) Network Diagram Example



1.2.5 Solution Benefits

This solution utilizes RDMA technology to build a lossless network for high-performance computing clusters, meeting the three core requirements of "low latency," "high throughput," and "zero packet loss" for AI/ML network construction. The main benefits are as follows:

- Massive Compute Power: Utilizing InfiniBand technology, the large-scale compute configuration, combined with tiered networking strategies, provides flexible, on-demand deployment capabilities for AI/ML applications. In this setup, each server is equipped with 8 GPU cards, 2 x 100G NICs, and 2 x 25G NICs. High-speed connections are achieved through InfiniBand switches, where a tier-2 network supports up to 2048 GPUs and delivers up to 256,000 TFLOPS of compute power, while expanding to a tier-3 network supports up to 16,000 GPUs, offering up to 2,048,000 TFLOPS of compute power.
- High Bandwidth Utilization: Supports 400G access, featuring industry-leading load balancing technology to build a 1:1 non-blocking CLOS network.
- Ease of Deployment: Supports one-click RDMA configuration deployment, compatible with cloud platform deployment, and has extensive experience in large-scale RDMA parameter tuning and multiple application cases.
- Visualized Intelligent Operations and Maintenance: Full process WEB operations, real-time telemetry, and visual management, with intelligent network-wide performance tuning.
- Rapid Project Delivery: Shortens project cycles, accelerates business deployment, and helps customers quickly capture the AI market.
- Ultra-Low Latency for AI Applications: The computing network built with InfiniBand (IB) switches achieves ultra-low network latency, meeting high-performance requirements for AI applications.
- Flexible Scalability and Adaptability: The Spine-Leaf two-tier network architecture meets current network operational requirements while providing flexibility and reliability for future business expansion.
- Customized Solutions: Offers customized solutions for storage, in-band management, and out-of-band management, effectively controlling project costs and achieving precise configurations based on customer budget needs.
- Professional Technical Team: Authorized by NVIDIA, with professional solution engineers and numerous successful cases, ensuring the best solutions for customers.
- High Security: Network isolation between different services, redundancy firewalls for critical services, ensuring high reliability and security of the network.

1.2.6 Solution Planning

Note

- The purpose of solution planning is to select the network architecture and determine the quantity of switches, optical modules, and cables based on the actual AI/ML business requirements, facilitating subsequent cost estimation and hardware procurement.
- This document addresses HGX H100 servers, each fully equipped with 8 H100 GPU cards. The following plan is for reference only and can be flexibly adjusted according to the actual situation.

1.2.6.1 GPU Computing Network

1. Network Architecture Selection

The choice of network architecture is primarily based on the comparison between the number of ports on the Spine switch and the total number of ports required by all switches in the network. If the number of ports on a single Spine switch exceeds the total port demand of all switches, a two-tier architecture is suitable. Conversely, if the Spine switch's port count is insufficient to meet the connection needs of all switches, a three-tier architecture is necessary.

For example, in a network design, if a Spine switch has 64 ports of 400G, it can directly connect to 64 Leaf switches. Considering a pod configuration of 8 Leaf switches, this allows for up to 8 pods within a two-tier architecture. Beyond this scale, such as adding a 9th pod, the Spine switch will lack sufficient ports for connections, necessitating a shift to a three-tier architecture.

2. Configuring Network Interface Cards Based on GPU Card Quantity

When considering server configurations, it is recommended to follow NVIDIA's suggestion of a 1:1 ratio between GPUs and NICs to maximize performance. Furthermore, the number of NICs determines the number of Leaf switches within a single pod.

For example, if a single GPU server is configured with 8 NICs, it indicates that one pod should consist of 8 Leaf switches to ensure adequate network connectivity.

3. Network Connection Ideas

In a two-tier architecture, the design is relatively simple, with scalability mainly limited by the port count on the Spine switch and the configuration of Leaf switches within a single pod. To maintain a 1:1 network convergence ratio, the number of ports on each Leaf switch connecting to servers should be half of its total port count, with the remaining ports used for connections to the Spine switch.

Overall, by considering the port capacity of Spine switches and the configuration needs of pods, it is possible to effectively plan network architecture and server configuration to meet specific performance and scalability requirements, while avoiding the inclusion of sensitive project or geographical information.

Figure 1-4 GPU Network Diagram Example

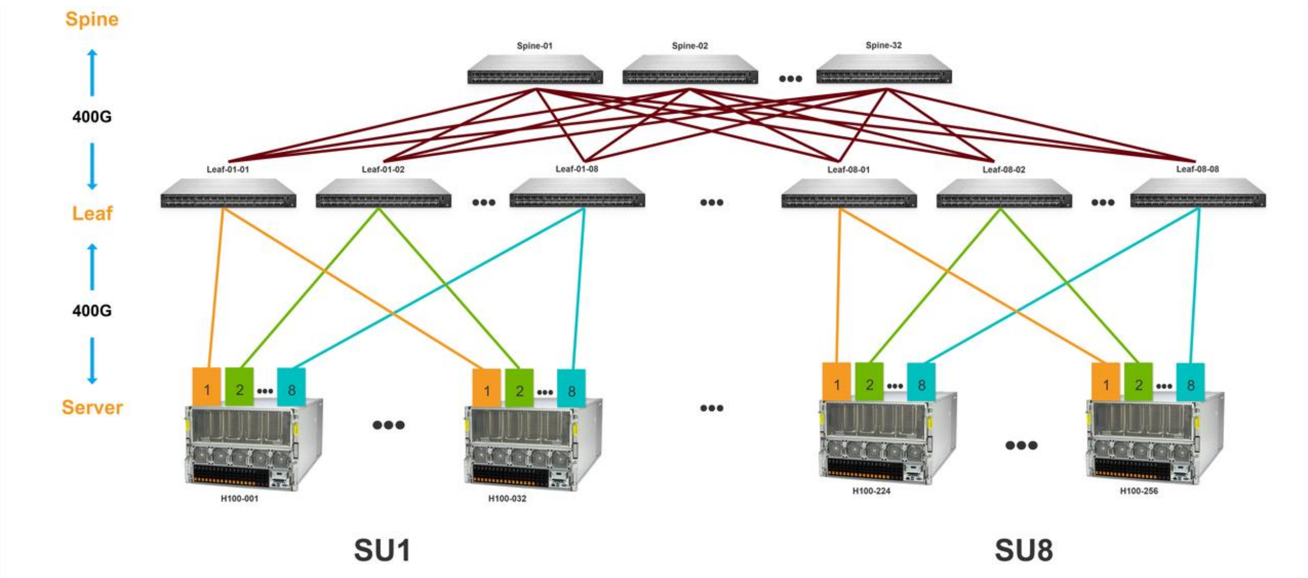


Table 1-5 GPU Network Planning Estimation Table

SU count	Node count (server num)	GPU count	NIC count	InfiniBand Switch Count(MQM9700/9790)		
				Leaf Switch count	Spine Switch count	Core Switch count
1	32	256	256	8	4	/
2	64	512	512	16	8	/
4	128	1024	1024	32	16	/
8	256	2048	2048	64	32	/
16	512	4096	4096	128	128	64
32	1024	8192	8192	256	256	128
64	2048	16384	16384	512	512	256
SU count	Node count (server num)	GPU count	NIC count	Cable Count		
				Node-Leaf	Leaf-Spine	Spine-core
1	32	256	256	256	256	/
2	64	512	512	512	512	/
4	128	1024	1024	1024	1024	/
8	256	2048	2048	2048	2048	/
16	512	4096	4096	4096	4096	/
32	1024	8192	8192	8192	8192	/
64	2048	16384	16384	16384	16384	/

SU count	Node count (server num)	GPU count	NIC count	Optical Modules Count		
				400G (node)	800G (Leaf)	800G (Spine)
1	32	256	256	256	256	128
2	64	512	512	512	512	256
4	128	1024	1024	1024	1024	512
8	256	2048	2048	2048	2048	1024
16	512	4096	4096	4096	4096	2048
32	1024	8192	8192	8192	8192	4096
64	2048	16384	16384	16384	16384	8192

1.2.6.2 Storage Network

In this network scenario, distributed storage is used, necessitating a separate storage network to ensure data exchange between storage devices and provide storage access for the front-end GPU servers.

The core network configuration is as follows:

- 45 storage servers, each equipped with 2 dual-port 100G NICs.
- 256 GPU servers, each equipped with 1 dual-port 100G NIC.
- 43 other servers, each equipped with 2 dual-port 100G NICs.

Spine Switches

- 4 N8550-64C switches serve as the Spine switches for the storage network, interconnecting the Leaf switches.

Storage Server Leaf Switches

- 2 N8550-64C switches act as one group of Leaf switches for storage servers, configured with MLAG. Each Leaf switch connects to the Spine switches using 4 links.

GPU Server & Other Server Leaf Switches

- 12 N8550-64C switches serve as 6 groups of Leaf switches for GPU and other servers, configured with MLAG. Each Leaf switch connects to the Spine switches using 1 link.

1.2.6.3 Storage Management Network

The core network configuration is as follows:

- 38 storage servers, each equipped with 2 dual-port 100G NICs.
- 256 GPU servers, each equipped with 1 dual-port 100G NIC.
- 43 other servers, each equipped with 2 dual-port 100G NICs.

Spine Switches

- 2 N8550-64C switches serve as the Spine switches for the storage network, interconnecting the Leaf switches.

Storage Server Leaf Switches

- 2 N8550-48B8C switches act as one group of Leaf switches for storage servers, configured with MLAG. Each Leaf switch connects to the Spine switches using 1 link.

GPU Server & Other Server Leaf Switches

- 14 N8550-48B8C switches serve as 7 groups of Leaf switches for GPU and other servers, configured with MLAG. Each Leaf switch connects to the Spine switches using 1 link.

1.2.6.4 Device Management Network

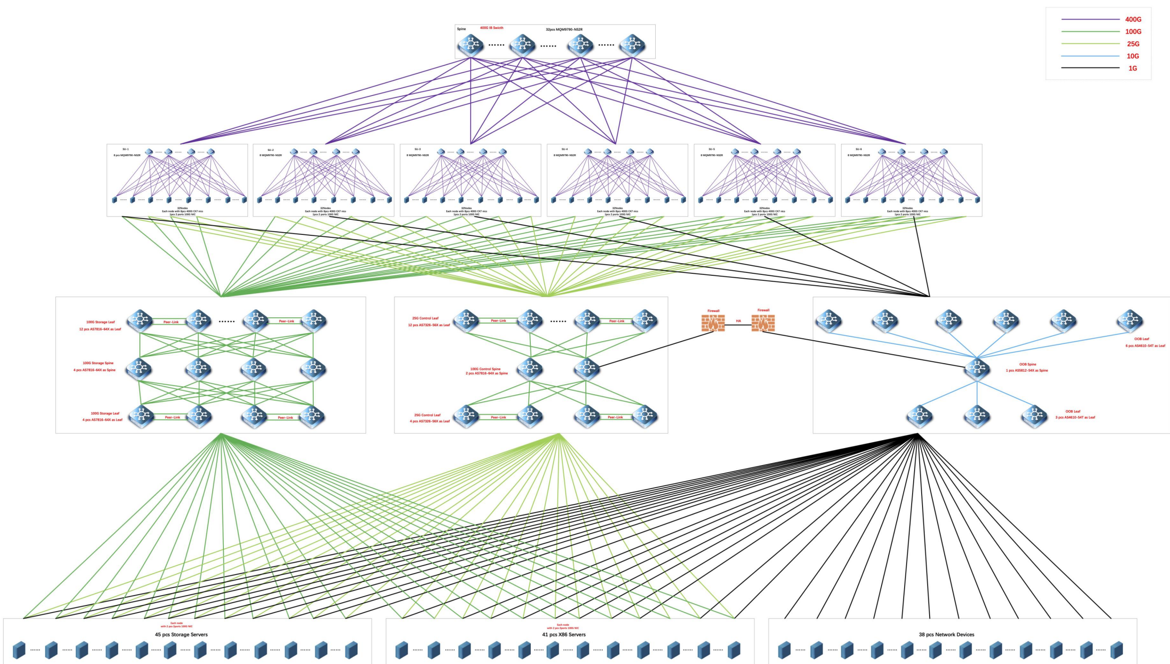
Spine Switch

- 1 N5850-48S6Q switch serves as the core switch for the device management network.

Leaf Switches

- 7 S5870-48T6S switches serve as Leaf switches, connecting all network devices.

Figure 1-6 Overall Network Topology Diagram Example



2. Product Overview

2.1 Switch

2.1.1 MQM9790-NS2F

- **Product Appearance**

Figure 2-1 MQM9790-NS2F Front View

Figure 2-2 MQM9790-NS2F Rear View

- **Product Specifications**

Table 2-3 MQM9790-NS2F Key Specifications

Specification Name	Specification Description
Ports	64 x NDR 400G 32 OSFP Connectors
CPU	x86 Coffee Lake i3
Switch Chip	NVIDIA QUANTUM-2
System Memory	Single 8GB
Storage	M.2 SSD SATA 16GB 2242 FF
Max. Power Consumption	1703w
Airflow	Back-to-Front(P2C)
Fan Number	6+1 Hot-swappable

2.1.2 N8550-64C

- **Product Appearance**

Figure 2-4 N8550-64C Front View

Figure 2-5 N8550-64C Rear View

- **Product Specifications**

Table 2-6 N8550-64C Key Specifications

Specification Name	Specification Description
Ports	64x 100G QSFP28
CPU	Cavium CN7230(Quad-core, 1.5 GHz)
Switch Chip	BCM56970
Switching Capacity	12.8 Tbps
MAC Address	72K
ARP Entries	90,000
IPv4/IPv6 Routes	128,000/64,000
Hot-swappable Fans	3(2+1 Redundancy), Front-to-Back

2.1.3 N8550-48B8C

- **Product Appearance**

Figure 2-7 N8550-48B8C Front View

Figure 2-8 N8550-48B8C Rear View

- **Product Specifications**

Table 2-9 N8550-48B8C Key Specifications

Specification Name	Specification Description
Ports	48x 25G SFP28, 2x 10Gb SFP+, 8x 100G QSFP28
CPU	Intel Xeon D-1518 processor quad-core 2.2 GHz
Switch Chip	Broadcom BCM56873 Trident III
Switching Capacity	4 Tbps full duplex
MAC Address	98K
Hot-swappable Fans	6 (5+1 Redundancy)

2.1.4 N5850-48S6Q

- **Product Appearance**

Figure 2-10 N5850-48S6Q Front View

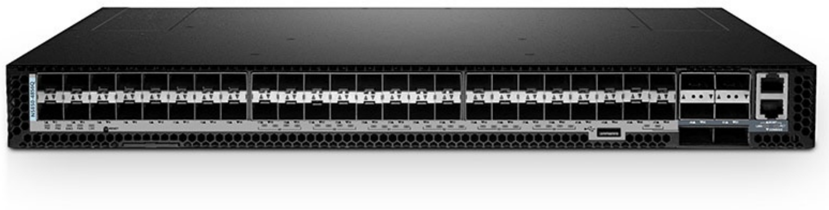


Figure 2-11 N5850-48S6Q Rear View

- **Product Specifications**

Table 2-12 N5850-48S6Q Key Specifications

Specification Name	Specification Description
Ports	48x 10G SFP+, 6x 40G QSFP+
CPU	Intel Atom C2538 processor quad-core 2.4GHz
Switch Chip	Broadcom BCM56864 Trident II+
Switching Capacity	1.44 Tbps full duplex
Hot-swappable Fans	5(4+1 Redundancy)

2.1.5 S5870-48T6S

- **Product Appearance**

Figure 2-13 S5870-48T6S Front View

Figure 2-14 S5870-48T6S Rear View

- **Product Specifications**

Table 2-15 S5870-48T6S Key Specifications

Specification Name	Specification Description
Ports	48x 10/100/1000BASE-T RJ45 | 6x 1G/10G SFP+
CPU	Intel Atom COMe Module with C3508 4-Core 1.6 GHz x86 Processor
Switch Chip	Broadcom BCM56277 Trident III
Switching Capacity	108 Gbps
MAC Address	64K min./112K max
Hot-swappable Fans	3(2+1 Redundancy)

2.2 Server

2.2.1 H100 Server

- **Product Appearance**

Figure 2-16 H100 Server Rear View

- **Product Specifications**

Table 2-17 H100 Server Key Specifications

Specification Name	Specification Description
Motherboard	
Motherboard	X13DEG-OAD
Processor	
CPU	5th Gen Intel® Xeon® / 4th Gen Intel® Xeon® Scalable processors and Intel® Xeon® Max Series processors
Core	Up to 64 cores
System Memory	
Memory Capacity	32 DIMM slots Up to 8TB 3DS ECC RDIMM, DDR5-5600MHz
Memory Type	5600/5200/4800/4400/4000 MT/s ECC DDR5 RDIMM(3DS) Up to 256GB of memory with speeds of up to 5600MT/s(1DPC) or 4400 MT/s(2DPC)

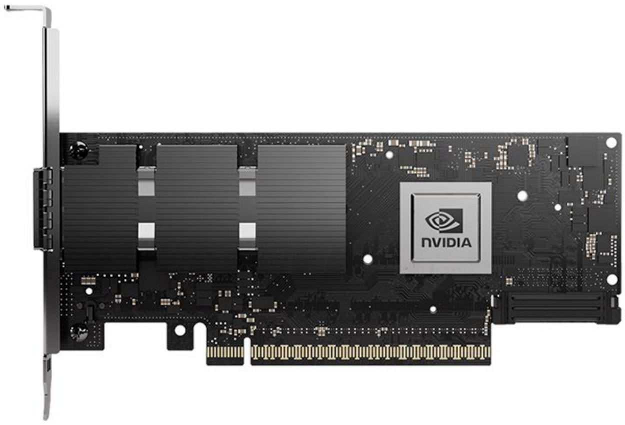
Specification Name	Specification Description
System Memory	
DIMM Sizes	16GB, 32GB, 64GB, 96(XCC Only)GB, 128GB, 256GB RDIMM: Up to 256GB
Memory Voltage	1.1V
Error Detection	Corrects single-bit errors Detects double-bit errors(using ECC memory)
Expansion Slots	
PCIe	8 PCIe Gen 5.0 X16 LP 4 PCIe Gen 5.0 X16 FHHL Slots
M.2	M.2 Interface: 2 PCIe 4.0 x4 Form Factor: 22110, 2280 Key: M-Key

2.3 Adapter

2.3.1 MCX75310AAS-NEAT

- **Product Appearance**

Figure 2-18 MCX75310AAS-NEAT Appearance



- **Product Specifications**

Table 2-19 MCX75310AAS-NEAT Key Specifications

Specification Name	Specification Description
Technology	InfiniBand (Default) / Ethernet
Ports	Single
Interface Type	OSFP
Host Interface	PCIe 5.0x 16
Data Rate	(IB) NDR/NDR200/HDR/HDR100/EDR/FDR/SDR; (Eth) 400/200/100/50/40/10/1Gb/s
SR-IOV	Yes
RDMA	Yes
Supported Operating Systems	Linux, Windows, VMware

2.4 Optical Module

2.4.1 OSFP-SR8-800G

- **Product Appearance**

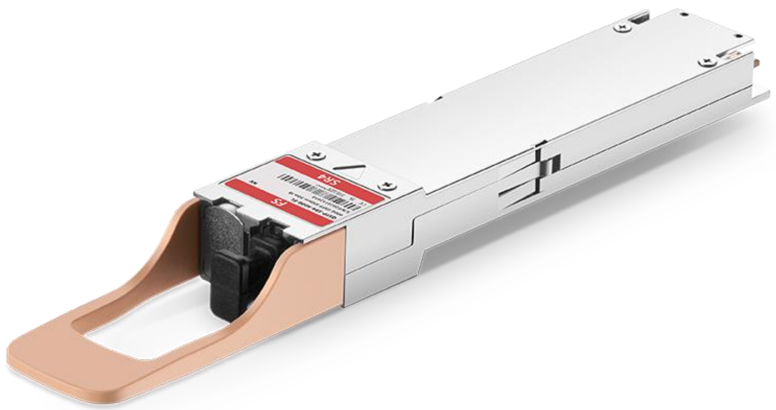
- **Product Specifications**

Table 2-20 OSFP-SR8-800G Key Specifications

Specification Name	Specification Description
Form Factor	OSFP
Max Data Rate	850Gbps (8x 106.25Gbps)
Wavelength	850nm
Max Cable Distance	30m@OM3/50m@OM4
Connector	Dual MPO-12/APC
Media	MMF
TX Power	-4.6~4dBm
Minimum Receiver Power	-6.4dBm
Powerbudget	1.8dB
Receiver Overload	4dBm

2.4.2 OSFP-SR4-400G-FL

- **Product Appearance**



- **Product Specifications**

Table 2-21 OSFP-SR4-400G-FL Key Specifications

Specification Name	Specification Description
Form Factor	OSFP
Max Data Rate	425Gbps (4x106.25Gb/s)
Wavelength	850nm
Max Cable Distance	30m@OM3/50m@OM4
Connector	MTP/MPO-12 APC
Media	MMF
TX Power	-4.6~4dBm
Minimum Receiver Power	-6.3~4dBm
Powerbudget	1.9dB

2.5 MTP Fiber Cable

2.5.1 MTP Jumper Cable

- **Product Specifications**

Table 2-22 MTP Trunk Cable Key Specifications

Specification Name	Specification Description
Connector A	US Conec MTP® Female (Pinless)
Connector B	US Conec MTP® Female (Pinless)
Polish Type	APC to APC
Fiber Mode	OM4 50/125μm
Wavelength	850/1300nm
40/100G Ethernet Distance	150m at 850nm
10G Ethernet Distance	400m at 850nm

2.6 Firewall

2.6.1 NSG-5220

- **Product Appearance**

Figure 2-23 NSG-5220 Front View

Figure 2-24 NSG-5220 Front View

- **Product Specifications**

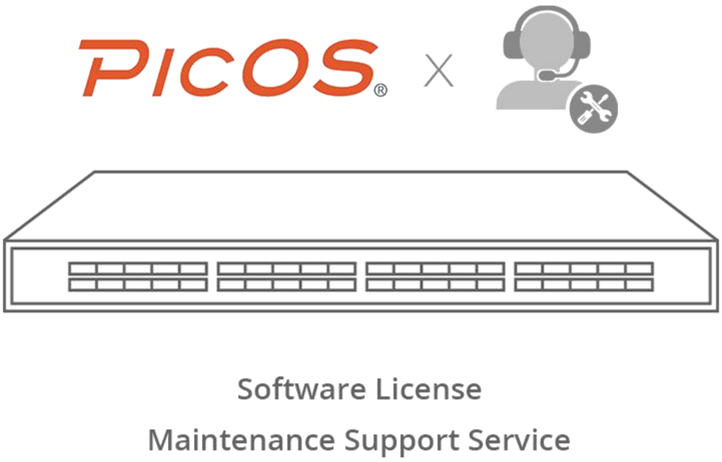
Table 2-25 NSG-5220 Key Specifications

Specification Name	Specification Description
Ports	16 x 1Gb Rj45, 8 × 1Gb SFP, 2× 10Gb SFP+
Firewall Throughput	Intel Atom C2538 processor quad-core 2.4GHz
Maximum Concurrent Sessions	3 Million
New Sessions/s	140,000
Power Specification	100W, Dual AC (default)

2.7 Software/License

2.7.1 PicOS

- **Software Interface**



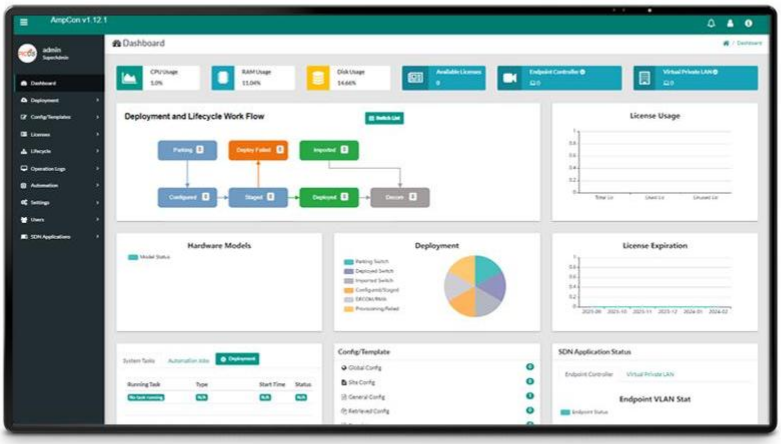
- **Product Specifications**

Table 2-26 PicOS Software Key Features

Function Description
Support MLAG, EVPN-VXLAN, LLDP, Voice VLAN
Support ACL, RADIUS, TACACS+, DHCP Snooping
Support A Full SUite of Spanning Tree Protocols and Security Mechanism
Easy to Use While Delivering Full Network Visibility
Interoperability with A Wide Range of Hardware Platforms
Provide Support for PicOS Installation, Configuration, or Operation
Assist in Identifying and Verifying the Causes of Suspected Errors
Growing Library of Online Support Resources and Documentation

2.7.2 AmpCon

- **Software Interface**



- **Product Specifications**

Table 2-27 AmpCon Software Key Features

Function Description
Zero-Touch Provisioning for Easy, Secure Scaling
Deployment and Configuration Management of Remote Switches at Scale
Support Configuration and License Management to Maintain Network Security
Scheduled Backups and Upgrades
Powerful, Agentless Automation with Ansible Playbooks
Infinite Expandability with Open APIs
Reduces Operating Expenses through Simplified Deployment and Lifecycle Management
Eliminates Human Error in Large-Scale Deployments and Upgrades

2.7.3 License for NSG-5220

- **Software Interface**

- **Product Specifications**

Table 2-28 License Key Features

Function Description
Hardware Maintenance & Software Update & Apply Recognition Library Update
Provide More Network Protection and Management Options
No Hidden Fees for Features

3. Implementation Preparations

Based on the equipment situation during project implementation, complete the exploration record form, provide inspection results according the evaluation criteria, and proceed with the commencement only after confirming that all meet the standards.

3.1 Engineering Exploration

Table 3-1 Inspection Demonstration Form

Classification	Serial Number	Check Item	Evaluation Criteria
Data Center Environmental Requirements	1	Ventilation	Leave enough space of 20cm on both sides of the chassis and on the front and rear panels to ensure normal heat dissipation.
	2	Anti-interference	Keep away from high-power (radar, radio transmitter), high-frequency and high-current equipment
	3	Aisle	The aisle width should be no less than 0.8 meters.
	4	Grounding	• Must be grounded via yellow-green safety ground wire. • The facility's lightning protection system is an independent system consisting of lightning rods, down conductors and connectors connected to the grounding system. • Grounding required for electromagnetic compatibility design, including: shielding grounding, filter grounding, noise and interference suppression, level reference. • The resistance between the chassis and the ground must be less than 0.1 ohm (can be confirmed with the data center administrator). Please refer to the hardware manual for different devices.

Classification	Serial Number	Check Item	Evaluation Criteria
Data Center Environmental Requirements	**5**	**Environmental Survey**	• Space Layout and Planning: Assess the size, shape, and available space of the data center to ensure there is sufficient room for equipment installation and maintenance. • Temperature and humidity control: The temperature and humidity in the computer room must be controlled within an appropriate range to ensure normal operation of the equipment. • Power Supply: Evaluate the data center's power requirements, including the need for UPS (Uninterruptible Power Supply) and backup generators. • Network and Data Communication Cabling: Assess the data center's network connectivity needs, including fiber optic and Ethernet cabling requirements. • Security and Monitoring: Ensure the data center's security measures, including access control, video surveillance, and fire protection systems. • Air Conditioning and Ventilation Systems: Evaluate appropriate air conditioning and ventilation systems to maintain temperature and humidity control within the data center.
	6	**Wind Duct Survey**	• Airflow Demand Analysis: Determine the amount of heat generated by the equipment in the data center and the level of heat that needs to be dissipated. Analyze the airflow requirements, including the volume and speed of incoming and outgoing air. • Evaluation of Existing System Capacity: Assess the capacity and efficiency of the current air duct system, including fans, air conditioning units, and other air handling equipment. Check whether the existing system can meet the demands of additional or anticipated loads.

Classification	Serial Number	Check Item	Evaluation Criteria
Data Center Environmental Requirements	6	Wind Duct Survey	• Air Duct Layout and Design: Examine the optimal air duct layout to achieve efficient air distribution and heat removal. Ensure that the air duct design supports even airflow distribution to avoid the formation of "hot spots."
Cabinet/Equipment Installation	7	Cabinet Brackets	The cabinet should have adjustable height "L" brackets on both sides. The specific load capacity should be determined according to the equipment involved in the project, with reference to the hardware installation manual of the corresponding product model for equipment weight.
	8	Cabinet Specifications	Should be a 4-post 19-inch standard cabinet. Check the cabinet's width (W - distance between the side perforated rails), depth (D - distance between the front and rear doors), and height (H), ensuring they meet the specifications of the equipment involved in the project. Specific equipment dimensions can be found in the hardware installation manual of the corresponding product model.
	9	Cabinet U Standard	Check whether the cabinet's U standard is 44.45mm, which is the standard IEC cabinet.
	10	Cabinet Grounding Point	There should be a grounding point within 3000mm for equipment grounding (grounding wire length L = 3000±30mm).
	11	Cabinet Power Outlets	Check the power supply mode of the cabinet PDU, take photos, and confirm whether it matches the equipment model involved in the project. If the on-site interface uses other standard interfaces, you need to inform us in advance.
	12	Cabinet Space	Sufficient space must be provided for equipment installation. The specific space will be evaluated based on the actual equipment involved in the project.

Classification	Serial Number	Check Item	Evaluation Criteria
Power Supply	13	**PDU Power Position**	The rear of the cabinet on both sides of the square hole strip along the depth of the equipment can not have PDU and other obstacles to prevent the depth of the equipment and interference, or impede the power module extraction and insertion, otherwise it will affect the installation of power supply modules
	14	**Power Cable Length**	• According to the default power cord specification of the equipment, check whether the length of the power cord at the site can meet the requirements, and the default length of the power cord of the equipment models involved in the project can be referred to the hardware installation manual of the corresponding models. • Need to pay attention to the site customer alignment specifications, the length of the power cord to consider whether to meet the requirements according to customer specifications winding. If the length of the power cord exceeds the default delivery length, you need to request the supply chain to increase the length of the power cord in advance.
	15	**Cabinet Power Consumption**	Cabinet PDU power supply to meet the equipment power consumption requirements, with the server room administrator to confirm the maximum load of the cabinet

3.2 Hardware Preparation

Evaluate server information involved in project implementation, including server model, quantity, network card configuration, etc.

Note

The following data is based on the 8 SU count in "Table 1-5 GPU Network Planning Estimation Table" for demonstration only. Please update the collection form according to the actual situation of the project.

Table 3-2 Product Confirmation Form

Name	Model	Quantity
GPU Server	HGX H100	256
800G Optical Module	OSFP-SR8-800G	3072
400G Optical Module	OSFP-SR4-400G-FL	2048
MTP Patch Cord	OM4 50/125μm	4096
InfiniBand Switch	MQM9790-NS2F	96
Storage Switch	N8550-64C	18
Storage Management Switch	N8550-64C	2
Storage Management Switch	N8550-48BC	16
Device Management Switch	N5850-48S6Q	1
Device Management Switch	S5870-48T6S	9
Device Management Switch	S5870-48T4S	9
Firewall	NSG-5220	2

3.3 Software Preparation

Collect business information and determine the software version that needs to be deployed (the following software versions are for reference only, actual deployment is subject to customer needs)

Table 3-3 AI/ML Network Software Version Confirmation Form

Product	Software Name	Version Number	Notes
Server	Ubuntu	22.04	The server OS version is determined according to the customer's needs
NIC Firmware Version	Firmware	20.36.1010	Based on the official MLX instructions, for MLNX_OFED version 4.6 and above, CX5 NICs (corresponding to firmware FW16.25.1010 and above) or CX6 (firmware version 20.25.1010 and above)
White Box Switch System	PICOS	/	/
AmpCon management platform	AmpCon	/	/
Firewall License	/	/	/

3.4 Tools and Materials Preparation

Collect customer business information and determine the software version that needs to be deployed (the following software versions are for reference only, actual deployment is subject to customer needs)

Table 3-4 On-site Deployment Tools and Data Confirmation Form

Tool Name	Requirement	Function Description
PC/Laptop	Required	Laptop with Win7/Win10 installed, Chrome (70+) or Firefox (50+) browser is required to be installed inside the computer for deployment operations. The PC/laptop also requires at least one network card with a speed of >=1000Mbps for fast uploading of mirrors, tarballs and software.
Network Cable	Required	Prepare your own Gigabit network cable for PC/laptop to connect to the on-site network.
Serial Cable	Required	Used to configure switches and debug.
USB Flash Drive	Required	For software version upgrade.
EasyCheck V2 Endface Tester	Required	For endface inspection.
MTP Optical Power Meter	Optional	For optical power detection.
ATC-NE-M1 MTP Cleaning Pen	Required	For endface cleaning.
FS Box V4.0	Required	Used for module recoding
Power supply device (switch or test board)	Required	For powering products
Delivery Guidance Material	Required	For materials involved in delivery, please refer to the following manuals: "Version Release Notes", "Configuration Guide", "Command Reference", "Hardware Installation Manual", "Product User Manual"
Toolbox	Optional	For flexible use in on-site installations
Visual Fault Locator	Optional	Used to test for fiber damage

3.5 Structured Cabling System Planning

Integrated deployment scenarios and equipment lists, and completion of integrated cabling connection planning forms

3.5.1 Cabling Planning

Table 3-5 GPU Computing Network Cabling Planning

Part I: GPU Computing Network Cabling Planning			
1. Every 8 Leafs form 1 SU, called SU1. 2. SU1 Leaf numbers are SU1-1 to SU1-8, SU2 Leaf numbers are SU2-1 to SU2-8 and so on. 3. SU1 Server are GPU1-1 to GPU1-32, and SU2 Server are GPU2-1 to GPU2-32, and so on. 4. Spine numbers are Spine01-Spine32 respectively.			
Interconnect			
SU Group	**Server Side**	**Leaf Side**	**Spine Side**
SU1	GPU1-1 to GPU1-32	SU1-1	Spine01-Spine32
		SU1-2	
		SU1-3	
		SU1-4	
		SU1-5	
		SU1-6	
		SU1-7	
		SU1-8	

SU Group	Server Side	Leaf Side	Spine Side
SU2	GPU1-1 to GPU1-32	SU2-1	Spine01-Spine32
		SU2-2	
		SU2-3	
		SU2-4	
		SU2-5	
		SU2-6	
		SU2-7	
		SU2-8	
SU3	GPU1-1 to GPU1-32	SU3-1	Spine01-Spine32
		SU3-2	
		SU3-3	
		SU3-4	
		SU3-5	
		SU3-6	
		SU3-7	
		SU3-8	
SU4	GPU1-1 to GPU1-32	SU4-1	Spine01-Spine32
		SU4-2	
		SU4-3	
		SU4-4	
		SU4-5	
		SU4-6	

SU Group	Server Side	Leaf Side	Spine Side
SU4	GPU1-1 to GPU1-32	SU4-7	Spine01-Spine32
		SU4-8	
SU5	GPU1-1 to GPU1-32	SU5-1	Spine01-Spine32
		SU5-2	
		SU5-3	
		SU5-4	
		SU5-5	
		SU5-6	
		SU5-7	
		SU5-8	
SU6	GPU1-1 to GPU1-32	SU6-1	Spine01-Spine32
		SU6-2	
		SU6-3	
		SU6-4	
		SU6-5	
		SU6-6	
		SU6-7	
		SU6-8	
SU7	GPU1-1 to GPU1-32	SU7-1	Spine01-Spine32
		SU7-2	
		SU7-3	
		SU7-4	

SU Group	Server Side	Leaf Side	Spine Side
SU7	GPU1-1 to GPU1-32	SU7-5	Spine01-Spine32
		SU7-6	
		SU7-7	
		SU7-8	
SU8	GPU1-1 to GPU1-32	SU8-1	Spine01-Spine32
		SU8-2	
		SU8-3	
		SU8-4	
		SU8-5	
		SU8-6	
		SU8-7	
		SU8-8	

Table 3-6 Storage Network Cabling Planning

<table>
<tr><th colspan="4">Part 2: Storage Network Cabling Planning</th></tr>
<tr><td colspan="4">1. Every 2 Leafs form 1 SU, called SU1.
2. SU1 Leaf number is SU1-1 to SU1-2 respectively, SU2 Leaf number is SU2-1 to SU2-2 respectively and so on.
3. SU1 Server are GPU1-1 to GPU1-32. SU2 Server are GPU2-1 to GPU2-32.
4. Storage Server is Storage01-Storage38 respectively.
5. Other Servers are Other1-43, respectively.
6. Spine numbers are Spine01-Spine02, respectively.</td></tr>
<tr><td colspan="4">Interconnect</td></tr>
<tr><th>SU Group</th><th>Server Side</th><th>Leaf Side</th><th>Spine Side</th></tr>
<tr><td rowspan="2">SU1</td><td rowspan="2">GPU1-1 to GPU1-32</td><td>SU1-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU1-2</td></tr>
<tr><td rowspan="2">SU2</td><td rowspan="2">GPU2-1 to GPU2-32</td><td>SU2-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU2-2</td></tr>
<tr><td rowspan="2">SU3</td><td rowspan="2">GPU3-1 to GPU3-32</td><td>SU3-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU3-2</td></tr>
<tr><td rowspan="2">SU4</td><td rowspan="2">GPU4-1 to GPU4-32</td><td>SU4-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU4-2</td></tr>
<tr><td rowspan="2">SU5</td><td rowspan="2">GPU5-1 to GPU5-32</td><td>SU5-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU5-2</td></tr>
<tr><td rowspan="2">SU6</td><td rowspan="2">GPU6-1 to GPU6-32</td><td>SU6-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU6-2</td></tr>
<tr><td rowspan="2">SU7</td><td rowspan="2">GPU7-1 to GPU7-32</td><td>SU7-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU7-2</td></tr>
<tr><td rowspan="2">SU8</td><td rowspan="2">GPU8-1 to GPU8-32</td><td>SU8-1</td><td rowspan="2">Spine01-Spine02</td></tr>
<tr><td>SU8-2</td></tr>
</table>

SU Group	Server Side	Leaf Side	Spine Side
Storage SU	Storage01 to Storage38	Sto-1	Spine01-Spine02
		Sto-2	
Other SU	Other01 to Other43	Oth-1	Spine01-Spine02
		Oth-2	

Table 3-7 Storage Management Network Cabling Planning

Part 3: Storage Management Network Cabling Planning			
1. Every 2 Leaf units form 1 SU, called SU1. 2. SU1 Leaf number is SU1-1 to SU1-2 respectively, SU2 Leaf number is SU2-1 to SU2-2 respectively and so on. 3. SU1 Server are GPU1-1 to GPU1-32. SU2 Server are GPU2-1 to GPU2-32. 4. Storage Server is Storage01-Storage38 respectively. 5. Other Servers are Other1-43, respectively. 6. Spine numbers are Spine01-Spine02, respectively.			
Interconnect			
SU Group	Server Side	Leaf Side	Spine Side
SU1	GPU1-1 to GPU1-32	SU1-1	Spine01-Spine02
		SU1-2	
SU2	GPU2-1 to GPU2-32	SU2-1	Spine01-Spine02
		SU2-2	
SU3	GPU3-1 to GPU3-32	SU3-1	Spine01-Spine02
		SU3-2	
SU4	GPU4-1 to GPU4-32	SU4-1	Spine01-Spine02
		SU4-2	
SU5	GPU5-1 to GPU5-32	SU5-1	Spine01-Spine02
		SU5-2	

SU Group	Server Side	Leaf Side	Spine Side
SU6	GPU6-1 to GPU6-32	**SU6-1**	Spine01-Spine02
		SU6-2	
SU7	GPU7-1 to GPU7-32	**SU7-1**	Spine01-Spine02
		SU7-2	
SU8	GPU8-1 to GPU8-32	**SU8-1**	Spine01-Spine02
		SU8-2	
Storage SU	Storage01 to Storage38	**Sto-1**	Spine01-Spine02
		Sto-2	
Other SU	Other01 to Other43	**Oth-1**	Spine01-Spine02
		Oth-2	

Table 3-8 Storage Management Network Cabling Planning

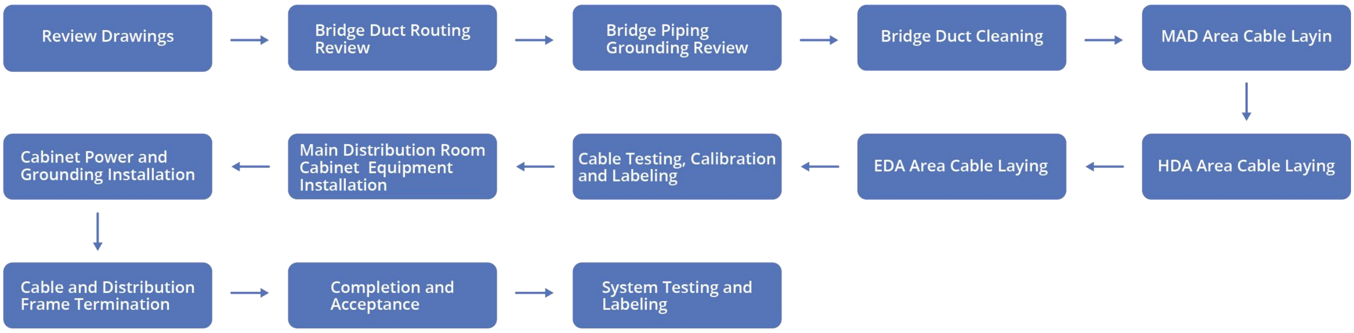
<table>
<tr><th colspan="4">Part 4: Equipment Management Network Cabling Planning</th></tr>
<tr><td colspan="4">1. Every 8 Leaf units form a SU, called SU1.
2. The Leaf number of SU1 is SU1, the number of SU2 is SU2, and so on. The Storage Server Leaf number is Sto Leaf, the Other Server Leaf number is Oth Leaf, and the network device Leaf number is Net Leaf.
3. SU1 Server is GPU1-1 to GPU1-32 respectively. SU2 Servers are GPU2-1 to GPU2-32 respectively.
4. The storage servers are Storage01-Storage38.
5. There is only 1 Spine, numbered IPMI Spine.</td></tr>
<tr><td colspan="4">Interconnect</td></tr>
<tr><th>SU Group</th><th>Server Side</th><th>Leaf Side</th><th>Spine Side</th></tr>
<tr><td>SU1</td><td>GPU1-1 to GPU1-32</td><td>SU1</td><td rowspan="4">IPMI Spine</td></tr>
<tr><td>SU2</td><td>GPU2-1 to GPU2-32</td><td>SU2</td></tr>
<tr><td>SU3</td><td>GPU3-1 to GPU3-32</td><td>SU3</td></tr>
<tr><td>SU4</td><td>GPU4-1 to GPU4-32</td><td>SU4</td></tr>
</table>

SU Group	Server Side	Leaf Side	Spine Side
SU5	GPU5-1 to GPU5-32	**SU5**	IPMI Spine
SU6	GPU6-1 to GPU6-32	**SU6**	
SU7	GPU7-1 to GPU7-32	**SU7**	
SU8	GPU8-1 to GPU8-32	**SU8**	

3.5.2 Construction Preparation

Familiarize with structured cabling construction requirements and complete integrated cabling construction preparation

3.5.2.1 Workflow



3.5.2.2 Pre-construction Survey

Before the installation project commences, an inspection of the building and environmental conditions of the telecommunication room and the data center must be conducted. The following conditions must be met before commencing work:

- The construction of the telecommunication room, data center, and work area civil engineering must be completed. The ground of the building should be flat and clean, with door height and width allowing for the movement of equipment and machinery without obstruction. Door locks and keys should be complete.
- The positions, quantities, and dimensions of cable trays and steel pipes must comply with the design and drawing requirements.
- Reliable construction power supply and grounding devices should be provided in the telecommunication room and data center.
- The area, environmental temperature, and humidity of the telecommunication room and data center must comply with design requirements and relevant regulations.

3.5.2.3 Pre-construction Material Inspection

Construction unit should be used in the project wire specifications, quantity, quality checks, no factory inspection materials or design does not match the use of the project.

3.5.2.4 Inspection Requirements for Cables

Pre-terminated patch cords should be checked first after opening to see if there is any damage to the appearance of the fiber optic cable and whether the encapsulation of the patch cord ends is good.

Length Testing: Each optical fiber should be tested, and the test results should be consistent. If there is a significant difference in the length of the optical fibers within the same cable reel, testing should be conducted from the other end or a through-light inspection should be performed to determine if there is any fiber breakage.

3.5.2.5 Inspection Requirements for Connectors

- The plastic material of terminal blocks, information outlets, and other connectors should be flame-retardant.
- The types, quantities, and positions of connectors used in fiber optic outlets should comply with the design specifications.
- The faceplates of fiber optic outlets should be clearly marked with "TX" for transmit and "RX" for receive.

3.5.2.6 Cable Installation Requirements

- Cables should be laid out naturally straight without twisting, looping, or joints, and should not be subjected to external pressure or damage.
- Both ends of the cables should be labeled with clear and properly aligned identifiers, including numbers. Labels should be made of durable materials.
- Cables should have sufficient slack to accommodate terminations, testing, and changes. For fiber optic cables, routing should include slack loops, with a recommended reserve length of 3 to 5 meters, or as specified in the design.
- When cables are vertically laid in cable trays, they should be secured to the tray's supports at the upper end and at intervals of 1.5 meters. For horizontal laying, cables should be fixed at the start, end, bends, and every 5 to 10 meters.
- Cable bundling should be done when laying cables on horizontal or vertical cable trays.
- Different types of cables, such as twisted pair cables, fiber optic cables, and other signal cables, should be bundled according to their category, quantity, diameter, and number of cores. The bundling spacing should not exceed 1.5 meters, and should be evenly distributed without being too tight or causing cable compression.

3.5.2.7 Bend Radius Requirements for Cables

The bending radius of MTP cables should be 7.5mm.

4. Hardware Installation

4.1 Unpacking and Acceptance

4.1.1 Equipment and Materials Entry Acceptance Form

Table 4-1 Acceptance Checklist Description

Acceptance Content	Acceptance Method	Standard
Product Quantity	Check against contract	Quantity matches
Outer Packaging	On-site observation	No damage to the outer packaging
Packing list, Equipment Installation and User Manual Certificate, Test Report	Check against contract and design documents, according to packing list	Complete information, fresh seal
Equipment Material, Model, Specifications	Check against contract and design documents	Comply with contract and design documents
Equipment Appearance	Unpacking and physical inspection of equipment	No damage, no rust
Accessories Inventory	Check relevant accessories against the contract and packing list	Quantity matches the packing list, no damage, no rust

4.2 Hardware Installation

4.2.1 Network Equipment: MQM9790 on the Shelf

4.2.1.1 Equipment Installation

- Tool-Less Rail Kit

1. 2x System Rails (A)
2. 2x Rack Rails (B)

- Rail Kit Parts

A	B
	

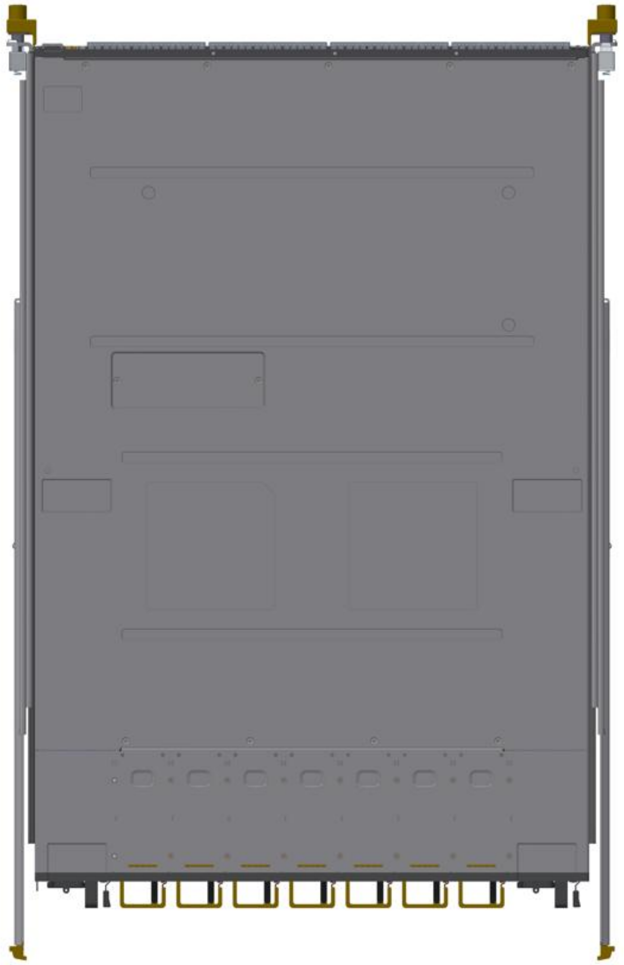
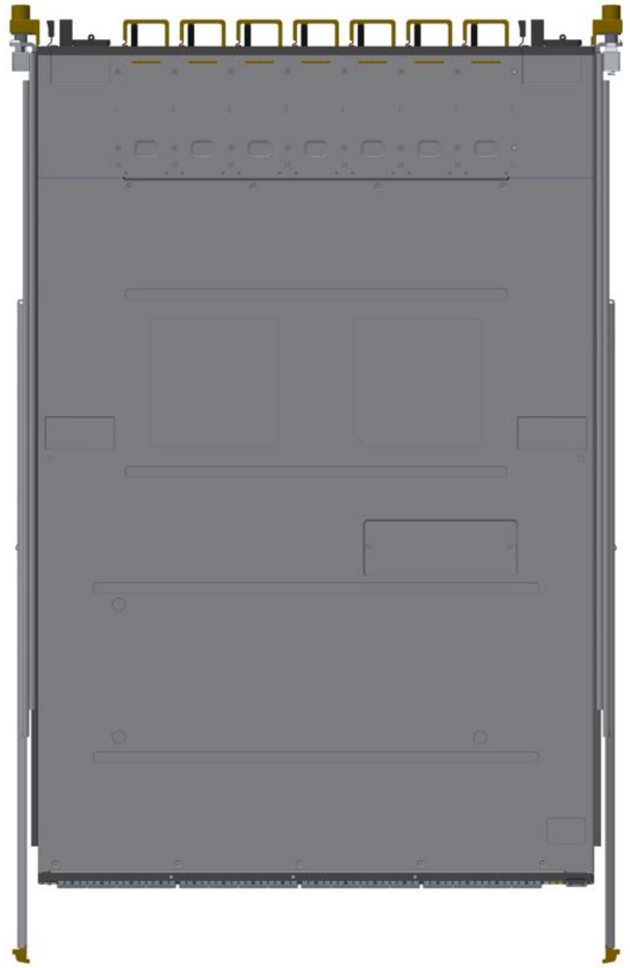
- Prerequisites

Before mounting the system to the rack, select the way you wish to place the system. Pay attention to the airflow within the rack cooling, connector and cablirng options.

While planning how to place the system, review the twoinstallation options shown in the table below, and consider the following points:

1. Make sure the system air flow is compatible with your installation selection. It is important to keep the airflow within the rack in the same direction.
2. Note that the part of the system to which you choose to attach the rails (the front panel direction, as demonstrated in Option 1 or the FRUs direction, as demonstrated in Option 2) will determine the system's adjustable side. The system's part to which the brackets are attached will be adjacent to the cabinet.
3. The FRUs, as well as high-speed and MNG cables, must be extracted for replacement as part of the switch service. Consider this when planning the switch installation.

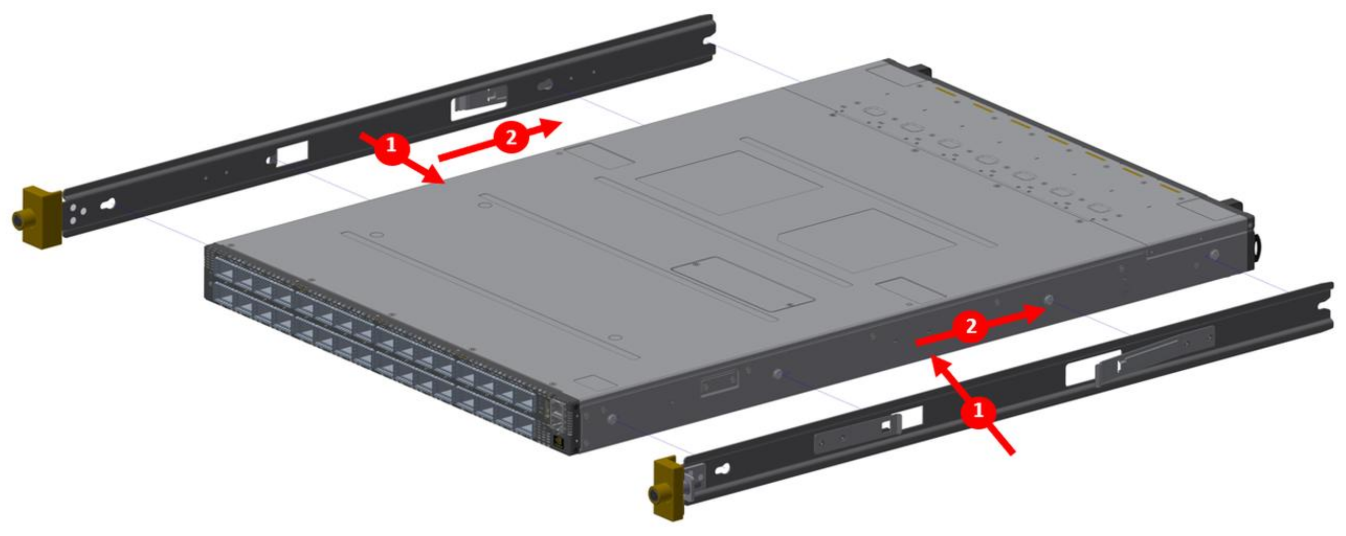
- Switch Rails Installation

Front Side (Ports)	Rear Side (FRUs)
	

Warning

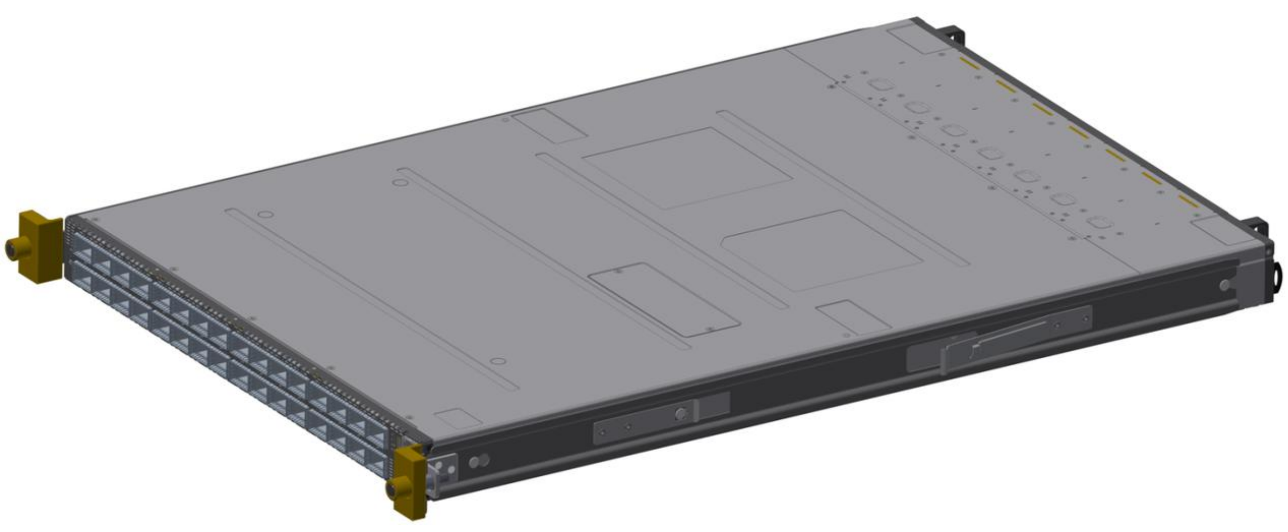
The following steps include illustrations that show front side (ports) installation, yet all instructions apply to all installation options.

1. Attach the left and right system rails (A) to the switch.



2. Secure the assembly by gently pushing the system chassis' pins through the slider key holes, until locking occurs

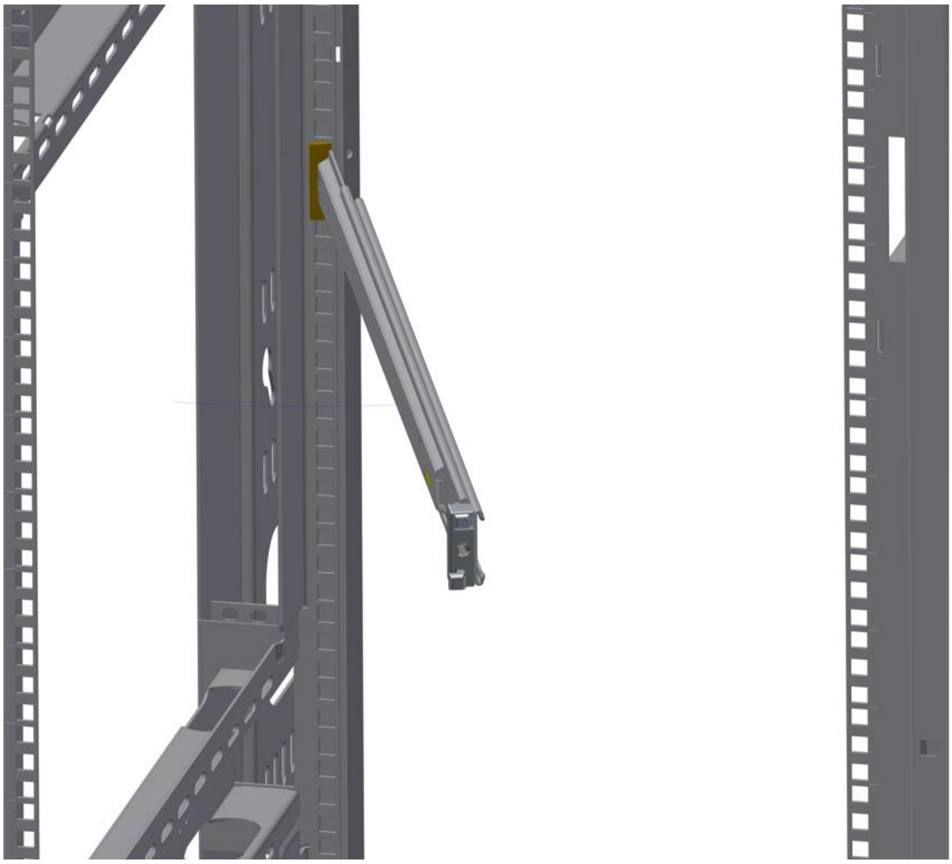
Securing the System in the Switch Rails (A)



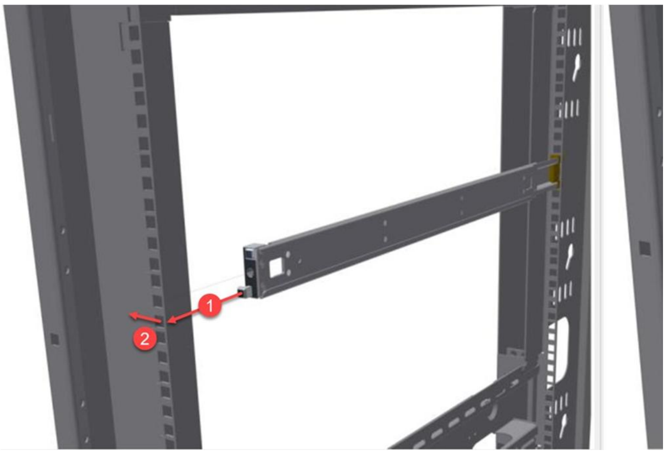
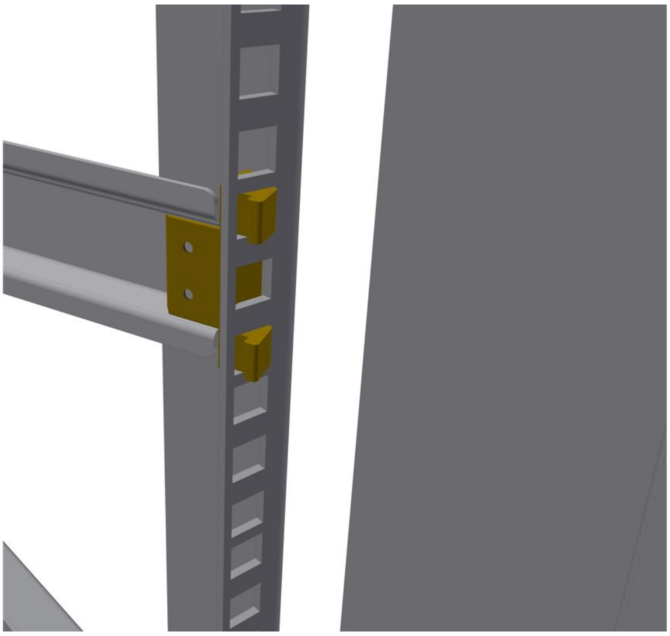
Chassis' Pins in the Rails' Slots Locking them in a Fixed Position

3. Mount both of the rack rails (B) into the rack by angularly inserting the brakes located at the rails edges into the designated slots in the rack unit, as shown in the following figure:

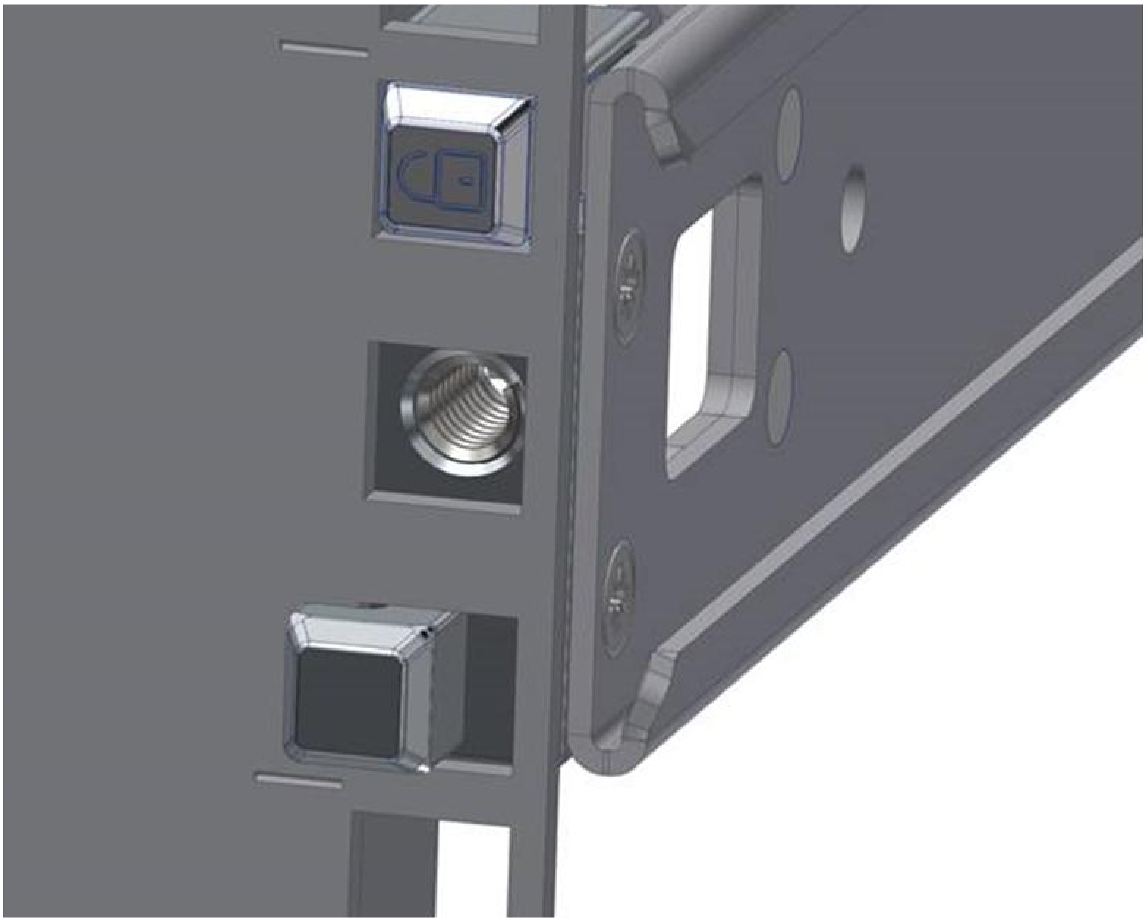
Inserting the Rack Rails (B)



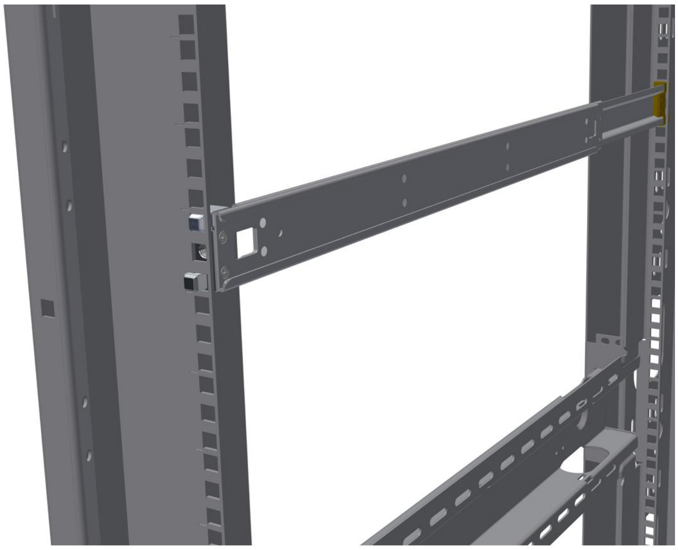
4. Align both rack rails (B) to sit horizontally in parallel to the rack assembly. By straightening the rails' angular position, their breaks will be caught and locked in the rack's slots.

Aligning the Rack Rails (B) Angular Position	The Breaks are Caught and Locked in the Rack's Slots
	

Rack Rails Fully Inserted and Locked in the Rack Assembly



5. Pull the rack rails' telescopic extensions all the way to the rack's opposite side, and insert the latches at the rails' free edges to the rack's slots. A click should be heard as the spring latches are fully inserted and locking occurs.

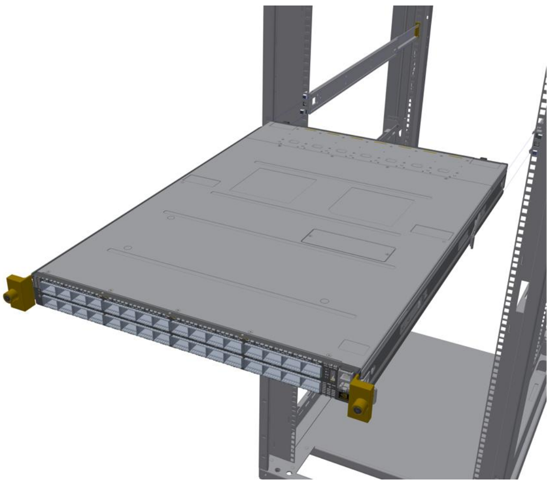
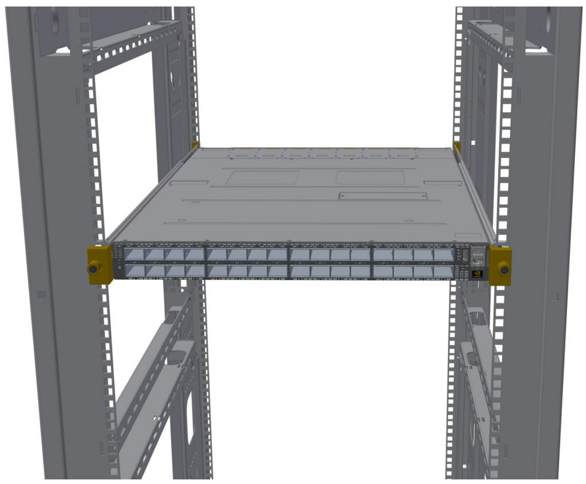
Pulling the Rails Telescopic Extensions	Inserting the Spring Latches to the Rack's Slots
	

To mount the system into the rack:

Note: At least two people are required to safely mount the system in the rack.

While your installation partner is supporting the system's weight, perform the following steps:

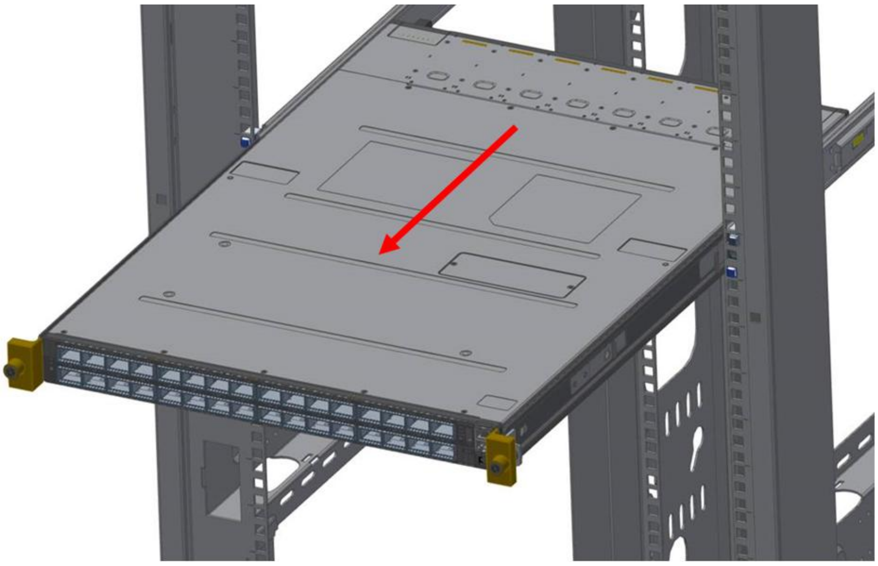
1. Slide the rails installed on the system into the channels in the rack rails. Push them forward until the locking mechanism is activated on both sides, and a click is heard.
2. Tighten the captive screws on both sides to further secure the system to the rack's posts.
3. Sliding the System's Rails (A) into the Rack Rails (B)

Sliding the System Rails (A) into the Rack Rails (B)	Tightening the Captive Screws
	

To mount the system into the rack:

1. Turn off the system and disconnect it from peripherals and from the electrical outlet. While your installation partner is supporting the system's weight:
2. Loosen the captive screws attaching the system's rails to the rack's posts.
3. Use two hands to pull the system out until the rails are stopped.

Pulling the System Out

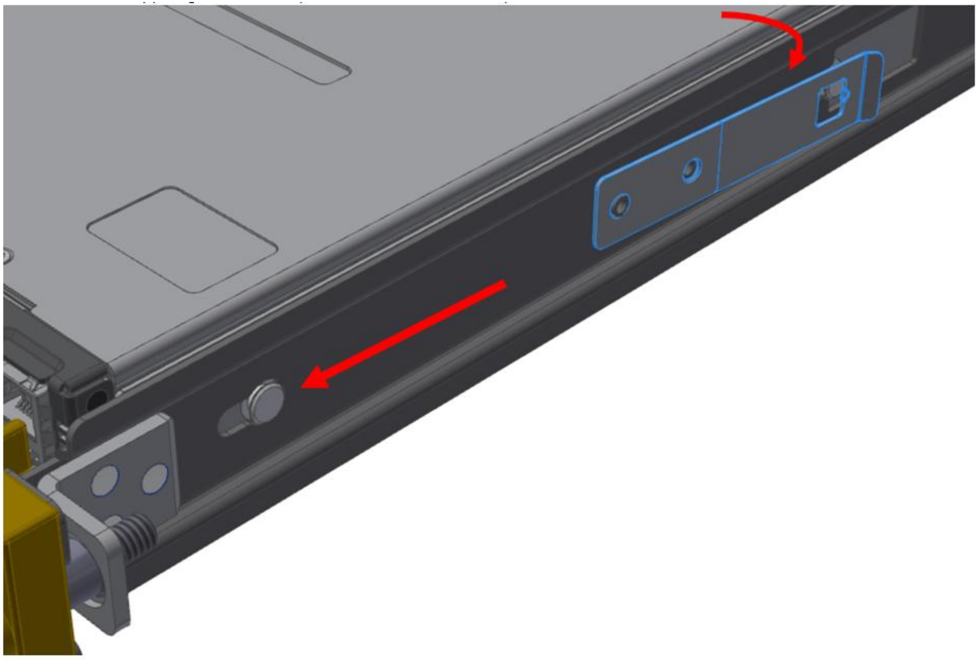


4. Press the spring latches on both sides of the rack, and continue to pull the system out until the rack rails are clear of the system's rails.

Pressing the Spring Latches on Both Sides

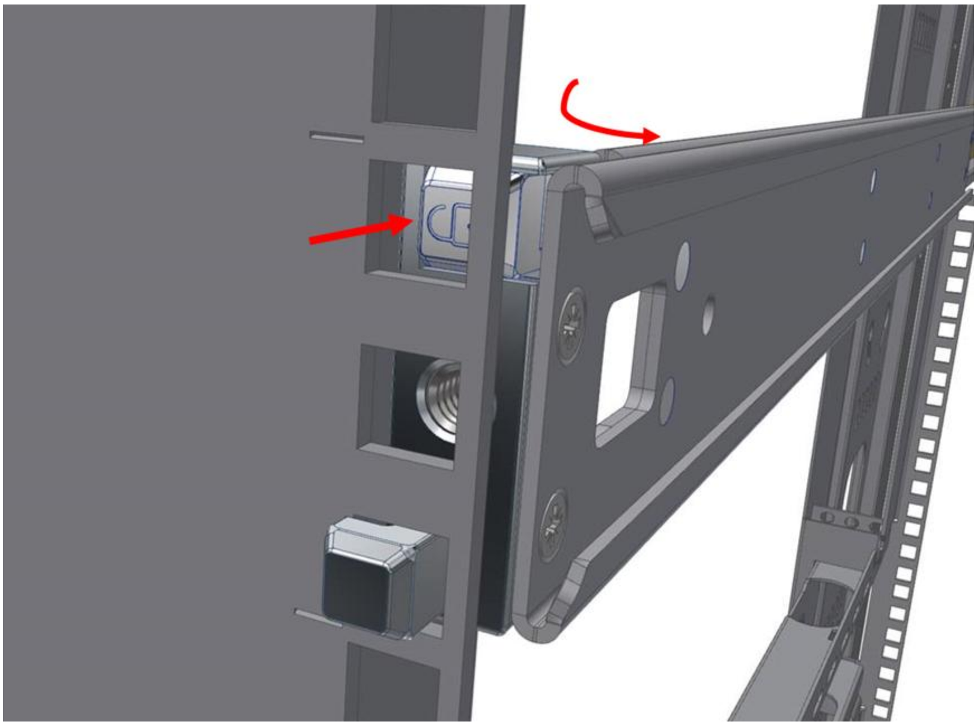
5. Remove the rails from the system. Release the metal latches and pull out the rails, so the system's pins will be removed out of the oval slots.

Removing the Rails from the System



6. Remove the rails from the rack by pressing the lock button, and pull the rails outside of the rack assembly.

Pressing the Lock Button to Remove the Rails from the Rack



4.2.1.2 Cable Installation

- Power Cable and Cable Retainer

In some switch models, the product's package includes cable retainers. It is highly recommended to use them in order to secure the power cables in place.

When installing retainers for the PSUs of the QM97x0 switch systems, please adhere to the following instructions:

1. Verify the integrity of the retainer assembly, as demonstrated in the below table: The snaps' push-pins must have visible edges with no broken or torn parts. The shoulders' pins should be in-tact and must not be bent inwards.

Proper Condition	Improper Condition
	

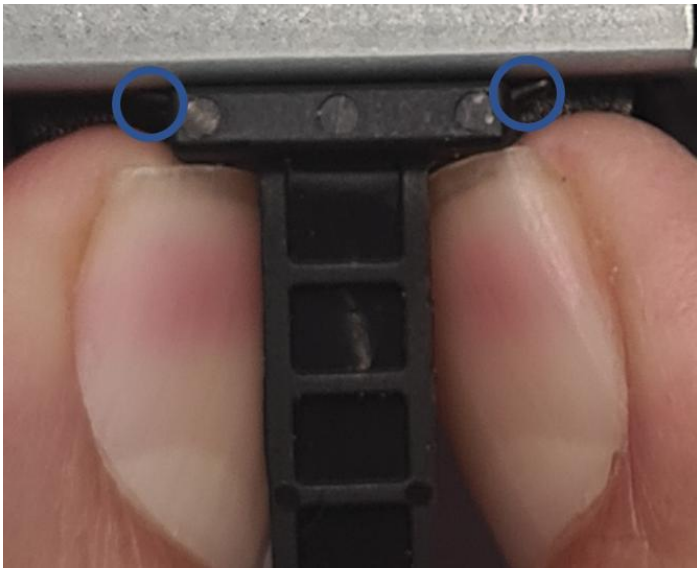
2. It is advised to place the PSU on a flat, stable surface. While you secure the PSU in place, use two thumbs to insert the retainer's two snaps into the designated holes located near the AC inlet. Make sure that the retainer's plastic loop is facing upwards, as demonstrated in the below table.

Warning

For demonstration purposes, the images in this document show C2P (Connector-toPower) airflow PSUs with red latches, yet the instructions apply to P2C (Power-toConnector) PSUs with blue latches as well.

Correct Insertion	Incorrect Insertion
	

3. Push the retainer until the shoulders' pins (in blue circles below) are open and aligned with the PSU front panel, as shown in the following table:

Fully Mated Retainer	
	

4. Make sure that the retainer is fully locked in place by gently attempting to pull it outwards.
5. Open the plastic loop and route the AC cord through it. Locate the loop over the AC cord, as shown in the following table, and fasten it tightly.

Proper Loop Placement	Improper Loop Placement
	

Warning

Each cable retainer can be used once only. Once the retainer has been fully inserted and the shoulders' pins have been adjusted, the retainer cannot be used again, and should be discarded if pulled out.

- **Port Cables**

All cables can be inserted or removed with the unit powered on.

To insert a cable, press the connector into the port receptacle until the connector is firmly seated. The LED indicator, corresponding to each data port, will light when the physical connection is established. When a logical connection is made, the relevant port LED will turn on.

To remove a cable, disengage the locks and slowly pull the connector away from the port receptacle. The LED indicator for that port will turn off when the cable is unseated.

For full cabling guidelines, ask your NVIDIA representative for a copy of NVIDIA Cable Management Guidelines and FAQs Application Note.

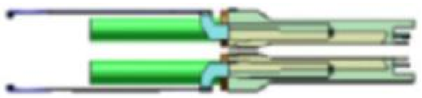
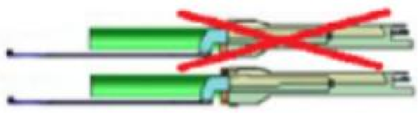
Important

Do not force the cable into the cage with more than 40 newtons/9.0 pounds/4kg force. Greater insertion force may cause damage to the cable or to the cage.

Warning

Unused OSFP cages must be closed with the thermal caps supplied with the system.

Cable Orientation

1. Splitter (Breakout) Cables and Adapters

In the QM9700 and QM9790 systems, a single OSFP cage contains 2 NDR ports, and a single NDR port (quad-lane) is divided into 2 dual-lane ports. This maximizes flexibility by enabling end users to use a combination of dual-lane and quad-lane interfaces according to the specific requirements of their network. For the systems splitting options, see QM9700/QM9790 Splitting Options below.

Splitting a port changes the notation of that port from x/y/z to x/y/z/i, with “x/y/z” indicating the previous notation of the port prior to the split, and “i” indicating the number of the resulting singlelane port (1,2). Each sub-physical port is then handled as an individual port. For example, splitting port 1/5/1 into 2 lanes results in ports 1/5/1/1 and 1/5/1/2. For full notation schematics, see Port Notation Schematics.

Warning

The following behavior should be expected when disconnecting a 1:2 splitter cable (from cages in both the upper and lower rows):

- When you disconnect a cable marked as “1”, the CLI <cage number>/1 will always go down, and the left LED of the cage will be turned off.
- When you disconnect the cable marked as “2”, the CLI <cage number>/2 will always go down, and the right LED of the cage will be turned off.

Breakout Cable Example

Warning

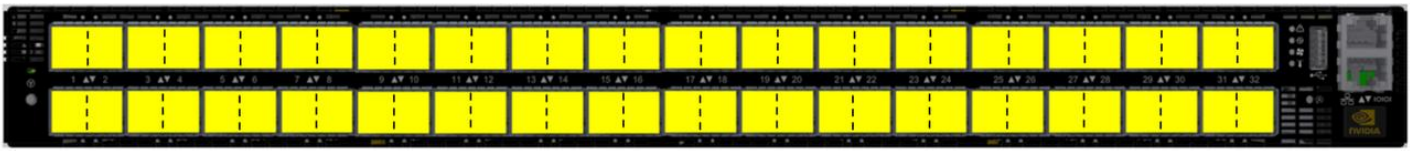
This feature is available only in Quantum/Quantum-2 based systems.

Tip

- Splitting the interface deletes all configuration on that interface.
- In order to be able to use this feature, the system profile command must be activated with split-ready configuration (cross-reference to system profile command).
- Changes will take effect after reset. In order to reset an unmanaged switch, please reboot the switch, or run flint -d <device> swreset .

For more information on how to change the system's profile to allow Split-Ready configuration, how to change the module type to a split mode, and how to unsplit a split port when using QM9700, please refer to the "InfiniBand Switching" chapter in the latest MLNX-OS® User Manual. For QM9790, please refer to latest NVIDIA Firmware Tools (MFT) Documentation.

QM9700/QM9790 Splitting Options



Split x2

All NDR ports are splittable. Each OSFP cage contains two NDR ports of 400G, and each NDR port can be split to two.

Port Notation Schematics

Two port notation profiles can be selected for the QM97x0 NDR switch systems. In both cases, each cage in the system's front panel holds two ports from the same ASIC, and the cage numbers are global:

1. ASIC/Cage/Port:

1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
1/1/1, 1/1/2	1/3/1, 1/3/2	1/5/1, 1/5/2	1/7/1, 1/7/2	1/9/1, 1/9/2	1/11/1, 1/11/2	1/13/1, 1/13/2	1/15/1, 1/15/2	1/17/1, 1/17/2	1/19/1, 1/19/2	1/21/1, 1/21/2	1/23/1, 1/23/2	1/25/1, 1/25/2	1/27/1, 1/27/2	1/29/1, 1/29/2	1/31/1, 1/31/2
1/2/1, 1/2/2	1/4/1, 1/4/2	1/6/1, 1/6/2	1/8/1, 1/8/2	1/10/1, 1/10/2	1/12/1, 1/12/2	1/14/1, 1/14/2	1/16/1, 1/16/2	1/18/1, 1/18/2	1/20/1, 1/20/2	1/22/1, 1/22/2	1/24/1, 1/24/2	1/26/1, 1/26/2	1/28/1, 1/28/2	1/30/1, 1/30/2	1/32/1, 1/32/2
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32

2. ASIC/Cage/Port/Split:

1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
1/1/1/1, 1/1/1/2 1/1/2/1, 1/1/2/2	1/3/1/1, 1/3/1/2 1/3/2/1, 1/3/2/2	1/5/1/1, 1/5/1/2 1/5/2/1, 1/5/2/2	1/7/1/1, 1/7/1/2 1/7/2/1, 1/7/2/2	1/9/1/1, 1/9/1/2 1/9/2/1, 1/9/2/2	1/11/1/1, 1/11/1/2 1/11/2/1, 1/11/2/2	1/13/1/1, 1/13/1/2 1/13/2/1, 1/13/2/2	1/15/1/1, 1/15/1/2 1/15/2/1, 1/15/2/2	1/17/1/1, 1/17/1/2 1/17/2/1, 1/17/2/2	1/19/1/1, 1/19/1/2 1/19/2/1, 1/19/2/2	1/21/1/1, 1/21/1/2 1/21/2/1, 1/21/2/2	1/23/1/1, 1/23/1/2 1/23/2/1, 1/23/2/2	1/25/1/1, 1/25/1/2 1/25/2/1, 1/25/2/2	1/27/1/1, 1/27/1/2 1/27/2/1, 1/27/2/2	1/29/1/1, 1/29/1/2 1/29/2/1, 1/29/2/2	1/31/1/1, 1/31/1/2 1/31/2/1, 1/31/2/2
1/2/1/1, 1/2/1/2 1/2/2/1, 1/2/2/2	1/4/1/1, 1/4/1/2 1/4/2/1, 1/4/2/2	1/6/1/1, 1/6/1/2 1/6/2/1, 1/6/2/2	1/8/1/1, 1/8/1/2 1/8/2/1, 1/8/2/2	1/10/1/1, 1/10/1/2 1/10/2/1, 1/10/2/2	1/12/1/1, 1/12/1/2 1/12/2/1, 1/12/2/2	1/14/1/1, 1/14/1/2 1/14/2/1, 1/14/2/2	1/16/1/1, 1/16/1/2 1/16/2/1, 1/16/2/2	1/18/1/1, 1/18/1/2 1/18/2/1, 1/18/2/2	1/20/1/1, 1/20/1/2 1/20/2/1, 1/20/2/2	1/22/1/1, 1/22/1/2 1/22/2/1, 1/22/2/2	1/24/1/1, 1/24/1/2 1/24/2/1, 1/24/2/2	1/26/1/1, 1/26/1/2 1/26/2/1, 1/26/2/2	1/28/1/1, 1/28/1/2 1/28/2/1, 1/28/2/2	1/30/1/1, 1/30/1/2 1/30/2/1, 1/30/2/2	1/32/1/1, 1/32/1/2 1/32/2/1, 1/32/2/2
2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32

Logical Port Numbering Schematic

Two profiles can be selected for the QM97x0 NDR switch systems. The first one defines the system as a pure 64-port NDR (32 cages) switch. The other profile permits any or all NDR ports to be split into two 2X (NDR200) ports. The following diagrams attempt to show how the logical ports map onto the physical NDR ports, as viewed by the IB tools (e.g. ibnetdiscover):

Switch Profile: Non-Splittable (Suitable for L2/Spine Switches)

Cage #	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
Logical 2X Port #	1	5	9	13	17	21	25	29	33	37	41	45	49	53	57	61
Logical 2X Port #	2	6	10	14	18	22	26	30	34	38	42	46	50	54	58	62
Logical 2X Port #	3	7	11	15	19	23	27	31	35	39	43	47	51	55	59	63
Logical 2X Port #	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64
Cage #	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32

Warning

The IB tools report 65 logical ports. Port 65 is an internal port used for the SHARP Aggregation Node when SHARP is enabled.

Switch Profile: Splittable

Cage #	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	31
Logical 2X Port # / Logical 2X Port #	1 2 3 4	9 10 11 12	17 18 19 20	25 26 27 28	33 34 35 36	41 42 43 44	49 50 51 52	57 58 59 60	65 66 67 68	73 74 75 76	81 82 83 84	89 90 91 92	97 98 99 100	105 106 107 108	113 114 115 116	121 122 123 124
Logical 2X Port # / Logical 2X Port #	7 8 5 6	15 16 13 14	23 24 21 22	31 32 29 30	39 40 37 38	47 48 45 46	55 56 53 54	63 64 61 62	71 72 69 70	79 80 77 78	87 88 85 86	95 96 93 94	103 104 101 102	111 112 109 110	119 120 117 118	127 128 125 126
Cage #	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32

Warning

The IB tools report 65 logical ports. Port 65 is an internal port used for the SHARP Aggregation Node when SHARP is enabled.

4.2.1.3 Power Supply

NVIDIA systems are equipped with two replaceable power supply units work in a redundant configuration. Either unit may be extracted without bringing down the system.

Important

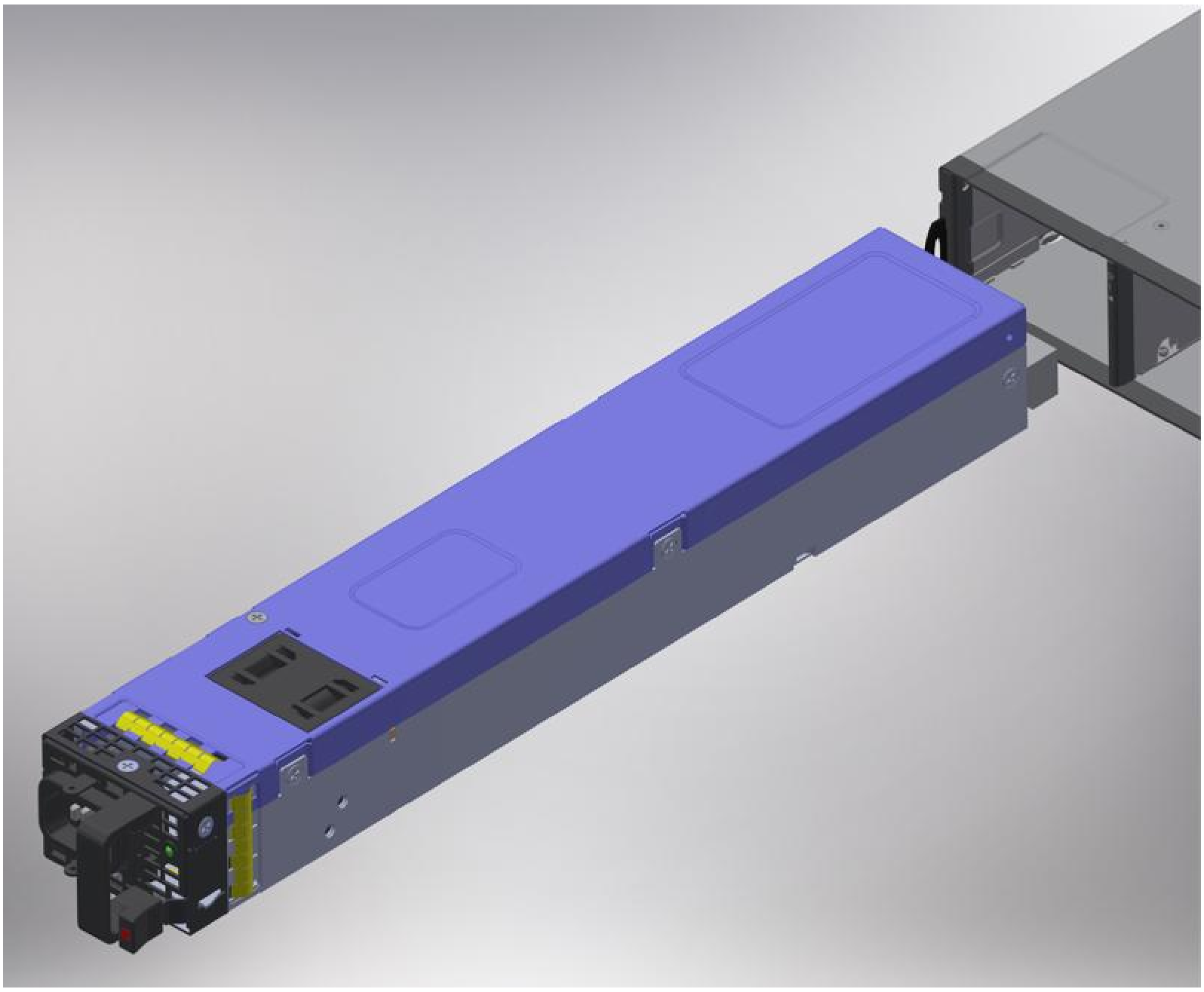
Make sure that the power supply unit that you are NOT replacing is showing all green, for both the power supply unit and the rear System Status LEDs.

Important

Power supply units have directional air flows similar to the fan module. The fan module airflow must coincide with the airflow of all of the power supply units. If the power supply unit airflow direction is different from the fan module airflow direction, the system's internal temperature will be affected. For power supply unit air flow direction, refer to Air Flow.

To extract a power supply unit:

1. Remove the power cord from the power supply unit.
2. Grasping the handle with your hand, push the latch release with your thumb while pulling the handle outward. As the power supply unit unseats, the power supply unit status LEDs will turn off.
3. Remove the power supply unit.



To insert a power supply unit:

1. Make sure the mating connector of the new unit is free of any dirt and/or obstacles.

> **Important**
>
> Do not attempt to insert a power supply unit with a power cord connected to it.

2. Insert the power supply unit by sliding it into the opening, until a slight resistance is felt.
3. Continue pressing the power supply unit until it seats completely. The latch will snap into place, confirming the proper installation.
4. Insert the power cord into the power supply connector.
5. Insert the other end of the power cord into an outlet of the correct voltage.

Important

The green power supply unit indicator should light. If it does not, repeat the whole procedure to extract the power supply unit and re-insert it.

4.2.1.4 Fans

The system can fully operate if one fan FRU is dysfunctional. Failure of more than one fan is not supported.

Important

Make sure that the fans have the air flow that matches the model number. An air flow opposite to the system design will cause the system to operate at a higher (less than optimal) temperature. For power supply unit air flow direction, refer to Air Flow.

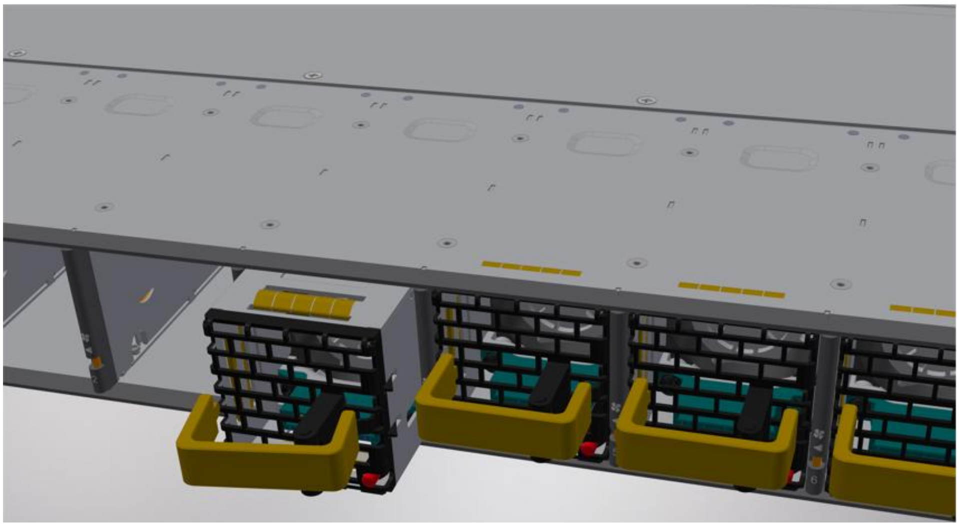
To extract a fan unit:

Warning

When replacing a faulty fan unit in an operational switch system, do not leave the slot unpopulated for more than 60 seconds.

1. Extract the fan by pulling the gold handle outwards. As the fan unit unseats, its status LEDs will turn off.
2. Remove the fan unit.

Fan Module Latches



To remove or replace a fan unit, gently pull out its handle while pushing the latch release with your index finger.

To insert a fan unit:

1. Make sure the mating connector of the new unit is free of any dirt and/or obstacles.
2. Insert the fan unit by sliding it into the opening until slight resistance is felt. Continue pressing the fan unit until it seats completely.

Important

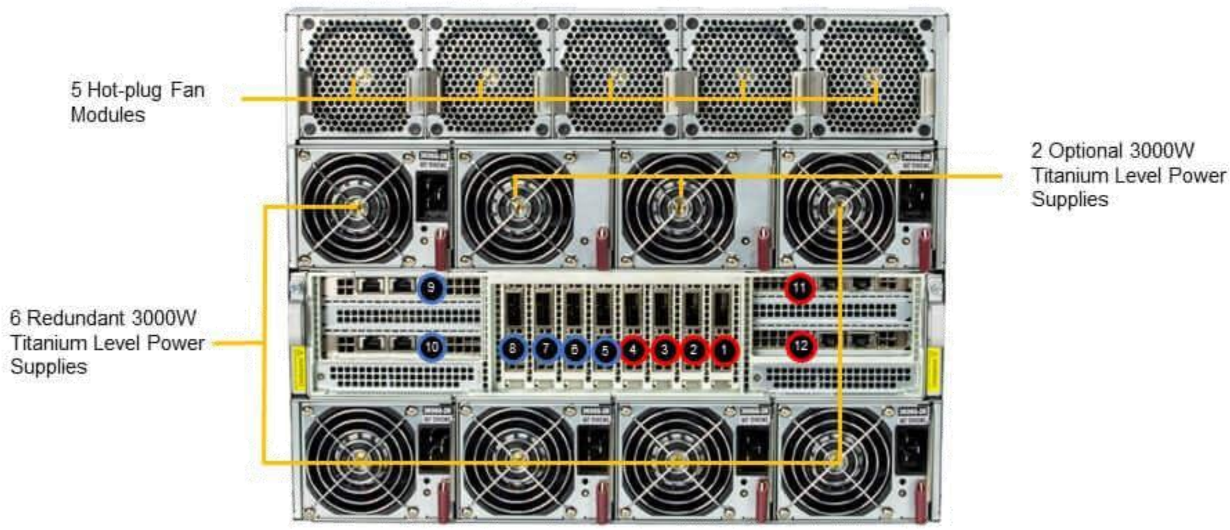
The green Fan Status LED should light. If not, extract the fan unit and reinsert it. After two unsuccessful attempts to install the fan unit, power off the system before attempting any system debug.

4.2.2 Server: H100 Server Network Card Installation

4.2.2.1 Pre-installation Preparation

1. H100 Server Rear View

(Rear View – System)

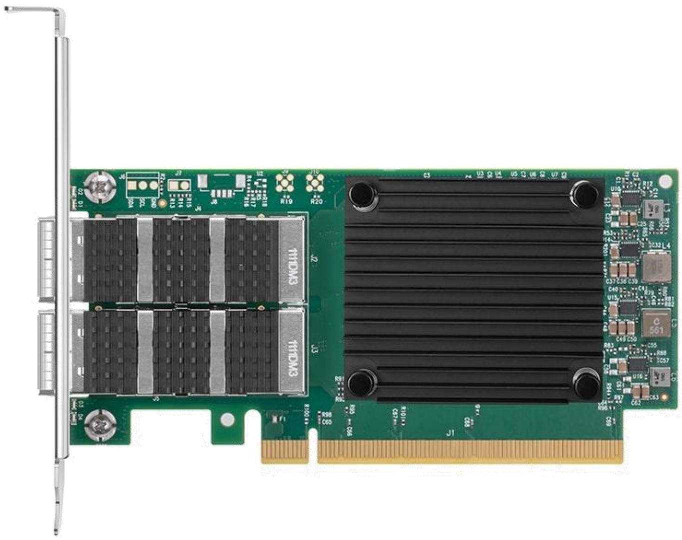


	Slot Description
1	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
2	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
3	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
4	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
5	PCIe 5.0 x16 (LP) from PLX switch linked to GPUs
6	PCIe 5.0 x16 (LP) from PLX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

	Slot 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

2. Picture of the MCX623106AN-CDAT network card to be installed (a total of 2 cards are required)

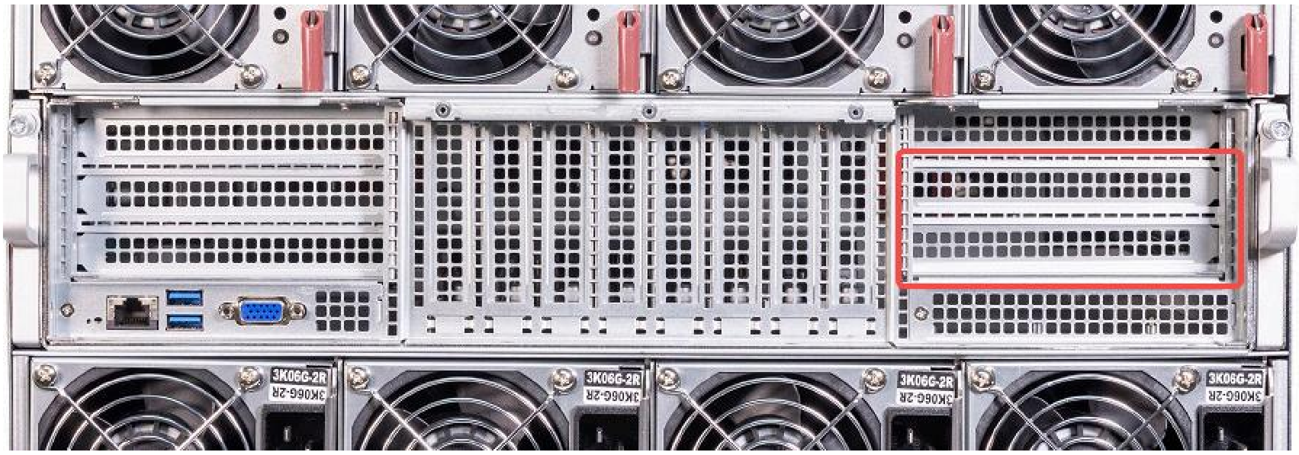


3. Disassembly Tools

A Phillips screwdriver is required.

4. Installation Location

Diagram of the installation location of two network cards: slot 10 & slot 12.

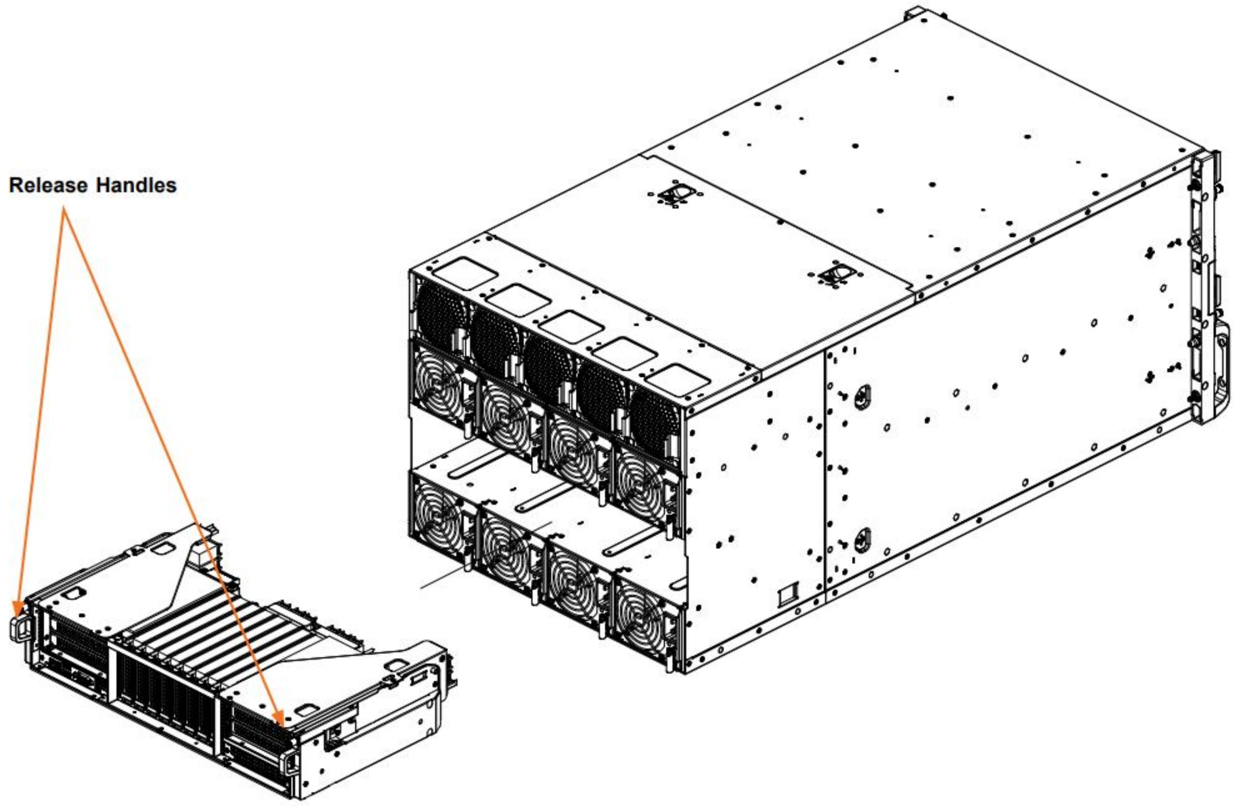


4.2.2.2 Installation Guide

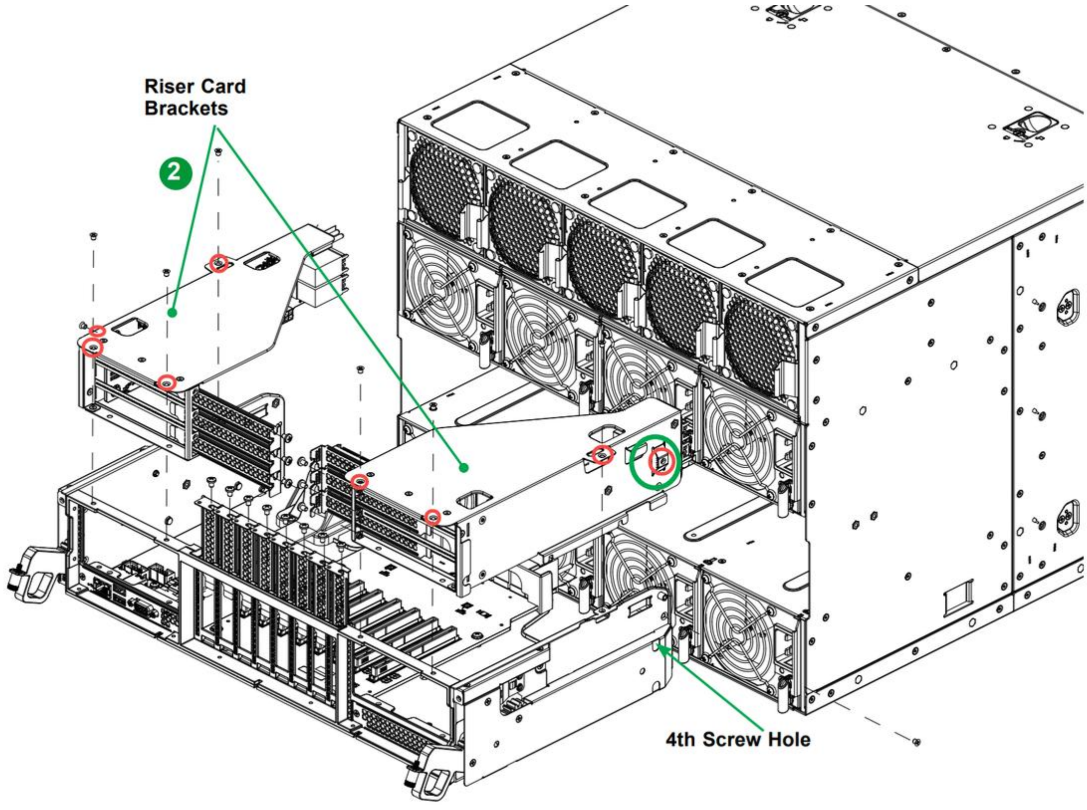
Note

The system consists of several trays. A release lever for each tray can be used to remove it from the chassis. To replace a tray, make sure the main handle is in the pull-down position. Once the tray has been fully inserted, lift and close the handle to lock the tray in place.

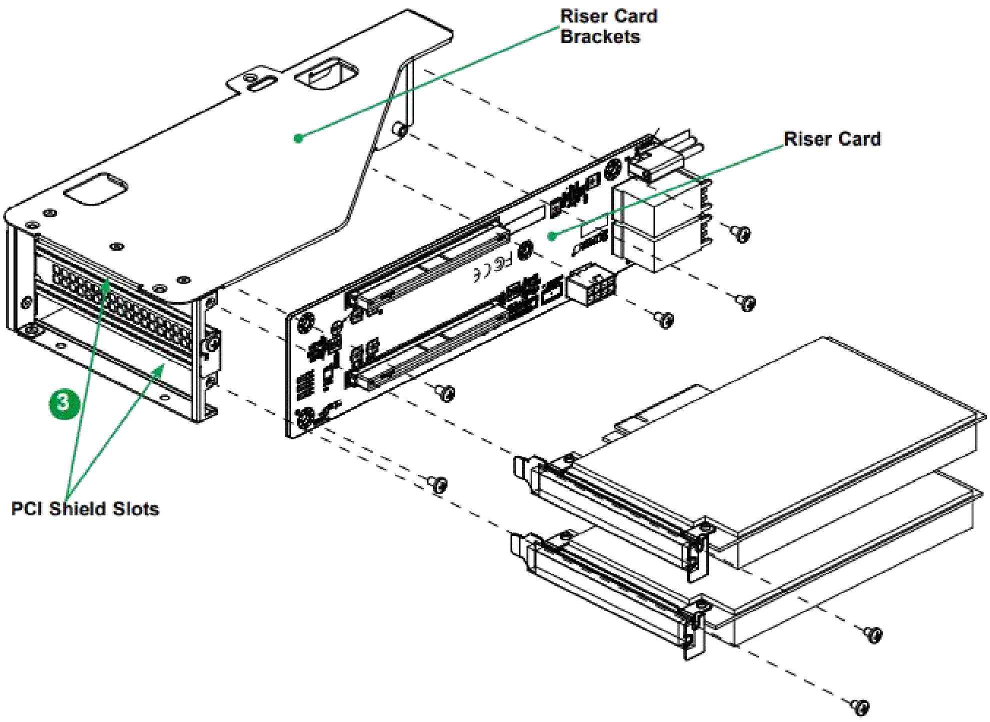
1. Turn off the system power, unscrew the screws at both ends, pull down the handle to remove the PCIe expansion tray, as shown in the figure.



2. Unscrew (four on each side) and remove the riser bracket to be used.



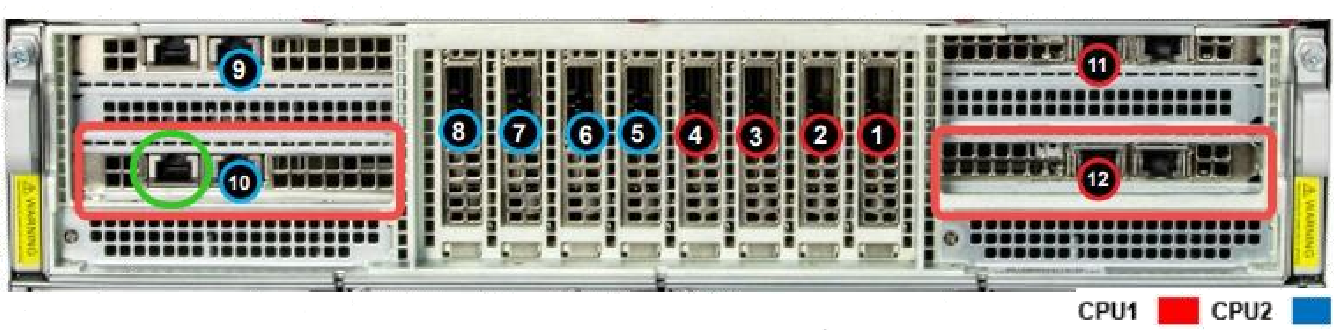
3. Remove the blank PCI shield covering the bracket slot. Slide the expansion card into the slot on the riser card while aligning the shield.



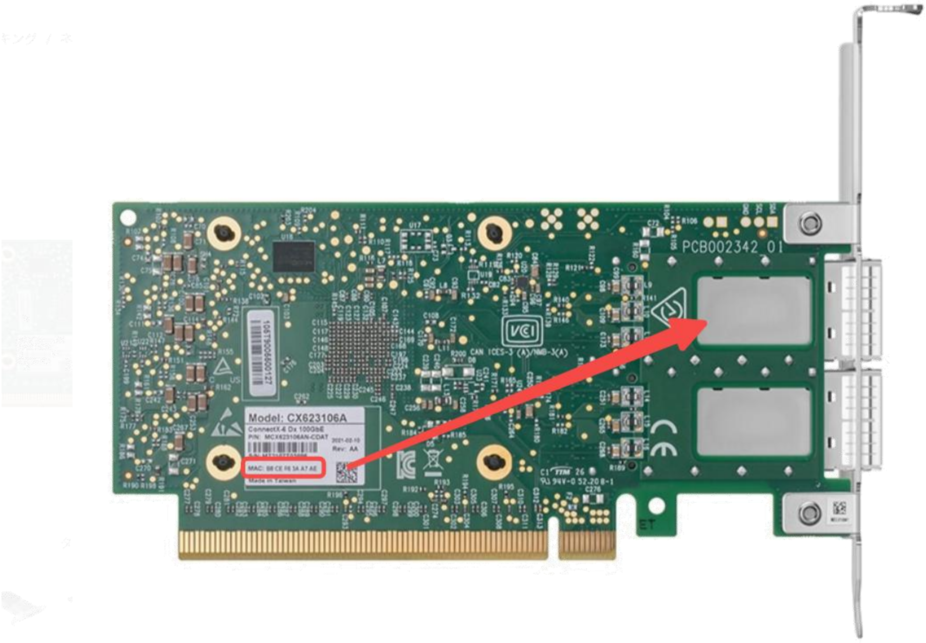
4. Push the PCIe expansion tray into the chassis.

4.2.2.3 Post-installation Verification

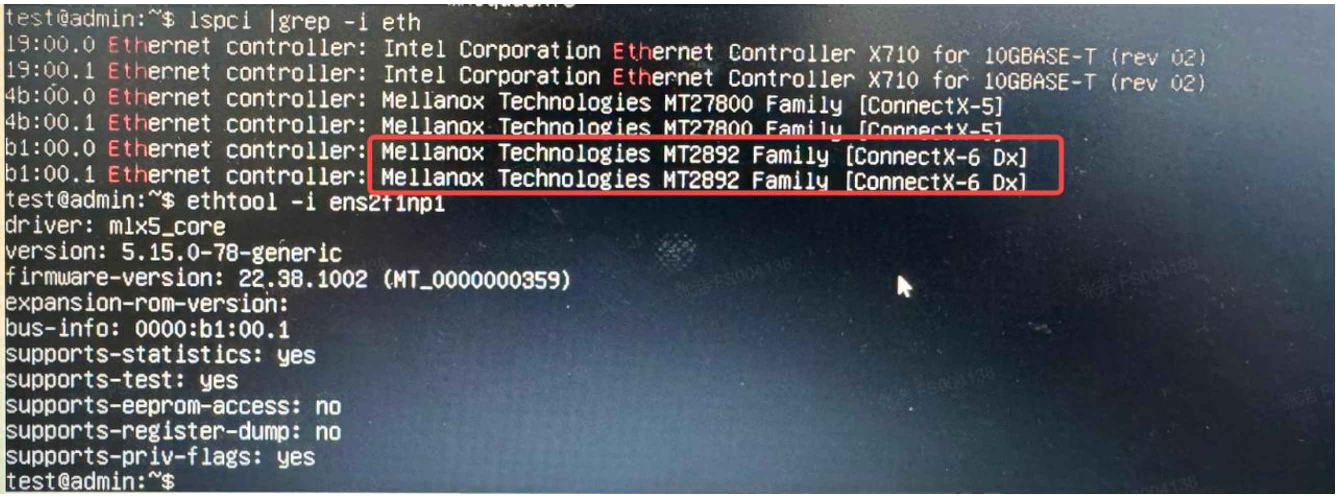
1. PXE network port wiring as shown in the figure of the green network port.



Need to record Slot10 NIC mac address for PXE installation system; the default MAC address is the uppermost port of the PCIe interface shown in the figure.



2. 4 records of [ConnectX-6 DX] were found under the installed system (Ubuntu22.04lts)



4.2.3 Optical Module: OSFP High-speed Module Installation

4.2.3.1 Pre-installation Preparation

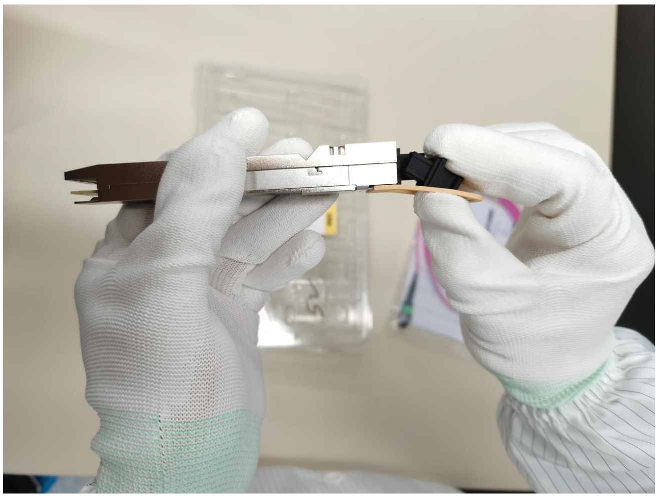
Attach the antistatic wrist strap to your wrist and an ESD grounding connector or bare metal surface on the chassis/wear antistatic gloves.

4.2.3.2 Unpacking and Inspection

Take out the OSFP module and MTP patch cord from the protective packaging and check whether they are broken or scratched.

4.2.3.3 Installation of OSFP Modules

Remove the optical port dust plugs from the OSFP module and MTP patch cord.



When installing the OSFP module, use your index finger and thumb to hold the left and right sides of the optical module with one hand, and use your other hand to push the rear of the fixed transceiver parallel to the device port until the transceiver contacts the device port socket and the locking device clicks. (If the transceiver does not slide in easily, make sure it is in the correct direction.)

- OSFP-SR8-800G module:

Matching with switches: install the optical module with the heat sink end facing up to the top row of ports on the switch (as shown in the left picture), and install the optical module with the heat sink end facing down to the bottom row of ports on the switch (as shown in the right picture) to ensure that the installation direction is correct.

- OSFP-SR4-400G module:

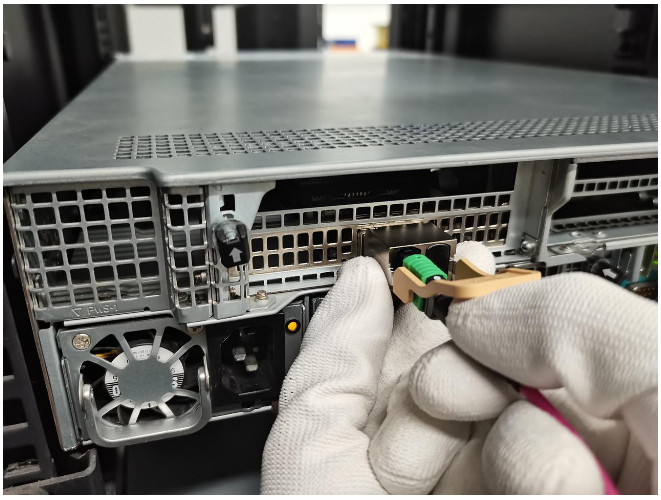
Matching with NIC: the labeled end of the optical module needs to be installed face down into the NIC port to ensure the correct orientation.


4.2.3.4 MTP Patch Cord Installation

To install a fiber optic patch cord with MTP, use your index finger and thumb to hold the top and bottom sides of the MTP connector, match the keys on the connector body to the keyway on the transceiver, and insert the MTP patch cord horizontally (not at a tilted angle) into the OSFP module until the locking mechanism makes a clicking sound. (Make sure the connector is clean)

- OSFP-SR8-800G module:

- OSFP-SR4-400G module:



4.2.3.5 Installation Completed Effect Diagram

When the link is fully established, the LED will be constantly green.

Note

In IB networks, MTP patch cords with APC interfaces are recommended.The APC end face has a unique slanted 8° design, an innovative design that helps to significantly improve the return loss at the connection. As the return loss increases, the optical signal reflected at the connection is dramatically reduced. This means that the reflected light will not significantly affect the signal source at the transmitter end and will not be incorrectly recognized by the receiver as a transmitter signal from the opposite end. As a result, this design is much better at reducing BER.

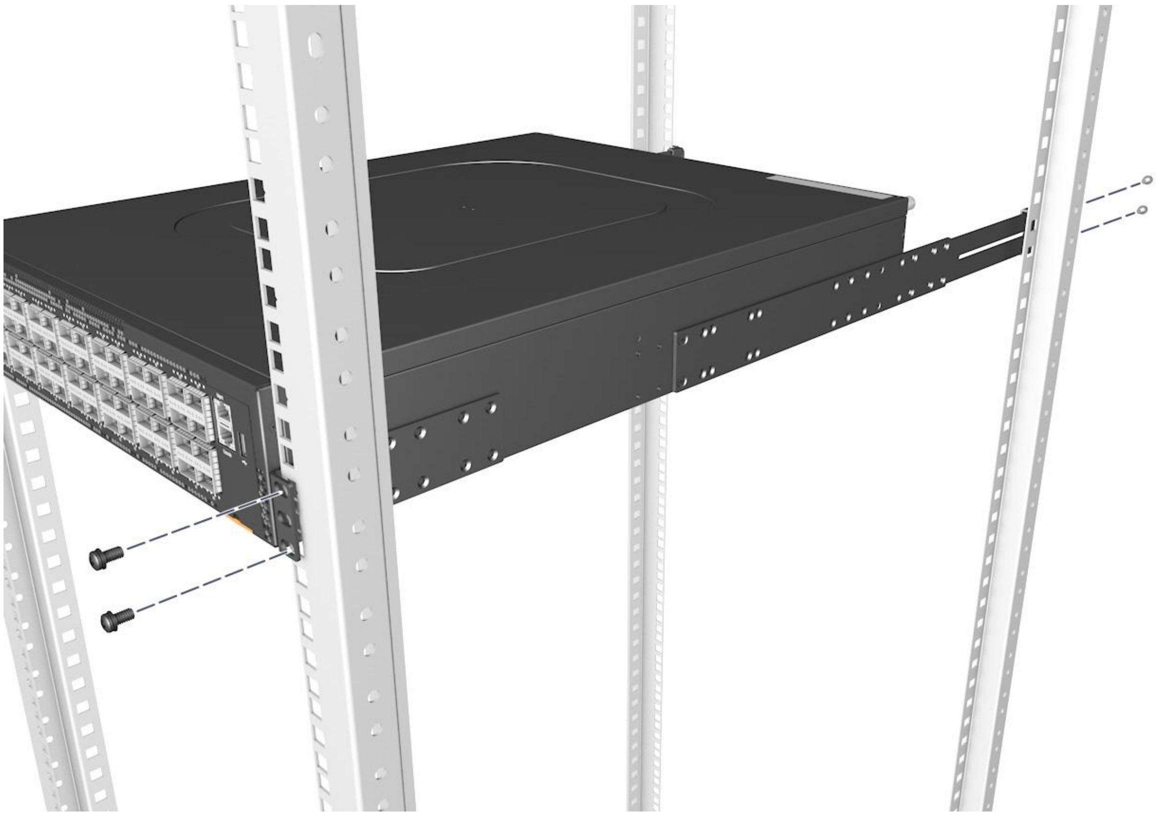
Note

Notably, the UPC fiber optic connector end face is an ultra-polished physical flat end face, and thus it is not recommended to directly mate UPC with APC. Such a combination can result in significant loss and potential damage to the fiber end faces. Based on these considerations, IB networks strongly recommend using APC connectors and avoiding direct interconnection between APC and UPC.

4.2.4 Network Equipment: N8550-64C Installation

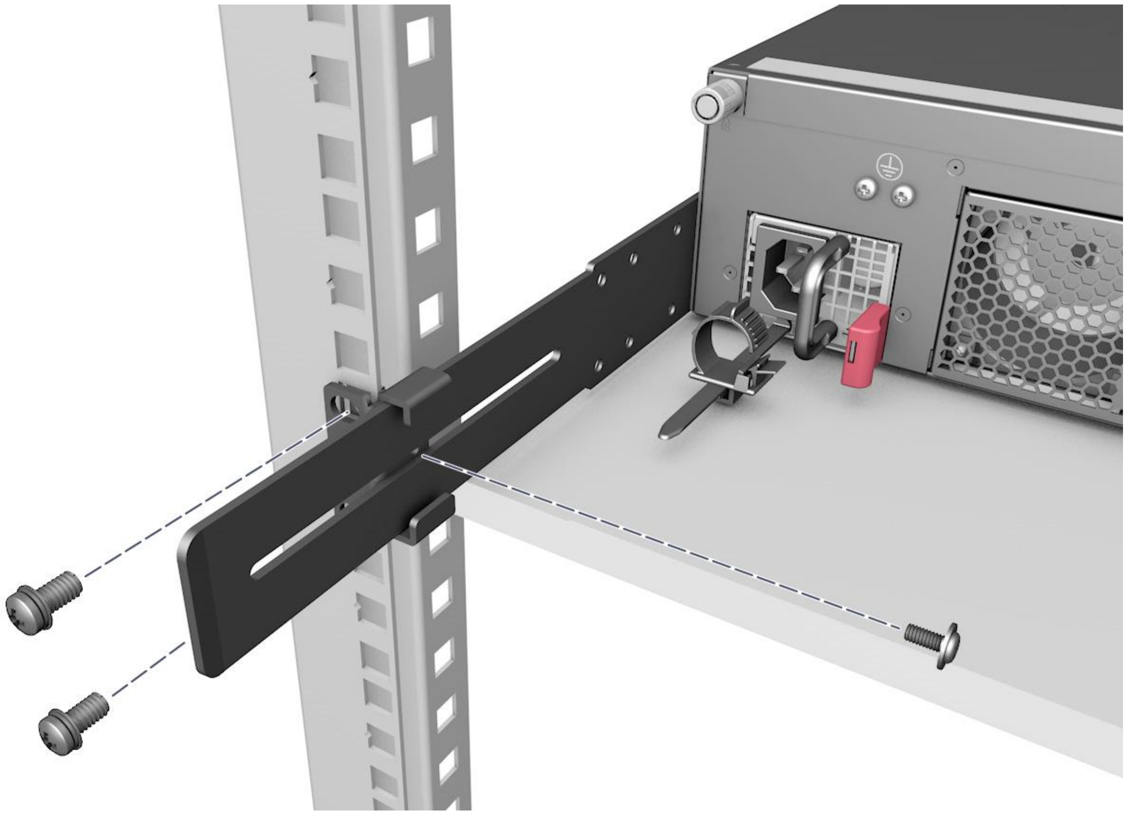
4.2.4.1 Rack Mounting

1. Install each front bracket and rear bracket onto the switch using the four included bracket screws.
2. Use two additional screws to secure each rear bracket at the midpoint on both sides of the switch.
3. Secure the switch in the rack with screws and cage nuts.



4.2.4.2 Adjust Rear-post Brackets

Use the provided locking device to lock the position of the rear bracket ear locking screws. You can also adjust the rear bracket ears to accommodate different rack depths from 56 cm to 75 cm.

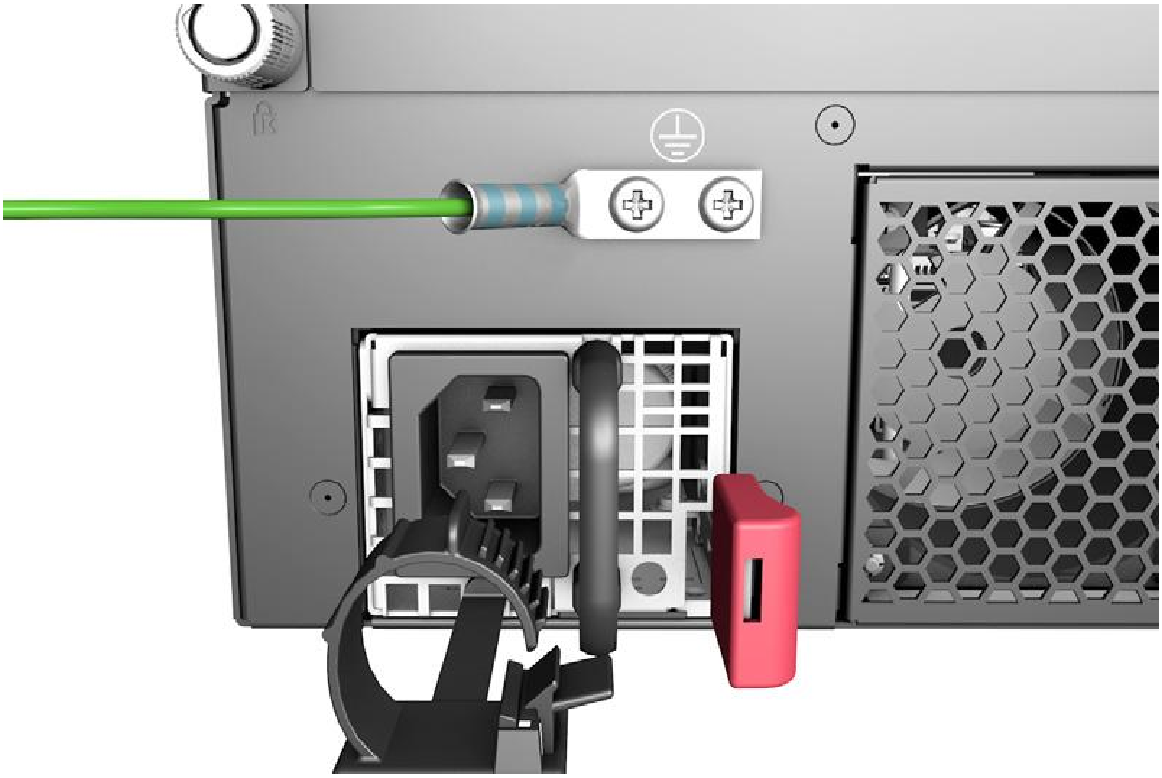


Note

In view of the diversity of the actual installation environment may exist, including ground level, space constraints and other factors, the installation and adjustment of the guide rail must be closely linked to the site-specific conditions for flexible implementation.

4.2.4.3 Grounding the Switch

1. Ensure the rack is properly grounded and complies with ETSI ETS 300 253. Verify that there is a good electrical connection to the rack ground point (no paint or isolating finishes.
2. Connect a No. 10 AWG ground wire (not included) to the ground point on the rear panel of the switch, and then connect the other end of the wire to the rack ground point.

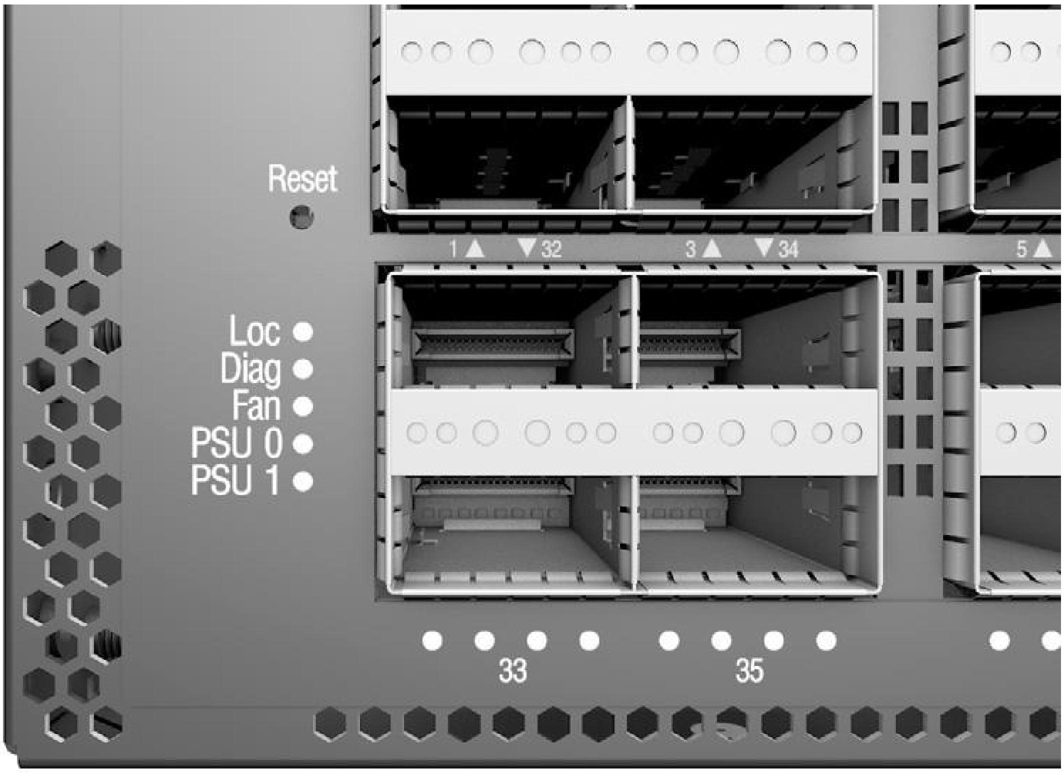


4.2.4.4 Connecting the Power

1. Install one or two AC or DC PSUs on the switch. The switch supports up to two PSUs, and their airflow direction must match the airflow direction of the installed fan tray.
2. Connect external AC power to the PSUs.

4.2.4.5 Verifying Switch Operation

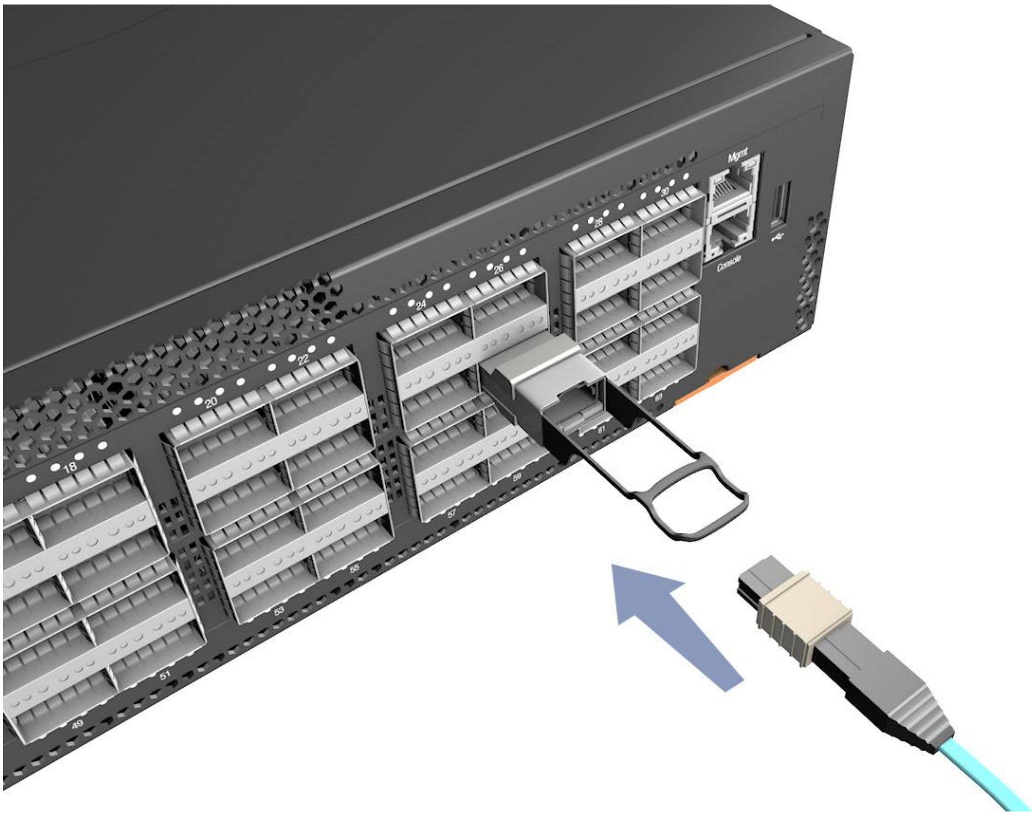
Verify basic operation of the switch by checking the system LEDs. For normal operation, the PSU1/PSU2, Diagnostic, and Fan LEDs should all be green.



4.2.4.6 Connecting Network Cable

1. Connect a 100-ohm, Category 5.5e or better twisted-pair cable to the RJ-45 management port, using Category 5e or better twisted-pair cable.
2. Connect the DAC cable to the QSFP28 slot. Or, first install the DAC cable to the QSFP28 slot. The transceiver ports support the following transceivers:

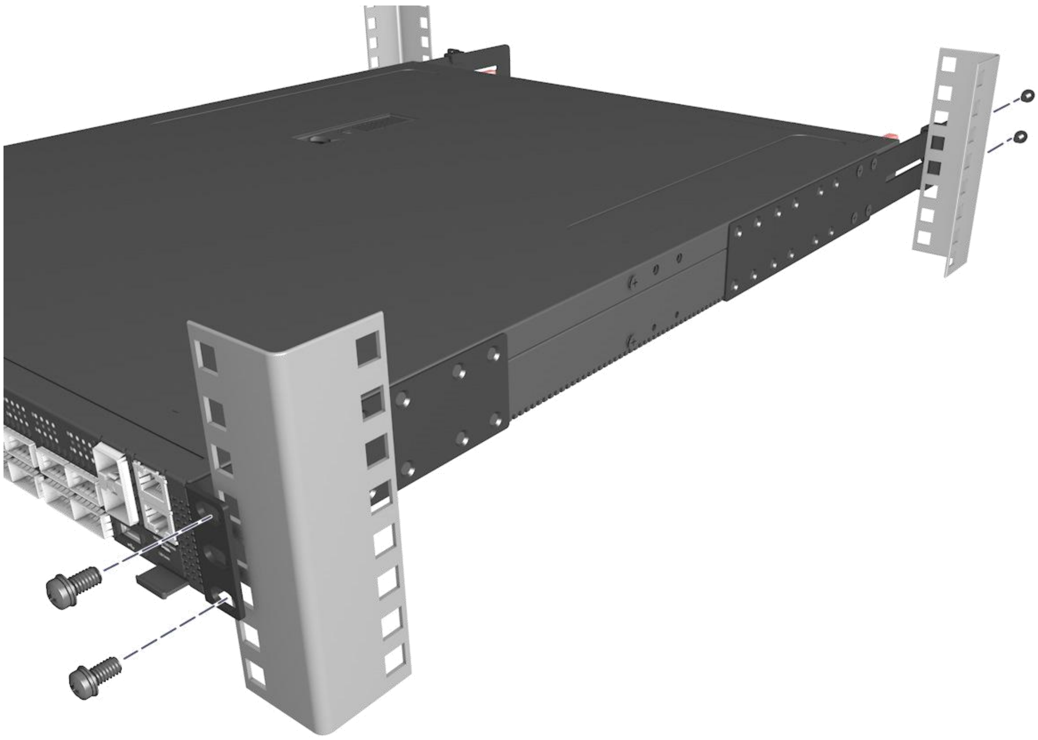
- 100GBASE-CR4、AOC、SR4、LR4 和 PSM4
- 40GBASE-CR4、SR4 和 LR4



4.2.5 Network Equipment: N8550-48B8C Installation

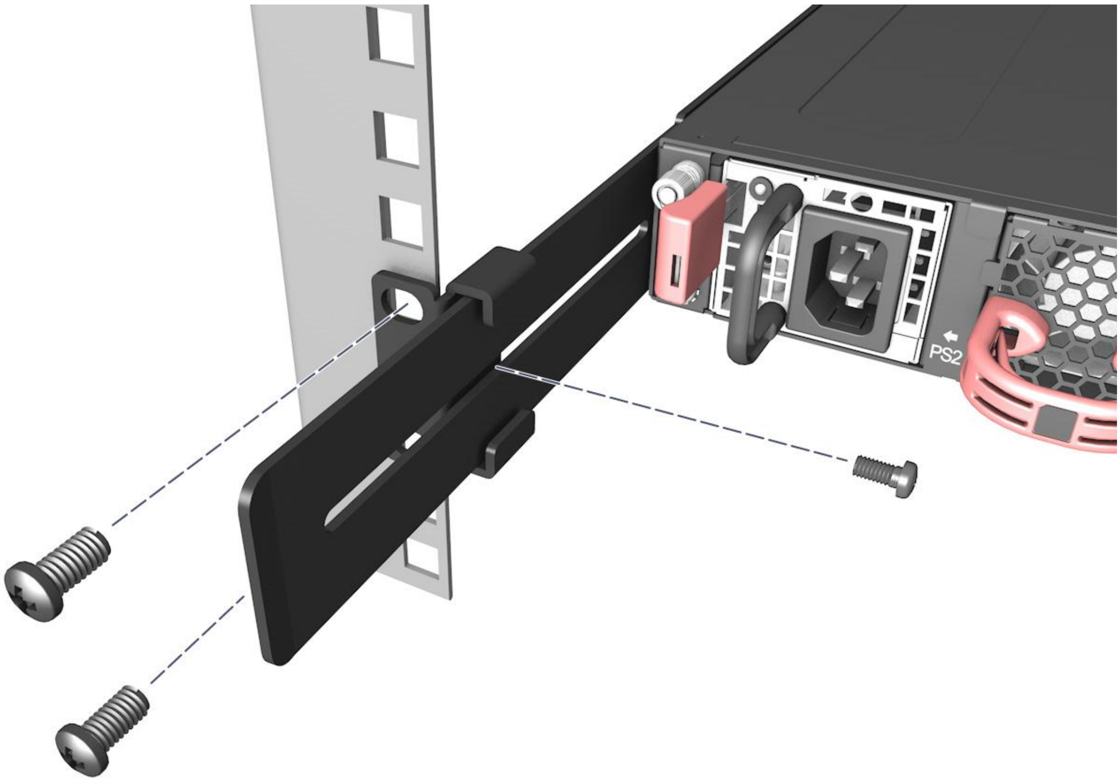
4.2.5.1 Rack Mounting

1. Install each front bracket and rear bracket onto the switch using the four included bracket screws.
2. Use two additional screws to secure each rear bracket at the midpoint on both sides of the switch.
3. Secure the switch in the rack with screws and cage nuts.



4.2.5.2 Adjust Rear-post Brackets

Use the provided locking device to lock the position of the rear bracket ear locking screws. You can also adjust the rear bracket ears to accommodate different rack depths from 56 cm to 75 cm.

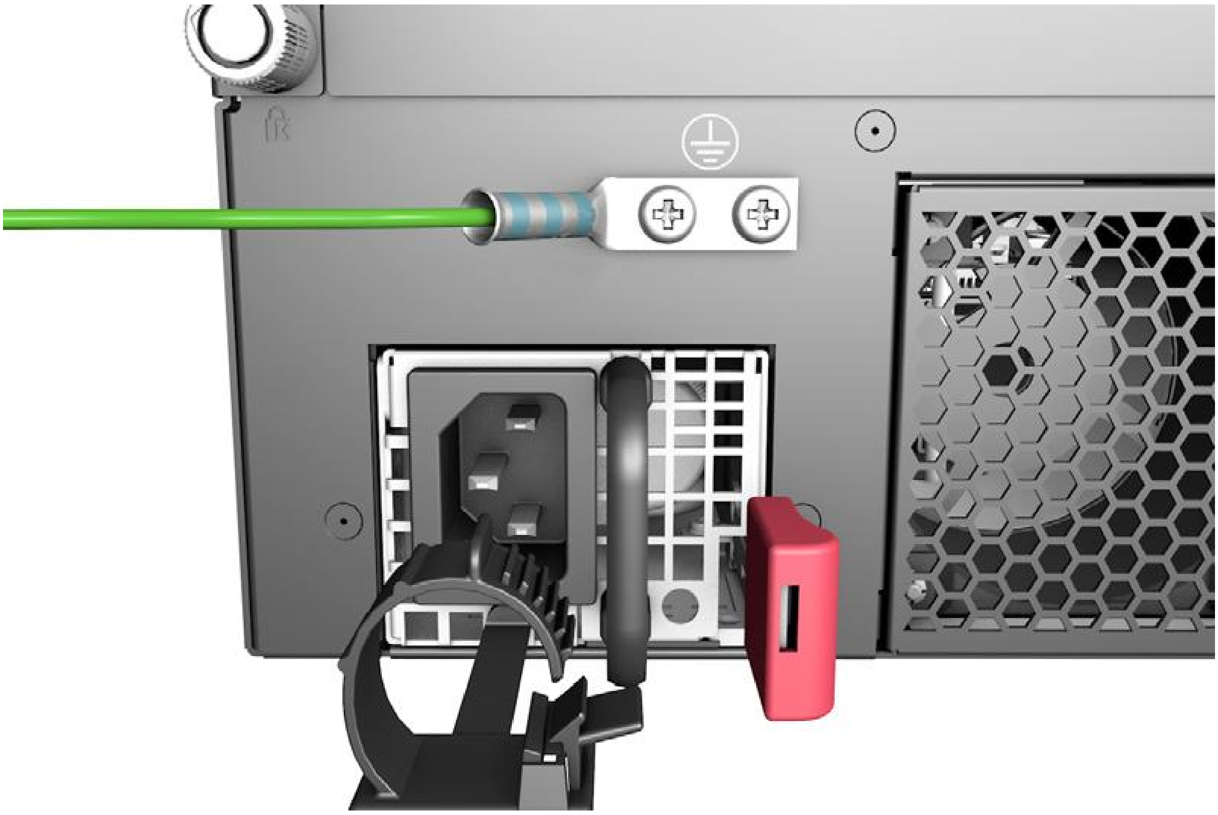


Note

The guide rail can be adjusted according to the actual situation on site to meet the installation requirements.

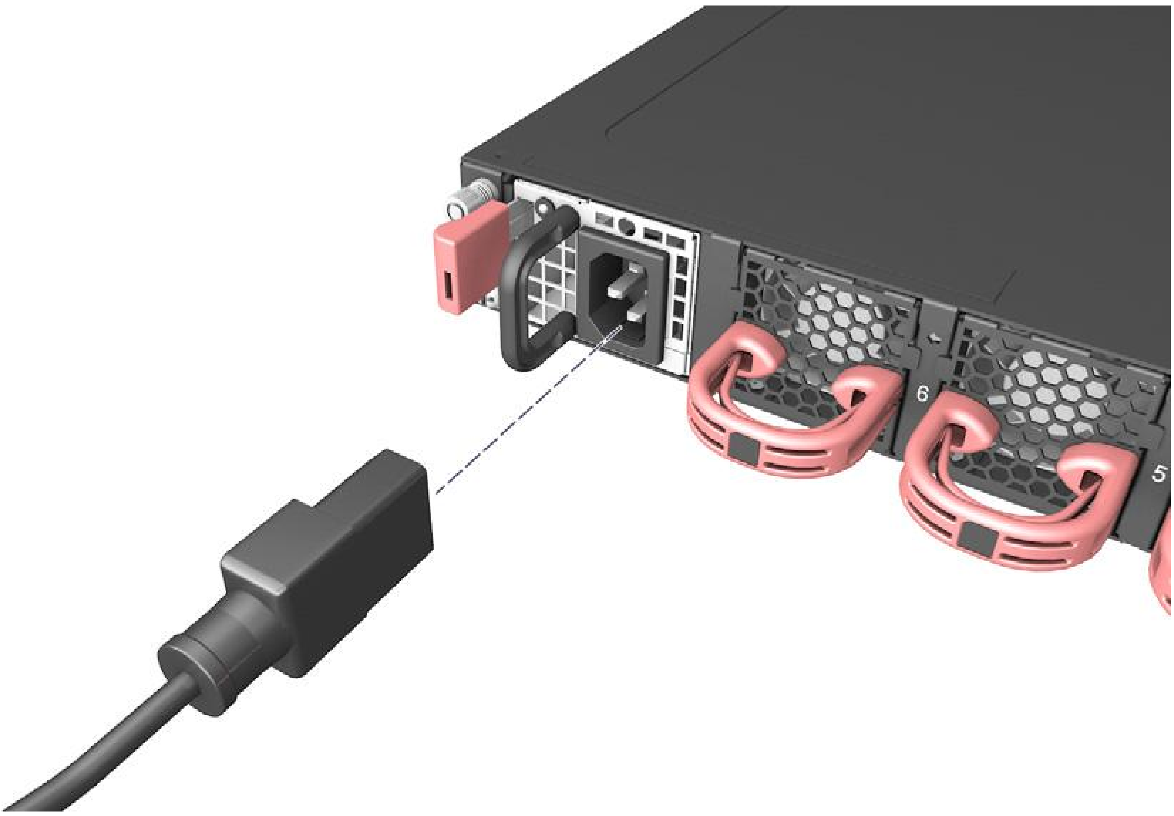
4.2.5.3 Grounding the Switch

1. Ensure the rack is properly grounded and complies with ETSI ETS 300 253. Verify that there is a good electrical connection to the rack ground point (no paint or isolating finishes).
2. Connect a No. 10 AWG ground wire (not included) to the ground point on the rear panel of the switch, and then connect the other end of the wire to the rack ground point.



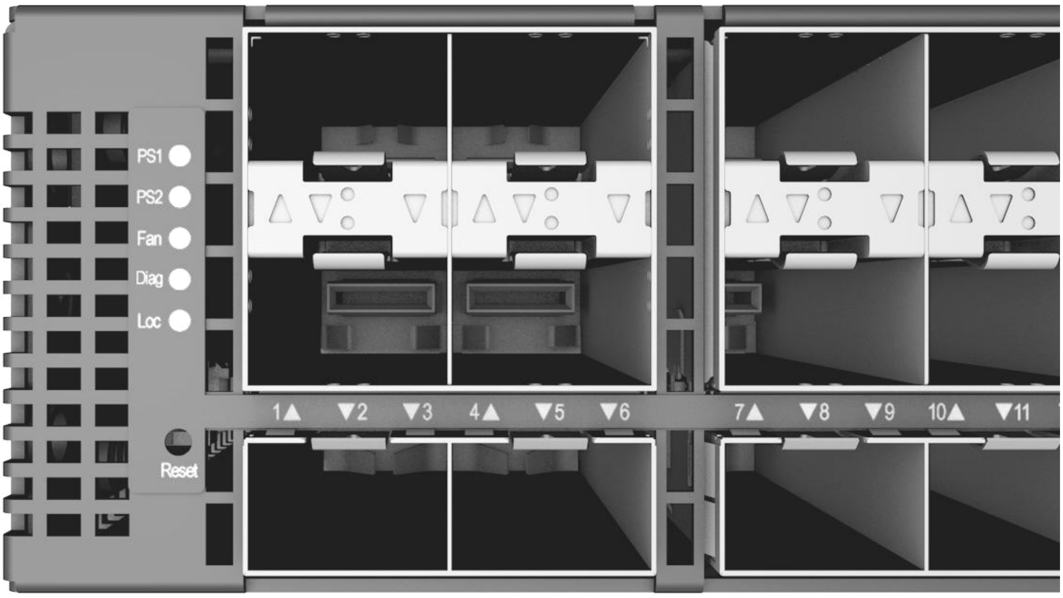
4.2.5.4 Connecting the Power

1. Install one or two AC or DC PSUs on the switch. The switch supports up to two PSUs, and their airflow direction must match the airflow direction of the installed fan tray.
2. Connect external DC power to the PSUs.



4.2.5.5 Verifying Switch Operation

Verify basic operation of the switch by checking the system LEDs. For normal operation, the PSU1/PSU2, Diagnostic, and Fan LEDs should all be green.



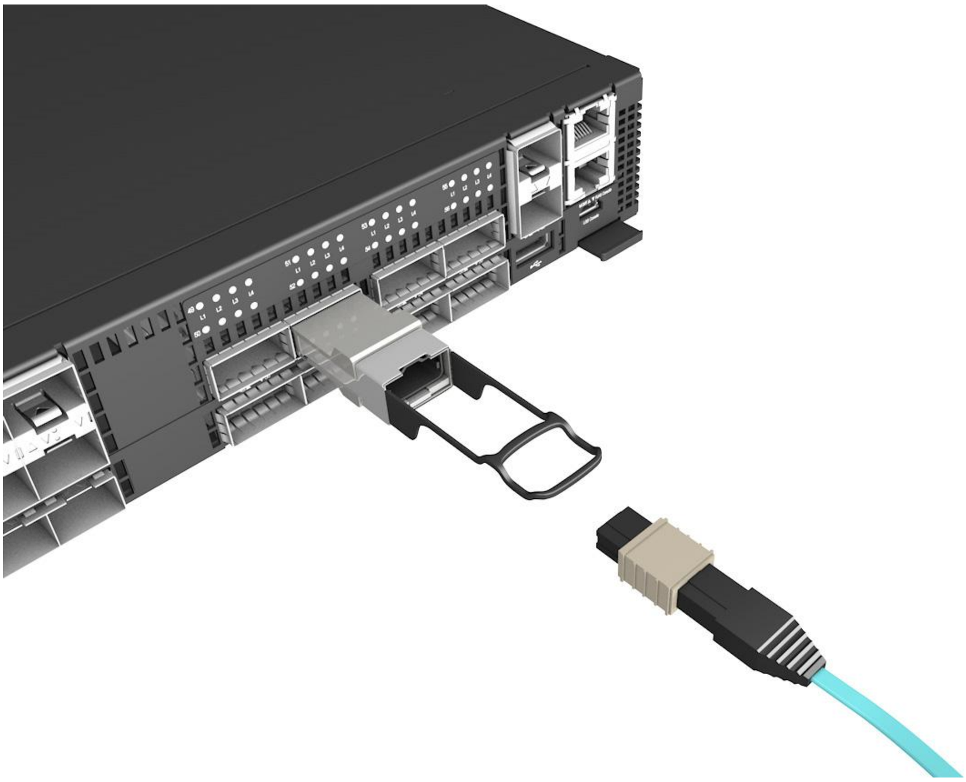
4.2.5.6 Connecting Network Cable

1. Connect a 100-ohm, Category 5.5e or better twisted-pair cable to the RJ-45 management port, using Category 5e or better twisted-pair cable.
2. Connect the DAC cable to the SFP28/QSFP28 slot. Or first install the SFP28/QSFP28 transceiver into the slot, and then connect the fiber optic cable to the transceiver port.
3. After the connection is established, check the port status LED to ensure that the link is valid. For SFP28 ports:

- Green — 25 Gbps mode
- Yellow — 10 Gbps mode

4. During connection, check the port status LED indicators to ensure the link is effective. Each QSFP28 port has four LED indicators that indicate a valid link in the following modes:

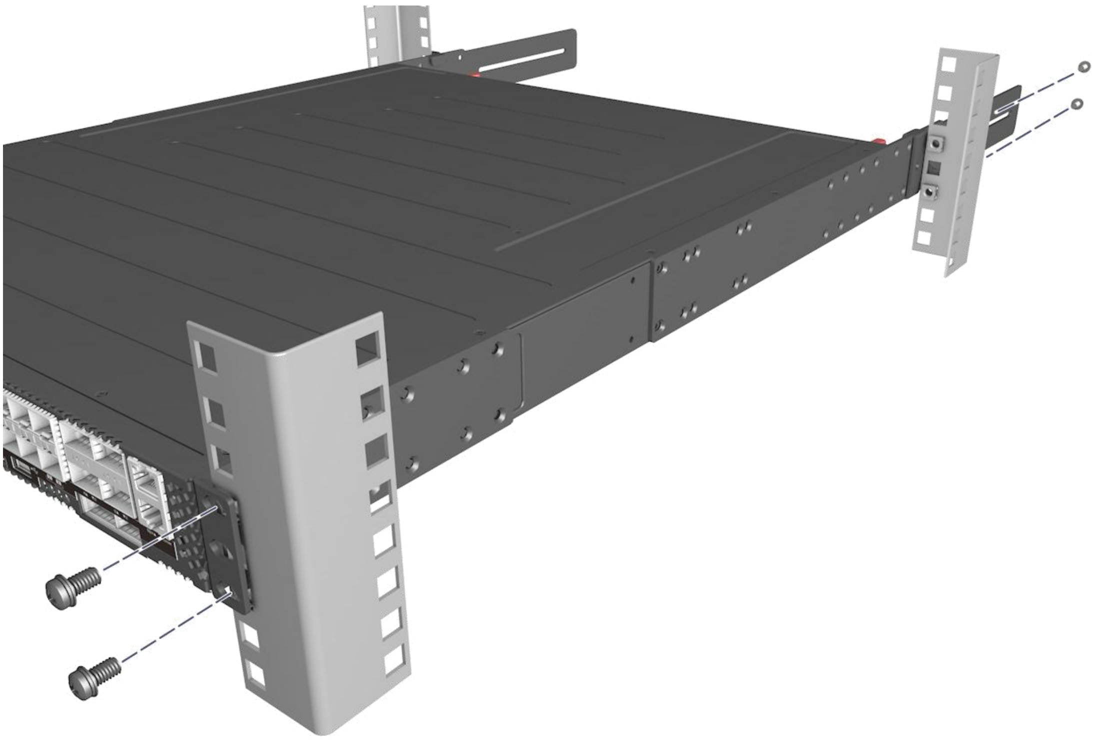
- 1 blue LED - 100 Gbps mode
- 1 orange LED - 40 Gbps mode
- 1-4 white LEDs - 25 Gbps grouped mode (four channels)
- 1-4 green LEDs - 10 Gbps grouped mode (four channels)
- 1 and 3 magenta LEDs - 50 Gbps split mode (two channels)



4.2.6 Network Equipment: N5850-48S6Q Installation

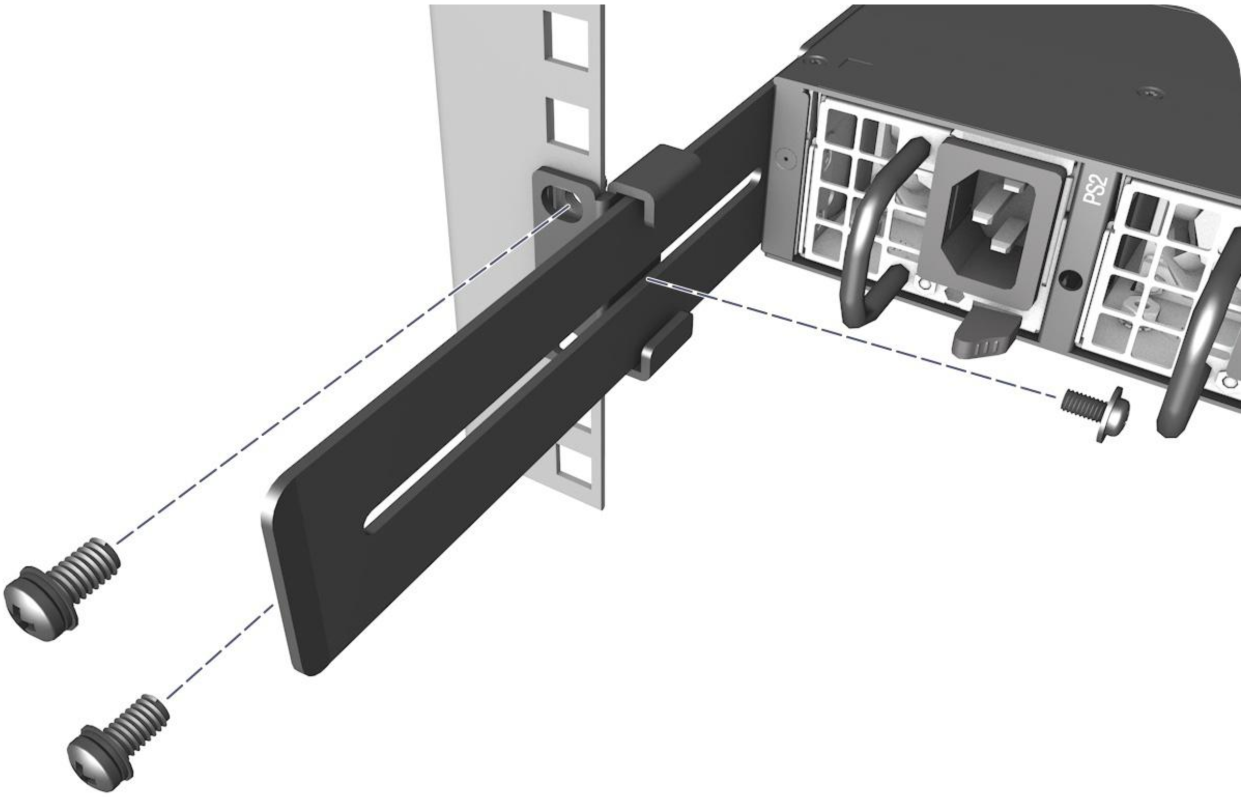
4.2.6.1 Rack Mounting

1. Install each front bracket and rear bracket onto the switch using the four included bracket screws.
2. Use two additional screws to secure each rear bracket at the midpoint on both sides of the switch.
3. Secure the switch in the rack with screws and cage nuts.



4.2.6.2 Adjust Rear-post Brackets

Use the provided locking device to lock the position of the rear bracket ear locking screws. You can also adjust the rear bracket ears to accommodate different rack depths from 56 cm to 75 cm.



Note

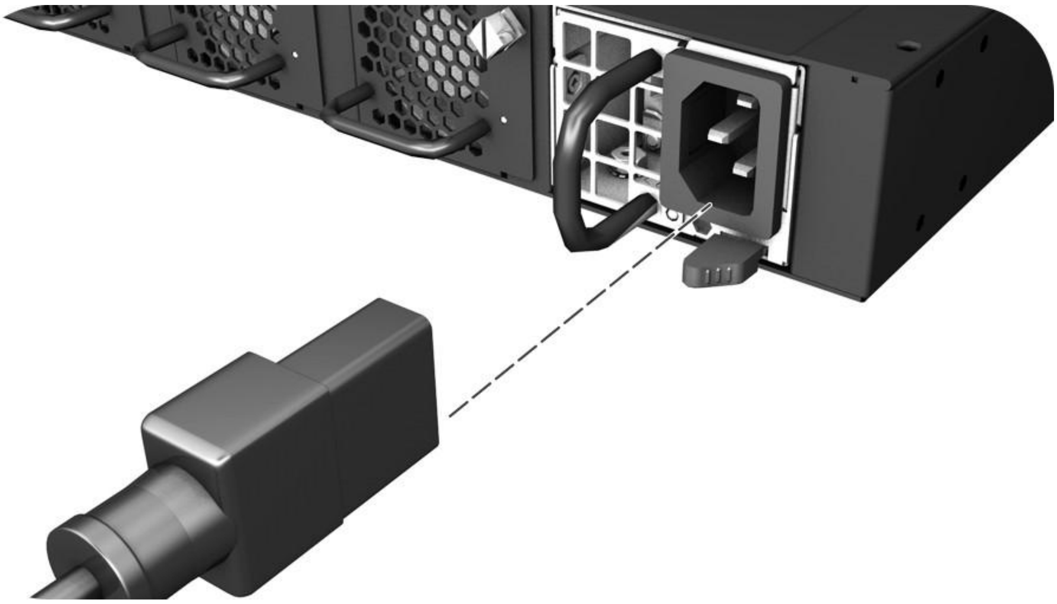
The guide rail can be adjusted according to the actual situation on site to meet the installation requirements.

4.2.6.3 Grounding the Switch

1. Ensure the rack is properly grounded and complies with ETSI ETS 300 253. Verify that there is a good electrical connection to the rack ground point (no paint or isolating finishes).
2. Connect a No. 10 AWG ground wire (not included) to the ground point on the rear panel of the switch, and then connect the other end of the wire to the rack ground point.

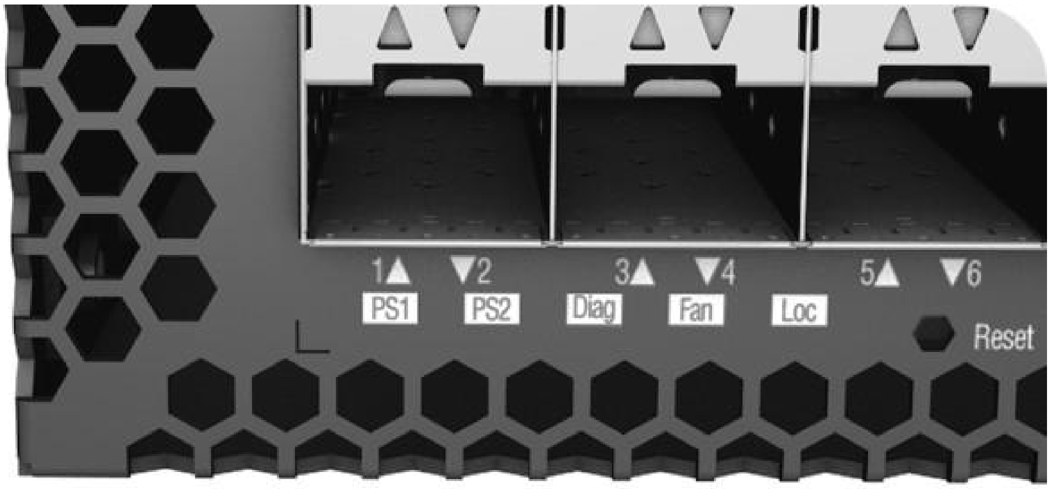
4.2.6.4 Connecting the Power

1. Install one or two AC or DC PSUs on the switch. The switch supports up to two PSUs, and their airflow direction must match the airflow direction of the installed fan tray.
2. Connect external DC power to the PSUs.



4.2.6.5 Verifying Switch Operation

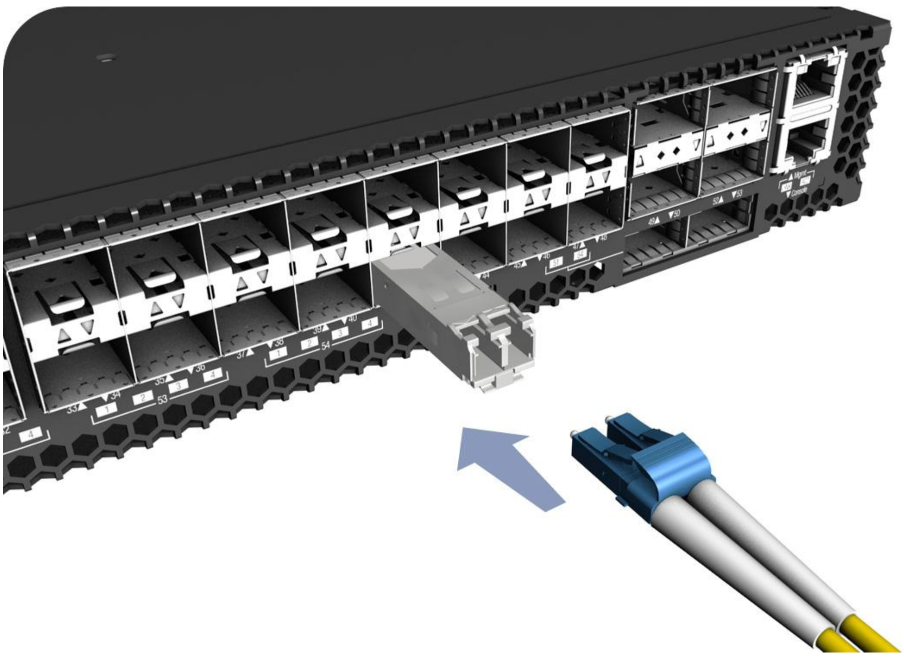
Verify basic operation of the switch by checking the system LEDs. For normal operation, the PSU1/PSU2, Diagnostic, and Fan LEDs should all be green.



4.2.6.6 Connecting Network Cable

1. Connect a 100-ohm, Category 5.5e or better twisted-pair cable to the RJ-45 management port, using Category 5e or better twisted-pair cable.
2. Connect the DAC cable to the SFP+/QSFP+ slot. Or first install the SFP+/QSFP+ transceiver into the slot, and then connect the fiber optic cable to the transceiver port. The following transceivers are available:

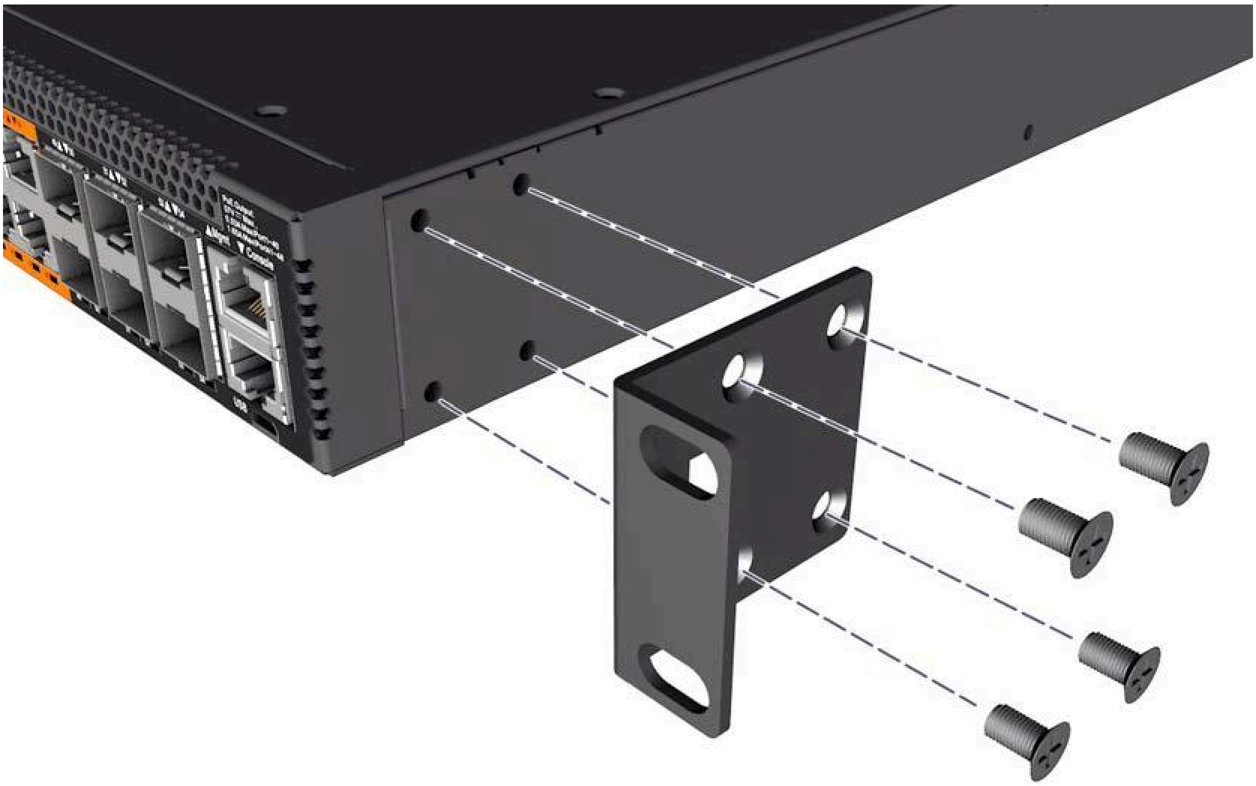
- 40GBASE-CR4 \ 40GBASE-SR4
- 10GBASE-CR \ 10GBASE-SR (ET5402-SR)
- 1000BASE-SX (ET4201-SX) \ 1000BASE-LX (ET4201-LX)



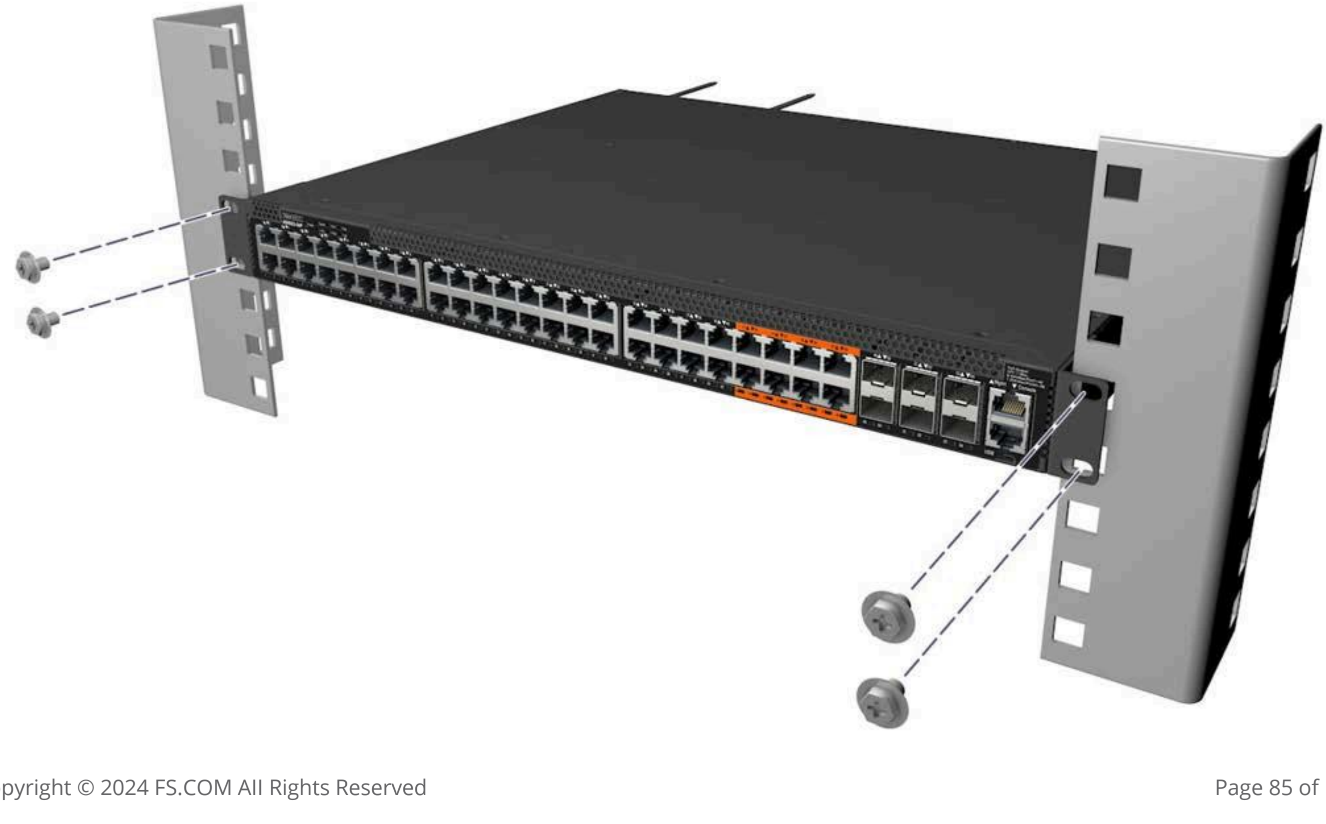
4.2.7 Network Equipment: S5870-48T6S Installation

4.2.7.1 Mounting the Switch

1. Install the brackets using the included flat-head M4x8 screws.



2. Secure the unit to the cabinet using the screws and bracket/spring clip nuts provided with the cabinet.



4.2.7.2 Grounding the Switch

1. Ensure Rack Grounding

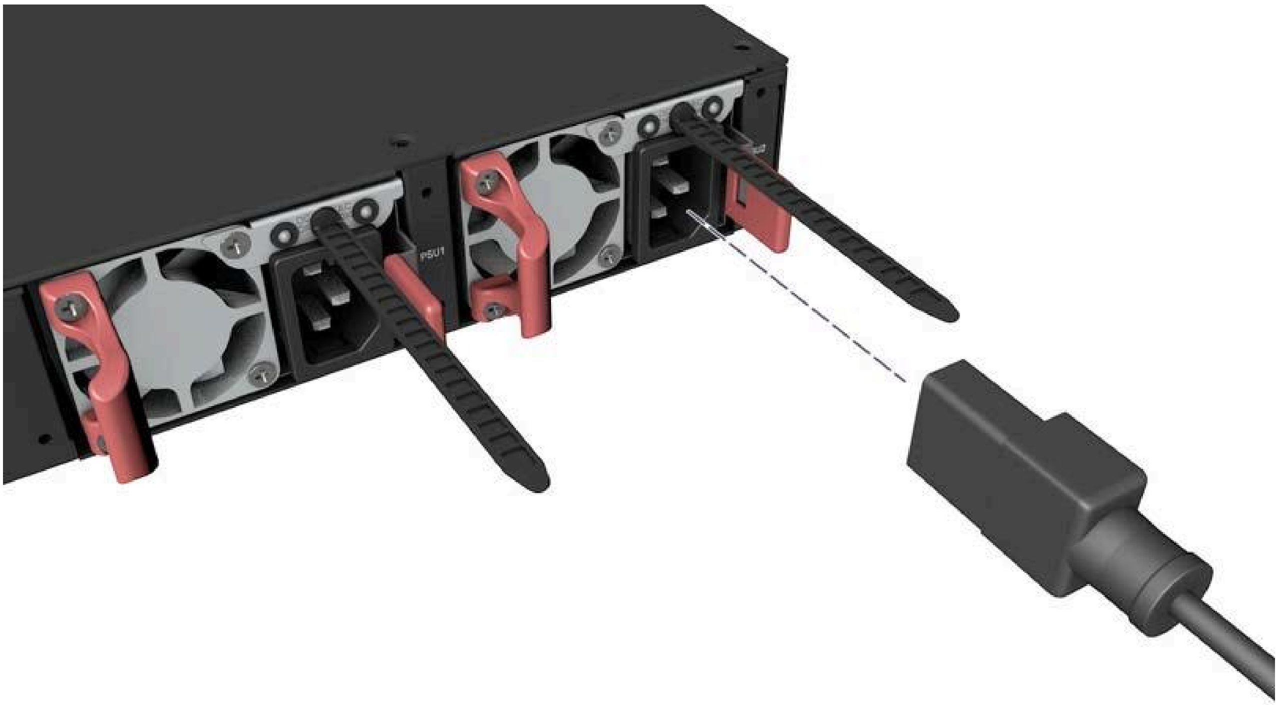
Verify that the rack is properly grounded and complies with international and local standards. Ensure a good electrical connection between the grounding point on the rack and the device (no paint or insulation surface treatment).

2. Install the Grounding Wire

- Attach a grounding wire (#14 AWG/1.5 mm², green with yellow stripes) to the grounding point on the rear or side panel of the device.
- Connect the other end of the grounding wire to the rack grounding point.

4.2.7.3 Connecting the Power

If not pre-installed, install one or two AC PSUs in the device. Then connect the external AC power to the PSUs.



4.2.7.4 Connecting Network Cable

1. RJ45 port:

Connect Category 5, 5e or better twisted pair cables.

2. SFP+ port:

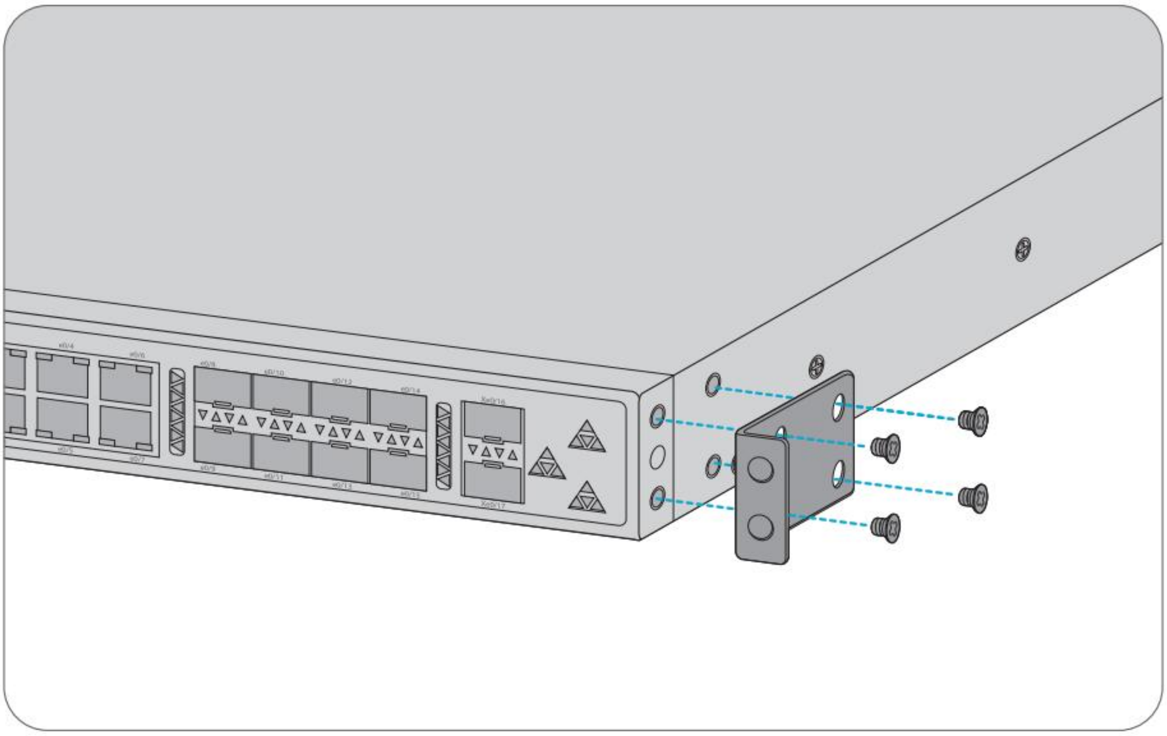
Install the transceiver and connect the fiber optic cable to the transceiver port. Or you can connect the AEC/AOC/DAC cable directly to the SFP+ slot. The SFP+ port supports the following transceivers:

- 10GBASE-SR, LR, CR, T
- 1000BASE-SX, LX, T

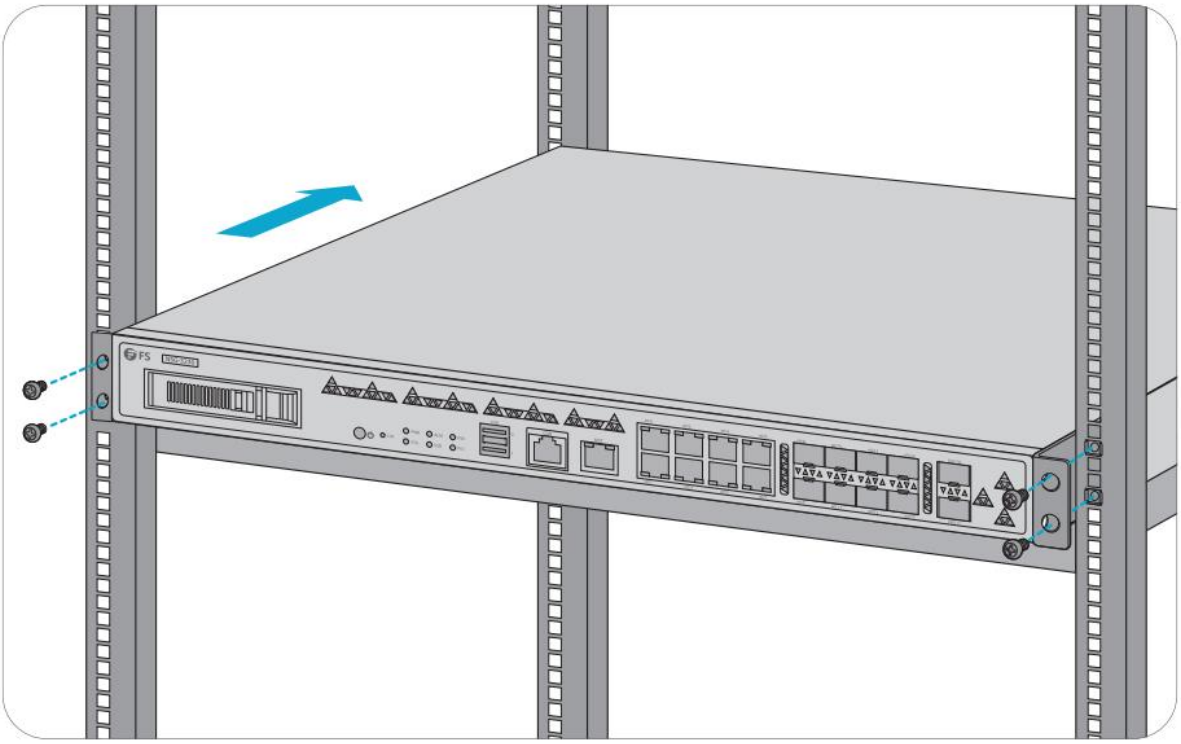
4.2.8 Network Equipment: NSG-5220 Installation

4.2.8.1 Firewall Mounting

1. Attach the brackets to the switch.



2. Secure the switch to the rack using the screws and cage nuts/clips provided with the rack.



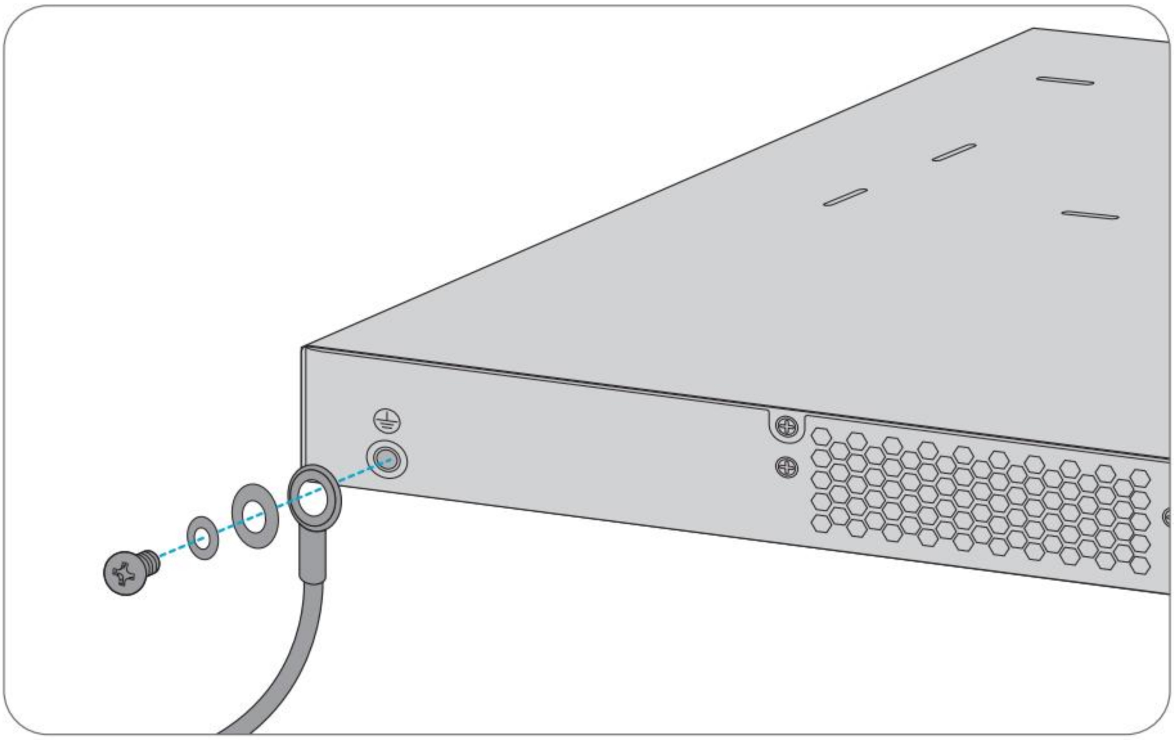
4.2.8.2 Grounding the Firewall

1. Ensure Rack Grounding

Verify that the rack is properly grounded and complies with international and local standards. Ensure a good electrical connection between the grounding point on the rack and the device (no paint or insulation surface treatment).

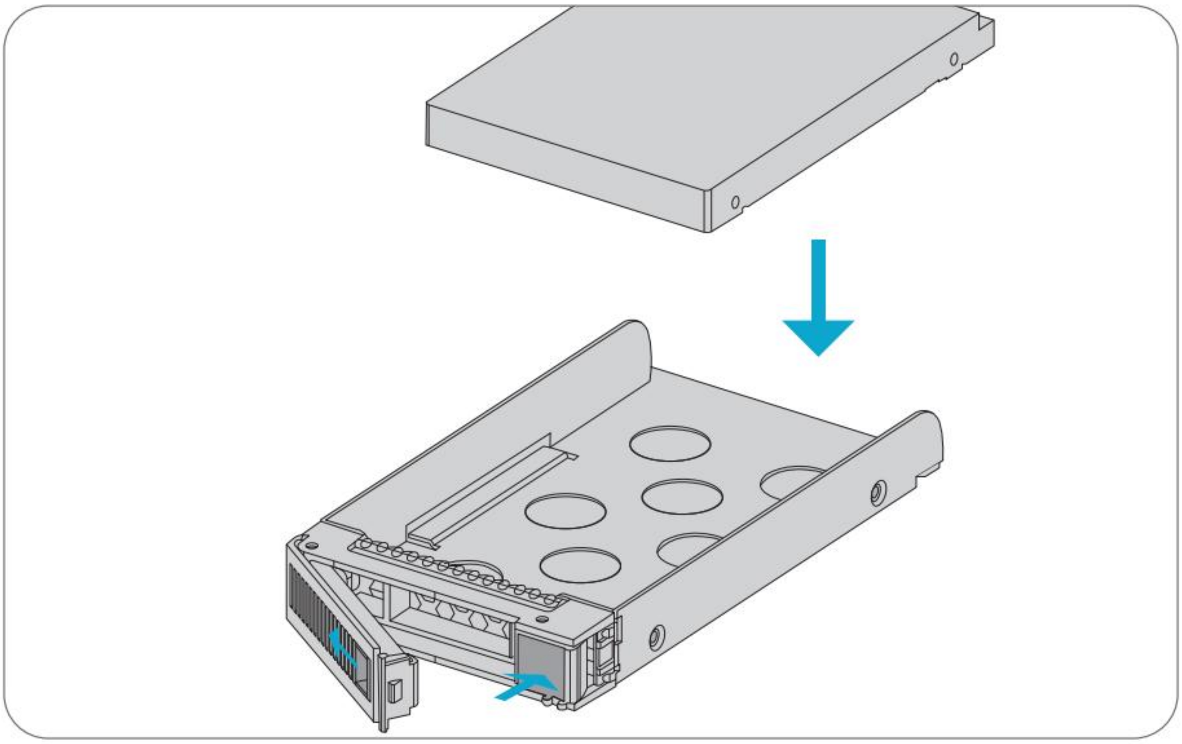
2. Install the Grounding Wire

- Attach a grounding wire (#14 AWG/1.5 mm², green with yellow stripes) to the grounding point on the rear or side panel of the device.
- Connect the other end of the grounding wire to the rack grounding point.

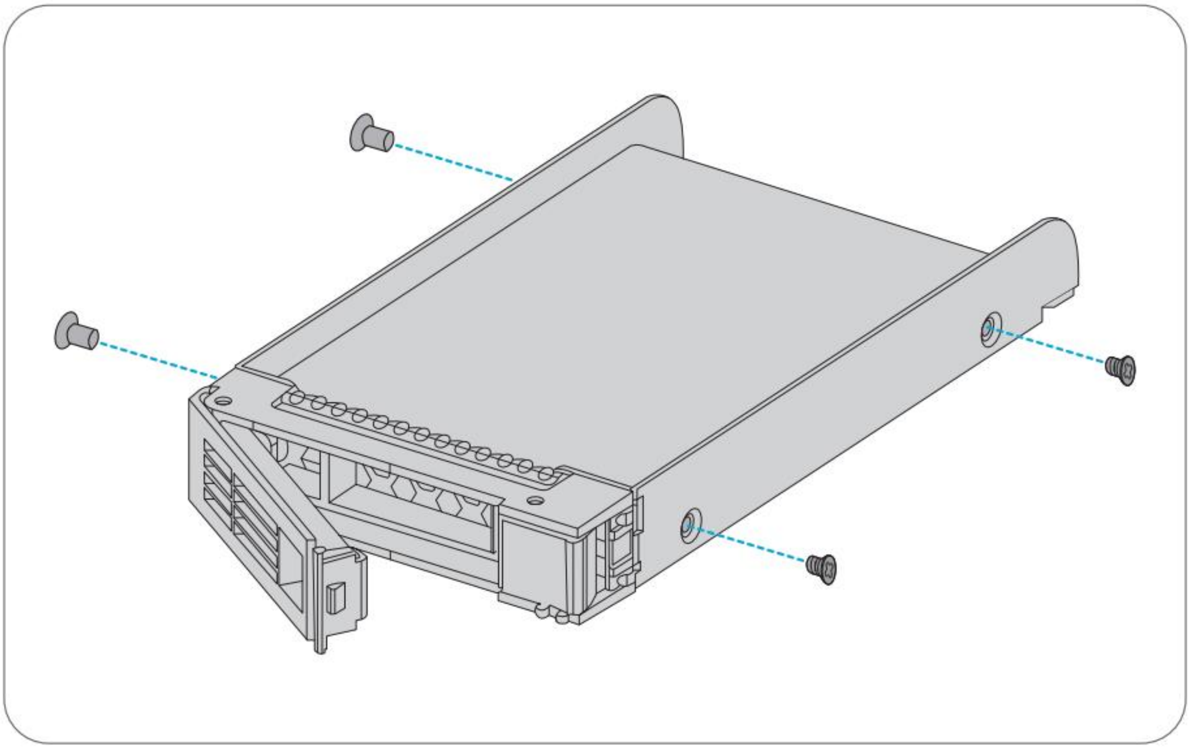


4.2.8.3 Installing the SSD

1. Move the red button on the SSD expansion slot to the unlocked state COP2, press the black button next to it, and then the slot cover will open automatically.
2. Pull the disk enclosure slowly out of the firewall.

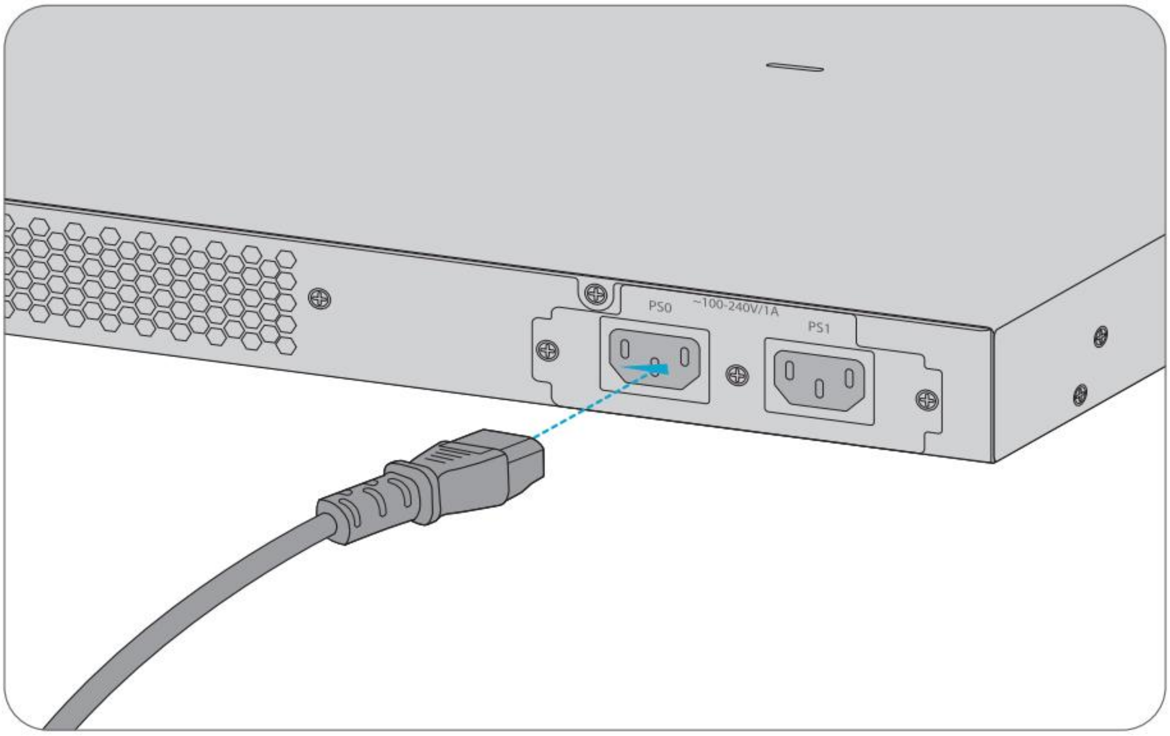


3. Put the SSD inside, align the screw holes on the SSD with thoseon the disk enclosure, and tighten the screws on both sides.



4.2.8.4 Connecting the Power

1. Plug the AC power cord into the power port on the firewall (A).
2. Connect the other end of the power cord to an AC power source.
3. Press the power button for about 5 seconds to turn on the firewall (B).



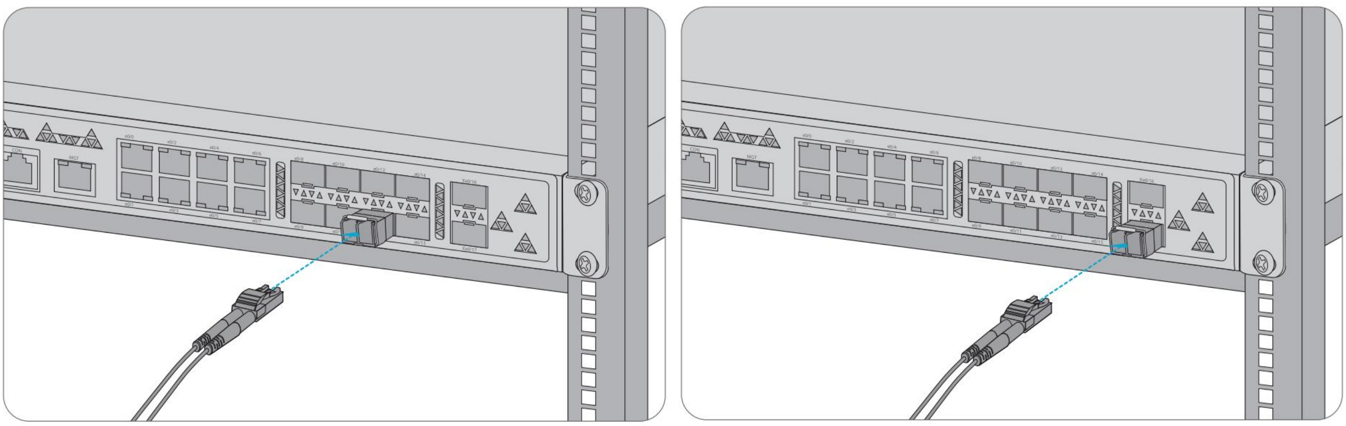
4.2.8.5 Connecting Network Cable

1. RJ45 port:

Connect Category 5, 5e or better twisted pair cables.

2. SFP/SFP+ port:

- Plug a compatible SFP/SFP+ optical module into the SFP/SFP+ port on the firewall.
- Connect a fiber optic cable to the optical module. Then connect the other end of the cable to another fiber optic device.



4.3 Powering on the Device

4.3.1 Pre-Power Checks

1. Ensure the device is adequately grounded.
2. Verify that fan modules and power modules are correctly installed.
3. Check power cord connections for correctness and ensure they match the input power requirements.
4. Confirm that the supplied voltage matches the device's requirements.
5. Verify that cable configurations are correct, the terminal (which could be a PC) intended for configuration is powered on, and configuration parameters are set.

4.3.2 Powering on the Device

After completing device installation and ensuring all checklist items are ready, close the breaker at the power supply module and turn on the cabinet switch.

4.3.3 Checks after power-on

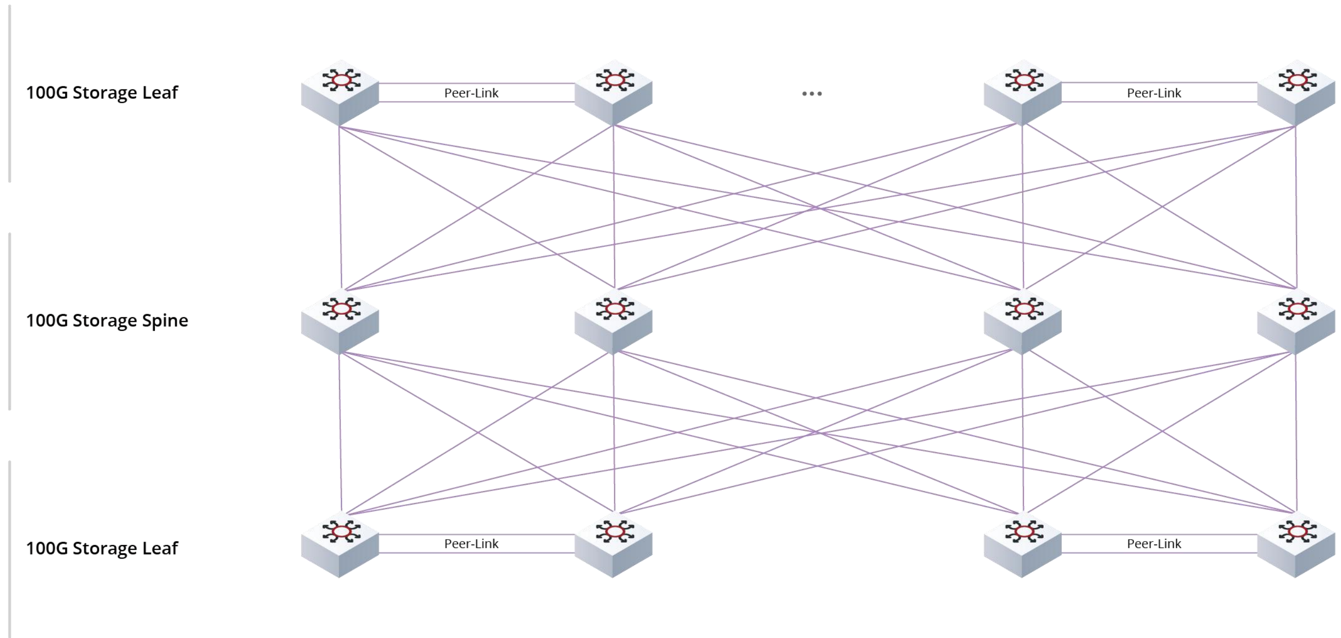
After powering on, perform the following checks to ensure proper configuration:

1. Ensure the ventilation system is functioning correctly.
2. Check for print information on the terminal interface used for configuration.
3. Verify that indicator lights on the device display normal operation.

5. Storage Network Debugging

5.1 Storage Network Topology Description

5.1.1 Example of Solution Topology



5.1.2 Solution Topology Description

- All Server San servers are interconnected with GPU servers.
- Use a 100G network for interconnecting GPU servers and storage servers.
- BGP protocol is used between Leaf and Spine. Through BGP's powerful routing control capabilities, the storage network forwarding path is optimized to ensure that storage does not have high latency.

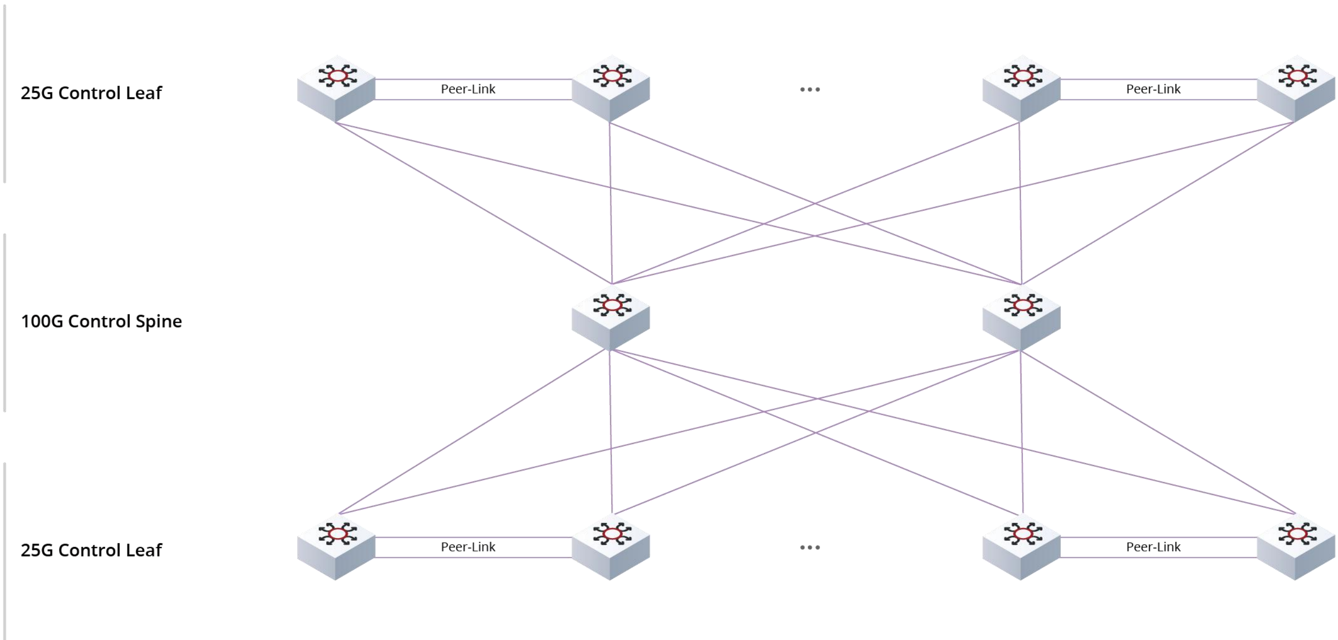
5.2 Switch Configuration

Please refer to the PicOS configuration document: PicOS® Configuration Guide V4.4.4

6. Storage Management Network Debugging

6.1 Storage Management Network Topology Description

6.1.1 Example of Solution Topology



6.1.2 Solution Topology Description

- All Server San servers are interconnected with GPU servers.
- All Server San storage management networks are interconnected using a separate 25G network card to ensure that the management network is isolated from other networks and does not affect each other.
- BGP protocol is used between Leaf and Spine. Through BGP's powerful routing control capabilities, the storage network forwarding path is optimized to ensure that storage does not have high latency.

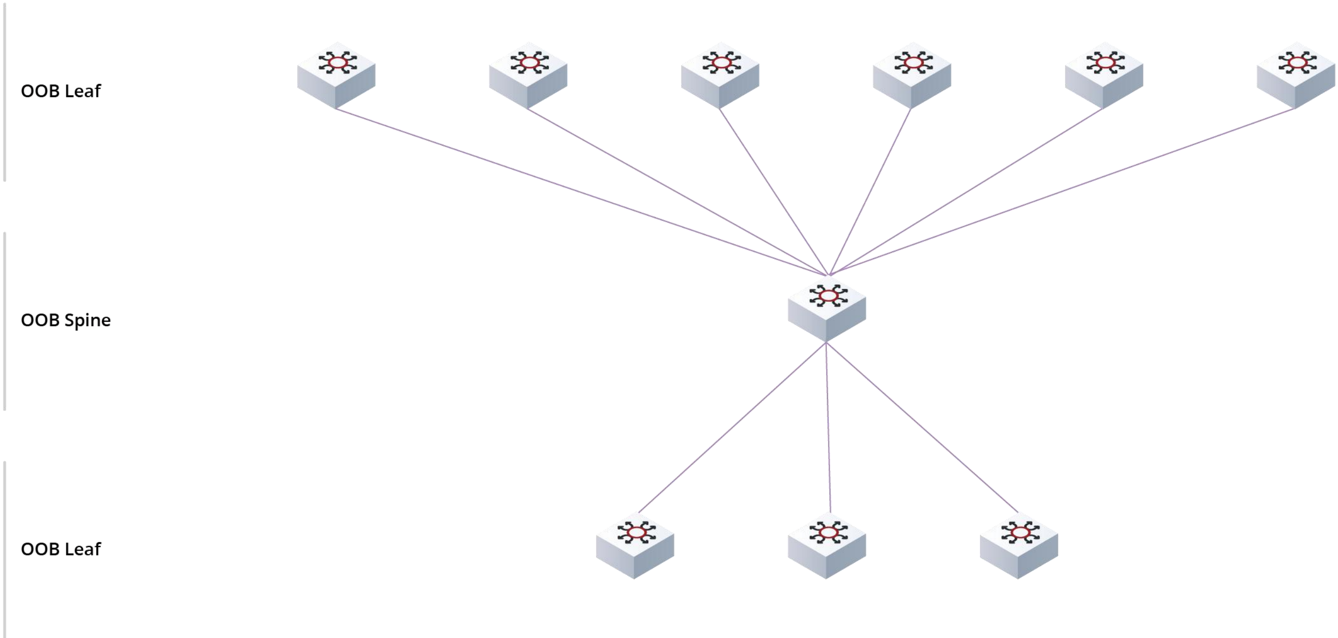
6.2 Switch Configuration

Please refer to the PicOS configuration document: PicOS® Configuration Guide V4.4.4

7. Device Management Network Debugging

7.1 Device Management Network Topology Description

7.1.1 Example of Solution Topology



7.1.2 Solution Topology Description

- Connect the management ports of all servers and switches to the management network.
- Use a 1G network for the switch downlinks connecting to each device, and employ 10G devices for the uplinks interconnecting the switches.

7.2 Switch Configuration

Please refer to the PicOS configuration document: PicOS® Configuration Guide V4.4.4

7.2.1 Configure the Spine Switch

7.2.2 Configure the Leaf Switch

8. GPU Server Network Debugging

8.1 Configuration Approach

Steps	Actions	Specific Actions	Purpose
Step 1	Server operating system installation	Prerequisite preparation	PXE is a prerequisite for installing an operating system. If the environment is not set up, the corresponding system installation cannot be completed through PXE. This includes the construction of DNS and DHCP servers.
		PXE operating system installation	Build the PXE installation environment, including the PXE client to be connected to the PXE server, to ensure that the client can pull the appropriate image files as well as configuration files on the server.
		Modify BMC IP address	Modify the BMC IP address to facilitate later server maintenance and management.
		Switch configuration modification	Shut down PXE traffic on the switch to avoid infinite reboot to the PXE installation interface after the server OS installation is completed.
Step 2	GPU and NIC driver installation	Prerequisite preparation	Let the motherboard scheduling GPU, network card resources to complete the hardware level integration.
Step 3	Operating system suite synchronization	Install rsyslog	Install the log server, when the server is working, when there is a corresponding error report, it can be found in time, as well as to avoid the server downtime, the log is lost, and the cause of the log can not be found in time.
		Automatic disk order formatting and mounting	Create an automatic formatting script, batch format and mount disks to support subsequent business.
		Install kernel and configure Use of corresponding kernel version drivers	Install the GPU server kernel to ensure that the server kernel is compatible with the client's current business.

8.2 Network Equipment Debugging

In this case, the configuration work involved in the IB Switch Configuration board is completed by NVIDIA engineers. To ensure the completeness of the documentation and the comprehensiveness of the user information, the content of this board is covered in this document. It is recommended that when implementing or adjusting IB-related configurations, please work or consult directly with the NVIDIA technical team to ensure that best practices can be achieved.

8.3 GPU Server Debugging

8.3.1 GPU Server System Installation

8.3.1.1 System Installation Prerequisites

1. Power supply

When the GPU server does not have an IB optical module installed, at least three PSUs are required for proper installation.

When the GPU server has an IB optical module installed, at least four PSUs are required for proper installation.

When the number of PSUs is not enough, it will not be able to boot up properly due to insufficient power supply.

When the GPU server is already powered up, at least 2 PSUs are required to ensure that it will not shut down.

2. Server

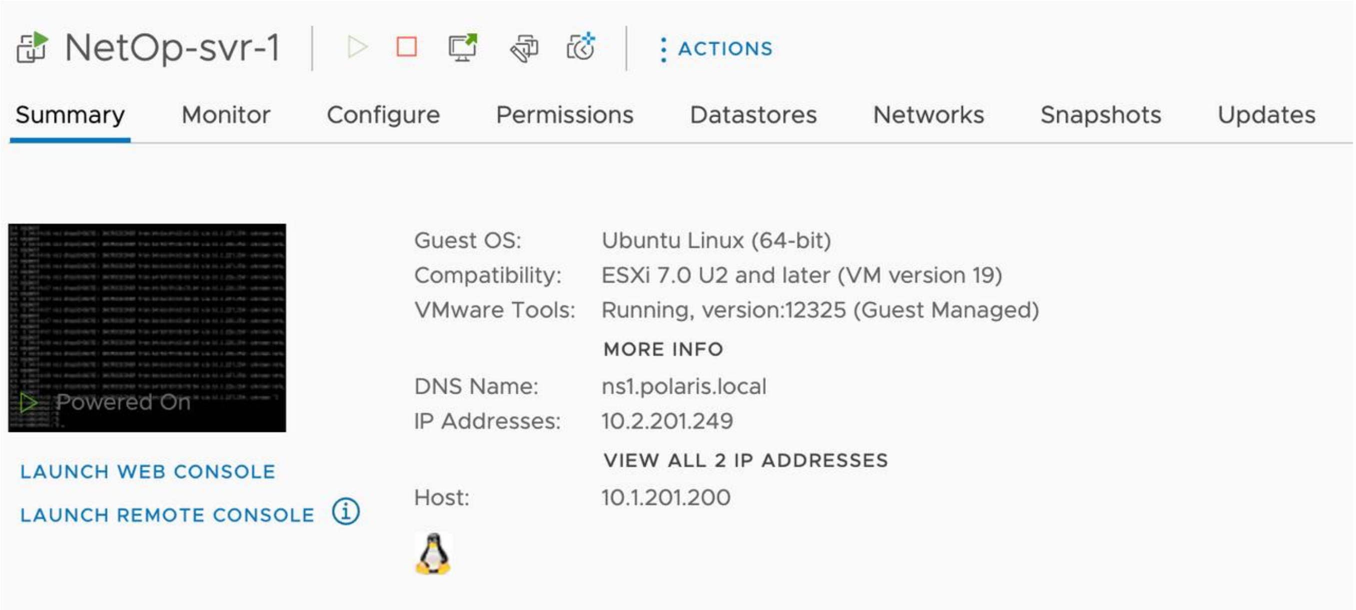
Please adjust the IP address according to the different PODs, and when the deployment is complete, it is recommended to delete the PXE related configuration.

After the DHCP configuration is done, you must restart the dhcp service with the command: SUdo systemctl restart isc-dhcp-server and check the status: SUdo systemctl status isc-dhcp-server The status should be as follows:

```
netop-admin@ns1:~$
netop-admin@ns1:~$ sudo systemctl status isc-dhcp-server
● isc-dhcp-server.service - ISC DHCP IPv4 server
     Loaded: loaded (/lib/systemd/system/isc-dhcp-server.service; enabled; vendor preset: enabled)
     Active: active (running) since Sat 2024-01-06 17:31:48 UTC; 1h 53min ago
       Docs: man:dhcpd(8)
   Main PID: 59651 (dhcpd)
      Tasks: 4 (limit: 38375)
     Memory: 5.2M
        CPU: 35ms
     CGroup: /system.slice/isc-dhcp-server.service
             └─59651 dhcpd -user dhcpd -group dhcpd -f -4 -pf /run/dhcp-server/dhcpd.pid -cf /etc/dhcp/dhcpd.conf ens160

Jan 06 18:36:35 ns1.polaris.local dhcpd[59651]: DHCPOFFER on 10.2.7.106 to b8:3f:d2:2a:27:3c ((none)) via 10.2.7.254
Jan 06 18:36:35 ns1.polaris.local dhcpd[59651]: DHCPREQUEST for 10.2.7.106 (10.2.201.249) from b8:3f:d2:2a:27:3c ((none)) via 10.2.7.254
Jan 06 18:36:35 ns1.polaris.local dhcpd[59651]: DHCPACK on 10.2.7.106 to b8:3f:d2:2a:27:3c ((none)) via 10.2.7.254
Jan 06 18:36:49 ns1.polaris.local dhcpd[59651]: DHCPDISCOVER from b8:3f:d2:2a:27:3c ((none)) via 10.2.7.254
Jan 06 18:36:49 ns1.polaris.local dhcpd[59651]: DHCPOFFER on 10.2.7.106 to b8:3f:d2:2a:27:3c (ubuntu-server) via 10.2.7.254
Jan 06 18:36:49 ns1.polaris.local dhcpd[59651]: DHCPREQUEST for 10.2.7.106 (10.2.201.249) from b8:3f:d2:2a:27:3c (ubuntu-server) via 10.2.7.254
Jan 06 18:36:49 ns1.polaris.local dhcpd[59651]: DHCPACK on 10.2.7.106 to b8:3f:d2:2a:27:3c (ubuntu-server) via 10.2.7.254
Jan 06 18:37:00 ns1.polaris.local dhcpd[59651]: reuse_lease: lease age 11 (secs) under 25% threshold, reply with unaltered, existing lease for 10.2.7.106
Jan 06 18:37:00 ns1.polaris.local dhcpd[59651]: DHCPREQUEST for 10.2.7.106 from b8:3f:d2:2a:27:3c (ubuntu-server) via 10.2.7.254
Jan 06 18:37:00 ns1.polaris.local dhcpd[59651]: DHCPACK on 10.2.7.106 to b8:3f:d2:2a:27:3c (ubuntu-server) via 10.2.7.254
netop-admin@ns1:~$
```

The state must be Active, if it is any other state then check for spelling and indentation errors.



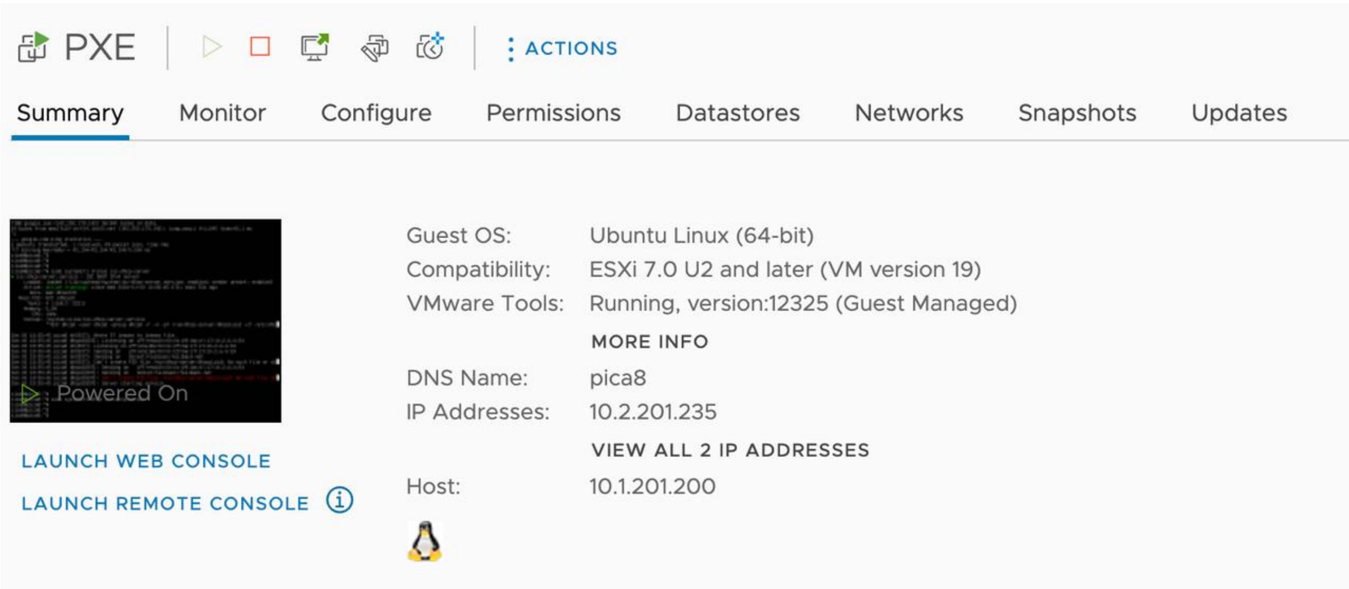
- PXE mirror server, this VM receives tftp and http traffic.

username: ********

password: ********

Image file path: /var/www/html/ubuntu-22.04.3custom.iso

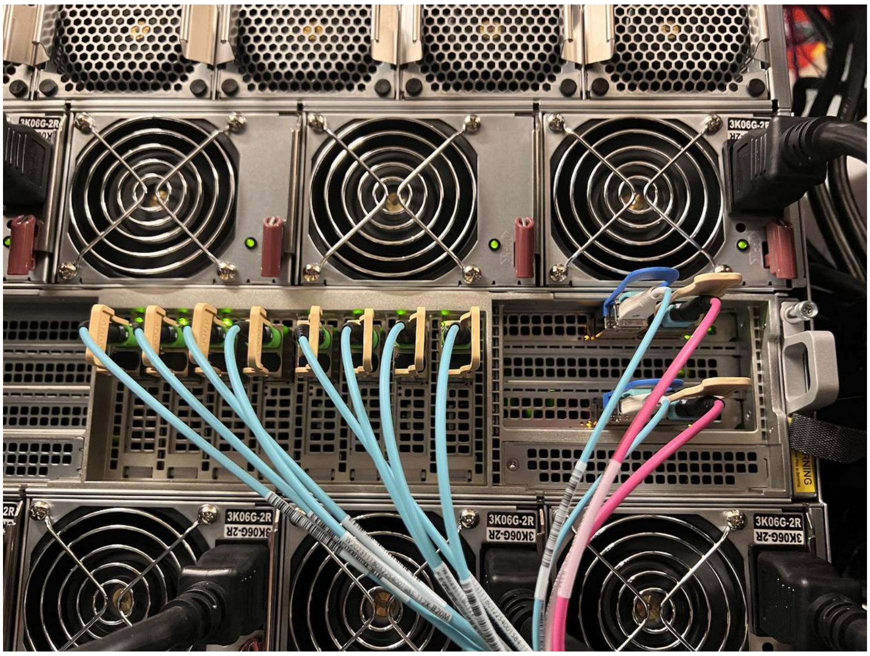
Automated deployment file path: /var/www/html/jammy/user-data



8.3.1.2 PXE Installation of Operating Systems

1. **PXE Environment Preparation**

- Please follow the figure below to install and connect the modules, and check whether the PXE server and PXE client can establish communication normally.



- PicOS is pre-configured with MLAG+LACP+LACP fallback. The LACP member ports of the PicOS switch will not forward any data when no aggregation message is received from the server, so you need to enable the LACP fallback function on PicOS. If you encounter that the server does not load the system successfully, that is, it cannot communicate with the PXE server, please check whether the switch is configured with the fallback function.
- Press the server's front panel power on button to boot up. If the server was booted up before the cable was connected, you need to reboot the server. If the server starts up normally and all networks are working, wait for about 15 minutes for the server to shut down automatically. If the server shuts down automatically, the installation has been successful. If the server does not shut down automatically, you need to use KVM switch or login to BMC to check the console output.

2. **BMC IP and password change instructions**

- UltraViolet servers do not have default BMC ip addresses, they are all DHCP clients, which means that you need a DHCP server to assign IPs to the BMCs before you can log in to the servers. 2. 1G switches have been configured with DHCP relay and are able to get IPs from the DHCP server above:
- The 1G switch has been configured with DHCP relay and is able to obtain IPs from the DHCP server above. to see what IPs have been assigned to each server, you need to log in to the DHCP server to view the log messages. the mac addresses of the BMCs can be viewed from the front panel Services tab, as shown in the following figure:

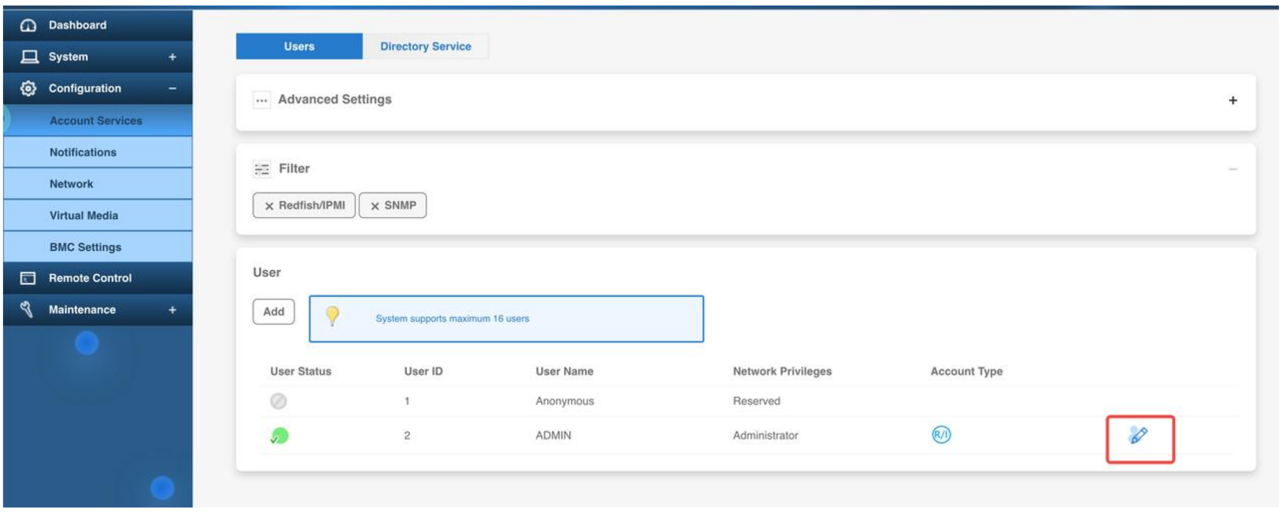
① ssh netop-admin@10.2.201.249 password:********

② journalctl -u isc-dhcp-server | grep fa:9b //This fa:9b is the last four digits of the mac address of the BMC in the above figure, and it matches directly, as shown below:

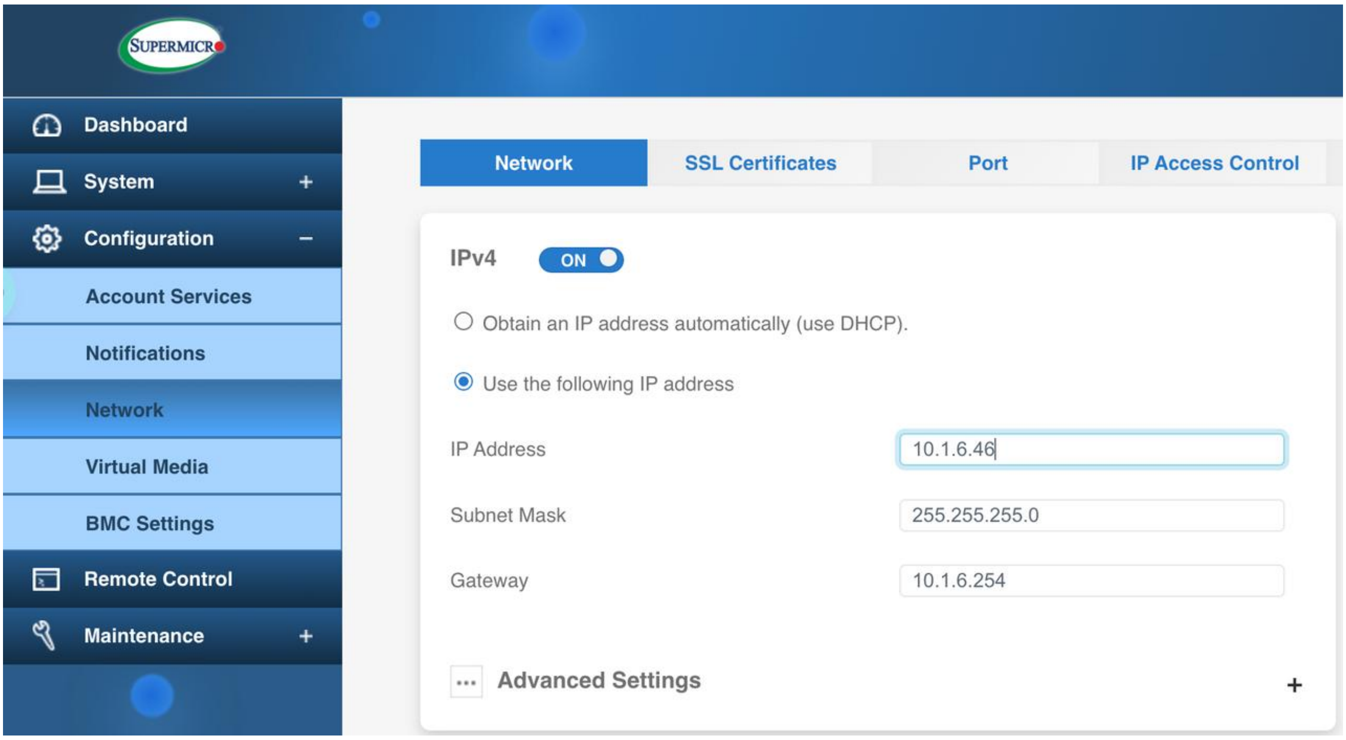
```
[netop-admin@ns1:~$ journalctl -u isc-dhcp-server | grep fa:9b
 Jan 03 15:09:03 ns1.polaris.local dhcpd[45808]: DHCPDISCOVER from 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 15:09:04 ns1.polaris.local dhcpd[45808]: DHCPOFFER on 10.1.6.100 to 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 15:09:04 ns1.polaris.local dhcpd[45808]: DHCPREQUEST for 10.1.6.100 (10.2.201.249) from 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 15:09:04 ns1.polaris.local dhcpd[45808]: DHCPACK on 10.1.6.100 to 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 15:39:04 ns1.polaris.local dhcpd[46060]: DHCPREQUEST for 10.1.6.100 from 7c:c2:55:50:fa:9b via ens160
 Jan 03 15:39:04 ns1.polaris.local dhcpd[46060]: DHCPACK on 10.1.6.100 to 7c:c2:55:50:fa:9b via ens160
 Jan 03 16:04:36 ns1.polaris.local dhcpd[46386]: DHCPDISCOVER from 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 16:04:37 ns1.polaris.local dhcpd[46386]: DHCPOFFER on 10.1.6.100 to 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 16:04:37 ns1.polaris.local dhcpd[46386]: DHCPREQUEST for 10.1.6.100 (10.2.201.249) from 7c:c2:55:50:fa:9b via 10.1.6.254
 Jan 03 16:04:37 ns1.polaris.local dhcpd[46386]: DHCPACK on 10.1.6.100 to 7c:c2:55:50:fa:9b via 10.1.6.254
```

You can see that the server is assigned the address 10.1.6.100, and then use this address to log in to the BMC

- After logging in to the BMC, the default user name is ADMIN in uppercase, there is no default password, each unit has a random password, the password is in the service paper on the front panel, as shown in the picture above with the letters after the PWD field.
- After logging in to BMC, please change the password to: ********

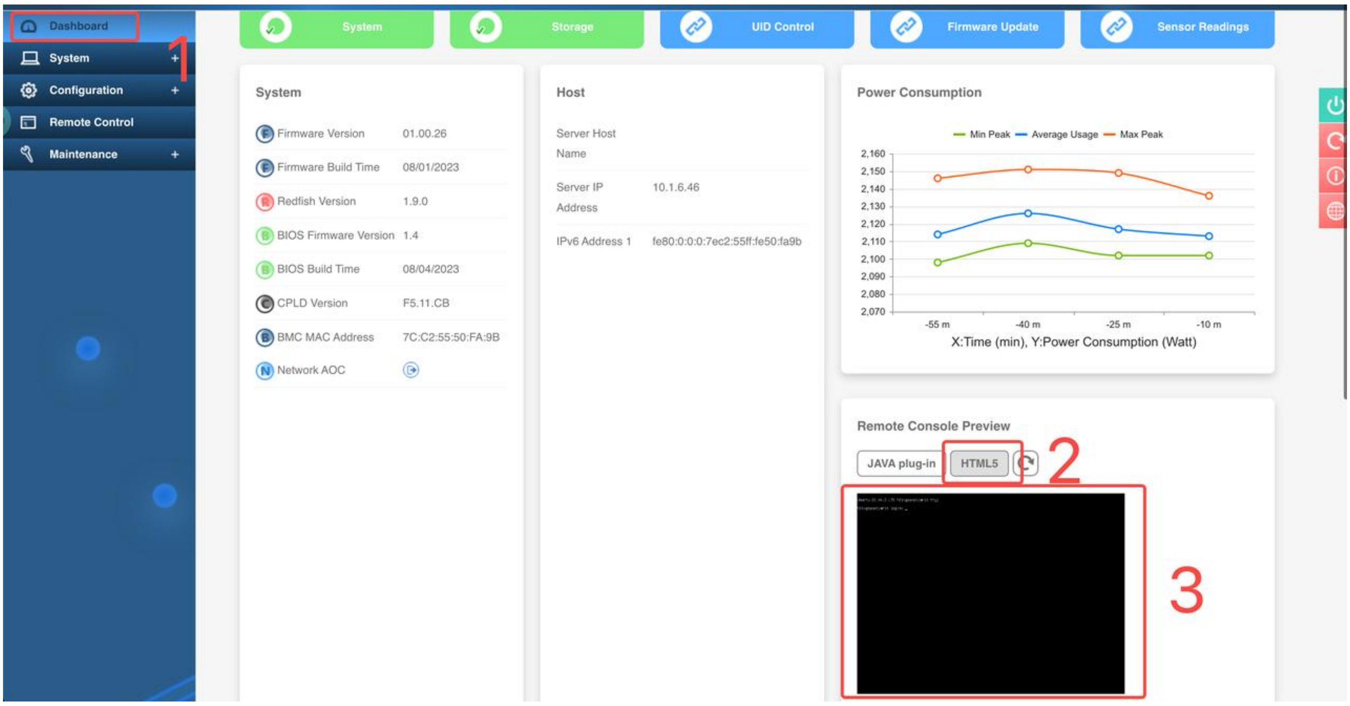


Along with this, change the default IP address of the BMC:



3. Modify the ip address and Hostname of the business system.

- Log in to BMC and open the console



Please refer to the AmpCon ZTP deployment document: [AmpCon Zero Touch Provisioning]

- Generate json file by manually logging SN number
- Bind the SN number to the IP address to generate a json file.
- Use Task Ansible to call the json file and execute the configuration to generate the playbook file.

Obtain device serial number through IPMI tool, then use serial number SN to find predefined IP address mapping. Update the hostname and network configuration via the json file, and finally verify that the new network configuration is in effect.

- Create a host file that traverses all dhcp addresses for different pods using a playbook file.

This script tells Ansible what IP address to use to connect to the device, what username and password to use, and how to handle SSH connections when connecting, and performs a login test.

- Design Ideas

When deployed, an intranet bonding interface with DHCP enabled obtains a specified network segment IP address from a DHCP server.

Iterate through all the IP addresses in this address pool, install ipmitool in the reachable server, and use this command to get the SN of this server.

Configure the corresponding IP address by using the SN --- Intranet bonding IP address --- Storage bonding IP address correspondence in sn_ip_map.json.

- Usage

Server: 10.20.2.254 netop-admin / Admin

Folder address: /home/netop-admin/bk/

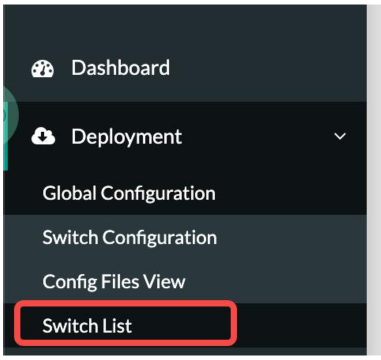
Under the folder, create the sn_ip_map.json file according to the POD, and create the host file according to the network segment of the corresponding POD.

```
1. cd /home/netop-admin/bk/
2. ansible-playbook -i host task.yml --extra-vars "ansible_SUdo_pass=<linux-default-passwd>" | tee -a reSUlt.txt
```

4. Modify the switch configuration to prevent PXE loops after PXE is completed.

- Login Ampcon : https://10.2.201.245 Username password:admin/admin
- Please click

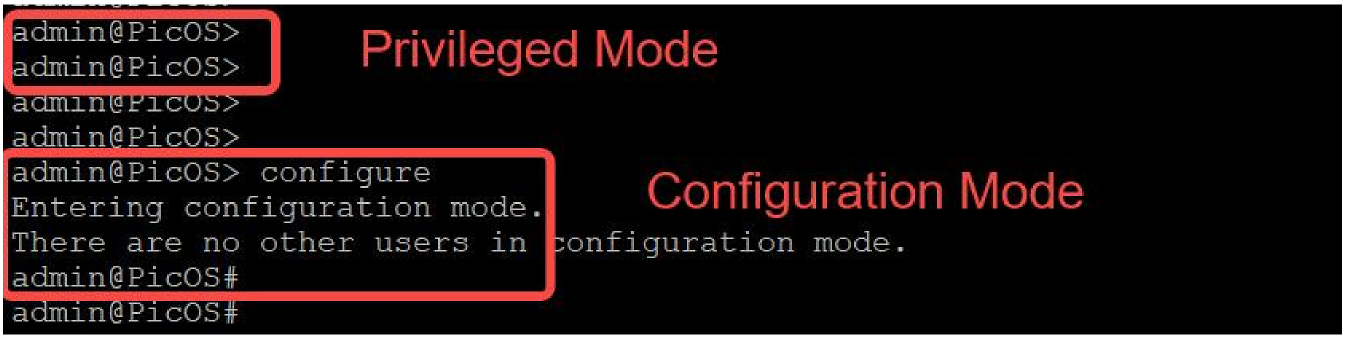
-



- Then find the corresponding POD intranet Leaf01 switch according to the switch name and SSH into it, switch username/password:admin/pica8.
- Please check whether the switch has fallback configuration first:

In privileged mode: show running-config | display set | match fallback

In configuration mode: run show running-config | display set | match fallback



- If there is no fallback, add the fallback command to all server-facing ports:

1. First enter configure to enter configuration mode
2. set interface aggregate-ethernet ae1 aggregated-ether-options lacp fallback enable true
3. set interface aggregate-ethernet ae2 aggregated-ether-options lacp fallback enable true
4. set interface aggregate-ethernet ae3 aggregated-ether-options lacp fallback enable true
5. set interface aggregate-ethernet ae4 aggregated-ether-options lacp fallback enable true
6. set interface aggregate-ethernet ae5 aggregated-ether-options lacp fallback enable true
7. set interface aggregate-ethernet ae6 aggregated-ether-options lacp fallback enable true
8. set interface aggregate-ethernet ae7 aggregated-ether-options lacp fallback enable true
9. set interface aggregate-ethernet ae8 aggregated-ether-options lacp fallback enable true
10. set interface aggregate-ethernet ae9 aggregated-ether-options lacp fallback enable true
11. set interface aggregate-ethernet ae10 aggregated-ether-options lacp fallback enable true
12. set interface aggregate-ethernet ae11 aggregated-ether-options lacp fallback enable true
13. set interface aggregate-ethernet ae12 aggregated-ether-options lacp fallback enable true
14. set interface aggregate-ethernet ae13 aggregated-ether-options lacp fallback enable true
15. set interface aggregate-ethernet ae14 aggregated-ether-options lacp fallback enable true
16. set interface aggregate-ethernet ae15 aggregated-ether-options lacp fallback enable true
17. set interface aggregate-ethernet ae16 aggregated-ether-options lacp fallback enable true
18. set interface aggregate-ethernet ae17 aggregated-ether-options lacp fallback enable true
19. set interface aggregate-ethernet ae18 aggregated-ether-options lacp fallback enable true

```
set interface aggregate-ethernet ae19 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae20 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae21 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae22 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae23 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae24 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae25 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae26 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae27 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae28 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae29 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae30 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae31 aggregated-ether-options lacp fallback enable true
set interface aggregate-ethernet ae32 aggregated-ether-options lacp fallback enable true
After execution, enter
commit
After the installation is complete and all servers are shut down, execute:
del interface aggregate-ethernet ae1 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae2 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae3 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae4 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae5 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae6 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae7 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae8 aggregated-ether-options lacp fallback
 del interface aggregate-ethernet ae9 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae10 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae11 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae12 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae13 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae14 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae15 aggregated-ether-options lacp fallback
del interface aggregate-ethernet ae16 aggregated-ether-options lacp fallback
```

```
53. del interface aggregate-ethernet ae17 aggregated-ether-options lacp fallback
54. del interface aggregate-ethernet ae18 aggregated-ether-options lacp fallback
55. del interface aggregate-ethernet ae19 aggregated-ether-options lacp fallback
56. del interface aggregate-ethernet ae20 aggregated-ether-options lacp fallback
57. del interface aggregate-ethernet ae21 aggregated-ether-options lacp fallback
58. del interface aggregate-ethernet ae22 aggregated-ether-options lacp fallback
59. del interface aggregate-ethernet ae23 aggregated-ether-options lacp fallback
60. del interface aggregate-ethernet ae24 aggregated-ether-options lacp fallback
61. del interface aggregate-ethernet ae25 aggregated-ether-options lacp fallback
62. del interface aggregate-ethernet ae26 aggregated-ether-options lacp fallback
63. del interface aggregate-ethernet ae27 aggregated-ether-options lacp fallback
64. del interface aggregate-ethernet ae28 aggregated-ether-options lacp fallback
65. del interface aggregate-ethernet ae29 aggregated-ether-options lacp fallback
66. del interface aggregate-ethernet ae30 aggregated-ether-options lacp fallback
67. del interface aggregate-ethernet ae31 aggregated-ether-options lacp fallback
68. del interface aggregate-ethernet ae32 aggregated-ether-options lacp fallback
```

8.3.2 GPU Network Card Driver Installation

```
1. #!/bin/bash
2. export DEBIAN_FRONTEND=noninteractive
3. apt-get remove --purge nvidia-* -y
4. apt-get autoremove -y
5. apt-get update
6. add-apt-repository ppa:graphics-drivers/ppa -y
7. apt-cache madison nvidia-driver-535-server
8. apt-get update
9. apt-get install nvidia-driver-535-server -y
10. cat /proc/driver/nvidia/version
11. tar -xvf /opt/packages/MLNX_OFED_LINUX-23.10-0.5.5.0-ubuntu22.04-x86_64.tgz -C /opt/packages
12. /opt/packages/MLNX_OFED_LINUX-23.10-0.5.5.0-ubuntu22.04-x86_64/mlnxofedinstall -q
```

8.3.3 Synchronization of "Development Kit" Versions

8.3.3.1 Install Rsyslog

```
#!/bin/bash
export DEBIAN_FRONTEND=noninteractive
systemctl disable rsyslog
apt-get update -y
rm -f /etc/apt/trusted.gpg.d/adiscon-ubuntu-v8-stable.gpg
rm -f /etc/apt/sources.list.d/adiscon-ubuntu-v8-stable-jammy.list
apt-get remove --purge rsyslog -y
apt-get autoremove -y
apt-get update -y
apt-get install rsyslog -y
```

8.3.3.2 Automatically Format and Mount Disks in Order

1. NVME platform

```
#!/bin/bash
set -e
apt-get update
apt-get install parted -y
unmounted_nvme_disks=()
nvme_disks=$(lsblk -lno NAME,TYPE | awk '/nvme/ && /disk/ {print "/dev/" $1}')
for disk in $nvme_disks; do
    if ! mount | grep -q $disk; then
        unmounted_nvme_disks+=("$disk")
    fi
done
IFS=$'\n' sorted_unmounted_nvme_disks=($(printf "%s\n" "${unmounted_nvme_disks[@]}" | sort))
unset IFS
echo
printf "%s\n" "${sorted_unmounted_nvme_disks[@]}"
```

```
declare -A disk_map
data_counter=1
for disk in "${sorted_unmounted_nvme_disks[@]}"; do
    mount_point="/data${data_counter}"
    disk_map["$disk"]=$mount_point
    ((data_counter++))
 done
echo
for disk in "${!disk_map[@]}"; do
    echo "$disk -> ${disk_map[$disk]}"
done
for disk in "${!disk_map[@]}"; do
    echo "$disk${!disk_map[@]}"
        parted -s ${disk} mklabel gpt mkpart primary xfs 0% 100%
        mkfs.xfs -f ${disk}p1
        mkdir -p ${disk_map[$disk]}
        mount ${disk}p1 ${disk_map[$disk]}
        UUID=$(blkid -s UUID -o value ${disk}p1)
        echo "UUID=${UUID} ${disk_map[$disk]} xfs defaults 0 0" | tee -a /etc/fstab
        echo "$disk -> ${disk_map[$disk]}"
done
df -h
cat /etc/fstab
```

2. Automatic mounting of SATA hard disks

```
#!/bin/bash
set -e
apt-get update
apt-get install parted -y
unmounted_disks=()
disks=$(lsblk -lno NAME,TYPE | awk '/^sd/ && /disk/ {print "/dev/" $1}')

for disk in $disks; do
```

```
    if ! mount | grep -q $disk; then
        unmounted_disks+=("$disk")
    fi
done

IFS=$'\n' sorted_unmounted_disks=($(printf "%s\n" "${unmounted_disks[@]}" | sort))
unset IFS

echo
printf "%s\n" "${sorted_unmounted_disks[@]}"

declare -A disk_map
data_counter=1

for disk in "${sorted_unmounted_disks[@]}"; do
    mount_point="/data${data_counter}"
    disk_map["$disk"]=$mount_point
    ((data_counter++))
done

echo
for disk in "${!disk_map[@]}"; do
    echo "$disk -> ${disk_map[$disk]}"
done

for disk in "${!disk_map[@]}"; do
    echo "$disk${disk_map[$disk]}"
    parted -s ${disk} mklabel gpt mkpart primary xfs 0% 100%
    mkfs.xfs -f ${disk}1
    mkdir -p ${disk_map[$disk]}
    mount ${disk}1 ${disk_map[$disk]}
    UUID=$(blkid -s UUID -o value ${disk}1)
    echo "UUID=${UUID} ${disk_map[$disk]} xfs defaults 0 0" | tee -a /etc/fstab
```

```
    echo "$disk -> ${disk_map[$disk]}"
done

df -h
cat /etc/fstab
```

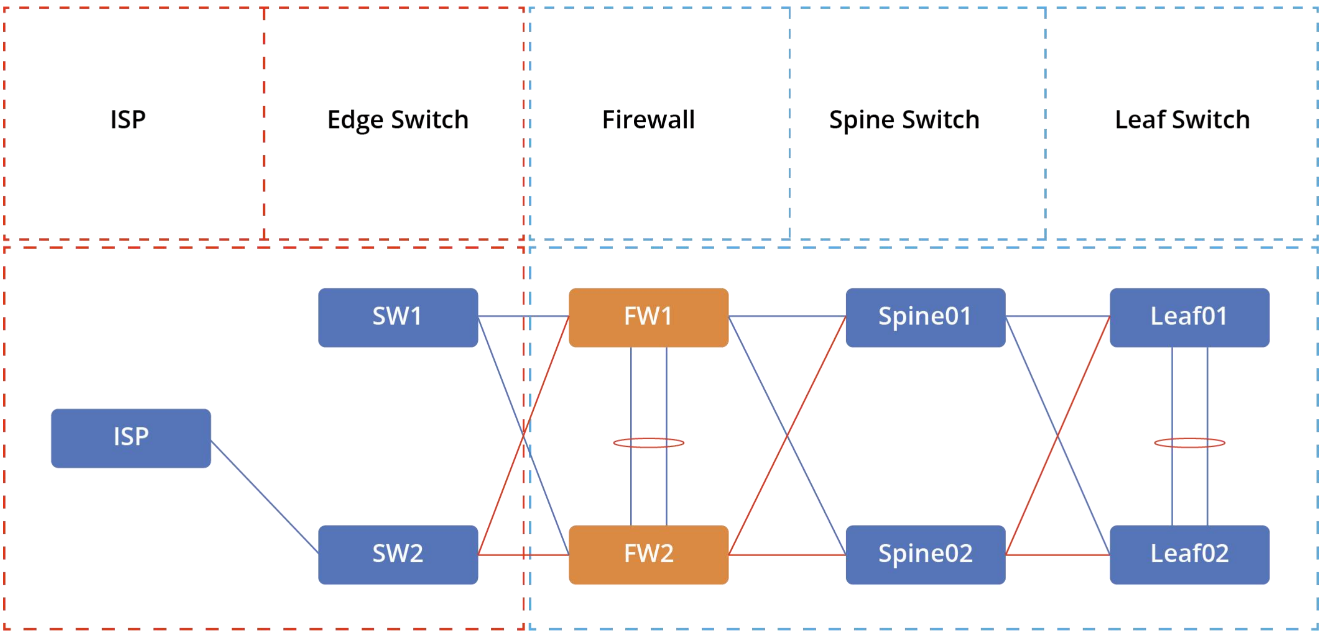
3. Automatic mounting of SATA hard disks

```
#!/bin/bash
apt-get update
DEBIAN_FRONTEND=noninteractive apt-get install linux-image-5.15.0-91-generic linux-headers-5.15.0-91-generic -y
dpkg --get-selections | grep linux-image
sed -i 's|^GRUB_DEFAULT=.*|GRUB_DEFAULT="Ubuntu, with Linux 5.15.0-91-generic"|' /etc/default/grub
update-grub
```

9. Firewall Network Debugging

9.1 Network Topology Description

9.1.1 Example of Solution Topology



9.1.2 Solution Topology Description

- Leaf01 and Leaf02 are interconnected with Server via MLAG, and Leaf and Spine are interconnected via BGP.
- Spine01 and Spine02 are interconnected with FW1 and FW2 respectively, with blue color as the primary link and red color as the backup link. When the primary link fails, the traffic goes to the backup link.
- FW1 and FW2 are connected to SW1 and SW2 respectively, and SW1 and SW2 are connected to different links of ISPs. 2 different IP addresses are configured on the outgoing interfaces of FW1 and FW2. The red color is the backup link. When the upstream and downstream links of FW1 fail, all traffic is switched to the backup wall.

9.2 Network Topology Description

The information on the form is for reference only. Actual situation is subject to on-site implementation.

Device	VLAN Number	Interface	Interface Description	IP Address/Mask	GW
Spine1	4000		Spine1_fw1	10.64.126.2/25	active：10.64.126.1
			Spine1_fw2		
Spine2	4000		Spine2_fw1	10.64.126.130/25	active：10.64.126.129
			Spine2_fw2		
FW1	NA		fw1_Spine1	10.64.126.1/25	NA
	NA		fw1_Spine2	10.64.126.129/25	
	4001		fw1_isp1	Adjusted for actual conditions on site	
	4002		fw2_isp2		
	HA				
FW2	NA		fw2_Spine1	10.64.126.1/25	NA
	NA		fw2_Spine2	10.64.126.129/25	
	4001		fw1_isp1	Adjusted for actual conditions on site	
	4002		fw2_isp2		
IP Allocation	HA Standby		NA	192.168.10.1	
	HA Primary			192.168.10.2	
	SSL VPN client			10.5.0.0/24	
	SSL VPN Internal Address			10.1.0.0/16, 10.2.0.0/16, 10.3.0.0/16, 10.4.0.0/16, 10.5.0.0/16	

9.3 Security Policy Planning

Traffic Direction	Service	Source IP	Destination IP	Port Number	Policy
inbond	SSL VPN	any	Firewall WAN port IP	443	permit
		SSL VPN Address Pool	Access service	Adjusted for actual conditions on site	permit
	OpenVPN	any	Adjusted for actual conditions on site		permit
	JumpServer	any			permit
outbond	any	any	any	any	permit

9.4 Policy Detailed Configuration

9.4.1 Firewall Internet Configuration

Connecting to the network in routing mode.

1. **Step 1: Connect the device**

- Connect the interface of the device (e.g., ethernet0/1) to the egress switch. In general, the external interface is selected to belong to the untrust security domain.
- Connect the other interface of the device (e.g., ethernet0/0) to the intranet. By default, ethernet0/0 belongs to the TRUST security domain.

2. **Step 2: Configure the interface**

- Select "Network > Interfaces".
- Double-click the ethernet0/1 interface.
- Double-click the ethernet0/1 interface. Click the OK button to complete the configuration.
- By default, the ethernet0/0 interface belongs to the trust security domain and has the IP address X.X.X.X/24, so there is no need to configure it again.

3. **Step 3: Configure the source NAT rule to convert the inside IP address to the outside IP address.**

- Select "Policy > NAT > Source NAT".
- Select "Policy > NAT > Source NAT". Click the "New" button.
- Configure the source security zone, destination security zone, inbound traffic, outbound traffic, and services.
- Click the "OK" button to complete the configuration.

4. **Step 4: Configure policy rules to allow internal users to access the external network.**

- Select "Policy > Security Policy > Policies".
- Click the "New" button on the upper left corner of the list and select "Policy".
- Configure the source security zone, destination security zone, source IP address, source destination IP address, and service.
- Click the "OK" button to complete the configuration.

5. **Step 5: Configure the default route.**

- Select "Network > Routing > Destination Route". Click "New" to complete the configuration.
- Click the "New" button.
- Select "Network > Routing > Destination Route". Click the "New" button. Configure the default route with destination address and subnet mask as "0.0.0.0" and the next hop as the IP address of the Internet.
- Click "OK" button to complete the configuration.

Please refer to pages 22-25 of the document: NSG Series Firewalls WebUI Configuration Guide

9.4.2 HA Active-Active Deployment

Select "System > HA" to enter the HA configuration page.

Specify the working mode as "Active-Passive" to indicate that the two devices in the HA cluster work in active mode and backup mode respectively.

1. Device A

Select "System > HA".

- Working Mode: Active-Passive
- HA control connection interface 1: ethernet0/4
- HA control connection interface 2: ethernet0/8
- IP address: 1.1.1.1/24
- HA Cluster ID: 1
- Node ID: 0
- HA Group Configuration: Set the priority to 10 and the monitoring object to track1.
- Click the OK button to complete the configuration.

2. Device B

Select "System > HA".

- Working Mode: Active-Passive
- HA control connection interface 1: ethernet0/4
- HA control connection interface 2: ethernet0/8
- IP address: 1.1.1.2/24
- HA Cluster ID: 1
- Node ID: 1
- HA Group Configuration: Set the priority to 100 and the monitoring object to track1.
- Click the OK button to complete the configuration.

Please refer to pages 530-533 of the document: NSG Series Firewalls WebUI Configuration Guide

9.4.3 Security Policy Deployment

1. Check the license and signature database information

Please refer to pages 558-560 of the document: NSG Series Firewalls WebUI Configuration Guide

Step 1: Enable abnormal behavior detection for intranet user hosts.

Step 2: Configure the core asset objects. Specify the Web server and Mail server as the core asset objects for abnormal behavior detection.

2. Blocking attacks on the intranet from the extranet through intrusion prevention

Step 2: Install the Intrusion Prevention License.

- Select "System > Device Management > Settings and Operations", click "Reboot Device", and the license will take effect after reboot.

Step 2: Enable the intrusion prevention function and upgrade the feature library.

- Select "Objects > Intrusion Prevention > Configuration > Upload License", check whether the intrusion prevention function is enabled, if not, check the "Enable" checkbox and reboot the device.
- Select "System>Upgrade Management>Feature Database Upgrade", and in the section of Intrusion Prevention Feature Database Upgrade, click "Online Upgrade Now" to upgrade the intrusion prevention feature database, so as to ensure the integrity and correctness of features.

Please refer to pages 560-561 of the document: NSG Series Firewalls WebUI Configuration Guide

Step 3: Bind the internal and external interfaces to the corresponding security domains.

- Bind the internal network interface ethernet0/2 to "trust". Select "Network > Security Zones", choose "trust" and click "Edit" to enter the "Security Zone Configuration" dialog box. Bind the interface: ethernet0/2.
- Use the same method to bind the internal interface ethernet0/1 to "dmz".
- Use the same method to bind the external interface ethernet0/3 to "untrust".

3. Blocking Attacks on the Intranet from Outside Networks through Intrusion Prevention

Step 4: Create intrusion prevention rules.

- Users can use the default intrusion prevention rules or create their own rules. Select "Objects> Intrusion Prevention>Templates", click "New" button to customize the Intrusion prevention rules. In this case, we use the system default intrusion prevention rule predef_default, which contains all IPS features and the processing behavior is reset.

Please refer to pages 424-444 of the document: NSG Series Firewalls WebUI Configuration Guide

Step 5: Create a security policy.

By default, the device denies information transmission between all security domains.

In this case, to allow the extranet and data center office to access the data center server farm and to invoke the default intrusion prevention rules, the following two need to be configured

Security policy: untrust to dmz.

- Select "Policy>Security Policy", click the "New" button to create a new security policy, and enter the "Basic Configuration" tab.

Create a security policy

- Select the "Protection Status" tab.
- Intrusion Prevention: Check the "Enable" checkbox.
- Template: predef_default

Security Policy: trust to dmz (same as above)

Please refer to pages 37-39 of the document: NSG Series Firewalls WebUI Configuration Guide

9.4.4 SSL VPN

Step 1: Create a local user.

Select "Objects > Users > Local Users", select the local server "local" and click "New > User".

- Name: test1
- Password: 123456
- Re-enter the password: 123456

Step 2: Configure SCVPN address pool.

Select "Network > VPN > SSL VPN" and click "Related Configuration > SSL VPN Address Pool" on the top right of the page.

- On the <Address Pool> page, click "New" to create an address pool entry.

Step 3: Create a Security Zone.

Select "Network > Security Domain" and click "New".

- Security Zone Name: VPN
- Type: Layer 3 Security Zone
- Virtual Router: trust-vr

Step 4: Create a tunnel interface.

Select "Network > Interfaces" and click "New > Tunnel Interface".

Note: The IP address of the tunnel interface must be in the same network segment as the IP address of the SSL VPN address pool.

Step 5: Configure the SCVPN device side.

Select "Network > SSL VPN" and click "New".

In the <Name/Access User> tab, do the following configuration:

- SSL VPN name: scvpn
- Access User: Click "New" and select "local" in the drop-down menu of "AAA Server".
- In the <Access Interface/Tunnel Interface> tab, do the following configuration:
- Outgoing interface: ethernet0/3
- Outgoing interface: ethernet0/3 Service port: 4433
- Tunnel interface: tunnel1
- Address pool: scpool

In the <Tunnel Routing Configuration> tab, click "New" and do the following configuration

- IP: X.X.X.X
- Subnet mask: 255.255.255.0

The IP address of the tunnel route must be in the same network segment as the IP address of the server.

Click the <Advanced Configuration> tab and do the following configuration in the <Parameter Configuration> tab:

- Advanced Parameters
- Data Port (TCP): 10000

After the configuration of the device side is completed, you need to enable the "Communication Stability Optimization" function of the client side at the same time, so that the client side of the PC can communicate with the device side through the TCP protocol.

Click the <Advanced Configuration> tab, and make the following configurations in the <Secondary Authentication> tab:

- SMS password authentication
- SMS gateway
- SMS authentication valid time
- Authentication code length

Step 6: Create a policy rule from VPN to Any.

Select "Policy > Security Policy" and click "New".

Please refer to pages 191-204 of the document: NSG Series Firewalls WebUI Configuration Guide

Step 7: Acceptance Results

Download SSL VPN Client: https://resource.fs.com/mall/file/user_manual/fs-secure-connect-v1.4.9.1275-for-windows-os.zip

9.4.5 IPsec VPN

In the configuration of IPSec VPN before you need to ensure that the two ends of the public network is interoperable, you can ping the interface IP of FW2 on the firewall wall of FW1, ping the interface IP of FW1 on the firewall wall of FW2 to test, provided that the interfaces of the two ends of the open ping management, to confirm that there are no problems in the public network environment, you can begin to configure the IPSec VPN①

> Please refer to pages 168-177 of the document: NSG Series Firewalls WebUI Configuration Guide

[FW1 Configuration]

1. Click Network-VPN-IPSec VPN-New to enter the IPSec VPN configuration page:
2. Click the "VPN Peer List" option, click the green plus sign "New" at the back of the selection box, and enter the "VPN Peer List" configuration, fill in all the required fields (Name/Interface/Interface/Interface/Interface), and enter the "VPN Peer List" configuration. Configuration, fill in all the required fields (Name/Interface/PeerIP/Proposal (the configuration of both ends of the firewall must be consistent)/Pre Share Key).

(When multiple IPsec VPN tunnels are used, configure multiple Peer Lists based on different source interfaces)

- Name: specifies the name of the ISAKMP gateway;
- Interface: specify the binding interface of the ISAKMP gateway;
- Authentication Mode: the default is the primary mode, and this configuration case is also configured according to the primary mode;
- Type: default is static IP, this configuration case is also configured according to static IP;
- Peer address: configure the public network address of the peer;
- Proposal: Specify the P1 proposal of the tunnel, just keep the same on both ends, usually only need to configure one proposal.
- Pre-shared key: fill in the encrypted password, and keep the same at both ends;
- <Advanced Configuration NAT traversal: if there is NAT on the upper layer of the device, you need to enable this function;
- DPD: Configure the DPD (secure tunnel peer state detection) function. By default, this function is disabled, click the "Enable" button to enable the function. When this function is enabled, the system will periodically send request messages to the other end according to the specified time interval to detect whether the ISAKMP gateway exists or not.

① DPD Interval - Specifies the time interval in seconds for sending DPD query requests to the peer. The range of values is 1 to 10 seconds, and the default value is 10 seconds.

② DPD Retry - Specifies the number of times to send a DPD query request to the peer. After sending a query request to the other end, if this end does not receive a message from the other end within the specified time interval, the system will send the query request again, and so on until the number of times specified in this parameter is completed. After the specified number of queries is completed, if the other end still cannot receive the message, then the ISAKMP gateway at the other end is judged to be down. The number of query requests ranges from 1 to 10, and the default is 3.

3. After the configuration of "VPN Peer List" is completed, click OK. Enter the "IKE VPN List" configuration, click the green plus sign "New" at the back of the selection box that appears, and complete the "IKE VPN List" configuration. For multiple IPsec VPN tunnels, you need to configure multiple IKE VPNs):

- Peer Options: Select the VPN peer we just created;
- Name: Specify the name of the tunnel;
- P2 protocol: specify the P2 proposal of the tunnel, just keep the same on both ends;
- Proxy ID: fill in the interested stream, if both devices are FS firewalls, you can choose Auto;
- <Advanced Configuration Auto-connect: configure the auto-connect function. By default, this function is disabled, click the "Enable" button to enable the function. The device provides two ways to trigger the establishment of SA: automatic and traffic-triggered. In the automatic mode, the device checks the status of SA every 60 seconds, and automatically initiates a negotiation request if the SA is not established; in the traffic-triggered mode, the negotiation request is initiated when there is data traffic to be transmitted through the tunnel. By default, the system uses the traffic trigger method. It is recommended to check this box.

[FW2 configuration]

1. FW2 configuration is consistent with that of FW1:
2. Click IPSec vpn monitoring (which may be different for different versions) in the upper right corner of the relevant configuration after the configuration is complete to check whether both ISAKMP SA and IPSec SA are up, mainly focusing on the status information of the IPSec SA phase:

Tunnel Interface and Routing Configuration

After confirming that the IPSec VPN configuration is complete and is in the state of normal establishment, you need to establish the tunnel interface.

[FW1 Configuration]

1. Click network-interface-New-Tunnel interface (multiple IPsec VPNs need to bind multiple Tunnel):
2. Configure routing, click network-routing-Destination-New, and select the new tunnel interface we just created as the next hop. 3:
3. Configure a default route (multiple IPsec VPNs need to configure multiple static routes for switching), the default route 0.0.0.0/0 to the outside network of the other end, the next hop is the carrier's IP address.

[FW2 Configuration]

The configuration of FW2 is consistent with that of FW1.

Security Policy Configuration

Click Policy-Security Policy-New. In order to realize mutual access, you need to establish a release security policy.

NAT Configuration (optional, configured when accessing ISP)

10. Opening Acceptance

According to the on-site environment, formulate an acceptance plan and confirm the acceptance criteria.

The customer's single-point network should have high availability to ensure a fast failover and recovery mechanism in the event of a network failure or single-point failure.

10.1 Optical Module Connectivity Testing

Note

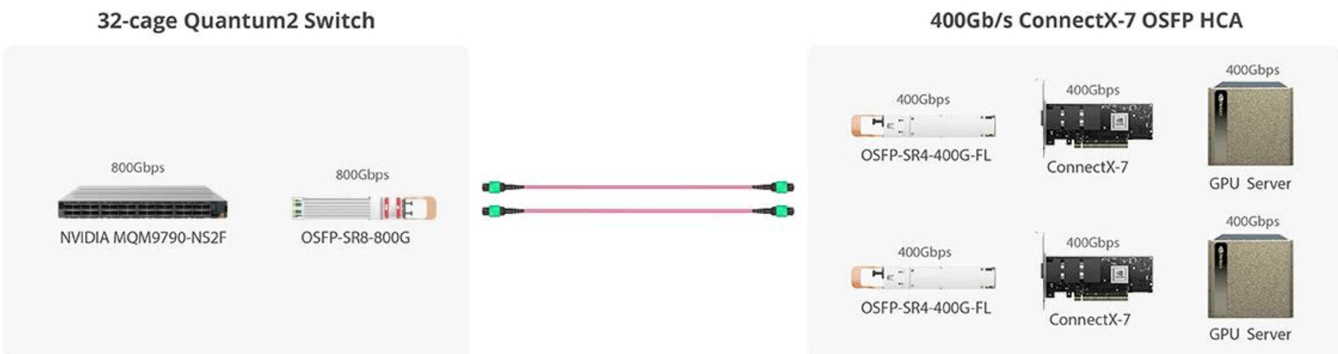
Before performing the compatibility test, please make sure that the optical modules and patch cords have been inspected and cleaned (refer to 3.5 for details) and plug-in and pull-out tests have been performed. (For details, please refer to 3.6)

In addition, when interconnecting two optical modules, please make sure that both 400G/800G OSFP optical modules are of the same model name and parameter specifications. (The electrical chips are both 5nm or both 7nm.)

10.1.1 Preparation Phase

- Download and install the SecureCRT software (as shown below on the right, you can download it yourself) and prepare the control cable (as shown below on the left);

- Prepare at least 1 Nvidia 9790 switch, 2 ME MCX75310AAS-NEAT NICs, a server or host computer (for loading X-7 NICs), and a PC;
- 1 800G SR8 module, 2 400G SR4 modules, and 2 MTP-12 APC patch cables;
- Set up the test environment in the manner shown:



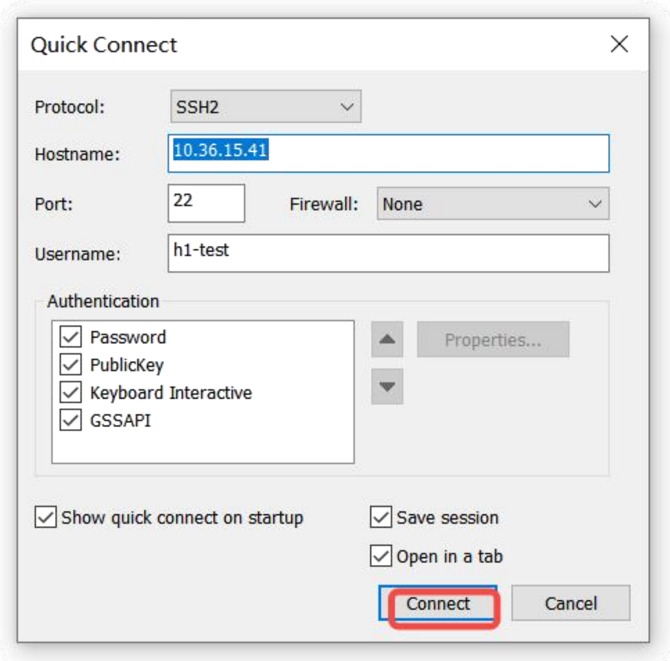
10.1.2 Compatibility Test for Devices without Console Port/Management Port

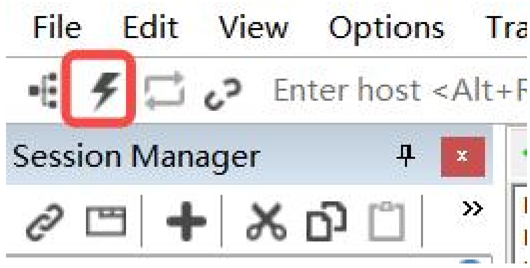
1. Log in to the server:

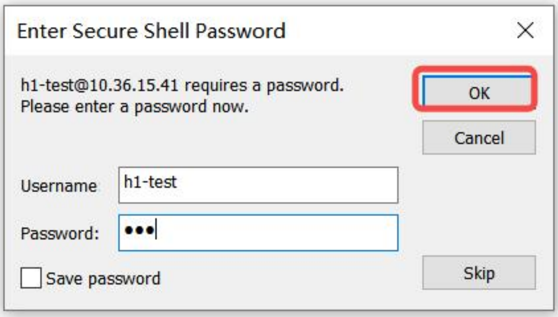
- As shown in the figure: Select the CRT software and select IP: 10.36.15.41 and IP: 10.36.15.45 for SSH access to the server;

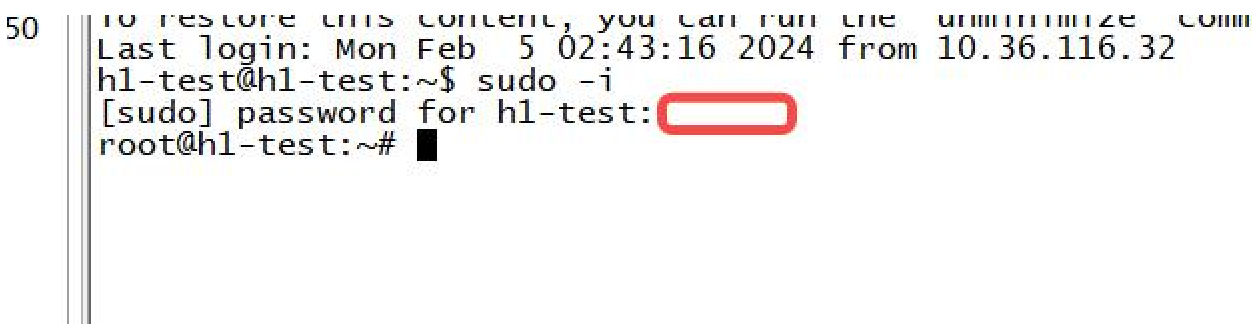
  Account: h1-test/h2-test

  Password: 123/123









2. After the module is connected, check the status of the device LEDs, and use test commands to check the module port information, module connectivity information, module DDM information, module Type type, and so on; refer to the related test commands:

1. [root@localhost ~]# SUdo ibswitches // Check lid number
2. [root@localhost ~]# SUdo mlxlink -d lid-15 -p 1/1 -c // Read switch port information
3. [root@localhost ~]# SUdo mlxlink -d lid-15 -p 1/1 -m // Read switch DDM
4. [root@localhost ~]# mlxlink -d mlx5_0 -c // Query module information on the network card
5. [root@localhost ~]# SUdo ib_write_bw -d mlx5_0 -D 160 10.36.15.95 --report_gbits

6. [root@localhost ~]# SUdo mlxlink -d lid-16 -p 12/2 --show_device // view device info
7. [root@localhost ~]# SUdo iblinkinfo -S 0xfc6a1c03009a2640 // check all port statuses
8. root@h2test:~# ibtracert 1 15 // View packet flow on port 1 of Lid-15 switch.
9. [root@localhost ~]# SUdo ib_write_bw -d lid-15 -D 160 10.36.15.95 --report_gbits // Streaming commands
10. [root@localhost ~]# ibqueryerrors -S 0xfc6a1c03009b24c0 // See how many times errors have been counted.
11. [root@localhost ~]#SUdo ibqueryerrors --clear-errors // clear error counts

- Device LED status after module connection

- Rate and Link situation when the module is connected

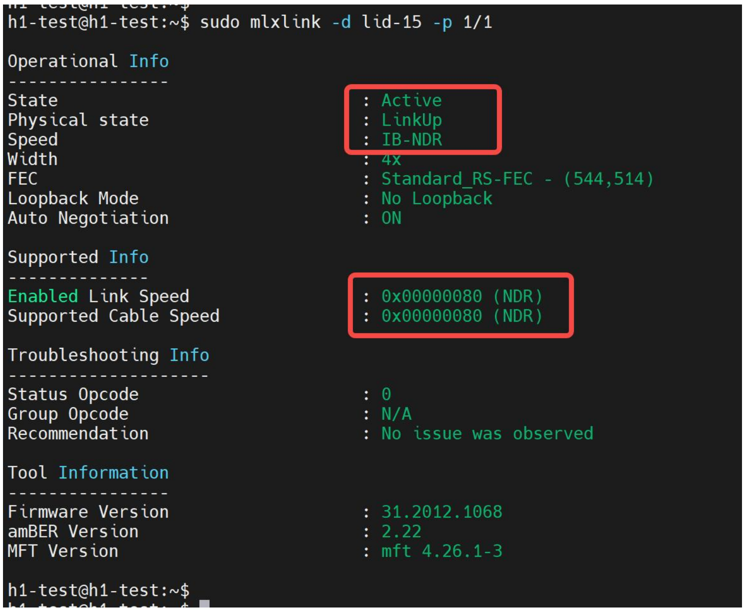
  State: Active

  Physical state: LinkUp

  Speed: IB-NDR

  Enable Link Speed: NDR

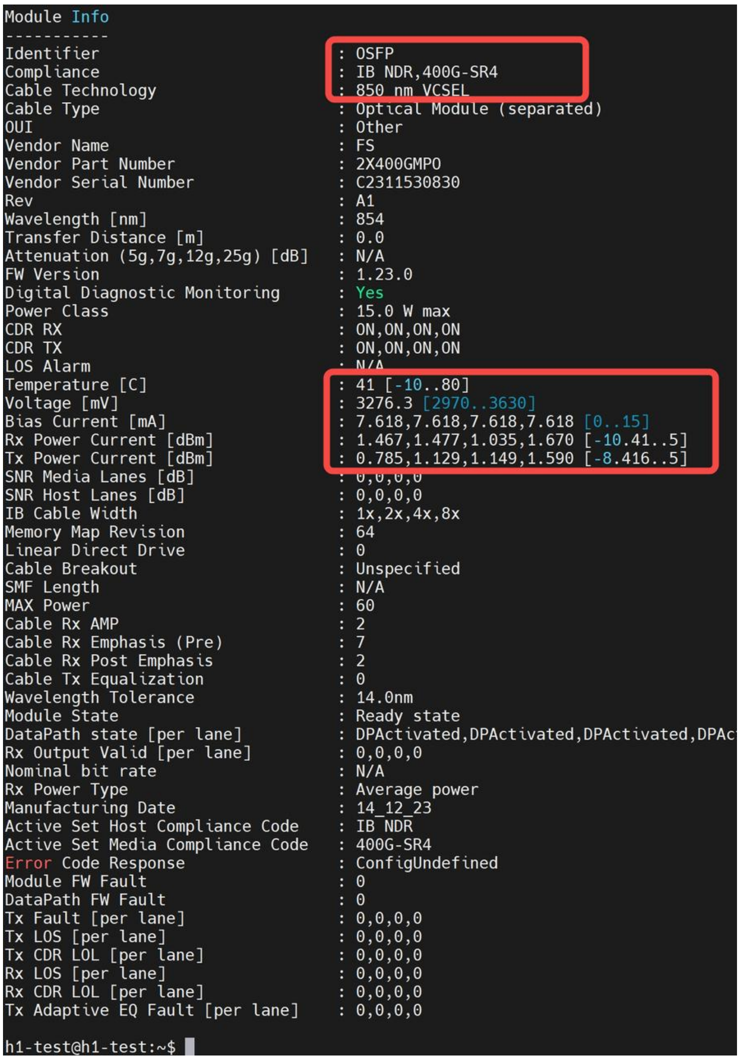
  Supported Cable Speed: NDR



- Basic module information (wavelength, Type type):

  check whether the basic information on both ends of the optical module is the same, module power consumption, DDM information (temperature, voltage, bias current, TX/RX Power)

  check whether the DDM parameters are within normal limits



```
Module Info
-----------
Identifier                          : OSFP
Compliance                          : IB NDR,400G-SR4
Cable Technology                    : 850 nm VCSEL
Cable Type                          : Optical Module (separated)
OUI                                 : Other
Vendor Name                         : FS
Vendor Part Number                  : 2X400GMPO
Vendor Serial Number                : C2311530830
Rev                                 : A1
Wavelength [nm]                     : 854
Transfer Distance [m]               : 0.0
Attenuation (5g,7g,12g,25g) [dB]    : N/A
FW Version                          : 1.23.0
Digital Diagnostic Monitoring       : Yes
Power Class                         : 15.0 W max
CDR RX                              : ON,ON,ON,ON
CDR TX                              : ON,ON,ON,ON
LOS Alarm                           : N/A
Temperature [C]                     : 41 [-10..80]
Voltage [mV]                        : 3276.3 [2970..3630]
Bias Current [mA]                   : 7.618,7.618,7.618,7.618 [0..15]
Rx Power Current [dBm]              : 1.467,1.477,1.035,1.670 [-10.41..5]
Tx Power Current [dBm]              : 0.785,1.129,1.149,1.590 [-8.416..5]
SNR Media Lanes [dB]                : 0,0,0,0
SNR Host Lanes [dB]                 : 0,0,0,0
IB Cable Width                      : 1x,2x,4x,8x
Memory Map Revision                 : 64
Linear Direct Drive                 : 0
Cable Breakout                      : Unspecified
SMF Length                          : N/A
MAX Power                           : 60
Cable Rx AMP                        : 2
Cable Rx Emphasis (Pre)             : 7
Cable Rx Post Emphasis              : 2
Cable Tx Equalization               : 0
Wavelength Tolerance                : 14.0nm
Module State                        : Ready state
DataPath state [per lane]           : DPActivated,DPActivated,DPActivated,DPAc
Rx Output Valid [per lane]          : 0,0,0,0
Nominal bit rate                    : N/A
Rx Power Type                       : Average power
Manufacturing Date                  : 14_12_23
Active Set Host Compliance Code     : IB NDR
Active Set Media Compliance Code    : 400G-SR4
Error Code Response                 : ConfigUndefined
Module FW Fault                     : 0
DataPath FW Fault                   : 0
Tx Fault [per lane]                 : 0,0,0,0
Tx LOS [per lane]                   : 0,0,0,0
Tx CDR LOL [per lane]               : 0,0,0,0
Rx LOS [per lane]                   : 0,0,0,0
Rx CDR LOL [per lane]               : 0,0,0,0
Tx Adaptive EQ Fault [per lane]     : 0,0,0,0

h1-test@h1-test:~$
```

10.2 Storage Network Simulation Testing

10.2.1 Test Environment Description

GPU nodes and storage nodes are dual uplink 100G NICs, with a maximum throughput of 2*100G per node.

GPU storage Leaf - Storage Spine is 8*100G uplinks

Storage Leaf - Storage Spine is 32*100G uplink.

In this test environment, with all six units running at full capacity, the theoretical rate of 6*150=900G should be reached or close to it, but this is just a theory, if it can be close to 700-800G, it will be fine. Observe the port counts in the storage Leaf and add up the final storage server iperf test results.

The test software is Iperf2, all nodes will have iperf2 installed, which is 384KB in size.

10.2.2 Test Result

- Test Host-GPU Node: 10.3.6.1
- Test Host-Storage Node: 10.3.210.19

GPU Test Throughput Results

```
[SUM] 29.0000-30.0000 sec  20.3 GBytes   175 Gbits/sec
[ 10] 0.0000-30.0163 sec  57.2 GBytes  16.4 Gbits/sec
[  1] 0.0000-30.0162 sec  34.5 GBytes  9.88 Gbits/sec
[  5] 0.0000-30.0168 sec  53.8 GBytes  15.4 Gbits/sec
[  6] 0.0000-30.0168 sec  34.0 GBytes  9.72 Gbits/sec
[  9] 0.0000-30.0165 sec  85.0 GBytes  24.3 Gbits/sec
[  7] 0.0000-30.0165 sec  54.8 GBytes  15.7 Gbits/sec
[  4] 0.0000-30.0159 sec  51.9 GBytes  14.9 Gbits/sec
[  8] 0.0000-30.0165 sec  57.9 GBytes  16.6 Gbits/sec
[  2] 0.0000-30.0162 sec  87.2 GBytes  25.0 Gbits/sec
[  3] 0.0000-30.0167 sec  87.7 GBytes  25.1 Gbits/sec
[SUM] 0.0000-30.0004 sec   604 GBytes   173 Gbits/sec
[ CT] final connect times (min/avg/max/stdev) = 0.422/0.579/0.717/0.090 ms (tot/err) = 10/0
fs@h161-gpu-polaris:~/scripts$
```

Storage Test Throughput Results

```
fs@s19-x86-polaris:~$ iperf -s
------------------------------------------------------------
Server listening on TCP port 5001
TCP window size:  128 KByte (default)
------------------------------------------------------------
[  1] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41480
[  2] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41486
[  3] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41540
[  4] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41498
[  5] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41554
[  6] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41524
[  7] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41552
[  8] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41570
[  9] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41510
[ 10] local 10.3.210.19 port 5001 connected with 10.3.6.1 port 41500
[ ID] Interval       Transfer     Bandwidth
[ 10] 0.0000-29.9979 sec  87.7 GBytes  25.1 Gbits/sec
[  2] 0.0000-30.0025 sec  51.9 GBytes  14.9 Gbits/sec
[  4] 0.0000-29.9998 sec  53.8 GBytes  15.4 Gbits/sec
[  6] 0.0000-29.9990 sec  85.0 GBytes  24.3 Gbits/sec
[  8] 0.0000-29.9996 sec  54.8 GBytes  15.7 Gbits/sec
[  1] 0.0000-30.0228 sec  87.2 GBytes  25.0 Gbits/sec
[  3] 0.0000-30.0014 sec  34.5 GBytes  9.89 Gbits/sec
[  7] 0.0000-29.9994 sec  34.0 GBytes  9.73 Gbits/sec
[  5] 0.0000-29.9995 sec  57.2 GBytes  16.4 Gbits/sec
[  9] 0.0000-29.9978 sec  57.9 GBytes  16.6 Gbits/sec
[SUM] 0.0000-30.0223 sec   604 GBytes   173 Gbits/sec
```

Test throughput results between Leaf-Spine

```
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 7080 bits/sec, 6 packets/sec
  5 sec output rate 24928192 bits/sec, 43928 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 60352 bits/sec, 6 packets/sec
  5 sec output rate 91935304 bits/sec, 163719 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 459864 bits/sec, 43 packets/sec
  5 sec output rate 49897256 bits/sec, 88588 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 78084743168 bits/sec, 6430569 packets/sec
  5 sec output rate 103240 bits/sec, 21 packets/sec
root@STOR-STOR01-LEAF01>
```

```
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 46987804672 bits/sec, 3869352 packets/sec
  5 sec output rate 30619368 bits/sec, 54104 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 50512789504 bits/sec, 4159844 packets/sec
  5 sec output rate 61222344 bits/sec, 108293 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 824 bits/sec, 0 packets/sec
  5 sec output rate 35439408 bits/sec, 62987 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 10680 bits/sec, 10 packets/sec
  5 sec output rate 8096 bits/sec, 11 packets/sec
root@STOR-STOR01-LEAF02>
```

- Test Host-GPU Node: 10.3.6.2
- Test Host-Storage Node: 10.3.210.20

GPU Test Throughput Results

```
[  9] 0.0000-30.0138 sec  65.2 GBytes  18.7 Gbits/sec
[  2] 0.0000-30.0134 sec  62.0 GBytes  17.7 Gbits/sec
[ 10] 0.0000-30.0129 sec  70.5 GBytes  20.2 Gbits/sec
[  5] 0.0000-30.0136 sec  55.5 GBytes  15.9 Gbits/sec
[  6] 0.0000-30.0139 sec  67.4 GBytes  19.3 Gbits/sec
[  8] 0.0000-30.0141 sec  38.6 GBytes  11.1 Gbits/sec
[  3] 0.0000-30.0131 sec  63.4 GBytes  18.2 Gbits/sec
[  7] 0.0000-30.0139 sec  65.6 GBytes  18.8 Gbits/sec
[  4] 0.0000-30.0137 sec  55.8 GBytes  16.0 Gbits/sec
[  1] 0.0000-30.0297 sec  42.6 GBytes  12.2 Gbits/sec
[SUM] 0.0000-30.0013 sec   587 GBytes   168 Gbits/sec
[ CT] final connect times (min/avg/max/stdev) = 0.391/0.550/0.694/0.093 ms (tot/err) = 10/0
fs@h162-gpu-polaris:~$
```

Storage Test Throughput Results

```
fs@s20-x86-polaris:~$ iperf -s
------------------------------------------------------------
Server listening on TCP port 5001
TCP window size:  128 KByte (default)
------------------------------------------------------------
[  1] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49022
[  2] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49040
[  3] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49038
[  4] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49064
[  5] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49072
[  7] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49094
[  8] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49110
[  9] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49078
[ 10] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49120
[  6] local 10.3.210.20 port 5001 connected with 10.3.6.2 port 49052
[ ID] Interval       Transfer     Bandwidth
[  9] 0.0000-29.9991 sec  65.6 GBytes  18.8 Gbits/sec
[  2] 0.0000-30.0021 sec  63.4 GBytes  18.2 Gbits/sec
[  4] 0.0000-29.9997 sec  55.8 GBytes  16.0 Gbits/sec
[  5] 0.0000-29.9993 sec  55.5 GBytes  15.9 Gbits/sec
[  8] 0.0000-29.9992 sec  70.5 GBytes  20.2 Gbits/sec
[  3] 0.0000-30.0012 sec  62.0 GBytes  17.7 Gbits/sec
[  7] 0.0000-29.9993 sec  65.2 GBytes  18.7 Gbits/sec
[  6] 0.0000-29.9992 sec  67.4 GBytes  19.3 Gbits/sec
[ 10] 0.0000-29.9979 sec  38.6 GBytes  11.1 Gbits/sec
[  1] 0.0000-30.0452 sec  42.6 GBytes  12.2 Gbits/sec
[SUM] 0.0000-30.0452 sec   587 GBytes   168 Gbits/sec
```

Test Throughput Results between Leaf-Spine

```
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 1448 bits/sec, 1 packets/sec
  5 sec output rate 161008 bits/sec, 21 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 62336 bits/sec, 7 packets/sec
  5 sec output rate 96235632 bits/sec, 171049 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 49584672768 bits/sec, 4083692 packets/sec
  5 sec output rate 54866816 bits/sec, 96615 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 38486237184 bits/sec, 3169743 packets/sec
  5 sec output rate 109240 bits/sec, 27 packets/sec
root@STOR-STOR01-LEAF01>
```

```
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 31814692864 bits/sec, 2619882 packets/sec
  5 sec output rate 51083104 bits/sec, 90534 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 64963571712 bits/sec, 5349548 packets/sec
  5 sec output rate 95488736 bits/sec, 169056 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 832 bits/sec, 0 packets/sec
  5 sec output rate 148984 bits/sec, 46 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 5168 bits/sec, 6 packets/sec
  5 sec output rate 11312 bits/sec, 15 packets/sec
root@STOR-STOR01-LEAF02>
```

- Test Host-GPU Node: 10.3.7.1
- Test Host-Storage Node: 10.3.210.21

GPU Test Throughput Results

```
[SUM] 29.0000-30.0000 sec  20.4 GBytes   175 Gbits/sec
[ 16] 0.0000-30.0029 sec  23.6 GBytes  6.77 Gbits/sec
[  7] 0.0000-30.0191 sec  32.3 GBytes  9.23 Gbits/sec
[ 17] 0.0000-30.0189 sec  31.3 GBytes  8.96 Gbits/sec
[ 18] 0.0000-30.0188 sec  23.5 GBytes  6.72 Gbits/sec
[  8] 0.0000-30.0023 sec  50.7 GBytes  14.5 Gbits/sec
[  2] 0.0000-30.0025 sec  29.8 GBytes  8.52 Gbits/sec
[  5] 0.0000-30.0190 sec  35.4 GBytes  10.1 Gbits/sec
[ 19] 0.0000-30.0027 sec  45.9 GBytes  13.1 Gbits/sec
[  1] 0.0000-30.0027 sec  29.8 GBytes  8.52 Gbits/sec
[ 12] 0.0000-30.0192 sec  15.0 GBytes  4.30 Gbits/sec
[  4] 0.0000-30.0030 sec  41.3 GBytes  11.8 Gbits/sec
[ 15] 0.0000-30.0026 sec  35.6 GBytes  10.2 Gbits/sec
[  3] 0.0000-30.0026 sec  29.6 GBytes  8.47 Gbits/sec
[ 10] 0.0000-30.0026 sec  30.1 GBytes  8.62 Gbits/sec
[ 13] 0.0000-30.0030 sec  25.4 GBytes  7.26 Gbits/sec
[  6] 0.0000-30.0023 sec  29.1 GBytes  8.32 Gbits/sec
[ 11] 0.0000-30.0026 sec  30.2 GBytes  8.64 Gbits/sec
[ 20] 0.0000-30.0190 sec  28.9 GBytes  8.28 Gbits/sec
[ 14] 0.0000-30.0193 sec  28.8 GBytes  8.24 Gbits/sec
[  9] 30.0000-31.0189 sec   128 KBytes  1.03 Mbits/sec
[  9] 0.0000-31.0189 sec  14.5 GBytes  4.01 Gbits/sec
[SUM] 0.0000-30.0022 sec   611 GBytes   175 Gbits/sec
[ CT] final connect times (min/avg/max/stdev) = 0.437/0.552/0.705/0.074 ms (tot/err) = 20/0
fs@h193-gpu-polaris:~$
```

Storage Test Throughput Results

```
[ 11] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 52942
[ 12] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 52988
[ 13] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 52952
[ 14] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 52968
[ 15] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 52986
[ 16] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 52978
[ 17] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53008
[ 18] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53020
[ 19] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53016
[ 20] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53022
[ 21] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53056
[ 22] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53046
[ 23] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53072
[ 24] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53092
[ 25] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53112
[ 26] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53098
[ 27] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53086
[ 28] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53118
[ 29] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53116
[ 30] local 10.3.210.21 port 5001 connected with 10.3.7.1 port 53126
[ ID] Interval       Transfer     Bandwidth
[ 30] 0.0000-29.9884 sec  45.9 GBytes  13.2 Gbits/sec
[ 15] 0.0000-29.9994 sec  41.3 GBytes  11.8 Gbits/sec
[ 19] 0.0000-29.9968 sec  30.2 GBytes  8.64 Gbits/sec
[ 23] 0.0000-29.9944 sec  30.1 GBytes  8.62 Gbits/sec
[ 26] 0.0000-29.9922 sec  35.6 GBytes  10.2 Gbits/sec
[ 28] 0.0000-29.9910 sec  23.6 GBytes  6.77 Gbits/sec
[ 12] 0.0000-30.0019 sec  29.8 GBytes  8.52 Gbits/sec
[ 20] 0.0000-29.9959 sec  50.7 GBytes  14.5 Gbits/sec
[ 27] 0.0000-29.9916 sec  25.4 GBytes  7.26 Gbits/sec
[ 13] 0.0000-30.0012 sec  29.8 GBytes  8.52 Gbits/sec
[ 17] 0.0000-29.9985 sec  29.1 GBytes  8.32 Gbits/sec
[ 14] 0.0000-30.0006 sec  29.6 GBytes  8.47 Gbits/sec
[ 21] 0.0000-29.9992 sec  23.5 GBytes  6.73 Gbits/sec
[ 24] 0.0000-29.9941 sec  31.3 GBytes  8.97 Gbits/sec
[ 29] 0.0000-29.9922 sec  28.8 GBytes  8.24 Gbits/sec
[ 16] 0.0000-29.9997 sec  35.4 GBytes  10.1 Gbits/sec
[ 22] 0.0000-29.9991 sec  15.0 GBytes  4.31 Gbits/sec
[ 11] 0.0000-30.0459 sec  32.3 GBytes  9.22 Gbits/sec
[ 18] 0.0000-29.9987 sec  14.5 GBytes  4.15 Gbits/sec
[ 25] 0.0000-29.9942 sec  28.9 GBytes  8.29 Gbits/sec
[SUM] 0.0000-30.0463 sec   611 GBytes   175 Gbits/sec
```

Test Throughput Results between Leaf-Spine

```
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 224 bits/sec, 0 packets/sec
  5 sec output rate 19943688 bits/sec, 34997 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 408 bits/sec, 0 packets/sec
  5 sec output rate 57033264 bits/sec, 101180 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 97776541696 bits/sec, 8051626 packets/sec
  5 sec output rate 23064920 bits/sec, 40720 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 6338392 bits/sec, 597 packets/sec
  5 sec output rate 57185224 bits/sec, 101401 packets/sec
root@STOR-STOR01-LEAF01>
```

```
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 3984216 bits/sec, 374 packets/sec
  5 sec output rate 41919528 bits/sec, 74485 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 12459998208 bits/sec, 1026030 packets/sec
  5 sec output rate 21798744 bits/sec, 38573 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 552 bits/sec, 0 packets/sec
  5 sec output rate 83651928 bits/sec, 148055 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 56655294464 bits/sec, 4665339 packets/sec
  5 sec output rate 14128 bits/sec, 20 packets/sec
root@STOR-STOR01-LEAF02>
^C
admin@STOR-STOR01-LEAF02:~$
```

- Test Host-GPU Node: 10.3.7.2
- Test Host-Storage Node: 10.3.210.22

GPU Test Throughput Results

```
[SUM] 29.0000-30.0000 sec  21.7 GBytes   186 Gbits/sec
[  2] 0.0000-30.0161 sec  32.6 GBytes  9.32 Gbits/sec
[ 18] 0.0000-30.0160 sec  32.4 GBytes  9.27 Gbits/sec
[ 17] 0.0000-30.0163 sec  33.4 GBytes  9.57 Gbits/sec
[ 16] 0.0000-30.0162 sec  30.0 GBytes  8.57 Gbits/sec
[ 14] 0.0000-30.0161 sec  35.7 GBytes  10.2 Gbits/sec
[  9] 0.0000-30.0164 sec  29.0 GBytes  8.30 Gbits/sec
[  8] 0.0000-30.0160 sec  26.4 GBytes  7.55 Gbits/sec
[ 20] 0.0000-30.0163 sec  28.7 GBytes  8.22 Gbits/sec
[  5] 0.0000-30.0161 sec  31.2 GBytes  8.92 Gbits/sec
[  7] 0.0000-30.0161 sec  30.5 GBytes  8.74 Gbits/sec
[  3] 0.0000-30.0165 sec  24.4 GBytes  6.98 Gbits/sec
[ 19] 0.0000-30.0159 sec  32.5 GBytes  9.30 Gbits/sec
[  6] 0.0000-30.0160 sec  30.9 GBytes  8.84 Gbits/sec
[ 10] 0.0000-30.0162 sec  27.4 GBytes  7.83 Gbits/sec
[  1] 0.0000-30.0161 sec  33.2 GBytes  9.50 Gbits/sec
[  4] 0.0000-30.0160 sec  27.1 GBytes  7.76 Gbits/sec
[ 13] 0.0000-30.0163 sec  17.4 GBytes  4.97 Gbits/sec
[ 12] 0.0000-30.0162 sec  27.6 GBytes  7.90 Gbits/sec
[ 15] 0.0000-30.0162 sec  33.9 GBytes  9.70 Gbits/sec
[ 11] 0.0000-30.0162 sec  26.4 GBytes  7.54 Gbits/sec
[SUM] 0.0000-30.0011 sec   590 GBytes   169 Gbits/sec
[ CT] final connect times (min/avg/max/stdev) = 0.405/0.616/0.839/0.103 ms (tot/err) = 20/0
fs@h194-gpu-polaris:~$
```

Storage Test Throughput Results

```
fs@s22-x86-polaris:~$ iperf -s
------------------------------------------------------------
Server listening on TCP port 5001
TCP window size:  128 KByte (default)
------------------------------------------------------------
[  1] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56250
[  2] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56236
[  3] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56292
[  4] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56270
[  5] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56286
[  6] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56306
[  7] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56234
[  9] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56300
[ 10] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56312
[ 11] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56328
[ 12] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56332
[ 13] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56350
[ 14] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56338
[ 15] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56390
[ 16] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56396
[ 17] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56366
[ 18] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56378
[ 19] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56408
[ 20] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56414
[  8] local 10.3.210.22 port 5001 connected with 10.3.7.2 port 56320
[ ID] Interval       Transfer     Bandwidth
[ 19] 0.0000-29.9968 sec  32.5 GBytes  9.31 Gbits/sec
[  1] 0.0000-30.0435 sec  33.2 GBytes  9.49 Gbits/sec
[  2] 0.0000-30.0019 sec  30.5 GBytes  8.74 Gbits/sec
[  4] 0.0000-30.0009 sec  33.9 GBytes  9.71 Gbits/sec
[  6] 0.0000-30.0012 sec  30.9 GBytes  8.84 Gbits/sec
[  7] 0.0000-30.0009 sec  26.4 GBytes  7.56 Gbits/sec
[ 10] 0.0000-29.9991 sec  27.1 GBytes  7.76 Gbits/sec
[ 12] 0.0000-29.9981 sec  30.0 GBytes  8.58 Gbits/sec
[ 14] 0.0000-29.9984 sec  27.6 GBytes  7.90 Gbits/sec
[ 16] 0.0000-29.9970 sec  35.7 GBytes  10.2 Gbits/sec
[ 18] 0.0000-29.9978 sec  28.7 GBytes  8.22 Gbits/sec
[  3] 0.0000-30.0008 sec  32.6 GBytes  9.32 Gbits/sec
[  8] 0.0000-30.0010 sec  17.4 GBytes  4.98 Gbits/sec
[ 11] 0.0000-29.9986 sec  31.2 GBytes  8.92 Gbits/sec
[ 15] 0.0000-29.9970 sec  33.4 GBytes  9.57 Gbits/sec
[ 20] 0.0000-29.9958 sec  32.4 GBytes  9.27 Gbits/sec
[  5] 0.0000-30.0011 sec  24.4 GBytes  6.98 Gbits/sec
[ 13] 0.0000-30.0002 sec  26.4 GBytes  7.55 Gbits/sec
[  9] 0.0000-30.0004 sec  27.4 GBytes  7.83 Gbits/sec
[ 17] 0.0000-29.9974 sec  29.0 GBytes  8.30 Gbits/sec
[SUM] 0.0000-30.0441 sec   590 GBytes   169 Gbits/sec
```

Test Throughput Results between Leaf-Spine

```
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 65898180608 bits/sec, 5426549 packets/sec
  5 sec output rate 19983240 bits/sec, 35643 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 2712 bits/sec, 3 packets/sec
  5 sec output rate 57426256 bits/sec, 102451 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 32562403328 bits/sec, 2681409 packets/sec
  5 sec output rate 58689072 bits/sec, 104385 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 4685952 bits/sec, 446 packets/sec
  5 sec output rate 28088336 bits/sec, 49670 packets/sec
root@STOR-STOR01-LEAF01>
^C
admin@STOR-STOR01-LEAF01:~$
```

```
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 2977432 bits/sec, 283 packets/sec
  5 sec output rate 47519184 bits/sec, 84723 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 58164002816 bits/sec, 4789804 packets/sec
  5 sec output rate 38648 bits/sec, 51 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 224 bits/sec, 0 packets/sec
  5 sec output rate 54420448 bits/sec, 97014 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 28885874688 bits/sec, 2378625 packets/sec
  5 sec output rate 54754256 bits/sec, 97049 packets/sec
root@STOR-STOR01-LEAF02>
^C
admin@STOR-STOR01-LEAF02:~$
```


- Test Host-GPU Node: 10.3.8.1
- Test Host-Storage Node: 10.3.210.23

GPU Test Throughput Results

```
[  8] 0.0000-30.0106 sec  31.9 GBytes  9.14 Gbits/sec
[  5] 0.0000-30.0105 sec  31.4 GBytes  9.00 Gbits/sec
[ 18] 0.0000-30.0103 sec  40.9 GBytes  11.7 Gbits/sec
[ 13] 0.0000-30.0105 sec  16.3 GBytes  4.67 Gbits/sec
[  2] 0.0000-30.0106 sec  29.3 GBytes  8.38 Gbits/sec
[  7] 0.0000-30.0102 sec  27.2 GBytes  7.78 Gbits/sec
[ 14] 0.0000-30.0103 sec  51.6 GBytes  14.8 Gbits/sec
[  1] 0.0000-30.0098 sec  37.0 GBytes  10.6 Gbits/sec
[  4] 0.0000-30.0103 sec  38.4 GBytes  11.0 Gbits/sec
[ 17] 0.0000-30.0101 sec  36.5 GBytes  10.4 Gbits/sec
[ 20] 0.0000-30.0106 sec  30.2 GBytes  8.64 Gbits/sec
[ 11] 0.0000-30.0102 sec  28.1 GBytes  8.05 Gbits/sec
[ 15] 0.0000-30.0102 sec  29.5 GBytes  8.45 Gbits/sec
[ 12] 0.0000-30.0099 sec  24.8 GBytes  7.10 Gbits/sec
[ 16] 0.0000-30.0106 sec  52.9 GBytes  15.1 Gbits/sec
[  3] 0.0000-30.0105 sec  45.1 GBytes  12.9 Gbits/sec
[ 19] 0.0000-30.0103 sec  24.0 GBytes  6.86 Gbits/sec
[  6] 0.0000-30.0101 sec  28.5 GBytes  8.16 Gbits/sec
[  9] 0.0000-30.0102 sec  18.5 GBytes  5.28 Gbits/sec
[ 10] 30.0000-31.0106 sec   128 KBytes  1.04 Mbits/sec
[ 10] 0.0000-31.0106 sec  27.1 GBytes  7.52 Gbits/sec
[SUM] 0.0000-30.0004 sec   649 GBytes   186 Gbits/sec
[ CT] final connect times (min/avg/max/stdev) = 0.208/0.523/0.706/0.156 ms (tot/err) = 20/0
fs@h225-gpu-polaris:~$
```

Storage Test Throughput Results

```
[ 61] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40428
[ 62] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40492
[ 63] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40526
[ 64] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40412
[ 65] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40480
[ 66] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40474
[ 67] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40442
[ 68] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40478
[ 69] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40540
[ 70] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40440
[ 71] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40468
[ 72] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40482
[ 73] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40472
[ 74] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40470
[ 75] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40512
[ 76] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40554
[ 77] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40560
[ 78] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40556
[ 79] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40558
[ 80] local 10.3.210.23 port 5001 connected with 10.3.8.1 port 40564
[ ID] Interval       Transfer     Bandwidth
[ 80] 0.0000-29.9876 sec  40.9 GBytes  11.7 Gbits/sec
[ 62] 0.0000-30.0011 sec  37.0 GBytes  10.6 Gbits/sec
[ 64] 0.0000-30.0016 sec  24.8 GBytes  7.10 Gbits/sec
[ 66] 0.0000-29.9977 sec  51.6 GBytes  14.8 Gbits/sec
[ 68] 0.0000-29.9982 sec  28.5 GBytes  8.16 Gbits/sec
[ 70] 0.0000-29.9986 sec  27.2 GBytes  7.78 Gbits/sec
[ 72] 0.0000-29.9979 sec  18.5 GBytes  5.29 Gbits/sec
[ 74] 0.0000-29.9935 sec  45.1 GBytes  12.9 Gbits/sec
[ 76] 0.0000-29.9929 sec  29.5 GBytes  8.45 Gbits/sec
[ 78] 0.0000-29.9919 sec  36.5 GBytes  10.4 Gbits/sec
[ 61] 0.0000-30.0474 sec  27.1 GBytes  7.76 Gbits/sec
[ 63] 0.0000-30.0036 sec  16.3 GBytes  4.67 Gbits/sec
[ 67] 0.0000-29.9994 sec  28.1 GBytes  8.05 Gbits/sec
[ 71] 0.0000-29.9948 sec  38.4 GBytes  11.0 Gbits/sec
[ 75] 0.0000-29.9932 sec  31.4 GBytes  9.00 Gbits/sec
[ 79] 0.0000-29.9892 sec  52.9 GBytes  15.1 Gbits/sec
[ 65] 0.0000-30.0006 sec  29.3 GBytes  8.39 Gbits/sec
[ 73] 0.0000-29.9944 sec  31.9 GBytes  9.14 Gbits/sec
[ 69] 0.0000-29.9999 sec  24.0 GBytes  6.86 Gbits/sec
[ 77] 0.0000-29.9919 sec  30.2 GBytes  8.64 Gbits/sec
[SUM] 0.0000-30.0472 sec   649 GBytes   186 Gbits/sec
```

Test Throughput Results between Leaf-Spine

```
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 49101606912 bits/sec, 4043283 packets/sec
  5 sec output rate 11451528 bits/sec, 20311 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 2640 bits/sec, 0 packets/sec
  5 sec output rate 40002400 bits/sec, 70783 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 49503510528 bits/sec, 4076466 packets/sec
  5 sec output rate 38458104 bits/sec, 68475 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 15687000 bits/sec, 1557 packets/sec
  5 sec output rate 62481168 bits/sec, 111112 packets/sec
root@STOR-STOR01-LEAF01>
^C
admin@STOR-STOR01-LEAF01:~$
```

```
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 3698216 bits/sec, 351 packets/sec
  5 sec output rate 26126992 bits/sec, 46121 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 47890460672 bits/sec, 3943573 packets/sec
  5 sec output rate 65473496 bits/sec, 116752 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 3048 bits/sec, 1 packets/sec
  5 sec output rate 70635192 bits/sec, 125116 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 50652786688 bits/sec, 4171027 packets/sec
  5 sec output rate 15088 bits/sec, 21 packets/sec
root@STOR-STOR01-LEAF02>
^C
admin@STOR-STOR01-LEAF02:~$
```


- Test Host-GPU Node: 10.3.8.2
- Test Host-Storage Node: 10.3.210.24

GPU Test Throughput Results

```
[ 15] 0.0000-30.0158 sec  31.2 GBytes  8.93 Gbits/sec
[ 19] 0.0000-30.0158 sec  45.5 GBytes  13.0 Gbits/sec
[ 11] 0.0000-30.0160 sec  15.0 GBytes  4.29 Gbits/sec
[  8] 0.0000-30.0160 sec  27.9 GBytes  7.98 Gbits/sec
[ 13] 0.0000-30.0156 sec  22.3 GBytes  6.37 Gbits/sec
[  1] 0.0000-30.0155 sec  8.73 GBytes  2.50 Gbits/sec
[  9] 0.0000-30.0150 sec  19.9 GBytes  5.71 Gbits/sec
[  6] 0.0000-30.0156 sec  21.1 GBytes  6.03 Gbits/sec
[  3] 0.0000-30.0153 sec  13.0 GBytes  3.71 Gbits/sec
[  2] 0.0000-30.0151 sec  49.1 GBytes  14.1 Gbits/sec
[ 10] 0.0000-30.0156 sec  36.2 GBytes  10.4 Gbits/sec
[ 16] 0.0000-30.0157 sec  50.7 GBytes  14.5 Gbits/sec
[ 18] 0.0000-30.0158 sec  27.8 GBytes  7.95 Gbits/sec
[ 14] 0.0000-30.0151 sec  20.4 GBytes  5.83 Gbits/sec
[ 12] 0.0000-30.0154 sec  46.2 GBytes  13.2 Gbits/sec
[ 20] 0.0000-30.0160 sec  19.4 GBytes  5.54 Gbits/sec
[ 17] 0.0000-30.0155 sec  14.3 GBytes  4.08 Gbits/sec
[  7] 0.0000-30.0156 sec  29.1 GBytes  8.34 Gbits/sec
[  4] 0.0000-30.0150 sec  27.5 GBytes  7.87 Gbits/sec
[  5] 0.0000-30.0322 sec  35.7 GBytes  10.2 Gbits/sec
[SUM] 0.0000-30.0017 sec   561 GBytes   161 Gbits/sec
[ CT] final connect times (min/avg/max/stdev) = 0.146/0.576/0.783/0.169 ms (tot/err) = 20/0
fs@h226-gpu-polaris:~$
```

Storage Test Throughput Results

```
[ 21] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60010
[ 22] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60006
[ 23] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60032
[ 24] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60012
[ 25] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60008
[ 26] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 59994
[ 27] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60056
[ 28] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60058
[ 29] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60078
[ 30] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60070
[ 31] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60096
[ 32] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60088
[ 33] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60170
[ 34] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60134
[ 35] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60190
[ 36] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60124
[ 37] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60154
[ 38] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60178
[ 39] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60110
[ 40] local 10.3.210.24 port 5001 connected with 10.3.8.2 port 60140
[ ID] Interval       Transfer     Bandwidth
[ 40] 0.0000-29.9908 sec  20.4 GBytes  5.84 Gbits/sec
[ 22] 0.0000-30.0037 sec  8.73 GBytes  2.50 Gbits/sec
[ 24] 0.0000-29.9998 sec  46.2 GBytes  13.2 Gbits/sec
[ 26] 0.0000-30.0029 sec  13.0 GBytes  3.71 Gbits/sec
[ 28] 0.0000-29.9972 sec  36.2 GBytes  10.4 Gbits/sec
[ 30] 0.0000-29.9969 sec  29.1 GBytes  8.34 Gbits/sec
[ 32] 0.0000-29.9948 sec  50.7 GBytes  14.5 Gbits/sec
[ 34] 0.0000-29.9950 sec  27.8 GBytes  7.96 Gbits/sec
[ 36] 0.0000-29.9951 sec  19.4 GBytes  5.54 Gbits/sec
[ 38] 0.0000-29.9950 sec  22.3 GBytes  6.37 Gbits/sec
[ 21] 0.0000-30.0528 sec  21.1 GBytes  6.02 Gbits/sec
[ 23] 0.0000-30.0038 sec  27.5 GBytes  7.87 Gbits/sec
[ 27] 0.0000-30.0028 sec  19.9 GBytes  5.71 Gbits/sec
[ 31] 0.0000-29.9951 sec  45.5 GBytes  13.0 Gbits/sec
[ 35] 0.0000-29.9953 sec  14.3 GBytes  4.08 Gbits/sec
[ 39] 0.0000-29.9902 sec  27.9 GBytes  7.99 Gbits/sec
[ 25] 0.0000-29.9996 sec  49.1 GBytes  14.1 Gbits/sec
[ 33] 0.0000-29.9960 sec  15.0 GBytes  4.29 Gbits/sec
[ 37] 0.0000-29.9915 sec  31.2 GBytes  8.93 Gbits/sec
[ 29] 0.0000-30.0171 sec  35.7 GBytes  10.2 Gbits/sec
[SUM] 0.0000-30.0717 sec   561 GBytes   160 Gbits/sec
```

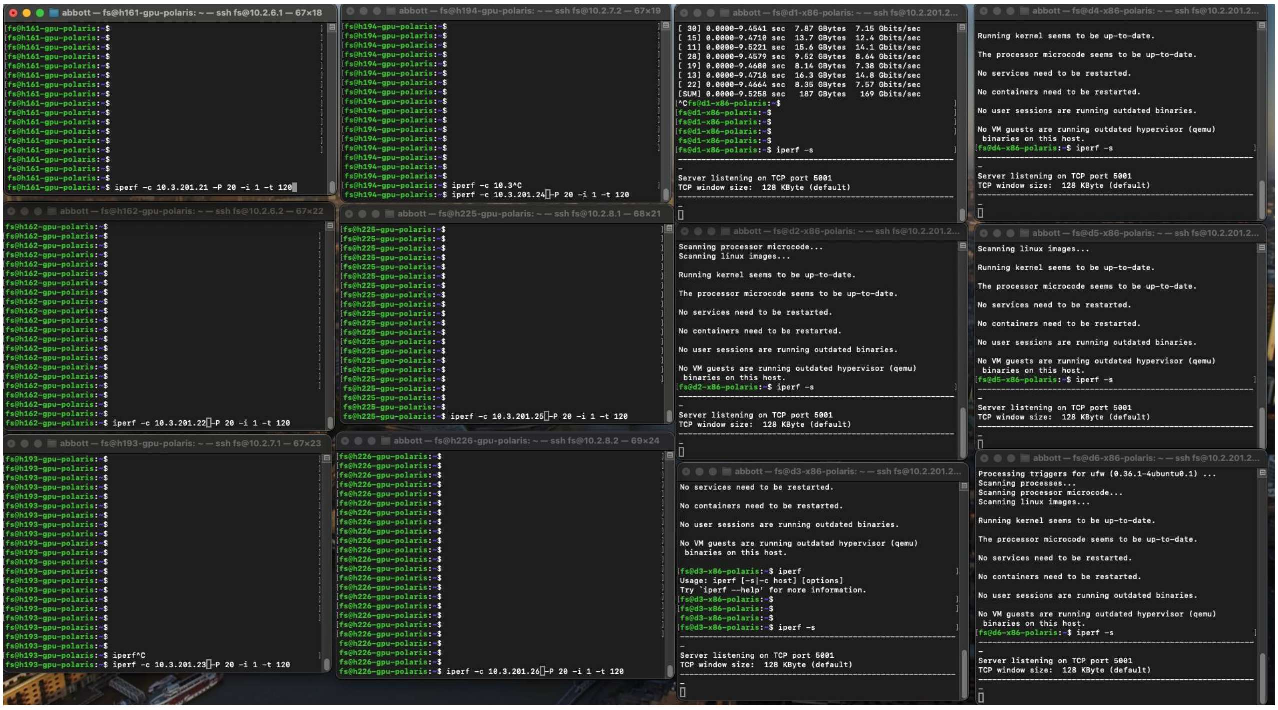
Test Throughput Results between Leaf-Spine

```
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 52220821504 bits/sec, 4300179 packets/sec
  5 sec output rate 47237048 bits/sec, 83604 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 536 bits/sec, 0 packets/sec
  5 sec output rate 16591944 bits/sec, 29157 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 46446968832 bits/sec, 3824775 packets/sec
  5 sec output rate 42852928 bits/sec, 76359 packets/sec
root@STOR-STOR01-LEAF01>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF01>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 5370264 bits/sec, 518 packets/sec
  5 sec output rate 60888888 bits/sec, 108313 packets/sec
root@STOR-STOR01-LEAF01>
```

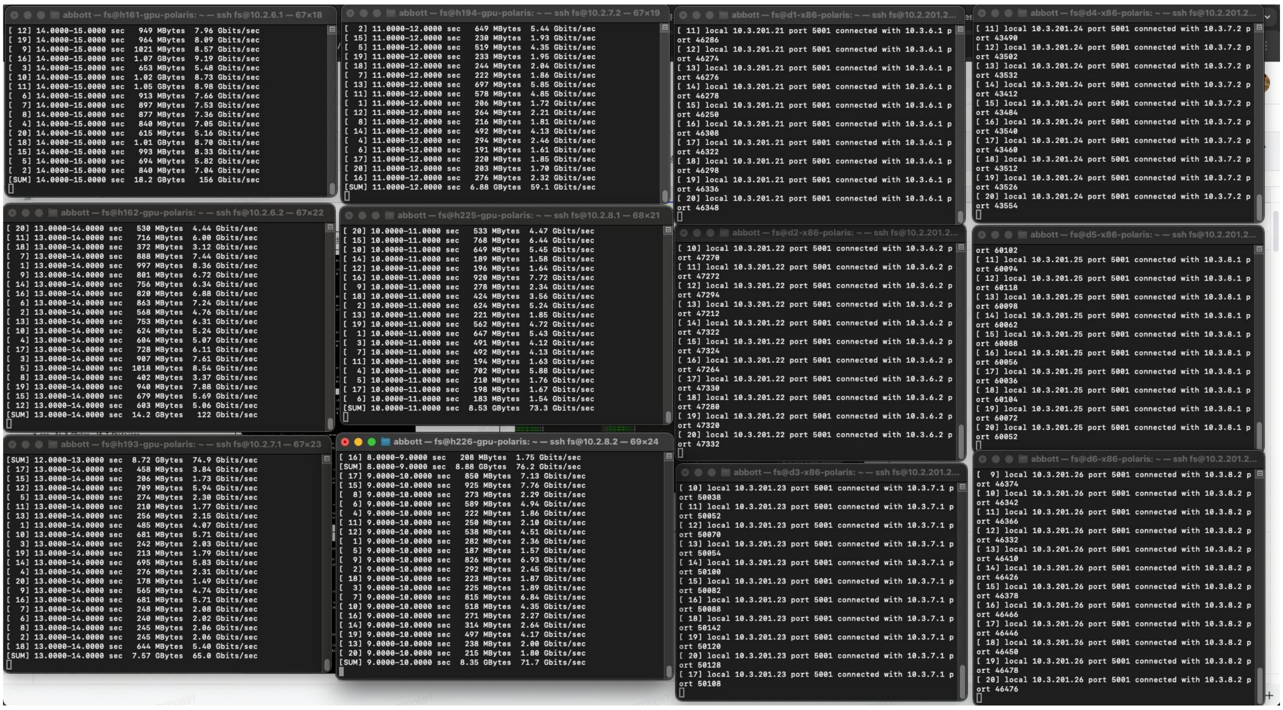
```
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae51 | match bits/sec
.
  5 sec input rate 2096248 bits/sec, 199 packets/sec
  5 sec output rate 51105184 bits/sec, 90905 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae52 | match bits/sec
.
  5 sec input rate 40282013696 bits/sec, 3317107 packets/sec
  5 sec output rate 29900312 bits/sec, 53067 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae53 | match bits/sec
.
  5 sec input rate 952 bits/sec, 1 packets/sec
  5 sec output rate 59621240 bits/sec, 105735 packets/sec
root@STOR-STOR01-LEAF02>
Synchronizing configuration...OK.
Welcome to PICOS
root@STOR-STOR01-LEAF02>
Execute command: show interface aggregate-ethernet ae54 | match bits/sec
.
  5 sec input rate 58121494528 bits/sec, 4786041 packets/sec
  5 sec output rate 38998048 bits/sec, 69275 packets/sec
root@STOR-STOR01-LEAF02>
```

- Simultaneous Speed Test of All Nodes, Overall iperf Throughput Results

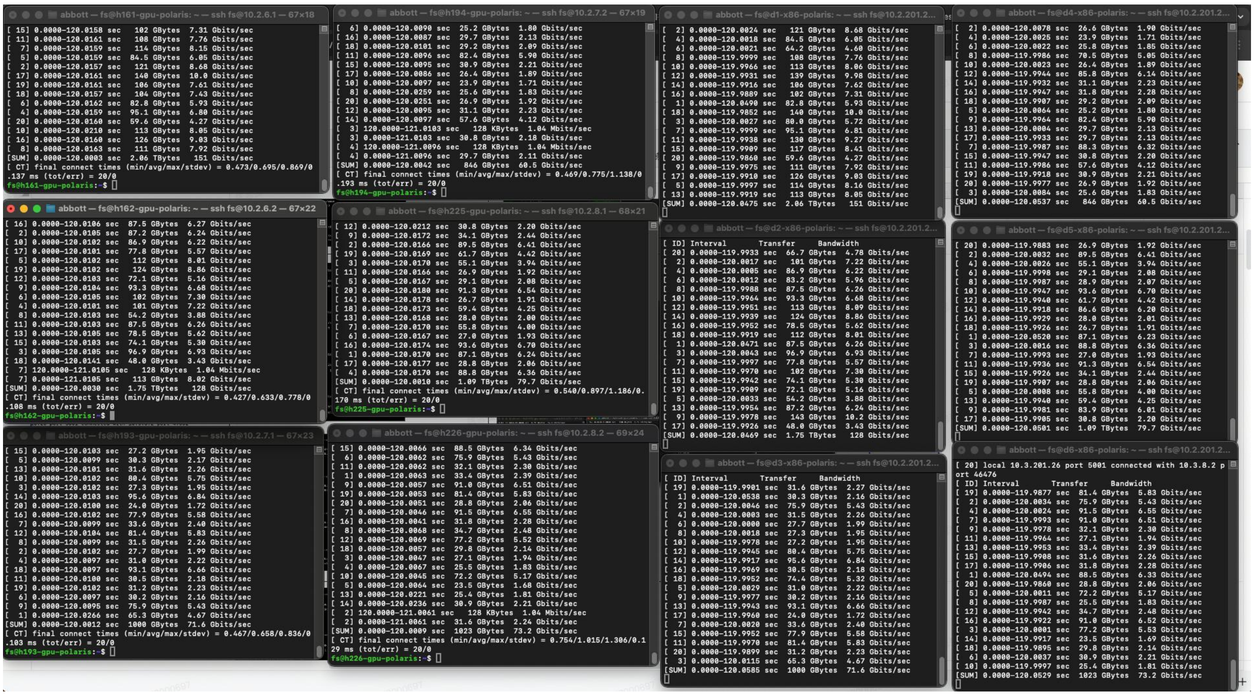
Not yet started speed test, environment screenshot



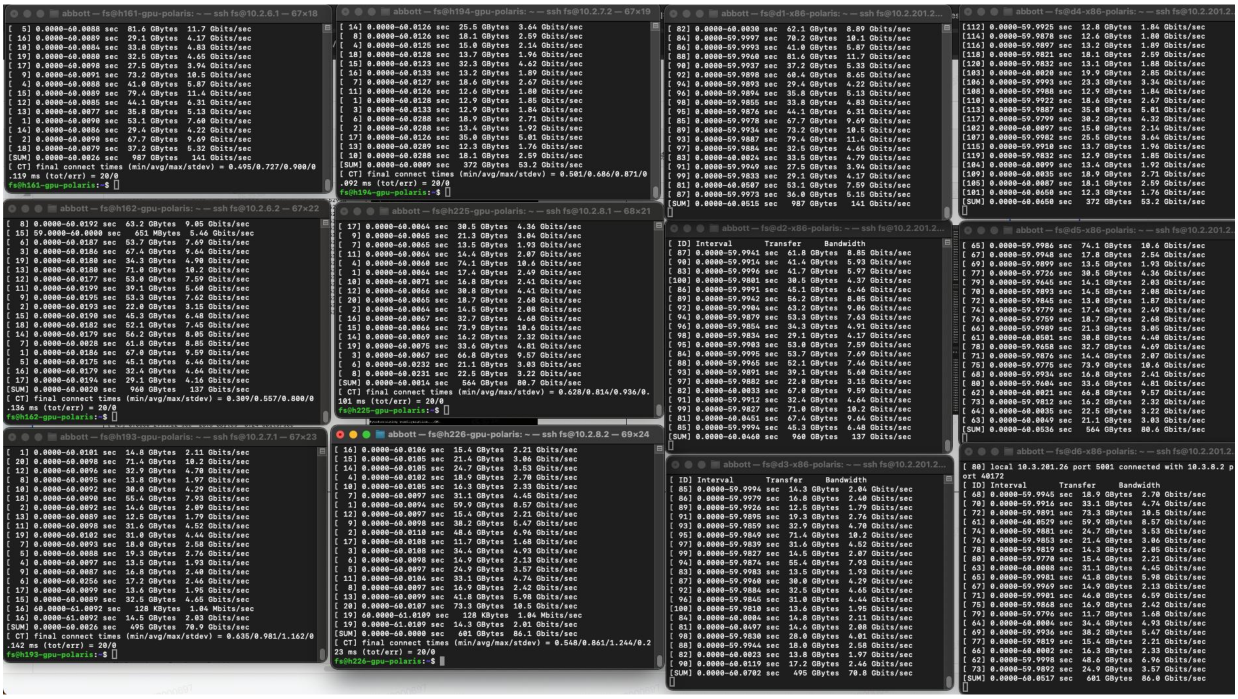
Screenshot during testing



Test results 1



Test results 2



11. Document Handover

11.1 Project Delivery Documents List

Document Name	Document Description
Project Equipment List	List of all equipment for program equipment
Project Implementation Plan	Including implementation plan, configuration list, device management information, software versions, etc.
Project Test Documentation	Connectivity Testing
Operation and Maintenance Guide	Including product manual, configuration command manual, etc.
Project Cabling Implementation Plan	Includes implementation planning, device location information, cabling routing, port routing table, label management, etc.

Note

Documents are updated periodically. Please visit the FS official website: www.fs.com.

12. Appendix

12.1 Server Image Files

Golden image link: https://releases.ubuntu.com/jammy/

12.2 Default System and Software Package Versions in the Image

- Ubuntu 22.04.3 - mini install
- Nvidia-driver-535.129.03 Offline Warehouse: nvidia-driver-local-repo-ubuntu2204-535.129.03_1.0-1_amd64.deb
- Driver installation package for Nvidia NIC:MLNX_OFED_LINUX-23.10-0.5.5.0-ubuntu22.04-x86_64
- Kubernets 1.26.9
- Kubeadm、kubelet、kubectl 1.26.9
- Nerdctl-full 1.7.2
- Containerd v1.7.11

12.3 Script for Image Packaging

```
# Install k8s
echo "deb https://mirrors.aliyun.com/kubernetes/apt kubernetes-xenial main" > /etc/apt/sources.list.d/kubernetes.list
curl https://mirrors.aliyun.com/kubernetes/apt/doc/apt-key.gpg | apt-key add
apt-get update -y
apt-get install kubectl=1.26.9-00 kubeadm=1.26.9-00 kubelet=1.26.9-00 -y
```

```
# Download nerdctl Integration Pack
cd /opt/packages
wget https://github.com/containerd/nerdctl/releases/download/v1.7.1/nerdctl-full-1.7.1-linux-amd64.tar.gz
# unzip and install
tar Cxzvvf /usr/local nerdctl-full-1.7.1-linux-amd64.tar.gz
# activation service
systemctl enable containerd
systemctl enable buildkit
```

```
# Creating a Default Configuration
mkdir -p /etc/containerd
containerd config default | tee /etc/containerd/config.toml
# Modify Configuration /etc/containerd/config.toml
# The modifications are
# SystemdCgroup = true
# sandbox_image = "registry.k8s.io/pause:3.9"
```

```
apt update -y
apt install nano vim iputils-ping net-tools dialog gcc make -y
mkdir -pv /opt/packages
```

13. Terms and Acronyms

Noun	Explanation
AIGC	AIGC (AI-Generated Content) refers to an approach based on AI technologies such as Generative Adversarial Networks (GANs), large pre-trained models, etc., which generates relevant content with appropriate generalization capabilities through learning and pattern recognition of existing data. Similar concepts include Synthetic media (Synthetic media), which mainly refers to AI-based generated media content such as text, images, and audio.
Leaf	The flat network architecture of the Spine Leaf two-tier appliance is derived from the CLOS network. the Leaf is the equivalent of an access switch in a traditional three-tier architecture and acts as a TOR (Top Of Rack) to directly connect to physical servers.
Spine	The flat network architecture of the Spine Leaf two-layer device is derived from the CLOS network. the Spine is the equivalent of a core switch. routing between the Leaf and Spine switches is done via ECMP or DLB.
NIC	A network interface controller (NIC), also known as a network interface controller, network adapter, network interface card, or LAN adapter, is a piece of computer hardware designed to allow computers to to communicate on a computer network.
ROCEv2	Since RoCEv1 data frames do not carry IP headers, they can only communicate within the L2 subnet. To solve this problem, IBTA proposed RoCE V2 in 2014, and RoCEv2 extends RoCEv1 by replacing GRH (Global Routing Header) with UDP header + IP header.
InfiniBand	InfiniBand is an industry-standard specification that defines an input/output architecture for interconnecting servers, communications infrastructure devices, storage and embedded systems. InfiniBand is a true fabric architecture that utilizes switched point-to-point channels for data transfer, whether in chassis backplane applications or over external copper and fiber optic connections, where data transfer is often industry leading. Reliable messaging (send/receive) and memory operation semantics (RDMA), without software intervention in the data movement path, ensure the lowest latency and highest application performance.
RDMA	RDMA (Remote Direct Memory Access) means Remote Direct Address Access, through RDMA, the local node can "directly" access the remote node's memory. The so-called direct, meaning that you can access the local memory, bypassing the traditional Ethernet complex TCP / IP network stack to read and write the remote memory, and this process is not perceived by the end, and most of the work of this read and write process is done by hardware rather than software.

Shenzhen (China)

Address: 24F, Yingfeng Center, Haitian 2nd Rd, Nanshan District, Shenzhen
Tel: +86 (755) 8357 1351
Email: sales@feisu.com

Wuhan (China)

Address: 9-14F, Optical Valley Software Park A7, Guanshan Ave, Wuhan
Tel: +86 (027) 8808 9195
Email: sales@feisu.com

Delaware (United States)

Address: 380 Centerpoint Blvd, New Castle, DE 19720, United States
Tel: +1 (888) 468 7419
Email: us@fs.com

Birmingham (United Kingdom)

Address: Regus Edmund House, 12-22 Newhall Street, Birmingham, B3 3AS
Tel: +49 (0) 8165 4099 260
Email: uk@fs.com

Munich (Germany)

Address: NOVA Gewerbepark Building 7, Am Gfild 7,85375 Neufahrn bei Munich, Germany
Tel: +49 (0) 8165 4099 260
Email: de@fs.com

Melbourne (Australia)

Address: 57-59 Edison Rd, Dandenong South, VIC 3175, Australia
Tel: +61 3 9693 3488
Email: au@fs.com

Singapore

Address: 30A Kallang Pl, #11-10/11/12 Singapore 339213
Tel: +65 6443 7951
Email: sg@fs.com

Tokyo (Japan)

Address: JS Progress Building, 4-1-23 Heiwajima, Ota-ku, Tokyo 〒143-0006
Tel: 03-5826-8305
Email: jp@fs.com

FS has several offices around the world. Addresses, phone numbers are listed on the FS Website at https://www.fs.com/contact_us.html