



LinkX User Guide for 400Gbps 100G-PAM4 OSFP & QSFP112-based Cables and Transceivers

Table of Contents

- Introduction..... 5**
 - Interconnect Documents..... 5
 - 400Gb/s Rate 5
 - Parts are Described by Use in the Switch 6
- NVIDIA Quantum, Spectrum, and LinkX Product Lines 7**
 - Accelerated Data Center Focus..... 8
 - Why AI Accelerated Networking is Different..... 9
 - Dual Protocol Capability 9
 - Unique Connectors in Switches and Adapters 10
- LinkX 100G-PAM4 Product Line Overview..... 11**
 - Same Electronics and Optics: Different Connector Shells 12
 - Backwards Compatibility 12
 - 100G-PAM4 Series: Twin-port OSFP, OSFP, and QSFP112 Plugs 12
 - Twin-port OSFP and OSFP Plugs 13
 - QSFP112 Plugs 14
 - Optical Connectors 15
 - MPO-12/APC Optical Connectors 16
 - LC Duplex Optical Connectors 17
- Optical Fiber Types 18**

Single-mode Fiber	18
Multimode Fiber.....	18
NVIDIA Supplied Single-mode and Multimode Crossover Fiber Cables	19
Optical Patch Panels.....	23
Linking DGXs, Switches, ConnectX-7 Adapters, and BlueField-3 DPUs Overview	25
100G-PAM4-based Switches.....	25
PCIe Adapters: NICs, HCAs, and DPUs.....	26
QSFP Cage Backwards Compatibility.....	27
DGX H100 Networking	27
Cedar7 ConnectX-7 Mezzanine Cards	28
PCIe Cards.....	28
Transceivers and Fiber Details: 100G-PAM4	31
2xSR4 Multimode, Twin-port OSFP, 50-meters	32
2xDR4 Single-mode, Twin-port OSFP, 100-meters	33
2xDR4 Single-mode, Twin-port OSFP, 500-meters	33
2xFR4, Single-mode, 800G Twin-port OSFP, 2km	34
SR4, DR4 400G Single-port OSFP Transceivers	35
SR4, Multimode, 400G Single-port OSFP or QSFP112, 50-meters	36
DR4, Single mode, 400G Single-port OSFP or QSFP112, 100-meters	36
200Gb/s Transceivers are Not Offered.....	37
Transceiver and Fiber Colors and Labeling.....	38

Copper DAC and ACC Cables Overview	40
Direct Attach Copper (DAC) Cables	40
Active Copper Cables (ACCs)	40
Backwards Compatible DACs and AOCs to 200Gb/s 50G-PAM4	41
50G-PAM4 DACs	41
50G-PAM4 AOCs	42
No AOCs for 100G-PAM4 Series	42
Where to Find More LinkX Documentation	43
Document Revision History	44

Introduction

Guide to Configuring 100G-PAM4 LinkX Interconnects for Quantum-2 InfiniBand and Spectrum-4 Ethernet OSFP-based Switches with ConnectX-7 Adapters and BlueField-3 DPUs based on twin-port OSFP, OSFP, and QSFP112 Cables and Transceivers

NVIDIA® LinkX® 400Gb/s Cable and Transceiver User Guides provides detailed information, figures, and ordering part numbers to assist in configuring cables and transceivers for use with network switches, BlueField® DPUs, and ConnectX® network adapters for both Ethernet and InfiniBand protocols. This document covers only cables and transceivers based on 100G-PAM4 modulation and a few specific parts for backwards compatibility linking to 50G-PAM4 and 25G-NRZ devices. Other user guides focus on 50G-PAM4 and 25G-NRZ cables and transceivers, switches, and network adapter interconnects.

Interconnect Documents

This guide is designed to be used in conjunction with additional documentation located in docs.nvidia.com/networking > Interconnect. This site is where the following LinkX cables and transceivers documents are provided.

- See [Where to Find More LinkX Documentation](#)

400Gb/s Rate

The 400Gb/s rate cables and transceivers are used for **both** InfiniBand NDR Quantum-2 QM9700 and 400GbE SN5600 Spectrum™-4 Ethernet OSFP-based switches. The 400Gb/s rate consists of 4-channels of 100G-PAM4 signaling and based on the octal small form-factor plug (OSFP).

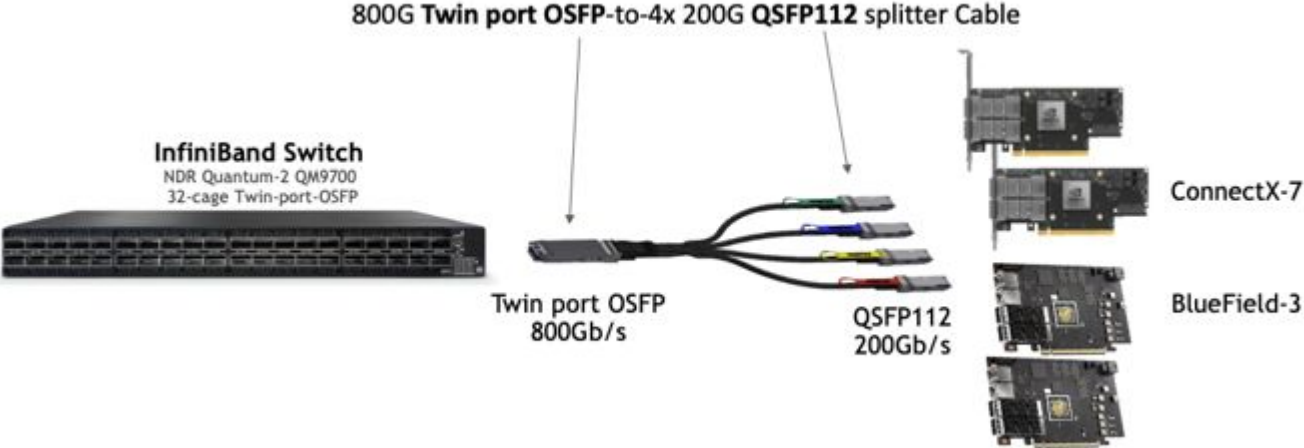
- For switches: **two** 400Gb/s ports are included in **one** 2x400G OSFP air-cooled switch cage in the switches and called a twin-port finned top OSFP. Since there are two 400G engines in the twin-port OSFP transceivers, this has an 800Gb/s aggregate electrical data rate and may also be referred to as 2x400G optical.
- For PCIe network adapters: ConnectX-7 network adapters and BlueField-3 Data Processing Units (DPUs) use 4-channel, 400Gb/s single-port flat top OSFP or QSFP112 cages with finned heat sinks on top of the board cages. The twin-port OSFP devices cannot be used with adapters and DPUs as the finned top is considerably taller than the single port cage.
- LinkX cables and transceivers tie them all together with copper cables and optical transceivers.

The 400Gb/s Spectrum-3 SN4000 series Ethernet switches are constructed from 8-channels of 50G-PAM4 signal modulation and use the quad small form-factor plug with double density (QSFP-DD) and are covered in another document. QSFP-DD is not used with InfiniBand.

Parts are Described by Use in the Switch

The DAC, ACC, and AOC cable part number descriptors are based on the connector *used in the switches*, as the devices used in network adapters may consist of multiple connector types such as the single-port OSFP, QSFP112, QSFP56, etc. These plugs or “form-factors” are used to contain optical transceivers and copper cables to form the switching networks linking CPU/GPU compute engines with storage subsystems and to other system clusters in the network. For example, an 800G twin-port OSFP-to-4x 200G QSFP112 splitter copper cable is listed in the parts lists as 800Gb/s twin-port OSFP, not 200Gb/s QSFP112.

Cable Descriptions Denoted by the Switch-side Connector



NVIDIA Quantum, Spectrum, and LinkX Product Lines

The NVIDIA LinkX line of cables and transceivers offers a wide array of products for configuring any network switching and adapter system. This product line is specifically optimized for 100G-PAM4 line rates.

The NVIDIA networking product line includes:

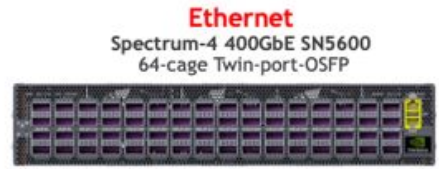
- Quantum InfiniBand switches
- Spectrum Ethernet switches
- ConnectX-7 PCIe-based network adapters
- BlueField-3 PCIe-based Data Processing Units (DPUs)

These are all interconnected using the LinkX cables and transceiver product including:

- Direct attached copper (DAC) cables
- Active copper cables (ACC)
- Multimode transceivers and active optical cables (AOCs)
- Single mode transceivers
- Crossover optical fibers

The above-mentioned product set creates a wide array of interconnect capabilities to build any computing and networking configuration from traditional, simple fat-tree, leaf-spine networks to complex HPC-style 4D-Torus supercomputer configurations.

NVIDIA's 100G-PAM4-based 400Gb/s networking product line



InfiniBand & Ethernet

LinkX

Passive Copper
Up to 3 meters

800G to 2x 400G or 4x 200G
OSFP or QSFP112

Active Copper
Up to 5 meters

800G to 2x 400G or 4x 200G
OSFP or QSFP112

OPTICS

Up to 50 meters

Twin port 2x400G SR8
Multimode Transceiver

Up to 500 meters

Twin port 2x400G DR8
Single mode Transceiver

Up to 2,000 meters

Twin-Port 800G 2xFR4
Single mode Transceiver

Multi-mode fibers:

Single-mode fibers:

Single Port 400G SR4 Multimode
QSFP112 and OSFP Transceivers

Single Port 400G DR4 Single mode
QSFP112 and OSFP Transceivers

InfiniBand & Ethernet



Accelerated Data Center Focus

NVIDIA's LinkX cables and transceivers 100G-PAM4 product line focuses exclusively on accelerated and artificial intelligence computing systems:

- Data center-oriented reaches up to 2 kilometers but generally <50-meters
- High data rates
 - 800Gb/s to the switch
 - 400Gb/s to the network adapter and DPU
- Low latency delays in transferring huge amounts of data using 100G-PAM4 line rates
- Crossover fibers to directly attach two transceivers together

Why AI Accelerated Networking is Different

To minimize data retransmissions, the cables and transceivers are designed and tested to extremely low bit error ratios (BER) required for low-latency, high-bandwidth applications. These are designed to work seamlessly with NVIDIA's Quantum InfiniBand and Spectrum Ethernet network switches, ConnectX PCIe network adapters, and BlueField PCIe DPUs. While the traditional IEEE Ethernet cables and transceivers offering may have hundreds of different parts optimized for specific use cases, for accelerated AI computing, NVIDIA focuses on only a few parts that are optimized for exceptionally high data rates and low latency.

For InfiniBand cables and transceivers, specific lengths, or reaches, are chosen in coordination with the switch buffer latency requirements, latency delays in transceiver and optical fiber. This is done to minimize latency delays in very high-speed systems - a critical factor for AI accelerated systems. Most AI and HPC systems are in size less than the area of two tennis courts. Distances are kept short between components to minimize latency delays.

NVIDIA designs and builds complete subsystems based on the above switches, adapters, and interconnects. All these subsystems, including the major integrated circuits, are designed, manufactured, and/or sourced, all from one supplier—NVIDIA, and tested to work optimally in NVIDIA end-to-end complete configurations—specifically for artificial intelligence and accelerated applications.

Dual Protocol Capability

Network switches have specific models for InfiniBand or Ethernet protocols. However, the LinkX cables and transceivers, ConnectX-7 network adapters, and BlueField-3 DPUs all contain **both** InfiniBand and Ethernet firmware and when inserted into a specific protocol switch, will automatically adopt, and support that protocol. This holds true for the 100G-PAM4 based product lines but fragments at slower line rates into some parts being dual-protocol and some being protocol-specific. See the [NVIDIA LinkX Cables and Transceivers Key Technologies](#) document for more information.

Unique Connectors in Switches and Adapters

For 100GbE, EDR, 200GbE, and HDR networking, the **same** QSFP28 or QSFP56 connector is **used on both** the switches and adapters. For 100G-PAM4 systems, the connector used in the switches is **unique** from those used in adapters and DPUs. This is an important point and set up the basis for the various products. This creates a wide variety of LinkX parts and specific configurations rules.

- Switches use **only** 800G twin-port OSFP finned top cables and transceivers.
- DGX H100 systems use a flat top version of the above.
- ConnectX-7 uses 400G flat top OSFP **or** QSFP112/56/28 cables and transceivers.
- BlueField-3 uses **only** QSFP112/56/28 cables and transceivers; no OSFP devices.
- QSFP112-based ConnectX-7 adapters and BlueField-3 DPUs accept legacy QSFP56 and QSFP28 slower devices. QSFP112 cannot be used in OSFP switches.

LinkX 100G-PAM4 Product Line Overview

The LinkX product line consists of direct attached copper (DAC), active copper cables (ACC), as well as both multimode and single-mode transceivers and crossover fibers.

- **DAC** consists of shielded copper wire with additional RFI foil shielding.
 - For 100G-PAM4 speeds, these can reach 3 meters. These have near zero latency and <0.1 Watts.
 - 30AWG is used for cables ≤ 2 m and 26AWG up to 3m.
 - Available in 800G straight and 400G or 200G and QSFP112 and OSFP split ends.
- **ACC** is a DAC cable that includes pre-emphasis signal booster ICs (not a DSP), aka Linear ACCs.
 - Extends the reach to 3, 4, and 5-meters.
 - 1.5 Watts on 800G end and 0.6W and 0.35W for split ends, with very low latency.
 - 30AWG is used for cables ≤ 2 m and 26AWG up to 3m.
 - Available in 800G straight and 400G or 200G and QSFP112 and OSFP split ends.
- **Multimode optics** convert electrical signals to drive 850nm laser light into 50-micron large diameter optical fibers.
 - Signals sent down the large diameter fiber take different paths (modes) to arrive at the end detector and become distorted as fiber length increases.
 - Multimode has a 50-meter maximum reach.
 - Large fiber 50-um diameter core is easy to align fibers with lasers and detectors resulting in lower costs to manufacture.
 - Fast and easy-to-align and manufacture minimizes costs compared to single mode optics.
 - 800G twin-port OSFP 17-Watt (max.) transceivers.
 - 400G QSFP112 or OSFP 9-Watt (max.) transceivers.
 - Accepts 50-meter straight or 1:2 splitters fiber cables.
 - 1:2 fiber splitters are used in both transceiver ports creating 4 links of 2x100G-PAM4 (200G).
 - 200Gb/s multimode transceivers are not offered and when used with 2-channel fiber split ends enables 200G and reduces power consumption.
- **Single-mode optics** main advantage is the long reach it provides using tiny 9-um diameter fiber core, which holds the signal together out to enormous lengths up to 2km.
 - The tiny core diameter is difficult to align and manufacture, so transceivers are more expensive than multimode transceivers.
 - 800G twin-port OSFP 17-Watt transceivers.
 - 400G QSFP112 or OSFP 9-Watt transceivers.
 - 200Gb/s multimode transceivers are not offered and when used with 2-channel fiber split ends enables 200G and reduces power consumption.

- The main advantage of single-mode optics is the long reach it provides using:
 - Accepts 50-meter 1:2 splitter fibers for switch to adapter links.
 - Accepts straight fibers up at 100m, 500m, and 2km reaches for switch-to-switch applications. NVIDIA supplies up to 100m cables.

Same Electronics and Optics: Different Connector Shells

- Twin-port OSFP 800Gb/s finned top cables and transceivers are used only in switches.
- Twin-port OSFP 800Gb/s **flat top** cables and transceivers are available for use in DGX-H100s and NVLink4 switching systems in single-mode and multimode optics.
- The flat-top versions have the same 800G twin-port OSFP transceiver internals as with finned-top devices and use the riding heat sinks in liquid-cooled systems and large DGX H100 chassis for the GPU links.
- The 4-channel 400G transceivers use the OSFP or QSFP112 form-factor which has the same electronics and optics in both.

Backwards Compatibility

A series of DACs and active optical cables (AOCs) are offered for backwards compatibility with 2x200G twin-port OSFP on one end and 2x 200G or 2x 100G QSFP56/28 on split ends for linking to HDR/200GbE and 100GbE/HDR100 adapters and switches. These DACs and AOCs are the only cables that can link NDR/400GbE to HDR/200GbE systems. Transceivers are not available.

Additionally, QSFP112 cages in ConnectX-7 adapters and BlueField-3 DPUs are backwards compatible and support QSFP56 200G with 4x50G-PAM4 and QSFP28 100G with 4x 25G-NRZ. The cages accept the different connector types, and the cards can down shift to slower line rates.

100G-PAM4 multimode and single-mode transceivers cannot downshift to slower line rates of 50G-PAM4 and 25G-NRZ.

100G-PAM4 Series: Twin-port OSFP, OSFP, and QSFP112 Plugs

Electronics, optics, and copper wires are housed in metal shell plugs called form-factor plugs. The metal plugs have many code name extensions based on the single-channel; small-form-factor plug (SFP).

SFP can be preceded by Q for quad or 4-channels (QSFP) and for 8-channels, quad -- double density (QSFP-DD) and octal (OSFP). NVIDIA created an 8-channel transceiver called the twin-port OSFP that has 8 electrical channels and two optical 4-channel ports. Numbers at the end indicate the maximum

Gb/s speed rating of the connector e.g., 28, 56, 112 for QSFP28, QSFP56, QSFP112. Also note that InfiniBand and Ethernet use slightly lower-speed ratings than the connector maximum speed, e.g. QSFP112 supports 100G line rates.

The 100G-PAM4 line uses only four connector types for DAC, ACC, AOC cables and transceivers:

• Twin-port OSFP finned top	800G	8-channels	Switches only: Quantum-2 InfiniBand and Spectrum-4 Ethernet
• Twin-port OSFP flat top	800G	8-channels	DGX H100 Cedar7 GPU links and liquid-cooled systems
• Single-port OSFP	400G	4-channels	ConnectX-7/OSFP adapters only
• Single-port QSFP112	400G	4-channels	ConnectX-7/QSFP112 and BlueField-3 DPUs

Connectors used in ConnectX-7 and BlueField-3 have flat tops and use the riding cage heat sinks.

The MPO-12/APC is used in all the transceivers, except the 2xFR4 2km twin-port OSFP which uses LC optical connectors.

Twin-port OSFP, OSFP, and QSFP112 Connector Plugs

			
<i>Twin-port 2x400G finned top OSFP</i>	<i>Twin-port 2x400G flat top OSFP</i>	<i>Single-port OSFP</i>	<i>Single-port QSFP112</i>

Twin-port OSFP and OSFP Plugs

The octal small form-factor plug, or OSFP, has become the preferred form-factor for high-speed applications, such as artificial intelligence and HPC networking, as it offers future expansion with more channels, more space for components, and higher power dissipation capabilities. The twin-port OSFP 800G plug has 8-channels of electrical signaling for the switch and two 400Gb/s engines inside the transceiver that exit to two 400G optical or copper ports. Extra cooling fins are used on top to support 17-Watt transceivers, hence the name “2x400G twin-port OSFP finned-top.”

❗ Twin-port OSFP “finned-top” cables and transceivers are **only** used in NVIDIA Quantum-2 NDR InfiniBand and Spectrum-4 SN5600 400GbE Ethernet systems.

The 800G twin-port OSFP is also offered in a flat top version and a 400Gb/s single-port, flat-top OSFP is offered. These are all the same size, but the twin-port OSFP finned top version is taller due to the heat sink.

The three OSFP versions are:

- 800G finned-top, twin-port, 8-channel, 2x400G OSFP for Quantum-2 and Spectrum-4 SN5600 Ethernet air-cooled switches.
- 800G flat-top, twin-port, 8-channel, 2x400G OSFP for linking DGX H100 Cedar7 GPU links which use internal cage riding, air-cooled, heat sinks for liquid-cooled systems. This has the same internals as the finned top version.
- 400G flat-top, single-port, 4-channel, OSFP for ConnectX-7/OSFP network adapters using cage riding heat sinks.

❗ Single-port, 400G OSFP or QSFP112 devices **cannot** be used in twin-port OSFP switch cages -- only adapters and DPUs.

800G twin-port OSFP finned-top connectors are also used to construct passive copper cables, active copper cables, and active optical cables using twin-port OSFP, OSFP, QSFP112, and QSFP56 connector ends. These parts are available in various combinations including 800G twin-port OSFP with flat tops instead of fins.

Lastly, a long reach single mode, 2km twin-port OSFP called 2xFR4 transceiver will be available at the end of 2023 that uses a finned-top but with a lid on top of the fins creating a closed channel for additional cooling.

Twin-port OSFP Open Finned Top



Closed Finned Top



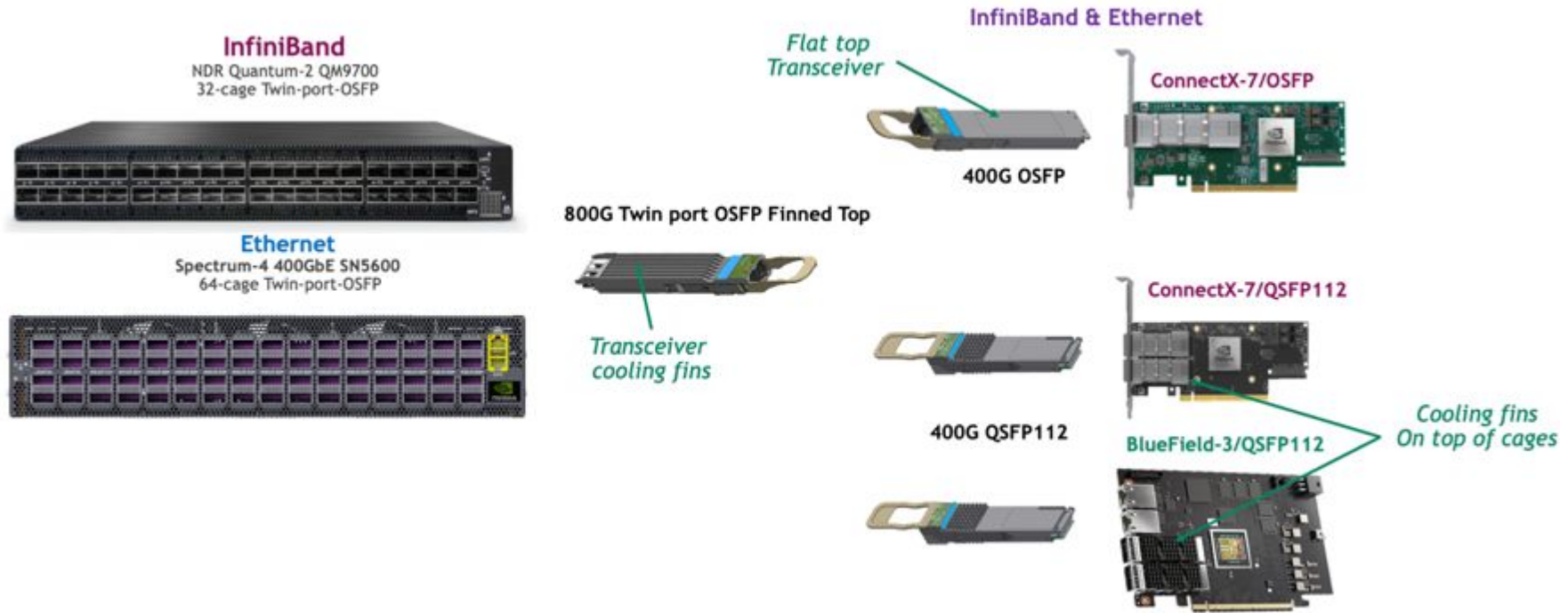
QSFP112 Plugs

QSFP112 form-factor is a single-port, 4-channel, 400G for ConnectX-7/QSFP112 and BlueField-3/QSFP112 DPUs.

- QSFP112 is without cooling fins on top and uses the cooling fins located on adapter and DPU connector cages.
- QSFP112 cannot be used in twin-port OSFP switch cages nor single-port ConnectX-7 adapters based on OSFP.

⚠ ConnectX-7 is offered in **both** OSFP and QSFP112 versions.
BlueField-3 DPUs **only** use the QSFP112.

Twin-port OSFP, OSFP, and QSFP112 Connector Cages for Switches, ConnectX-7, and BlueField-3 DPUs



Optical Connectors

Two optical connectors are used for 100G-PAM4:

- MPO-12/APC 8-fiber: Multiple-Push-On, 12 fiber, Angled Polished Connector for single-mode and multimode optics.
- LC duplex: 2-fiber Lucent Connector for 2km 2xFR4 transceivers for single-mode optics only.

The MPO-12/UPC, Ultra-flat Polish Connector, using 8-fibers is used for 50G-PAM4 backwards compatibility links for 200GbE, HDR, 100GbE, and EDR SR4 transceivers.

MPO-12/APC Optical Connectors

The MPO-12 optical connector is a ceramic block with holes that contain the ends of multiple optical fibers in either single-mode or multimode types. The ceramic blocks are made with different numbers of holes 8, 12, 16, 24, etc. but the 100G-PAM4 series uses 8-fibers but is labeled MPO-12.

Some of the light sent into a fiber reflects backwards from the fiber end face. Slower speed electronics and optics are **less sensitive** to back reflection created inside the optical fiber. Hence, 25G-NRZ (4x25G-NRZ SR4) and some 50G-PAM4 transceivers (4x50G-PAM4 SR4) use a different polish: MPO-12/UPC or **Ultra-flat Polished Connector**.

100G-PAM4 transceivers are **more sensitive** to back reflections and use the MPO-12/APC or **Angled Polished Connector**. This has an 8-degree polish on the end that causes the back reflections to be diverted into the fiber side cladding and away from the transmitter.

For 100G-PAM4 optics:

- MPO-12/APC is used for both single-mode and multimode optics.
- Two MPO-12/APC optical connectors are used with the twin-port 2x400G OSFP transceivers for two ports of 4-channels of 100G-PAM4 each.
- A single MPO-12/APC optical connector is used with 4-channel, 400G OSFP and QSFP112 transceivers.
- Only 8-fibers are used with 4 fibers for transmission and 4-fibers for receiving.
- NVIDIA supplied fiber cables have NVIDIA green plastic shells on the MPO ends to denote APC type.
- Transceivers have alignment pins and MPO-12/APC connectors have alignment holes. This is important to know when using trunk cables which may have pins or holes in the connector end. Inserting a pinned connector into a transceiver will damage both. A white dot or hole in the plastic side indicates fiber pin-1 side.
- The MPO-12/APC (green) is not compatible with the MPO-12/UPC (blue) ultra-flat polished connector used with 40G-200Gb/s transceivers with 50G-PAM4 and 25G-NRZ modulation as it has a flat polish.
- 400Gb/s QSFP-DD single mode DR4 transceivers used in Spectrum-3 SN4000 and Spectrum-4 SN5400 switches also use the 8-fiber MPO-12/APC. The DR4 uses 8-channels of 50G-PAM4 electrical and 4-channels 100G-PAM4 optical. This is the **same** optical connector used with twin-port and single-port 100G-PAM4 transceivers and can be used to link QSFP-DD with twin-port OSFP transceivers linking QSFP-DD 8x50G-PAM4 systems to 100G-PAM4 systems.

LC Duplex Optical Connectors

The 2-fiber Lucent Connector, or LC duplex, is typically used for single channel links using a transmit fiber and a receive fiber. This applies to multimode and single-mode optics.

Running parallel fibers over long reaches becomes expensive so transceivers use multiple lasers with different wavelengths which are combined or “multiplexed” into a single fiber for transmission and filtered back out at the receiver. One fiber for transmission and one for receiving but carrying 4 or 8 channels.

For long reaches up to 2-kilometers, two 2-fiber LC duplex optical connectors are used in the Far Reach 4-channel 2xFR4 twin-port OSFP transceiver. These are used to link 2x400Gb/s NDR InfiniBand and Spectrum-4 Ethernet switches together and linking clusters across campuses by using only two fibers in each optical connector.

MPO-12/APC, MPO-12/UPC, and Duplex LC Optical Connectors

			
<i>MPO-12/APC</i>		<i>MPO-12/UPC</i>	<i>Duplex LC</i>

Optical Fiber Types

Single-mode Fiber

Single-mode fiber has a tiny 9-um light carrying core. This is small enough to make the light bend inside at a very shallow angle which keeps the data pulse light photon packets together as a group or “single mode” which can travel over great distances. Single-mode fiber uses 1310nm wavelength and is typically used for long reaches of 50-meters to 2km to link switches together.

- The fiber cable jacket is colored yellow

Multimode Fiber

Multimode fiber is used for short reaches of less than 50-meters. Multimode fiber has a large 50-um light carrying core and some of the photons in a single data pulse takes steeper angles traveling down the fiber. This causes the light pulse to scatter inside the fiber into multiple paths or “multi(ple)modes” with some parts of the pulse taking longer to arrive at the end at the photodetector. This results in the pulse intensity weakening and spreading out in time enough to collide with the next following data pulse which limits the maximum reach.

- Multimode fiber is optimized to be most transparent at 850-nm wavelength.
- Multimode 850nm transceivers cannot operate with 1310nm single-mode fibers and transceivers.
- The fiber cable jacket is colored aqua.
- Fiber type is OM4 up to 50-meters and OM3 for <30m.

NVIDIA Supplied Single-mode and Multimode Crossover Fiber Cables

NVIDIA offers only crossover optical fiber cables based on 4-channels of 100G-PAM4 only using MPO-12/APC optical connectors. These cables are specific to the 100G-PAM4 NVIDIA offering for NDR InfiniBand and Spectrum Ethernet. Linking up to five transceivers directly together over short reaches without intermediate optical patch panels is standard practice for accelerated AI accelerated systems.

NVIDIA fiber cables are crossover, Type-B fibers that enable directly linking two transceivers together. This enables directly connecting transceivers together and aligning transmit lasers with receiver photodetectors by crossing over the fibers pin arrangement inside the cable with both optical connectors. For example, pin 1 in transmit laser side is crossed over to pin 12 in the receive transceiver's photodetector side optical connector.

NVIDIA's fiber cable offering includes:

- Straight, single mode, crossover fibers up to 100m, and 1:2 splitter fibers up to 50-meters.
- Straight and 1:2 splitter multimode crossover fibers up to 50-meters.
- These fibers are available from NVIDIA and designed and qualified specifically for NVIDIA systems with the highest quality and tolerances.

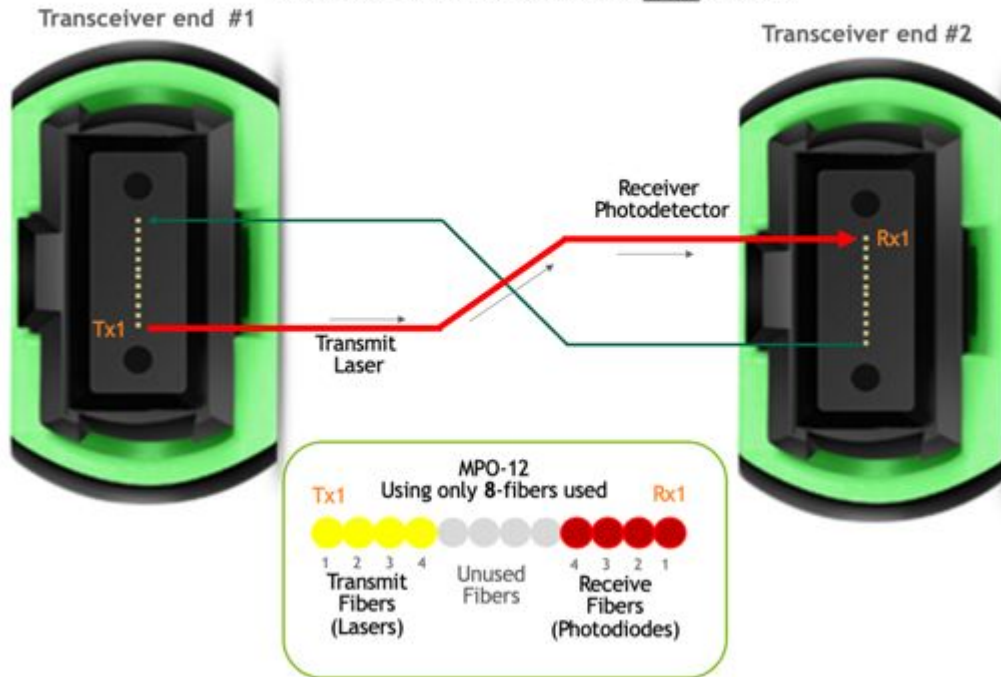
NVIDIA does *not* offer fiber cables based on:

- MPO-12/**UPC** or MPO-**16**/APC or LC duplex.
- Non-crossover, Type-A parallel straight or splitter fiber cables, trunk cables, 2-fiber single mode LC cables, and crossover fibers longer than 100-meters.
- NVIDIA recommends sourcing these from numerous third-party suppliers.

Crossover fibers

Type-B “Crossover” cable

Transmitter lasers 1,2,3,4 have to align with receiver photodetectors 1,2,3,4
Transmit and receive fibers are crossed inside the cable



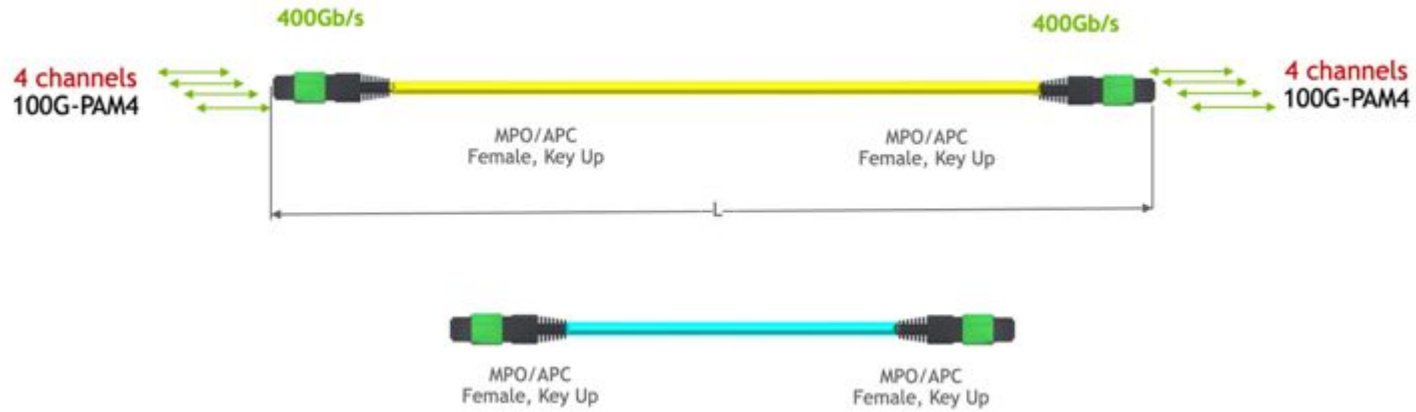
Fiber splitter 1:2 cables using three connectors with 4-channels for the twin-port OSFP transceiver end and two 2-channel ends for single port transceivers are offered for 3m to 50-meters in both single-mode and multimode. Twin-port OSFP transceivers use two 1:2 splitter cables for an effective 1:4 split to 4x200G.

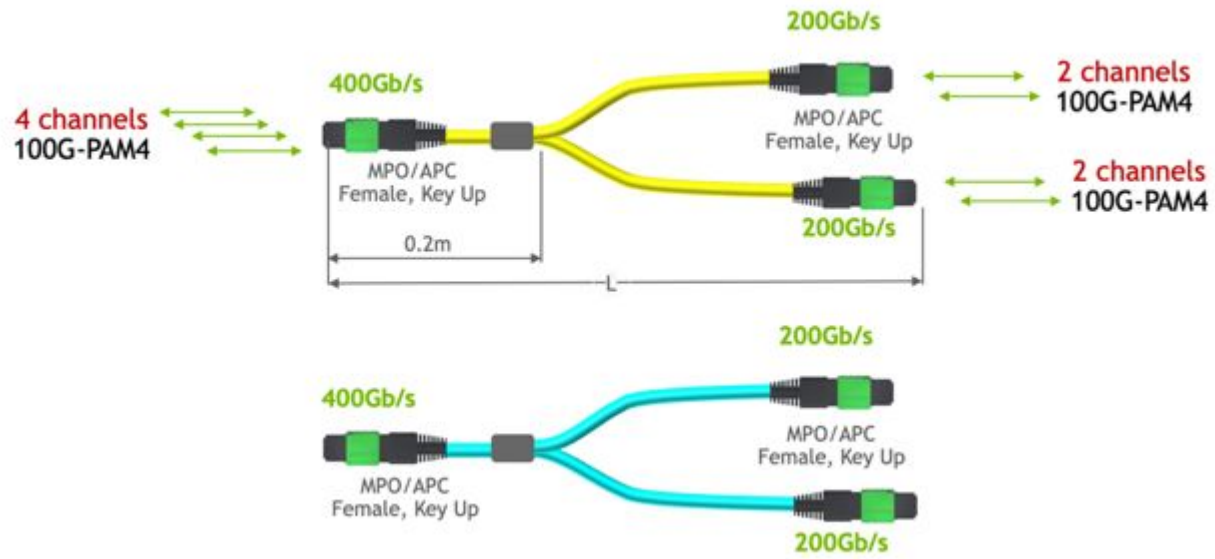
Length is measured from optical connector end faces.

- Split points are 0.2-meters from the single optical connector face.
- Optical connector keys - small bumps on top of the connector ends are used to orient the fiber to the transceiver and align the fiber 1 to laser 1, etc.
- Bend radius of the fibers is 3mm.

- Split point is 0.2m from connector end face.
- All NVIDIA-supplied fiber cables are female on both ends -- pins are on the transceivers.

NVIDIA-supplied, Straight and Splitter Fiber Cables





Straight and Splitter Fiber Cables in Twin-port OSFP, OSFP, and QSFP112 Transceivers





Optical Patch Panels

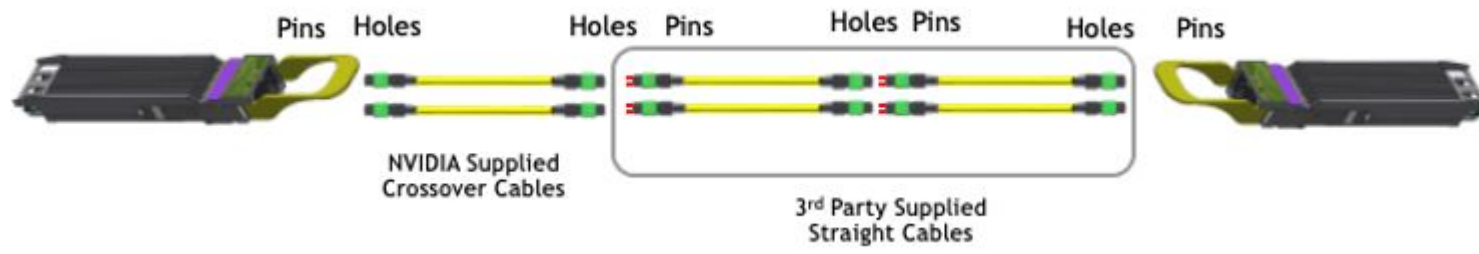
The reaches specifications used with NVIDIA transceivers (50m, 100m, 500m, and 2kilometers) are calculated assuming two optical patch panels are used in the link and assuming each employ 2 optical connectors each, for a total of 4 optical connectors. Transceivers are tested in the factory with this worst-case configuration. Complex cabling infrastructures or lengths longer than specified with multiple optical connectors in the link should have the link budget calculated to ensure sufficient light can reach the end points.

When Type-A, parallel fiber trunk cables are used, at least one segment should have a crossover Type-B cable in the link to align lasers with photodetectors. For lengths longer than 100-meters, we recommend the customer sourcing high-quality third-party single-mode fiber cables for parallel MPO-12/APC and LC based cables. Attention must be paid to connector pins and holes and fiber polarity to align transmit lasers with receive transceiver photodetectors.

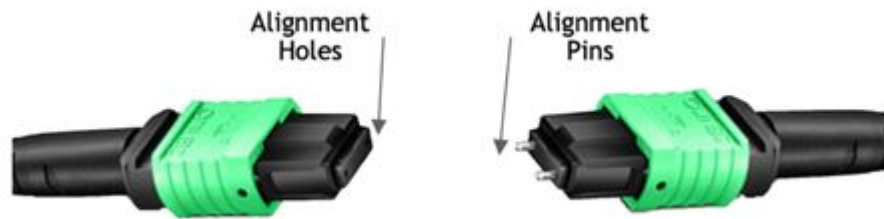
A single crossover Type-B cable can be used with multiple Type-A straight fiber cables The end goal is to align pin 1 of the transmit laser transceiver to pin 12 of the photodetector in the receive transceiver.

Using three NVIDIA supplied, crossover fiber cables will not work as the fiber cables have holes for alignment pins on both ends and need to mate with another cable pins on connector ends.

Two Connector Junctions or Two Patch Optical Panel Links for NVIDIA Tested Transceivers Using Type-A and B Cables



MPO-12/APC Alignment Holes and Pins



Linking DGXs, Switches, ConnectX-7 Adapters, and BlueField-3 DPUs Overview

100G-PAM4-based Switches

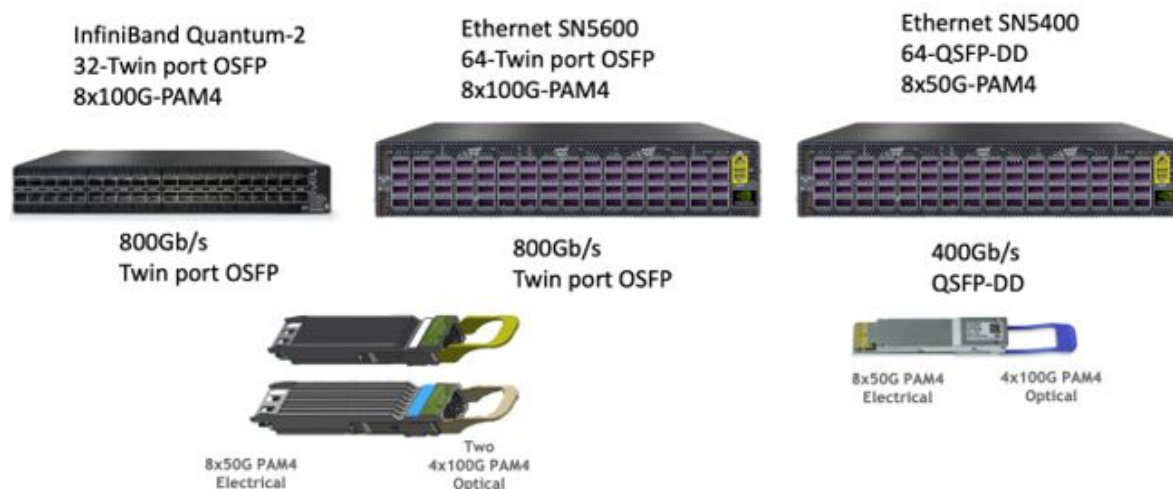
The Quantum-2 NDR InfiniBand and Spectrum-4 SN5600 Ethernet line of OSFP air-cooled, 400Gb/s network switches are based on 100G-PAM4 signal modulation. Although 400Gb/s is the speed rating, these switches use connector cages that house two 400Gb/s ports in a single cage called 2x400G twin-port OSFP and are used exclusively in these air-cooled switches.

SN5400 and SN4000 series Ethernet switches, based on 50G-PAM4 modulation, use QSFP-DD connectors and accept QSFP56 and QSFP28 connectors.

ConnectX network adapters and BlueField data processing units (DPUs) employ only flat-top connector shells of single-port OSFP, QSFP112, QSFP56 form-factors as flat tops are assumed and not denoted in part number descriptions. QSFP-DD is used in Ethernet switches only and not in ConnectX-7 or BlueField-3.

The twin-port OSFP cages enables doubling the number of 400Gb/s ports. 1RU, 32-cage Quantum-2 and 2RU, Spectrum-4 switches support 64x 400Gb/s ports and the 64-cage Spectrum-4 switch supports 128x 400Gb/s ports. This feature enables a single switch to create an enormous network. For example, a single SN5600 Ethernet switch with 64-cages (128-ports) has the same networking capability as 6-to-7 SN4000 switches.

Quantum-2 NDR InfiniBand and Spectrum-4 Ethernet Switch Cages and Twin-port, Finned-top OSFP Plugs



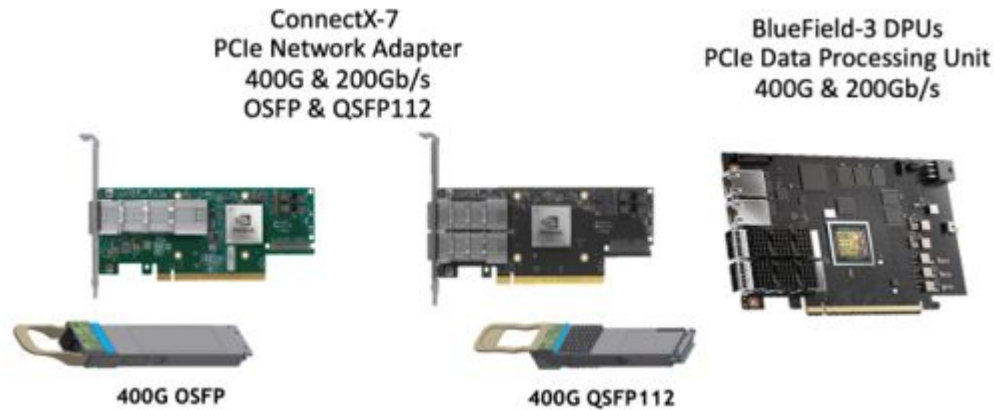
PCIe Adapters: NICs, HCAs, and DPUs

The ConnectX-7 Ethernet Network Interface Cards (NICs) and InfiniBand Host Channel Adapters (HCAs) are PCIe bus-based network adapters. The 100G-PAM4-based ConnectX-7 supports both InfiniBand and Ethernet in the same card. These are inserted into compute servers and storage subsystem PCIe Gen5 bus slots which act as the backbone of a subsystem chassis. PCIe-based adapter cards may have 8, 16, or 32-channels of 16 or 32Gb/transfer at the PCIe bus. The adapter ICs translate the large number of slower PCIe lanes and modulation type into 4-8 lanes at 100G-PAM4 and into the Ethernet or InfiniBand protocols. These signals are presented to the OSFP or QSFP112 connector cages that the cables or transceiver plugs are inserted into.

- 200G/400G ConnectX-7 are offered in QSFP112 and OSFP
- 200G/400G BlueField-3 are offered in QSFP112 only
- Based on PCIe Gen 5 at 32G transfers/sec
- Some ConnectX-6 and BlueField-2 adapters are protocol-specific and have different part numbers for InfiniBand and Ethernet and may require specific-protocol and part number cables and transceivers.

This creates a wide array of product SKUs with a twin-port finned-top OSFPs and QSFP-DD for the switch side and OSFP, QSFP112, QSFP56, or QSFP-DD cables and transceivers for everything else.

ConnectX-7 and BlueField-3 Adapter Cages and Single-port, Flat-top OSFP and QSFP112 Plugs



QSFP Cage Backwards Compatibility

QSFP cages in switches and adapters have a unique backwards compatibility feature enabling older versions to be inserted into newer models. For example:

- 4-channel QSFP112 (4x100G-PAM4) cages can support QSFP56 (4x50G-PAM4) and QSFP28 (4x25G-NRZ) devices.
- QSFP-DD (8x50G-PAM4) cage has 8-channels by using 2 rows of electrical pins in the cage supports QSFP112, 56, and 28.
- Modulation speeds are negotiated by the switches and adapters. Many, but not all transceivers, can down shift to slower modulation speeds.
- Twin-port OSFP and single-port OSFP cages are not compatible with the QSFP line of devices.
- No QSFP-to-OSFP adapter (QSA) is offered.

DGX H100 Networking

The DGX H100 system has two networking areas: one for GPUs and another for general networking.

- Cedar7's eight ConnectX-7 mezzanine cards are linked to switches.
- Four PCIe Gen5 bus connectors for ConnectX-7 and BlueField-3 adapters using standard LinkX cables and transceivers, either OSFP or QSFP112, QSFP56.

- The PCIe bus cards mounted in the DGX H100 chassis are network adapters and DPUs and must link to Ethernet or InfiniBand switches.

Cedar7 ConnectX-7 Mezzanine Cards

The DGX H100 is essentially a CPU and GPU server and ConnectX adapter and BlueField-3 DPU complex that must be connected to a network switch. It has eight ConnectX-7 ICs mounted on two mezzanine Cedar7 cards located in the back of the CPU server chassis. One side links directly to the eight GPU cluster located on the GPU sled above and the other side uses an internal DAC cable to link to the front of the server chassis four 800G cages. These cages have internal, air-cooled, riding heat sinks so use a twin-port, flat-top, 800G OSFP transceiver cage.

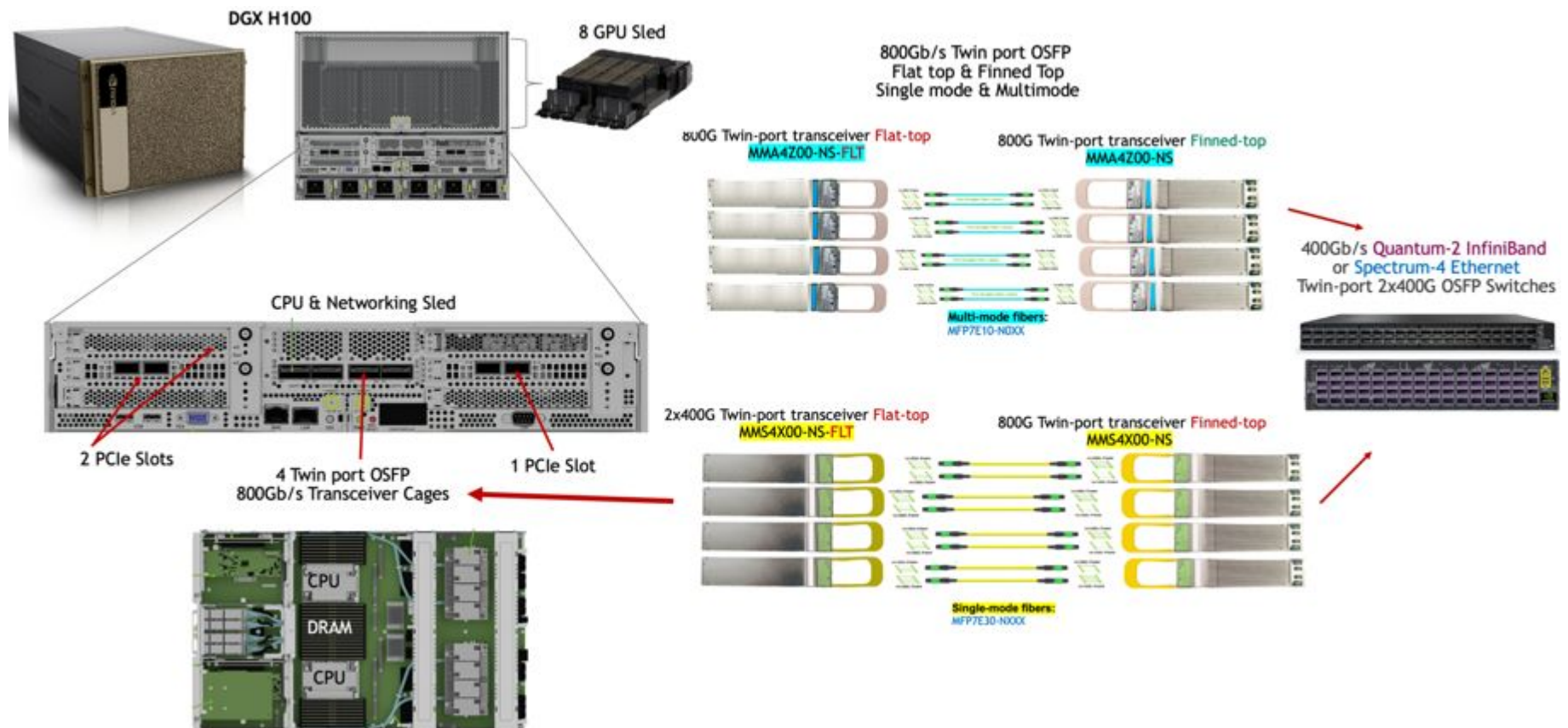
Each 800G cage supports two 400G ConnectX-7 ICs on the internal Cedar7 cards at 400G each. The four, twin-port OSFP flat-top cages support 800G single-mode or multimode transceivers and are usually linked to air-cooled switches using twin-port OSFP finned-top transceivers. This approach reduces the redundant PCB, cards, brackets, cages, etc. and reduces the overall size.

DACs and ACCs are generally not used due to the internal chassis cabling losses between the ConnectX-7 IC Cedar cards in the chassis backside and four OSFP cages on the chassis front creates too significant signal losses, resulting in very short cable lengths. Only one 3-meter ACC is offered.

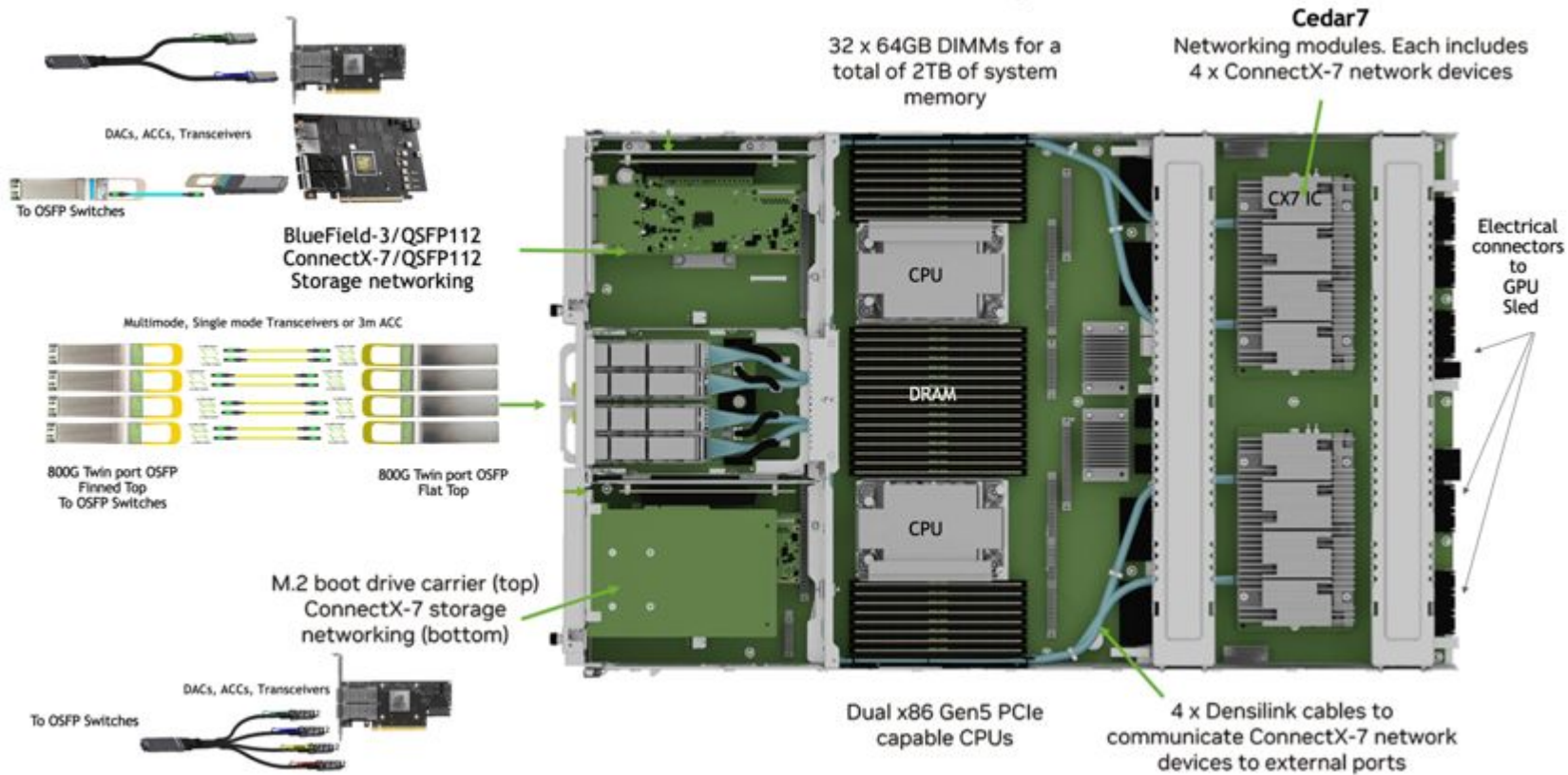
PCIe Cards

The DGX H1000 also has four PCIe Gen5 connectors for ConnectX-7 and BlueField-3 adapters and uses standard LinkX DAC, ACC cables, or transceivers in multimode or single-mode optics. The cards mounted in the DGX H100 chassis are network adapters and DPUs and must link to twin-port OSFP-based switches in Ethernet or InfiniBand. These cards receive the cable split ends from switches for any set up.

DGX H100 GPU-Cedar7 Flat-top OSFP Networking



CPU Sled Top-View Twin-port OSFP Cages on Left; Cedar7 Mezzanine Cards on Right



Transceivers and Fiber Details: 100G-PAM4

Twin-port OSFP single-mode transceivers house two complete multimode or single-mode optical engines inside that exit to two, 4-channel MPO-12/APC optical connectors creating the twin-ports. Twin-port transceivers can be linked to each other forming an 800Gb/s link and can be linked to two or four 400G QSFP112 or OSFP transceivers using two straight or 1:2 splitter fiber cables forming either two 400G or four 200Gb/s links.

- Twin-port OSFP transceivers can link to any combination of 2 or 4 OSFP and/or QSFP112 at the same time, as the internals are the same for both OSFP and QSFP112.
- The fiber types (straight or splitter) **must be the same in both twin-port transceiver ports** as the switch and transceiver combination can only operate in either 2x400G or 4x200G split mode and **cannot be configured with one straight and one 1:2 splitter fiber**.
- Twin-port 800G AOCs are not offered for 800G but are offered for 2x200G backwards compatibility to HDR and 200GbE.

Offered in multimode optics up to 50-meters and single-mode up to 100-meters and 500m and 2k-meters for long switch to switch links. The lengths chosen are related to not just the optics capabilities but also minimizing the overall link latencies including the switch buffer timing, optics latencies and capabilities, adapter timings. A 2km 2xFR4 version is also offered using two LC duplex optical connectors.

Transceiver reach and optics type nomenclature is as follows:

Twin-Port OSFP

- 2x SR4: Short Reach 4-channel multimode optics; 800Gb/s; 50-meters maximum reach, aka SR8
- 2x DR4: Datacenter Reach 8-channel; 800Gb/s; 100 and 500-meters maximum reach, aka DR8
- 2x FR4: 2x Far Reach 4-channel; 800Gb/s; 2km maximum reach

Single-Port OSFP or QSFP112

- SR4: Short Reach 4-channel multimode optics; 400Gb/s; 50-meters maximum reach
- DR4: Datacenter Reach 4-channel single-mode optics; 400Gb/s; 100m and 500-meters maximum reach
- FR4: Far Reach 4-channel, 400Gb/s, 2km maximum reach -- is not offered

2xSR4 and 2xDR4 are sometimes referred to in datasheets as 800G DR8 and 800G SR8. These are twin-port OSFP transceivers with 8-channels electrical (DR8) with two ports of 400G optical (2x400G).

The 400G InfiniBand and Spectrum-4 SN5600 Ethernet switches use exclusively the finned-top, twin-port, 2x400G OSFP connector. The twin-port OSFP employs 8 electrical pins all in one row, straight across for connection to the switch. Twin-port devices used in air-cooled switches have additional cooling fins on top of the connector as the transceivers dissipate 15-17 Watts each, and there are very few air-cooling holes in switch front panels.

Twin-port transceivers support two ports of 4x100G-PAM4 using two MPO-12/APC optical connectors. Each twin-port transceiver supports:

- Two straight fiber cables with two 400G transceivers
- Two 1:2 splitter cables supporting four transceiver ends. The 1:2 splitter fiber ends have 2 fiber sets (2-channels) and when used with 400Gb/s transceivers results in 200Gb/s rate.
- 400G transceiver cannot be downshifted to 50G-PAM4 for 200G. 200G transceivers are not offered.

2xSR4 Multimode, Twin-port OSFP, 50-meters

As most AI accelerated and HPC systems have a typical reach of less than 50-meters to limit the latency delay due to the light traveling in the fiber at 4.5 ns per meter, the 2xSR4 (aka SR8) is the most popular transceiver in addition to it being a less expensive than single mode optics. The 2xSR4 (aka SR8) consists of two VCSEL-based 850nm SR4 transceivers inside the twin-port OSFP shell. Two MPO-12/APC optical connectors exit the back of the transceiver supporting three configurations:

- 800G-to-800G switch-to-switch using two 2xSR4s and two 4-channel fibers.
- 800G-to-2x 400G switch-to-2x 400G adapter switch using a 2xSR4 and two 4-channel straight fibers to two 400G transceivers.
- 800G-to-4x 200G switch-to-4x 200G adapters switch using four 2xSR4s and two 2-channel fibers splitters.

800G-to-2x 400G where each 400G port supports a 400G transceiver or be split by two creating a total of four 200Gb/s runs.

Both fibers must be the same straight or splitter and not one of each as the transceiver can only act as straight or split mode for both ports.

- 50-meter transceiver design assumes two optical patch panels in the link budget.
- Pull tab color: tan
- Transceiver label: NVIDIA green
- Reach marking: Blue stripe indicating multimode 50-meters
- 17 Watts max.

2xSR4 800Gb/s Twin-port OSFP Finned and Flat Top, 50-meter, Multimode Transceivers



2xDR4 Single-mode, Twin-port OSFP, 100-meters

Single mode optics is all about long fiber reaches. For large system, the 2xDR4 single mode transceiver (aka DR8) offers long lengths and is typically used from 50-meters to 100-meters although supports upwards from 1-meter. The 2xDR4 consists of two transceiver engines using 1310nm externally modulated lasers (EML).

- Transceiver design assumes two optical patch panels in the link budget.
- Pull tab color: Yellow
- Transceiver label: NVIDIA green
- Reach marking: White indicating 100-meter single mode
- 17 Watts max.

2xDR4 800Gb/s Twin-port OSFP Finned and Flat Top, 100-meter, Single-Mode Transceivers



2xDR4 Single-mode, Twin-port OSFP, 500-meters

A 500-meter twin-port OSFP single-mode transceiver is offered using two parallel 8-fiber MPO-12/APC optical connectors. This transceiver has the same design and specifications as the 100-meter version but tested to 500-meters and may use different error correction schemes.

- Used to link two switches at 800Gb/s or three switches together at 400Gb/s
- Crossover fibers cables longer than 100-meters are not offered by NVIDIA
- Transceiver design assumes two optical patch panels in the link budget
- Pull tab color: Yellow
- Transceiver label: NVIDIA green
- Reach marking: Purple indicating 500-meter single mode
- 17 Watts max.
- NVIDIA does not supply fibers longer than 100-meters

2xDR4 800Gb/s Twin-port OSFP Finned and Flat Top, 500-meter, Single-Mode Transceiver



2xFR4, Single-mode, 800G Twin-port OSFP, 2km

A 2-kilometer twin-port OSFP single mode 2x Far Reach 4-channel transceiver (2xFR4) uses two 2-fiber LC optical connectors each carrying 400Gb/s multiplexed optical signals. These are used to link two Quantum-2 or Spectrum-4 OSFP switches together over long distances and complex networking infrastructures.

Four data signal channels, each using different laser wavelength light, are multiplexed into a single fiber for transmission and filtered out separately in the receiver. This saves costs associated with long fiber reaches, installation, and maintenance.

- The fibers cannot be split into separate transceivers as with parallel fiber 2xDR4 transceiver designs.
- Twin-port OSFP 2xFR4 100G-PAM4 cannot interface with a 4x50G-PAM4 FR4 transceiver due to different modulation schemes and the 2xFR4 cannot downshift to 50G-PAM4.
- Used to link two switches at 800Gb/s or 3 switches at 400Gb/s.
- Transceiver design assumes two optical patch panels in the link budget.
- Pull tab color: Yellow
- Transceiver label: NVIDIA green
- Reach marking: Bright green indicating 2km single mode

- 17 Watts max.
- NVIDIA does not supply the LC-to-LC fiber cables, but they are based on standard optical industry practices and readily available from 3rd party suppliers.

Note: To link two transceivers directly together, one end should have the transmit fiber reversed with the receive fiber to align the transmit lasers with the receive photodetectors. Alternately, introduce and cross over fiber somewhere in the link.

Twin-port 2x400G OSFP 2xFR4 Transceiver and LC Optical Connector



- The 2xFR4 transceiver uses a finned-top but with a lid on top of the fins creating a closed channel for additional cooling.
- **The closed finned top 2xFR4 transceivers can only be used in switch configurations with reverse air flow** where the cooling air enters at the transceiver side of the chassis providing additional cooling.

Twin-port OSFP Open Finned Top

Closed Finned Top



SR4, DR4 400G Single-port OSFP Transceivers

The single-port 4-channel transceivers are offered in both OSFP and QSFP112. QSFP112 was industry standardized much later than the OSFP, is significantly smaller, and fits on the restricted space of the BlueField-3 card with many components.

Single port, 4-channel, **OSFP transceivers are used in network adapters only** and connect to twin-port, 8-channel, OSFP transceivers inserted in switches.

The single-port OSFP is a flat-top, 4-channel version with cooling fins located on top of the adapter board connector cages called riding heat sinks (RHS). The twin-port OSFP and single-port OSFP are both physically the same size, but the single-port supports only one, 4-channel, MPO-12/APC optical connector and contains one, 400Gb/s, optical engine..

- OSFP and QSFP112 have the same technical and internal specifications only with different connector metal shells.
- 400G OSFP transceivers are only used in ConnectX-7/OSFP network adapters -- not used in BlueField-3/QSFP112 DPUs or Quantum-2, or Spectrum-4 switches.
- 400G QSFP112 is used in ConnectX-7/QSFP112 and BlueField-3/QSFP112 -- not used in switches.
- Available multimode up to 50-meters, and single-mode transceiver up to 100-meters.
- Transceiver design assumes two optical patch panels in the link budget.
- No QSA port adapter is offered for inserting QSFP112 into OSFP port adapter for use in OSFP cages.

SR4, Multimode, 400G Single-port OSFP or QSFP112, 50-meters

- 4-channel 100G-PAM4
- MPO-12/APC 4-channel optical connector
- Pull tab color: Tan
- Transceiver label: NVIDIA green
- Reach marking: Blue
- 9 Watts max.

DR4, Single mode, 400G Single-port OSFP or QSFP112, 100-meters

- 4-channel 100G-PAM4
- MPO-12/APC 4-channel optical connector
- Pull tab color: Yellow
- Transceiver label: NVIDIA green
- Reach marking: White
- 9 Watts max.

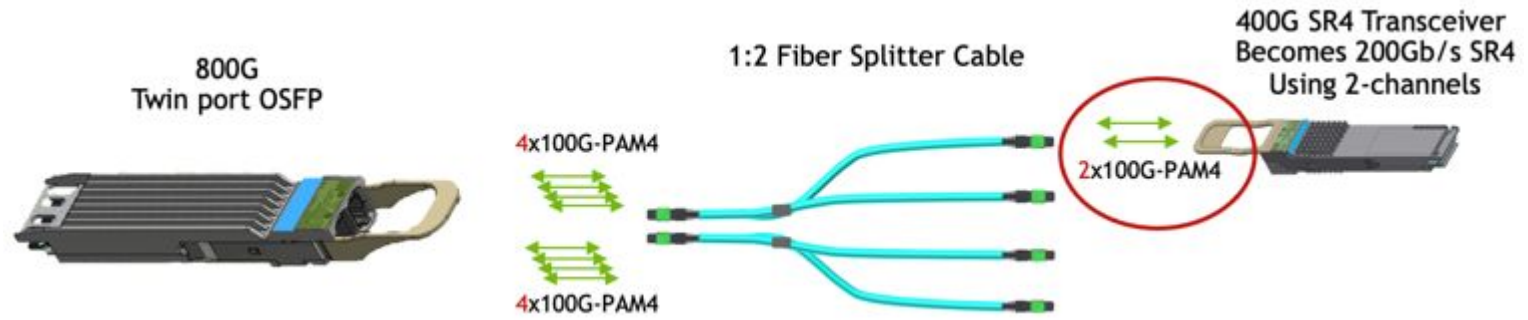
Multimode and Single-mode OSFP and QSFP112 Single-port Transceivers



200Gb/s Transceivers are Not Offered

- 200Gb/s OSFP or QSFP112 transceivers (NDR200) are **not** offered as the 100G-PAM4 electronics cannot downshift to 50G-PAM4.
- When 400Gb/s transceivers are used with 1:2 splitter fibers, only two channels are activated creating a 200Gb/s transceiver.
- Power consumption drops from 8 Watts multimode to 5.5 Watts; single mode from 9 Watts to 6.5 Watts.

Two Splitter Cables Create 200G Transceivers from 400G Transceivers



Transceiver and Fiber Colors and Labeling

NVIDIA transceivers have an NVIDIA green label on the connector body:

- 50-meter multimode transceivers have a blue stripe on the connector body and tan pull tabs.
- 100-meter single-mode transceivers have a white stripe and yellow pull tabs.
- 500-meter single-mode transceivers have a purple stripe and yellow pull tabs.
- 2k-meter single-mode transceivers have a bright green stripe and yellow pull tabs.

Multimode fiber: aqua colored cable jacket

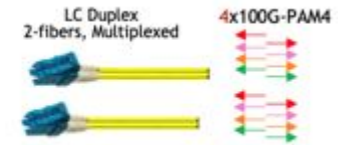
Single-mode fiber: yellow cable jacket

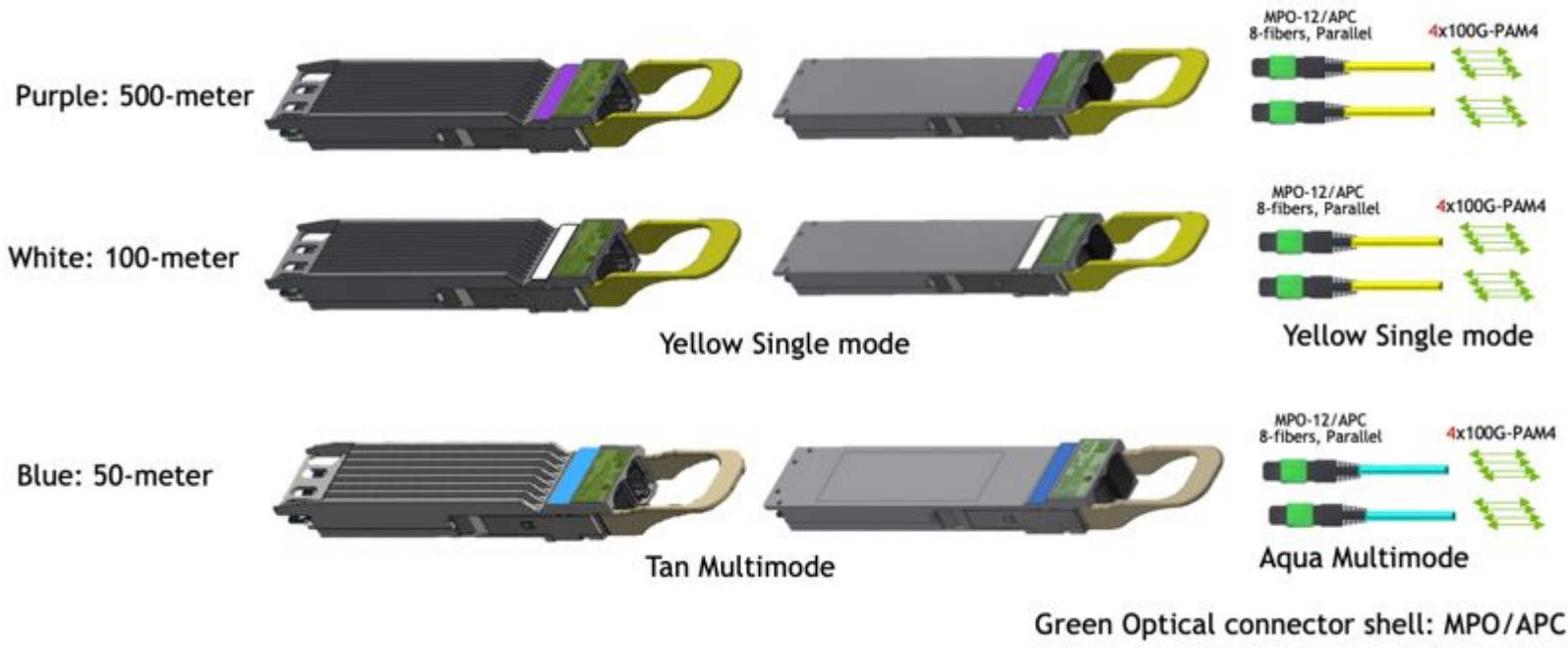
MPO-12/APC optical connector: green shell

MPO-12/UPC optical connector: blue shell

Cable and Transceiver Labeling and Colors

Green: 2Km





Copper DAC and ACC Cables Overview

Direct Attach Copper (DAC) Cables

Copper cabling is the lowest cost, lowest latency, and lowest power way to interconnect high-speed systems together. DAC cables “directly attach” the electrical subsystems together, hence the name Direct Attach Copper cables (DACs). DAC copper cables are mainly used for inside system racks linking compute servers to storage subsystems within the 3-meter maximum length.

- Offered in twin-port OSFP, OSFP, and QSFP112 ends
- DACs offer length of 0.5m to 2-meters for straight cables, and up to 3-meters for splitters
- Power consumption is 0.1 Watts per end
- Thin 30AWG wire is used for short lengths up to 1-meter for 800G-to-800G cables
- Thicker 26AWG wire with more shielding is used for 2- to 3-meters and splitters
- Two different splitters cables are offered up to 3-meters for 1:2 400Gb/s 4-channel and 1:4 200Gb/s 2-channel ends for use in ConnectX-7 and BlueField-3 adapters
- BlueField-3 uses only QSFP112, and ConnectX-7 is offered in both QSFP112 and OSFP

DAC Twin-port OSFP Straight and Splitters using OSFP and QSFP112 Ends



Active Copper Cables (ACCs)

ACC cables are used to extend the low-cost, low-power, and low latency of copper cables from 3- to 5-meters which can span several racks. ACCs are essentially DAC cables with an additional IC in each end to boost the signal power and noise reduction. This cable uses a pre-emphasis type of circuit, aka linear ACC, that learns the noise signal and inverts it before sending down the cable creating an anti-noise. The noise and anti-noise energies cancel each other out preserving the data signal. Unlike HDR/200GbE ACCs that use DSPs, pre-emphasis ACCs have very low power and low latency.

- ACC configuration shadows the DAC line at longer lengths.
- Offered in twin-port OSFP, OSFP, and QSFP112 ends.
- ACCs offer length of up to 3-meters for the straight cables, and 4- and 5-meters for splitters.
- Power consumption is 1.5 Watts on the 8-channel ends, and 0.6W and 0.3W on the split ends.
- Thin 30AWG wire is used for short lengths of 3- and 4-meters.
- Thicker 26AWG wire with more shielding is used for 5-meter splitters.
- Two different splitter cables are offered up to 3-meters for 1:2 400Gb/s 4-channel and 1:4 200Gb/s 2-channel ends for use in ConnectX-7 and BlueField-3 adapters.
- BlueField-3 uses only QSFP112, and ConnectX-7 is offered in both QSFP112 and OSFP.

ACC Twin-port OSFP Straight and Splitters OSFP and QSFP112 Ends



Backwards Compatible DACs and AOCs to 200Gb/s 50G-PAM4

DACs and AOC splitters are offered for backwards compatibility by downshifting the modulation rate of switches to 50G-PAM4 enabling links Quantum-2 and SN5600 OSFP Ethernet switches to link to 200G HDR Quantum QSFP56 and Spectrum-3 switches and to 50G-PAM4 based ConnectX-6 and BlueField-2 DPUs using QSFP56 connector.

50G-PAM4 DACs

1:2 DAC splitters using twin-port OSFP 2x200G-to-two 200G QSFP56s based on 4x50G-PAM4 are offered from 1- to 2-meters.

As DACs are simply a bunch of wires, they can be downshifted to 2x 100GbE or HDR100 rates by the switch.

The 1:2 DAC splitter also supports 2x100G-to-two 100Gb/s QSFP56 links as 4x25G-NRZ ends for standard 100GbE and EDR rates.

A 1:4 DAC splitter is offered with 2x200G-to-four 100Gb/s using QSFP56 based on 2-channel 50G-PAM4 ends for 100G but does not support 4x25G-NRZ.

50G-PAM4 AOCs

1:2 AOC splitters using twin-port OSFP 2x200G-to-two 200G QSFP56s based on 50G-PAM4 is offered up to 30-meters for linking 400GbE/NDR switches to HDR/200GbE switches, adapters. In the Quantum-2 and Spectrum SN5600 switches, the 8-channel cages are down shifted to 50G-PAM4 modulation and split into 1:2 200Gb/s using QSFP56. This AOC can also support 200G-to-two 100Gb/s QSFP56 links as 2x 50G-PAM4.

These 2 AOCs are the only optical links available to link between 100G-PAM4 OSFP and 50G-PAM4 QSFP56 systems as separate transceivers are not available.

DACs and AOCs for backwards compatibility to HDR/200GbE, HDR100 and EDR/100GbE



No AOCs for 100G-PAM4 Series

AOCs consist of two or more transceivers with the fibers bonded inside and not detachable. This creates an entire, plug-and-play cable assembly like copper cables but at much longer lengths. While used extensively in EDR, HDR and 100GbE, 200GbE systems, **AOCs are not offered for 100G-PAM4** Ethernet or InfiniBand systems. The twin-port OSFP and OSFP ends are heavy and very likely to break the fibers during installation. An 800G-to-4x 200G AOC would have five large OSFP connectors.

Additionally, complex fiber infrastructures are better addressed with connectorized fibers, and they can support future line rate upgrades with minimal changes.

Where to Find More LinkX Documentation

This user guide is to be used in conjunction with other documents located in folders in docs.nvidia.com/networking/ > Interconnect. This site is where the following LinkX cables and transceivers documents are provided.

LinkX Overview Documents:	Review of parts, important notes, and configuration details for linking to NVIDIA switches and adapters <ul style="list-style-type: none">• LinkX Cables and Transceivers Guide to Key Technologies• LinkX User Guide for 400Gb/s 100G-PAM4 OSFP & QSFP112-based Cables and Transceivers (this document)• LinkX User Guide for 400Gb/s and 200Gb/s using 50G-PAM4 and 100Gb/s using 25G-NRZ Modulation Cables and Transceivers
Configuration Maps:	Picture and part number-based PowerPoint® slides for every configuration with NVIDIA switches, network adapters, and DGX GPU systems for 100G-PAM4, 50G-PAM4, 25G-NRZ cables and transceivers <ul style="list-style-type: none">• Configuration Maps
Parts Lists:	Tables summarize by speed, form factor, connector, power, reach, etc. and hyperlinks to individual products specs <ul style="list-style-type: none">• 400Gb/s (100G-PAM4) Transceivers and Fiber Parts List• 400Gb/s and 200Gb/s (50G-PAM4) and 100Gb/s (25G-NRZ) Cables and Transceivers Parts List using QSFP-DD, QSFP56, QSFP28, SFP28
Product Specifications:	10-to-16-page detailed hardware datasheets with physical, thermal, electrical, and optical specifications for each product <ul style="list-style-type: none">• docs.nvidia.com/networking/ > Interconnect > <i>select speed and type</i>
Additional Docs:	<ul style="list-style-type: none">• NVIDIA Cable Management Guidelines and FAQ

Document Revision History

Version	Date	Description of Change
1.0	August 2023	Initial release

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. Neither NVIDIA Corporation nor any of its direct or indirect subsidiaries and affiliates (collectively: "NVIDIA") make any representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice. Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete. NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer ("Terms of Sale"). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer's own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer's sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer's product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT,



INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA's aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, and Mellanox are trademarks and/or registered trademarks of NVIDIA Corporation and/or Mellanox Technologies Ltd. in the U.S. and in other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

© 2023 NVIDIA Corporation & affiliates. All Rights Reserved.

