The Journey to 400/800G has Begun

Which road will you take?

Author: Gary Bernstein, RCDD & CDCD - Global Data Center Specialist, Siemon

Reading Time: 15 minutes

The Road to 400G and Beyond

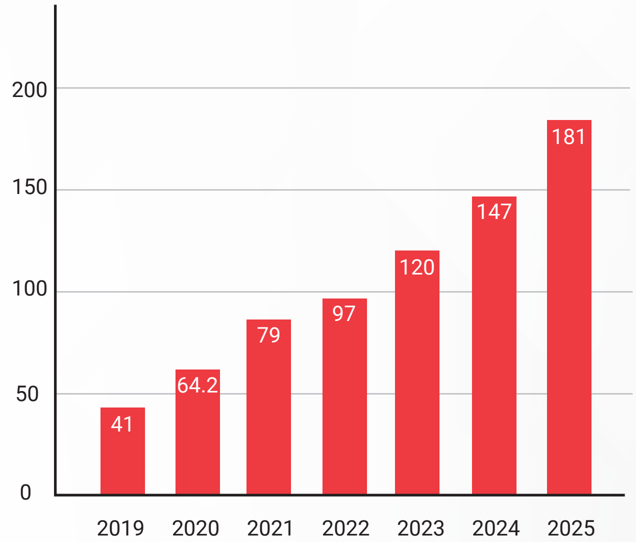
While enterprise data centers are just starting to adopt 100 Gigabit speeds for switch uplinks, industry standards and active equipment for 400 Gigabit applications are now available. These speeds are still a few years out for many enterprise data centers but are already rapidly gaining ground across large hyperscale cloud providers who define the market and propel technology advancements that ultimately will trickle down to the enterprise in years to come. In fact, these trendsetters are already pushing speeds to 800 Gigabit and working to develop fiber optic technologies that enable the viability of future 1.6 and 3.2 Terabit speeds. Industry standards aren't far behind with the IEEE Beyond 400 Gb/s Ethernet Study group already defining physical objectives for both 800 Gigabit and 1.6 Terabit fiber optic applications.

With emerging technologies continually driving the need for higher bandwidth and the road to 400 Gigabit already laid, data center owners and operators of all types and sizes need to be aware of the current standards-based options and how implementations are taking shape, while keeping an eye on developments for 800 Gigabit and beyond. This will help them determine which road towards 400 Gigabit will ultimately best align with their organization's needs and how to optimize their infrastructure to harness the full potential of the available technologies that will digitally transform and sustain the growth of their business.

With demands for faster transmission rates and larger bandwidth constanly increasing, are you on the Road to 400G?

Key Drivers, Trends and Technology

Data centers have cemented their position as the fulcrum for every organization's critical operations. Global digitization continues to expand, resulting in evermore need for greater bandwidth to support the resulting additional data traffic. This trend has led to predictions that data center traffic will grow by more than 180 zettabytes (ZB) by 2025. To put that into perspective, 1 ZB of data is equivalent to a billion Terabytes (TB) or a trillion Gigabytes (GB). To support that traffic, Frost & Sullivan forecasts $432 billion will be invested in the data center market by 2025. There are several key drivers and trends driving this growth and advancements in technology that are fueling the need for 400 Gigabit.

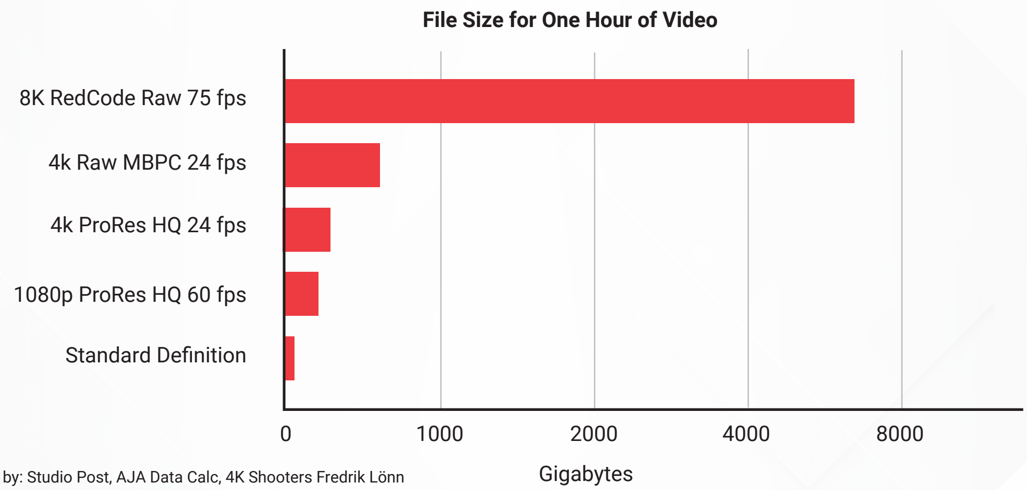


Source: Statista 2021

Enterprise businesses embarking on digital transformation and adopting emerging applications are driving the need for 400 Gigabit applications in cloud data centers.

Emerging Applications

While the overall amount of data is increasing, the following emerging applications are also driving file sizes to an all-time high and demanding additional compute power, bandwidth, and low-latency transmission. The use of uncompressed high-resolution video, such as RedCode Raw is on the rise in the creation of high quality content in the film and media industry. This graph demonstrates how this format compares to other standard video formats for data requirements and reinforces the growing data requirements as this technology evolves.



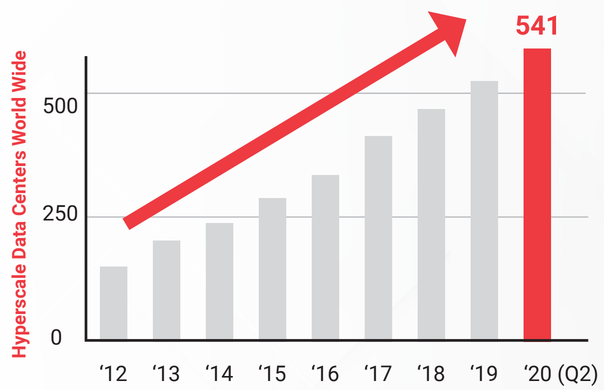
Calculated by: Studio Post, AJA Data Calc, 4K Shooters Fredrik Lönn

- Large enterprise intelligent building and Industry 4.0 initiatives are generating increasing amounts of IoT/IIoT data, with the number of connected devices projected to exceed 30 billion by 2025.
- Applications like virtual reality (VR), artificial intelligence (AI), machine learning (ML), autonomous cars and advanced data analytics are expediting the demand for extremely low latency and higher-speed server connections to support them.
- Applications in healthcare and finance, including high-definition MRI imaging, virtual telehealth, high-frequency trading, and online banking are demanding higher bandwidths.
- Uncompressed high-resolution video, computer animation, and visual effects in professional media and gaming are increasing file sizes exponentially.
- The widespread adoption of work-from-home strategies, distance learning, online retail, and video streaming are changing work patterns, buyer behaviors, and business processes that demand advanced processing requirements.
- Rollout of 5G mobile technology and edge computing, along with increasing virtualization and software-defined networking (SDN) are placing new demands on data center environments to deliver high performance connectivity in challenging environments.

Increasing Cloud Adoption

Enterprise businesses embarking on digital transformation and adopting emerging applications are driving the need for 400 Gigabit applications in cloud data centers. The increased usage of work-from-home strategies and increased online transactions and video traffic has also significantly fueled cloud adoption, with cloud spending rising 25% during the first quarter of 2020 according to findings from Synergy Research Group. As a result, spend in data center hardware has increased 25% in the public cloud and decreased by 3% in enterprise and non-cloud-service provider markets.

Even in the face of economic downturn from the pandemic, cloud spending is expected to continue increasing with Cisco predicting 94% of all workloads running in some form of cloud environment by the end of 2021 and Gartner predicting spending on public cloud services to reach nearly $700 billion by 2025. With the increase in cloud adoption, the number of hyperscale data centers has also increased, reaching 541 by the second quarter of 2020.

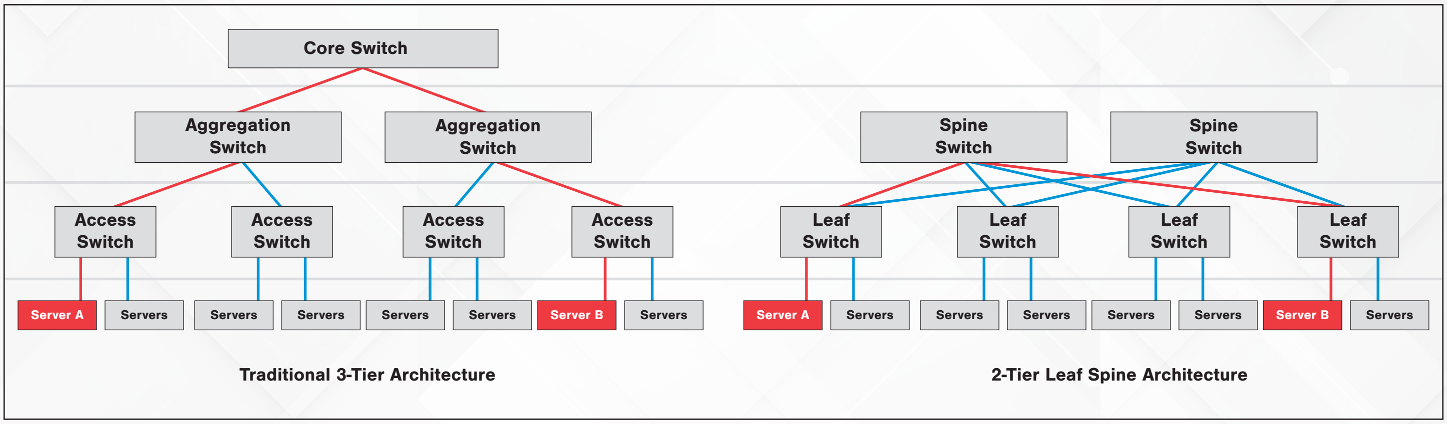


Source: Synergy Research Group

Highly-Virtualized, Low-Latency Architecture

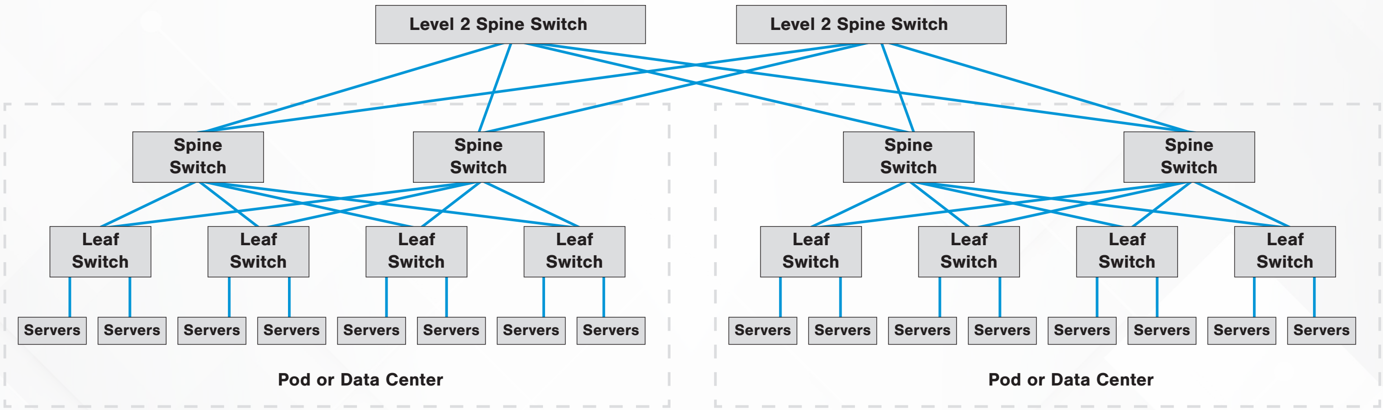
To support these changes, data centers continue to become more virtualized, with compute resources able to reside anywhere and implementing SDN to manage virtual machines and increased network traffic. This has driven a shift from traditional three-tier switch architecture with a north-south data traffic pattern to newer leaf-spine switch fabric architectures that enable east-west traffic and lower latency.

In a leaf-spine architecture, every leaf switch connects to every spine switch, reducing the number of switch hops that traffic must traverse between virtualized servers and significantly lowering latency, while providing a superior level of redundancy as shown below where Server A needs to "talk" to Server B.



Large high-performance computing cloud data centers are increasingly interconnecting multiple spine-leaf networks via second level spine switches. Referred to as a super-spine architecture, this approach connects separate functional areas (sometimes referred to as pods or halls) and enables data center interconnects (DCIs) that connect multiple dispersed data centers. Super-spine architecture results in massive amounts of data passing between interconnected networks, driving the need for even more bandwidth in these environments. Because each interconnected network can be easily replicated in a modular manner, super-spine architecture is also highly scalable, enabling faster turn-up to meet the exploding demand for cloud-based services.

Super Spine Architecture

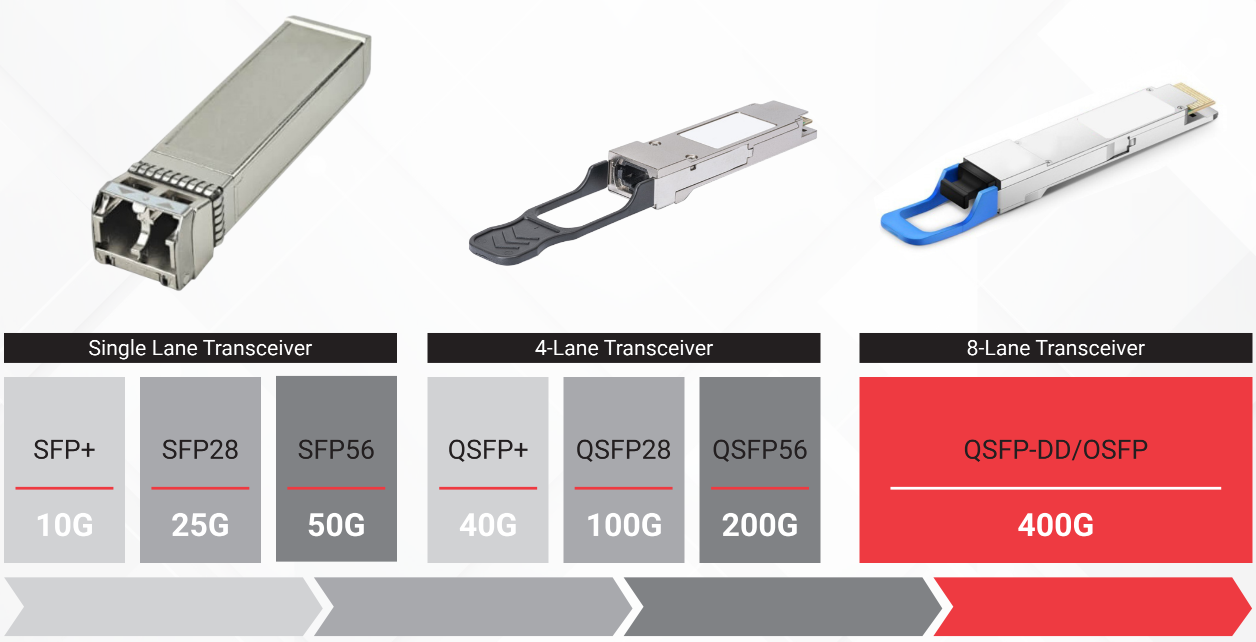


To support emerging real-time and cloud-native applications and 5G mobile technology that require extreme low-latency transmission, edge data centers that process and analyze data closer to devices and users are rapidly being deployed at customer locations, 5G access sites, central offices, and regional colocation data centers. Edge data centers also require higher bandwidth computing for local processing and are driving the need for next-generation 400 Gigabit speeds to send aggregated data back to core and cloud data centers and to interconnect with other edge data centers.

Technology Advancements

Advancements in signaling, transceiver, and optical technology are also key drivers of 400 Gigabit, providing the technical and economical means that make these speeds viable in the market. One significant advancement is the development of the four-level pulse amplitude modulation (PAM4) signaling scheme that offers twice the bit rate of previous non-return-to-zero (NRZ) signaling, enabling 25, 50, and 100 Gb/s per lane. This enables a more efficient migration path of 25-50-100-200-400-800 Gigabit and reduces the amount of infrastructure required to support them. For example, the 40 Gigabit 40GBASE-SR4 application that uses NRZ 10 Gb/s per lane encoding requires 8 fibers, with 4 transmitting and 4 receiving at 10 Gb/s. PAM4 signaling at 50 Gb/s and 100 Gb/s per lane support 200 Gigabit and 400 Gigabit respectively over 8 fibers.

Transceiver technology has also advanced, migrating from pluggable transceiver modules based on 10 Gb/s per lane NRZ signaling to 50 Gb/s per lane with PAM4. The latest QSFP-DD and OSFP pluggable transceivers are 8-lane 50 Gb/s pluggable transceivers that support 400 Gigabit. There are now several transceivers available on the market for a variety of multimode and singlemode applications based on the QSFP-DD and OSFP form factors. These pluggable transceivers with fiber cabling and connectivity allows the flexibility to support different transmission speeds and utilize existing hardware.



Optical advancements that also enable the viability of 400 Gigabit include short-wave division multiplexing (SWDM) that transmits data over multiple wavelengths on a single multimode fiber. Rather than transmitting at 50 Gb/s over 16 fibers to achieve 400 Gigabit speeds (i.e., 8 fibers receiving and 8 fibers transmitting at 50 Gb/s), SWDM technology transmits 50 Gb/s on two different wavelengths over a single fiber, thereby cutting the number of fibers in half.

In singlemode applications, another technology enabler is new lower-power, cost-effective short-reach transceiver technology that supports 400 Gigabit to 500 meters (m). Because link lengths in a data center are typically 500 m or less, it does not economically make sense for data centers to deploy expensive high-power lasers used in long-haul outside plant singlemode applications.

In conjunction with these technology advancements, NICs that provide the interface between servers and the network have evolved from 10 to 25 to 50 and now 100 Gigabit speeds. With data centers needing highly virtualized, low-latency environments for emerging technologies, migration to faster server speeds is happening in both enterprise and in hyperscale cloud data centers. Today's enterprise data centers are migrating to 25 and 50 Gigabit server NICs and cloud data centers are migrating to 50 and 100 Gigabit server NICs.

With key drivers, trends, and technology in place, industry standards for 400 Gigabit applications have already been developed.

The 400 Gig Landscape

With key drivers, trends, and technology in place, industry standards for 400 Gigabit applications have already been developed. There are various ways to configure these applications within switch-to-switch and switch-to-server links. While deployments are happening first within cloud data centers, there are options for the enterprise that are worth keeping an eye on for the future.

Application Standards

Current IEEE Ethernet standards for 400 Gigabit encompass both multimode and singlemode applications for various distances and forms of transmission via existing fiber cabling and connectivity as shown in the table below.

Transceiver	STD	Commercially Available	Form Factor	Breakout Option	Fiber Type	Distance (meters)	# of Fibers	Connector
400G-FR4	IEEE802.3cu/MSA	Yes	QSFP-DD,OSFP	No	OS2	2000	2	LC
400G-DR4	IEEE802.3bs	Yes	QSFP-DD,OSFP	Yes	OS2	500	8	12F MTP
400G-SR8	IEEE802.3cm	Yes	QSFP-DD,OSFP	Yes	OM3/OM4	70/100	16	16F/24F MTP
400G-SR4.2(BD)	IEEE802.3cm/MSA	Q1-2022	QSFP-DD	Yes	OM3/OM4/OM5	70/100/150	8	12F MTP
Coming Soon: 400G-VR4	802.3cu	Q2-2022	TBD	Yes	OM3/OM4	30/50	8	12F MTP
Coming Soon: 400G-SR4	802.3db	Q2-2022	TBD	Yes	OM3/OM4	70/100	8	12F MTP

Short-reach DR4 singlemode and SR4 multimode applications use parallel optics with 4 fibers transmitting and 4 receiving, which are supported by familiar 8- or 12-fiber MTP connectivity (i.e., Base-8 or Base-12) that has been used for several years in 40 and 100 Gigabit parallel optic applications. It's important to note that while 400GBASE-SR4.2 uses parallel optics, it also uses more expensive SWDM technology with two 50 Gb/s wavelengths per fiber to enable 100 Gb/s per fiber. While it is supported via OM3, OM4, and OM5 multimode, 400GBASE-SR4.2 is also the only application currently targeted to OM5 multimode.

400GBASE-SR8 requires 16-fiber multimode MTP connectivity that has no installed base and limited market availability. Industry standards bodies are however looking more closely at the viability of 16-fiber MTP connectivity and IEC is developing connector specifications, as these connectors offer the ability to breakout 400 Gigabit to eight 50 Gigabit links (i.e., 8X50) and will likely support 8X100 Gigabit breakouts in future 800 Gigabit applications. It should be noted that due to 400 Gigabit systems exhibiting a poor signal-to-noise ratio and therefore requiring better return loss performance, it is recommended that these deployments use angled physical contact (APC) connectors that feature an 8-degree angled end face that reduces signals reflecting to transmitters.

In addition to the current standards above, IEEE is working on new lower-cost multimode options in 802.3db, including 400GBASE-SR4 (short-reach) and 400GBASE-VR4 (very short reach) that will be based on 100 Gb/s per lane PAM4 signaling. These applications will likely become popular due to their simplicity and lower cost compared to existing 400GBASE-SR8 and 400GBASE-SR4.2 applications. 400GBASE-SR4 and 400GBASE-VR4 will be easily and cost-effectively supported using existing Base-8 or Base-12 MTP OM4 multimode fiber, eliminating the need for 16-fiber multimode MTP connectivity and more expensive SWDM technology and OM5 multimode.

Migration Path

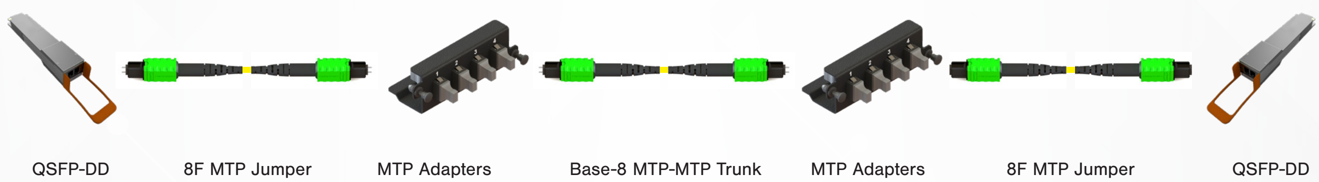
Enterprise data centers and cloud data centers have become quite different over the past few years. Server and uplink speeds in these environments are at distinct phases and require diverse solutions to support their applications. Enterprise data centers predominantly deploy multimode fiber and have lower bandwidth requirements and cloud data centers are singlemode fiber based and demand much higher bandwidth. These distinct environments have different migration paths using current 400 Gigabit applications as shown below.

	Enterprise Data Centers		Cloud Data Centers	
	Server	Uplinks	Server	Uplinks
Current	1/10G	10/40G	10/25G	40/100G
Future	25G	100G	50G	200G
	OR		OR	
	50G	200G	100G	400G
	OR		OR	
	100G	400G	200G	800G

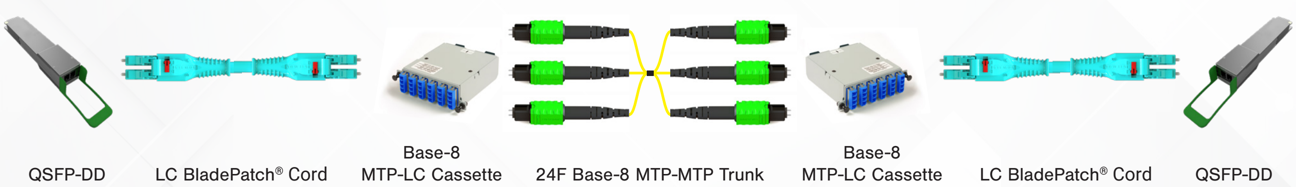
Switch-to-Switch Links

As previously mentioned, 400 Gigabit applications have initially been deployed in cloud data centers for switch-to-switch uplinks, either leaf to spine or spine to super-spine. These switch-to-switch links can be deployed using singlemode Base-8 400GBASE-DR4, singlemode duplex 400GBASE-FR4/FR8, or multimode Base-8 400GBASE-SR4.2 as shown below.

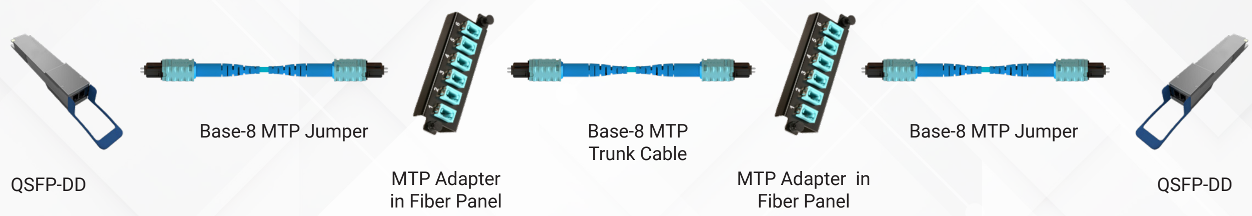
8-Fiber Singlemode 400 Gigabit Switch-to-Switch Channel
400GBASE-DR4



Duplex WDM Singlemode 400 Gigabit Switch-to-Switch Channel
400GBASE-FR4/FR8



8-Fiber 400 Gigabit Multimode Switch-to-Switch Channel
400GBASE-SR4.2

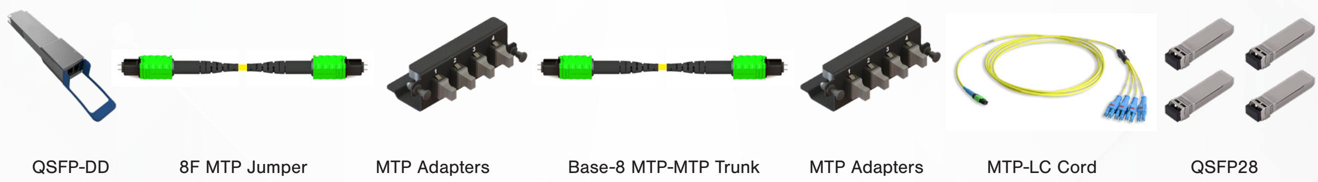


Breakout Configurations

Large enterprise companies are starting to evaluate migration to 400 Gigabit. The primary deployment will be breakout configurations where a single 400 Gigabit switch port breaks out to eight 50 Gb/s connections or four 100 Gigabit connections to optimize port utilization and switch density for reduced cost.

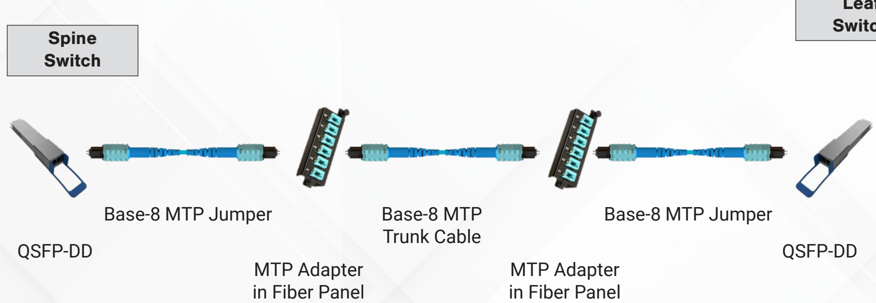
The installed base of plug-and-play Base-8 MTP connectivity can be leveraged in breakout applications such as 4X100 Gig or 8X50 Gig. Rather than using MTP-LC cassettes and duplex jumpers, breakouts can also be achieved via MTP adapters and MTP-to-LC hybrid breakout assemblies as shown in the example below of a 400GBASE-DR4 4X100 Gig breakout application.

8-Fiber Singlemode 400GBASE-DR4 4X100 Breakout

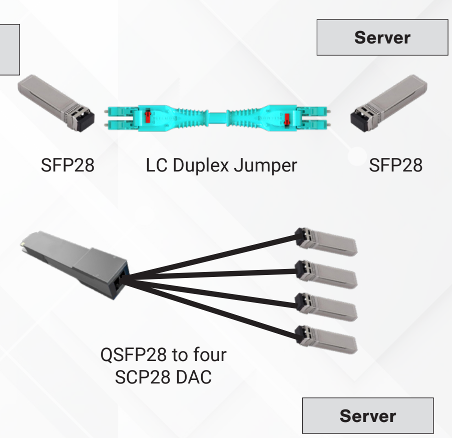


In enterprise data centers where server connections are just beginning to migrate to 25 Gigabit, switch-to-server links can be achieved as direct connections using QSFP-DD to QSFP-28 Direct Attach Cables (DACs) or active optical cables (AOCs), or via structured cabling using transceivers and duplex fiber jumpers as shown below.

8-Fiber 400 Gigabit Multimode Switch-to-Switch Channel 400GBASE-SR4.2



25 Gigabit Switch-to-Server Channel LC Dulex Jumper Option

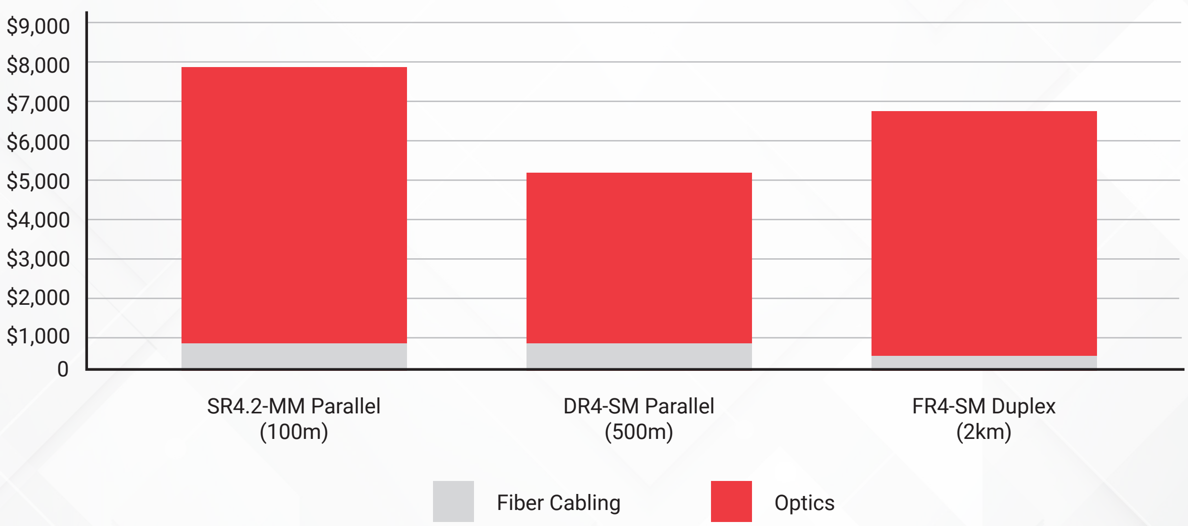


400G Cost Considerations

While there are several options for 400 Gigabit deployments, there is no single infrastructure design for every data center. Data center managers need to carefully weigh their options against future plans and consider their ability to leverage existing infrastructure based on cable and connectivity, as well as channel lengths.

While singlemode deployments have traditionally cost more due to more expensive higher-power lasers compared to multimode that uses low-cost vertical-cavity surface-emitting laser (VCSEL) technology, there may be situations where short-reach singlemode is ultimately more cost effective than using SWDM multimode technology. Due to high volume usage by hyperscalers, the cost of singlemode optics have come down considerably to match or even be less than multimode solutions. As shown in the chart below, this is making singlemode much more attractive for many data centers. As previously mentioned, because the upcoming 8-fiber 400GBASE-SR4 parallel optics application eliminates the need for more expensive SWDM technology and OM5 multimode and can be supported over the installed base of OM4 multimode, it is expected that this application will eventually provide a much more cost-effective option for the enterprise compared to 400GBASE-SR4.2.

400 Gigabit Channel Pricing Comparison
Fiber Cabling & Optical Transceivers



Notes:
- Uses estimated Average List Sell Prices of fiber cabling and transceiveres in the market.
- Includes 30 m fiber cabling channel with OM4 Multimode and OS2 Singlemode fiber trunk, cassettes/adapters and jumpers.
- Includes 2x transceivers per channel.

Beyond 400G

While 400 Gigabit is now a reality, it is a waypoint on a longer journey. The EEE Beyond 400 Gb/s Ethernet Study Group is already working to define objectives for 800 Gigabit based on 400 Gigabit logic, including the following objectives targeted to the data center and slated for 2025:

- Over 8 pairs of multimode fiber for short-reach to at least 50 m
- Over 8 pairs of multimode fiber to at least 100 m
- Over 8 pairs of singlemode fiber for short-reach to at least 500 m
- Over 8 pairs of singlemode fiber to at least 2 km
- Over 4 pairs of singlemode fiber for short-reach to at least 500 m
- Over 4 pairs of singlemode fiber to at least 2 km
- Over 4 wavelengths on singlemode fiber in each direction to at least 2 km

In addition, the IEEE is working to define 1.6 Terabit applications, with objectives that currently include over 8 pairs of singlemode fiber with lengths up to 500 m and 2 km. The 8-pair objectives for both multimode and singlemode applications may ultimately drive the need for Base-16 MTP connectivity.

Hyperscale cloud providers like Google, Microsoft, and Facebook are also working to develop technologies through multi-source agreements (MSAs) for 800 Gigabit and 1.6 Terabit speeds, including a new 16-lane pluggable transceiver module and other optical solutions with key challenges being power consumption.

While no one knows exactly how applications beyond 400 Gigabit will shape up, 800 Gigabit and beyond applications may drive further advancements in both MTP and duplex connectivity, such as small-form-factor options. Until a 200 Gb/s PAM4 is achieved (also within scope of the IEEE Beyond 400 Gb/s Ethernet Study Group), these applications will require a lot more fiber and performance of the cabling infrastructure and proper cable management will be more critical.

The Road Ahead

With any new technological development, there are always challenges to be overcome and unmapped routes to be travelled. However, the road to 400G is here and ready and waiting for those organizations ready to take that step. If these speeds are not currently in your organizations mid-to-long-term view, as we've seen from other similar developments over the years, these advancements will only improve your options as you migrate to higher speeds along this route. This will provide you with a variety of roads to travel using tried and tested solutions that help you reach your planned destination.

Regardless of the journey you are on, data centers of all types and sizes can rest assured that Siemon's superior stability as a fifth-generation company, with high-performance solutions recognized for world-class quality and expert technical data center design services, combined with a comprehensive data center partner ecosystem and global sales, logistics, and installation capabilities, makes us the go-to industry leader to support your organization's migration from 10 to 400G and beyond.

Meet The Author

Gary Bernstein is the Global Data Center Solutions Specialist at Siemon with more than 25 years of industry experience and extensive knowledge in data center infrastructure, telecommunications, and copper and fiber structured cabling systems. He has been a member of the TIA TR42.7 Copper Cabling Committee, TIA TR42.11 Optical Fiber Committee and various IEEE802.3 task forces and study groups including 40/100G "ba", 50/100/200G "cd", 200/400G "bs" Task Force and Beyond 400G.

North America
P: (1) 860 945 4200

Asia Pacific
P: (61) 2 8977 7500

Latin America
P: (571) 657 1950/51/52

Europe
P: (44) 0 1932 571771

China
P: (86) 215385 0303

India, Middle East & Africa
P: (971) 4 3689743

Siemon Interconnect Solutions
P: (1) 860 945 4213
www.siemon.com/SIS

Mexico
P: (521) 556 387 7708/09/10

WWW.SIEMON.COM

WP_400G_RevC 2/22