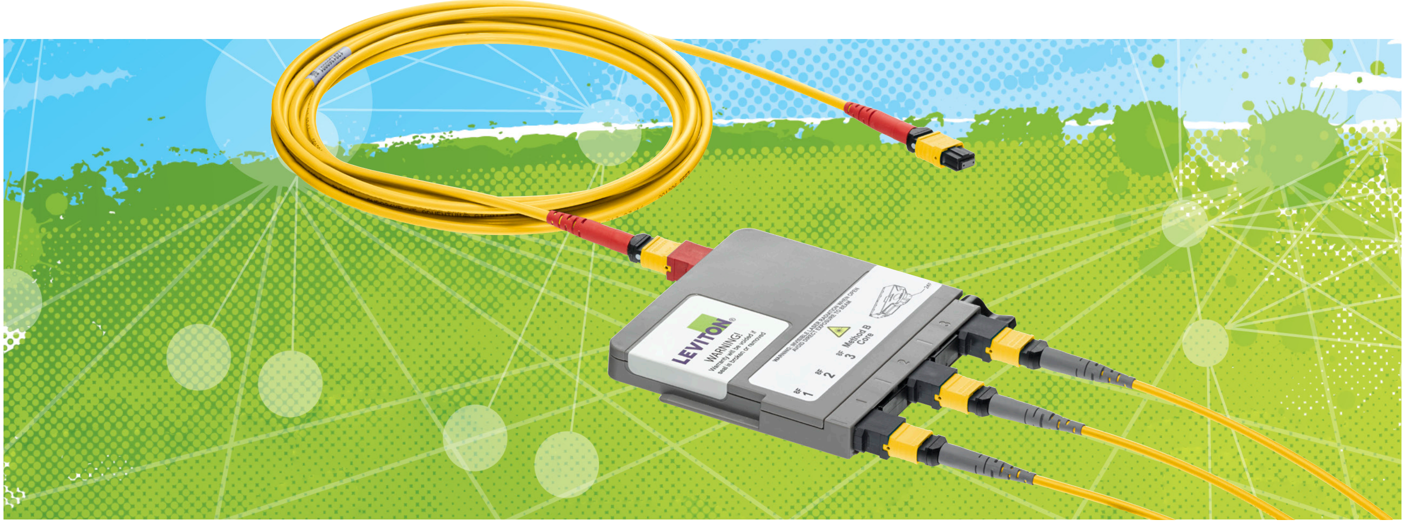
WHITE PAPER



Network Design for Greenfield Data Centers

Key considerations for new data centers, including general architecture, distances, data rates, fiber types, and more.

Mike Connaughton, RCDD, CDCD
Senior Product Manager, Leviton Network Solutions

TABLE OF CONTENTS

In recent years, we have witnessed economic volatility due to the global pandemic, but the data center market remained largely unaffected. Demand for supporting digital business infrastructure never wavered, and while some new data center construction slowed initially, 2021 saw robust expansion. In primary data center markets in North America, new data center construction at the midpoint of 2021 was up 42% year-over-year, with 527 megawatts of capacity under construction, according to the North American Data Center Trends Report from the CBRE Group, Inc. Much of the new global data center or “greenfield” construction comes from hyperscale cloud providers and colocation data centers expanding their regional footprints, or new construction in regions of the world where there is less existing infrastructure in place.

With a greenfield project, network designers may be starting from scratch with only an empty room or even just a concrete foundation. This means that there will be a significantly higher up-front investment required compared to updating an existing network, but there is an opportunity to create the right network architecture from the start since there are fewer existing impediments when making technology and design choices. The only real constraints in these projects are size, power, and money. That differs from brownfield data center upgrades, which face added constraints from the network infrastructure in place, such as the existing cable type or cabling layout. Brownfield upgrades can also be more constrained by time since updates to an operating data center may involve network disruption and downtime.

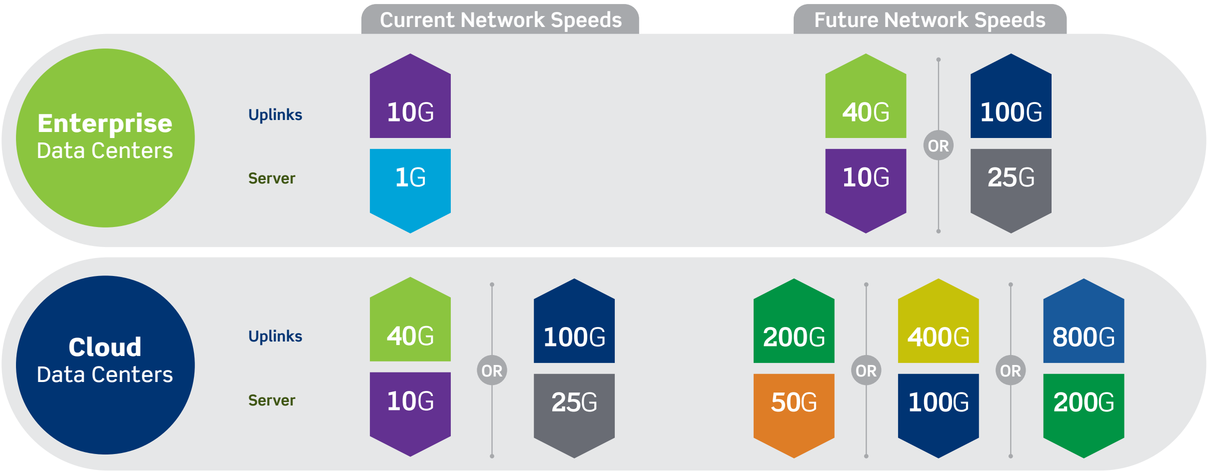
When data center operators begin the process of defining the makeup of their network, there are some important practical design questions they need to answer, including general architecture and distance needs, data rate requirements, fiber types, and cost. This article will address these and other key considerations.

HIGH-LEVEL QUESTIONS

The first decision to make for a new data center network will be to choose the right architecture. This involves asking questions like “What are the business needs of the data center? What are the workloads and processes that will run in the data center?” While these questions are bigger than just the physical layer, they do lead to size and speed decisions that shape the makeup of the cabling system.

The type of data center will affect the speed requirements of the network and the transceiver types most applicable to the design. Over the past 10 years, a split has formed between enterprise and cloud provider data centers when it comes to network migration patterns. For example, in previous years, cloud provider networks have operated at 40 Gb/s uplinks to the switch and 10 Gb/s from switch to server. These networks are now moving to 100 Gb/s uplinks and 25 Gb/s downlinks to the server, with some even preparing to migrate to 200 and 400 Gb/s uplinks and 50 and 100 Gb/s at the server. Enterprise-class switches tend to lag cloud data centers in these speeds, as shown in Figure 1. The different business purposes of enterprise, cloud, and hyperscale data centers have influenced different cabling architectures as a result.

Figure 1: Current vs Future Data Center Network Configurations

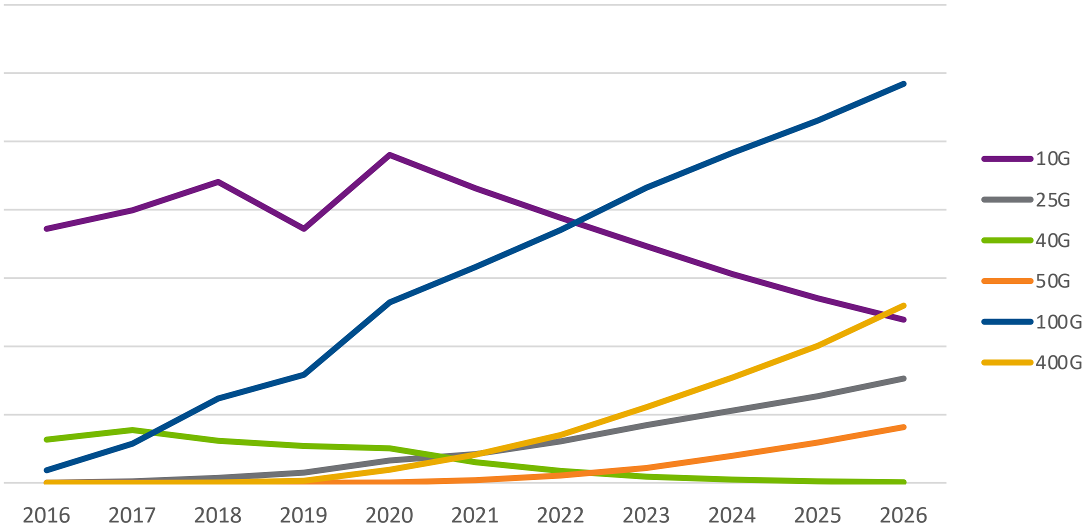


DATA RATE AND TRANSCEIVER CHOICES

As data centers move to more 100 Gb/s speeds for switches and servers, 10 Gb/s and 40 Gb/s transceivers will continue to decline. This trend is reflected in Figure 2 below, from the LightCounting March 2021 High Speed Ethernet Optics Report. It is not a coincidence that these two speeds are both on the decline. Uplink speeds are often faster than downlink speeds by a factor of four, so as 10 Gb/s transceiver ports have declined at the server so have 40 Gb/s speeds at the switch.

The Figure 2 forecast also shows that 100 Gb/s will soon surpass 10 Gb/s as the most widely deployed optical Ethernet connection speed, if it has not already. And it will remain the most popular transceiver speed in the foreseeable future.

Figure 2: High Speed Transceiver Forecast



Source: LightCounting March 2021 High speed Ethernet Optics Report

Server connection speeds

When evaluating transceivers at the server, 10 Gb/s has been declining for some time, as mentioned earlier. But they are still popular. Of those in use, 10G-SR and 10G-LR make up the main options for data centers. 10G-SR delivers a reach of 300-400 meters over multimode, while 10G-LR is the single-mode solution, with a reach of 10 kilometers. Both use LC connections.

25 Gb/s tranceivers are seeing greater adoption amongst all types of data centers when addressing downlinks to servers. With 25 Gb/s transceivers, there are two IEEE compliant options: 25G-SR and 25G-LR, similar to 10 Gb/s transceivers. An “extended” option is available in 25G-eSR, which extends the reach of short reach multimode links, as listed in Table 1 below. A 25G-ER single-mode transceiver is available for distances up to 40 kilometers, but this distance is targeted for applications outside of the physical data center.

Table 1: 25G Optical Transceivers Q1-2021

	Transceiver	Form Factor	IEEE Compliant	Fiber Type	Distance (meters)	# of Fibers	Connector
1	25G-SR	SFP28	Yes	OM3 / OM4 / OM5	70 / 100 / 100	2	LC
2	25G-eSR	SFP28	No	OM3 / OM4 / OM5	300 / 400 / 400	2	LC
3	25G-LR	SFP28	Yes	OS2	10,000	2	LC
4	25G-ER	SFP28	No	OS2	40,000	2	LC

While 40 Gb/s was initially popular as a switch-to-switch connection speed, it was never popular as a server speed, and it has been losing share to higher speeds that are more economical, such as 100 Gb/s. As we saw above in figure 2, the 100 Gb/s transceiver is on the verge of becoming the most common speed, and there are many options at that data rate. 100G-SR10, 100G-SR4, 100G-LRL4, and 100G-LR4 are the IEEE compliant options — the rest are all defined by multi-service agreements (MSAs), as listed in Table 2.

These MSAs have created an explosion of options. For some data center managers, it is a best practice to follow existing industry standards. But with data center demands changing so rapidly, the latest technology will often reach the market well ahead of new standards where they can be addressed. And many cloud and hyperscale data centers have requirements for distance, cost, or custom solutions that simply don't fit into standard specifications. Many of the MSA-developed transceivers can address these requirements, with a development process that can go from thought to reality within a year. For example, 100G-SR-BD, 100G-FR, and 100G-DR were quick to market in 2021, providing breakout options to 400 Gb/s for applications both in the server and the switch.

Table 2: 100G Optical Transceivers Q1-2021

	Transceiver	Form Factor	IEEE Compliant	Fiber Type	Distance (meters)	# of Fibers	Connector
1	100G-SR10	CFP / CFP2 / CPAK	Yes	OM3 / OM4	100 / 150	20	24F MTP
2	100G-SR10 MXP	Embed. Optics	No	OM3 / OM4	100 / 150	24	24F MTP
3	100G-SR4	QSFP28	Yes	OM3 / OM4 / OM5	70 / 100 / 100	8	12F MTP
4	100G-XSR4	QSFP28	No	OM3 / OM4	150 / 300	8	12F MTP
5	100G-LRL4	QSFP28	Yes	OS2	2,000	2	LC
6	100G-CWDM4	QSFP28	No	OS2	2,000	2	LC
7	100G-LR4	CFP2 / CPAK / QSFP28	Yes	OS2	10,000	2	LC / SC
8	10x10-LR	CPAK	No	OS2	10,000	20	24F MTP
9	100G-PSM4	QSFP28	No	OS2	500	8	12F MTP
10	100G-SWDM4	QSFP28	No	OM3 / OM4	70 / 100	2	LC
11	100G-SR-BD	QSFP28	No	OM3 / OM4 / OM5	70 / 100 / 100	2	LC
12	100G-FR	QSFP28	No	OS2	2,000	2	LC
13	100G-DR	QSFP28	No	OS2	500	2	LC

Of the options above, the 100G-SR4 and 100G-CWDM4 transceiver types are currently the most popular options. The "SR4" suffix denotes short-reach 100 meter (≈328 feet) multimode over 4 optical channels (8 fibers), using MPO connections. This is sometimes referred to as space division multiplexing, or SDM. "CWDM4" denotes course wavelength division multiplexing over four wavelengths, using 2 single-mode fibers and LC connections, with a reach of up to 2 kilometers (≈1.2 miles). There are some other options starting to emerge, such as 100G-DR and 100G-FR. These different transceiver choices will influence the type of fiber and connectivity chosen for a new network, whether for 100 Gb/s or any other data rate.

While there are fewer 400 Gb/s transceiver types than for 100 Gb/s, these options continue to grow. IEEE-defined 400G-DR4 and 400G-SR8 transceivers, using MPO connections, are anticipated to become the most popular 400 Gb/s transceiver types. The most common use of the 400 Gb/s port will be to aggregate multiple 100 Gb/s downlinks. Single-mode 400G-DR4 aggregates four lanes of 100G-DR1, while multimode 400G-SR4.2 will eventually aggregate four lanes of 100G-SR1.2 (the details for multimode compatibility are still being finalized). More MSA-driven transceiver options are in development for 400 Gb/s, and there are groups currently working on 800 Gb/s and higher speeds.

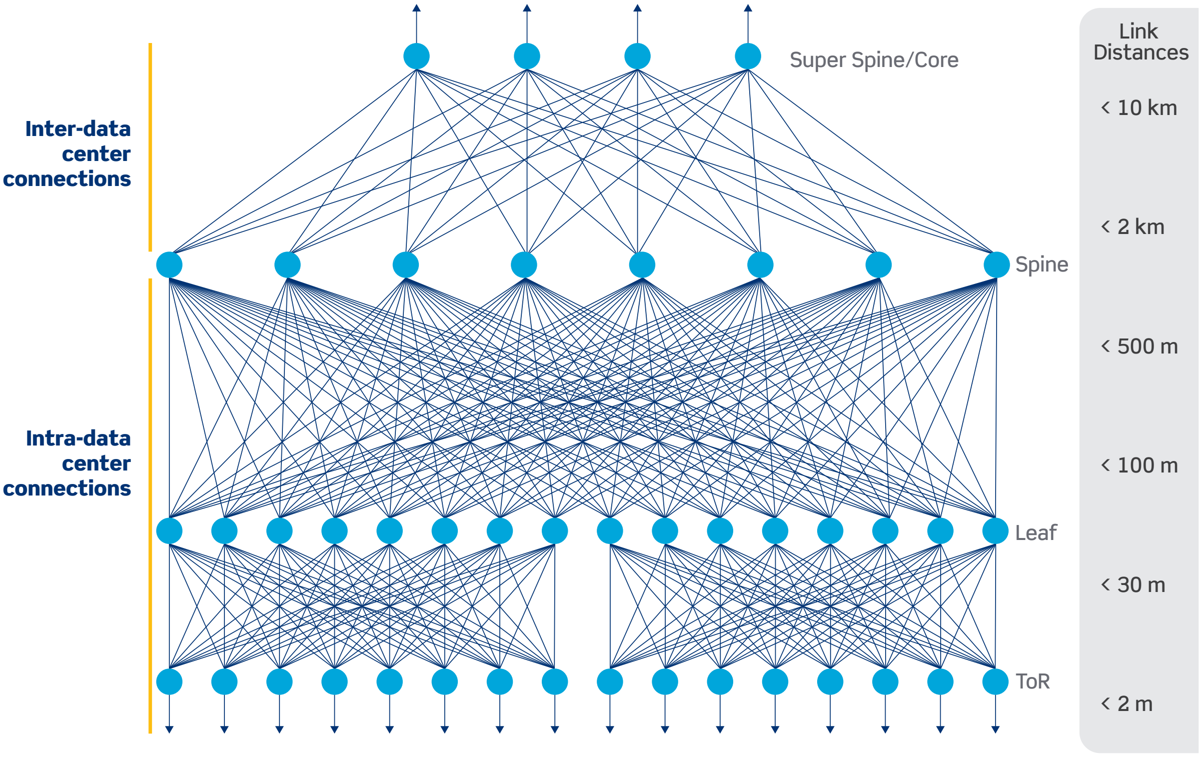
REACH REQUIREMENTS AND FIBER TYPES FOR EACH TIER

Data center managers can't effectively move forward with a cabling design without knowing the types of transceivers that will be used. Part of that transceiver decision process is determining cost — transceivers are the costliest part of the network — but another important factor is the reach required.

In a data center architecture, there will be different layer 1 network tiers, and each will have different reach requirements. For example, the transceiver and fiber cabling choices for a Top-of-Rack (ToR) design — which typically have short 2-3 meter connections — can be very different than the connections for an End-of-Row or Spine design.

The interconnected mesh architecture is a popular design for cloud data centers, and in this design, every spine switch is connected to every ToR switch by running through one leaf switch, as shown in Figure 3. At the bottom of this architecture, the connections are much shorter, and become longer when moving toward the Spine and Core connections at the top. The shorter reach connections from the ToR switches to servers are typically multimode fiber or even copper. The type of connections used in ToR to Leaf switches can vary, but OM4 is by far the most popular choice today. Single-mode OS2 fiber dominates the Leaf to Spine and Super Spine connections, as only single-mode will meet the longer distances required. The Super Spine environment may often include inter-data center connections, where they are located in a separate data center or another building in a data center campus. This is often called the "Data Center Interconnect, or DCI.

Figure 3: Logical Data Center Connections — Mesh Architecture



In some very large data centers, the distances from leaf to spine switches can extend beyond 500 meters. These longer link distances can push up against industry standards parameters for length and performance, depending on the transceivers and fiber type used. Link distances that go beyond the standards defined by IEEE, TIA, and ISO/IEC are commonly referred to as "extended reach" applications. Extended reach can also refer to a higher number of interconnections used in a link. For example, a higher number of channel connections can become common in data centers with multiple data halls inside one building, where interconnection panels are added to provide connectivity.

How is extended reach accomplished?

Standards-based links set limits on factors such as maximum allowable loss in the channel, individual components, and connector matings. These standards assume the worst-case scenarios when it comes to the performance of components and link lengths. Extended reach can be accomplished through "engineered links," or links specifically designed to allow additional mating or components that are evaluated and configured to meet a target application. With today's lower loss factory terminated products, the impact of each mating in a channel is greatly reduced. While loss is added with each additional mating, factory terminated components have a relatively small increase in total attenuation as additional components are added to a channel.

In some greenfield data center network designs, the initial layer 1 installation may not require any engineered links that perform beyond industry-standard limits for distance. However, these limits may become a concern with tech upgrades in the future. Reach typically decreases as data rates increase, largely due to bandwidth limitations in the fiber and transceiver-specific factors by application. For example, the TIA-defined supportable distance for 10GBASE-S using OM4 is 400 meters, but the supported distance drops to 100 meters when upgrading to 100GBASE-SR4. A good data center design will anticipate migration to those higher speeds. If the need is for 10 Gb/s server connections today, it is wise to consider cabling and connectivity that can support at least 25 Gb/s in the future.

When considering extended reach applications, it is prudent to have a full understanding of current and future data applications, target application transceiver specifications, and a careful analysis of the performance capabilities of the passive cabling infrastructure. The fiber cabling system manufacturer should be able to model the anticipated network designs and calculate maximum supported distances.

For example, Leviton has the capability to calculate extended reach with its Optical Link Verification Tool. Calculations completed in this tool are generated based on the IEEE engineered link models that were originally developed to assist Ethernet and Fibre Channel committees with specifications and evaluating the impact of various link penalties.

The Optical Link Verification Tool can reinforce design work by adding a high level of confidence in the channel performance and application support. The tool calculations provide a snapshot of how any given topology using Leviton end-to-end fiber solutions will outperform the standard.

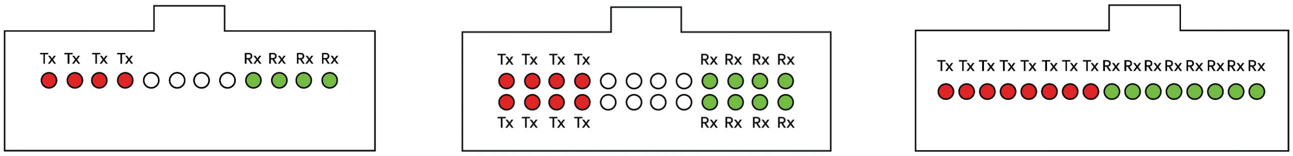
FIBER CONNECTION CHOICES

Once data center managers choose fiber types for the new network, they can then consider connection types. Today, the large majority of data center connectors are either LC or MPO connectors.

LC duplex connectors are currently the most popular fiber connection type. The duplex connector is easy to manage from a polarity perspective, and its established popularity makes it readily available. A common question we hear is "Will LC connections work when upgrading to data rates beyond 25 Gb/s?" While there are LC solutions for connecting networks at 40 Gb/s and 100 Gb/s, the duplex options that use LC connections will typically require multiplexing technologies like CWDM, which can raise the price of transceivers.

Moving beyond 10 Gb/s, parallel optics with MPO connections have become a popular choice. MPOs allow for breakout options that create easy connections between a higher speed ports and multiple lower speed ports. For example, in the Base-8 MPO connector shown in Figure 4, one MPO connection delivering 100 Gb/s can breakout out to four 25 Gb/s ports at the server rack. MPO assemblies also consolidate trunk cable connections into a denser package, freeing up more pathway space and creating backbone connections that will require fewer potential changes throughout future tech refreshes. This is important, since long trunk runs from the-spine-to-leaf tiers are not easily accessible after they are installed and can be time and labor intensive to rip and replace.

Figure 4: MPO connector options



Base-8/12 | **12-fiber MPO**

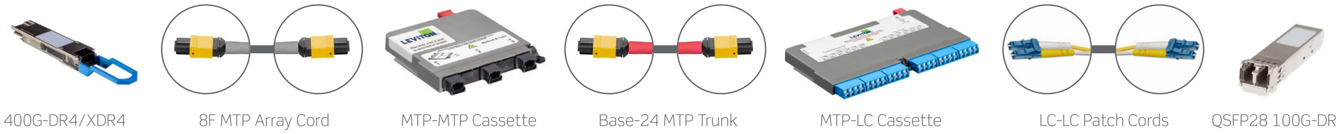
Base-24 | **24-fiber MPO**

Base-16 | **16-fiber MPO**

There are trunk cable options that can be terminated with any of the connector types shown above, regardless of the connectivity used at the transceiver. For example, a Base-24 MPO trunk cable is a flexible option as a backbone connection in the channel. While the connector used for the trunk cable does not need to match the transceiver format, it should contain at least as many fibers as the interface needs. Cassettes can be used at the end of the trunk to convert to the corresponding transceiver connectivity requirements. One Base-24 trunk can contain 2-, 8-, 16-, 20- or 24-fiber strand transceiver connections. That means it can be used to support all the transceiver options for 25 Gb/s and 100 Gb/s previously listed in Tables 1 and 2.

Figure 5 shows how a Base-24 MPO trunk can be used in a channel with both parallel transceivers at the uplink and duplex transceivers at the downlink. This type of configuration could be used to breakout 100 Gb/s to 25 Gb/s, or 400 Gb/s to 100 Gb/s in the future — without making any changes to the cabling system. While Figure 5 shows one channel breakout, three times the number links can be supported with the same cassettes and Base-24 trunk when fully populated.

Figure 5: 8-Fiber Channel Breakout with a Base-24 MPO Trunk



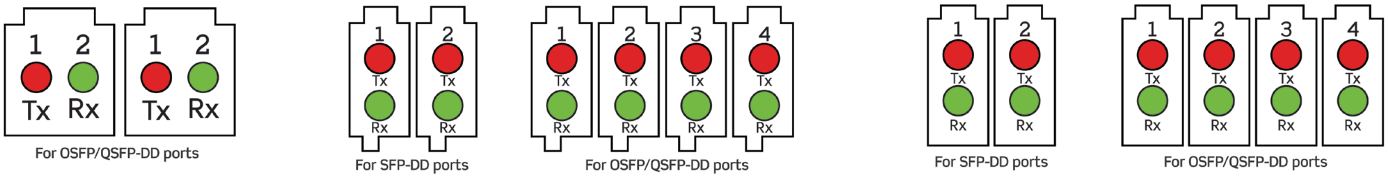
As MPOs allow for more lane assignments in one connection, managing polarity can become more critical. While there are different methods for maintaining proper polarity, network managers should stick to only one method to avoid future confusion.

Emerging connection options

Some new connector types have been introduced in recent years for addressing data rates of 400 Gb/s and higher, known as very small form factor (VSFF) connectors. These duplex connectors are used in OSFP, QSFP-DD and SFP-DD transceivers to increase port density at switch ports over traditional LC duplex connectors. Based on the multi-vendor MSA groups, three VSFF variants have been incorporated. They are the CS®, SN®, and MDC connectors.

The CS connectors allow for four fibers in almost the same footprint as a 2-fiber LC connector. The SN and MDC connectors, smaller than the CS connector, have vertically oriented transmit and receive fibers. They allow for two duplex connections in an SFP-DD interface and four duplex connections in a QSFP-DD interface as shown in Figure 6. While similar, the SN and MDC have some key differences and are not interchangeable. These new VSFF connectors are relatively new to the market, and it is currently unclear which will gain the largest market acceptance in the structured cabling space.

Figure 6: VSFF Connector Types



CS® Connector **SN® Connectors** **MDC Connectors**

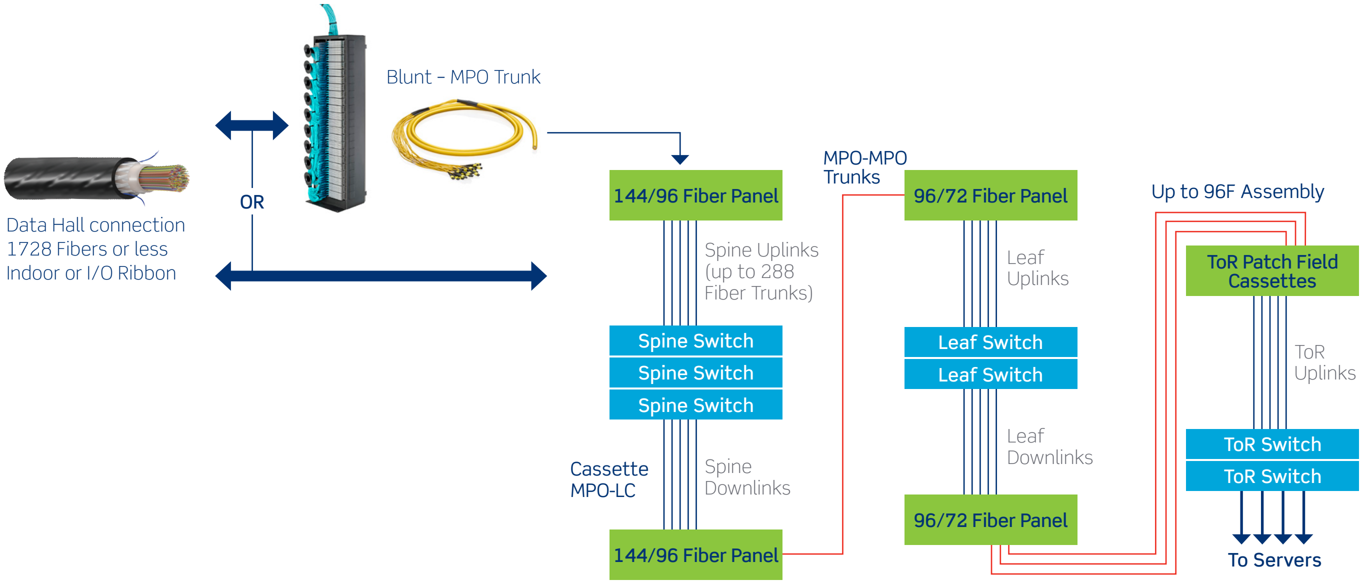
PHYSICAL CABLING TOPOLOGY

Once decisions have been made for transceiver type, fiber type and connector type, the next step is to envision the physical makeup of the channels. Following the mesh architecture introduced earlier, Figure 7 shows an end-to-end cabling topology, with the passive infrastructure and density for each tier of the design. Moving from one tier to the next, patch panels are needed to make the mesh connections. The number of tiers required in this mesh architecture can be altered depending on the number of servers to be connected and the switch radix.

In the entrance and spine MDA tiers shown on the left, the density will generally be much greater than the tiers on the right. They may contain the same number of total fibers as the tiers on the right, but they will be consolidated into a smaller number of cabinets or square footage in the MDA or entrance. Network designers need to consider how they will interface with the fiber connections at each location. Issues such as clear labelling and individual connector access can have a significant impact on the time spent at the patch field.

Figure 7: End-to-End Connections for Mesh Architecture

ENTRANCE	SPINE MDA	LEAF IDA	TOR HDA
Ribbon Cable (up to 1728 Fibers) Splice Enclosure Splice on MPO	196-288 Fiber Trunks 144/96 Fibers per RU Mesh Achieved	48-144 Fiber Trunks 96/72 RU Panels Cassettes or Adapters Mesh Achieved	12-96 Fiber Trunks 96/72 Fibers per RU Cassettes Patch Cords



TYPICAL LINK COSTS

The key to deciding on the best physical layer design will come down to choosing solutions that address the performance demands required while economical enough to meet budget constraints. When finding this balance, there are practical reasons for implementing different technologies at different tiers based on cost.

For example, copper or multimode fiber will often still be the technology of choice in data center rows and servers. Copper is the least expensive solution to implement but supports shorter link distances and lower maximum data rates — currently up to 10 Gb/s. Multimode can handle higher speeds and support longer distances than copper but will come at a higher cost. Single-mode optics will be the technology of choice for the spine tier of the network. In some cases, it still comes at a higher cost than multimode, but it is typically the only solution capable of delivering higher data rates such at 100 Gb/s (and higher) at the longer link distances required.

When considering costs, network designers should always consider the cost of the entire link. For example, transceivers for parallel optics will typically cost less than duplex transceivers, but the associated parallel MPO cabling comes at a higher cost, considering it is an 8-fiber solution versus a 2-fiber solution. For longer reach distances, the impact of the cable cost becomes more significant. However, in general, link costs are largely shaped by the transceivers, as they are the most expensive components, as shown in Table 3. This is a key reason why transceivers are the most important decision to make first when designing the network.

When evaluating the passive infrastructure, most of the costs are in the cabling and connectivity. Components such as splice or patch enclosures - categorized as “hardware” in Table 3 - do not add as much to the material cost of the overall structured cabling system.

Table 3: Share of link cost by component

Item	Share of Cost
Transceiver	60-75%
Cabling	20-30%
Hardware	5-10%

While the choices for patching hardware have little effect on capital expenditures, they can significantly influence operational costs, as this is the area of the network where technicians will access the cabling after installation. Questions about density, accessibility, manageability, and physical protection should be considered, as they have an impact on the efficiency of MAC work in the future. As with fiber type and connection type, the ideal choices for enclosures, panels, or fiber cassettes may be different based on the area or tier of the data center. However, the deciding factors will have more to do with functional efficiency than product cost.

Beyond the horizon

As new data centers are designed to scale out in the future to add more servers, network designers should plan for a fiber infrastructure that anticipates higher data rates and longer distances down the line. The greenfield data center will become a brownfield when it reaches its first tech refresh. For many data centers, that refresh cycle can occur every three to five years, so planning now will ensure the most cost-effective design with the least disruption. That means if you want your installed fiber cabling system to last 20 years, it would need to support at least three network migrations. This involves anticipating not only higher speeds with each refresh, but also anticipating potential distance reductions as a result. As mentioned earlier, an installed OM4 system can support 10GBASE-S over 400 meters, but that supported distance drops to 150 meters when moving to 40GBASE-SR4, and 100 meters when upgrading to 100GBASE-SR4. A good data center design will anticipate migration to those higher speeds and ways to address distance requirements.

The Leviton Opt-X® product family is a single, simple connectivity solution that makes it easy to migrate from 10 to 40, 100, 200 and 400 Gb/s networks. The system employs a MTP backbone that can stay in place over multiple tech refreshes, saving time, reducing labor, and minimizing network downtime. The system’s ultra-low-loss connectivity helps manage power budgets while allowing multiple connection points and extended distances. When using the Opt-X Unity System, performance assurance and extended distance calculations can be provided through extensive testing and analysis with the Leviton Optical Link Verification Tool.

And remember, Leviton can provide assistance when you are creating a forward-looking network migration plan. We understand the active equipment trends and next-generation standards in development to ensure that a network design built from the ground up today will be ready for change in the years to come.

Learn more about Leviton data center solutions at Leviton.com/datacenter.

LEVITON®
Network Solutions

Today's networks must be fast and reliable, with the flexibility to handle ever-increasing data demands. Leviton can help expand your network possibilities and prepare you for the future. Our end-to-end cabling systems feature robust construction that reduces downtime, and performance that exceeds standards. We offer quick-ship make-to-order solutions from our US and UK factories. We even invent new products for customers when the product they need is not available. All of this adds up to the **highest return** on **infrastructure investment**.

USA — NETWORK SOLUTIONS HEADQUARTERS

2222 - 222nd Street S.E., Bothell, WA, 98021, USA
+1 (800) 722 2082   infousa@leviton.com   leviton.com/ns

Customer Service
+1 (800) 722 2082
insidesales@leviton.com

International Customer Service
+1 (425) 486 2222
intl@leviton.com

Leviton Berk-Tek Cable Customer Service
+1 (800) 237 5835
berktek.info@leviton.com

Technical Support
+1 (800) 722 2082
+1 (425) 486 2222
appeng@leviton.com

APAC

+85 (2) 3620 2602   infoapac@leviton.com   leviton.com/ns

Customer Service
+1 (631) 812 6228
infoasean@leviton.com

China
+85 (2) 2774 9876
infochina@leviton.com

South Korea
+82 (2) 3273 9963
infokorea@leviton.com

CANADA

+1 (800) 461 2002   infocanada@leviton.com   leviton.com/ns

Customer Service
+1 (514) 954 1840
pcservice@leviton.com

Network Solutions products are **available worldwide** in **over 100 countries**. Visit us online at leviton.com/ns to learn more.

© Copyright Leviton Manufacturing Co., Inc.

EUROPE

Viewfield Industrial Estate, Glenrothes, KY6 2RS, UK
+44 (0) 1592 772124   infoeurope@leviton.com   leviton.com/ns/emea

Customer Service
+44 (0) 1592 772124
customerserviceeu@leviton.com

Technical Support
+44 (0) 1592 778494
appeng.eu@leviton.com

Benelux
+44 (0) 1592 772124
infobenelux@leviton.com

Nordics
+46 (70) 9675033
infonordics@leviton.com

Central & Eastern Europe (CEE)
+44 (0) 1592 772124
infocee@leviton.com

Portugal
+351 (21) 421 4133
infoportugal@leviton.com

DACH
+49 (0) 173 272 0128
infodach@leviton.com

Spain
+34 (91) 490 59 19
infospain@leviton.com

France
+33 (0) 1709 87826
infofrance@leviton.com

UK & Ireland
+44 (0) 1592 772124
infouk@leviton.com

Italy
+39 (02) 3534896 (Milan)
+39 (06) 68584613 (Rome)
infoitaly@leviton.com

LATAM

infolatam@leviton.com   leviton.com/ns

Customer Service
+52 (55) 2333 5963
infolatam@leviton.com

Caribbean
+1 (954) 593 1896
infocaribbean@leviton.com

Mexico
+52 (55) 2128 6286
lsamarketing@leviton.com

Colombia
+57 (1) 743 6045
infocolombia@leviton.com

MIDDLE EAST & AFRICA

Bay Square, Building 3, Office 205, Business Bay, Dubai, UAE
+971 (4) 247 9800   infomea@leviton.com   leviton.com/ns

Customer Service
+971 (4) 247 9800
lmeinfo@leviton.com