Design Guide

Fabric Networks

Designing your network for the future—from 10G through 400G and beyond

Contents

Executive summary

The widespread adoption of virtualization and cloud computing has led to the need for new data center switching architectures that provide lower latency and higher throughput. These new architectures are based on fabric network switches, and are different from the traditional three-layer switching topologies.

These fabric switches can take many forms—from fabric extensions in a top-of-rack deployment, to centralized fabric at the HDA or IDA, to a full mesh architecture. As such, consideration must be given as to how the physical layer infrastructure is designed and implemented to ensure the switch fabric can scale easily and efficiently.

This white paper provides an overview of fabric technology, along with design considerations and a practical look at implementing fiber connectivity that can accommodate changes in architecture as well as higher line rates as the network grows.

Practical examples of fabric network design with the SYSTIMAX® InstaPATCH® 360 preterminated fiber solution are also given to highlight the importance of designing infrastructure that supports higher speeds and network growth.

New data center architectures

Data center designs and architectures have evolved to accommodate the growth of cloud-based storage and compute services. Traditional private enterprise data centers are adapting their current architectures to prepare for new, agile, cloud-based designs. These new enterprise architectures resemble "warehouse-scale" facilities but are designed to support many varied enterprise applications.

To prepare for cloud architectures, an optimized direct path for server-to-server communication is achieved with "leaf-spine" architecture (see Figure 1). This design allows applications on any compute and storage device to work together in a predictable, scalable way regardless of their physical location within the data center.

"Cloud networks" are based on an architecture consisting of meshed connections between leaf and spine switches. The mesh of network links is often referred to as a "network fabric." The performance of the fabric is well suited to establishing universal "cloud services": enabling any-to-any connectivity with predictable capacity and lower latency. The fabric has inherent redundancy, as multiple switching resources are spread across the data center, thus helping assure better application availability. These distributed network designs can be much more cost-effective to deploy and scale when compared to very large, traditional centralized switching platforms.

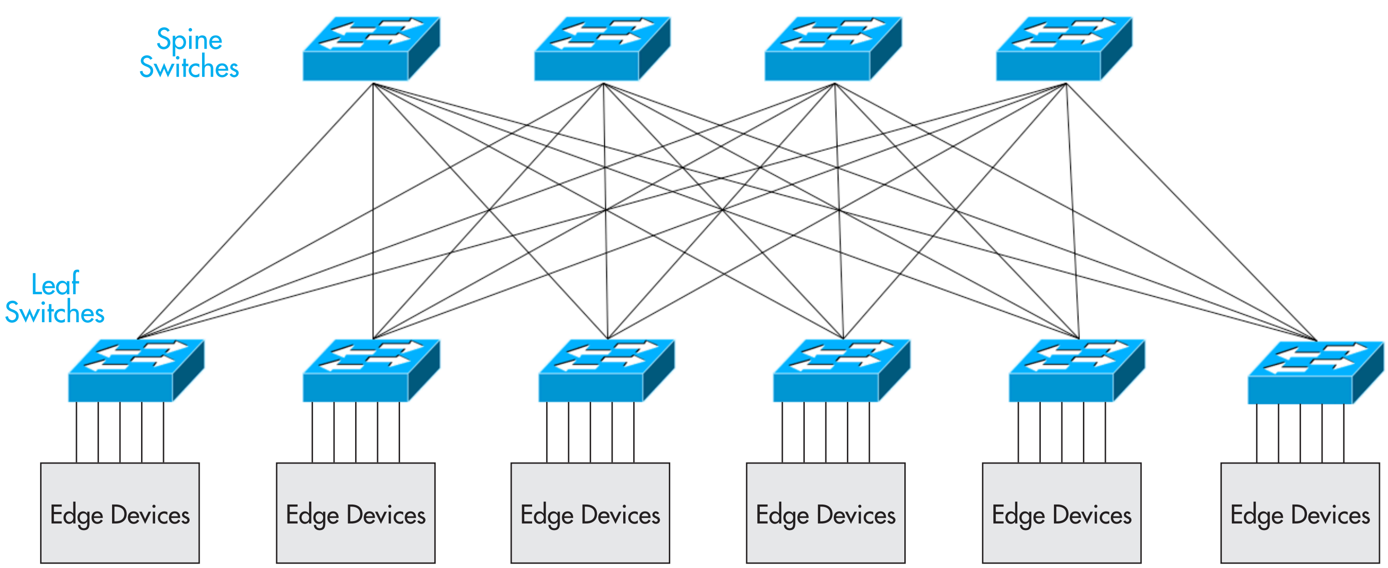

Figure 1. A "mesh" network fabric architecture with connections between each leaf and spine switch, providing any-to-any connectivity

Fabric topology design—capacity

The traditional way to design networks is to total up traffic through aggregation (e.g., if a network needs to support 10 data streams of 1 Gb each, adding the streams would indicate that 10 Gb network links would be needed). Leaf-and-spine networks work differently, however. To scale fabric networks, designers need to consider the following factors:

- The speed (or bandwidth) of the fabric links
- The number of compute/storage device ports (also known as edge ports)
- The total fabric bandwidth needed to service all data center applications

The speed of the fabric is not the total carrying capacity between every pair of leaves in the fabric; it is the total bandwidth between each leaf and all of the spine switches.

In the example above, there are four spine switches. If each leaf switch has one 40 Gb link to each spine, the result is a 160 Gb fabric. Note that each leaf must have the same link speed to every spine switch. Also note that there are no device connections directly to the spine switches.

The speed of the fabric needs to be sized to support the largest amount of traffic any single leaf switch could send. For example, if there are 48 10 Gb ports connected to high-speed servers, the fabric would need to support 48 X 10 Gb—or 480 Gb of bandwidth.

The total number of edge ports is the next important consideration. This is a function of the number of leaf switches in the fabric. For example, if a leaf switch provides 24 10 Gb ports, each additional leaf adds another 24 ports to the total fabric. A new leaf can be added only if each spine switch has an additional port available for the new leaf switch.

When links between leaf-spine switches (at 40 Gb, for example) have more capacity than edge-port links (at 10G, for example), the design is referred to as a fat tree. If the links are kept at the same speed (e.g., 10G edge: 4 X 10G leaf-spine) the design would be referred to as a skinny tree. Fat tree designs have obvious benefits for scaling up the data center fabric. Each leaf-and-spine switch must have enough ports to enable the any-to-any mesh connections. The number of ports and the capacity of each port predetermine the maximum size and bandwidth up to which the fabric can scale.

A typical fat tree might use 40 Gb fabric links. In our example above, we have four spine switches, with each spine supporting six 40 Gb ports—for a total bandwidth of 240 Gb. Assuming each leaf switch has 48 10G ports, this yields a total of 288 10 Gb edge ports. However, most devices will be dual attached, requiring two 10 Gb ports per edge device. This configuration will support 144 fully-redundant edge devices.

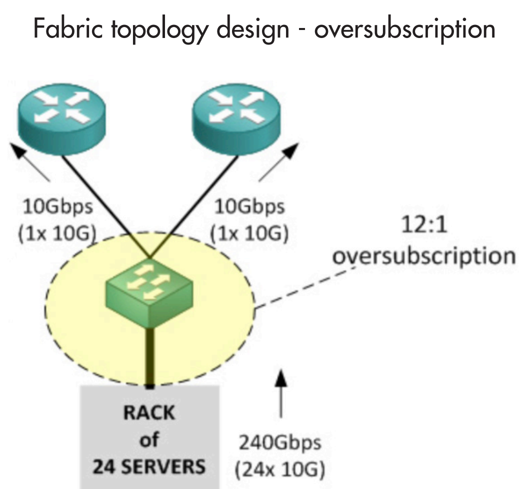

Figure 2. Leaf-spine link speeds and oversubscription

The total fabric bandwidth can be calculated by multiplying the number of edge ports by the speed of the edge ports—or the number of spine ports by the speed of the spine ports. If there is no oversubscription, these two numbers will be the same. Network fabrics are intended to be nonblocking, where all the traffic produced by edge devices can run over the fabric without delay or "blocking." Oversubscription means some traffic can be blocked or delayed because resources are fully utilized by others. Blocking can severely impact data center applications—especially applications such as FCoE, which depend on a nonblocking environment. Many fabric network architectures maintain separate storage networks—some with Fibre Channel, some with IP-based storage, and some with distributed software-defined storage.

Designers consider how applications communicate and calculate the overall capacity requirements, which equates to the fabric size of the network. Some network designs include compromise that suits the budget and service quality appropriate for the services to be delivered, which means an acceptable level of blocking or contention for network resources is designed into the overall network architecture. The oversubscription ratio describes the level of resource contention that exists for edge devices. An example is shown in Figure 2, with an oversubscription ratio of 12:1.

If the oversubscription ratio is too high, performance of the applications suffers. If the oversubscription ratio is kept very low, the number of servers—and therefore the number of applications that can be supported on the fabric—is reduced. This balance between capital cost and application capacity is a critical design factor. It is also a factor that is very likely to change quickly over time as the demand for applications increases. Server hardware capacity tends to increase, which means fabric link capacity will be stressed.

It is clear from the discussion above that higher leaf-spine link capacity can improve the service level by minimizing the oversubscription ratio and increasing the number of servers that can be supported by the network fabric. The capacity of these links would ideally be as high as practically possible.

As the fabric expands, connections must be made to every other peer device. The number of connections grows quickly as leaf switches are added. Physical layer connectivity must adapt to support these network fabrics with higher density, greater network link speed, and multifiber modularity—which, in turn, helps speed deployment and network availability. Pictured below are the MPO equipment cords that might be used to provide the physical connections for QSFP (4 X 10G)—40G Ethernet network link

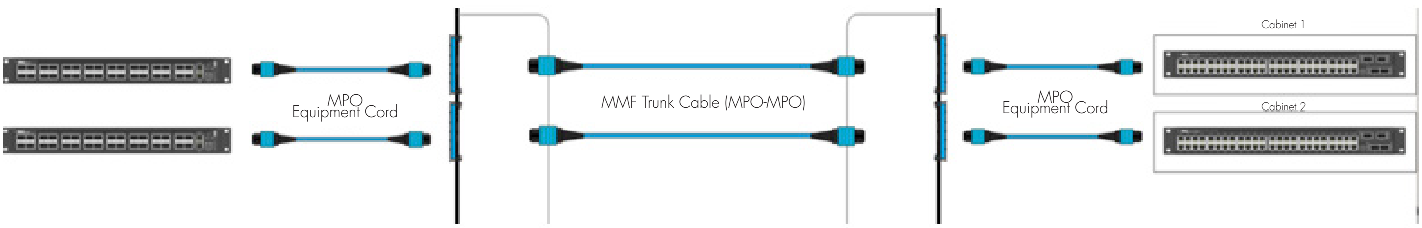

Figure 3. 40G switch connectivity with 40GBASE-SR4 (note: for simplicity, pins not shown)

To optimize the fabric capacity, the optical components must provide high bandwidth and low loss—stepping up to the next network speeds. 40G, 100G or even 400G should be part of the day-one design requirements to prevent the need for a redesign of the cabling infrastructure.

The optical network technology supporting these links is progressing rapidly. Speeds are increasing very quickly—in some cases, the solutions being offered in this area are well ahead of the industry standards. The cost benefit of these various options is key to keeping pace with the overall data center capacity requirements and is a key element in the balance between CapEx and availability risk.

Network fabric—the physical network

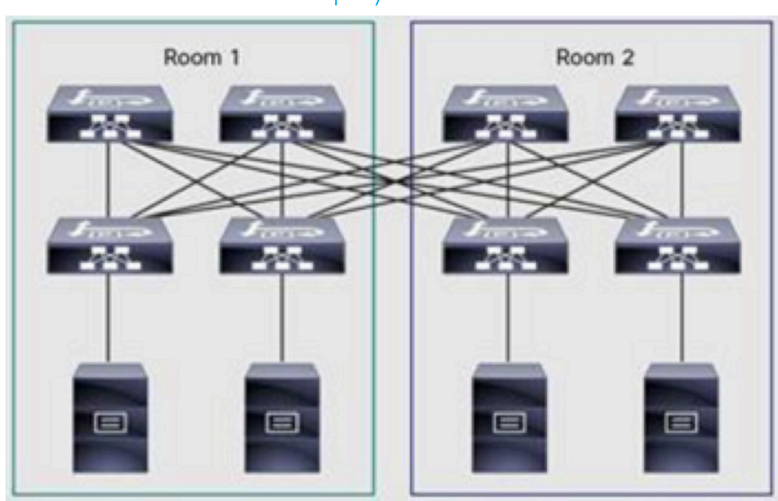

Figure 4. Leaf-spine design spanning multiple halls

Fabric network implementation is similar to the traditional three-tier networks deployed in the past in several ways—they must be scalable, manageable and reliable. The use of structured cabling designs remains equally valid and valuable when implementing fabric network topologies. The cabling pathways and spaces remain the same. Fabrics may span multiple halls in a data center. Fabric elements also require out-of-band management network support. These physical design requirements are incorporated into a floor layout. A typical layout is shown below.

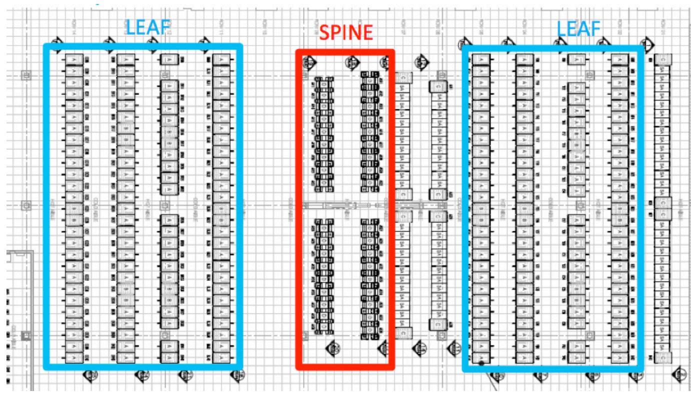

Figure 5. Plan view of leaf-spine cabinet layouts

Figure 6 illustrates a typical data center topology[1] with a cross-connect at the IDA: the spine switches are shown at the main distribution area (MDA) and the leaf switches shown at the horizontal distribution area (HDA).

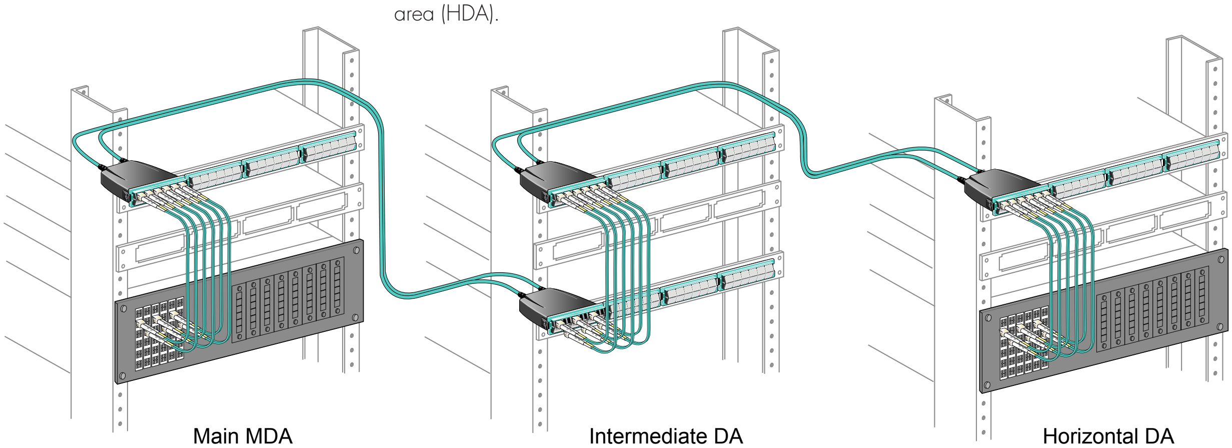

Figure 6. Leaf-spine with InstaPATCH 360 cross-connect at IDA

Designing the fabric network with a cross-connect greatly increases flexibility and agility and simplifies what can be a very complex channel. The use of cross-connects in data centers has already become mandatory per the CENELEC EN 50600-X standard in Europe. Based on the ability to provide any-to-any port connectivity, the cross-connect architecture is recommended.

[1]For more information on network topologies see BICSI Data Center Standard.

Data center pathways and spaces supporting fabric network architectures

Designing the physical network links will depend largely upon the overall network topology and upon the networking vendor. Some switch vendors provide proprietary optics with a preference for either singlemode or multimode media. Still others favor larger chassis-based switches and zone-based horizontal area cabling. The future view of the next-generation networks will also vary. In many cases, the cabling design team is the last to know which particular network hardware must be supported on day one. The ideal tool set will support the variety of options that may come their way and will make it easy to evaluate future network options—supporting new initiatives while avoiding vendor lock-in.

The InstaPATCH 360 preterminated fiber cabling system shown in figure 6 is perfectly suited to provide a high-performance factory-manufactured structured cabling platform that easily handles the broad set of fiber applications needed to support fabric networks. InstaPATCH 360 trunk cables, distribution modules and patch cords are configured to suit trunk, switch and compute equipment requirements for day one, as well as to provide a migration path to support day two requirements.

In the fabric network example, we discussed how leaf-spine connectivity might look like those illustrated below—a parallel multimode fiber trunk. This design uses lower cost multimode optics, maintains backwards compatibility with previous network technologies and can provide an upgrade path to 100G capacity in the future provided the link design is valid.

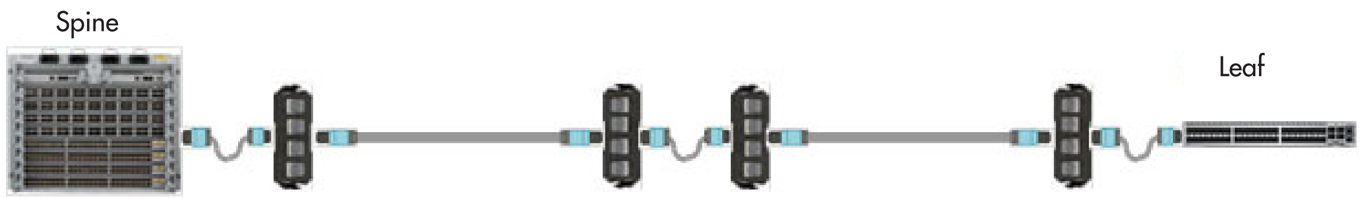

Figure 7. Leaf-spine with parallel fiber trunk and intermediate DA

Fabric topology design—switch port density

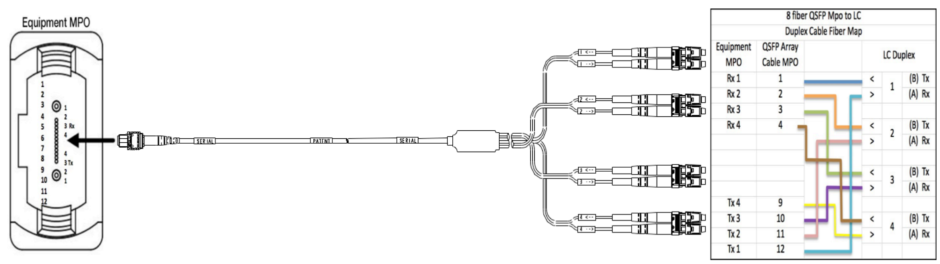

Figure 8. QSFP to LC assembly

Multiple 10G ports can be collected together to support higher capacity links. The IEEE standards provide for a group of four 10G ports that are taken together, combined onto a single 12-fiber MPO connector to form one 40G link. This QSFP standard is used to build higher capacity links (40G) but is also often used to connect a single port on a leaf switch to four servers—increasing the trunk density and the panel capacity of the leaf switch. Combining four LC ports into one QSFP yields roughly 4:1 increased panel density on a leaf switch when compared with using separate serial ports designed for SFP+ interfaces.

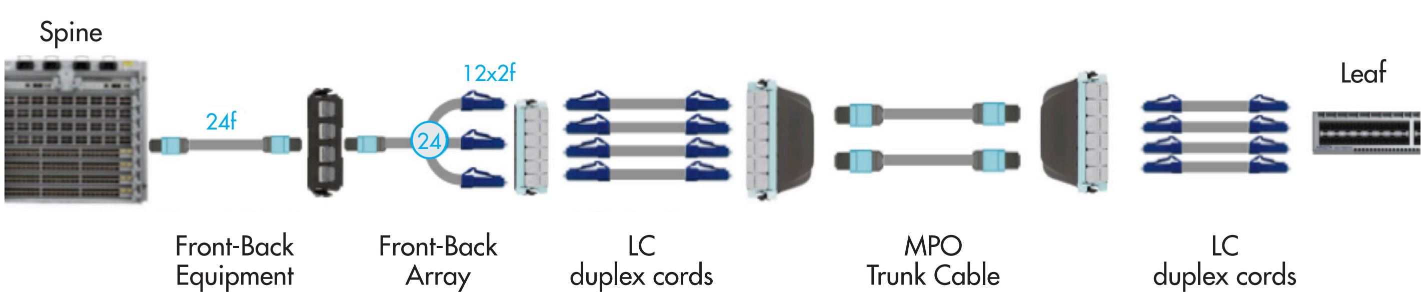

Figure 9. Server-to-leaf connectivity with InstaPATCH 360

Equipment vendors make use of MPO connectors to provide higher fiber density. For example, some vendors provide 12 10 Gb ports on one 24-fiber MPO connector. These can be grouped into three (at 40G each) or perhaps be broken down to 12 10G ports for device connections. Grouping fibers into higher density connection systems saves panel space on network devices and provides convenient management via parallel fiber trunk arrangements.

Fabric topology design—reach

Data centers are often large in terms of physical space and the number of compute and storage devices they contain. There are several data center standards detailing the best practice for network cabling and space design. Examples include ANSI/TIA-942-B, ISO/IEC 11801-5, and CENELEC EN50173-5 and EN50600-X.

Structured cabling designs enable scale and improve mean-time to repair (MTTR) and overall availability. It is therefore highly desirable to maintain this cabling structure with fabric network topologies. Leaf-spine links must also use proper communication pathways in designated communication spaces—just as it was with previous network topologies.

Providing high-capacity links at a reasonable cost is a key design element in fabric networks. Multimode optic devices are typically less expensive than the equivalent singlemode optic devices—especially as network speeds increase. There are currently a wide variety of choices available to the network designer: both standards-based and proprietary solutions, which offer different combinations of capacity, cost and operational risk/benefits. New data link interfaces are emerging that will offer still more choices for link designs. The cabling technology must enable near-term network capacity and make way for fabric designs with increased size and capacity.

Engineering the network links is, therefore, an important consideration in the design of fabric networks. Patching zones can be useful in each of the distribution areas, as shown in the example configuration below. Support for networking applications varies but, generally, the higher the speed, the shorter the distance that can be supported by the structured cabling link. Increasing the number of patches also reduces the signal on the link and, therefore, reduces the usable link distance. Quite often application specifications are provided by the network hardware manufacturers in terms of the maximum point-to-point distance supported. It is important to understand the relationship these specifications will have as we look at actual structured cabling designs.

Suppose you were planning the implementation of new data center services for the topologies shown in Figure 10. The applications that must be supported on day one include 10 Gb Ethernet and 8G Fibre Channel (FC). The data center has been organized into manageable data halls within the data center. Will there be issues supporting the link lengths required by your design?

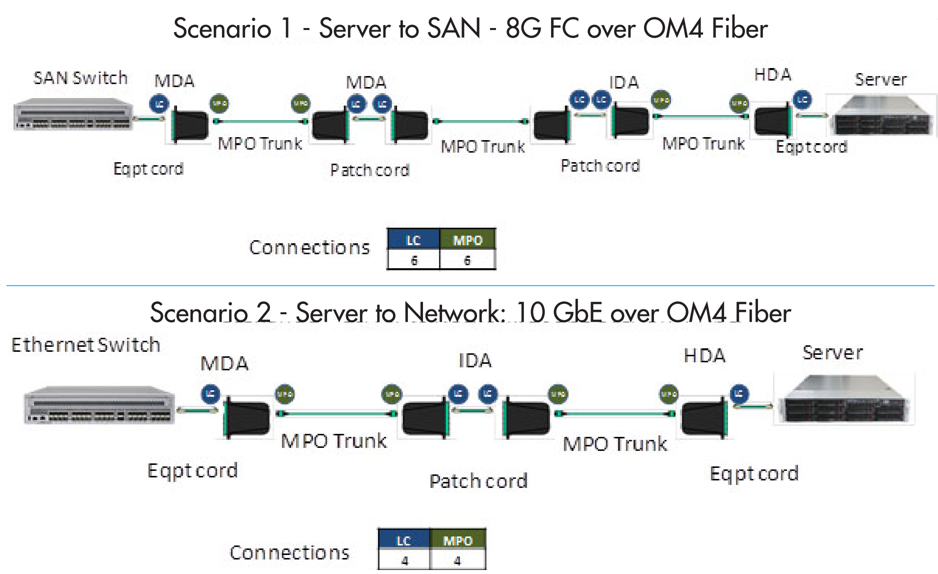

Figure 10. Example data center topologies

Part of the design intent requires that, to ensure future capacity upgrades, the network infrastructure must be able to support at least the next highest data link speeds. The vendor community offers several choices: some singlemode, some multimode, some standards based and still others are new proprietary solutions. Which of the potential future network applications will your day-one design support going forward?

The process to identify the best solution starts with understanding the design options under consideration. Will the topology proposed work reliably with the network gear being considered? If there are options available, which strategy seems to offer the best commercial cost and highest reliability? To answer these questions, we first look to the industry standards that detail the choices for our Ethernet data links. This table includes completed standards as well as those in the task force stage. There are additional applications, including 50G and 200G, which are currently in the working group phase in IEEE 802.3.

Application	Standard	IEEE Reference	Media	Speed	Target distance
10-Gigabit Ethernet	10GBASE-SR	802.3ae	MMF	10 Gb/s	33 m (OM1) to 550 m(OM4)
	10GBASE-LR		SMF		10 km
	10GBASE-LX4		MMF		300 m
	10GBASE-ER		SMF		40 km
	10GBASE-LRM	802.3aq	MMF		220 m (OM1/OM2) to 300 m (OM3)
25-Gigabit Ethernet	25GBASE-SR	P802.3by	MMF	25 Gb/s	70 m (OM3) 100 m (OM4)
40-Gigabit Ethernet	40GBASE-SR4	802.3bm	MMF	40 Gb/s	100 m (OM3) 150 m (OM4)
	40GBASE-LR4		SMF		10 km
	40GBASE-FR		SMF		2 km
	40GBASE-ER4		SMF		40 km
100-Gigabit Ethernet	100GBASE-SR10		MMF	100 Gb/s	100 m (OM3) 150 m (OM4)
	100GBASE-LR4		SMF		10 km
	100GBASE-SR4		SMF		70 m (OM3) 100 m (OM4)
	100GBASE-ER4		SMF		40 km
400-Gigabit Ethernet	400GBASE-SR16	P802.3bs	MMF	400 Gb/s	70 m (OM3) 100 m (OM4)
	400GBASE-DR4		SMF		500 m
	400GBASE-FR8		SMF		2 km
	400GBASE-LR8		SMF		10 km

Figure 11. Ethernet fiber applications standards (standards in progress are shown in red)

Industry standards provide design rules we can use to determine if our data center topology will reliably support the application design requirements. Looking back at the data center design requirements, we can assess each link topology to determine maximum link lengths and maximum signal losses. What is the total loss from all of the connectivity in the link? How does the length and loss combination compare with the limits set by that application standard? Comparing each case with the standards will yield a go/no-go decision for our design.

Determining the losses for the system links requires an understanding of the components deployed. These characteristics vary from vendor to vendor and even within any given production lot. We are obviously interested in the worst-case insertion loss values so as to ensure we do not exceed the tolerances allowed by the networking gear. High-bandwidth fiber media can support much longer links, whereas lower quality fiber will require shorter lengths to function reliably. Basing your design on standards and the vendor-supplied component performance data leaves all of these link calculations to you, the cabling system designer.

The day-two design requirements will require at least the next highest network speed must also be supported based on the initial design topology. There are a number of combinations to consider.

We look for maximum (not average or typical) loss that any cabling element will contribute to the link we are designing. The bandwidth of the fiber media must be considered—OM3 having less bandwidth than OM4, for example. We can consider the possibility of parallel multifiber links in the future. Finally, we can consider the impact of the scale and size of the data center—how does the length of the links required limit the choices we have for next-generation network speeds?

Performing the analysis on the two scenarios outlined above with standard components requires the worst-case insertion loss values for all components in the channel. In this example, the LC/MPO modules have 0.50 dB insertion loss, and the fiber trunk cables are rated at 3.5 dB/km insertion loss. The duplex fiber patch cords are assumed to be a few meters in length; as such, they will not contribute materially to the overall insertion loss.

Based on these values, the total insertion loss is 3.34 dB, which exceeds the maximum 2.19 dB insertion loss for 8G Fibre Channel. As designed, this link would likely fail or experience excessive bit errors.

Scenario 1 - Server to SAN - 8G FC over OM4 Fiber

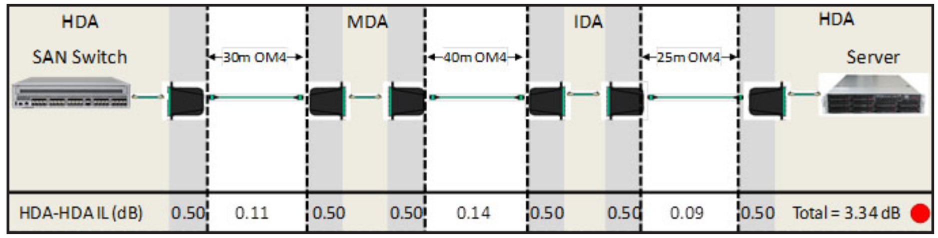

Figure 12. Loss budget calculations for server-to-SAN link over OM4 fiber

A similar analysis is performed for the second scenario, which outlines a 10G Ethernet server-to-network link operating over 130 meters of OM4 fiber. In this scenario, the total loss budget is 2.39 dB, which is below the loss limit for this application over OM4 fiber. Based on this analysis, the link should operate properly.

A common occurrence in data center operations is the insertion of additional connections as the network grows and new data halls come online. In this example, an additional connection at the IDA has been added, increasing the total channel length to 150 meters and adding two more LC/MPO modules. As shown below, the new total insertion loss is now 3.53 dB, which exceeds the maximum allowable value. This link, as designed, would either fail or experience excessive bit errors.

Scenario 2 - Server to Network: 10 GbE over OM4 Fiber

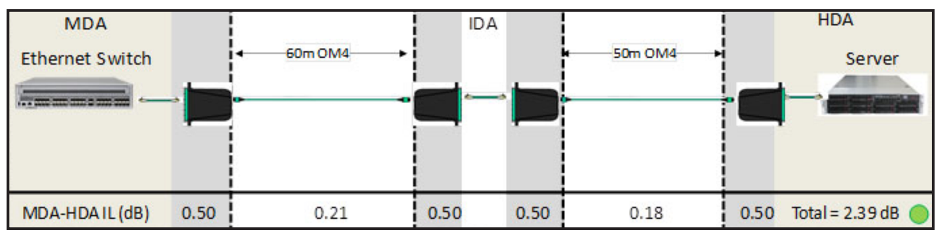

Figure 13. Loss budget calculations for server-to-network link over OM4 fiber

Scenario 2a - Server to Network - Additional Span Added

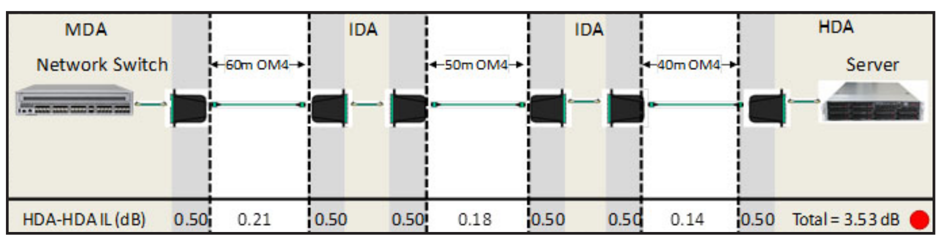

Figure 14. Loss budget calculations for server-to-network link over OM4 fiber—additional connections added

If we consider the possibility of upgrading this link from 10 GbE to 40 GbE by using 40GBASE-SR4 optics, the insertion loss calculation is as follows, shown below. Note: the overall insertion loss has been reduced when upgrading from 10 GbE serial to 40 GbE with parallel optics, based on the substitution of the LC/MPO modules with simple MPO adapter panels. However, in spite of the lower insertion loss, the link has exceeded the overall loss budget for 40GBASE-SR4 of 1.5 dB for applications on OM4 fiber. As such, this link would likely fail or experience errors when upgraded to 40 GbE.

Scenario 2b - Server to Network - Upgrade to 40 GbE over OM4 Fiber

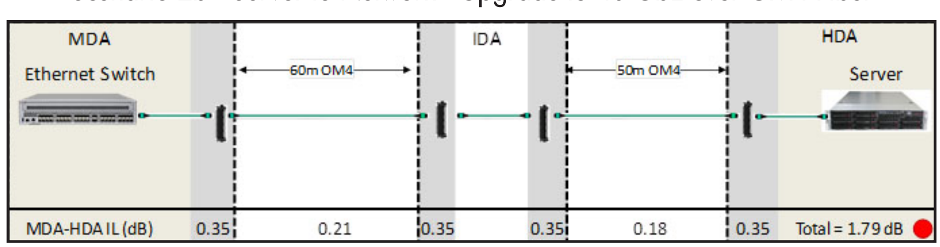

Figure 15. Loss budget calculations for server-to-network link over OM4 fiber—40G upgrade

The previous illustrations show the basic calculations considering the loss from connectors, cabling and the specification limits. Our actual data centewr design includes a variety of patching requirements, so the total of all connections and sum of the cable lengths must be considered for the various link combinations we need to support.

To answer these questions, we can take a simplistic approach. Adding the total losses and comparing the results to the standards requirements yields a go/no-go decision. Repeating this process for each link topology and application type gives us an overall understanding for the entire data center. This process is time consuming. Change vendors, for example, and the maximum loss for each component might change. Still, other vendors do not quote maximum values—rather typical loss values, which are not useful for this exercise.

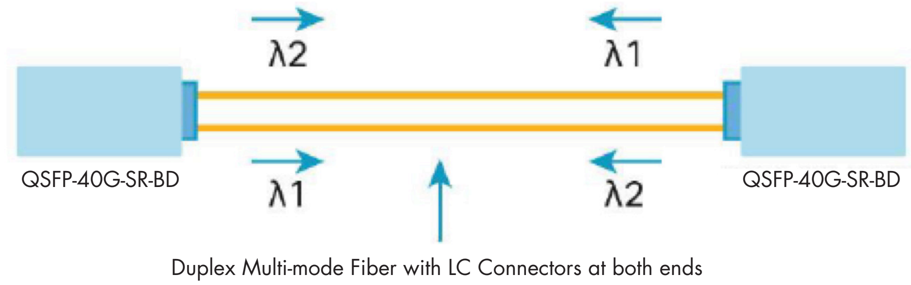

Figure 16. Bidirectional transmission

There are other vendor-specific proprietary technologies that have specific link limits, but they are not defined by industry standards. The Cisco BiDi scheme QSFP-40G-SR-BD (illustrated above) is an example of a new link design utilizing two wavelengths on each of two fibers, for an overall capacity of 40 Gb/s. In this case, there are no standardized limits with which to compare the link designs. The link design depends on the vendor claims and is subject to the design information they supply for various cabling topologies.

There are many options, given the volatility in network design and the myriad media types from which to choose. Engineering fabrics based on link criteria is not an easy task. While fabric networks have inherent fault tolerance, the physical links should not introduce a point of risk. Opting to purchase standards-based components requires the end user to evaluate the overall link designs and then determine if they are fit for the purpose. There are no vendor guarantees that the end user's design will function as required. These vendors only certify the component performance—not the overall link function.

Fabric topology design—application support

The previous elements of capacity, cabling topology, density, reach and network hardware requirements all sum into support for a particular link design or network application. Keeping options open means considering the permutations and combinations that make sense for your data center. Will a proprietary solution lock-in limit your options in the future?

CommScope has developed the InstaPATCH 360 solution to provide a modular plug-and-play solution that supports all of the combinations of fiber types, channel counts and topology strategies supporting fabric networks. To further support this, CommScope offers application design support. You can determine how to support any network application—standards-based or not—and match it to the modular topology your data center requires. Mix, match, and compare the performance of network hardware cost and performance quickly and easily. Reduce design errors and anticipate future capacity plans based on the CommScope Application Assurance Guide. CommScope provides supported distance for the link based on fiber type, number of connectors and application.

The charts below illustrate the guaranteed engineered link support for a few common fabric link choices, including the applications outlined in scenarios 1, 2 and 2a above. Based on the following chart, scenario 1 (8G Fibre Channel over 95 meters of LazrSPEED® 550 [OM4] fiber, with six MPO and six LC connectors) would be fully supported. As indicated, this topology could be supported for up to 150 meters.

8 Gigabit Fibre Channel, 850 nm Serial “limiting receiver” (FC-PI-4 800-MX-SN)

Supportable Distance ft (m)

LazrSPEED 550 with LC Connections

#LC Connections with:	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
0	790 (240)	740 (225)	740 (225)	690 (210)	690 (210)	640 (195)
1	740 (225)	740 (225)	690 (210)	690 (210)	640 (195)	640 (195)
2	740 (225)	740 (225)	690 (210)	640 (195)	640 (195)	590 (180)
3	740 (225)	690 (210)	690 (210)	640 (195)	640 (195)	590 (180)
4	690 (210)	690 (210)	640 (195)	640 (195)	590 (180)	540 (165)
5	690 (210)	640 (195)	640 (195)	590 (180)	590 (180)	540 (165)
6	690 (210)	640 (195)	590 (180)	590 (180)	540 (165)	490 (150)

Figure 17. 8G Fibre Channel performance over LazrSPEED 550

40GBASE-SR4 utilizes MPO connectors. Standards-based transceivers and the corresponding maximum reach with a given cabling topology can be read directly from the table. A link with six MPO connections can be configured with a maximum of 140 meters of LazrSPEED 550 OM4 trunk cable. Comparing this to the 100GBASE-SR4 table shows a maximum link length of 115 meters. Designing the day-one reach to a maximum of 115 meters would provide a supported upgrade path to 100G utilizing the same cabling infrastructure.

40 Gigabit Ethernet, 850 nm Parallel (40GBASE-SR4)

Supportable Distance ft (m)

LazrSPEED 550

#MPO Connections	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	570 (175)	560 (170)	540 (165)	510 (155)	490 (150)	460 (140)

LazrSPEED 300

#MPO Connections	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	460 (140)	440 (135)	430 (130)	410 (125)	390 (120)	380 (115)

Figure 18. Applications performance—40GBASE-SR4 over LazrSPEED fiber

100 Gigabit Ethernet, 850 nm 4-lane Parallel (100GBASE-SR4)

Supportable Distance ft (m)

LazrSPEED 550 WideBand and LAZRSPEED 550

#MPO Connections	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	390 (120)	390 (120)	370 (114)	370 (114)	350 (108)	350 (108)

LazrSPEED 300

#MPO Connections	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	280 (85)	280 (85)	280 (85)	260 (80)	260 (80)	250 (75)

Figure 19. Applications performance 100GBASE-SR4 over LazrSPEED fiber

CommScope application assurance also extends to nonstandard vendor-specific networking options. The tables below show the engineered link support for Cisco BiDi 40G technology. Extended-reach CSR4 options are also shown below. A comparison of the two options provides the designer with the maximum reach for these alternatives over LazrSPEED 550 OM4 fiber (150 meters for Cisco BiDi vs. 420 meters for 40GBASE-SR4).

40 Gigabit Ethernet, Cisco "BiDi" (QSFP-40G-SR-BD)

Supportable Distance ft (m)

LazrSPEED 550 WideBand and LazrSPEED 550 with LC Connections

#LC Connections with:	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
0	490 (150)	490 (150)	490 (150)	480 (145)	480 (145)	460 (140)
1	490 (150)	490 (150)	490 (150)	480 (145)	460 (140)	460 (140)
2	490 (150)	490 (150)	480 (145)	480 (145)	460 (140)	440 (135)
3	490 (150)	480 (145)	480 (145)	460 (140)	460 (140)	440 (135)
4	490 (150)	480 (145)	460 (140)	460 (140)	440 (135)	430 (130)
5	480 (145)	460 (140)	460 (140)	440 (135)	440 (135)	430 (130)
6	480 (145)	460 (140)	440 (135)	440 (135)	430 (130)	410 (125)

40 Gigabit Ethernet, 850 nm Parallel Extended Reach for Cisco (QSFP-40G-CSR4 TRANSCEIVERS)

Supportable Distance ft (m)

LazrSPEED 550

#MPO Connections*	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	1380 (420)	1310 (400)	1310 (400)	1310 (400)	1310 (400)	1310 (400)

LazrSPEED 300

#MPO Connections*	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	980 (300)	980 (300)	980 (300)	980 (300)	980 (300)	980 (300)

LazrSPEED 150

#MPO Connections*	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distance ft (m)	490 (150)	490 (150)	490 (150)	490 (150)	490 (150)	490 (150)

*Number of connections excludes the connection to the active device of each end of the channel

Figure 20. Applications performance—Cisco applications

A review of all four scenarios shows a comparison of application support and guarantee with the standard, components-based method, which uses manual analysis, and with CommScope's guaranteed applications performance. In the case of the components-based designs, only one out of four scenarios would have met the design requirements and loss budgets. Using the CommScope InstaPATCH 360 system, each of the four design scenarios would be met.

Scenario	Application	Total Length	LC	MPO	Supported with standard components	CommScope support
1	8G Fibre Channel	95m	6	6	No	Yes
2	10GbE	110m	4	4	Yes	Yes
2a	10GbE	150m	6	6	No	Yes
2b	40GbE	110m	0	4	No	Yes

Figure 21. Comparison of scenarios

Fabric links—Ethernet options

Singlemode, multimode, parallel, or duplex cabling options—the right choice for your data center will depend on its size, the growth pace of the services it supports, hardware budgets, vendor technologies and more. The design for network cabling must integrate with the network architecture, topology and road map.

Consider the two alternatives for 40G Ethernet links (above):

- 40GBASE-SR4, four lanes of 10G over eight fibers
- The proprietary BiDi duplex fiber design, which multiplexes two lanes of 20G onto the single pair of fibers

The reach these solutions offer is quite different, as you can see by comparing the application solution tables for each case. The CSR4 can range up to 420 meters on LazrSPEED 550 OM4 fiber—versus 150 meters for the BiDi. These distances and topologies shown in the application tables are based on use of the CommScope InstaPATCH 360 pre-terminated fiber solution. This example allows the designer to compare and design links with two nonstandard transceivers while also comparing the reach and topology capabilities to standards-based transceivers.

The InstaPATCH 360 system supports the 40GSR4 standards-based transceiver through six connections over 140 meters of fiber compared to the standards-based requirement of 125 meters—much greater reach and topology flexibility for fabric designs. Looking ahead to 100GSR4, 108 meters with six connections is supported.

Fabric links—next steps

We previously discussed the advantage of high-capacity fabric links (40G and above) enabling more servers and storage devices to share in higher overall network capacity. While today, 40G fabric links are cost-effective and efficient for many fabric designs, it is likely only a matter of time before the access port speed for servers will climb to 25G or perhaps 50G. In the coming years, fabric links will, of course, need to increase to 100G or perhaps 400G.

Looking ahead, there are a number of choices for higher-speed links. Some vendors advocate singlemode optic solutions. Others advocate multimode optics. For each of these media choices, there are potential duplex and parallel channel options. The relative cost of these options continues to evolve rapidly. Some basic relationships still hold—multimode optic systems continue to be lower in capital cost and are perhaps easier to maintain and operate compared to singlemode optic systems.

Fabric links—new media options

OM3 and OM4 multimode fiber supports 40G links with a reach and topology flexibility that meets the needs of all but the very large “warehouse-scale” data centers. Looking ahead to 100G and beyond, one of the more promising methods of increasing network capacity involves adding more communication channels to each duplex pair of multimode fibers. While wavelength division multiplexing (WDM) technology has been available for singlemode optics at a high cost, new shortwave WDM (SWDM) transceivers will combine four channels on a fiber pair, gaining four times the capacity on multimode fiber. This lower-cost alternative combines the ease of installation and operations of multimode fiber with a bandwidth scaled up to support the growth fabric networks will require.

Figure 22. Shortwave division multiplexing over wideband MMF

In support of SWDM, CommScope—together with other members of the SWDM Alliance—have developed a new "wideband" multimode fiber media, or WBMMF. This fiber is designed to expand the available capacity of multimode media, enabling more communication channels per fiber at longer reach. WBMMF will be used to deliver higher data rates of 100 GB/s and 400 Gb/s while reducing the number of fibers required to support these future high-capacity fabric networks.

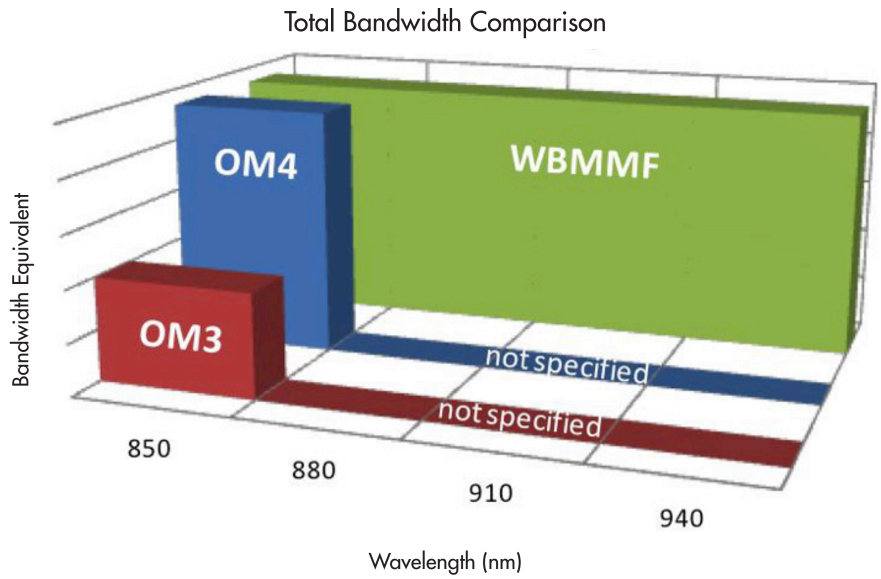

Figure 23. A comparison of total bandwidth between OM3, OM4, and WBMMF

WBMMF is backward compatible with OM3 and OM4 fibers. WBMMF standardization has progressed rapidly within the Telecommunications Industry Association (TIA) TR-42 committee with publication of a standard anticipated in 2016. CommScope has released the LazrSPEED 550 Wideband cabling solution across all components of the InstaPATCH platform—the next-generation path to low-cost high-capacity networks is available today.

Conclusions

In response to the demand for lower costs and higher capacities, data centers are adapting new fabric network-based systems to support cloud-based compute and storage systems. Data center cabling topologies are increasing in density to support the any-to-any low-latency communications typically required by distributed cloud applications.

The design of high-capacity links is more complex given that the number of network links must increase and the network speeds are increasing. Providing more data center capacity means pushing the limits of existing media and communication channel technologies. Fiber apparatus designs and WBMMF are also evolving to provide next-generation capacity and physical density perfectly suited to fabric network architectures. Singlemode fiber provides support for longer channel lengths.

Application design and engineered link solutions from CommScope ensure reliable high-speed networks that are designed to meet the rigorous demands of current and future network capacity requirements. InstaPATCH 360 systems provide greater reach for high-capacity links, design topology freedom to scale to very large complex environments, and guaranteed application performance for both standards-based and emerging proprietary systems.

Engineered solutions make complex fabric network designs easy to design, implement and manage. Preterminated high-performance systems support the next-generation network media and duplex and multifiber modular applications, while reducing deployment management time and expense.

COMMSCOPE®

www.commscope.com

Visit our website or contact your local CommScope representative for more information.

TP-110117.1-EN (04/16)