

Structured Cabling Makes a Comeback With High-Speed, Low-Latency Deployments

Switch-to-server connections cannot be a weak link



Emerging IoT and IIoT technologies, like self-driving cars, virtual and augmented reality (VR and AR), AI, machine-to-machine (M2M) communication, and advanced data analytics require faster transmission speeds and low-latency communication between servers and switches at the edge of the network. As modern hyperconverged data center environments look to support these emerging technologies, switch-to-server connections cannot be a weak link in the chain.

Migration strategies for switch-to-server connections that were once projected to advance from 10 Gbps to 40 Gbps have now evolved due to innovations in switching tech-

nology. To minimize latency and bottlenecks, data centers have also shifted away from a traditional three-tier architecture to a more efficient, full-mesh leaf-spine fabric architecture. When it comes to switch-to-server connections, these technology trends indicate that fiber optic structured cabling with an “all-to-all” cross connect scenario may ultimately make the most sense for modern data centers. Data center managers embarking on modernization projects to support high-speed, low-latency performance would therefore be wise to keep their eyes on the evolving landscape and timeless benefits of structured cabling.

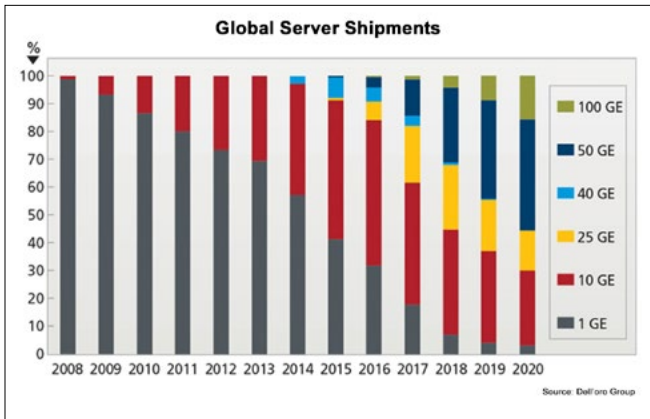


FIGURE 1: The shift to 25 Gbps per lane made possible via PAM4 encoding is driving an increase in 25, 50, and 100 Gbps server ports and a decline in 10 and 40 Gbps.

A BETTER TRAFFIC PATTERN

Historically, data centers have been designed using a three-tier hierarchical architecture with core, distribution, and access layer switches. While this design accommodated traffic between servers that connected to the same access switch, any traffic between different access switches needed to transmit in a north-south pattern through higher-level switch tiers. Unfortunately, a north-west traffic pattern introduces latency due to additional connections between switches, which can be a major impediment for modern hyperconverged data centers with increased virtualization, software-defined networking, and shared compute and storage resources to support emerging data-intensive and time-sensitive applications.

To reduce latency and optimize east-west traffic for server-to-server communications, many modern data centers are shifting to a full-mesh leaf-spine fabric architecture where every leaf switch connects to every other leaf and spine switch within the fabric. This limits the number of switches that data needs to traverse, enabling more direct pathways between devices. Reducing latency with an east-west traffic

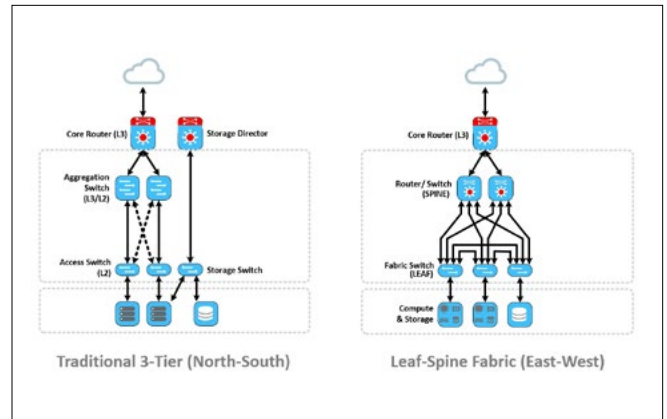


FIGURE 2: Traditional three-tier architecture versus leaf-spine fabric.

pattern significantly improves performance for virtualized server environments where resources for a specific application are often distributed across multiple servers.

Despite the benefits, a leaf-spine architecture is more costly to implement than a traditional three-tier architecture — connecting every switch to every other switch increases the number of cables and can lead to more expensive spine switches with higher port counts. In smaller enterprise data centers where speed and latency are not critical considerations, a traditional three-tier architecture may continue to make sense for the time being. However, due to the benefits, leaf-spine architecture is now trending among most mid-to large-scale hyperconverged enterprise data centers, as well as for larger cloud, colocation, and hyperscale data centers.

25 IS THE NEW 10

For switch-to-server connections, data centers have historically deployed 10 Gbps using either a top-of-rack (ToR) deployment where switches in each cabinet connect directly to the servers in that cabinet via short-length, high-speed twinax direct attach cables (DACs) or a middle-of-row (MoR) or end-of-row (EoR) deployment where switches are placed in one

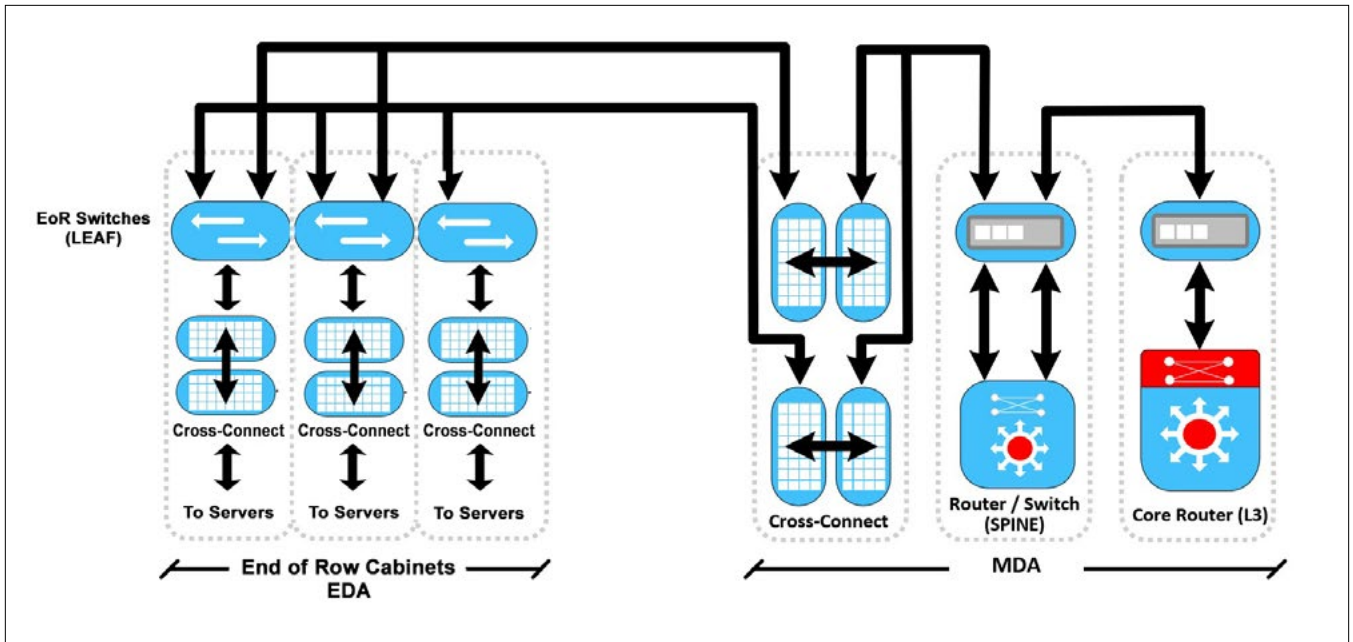


FIGURE 3: Leaf-spine architecture with cross-connects between active equipment.

cabinet and connected to servers across an entire row using balanced twisted-pair copper category 6A structured cabling. EoR deployments require distances of about 30 meters to reach servers in adjacent cabinets, while MoR deployments require about half of that. For simplicity sake, this article will focus only on ToR and EoR deployments.

While smaller enterprise data centers will still use 10 Gbps copper links for quite some time, emerging applications in larger data centers are beginning to tax the capacity of 10 Gbps links. It was originally anticipated that 10 Gbps would migrate to 40 Gbps based on a 10 Gbps per lane approach using non-return-to-zero (NRZ) encoding technology. However, the latest four-level pulse amplitude modulation (PAM4) encoding technology offers twice the bit rate and enables a 25 Gbps per lane approach. This technology has shifted the previous 10/40/100 Gbps migration path for server connections to a less-disruptive 25/50/100/200 Gbps path, which is well-suited for leaf-spine fabric architectures because it can significantly

reduce the number of physical ports required. As switches and servers are refreshed, enterprise data centers are therefore migrating to 25 Gbps, and it is expected that 40 Gbps applications will eventually phase out. As shown in Figure 1, a study published by Dell'oro Group on global server port shipments confirms the growth of 25, 50, and 100 Gbps switch ports and the decline of 10 and 40 Gbps.

For ToR deployments in enterprise data centers, high-speed DACs capable of supporting 25 Gbps are already gaining ground. However, the short 5-meter distance supported by these DACs cannot accommodate EoR deployments. In 2016, category 8 cabling was ratified by industry standards to support 30-meter 25 Gbps and 40 Gbps applications (i.e., 25GBASE-T and 40GBASE-T) in EoR deployments. However, due to the introduction of PAM4 technology and higher cost and power consumption, 25GBASE-T and 40GBASE-T PHY development never fully came to fruition, essentially preventing the adoption of BASE-T and category 8 cabling for speeds

Application	Media	Distance
25GBASE-CR	SFP28 passive copper DACs	5 meters
25GBASE-SR	OM4 multimode fiber	100 meters
25GBASE-LR	singlemode fiber	10 kilometers

TABLE 1: Options for 25 Gbps switch-to-server connections.

Application	Media	Distance
50GBASE-CR	SFP56 passive copper DACs	3 meters
50GBASE-SR	OM4 multimode fiber	100 meters
50GBASE-FR	singlemode fiber	2 kilometers
50GBASE-LR	singlemode fiber	10 kilometers
100GBASE-CR2	QSFP28 passive copper DACs	3 meters
100GBASE-SR2	OM4 multimode fiber	100 meters
100GBASE-FR1*	singlemode fiber	2 kilometers
100GBASE-LR1*	singlemode fiber	10 kilometers
100GBASE-DR	singlemode fiber	500 meters

TABLE 2: Options for 50 and 100 Gbps switch-to-server connections. *Pending IEEE Standard.

beyond 10 Gbps. Aside from backplane and long-haul (40 km) single-mode deployments, current 25 Gbps applications for switch-to-server connections are shown in Table 1.

To reduce cost via lower port counts and space savings, many data center managers are choosing to deploy four lanes of 25 Gbps via 100 Gbps switch ports using 4X25 Gbps breakout assemblies connecting one QSFP28 to four SFP28 connections with one MPO connected to four duplex fiber connectors.

For longer-distance 25 Gbps EoR deployments, data center managers can choose between active optical cable (AOC) assemblies or use fiber optic structured cabling with separate transceivers. AOCs essentially embed an optical transceiver into the connector and use fiber optic cable to support up

to 100-meter distances. While AOCs appear less complex and offer slightly lower power consumption, they do not have to comply with industry standards for interoperability. With embedded transceivers, they also do not offer the flexibility of supporting multiple generations of applications and ultimately will need to be replaced should an upgrade be needed. The use of embedded transceivers also makes the use of breakout applications more challenging, limiting the flexibility of using higher-speed 100-, 200-, and 400-Gbps switch ports to connect to multiple 25-, 50-, and 100-Gbps server ports.

HYPERSCALE TREND SETTERS

The same PAM4 advancement that shifted 10 Gbps per lane to 25 Gbps per lane now

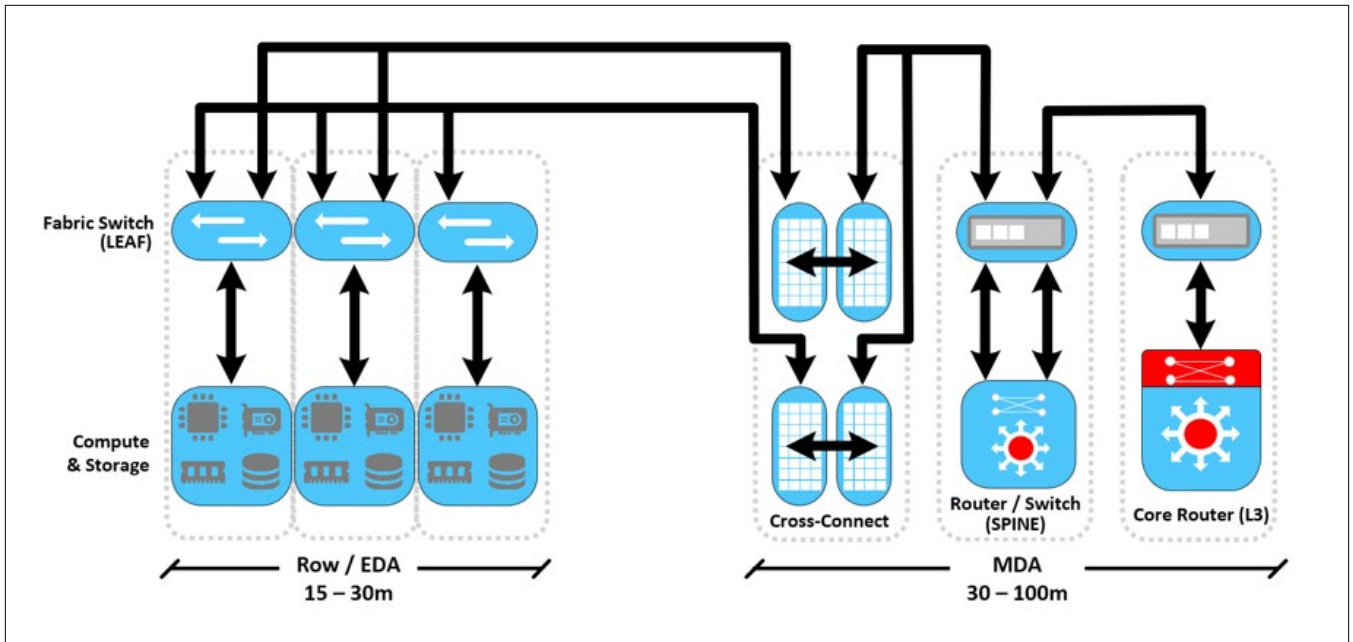


FIGURE 4: Leaf-spine architecture with cross-connect.

also enables 50 Gbps and 100 Gbps per lane. In large cloud and hyperscale data centers where 25 Gbps does not offer enough bandwidth, the deployment of 50- and 100-Gbps switch-to-server links has begun. Within the hyperscale environment, these data center links are primarily deployed using the media types shown in Table 2.

Just like with 25GBASE-SR, 50GBASE-SR and 100GBASE-SR2 deployments can also use either AOC assemblies or a fiber optic structured cabling approach with transceivers. To reduce cost via lower port counts and space savings, hyperscale trend setters are also starting to deploy multiple lanes of 50 Gbps or 100 Gbps via 200- and 400-Gbps switch ports using fiber breakout assemblies and next-generation fiber applications that are also based on PAM4 technology. In fact, the IEEE Standard 802.3cm approved in January of 2020 creates additional options for 400 Gbps operation, including 400GBASE-SR8

that transmits 50 Gbps per lane over eight lanes of multimode fiber up to 100 meters. 400GBASE-SR8 has broad market potential as it will be ideal for cost-effectively connecting a single 400 Gbps switch port to up to eight 50 Gbps server ports, four 100 Gbps server ports, or two 200 Gbps server ports. The IEEE is already working on 800 Gbps applications that will transmit 100 Gbps per lane over eight lanes.

While singlemode fiber deployments have long been considered more costly due to transceiver cost, larger hyperscale data centers are also trending toward singlemode fiber that supports longer distances and virtually unlimited bandwidth. Some of these hyperscale data centers can be upwards of 1 million square feet (nearly 100,000 square meters), which introduces distances greater than 100-meter AOCs and multimode solutions can support. While transceiver costs have come down in recent years, transceivers

for shorter-reach single-mode applications, such as 100GBASE-DR for up to 500 meters, can be more economically developed compared to those for traditional, longer-reach, 10-km, single-mode applications.

Hyperscale data centers tend to be the early adopters that ultimately shape the data center industry, and technology emerging within these environments will eventually become attractive options for enterprise data center operators as they migrate to speeds beyond 25 Gbps over the next decade.

BACK TO STRUCTURED CABLING

With the new 25/50/100 migration path and shift to a leaf-spine architecture, the overall design of the cabling infrastructure is also experiencing a shift. While 25-Gbps ToR deployments using single-lane 25GBASE-CR over SFP28 passive copper DACs for distances up to 5 meters or four lanes of 25 Gbps via 100GBASE-CR4 over breakout assemblies (i.e., QSFP28 to four SFP28) will remain popular for smaller data centers, large hyperconverged modern data centers are now shifting back toward the use of EoR deployments and structured cabling.

Well-suited for supporting east-west data traffic via leaf-spine architectures in virtualized distributed server environments, EoR deployments with structured cabling allows any two servers in a row to experience low-latency communication because they are connected to the same switch. In contrast, ToR deployments require an extra switch-to-switch transmission, introducing greater latency when a server in one cabinet needs to communicate with a server in another cabinet.

EoR deployments using structured cabling are easily achieved for advanced 25 Gbps and beyond speeds using fiber optic structured cabling. The trend for these deployments is to

use cross-connects where patch panels that mirror leaf switch ports connect via permanent cabling to patch panels that mirror the spine switch ports. Connections between the leaf and spine are made at the patch panels via patch cords, enabling “all-to-all” connectivity where any spine port can connect to any leaf port.

With the ability to be located anywhere in the data center, cross-connects support hyperconverged virtualized environments, enabling shared compute and storage resources to be clustered together and connected to a single system for better availability and reliability. Cross-connects are also popular among cloud and colocation data centers because they enable connections between any carrier or cloud service to any customer network to be made at a cross-connect outside of secure meet-me rooms.

Compared to ToR deployments, EoR deployments with “all-to-all” cross-connects also offer the benefit of structured cabling that reduces data center operating expense, including maximum port utilization, scalability, and improved manageability.

PORT UTILIZATION

When using a ToR deployment, data center managers may discover they are not fully utilizing all switch ports due to power and cooling concerns that often limit the number of high-performance virtualized servers per cabinet. These unused switch ports across several cabinets can add up and ultimately equate to unnecessary switch purchases and related maintenance and power.

Even when enough power and cooling can be supplied to a cabinet to support a full complement of servers, switch port utilization can be a concern with ToR — if the number of servers surpasses the number of available switch ports in the cabinet, the only options

are to locate the server elsewhere or add another ToR switch, further resulting in poor port utilization.

In contrast, using an EoR deployment with structured cabling allows virtually all active switch ports to be fully utilized because they are not confined to single cabinets. Via cross-connects, switch ports on higher-density leaf switches can be divided up, on demand, to any of the servers across several cabinets in a row.

SCALABILITY

In ToR deployments, a single switch upgrade from 10 Gbps to 25 Gbps improves connection speed to only the servers in that cabinet, and a widespread switch upgrade impacts many more switches. Having a greater number of ToR leaf switches also demands higher port densities for spine switches, which can cause scalability constraints. On the other hand, a single switch upgrade in an EoR deployment can increase connection speeds to multiple servers across several cabinets in a row, and fewer EoR leaf switches can reduce density requirements on spine switches.

ToR deployments can also land-lock equipment placement due to the short cabling lengths of DACs, which can prevent placing new equipment where it makes the most sense for power and cooling within a row or set of rows. For example, if budgets do not allow for outfitting another cabinet with a ToR switch to accommodate new servers, placement of the new servers may be limited to where network ports are available. This can lead to hot spots, which can adversely impact neighboring equipment within the same cooling zone and, in some cases, require supplemental cooling. EoR deployments with structured cabling avoids these problems.

Through backward compatibility and interoperability, and due to advancements like

PAM4 and SWDM, standards-based fiber structured cabling used with EoR deployments can also support multiple generations of switch-to-server connections ranging from 10 Gbps to 400 Gbps. This allows data center operators to leverage their existing cabling investment during upgrades regardless of which vendor's equipment is selected.

Unlike standards-based fiber optic solutions that work with all switches, regardless of speed or vendor, proprietary DACs may be required by some equipment vendors for use with their ToR switches. While this helps ensure that vendor-approved cable assemblies are used with corresponding electronics, proprietary cabling assemblies are not interoperable and can require cable upgrades to happen simultaneously with equipment upgrades. In other words, vendor-specific DACs will likely need to be swapped out if another vendor's switch is deployed.

MANAGEABILITY

ToR deployments can quickly become a management burden in large data centers where many more switches need to be individually managed. The cabling distance limitations of ToR deployments can also further limit manageability. For 25-Gbps ToR deployments, point-to-point cabling using DACs is limited to 5 meters, and, as speeds increase to 50 Gbps and 100 Gbps via 50GBASE-CR or 100GBASE-CR2, distances are even shorter for DACs at 3 meters. As distances decrease, ToR switches may have to move to an MoR position to effectively reach every server in the cabinet.

The cabling distance supported by EoR deployments with structured cabling using multimode or single-mode fiber solutions with transceivers can range between 100 meters and 10 km depending on the

application, which allows for more-flexible equipment placement throughout the life of the data center.

With structured cabling, moves, adds, and changes (MACs) can be accomplished within the “all-to-all” cross-connect by simply repositioning patch cord connections at corresponding patch panels, enabling a highly reconfigurable leaf-spine fabric. In contrast, because ToR switches connect directly to the servers in the same cabinet, all changes must be made within each individual cabinet. While this is doable in smaller edge and micro-edge data centers where rack counts range from about 1 to 50 cabinets, in larger modern data center deployments, making changes in each cabinet is complicated and time-consuming.

With structured cabling and cross-connects, the permanent portion of the channel between switches and patch panels remains unchanged, allowing active equipment to be left untouched and secure. This is also ideal for environments where switches and servers are managed by separate resources or departments. ToR deployments do not allow for physically segregating switches and servers into separate cabinets, and MACs require touching critical switch ports.

QUALITY AND RELIABILITY

While the argument can certainly be made that advanced 25 Gbps and beyond server connections in today’s hyperconverged data centers benefit from the use of structured cabling from a port utilization, scalability, and manageability standpoint, there are additional considerations surrounding quality and reliability that should always be on every data center manager’s radar.

First of all, while standards-based multimode and single-mode structured cabling work with any vendor’s equipment, some ToR

switches are designed to check vendor-security IDs on the DACs connected to each port and either display errors or prevent ports from functioning when connected to an unsupported vendor ID. Data centers may therefore be wise to consider purchasing DACs from reputable cable manufacturers that are third-party tested to work with all vendor equipment. DACs from reputable cable manufacturers often also come in more lengths and colors with a longer warranty than the average 90 days offered by some switch vendors. It should however also be noted that multimode and single-mode structured cabling from reputable manufacturers typically carry a much longer 15- to 25-year warranty.

For longer-distance solutions, structured cabling also offers advantages over AOCs. Not only are AOCs not recommended for distances greater than 100 meters, but they do not have to comply with industry standards for interoperability and can lock data center managers into a single application. AOCs have also often been the cause of errors due to defects, such as incorrect polarity, excessing bends, and damaged optical fibers, and they tend to be more susceptible to damage during installation, cannot be repaired in the field, and their fused optics makes testing difficult.

Regardless of which solution is selected for switch-to-server links — DACs, AOCs, or structured cabling with transceivers — data center managers also need to ensure that they are purchasing a quality product that has been fully verified for performance through third-party testing facilities, such as Intertek ETL and UL. This is especially important when investing in structured cabling multimode and single-mode cabling solutions that may be in place for up to 25 years, as they can support multiple generations of transceivers.

Purchasing fiber cabling from lower-cost, unproven sources introduces the possibility of installing potentially substandard or even counterfeit components where the actual fiber used in the cable is not as advertised. This doesn't just pertain to the permanent links between patch panels but also the fiber patch cords used to make connections between patch panels and active equipment. Too often, data centers will regard patch cords as an afterthought and purchase lower-cost cords from unproven sources. Not only can patch cords also have the incorrect glass type, but they have been shown to have subpar fiber end-face geometry, lower performance, and poor mechanical reliability, often failing TIA-568-C.3 Fiber Optic Test

Procedures (FOTPs) for cable pull, flex, torsion, and retention.

SUMMARY

There is no single infrastructure design for every data center. However, it is clear that modern hyperconverged data center environments will benefit from the use of standards-based fiber optic structured cabling. While this timeless approach appeared to be going by the wayside with the popularity of ToR deployments, as we look toward 25 Gbps and beyond and the deployment of leaf-spine fabrics to support high-speed, low-latency communications for emerging technologies, structured cabling is well-positioned to make its comeback. ■

ADVERTISERS' INDEX

Kohler

www.missioncriticalmagazine.com/webinars

Page 2

MISSION CRITICAL (ISSN: Digital 1947-153X) is published 6 times annually, bi-monthly, by BNP Media, Inc., 2401 W. Big Beaver Rd., Suite 700, Troy, MI 48084-3333. Telephone: (248) 362-3700, Fax: (248) 362-0317.

Copyright 2020, by BNP Media II, L.L.C. All rights reserved. The contents of this publication may not be reproduced in whole or in part without the consent of the publisher. The publisher is not responsible for product claims and representations.

Change of address: Send old address label along with new address to MISSION CRITICAL, P.O. Box 2145, Skokie, IL 60076.

Email: missioncritical@omeda.com

For subscription information or service, please contact Customer Service at: Phone: (800) 952-6643 Fax: (847) 291-4816.

A **bnp** PUBLICATION
media

INDUSTRY ALLIES

