COMMSCOPE®

SYSTIMAX® Solutions

Top of Rack: An Analysis of a Cabling Architecture in the Data Center

White paper

Matthew Baldassano, Data Center Business Unit
CommScope, Inc, June 2010

at work

Contents

I. Executive Summary

Cabling architectures lay the infrastructure foundation for data center networks. Sometimes overlooked, physical layer connectivity is a necessity that requires careful planning and consideration when choosing a cabling architecture. There are several types of cabling architectures found in the data center; direct-connect, zone-distribution, and top-of-rack (ToR). Each architecture offers advantages over the others depending on DC function, uptime tier level, growth (planned and unplanned), and equipment density. A "one-size-fits-all" approach to each of these architectures also may not fit everyone's needs and thus hybrid approaches are commonly used.

This white paper seeks to give a detailed analysis of ToR cabling architectures. Advances made in the networking world, including the introduction of 10 Gigabit Ethernet (10GbE) and Fibre Channel over Ethernet (FCoE) to the server, have leveraged ToR architectures as a solution.

TABLE 1: ADVANTAGES AND DISADVANTAGES OF THE TOP OF RACK CABLING ARCHITECTURE

	Pros	Cons
Top of Rack	- Most efficient use of cables - Scalable solution for small to large Data Centers - Easier cable management as compared to Zone-Distribution or Direct-Connect - Efficient use of cabinet space if not completely filled with servers - Modularity at server cabinet level	- Highest electronics cost as compared to Zone-Distribution and Direct-Connect - Implementation of redundant architecture mandates entirely redundant switches - Difficult to perform physical connection move, add or change - Increased network management overhead - Poor port utilization if not all ToR ports are used

Data centers that are built in ToR networks use low port count switches (typically ≤48 ports) placed in each server cabinet. Designation for the switch to be placed at the top of the cabinet allows for easier management of intra-cabinet patch cords routed to each server; however implementations where the switch is placed at the bottom or middle of the cabinet are not uncommon.

Inter-cabinet cabling (cables run between server cabinets) for ToR architectures typically consists of a lower ratio of connections (fiber or copper) exiting the cabinet than are needed to support each server within the cabinet. This allows the ToR switch to act as a cabling consolidation point. Aggregation switches are used to feed each ToR switch and most times are located in a central physical area.

II. Top of Rack Architecture

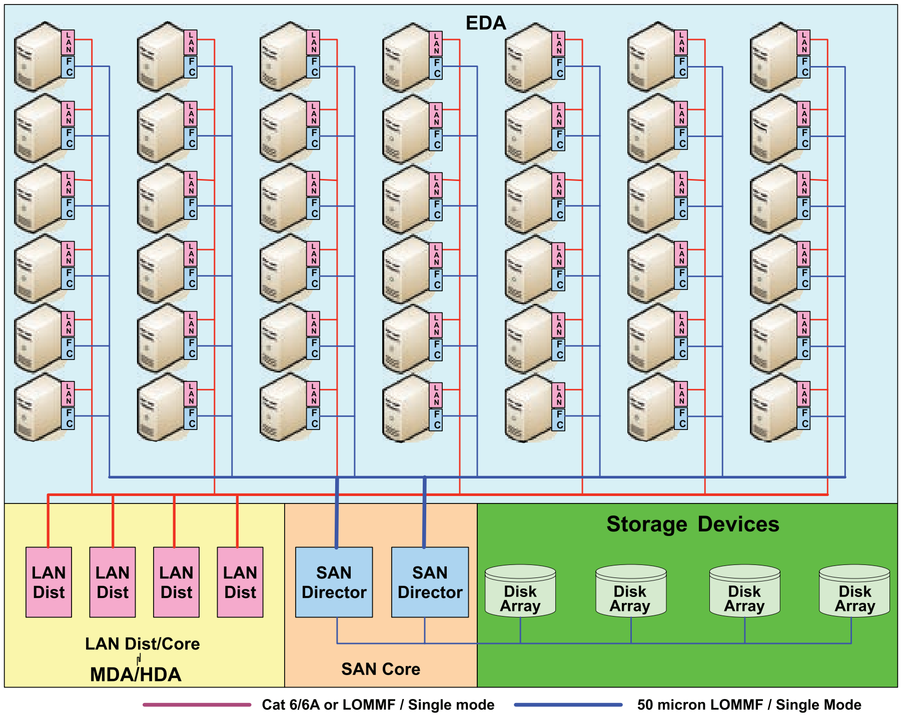
ToR architectures are typically used when deploying dense server environments that require several network connections per server. This is commonly found in clustered environments, load balancing appliances, web search engines, or in virtualized environments that share processing of several separate physical servers located within a single cabinet.

Intra- and inter-cabinet (Figure 1) cabling is easier to manage as ToR switches, which aim to reduce connections, will see a cable management benefit with ToR architectures. Within the cabinet it is common to find a vertical bundle of patch cords that are dressed properly and fanned out to each server port. With non-ToR architectures, managing these cable bundles in a centralized cross-connect or inter-connect area may present challenges.

FIGURE 1: TOP OF RACK INTER-CABINET CABLING

ToR architecture has existed in data centers for many years. Figure 2 represents the logical topology for both twisted pair copper and multimode fiber cabling. Prior to 10GbE, GbE connections were available, but all servers connected to the ToR switch would have to share a smaller bandwidth pipe. Uplinks capable of running 10GbE allow for GbE servers to share a much higher capacity channel, thus helping to limit bandwidth bottlenecks. ToR switches also rely on copper patch cords for the server connection and multimode fiber for the 10GbE uplink connection. Migration to 40/100GbE technologies will replace the existing 10GbE connections as ToR uplinks, but may require more dark fiber to be installed or running additional fiber when these speeds are required.

FIGURE 2: LOGICAL TOPOLOGY OF TOP OF RACK CABLING ARCHITECTURE



Notable benefits of ToR architectures include:

- Most efficient use of cables
- Scalable solution for small to large Data Centers
- Easier cable management as compared to Zone-Distribution or Direct-Connect
- Efficient use of cabinet space if not completely filled with servers

Notable drawbacks of ToR architectures include:

- Highest electronics cost as compared to Zone-Distribution and Direct-Connect
- Redundant design will double the number of ToR switches
- Difficult to perform physical connection move, add or change
- Increased network management overhead
- Poor port utilization if not all ToR ports are used
- Network stability risks due to potential layer-2 loops that cause broadcast storm

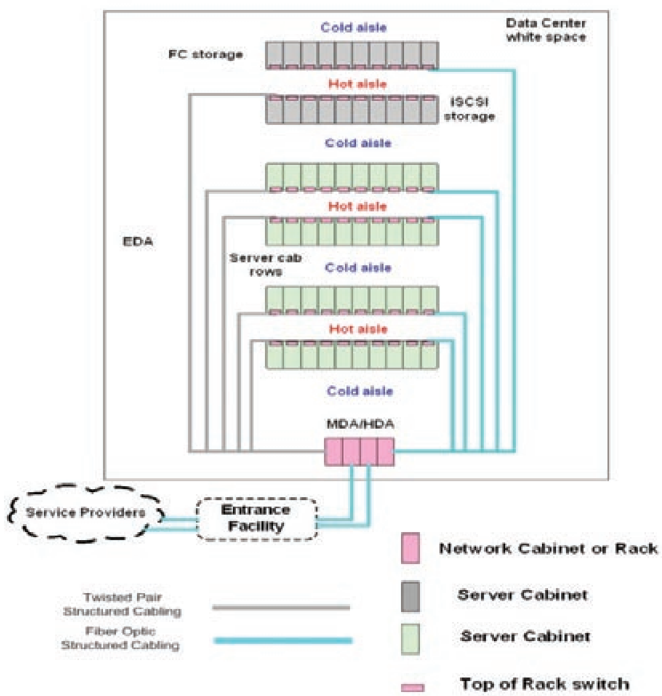
III. Top of Rack Examples and Considerations

Data center networks vary in size, complexity, and function however they sometimes may share resources such as a switches and storage. Architecting the physical layer of networks should be done to ensure that service reliability meets the availability requirements. For example, due to the critical nature of Fibre Channel (FC) protocol used in storage area networks (SANs) it has been best practice to have two sets of networks, also called fabrics with redundant pieces and parts down to the physical layer. However, LAN networks may not need the extra redundancy used in SANs based on it's ability to recover from interruptions. Out of band (OOB) management networks usually don't require GbE networks nor redundancy, however 10/100 switches and NIC cards have all but become obsolete and the price difference is not substantial.

In the case of OOB management networks GbE throughput may not be needed because only simple control functions and remote commands are issued over them. What is important for OOB management is to create a physically independent network dedicated for these functions. This physically independent network is to ensure that if both the primary and secondary paths to production LAN are offline remote access to the server is still possible.

ToR architecture for OOB management is possible using 10/100 or GbE switches and does not require 10GbE uplinks. Either copper or fiber GbE uplinks are sufficient to handle a server cabinet loaded with servers, but as mentioned previously they should be segregated at the physical layer.

FIGURE 3: PHYSICAL TOPOLOGY OF TOP OF RACK CABLING ARCHITECTURE



When deploying OOB management for blade servers that do not require 1 OOB port per server, it is suggested not to use ToR switches as the port utilization would be very low. A blade chassis uses fewer OOB connections to manage several blades relative to the deployment of individual rack mounted servers. Stacking 3 blade servers would still only require a limited number of OOB connections. Direct-connect architectures for blades are recommended for remote management functions.

Additional considerations exist for production traffic in blade servers. When using passthru mode in the blade enclosures Top of Rack would be ideal. When using built in blade switches the additional layer of networking inserted with the ToR switch can cause higher latency times, bottlenecks, and network management concerns.

10GbE has recently made its way into the data center networking world and has been driven by virtualization as well as higher data rates that are required for transaction oriented networks. Today 10GbE exists in several forms in data center networking equipment with the most popular being SFP+. Many network vendors have chosen to release 10GbE switches with SFP+ interface for uplink and server connections. SFP+ interface accepts SFP+ optical modules (10GBASE-SR) or CX1 direct-attach cables (10GBASE-CR). For the purpose of supported distances, SFP+ switches used in ToR implementations suggest using CX1 cables for server connections and SFP+ optical modules for uplink connections. The recommendation is based on the lower cost of CX1 cables which are pre-terminated with the SFP+ module and limited to several meters. They are also considerably less expensive than SFP+ optical modules needed for longer runs.

As power consumption of 10GBASE-T is driven lower through the evolution of PHY chips, Wake on LAN (WoL), and energy efficient Ethernet (EEE), 10GbE over twisted pair copper competes with the same capabilities of CX1 from a power efficiency perspective. Data centers currently using a direct-connect or zone-distribution cabling architecture will be able to leverage existing Category 6 and Category 6A installations for 10GbE. It should be noted that WoL and EEE do not include support for CX1 cables.

ToR switches with a fixed number of fiber uplink connections may become a bottleneck as the aggregate cabinet bandwidth increases or when adding additional servers. It is suggested as a best practice to run additional dark fiber in advance to avoid running more cables later.

Chassis switches, typically installed with n+1 active components, support 10GBASE-T port cards and allow for a more flexible and scalable data center cabling infrastructure. Chassis switches are also considered more reliable than ToR switches as they are designed with more advanced troubleshooting features and can support non-disruptive upgrades. If designing a ToR data center with the intent to support high service reliability it is important to provide redundant server connections to two or more ToR switches.

ToR architectures are not limited to LAN applications. FCoE protocol was designed to converge FC SANs to run over Ethernet protocol. This new Ethernet protocol is an industry standard referred to as converged enhanced Ethernet (CEE). The benefit of running once disparate physical networks over CEE allows for a single cable to support both. SAN vendors have adopted 10GbE ToR switches capable of supporting FCoE over a single cable to a server.

FCoE was designed to run over 10GbCEE and be can be managed transparently by a SAN administrator. The only change that will occur is the physical layer will no longer be dedicated to FC storage. Converging at the access layer (ToR switch) to support LAN and SAN data traffic requires a converged network adapter (CNA) to be installed for each server connection. CNAs support SFP+ optical transceivers and CX1 active cables and have independent Ethernet and FC onboard chipsets that split the traffic up prior to delivering each virtual interface to the operating system.

Collapsing LAN and SAN traffic at a ToR switch may become problematic when LAN or SAN upgrades, troubleshooting, security, or standard maintenance require the switch to reboot or take ports offline. Coordination between LAN and SAN groups was not needed until FCoE was introduced. FCoE capable director class switches would allow for easier change control processes to be put in place for both independent groups.

As mentioned previously, data centers that require high service reliability and are running FCoE using ToR must make sure that each server connects to two physically separate switches or risk the switch becoming a single point of failure for both LAN and SAN network. Security may also be at risk when combining once separate physical networks over a common network device.

IV. Total Cost of Ownership

Cabling architecture choices have a direct impact on the total cost of ownership (TCO) of a data center network. Factors such as white space dimensions, power/cooling capacity, anticipated growth, and refresh rates should be considered when choosing a network architecture.

Although not addressed in detail in this paper, direct-connect and zone-distribution architectures are alternative choices to ToR. The differences of cabling to electronic equipment varies significantly between the three architectures. In general direct-connect requires the most cabling as each server port is a channel directly run to a central consolidation point in the data center, however it has the lowest cost for electronic equipment. Zone-distribution is the best balance of cabling and electronics, but benefits are best realized when deploying a large or mega-data center, typically larger than 20,000 square feet.

ToR architectures require the least amount of physical layer cabling, but the highest electronics cost. When deciding which architecture is best suited to provide an optimized TCO model understanding cabling and electronics costs is important.

In a 10GbE data center network study CommScope found that cabling CAPEX accounted for 5% of a total network spending in a ToR architecture and the network and server electronics accounted for 95%. It should also be noted that network electronics typically are refreshed at much higher rate than the underlying infrastructure. Cabling infrastructure can be designed to support future data rates and can be used through technology refreshes thus extending the lifespan and reducing TCO.

TABLE 2: TOTAL COST OF OWNERSHIP CONSIDERATIONS OF THE TOP OF RACK CABLING ARCHITECTURE

CAPEX	Direct-Connect	Zone-Distribution	Top-of-Rack
Cable Cost	3x	2x	1x
Cable Type	OM4 Fiber CAT6A UTP	OM4 Fiber CAT6A UTP	OM4 Fiber CX1 – Twinax CAT6A UTP
Port / Transceiver Types	SFP+ 10GBASE-SR RJ45 10GBASE-T	SFP+ 10GBASE-SR RJ45 10GBASE-T	SFP+ 10GBASE-SR SFP+ 10GBASE-CR RJ45 10GBASE-T
Electronics Costs	1x	2x	3x
Cabling vs Electronics	25% Cabling 75% Electronics	7% Cabling 93% Electronics	5% Cabling 95% Electronics
OPEX			
Electronics Refresh Rate	3x		
Cabling Refresh Rate	1x		
Network Maintenance / Management	Low	Moderate	High

Recent advancements in both fiber and copper cabling media support this claim. In 2009 the OM4 specification for fiber was set as a standard and incorporated the specifications of pre-existing laser optimized cabling that had been in the market since 2002. Essentially any end-user who "future-proofed" their cabling infrastructure with higher quality fiber can leverage it to support longer distances for 10, 40 and 100G applications.

Category 6A can be used to preserve the 10GbE requirements set by 10GBASE-T network electronics. Green advancements such as Energy Efficient Ethernet (EEE) and Wake on LAN (WoL) are specific to 10GBASE-T and allow for pre-existing Category 6A cabling to be used as the network electronics make their way into server and networking equipment. Older twisted pair cabling, such as Category 5E and Category 6 do not meet the 100 meter distance requirement at 10GBASE-T data rates, so 10GbE distance would be limited and would not conform to industry standards.

V. Conclusions

It is important to understand all aspects of your network architecture prior to deciding on the cabling media and design needed to support it. Next generation data centers will be expected to sustain several iterations of technology refreshes that can leverage existing cabling infrastructure if provisioned for at the beginning.

Technologies such as 10GbCEE, 40/100GbE, FCoE, and 10G iSCSI are all dependent on high bandwidth media and structured cabling. ToR network architectures provide a solution adequate for today's standards, but questions regarding growth and change within a data center surround this approach. End-user costs for optical and electronic network equipment as well as cable media type play a role in the selection of a data center's cabling architecture; future technologies are dependent on selecting an approach that can support today's and tomorrow's data center equipment.

Although a "one-size-fits-all" approach is not recommended for all data center connectivity infrastructures, CommScope supports zone-distribution cabling architecture as the best balance to provide for today's capabilities with minimal change required to support technologies in the future.

Key for acronyms

10 Gigabit Ethernet **(10GbE)**
Converged enhanced Ethernet **(CEE)**
Converged network adapter **(CNA)**
Energy efficient Ethernet **(EEE)**
Fibre Channel **(FC)**
Fibre Channel over Ethernet **(FCoE)**
Out of band **(OOB)**
Storage area networks **(SANs)**
Top-of-rack **(ToR)**
Total cost of ownership **(TCO)**
Wake on LAN **(WoL)**

COMMSCOPE®

www.commscope.com

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

04/11