COMMSCOPE®

Designing infrastructure for the cloud

Why having the right physical layer is key to cloud performance

Global cloud IP traffic is expected to continue growing at a CAGR of 30 percent and will account for more than 92 percent of total data center traffic by 2020.[1] With its promise to maximize resources and minimize spending by delivering services over a common platform, cloud computing is becoming a universal model. Having proven its value in the enterprise data center, it is quickly making its way into a variety of industries and applications. Today, cloud-based networks manage everything from how consumers store and access their music and data, to how students on opposite ends of the globe collaborate and learn. It is also being used widely in wireless telecom as operators deploy cloud radio access networks (C-RAN) and network functions virtualization (NFV).

Until now, much of the attention has been rightfully placed on the architecture necessary to enable cloud computing to flourish and grow, including management systems and server platforms. But the fact remains that, beneath the virtualized network, there is a very real cabling infrastructure connecting the physical layer upon which the cloud model depends. How this infrastructure is designed and deployed is critical to the successful implementation of cloud-based applications.

VIRTUALIZATION ALTERS TRADITIONAL NETWORK DESIGN

Cloud computing is based on a high degree of abstraction, with virtual machines not being tied to physical assets. This enables cloud-based networks to easily and quickly scale up or down—simplifying administration.

Surprisingly, the first wide-scale introduction of cloud computing occurred in the 1970s, when IBM released its VM operating system that enabled enterprise administrators to operate multiple virtual machines (VMs) on a single physical node.[2] It would take 25 years and the explosive growth of the Internet before computer virtualization would become the norm. Since then, the swift movement to the cloud has significantly changed network architecture and infrastructure cabling.

> *"Data center network infrastructure is set to undergo major transformation to meet the needs of continuously growing data traffic while supporting the internet of things, the advent of 5G networks, and the influence of Network Function Virtualization."*
>
> **RCR Wireless**
>
> January 19, 2017

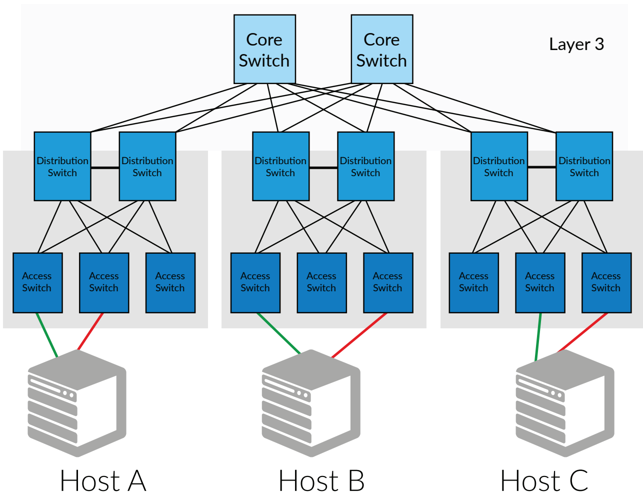

Figure 1: This model shows a traditional three-tier network.

Figure 1 shows a traditional three-tier network consisting of the core, distribution and access switch layers. In this model, traffic flows "north-south" between the core and access layers. It is important to note that the path through the core switches has capacity limits and high latency. In order to add data center capacity, the operator must add more fan-out distribution switches, which usually require higher capacity core switches and routers.

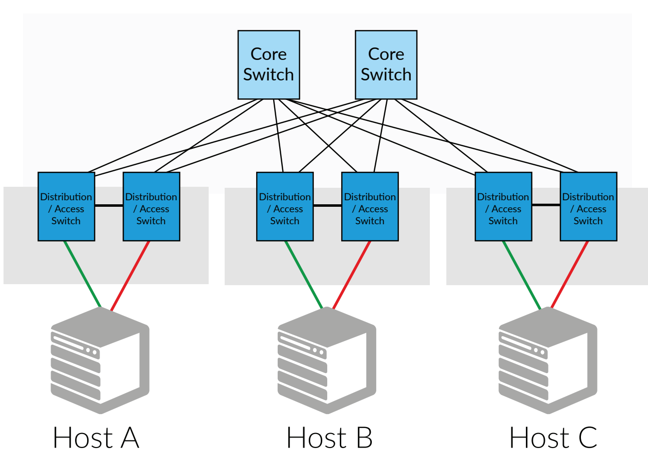

Figure 2: This illustrates a collapsed two-tier network.

The first significant improvement in this design was to merge the access and distribution switch layers, as shown in figure 2. By eliminating the separate physical access layer, the improved two-tier design streamlined the flow of traffic. In addition to reducing CapEx, this also reduces network latency, which, as traffic across the network continues to grow, becomes a more costly issue. Studies at Amazon showed that, for every 100 ms of application latency, sales revenue decreased by 1 percent.[3]

NETWORK DESIGN HEADS EAST-WEST

The implementation of a cloud-based system, however, requires a fundamental shift in our approach to network design. A Cloud environment requires high bandwidth between servers.

In other words, the network's ability to move massive amounts of traffic between distributed servers becomes critical. Whereas the traditional model was optimized for north-south traffic, today's network designs are optimized for east-west traffic flows.

> *"Post-cloud, most software functionality is implemented by large clusters that talk mostly to each other, that is east-west, with only a few tendrils of communication shooting north-south."*
>
> **HighScalability.com**
> September 4, 2012

The IT industry is now aligned on architectures that are more cloud-specific. The leaf-spine model is shown in figure 3. This model utilizes a mesh architecture between the leaf and spine switches.

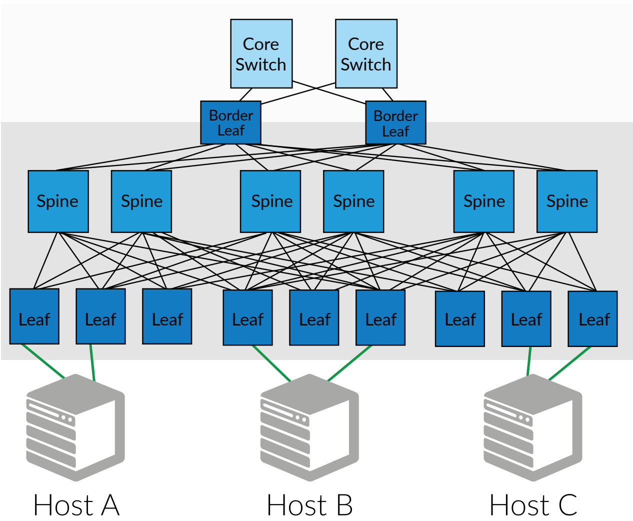

Figure 3: The IT industry is now deploying cloud-friendly leaf-spine networks.

Cloud application architectures requires high bandwidth and low predictable latency between servers. The number of virtual servers per physical host and the average I/O requirements are rising at 10 to 20X per year. This is driving the adoption of 25/50 Gbps Ethernet connectivity to the server.

Moreover, increasing I/O demands from the server require higher uplink capability from the access switches to the Spine Switches. This, in turn, drives the adoption of 100/200/400 Gbps Ethernet ports in the mesh network links. Initially, 100 Gbps uplinks will be required to support 25/40/50 Gbps access capacity to the hosts. Eventually, access capacity requirements to the host will hit 100 Gbps, requiring 200/400/800 Gbps uplinks to the spine switches. All of this has significant implications on the supporting cable infrastructure.

INCREASING DEMANDS ON CABLING INFRASTRUCTURE

High-bandwidth connectivity in the data center has a direct bearing on performance. New 25GbE chips use the same class of cables as 10GbE but switching ASICs deliver 2.5 times the performance of the previous 10G lanes.[4]

Driven by increasing I/O demands, cabling suppliers like CommScope are recommending fiber cabling in the data center. Current speeds, depending on distances, are demanding SM and/ or OM4-5 cabling in either serial or parallel form depending on the fiber transmission methods being used. This is because 25/40/50 and 100/200/400 Gbps will be delivered across a mix of parallel and serial fiber-optic links, for which preterminated fiber solutions offer the fastest and easiest deployment. In addition to the proper MPO-based connectorization, preterminated fiber will dramatically reduce installation time while allowing for easy and rapid redeployment of fiber infrastructure when reconfiguring or refreshing the data center layout.

Suppliers also recommend data center cabling move from category 6 to category 6A. Current-generation 10GBASE-T PHYs have significantly improved latency performance, which is low enough to support all but a very few applications.

TOWARD A HIGH-AVAILABILITY, LOW-LATENCY FUTURE

The 2016 Cloud Computing Survey by IDG illustrates the degree to which cloud-based computing has transitioned from being an emerging technology to a mainstream powerhouse going beyond basic storage. 70 percent of all enterprises surveyed are using the cloud for at least one cloud-based application or a portion of their IT infrastructure, with the average company having deployed 45 percent of its IT environment in the cloud. Cloud budgets, which average 28 percent of the total IT. Preferred cloud environments include private cloud (28 percent), public cloud (22 percent) and hybrid cloud (10 percent).[5]

Driving the move to cloud-based computing is the ever-growing increase in user expectations. Whether it's a consumer downloading music to their home computer or an enterprise user accessing an application on their virtual desktop, speed and availability are what matter. To ensure application performance meets user expectations, both network availability and latency must continue to improve. At the physical layer, this means implementing infrastructure solutions and cabling standards capable of providing higher bandwidth and lower latency without jeopardizing availability.

At the same time, network management must evolve as well. As the infrastructure grows more virtualized, the need for automated management tools and better visibility of the physical infrastructure will increase. According to forecasts from the International Data Corporation (IDC) total spending on cloud infrastructure will increase by 18.2% in 2017 to reach $44.2 billion.[6] As the industry continues to ramp up deployment of cloud-based applications, that number is expected to grow significantly by 2020. Of course, the type and level of performance we'll see by then is anybody's guess.

References

1 Source: Cisco Global Cloud Index: Forecast and Methodology, 2015–2020; Cisco Public; November 2016

2 A Brief History of Cloud Computing; IBM SoftLayer Blog; July 29, 2013

3 Latency is Everywhere and it Costs You Sales—How to Crush It; Highscalability.com; July 25, 2009

4 Hadoop* Clusters Built on 10 Gigabit Ethernet; Intel white paper; 2012

5 2016 Cloud Computing Survey, IDG, November 1, 2016

6 Growth In Cloud IT Infrastructure Spending Will Accelerate In 2017; Cloud Strategy Magazine; January 16, 2017

COMMSCOPE®

WP-108064.3-EN (05/18)