The Next Generation of Wide Area Networking

Introduction

As pointed out in *The 2014 State of the WAN Report*[1], the vast majority of WAN traffic currently uses either the Internet or MPLS. Since the Internet came into being in the early 1970s and MPLS services have been in place for over a decade, it's easy to conclude that neither the Internet nor MPLS are new services. Looked at from just this vantage point it's also possible to conclude that little if any change is occurring in the WAN.

While there haven't been any fundamentally new WAN services introduced in the last decade, there have been significant changes during that time period both in terms of the role of the WAN and the added functionality that IT organizations have deployed on top of WAN services such as the Internet and MPLS. For example, a decade ago the primary role of the WAN was to enable a company's branch office employees to access applications that were housed at the company's corporate data center(s) and to provide backup and Disaster Recovery (DR) for the applications and data stored in those data centers.

However, accessing applications over a WAN has always presented challenges. Beginning around 2006 many IT organizations began to deploy hardware-based WAN Optimization Controllers (WOCs). These WOCs provided functionality that was incremental to the functionality provided by the Internet and MPLS services and they were designed to overcome the first generation of application delivery challenges[2], such as supporting chatty protocols. While IT organizations continue to deploy hardware-based WOCs, for the last few years it has become increasingly common for IT organizations to implement software-based WOC functionality.

The way that IT typically evolves is that as new ways of doing business become common, they don't completely eliminate the traditional ways. As a result, IT organizations need to simultaneously support both the emerging and the traditional approaches. What this means relative to the WAN is that IT organizations need to still ensure that the WAN can fulfill its traditional role of enabling a company's branch office employees to access applications housed at the company's data center(s). However, at the same time, IT organizations also need to ensure that the WAN can fulfill its emerging role of enabling a company's employees to access applications and services provided by either a Software-as-a-Service (SaaS) or an Infrastructure-as-a-Service (IaaS) provider.

One goal of this white paper is to identify some of the challenges that are associated with ensuring the acceptable delivery of applications and services acquired from either a SaaS or an IaaS provider. Another goal of this white paper is to help IT organizations understand how emerging functionality enables them to overcome these challenges.

---

[1] http://www.ashtonmetzler.com/Metzler-WAN-Survey-2014.pdf

[2] The first generation of application delivery challenges is described in *Traditional Application and Service Delivery Challenges* http://www.ashtonmetzler.com/Traditional%20App%20Delivery%20Challenges%20V2.0.pdf

The Growing Use of Cloud Computing

There are three primary classes of cloud computing:

- *Public cloud* focuses on organizations acquiring applications and services from SaaS and IaaS providers.
- *Private cloud* focuses on companies internally implementing the same techniques that IaaS and SaaS providers deliver; e.g., virtualization, automation and self-service.
- *Hybrid cloud* computing focuses on creating applications and services that are comprised of components from both private and public clouds.

IDC recently quantified the dramatic growth in the use of public cloud services. According to IDC[3], worldwide spending on public cloud services is expected to be more than $107 billion in 2017. IDC also stated that over the 2013–2017 forecast period, public cloud services will have a compound annual growth rate (CAGR) of 23.5%, five times that of the IT industry as a whole. The dramatic growth in the use of public cloud services highlights the fact that the role of the data center is fundamentally changing.

The growing use of public cloud services is having a very significant impact on a company's WAN. One impact is that accessing public cloud applications and services is seen by IT organizations as the biggest driver of increased Internet traffic (see Table 1).

Application	Biggest Driver	Second Biggest Driver	Combination of Biggest and Second Biggest
Public Cloud Applications and Services	35.7%	20.8%	56.5%
Support for Mobile Users	22.7%	21.3%	44.0%
Enterprise Applications; e.g., CRM, SCM and ERP	11.1%	14.0%	25.1%
Video	10.6%	8.7%	19.3%
Disaster Recovery/Business Continuity	9.2%	7.7%	16.9%
Virtual Desktops	2.4%	8.7%	11.1%
Voice	4.3%	2.4%	6.7%
Don't Know/Not Applicable	3.9%	13.5%	6.7%

[3] http://www.idc.com/getdoc.jsp?containerId=prUS24298013

Movement of Virtual Machines between Data Centers	0.0%	2.9%	2.9%

Table 1: Drivers of Increased Internet Traffic Source: 2014 State of the WAN Report

The growing adoption of public cloud services is having more of an impact on the WAN than just increased traffic. More than a third of IT managers indicated that the adoption of public cloud services had caused them to re-evaluate their WAN strategy[4].

The Evolving WAN

Figure 1 shows the design of today's typical enterprise WAN. The WAN shown in Figure 1 supports the traditional role of the WAN; e.g., enabling a company's branch office employees to access applications at the company's data center(s) and providing for DR for the applications and data stored in those data centers.

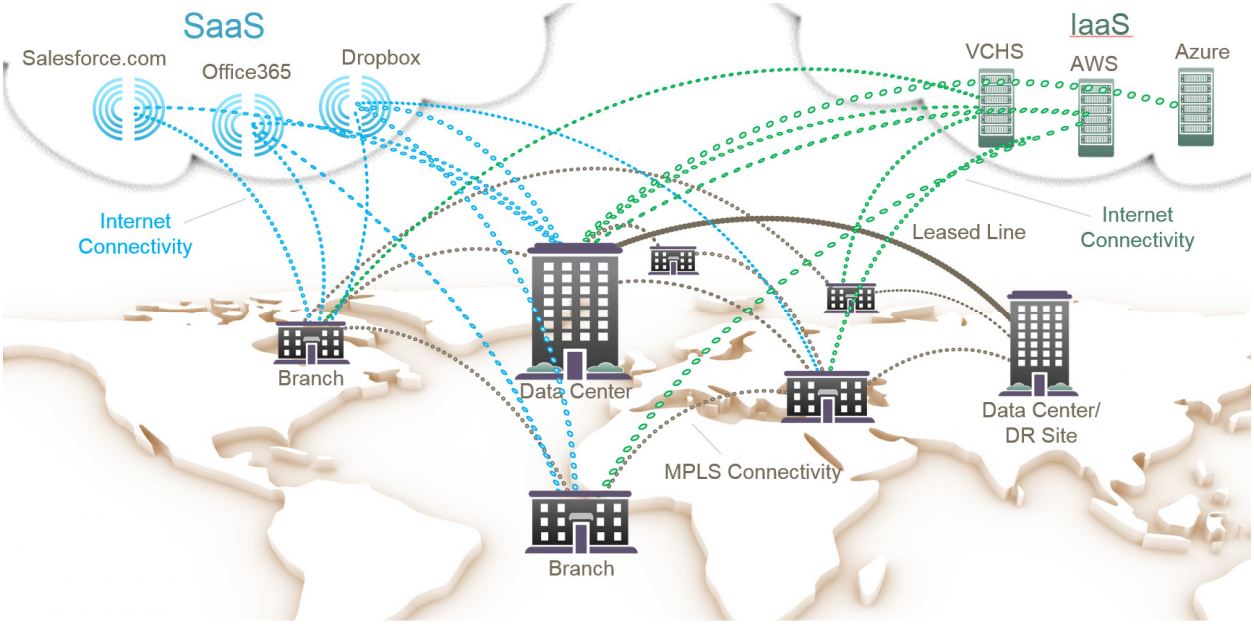

Figure 1: Today's Typical WAN Design

The WAN in Figure 1 also depicts how the growing use of cloud services has made the WAN more complex. One example of that complexity is that today IT organizations usually backhaul their Internet traffic on the organization's enterprise network (e.g., their MPLS network) to a central site where that traffic is then handed off to the Internet. The advantage of this approach is that it enables IT organizations to exert more control over their Internet traffic and it simplifies management in part because it centralizes the complexity of implementing and managing additional security policies. One disadvantage of this approach is that it results in extra traffic

[4] http://reports.informationweek.com/abstract/5/11755/Cloud-Computing/2014-Next-Gen-WAN-Survey.html

transiting the enterprise WAN which adds to the cost of the WAN. Another disadvantage of this approach is that it adds additional delay to the Internet traffic. While most Internet traffic is backhauled, in some cases IT organizations route traffic from branch offices to cloud services directly over the Internet. This approach creates visibility and control gaps for IT managers and it also creates sporadic performance issues due to the characteristics of the Internet that are described below.

Second Generation Application Delivery Challenges

There are three key functions associated with ensuring acceptable application delivery, which apply not only to enterprise-hosted applications, but also to the growing universe of public cloud applications. Those functions are:

- Management

  Effective management requires the ability to both understand what is happening on the network and to be able to control the network by performing tasks such as giving priority to business critical applications.

- Optimization

  There is a wide range of optimization functionality associated with acceptable application delivery. This includes the ability to significantly reduce the amount of data that is transmitted over the WAN and the ability to mitigate the effect of latency, congestion and packet loss.

- Security

  There are many components of an effective security strategy, including the ability to encrypt the traffic. As Internet use increases this functionality must be applied universally so that data in-flight is never visible nor compromised.

Most of the second generation of application delivery challenges are associated with the increased use of cloud computing. This includes the dramatic increase in the volume of Internet traffic and the growing negative impact of backhauling Internet traffic. There are, however, a number of other characteristics of the Internet in general, and of SaaS and IaaS services in particular, that create some additional challenges. Some of the characteristics of the Internet that create additional application delivery challenges include:

- Characteristics of TCP

  The TCP (Transmission Control Protocol) slow start algorithm is part of the TCP congestion control strategy and it calls for the initial data transfer between two communicating devices to be severely constrained. The algorithm calls for the data transfer rate to increase if there are no problems with the communications, which would

rarely be the case when accessing SaaS or IaaS applications applications over an Internet connection. In addition to the initial communications between two devices, or between a device and the cloud service, the slow start algorithm is also applied in situations where a packet is dropped.

- Internet Latency

  The latency of the Internet tends to be larger than the latency associated with an enterprise WAN, such as MPLS. One of the reasons for that is the use of the Border Gateway Protocol (BGP) within the Internet. BGP is an exterior routing protocol and is used within the Internet to route the traffic that goes between the millions of networks that comprise the Internet. However, since BGP doesn't know the specific, real-time characteristics of the networks that comprise the Internet, the end-to-end path chosen by BGP isn't likely to be the path with the least latency.

- Packet Loss

  Part of the reason why the Internet has high packet loss is the Internet is a "network of networks" that consists of millions of private and public, academic, business, and government networks and there isn't an organization that has control of the end-to-end design or performance of the Internet. Another factor that drives high Internet packet loss is there isn't any direct financial or governance incentive for Internet Service Providers (ISPs) to design peering points for low packet loss, and so they often don't.

  When there is packet loss, or when packets are delivered out-of-order, TCP will re-transmit packets and the TCP slow start algorithm (see above) assumes that the loss is due to congestion and takes steps to reduce the offered load on the network. These actions can dramatically reduce throughput and result in delayed screen refreshes and lengthy file transfers. Unfortunately, similar to BGP, OSPF does not make routing decision based on issues that impact throughput such as high levels of packets either being dropped or delivered out-of-order.

- Availability

  Within a single network, an interior routing protocol such as OSPF is used to communicate IP address ranges to all the routers within the network. These routing protocols can detect a network link failure and update the routing table on all routers within a few seconds. As previously discussed, BGP is used to exchange information between the millions of networks that comprise the Internet. The size and complexity of the Internet as well as the inherent characteristics of BGP mean that a failed network link and the resulting routing path change may take several minutes before all routing tables are updated.

An organizational challenge associated with using the Internet to access SaaS and IaaS services comes from the fact that these services typically share TCP port 443 with other Internet traffic. Since most monitoring systems cannot differentiate between the traffic types that share port 443, in many cases, the IT organization isn't even aware that the cloud services that are being used.

In the case of storage-as-a-service applications or broader IaaS use cases where the enterprise might encounter the need for more capacity, IT organizations can work with the cloud providers to add more processing and/or storage capacity. They can also increase the bandwidth of their Internet connections. However, the effectiveness of adding Internet bandwidth is limited by basic TCP operations including the TCP window size and the previously discussed TCP slow start algorithm. The result of combining these factors, commonly referred to as TCP's "long fat pipe problem[5]," is that as the length of the Internet connection increases, the network throughput decreases.

One of the reasons why the growing use of SaaS and IaaS services exacerbates the long fat pipe problem is because in many cases these services are often delivered from data centers far away from the users. One such example of this phenomena involves a European company that subscribes to a SaaS service in California. In most cases, if the SaaS provider does not have a data center in Europe, the data for the company's European users will be located in or near California, not in Europe. Further complicating this situation is the fact that in many instances, IT organizations are not aware of where the company's data is being stored or where the SaaS data center is located.

Another reason why the growing use of public cloud services exacerbates the long fat pipe problem is because the use of these services can increase the amount of traffic that transits the network. For example, SaaS providers typically rely on excessive use of Cascading Style Sheets[6] and JavaScript to provide a "desktop-like" experience. Doing so often means downloading large (~5 Megabyte) JavaScript files, which takes an inordinate amount of time over long Internet connections. The result of these factors is that users of public cloud services are often forced to wait for files to transfer and screens to refresh.

In addition to potentially increasing the amount of traffic that transits the network, IaaS services create some additional networking challenges. One example of these additional challenges is that neither Amazon Web Services (AWS) nor Microsoft Azure support multicast or broadcast traffic. Another example is that it is typically very difficult to set up a VPN between a company's firewall and an IaaS provider's Virtual Private Gateway (VPG). A third example is that in most cases if an IT organization wants to move data between two IaaS providers such as VMware vCloud and AWS, or from AWS to Azure, it must be done in two steps. The first step is to move the data from the originating IaaS provider to a data center belonging to the IT organization. The second step is to move the data from the IT organization's data center to the receiving IaaS provider. Having two steps adds time to the overall process and it usually adds cost as each IaaS provider is likely to charge for data transfer.

---

[5] http://flylib.com/books/en/3.223.1.244/1/

[6] http://en.wikipedia.org/wiki/Cascading_Style_Sheets

Requirements for the Next Generation WAN

To address the application delivery challenges associated with cloud computing, a new service delivery fabric must be implemented that secures, manages and optimizes data delivery. This fabric must create an overlay that integrates existing investments in the enterprise WAN infrastructure, as well as bring Internet connectivity and cloud services under the control of the IT organization.

As previously discussed, there are three categories of functionality that are closely associated with ensuring acceptable application delivery: Management, Optimization and Security. Below is a discussion of the additional functionality that is required within each of those categories to address the new cloud-driven WAN.

Rethinking WAN Routing

The dynamic nature of a cloud-driven WAN is forcing IT organization to rethink enterprise routing. As mentioned earlier, traditional routing protocols fall into two classes: interior routing for finding routes within a single network, and exterior routing for passing routing information between networks. As pointed out earlier, older routing protocols typically rely on manually configured metrics to find the best end-to-end path or to find the best peering point.

The growing use of SaaS and IaaS services is forcing enterprises into a situation where they are using multiple WAN services, typically an MPLS service and Internet access from one or more ISPs. The traditional static methods for determining the best path no longer work well in an environment where a growing percentage of the WAN traffic transits the Internet.

Addressing these issues requires new interior and exterior routing techniques that use real time measurements of latency, loss and other path metrics to determine the best path or paths to use for a particular service or application. To determine the best interior paths, measurements should be made continuously and paths must be selected based on a tailored objective for each traffic class, such as lowest latency or highest throughput. For example, VoIP and video may be routed over paths on the MPLS network and bulk data may be routed over Internet paths, with fail-over in either direction in the event of a service disruption or changing conditions. This capability also allows enterprises to leverage multiple network connections to a particular destination as if it were a single large network.

To determine the best exterior path, or transit gateway for an Internet hosted service such as a SaaS service like Salesforce.com or Office 365, techniques must be used to measure the loss, latency and other metrics from each potential gateway to the service. This information should complement software strategically placed at optimal egress points across the enterprise and within cloud hubs (i.e., Equinix, etc.) or IaaS (i.e, AWS, Azure) in order to not only perform the optimal transit routing to SaaS applications, but to also distribute information about adverse Internet conditions to other software instances across the network. This allows for an optimal end-to-end path to be determined for each user accessing each service on a completely dynamic basis.

However, in order to achieve new levels of optimal WAN routing and managing of cloud traffic, visibility is required into the ever-changing SaaS services and Internet weather. New SaaS services are introduced frequently and service components and data centers are regularly added to existing SaaS services. Some level of cloud intelligence is required to aggregate information about the changing subnets and IP addresses for SaaS services so that the software functions that perform the advanced routing have the most current information about those services.

Extending Optimization to the Cloud

For the reasons mentioned, routing traffic in new ways is essential. However, when dealing with Internet bandwidth and quality challenges, optimizing the connections themselves becomes equally, if not more important. WAN de-duplication and compression are needed to maximize the available Internet bandwidth. In particular, traffic must be inspected constantly and repetitive transmissions of duplicate data must be eliminated. This functionality must work across all IP applications and protocols, including SaaS applications, and it must be done at the byte-level so that all duplicate data can be found and processed.

As referenced earlier, the Internet often experiences high levels of latency as well as dropped and out-of-order packets. Path conditioning must be extended to the Internet, where techniques such as Forward Error Correction can reconstitute dropped packets and Packet Order Correction can re-sequence packets that traverse multiple WAN paths.

Finally, traditional traffic shaping must also be applied to the growing volume of cloud-driven traffic. This enables SaaS applications to be classified appropriately, and it enables IT organizations to prioritize critical traffic classes (e.g. Salesforce.com traffic) while eliminating or constraining traffic classes that typically include recreational use of the Internet (e.g. YouTube).

Building a Secure Fabric

Building out a new WAN fabric to address enterprise and cloud applications must be done using high performance edge-to-edge encryption based on IPsec VPN technology. However, most IPsec VPN implementations suffer from poor performance and have often been complicated and inflexible. New approaches, however, that have recently entered the market offer higher performance and simplified and automated enablement of encryption for all traffic in a consistent manner, spanning the enterprise locations and cloud infrastructure.

Delivering on the New Approach

Silver Peak has introduced a new solution onto the market that delivers on the promise of next-generation wide area networking. Silver Peak's new Unity WAN fabric (Figure 2) brings together the traditional enterprise WAN, the Internet and cloud services onto a single fabric. With Silver Peak Unity, the enterprise now becomes SaaS-aware and SaaS-optimized.

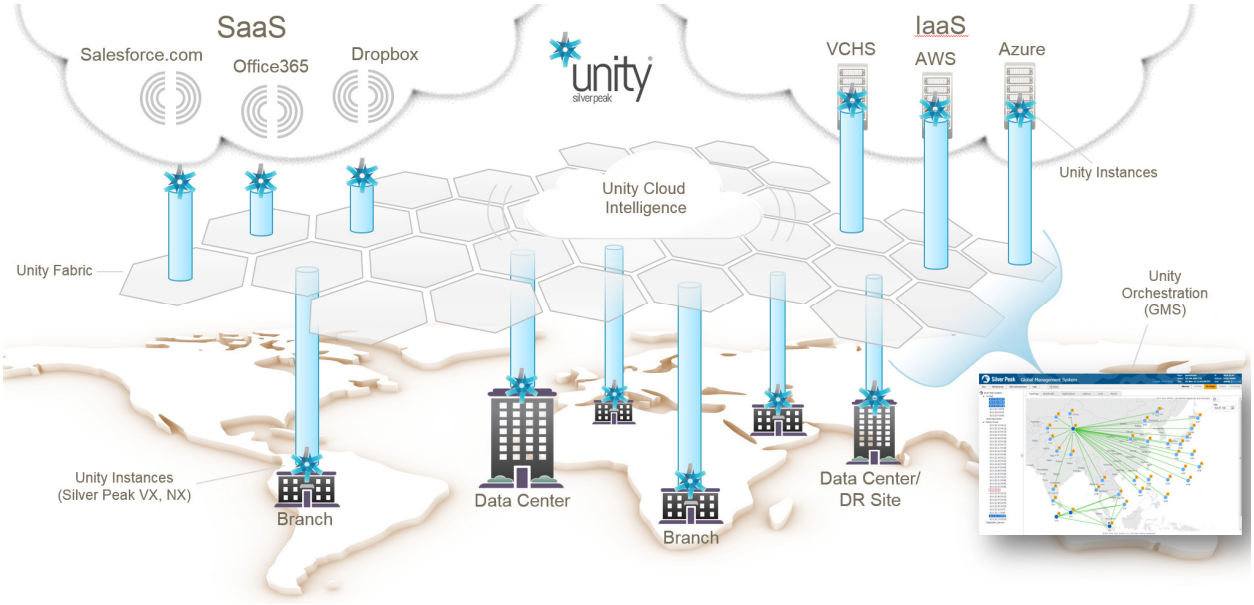

Figure 2: Silver Peak's Unity Fabric

The Silver Peak Unity fabric is built by installing Silver Peak software in data centers, branch offices and cloud interconnection hubs. It is a network overlay that can control and accelerate connectivity to any combination of enterprise services, IaaS resources and SaaS applications. Each Silver Peak instance on the Unity fabric communicates with Silver Peak's new Cloud Intelligence service, which aggregates constantly changing information about cloud providers and Internet weather. Unity uses this information, along with calculations from each software instance, to dynamically route traffic to the cloud over the optimal path. Orchestration is handled by Silver Peak's Global Management System (GMS), which provides IT managers with complete visibility and control over the deployment and use of cloud services.

For existing Silver Peak customers, building a Unity fabric is fairly straight-forward. They need to upgrade to release 7 of Silver Peak's software and subscribe to Silver Peak's Cloud Intelligence service. Any new customers simply purchase Silver Peak software for their data center or remote offices, and subscribe to the Cloud Intelligence service. Enterprises can also expand their Unity fabric by adding Silver Peak instances in cloud hubs or IaaS providers like AWS or VMware's vCloud.

Summary and Call to Action

In the traditional IT environment, users access applications and services that are housed in a corporate data center often using an enterprise WAN service such as MPLS. However, the dramatically increasing adoption of cloud computing means that the role of the data center is changing and that increasingly users will access applications and services both from a corporate data center as well as from a growing set of public cloud service providers.

The changing role of the data center is driving IT organizations to rethink their WAN strategy. As part of their new WAN strategy, IT organizations need to eliminate aspects of the traditional

WAN design, such as the backhauling of Internet traffic, that increase the complexity of the WAN. IT organizations also need to overcome the traditional performance impairments associated with the use of the Internet (i.e., latency, packet loss, low availability, out of order packets, reduced throughput) as well as the specific challenges associated with accessing SaaS or IaaS services; i.e., the long fat pipe problem, the lack of easy to use, sophisticated network functionality on the part of public cloud providers.

The management component of the next generation WAN must include new interior and exterior routing techniques that use real time measurements of latency, loss and other path metrics to determine the best path or paths to use for a particular service or application. It must also provide visibility into the ever-changing SaaS services and minute-to-minute Internet weather. The optimization component of the next generation WAN must include functionality such as de-duplication, compression, path conditioning and traffic shaping. The security component of the next generation WAN must include edge-to-edge encryption based on IPsec VPN technology that is both easy to implement and which provides high levels of performance.

Silver Peak has introduced a new solution onto the market that delivers on the promise of next-generation wide area networking. Silver Peak's new Unity WAN fabric brings together the traditional enterprise WAN, the Internet and cloud services onto a single fabric. For more information on Unity, go to http://www.silver-peak.com/products-solutions/unity.