The 2012 Cloud Networking Report

Part 4: The Wide Area Network (WAN)

By Dr. Jim Metzler
Ashton Metzler & Associates
Distinguished Research Fellow and Co-Founder
Webtorials Analyst Division

Platinum Sponsors:

Gold Sponsors:

Produced by:

Table of Contents

Executive Summary

The **2012 Cloud Networking Report** (The Report) will be published both in its entirety and in a serial fashion. This is the fourth of the serial publications. The first publication in the series described the changes that are occurring in terms of how cloud computing is being adopted, with a focus on how those changes are impacting networking. The second publication in the series focused on data center LANs. The third publication discussed Software Defined Networks (SDNs) and included the results of a survey that was done in conjunction with Information Week. The focus of this publication is Wide Area Networking.

The focus of the next publication of The Report will be security and management. The Report will then be published in its entirety and there will be a separate executive summary that covers the totality of The Report.

This section of The Report includes the results of surveys that were recently given to the subscribers of Webtorials.com. Throughout this report, the IT professionals who responded to those surveys will be referred to as the ***Survey Respondents***. In some cases, the results of the surveys given to the Survey Respondents will be compared to the results of surveys given in 2011.

The Wide Area Network (WAN)

Introduction

Background

The modern WAN got its start in 1969 with the deployment of ARPANET which was the precursor to today's Internet. The technology used to build the Internet began to be commercialized in the early 1970s with the development of X.25 based packet switched networks. While the early use of the Internet was strictly for academic and research purposes, the use of the Internet for commercial purposes started in the early 1990s with the development of the World Wide Web.

In addition to the continued evolution of the Internet, the twenty-year period that began in 1985 saw the deployment of four distinct generations of enterprise or private WAN technologies[1]. For example, in the mid to late 1980s, it became common for enterprise IT organizations to deploy integrated TDM-based WANs to carry both voice and data traffic. In the early 1990s, IT organizations began to deploy Frame Relay-based WANs. In the mid to late 1990s, some IT organizations replaced their Frame Relay-based WANs with WANs based on ATM (Asynchronous Transfer Mode) technology. In the 2000s, many IT organizations replaced their Frame Relay or ATM-based WANs with WANs based on MPLS. Cost savings was the primary factor that drove the adoption of each of the four generations of WAN technologies. The cost savings, however, were very modest when compared to the price performance improvements that are associated with local area networking that were discussed in a preceding section of The Report.

However, in contrast to the volatility of this twenty-five year period:

Today there is not a fundamentally new generation of technology under development that is focused on the WAN.

Relative to the deployment of new WAN services, what sometimes happens in the current environment is that variations are made to existing WAN technologies and services. An example of that phenomenon is Virtual Private LAN Service (VPLS)[2]. As described later in this section of the report, within VPLS an Ethernet frame is encapsulated inside of MPLS. While creating variations on existing services can result in significant benefits, it does not produce fundamentally new WAN services.

[1] An enterprise or private WAN is designed to provide for connectivity primarily within the enterprise and between the enterprise and key contacts such as partners. This is in contrast to the Internet that is designed to provide universal connectivity.
[2] http://vlt.me/vpls-0810

Contrasting the LAN and the WAN

The WAN is notably different than the data center LAN. These differences include the fact that:

- After a lengthy period in which there was little if any fundamental innovation, the LAN is experiencing broad fundamental change. In contrast, after a lengthy period in which the WAN underwent repeated fundamental change, there are currently no fundamentally new WAN specific technologies under development.

- While there are no fundamentally new WAN specific technologies under development, there are new WAN architectures being developed and implemented.

- In the vast majority of instances, the latency, jitter and packet loss that the LAN exhibits doesn't have an appreciable impact on application performance. In many instances, the latency, jitter and packet loss that public and private WANs exhibit has an appreciable impact on application performance. This is particularly true of 3G/4G networks.

- One of the primary design criteria for designing a data center LAN is scalability. A manifestation of the ongoing improvements in LAN scalability is that over the last fifteen years the speed of a data center LAN has increased from 10 Mbps to 10 Gbps – which is a factor of a thousand. In contrast, in many cases the primary design criterion for designing a WAN is to minimize cost. For example, in many parts of the world it is possible to get high-speed WAN links such as an OC-192 link. These links, however, are usually not affordable.

- The LAN follows Moore's Law. In contrast, the price/performance of enterprise WAN services such as MPLS doesn't come close to doubling every two years.

The WAN doesn't follow Moore's Law.

WAN Budgets

Both in 2011 and again in 2012, The Webtorials Respondents were asked how their budget for the forthcoming year for all WAN services compares to what it is in the current year. Their responses are contained in **Table 1**.

Table 1: WAN Budget Increases		
	Responses in 2011	**Responses in 2012**
Reduced by more than 10%	3.2%	7.2%
Reduced by 1% to 10%	11.1%	18.5%
Basically static	34.9%	40.5%
Increased by 1% to 10%	32.8%	21.2%
Increased by more than 10%	18.0%	12.6%

The change in the budget for WAN services in 2012 is notably different than what the corresponding change was in 2011. For example, in 2011 half of the ***Survey Respondents*** indicated that their WAN budget was increasing while in 2012, only a third did.

WAN budgets are notably more constrained than they were a year ago.

As is explained in the next subsection, the adoption of cloud computing will increase the rate of growth in the amount of traffic that transits the WAN. As such:

IT organizations need to make changes relative to how they use WAN services in order to support a significant increase in WAN traffic while experiencing a highly constrained WAN budget.

Drivers of Change

As mentioned in the section of this report entitled *The Emergence of Cloud Computing and Cloud Networking,* one of the characteristics of cloud computing is increased reliance on the network. The increased reliance on the WAN in particular stems from the fact that the resources that support cloud computing solutions are centralized in a small number of data centers and the vast majority of users access these solutions over the WAN. Hence, the more use that organizations make of cloud computing in general, the more traffic transits both the Internet and enterprise WAN services. When looking just at public or hybrid cloud services, the adoption of these services primarily results in more traffic transitting the Internet.

Below are some of the specific factors that are putting more traffic onto the WAN and hence, driving the need for IT organizations to change their approach to wide area networking.

Virtual Machine Migration

The section of this report entitled *The Emerging Data Center LAN* quantified the great interest that IT organizations have in server virtualization in general and in moving virtual machines (VMs) between data centers in particular. That section of the report also discussed the fact that one of the requirements associated with moving VMs between data centers is that the data storage location, including the boot device used by the VM being migrated, must be accessible

by both the source and destination physical servers at all times. If the servers are at two distinct locations and the data is replicated at the second site, the two data sets must be identical. One approach to enabling data access is to extend the SAN to the two sites and to maintain a single data source. Another option is migrate the data along with the VM to the secondary site. In either case, it is necessary to coordinate VM and storage migrations and to be able to move large data sets efficiently between data centers, which will have a significant impact on the WAN.

Virtual Desktops

Another form of virtualization that will drive a further increase in WAN traffic is desktop virtualization. In order to quantify the interest that IT organizations have in desktop virtualization, The ***Survey Respondents*** were asked to indicate the percentage of their company's desktops that have either already been virtualized or that they expected would be virtualized within the next year. Their responses are shown in **Table 2**.

Table 2: Deployment of Virtualized Desktops

	None	1% to 25%	26% to 50%	51% to 75%	76% to 100%
Have already been virtualized	44%	49%	6%	1%	0%
Expect to be virtualized within a year	24%	53%	20%	1%	1%

The data in **Table 2** indicates the growing interest that IT organizations have in desktop virtualization. For example:

> ***Over the next year, the percentage of IT organizations that have not implemented any desktop virtualization will be cut roughly in half.***

Part of the challenge in supporting virtualized desktops is that the implementation of virtualized desktops puts more traffic on the WAN, which typically leads to the need for more bandwidth. In addition to the bandwidth challenges, as explained in The 2012 Application and Service Delivery Handbook[3], there are performance challenges associated with each of the two primary form of desktop virtualization; e.g., client side (a.k.a., streamed desktops) and server side (a.k.a., hosted desktops).

Collaboration

As was described in the section of this report that is entitled *The Emergence of Cloud Computing and Cloud Networking,* many organizations are beginning to acquire services such as collaboration from a cloud computing service provider (CCSP). Independent of whether the collaboration service is provided by a CCSP or by the IT organization, it stresses the WAN. This stress comes in part from the fact that the performance of applications such as video and telepresence is very sensitive to delay, jitter and packet loss. The stress also comes in part because video and telepresence consume considerable WAN bandwidth. It is common, for example, to allocate several megabits per second of WAN bandwidth to a single telepresence session.

[3] http://www.webtorials.com/content/2012/07/2012-application-service-delivery-handbook-1.html

Mobile Workers

In the last few years there has been an explosive growth in the number of mobile workers. There are a number of concerns relative to supporting mobile workers. One such concern is that up through 2010, the most common device used by a mobile worker was a PC. In 2011, however, more tablets and smartphones shipped than PCs[4]. Related to the dramatic shift in the number and types of mobile devices that are being shipped, many companies have adopted the BYOD (Bring Your Own Device to Work) concept whereby employees use their own devices to access applications.

The ***Survey Respondents*** were asked to indicate the types of employee owned devices that their organization allows to connect to their branch office networks and which of these devices is actively supported, Their responses are shown in **Table 3**.

Table 3: Support for Employee Owned Devices

	Not Allowed	Allowed but not Supported	Allowed and Supported
Company managed, employee owned laptop	22%	24%	54%
Employee owned and managed laptop	38%	38%	25%
Blackberry	17%	24%	58%
Apple iPhone	14%	30%	55%
Android phone	19%	33%	48%
Windows mobile phone	26%	40%	34%
Apple iPad	18%	40%	52%
Android based tablet	28%	37%	35%
Windows based tablet	28%	36%	37%

The data in **Table 3** indicates that there is wide acceptance BYOD. In particular:

IT organizations are required to support a wide range of end user devices.

As a result of the movement to adopt BYOD, the typical branch office network now contains three types of end user devices that are all accessing business critical applications and services. This includes PCs as well as the new generation of mobile devices; i.e., smartphones and tablet computers. Because of their small size, this new generation of mobile devices doesn't usually have wired Ethernet ports and so they are typically connected via what is hopefully a secure WiFi network in the branch office or a 3G/4G service when WiFi isn't available.

Another key concern relative to supporting mobile workers is how the applications that these workers access has changed. At one time, mobile workers tended to primarily access either recreational applications or applications that were not delay sensitive; e.g., email. However, in the current environment mobile workers also need to access a wide range of business critical applications, many of which are delay sensitive. This shift in the applications accessed by mobile

[4] http://gizmodo.com/5882172/the-world-now-buys-more-smartphones-than-computers

workers was highlighted by SAP's announcement[5] that it will leverage its Sybase acquisition to offer access to its business applications to mobile workers.

One of the technical issues associated with supporting mobile workers' access to delay sensitive, business critical applications is that because of the way that TCP functions, even the small amount of packet loss that is often associated with wireless networks results in a dramatic reduction in throughput. A related issue is that typically there is a large amount of delay associated with 3G and 4G networks.

WAN Requirements

This subsection of The Report will summarize some of the emerging requirements that WANs must satisfy on a going forward basis. For example, whether providing connectivity from the branch office to the corporate data center or from the branch office to a CCSP's facility, the WAN must be able to prioritize applications in accordance with business priorities. MPLS provides built-in Class-of-Service (CoS), but the Internet does not have that type of capability. Nevertheless, in order to meet business demands the network must be able to examine, recognize and classify network traffic is a way that reflects business priorities and not just the network protocol or TCP/UDP port numbers. The ability to recognize and differentiate different applications on the network requires Deep Packet Inspection (DPI), network fingerprinting and pattern matching. Once the network traffic has been recognized, it must be placed into queues that reflect the different Quality of Service (QoS) demands that are associated with the varying traffic types and business priorities. As applications on the network change, the network must recognize these changes and adapt to them. This capability is needed whether MPLS or the Internet is being used and whether the branch office user is accessing resources in the corporate data center or in a CCSP's facilities.

Given the complexity of contemporary applications combined with the impact of poorly performing business critical applications, it is critical that IT organizations have visibility into each component of the network in order to understand both the underlying cause of poor application performance and to identify what is needed to remedy the problem and restore full application performance. In order to have this visibility, IT organizations must be able to determine what path or paths between the end users and the applications are associated with degraded application performance. IT organizations must also have the capability to isolate the performance problem to a particular network segment and then be able to use effective diagnostic tools to rapidly determine the root cause of the degradation.

As previously mentioned, in the current environment some end users are mobile and some reside in branch offices. In addition, some applications are hosted in a corporate data center and some at a CCSP's facility. These two factors mean that visibility must be provided across MPLS networks as well as across the Internet and that the visibility must extend from the end user to both the corporate data center and to the CCSP's facilities. In addition, performance data needs to be collected from various points in the network in order to create a comprehensive view of the end-to-end performance. Performance data can be collected from routers and/or dedicated appliances. Routers and dedicated appliances typically send performance data to collection and reporting systems that analyze the data, frequently using Flexible Netflow or the IETF IP Flow Information Export; i.e., IPFIX, RFC 5101, 5102 and 5103.

[5] Wall Street Journal, May 17, 2011, page B7

Traditional routing protocols can be used to reroute traffic when network segments fail, but they are not capable of identifying the optimal or best path for network traffic. In addition, when a network link is overloaded routing protocols continue to send traffic to it even though a better performing path is available. These limitations of routing protocols often result in unnecessary application degradation. With the growing adoption of cloud computing and the increasing reliance on the Internet to provide connectivity from branch offices to facilities provided by CSPs, it is critical that the best possible path (e.g., some combination of low delay, jitter and packet loss) from the users to the applications be identified in order to improve application performance. Identifying the best path from point A to point B enables IT organizations to place business critical applications on the path that exhibits the best performance while the remaining applications transit a different path.

One of the implications of the growing adoption of cloud computing is that the network must be able to dynamically provision secure and reliable connectivity between branch offices and a CCSP's facilities to support functions such as cloud bursting and failover/disaster recovery between a corporate data center and a CSP. While some applications use SSL encryption, for those that don't, network level encryption is typically needed and in many cases it is required by regulations or industry standards; e.g. HIPPA and PCI DSS. There are several technologies that are used to encrypt network traffic including the IETF's RFC 4301 IPSec and Group Encrypted Transport VPN (GET VPN).

The previously mentioned growing adoption of a new generation of mobile devices has transformed how end users access applications. It has also created challenges for application architects as well as for those responsible for the IT infrastructure and for security. One way to respond to these challenges is to take the software that previously ran on desktop and laptop PCs and which provided key network functionality, and run this software on the new generation of mobile devices. One limitation of this approach is that these devices have limited processing power – typically one tenth the processing power of traditional desktops and laptops. In addition, this new generation of mobile devices often has significant limitation on the size and resolution of their display and can have limitations on the functionality of their web browsers, such as not supporting Flash. The network must be able to adapt to these restrictions and support the new generation of mobile devices that already outsell PCs and will soon become the dominant form of end user device. This includes being able to assess the device's security posture, VPN compatibility and wherever possible, providing WAN optimization.

The last few years has seen the evolution of a new generation of very sophisticated hacker. For example, it is somewhat common for crime families, hactivists and national governments to take advantage of Internet connectivity to gain access to applications, servers and end user devices. They use this access to achieve a variety of ideological and political goals as well as to extort money. The network must be able to automatically block any and all attempted security attacks and to allow only legitimate traffic to transit the network. Network security must be effective and pervasive at all points of the network. This includes network access and egress links as well as end user devices, whether those devices are mobile or stationary. With the movement to adopt cloud computing and to allow branch office networks to connect directly to CCSPs, network security in the corporate data center is no longer sufficient. In order to be effective, network security must be distributed to branch office networks as well as to CCSPs. This means that virtualized network firewalls and network access control systems must be cost effective enough so that IT organizations can afford to deploy them in branch office networks. Network firewalls and network access support systems that are deployed at CCSPs' facilities must work together with the network security systems that reside both in enterprise branch office networks and in corporate data centers in order to provide a comprehensive, defense in depth network security.

In order to reduce or eliminate the backhauling of Internet traffic, the branch offices' network equipment must also be sophisticated enough to provide the same level of security as is traditionally provided at the corporate data center. The final section of The Report entitled *Management and Security* will discuss the varying ways that IT organizations are implementing security and will also discuss the role of cloud based security.

The broad and rapidly growing movement to adopt both cloud computing and a new generation of mobile devices makes it significantly more difficult to achieve some of the goals of effective service delivery; e.g., effective management, appropriate levels of security. That increased difficulty stems from the fact that while it will still be common for the IT organization to own and manage the IT infrastructure that supports business critical applications, on an increasing basis that infrastructure will be owned and managed by one or more CSPs. Despite being owned by the CSP, the IT organization still needs to have end-to-end visibility and to be able to direct security policies. Standards such as NetFlow v9 and IETF IPFIX are key building blocks that can be leveraged to provide that functionality.

Traditional WAN Services

Background

The ***Survey Respondents*** were given a set of eleven WAN services and asked to indicate the extent to which they currently utilize each WAN service. The survey question included Frame Relay and ATM among the set of WAN services. In the not too distant past, these services were widely deployed. However, over half of The Webtorials Respondents don't have any Frame Relay in their networks and almost two thirds of The Webtorials Respondents don't have any ATM in their networks. In addition, few IT organizations are increasing their use of these technologies[6], while many IT organizations are decreasing their use of these technologies[7].

One of the observations that can be drawn from the response to this survey question is that:

The primary WAN services used by IT organizations are MPLS and the Internet.

Because of the prevalence of MPLS and the Internet and the lack of development of new WAN technologies, the majority of the rest of this section of The Report will discuss how functionality is being added to MPLS and the Internet to respond to the emerging requirements. This section will also briefly discuss the possible use of software defined networking in the WAN.

WAN Design Criteria and Challenges

The ***Survey Respondents*** were given a list of possible concerns and were asked to indicate which two were their company's primary concerns relative to its use of MPLS and the Internet. The set of concerns that were presented to the ***Survey Respondents*** is shown in the left hand column of **Table 4**. The second and third columns from the left in **Table 4** show the percentage of the ***Survey Respondents*** who indicated that the concern is one of their company's two primary concerns with MPLS and the Internet respectively. The right hand column is the difference between the second and third columns from the left. This column will be referred to as the delta column.

The delta column contains positive and negative numbers. A positive number means that that concern was mentioned more often relative to MPLS than it was mentioned relative to the Internet. For example, the ***Survey Respondents*** mentioned cost as one of their primary concerns about the use of MPLS 22.1% more often than they mentioned cost as one of their primary concerns about the use of the Internet. Analogously, a negative number means that that concern was mentioned more often relative to the Internet than it was relative to MPLS. For example, the ***Survey Respondents*** mentioned latency as one of their primary concerns about the use of the Internet 19.3% more often than they mentioned latency as one of their primary concerns about use of MPLS.

[6] Roughly 2% of IT organizations are increasing their use of Frame Relay and 6% of IT organizations are increasing their use of ATM.
[7] Roughly 34% of IT organizations are decreasing their use of Frame Relay and 22% of IT organizations are decreasing their use of ATM.

Table 4: Concerns about MPLS			
Concern	**MPLS**	**Internet**	**Delta**
Cost	60.1%	38.0%	22.1%
Lead time to implement new circuits	32.2%	11.4%	20.8%
Uptime	30.1%	46.3%	-16.2%
Latency	27.0%	46.3%	-19.3%
Lead time to increase capacity on existing circuits	23.5%	13.1%	10.4%
Jitter	14.8%	18.8%	-4.0%
Packet Loss	12.2%	26.2%	-14.0%

The primary concerns that IT organizations have with the use of MPLS are cost, the lead time to implement new circuits and uptime. The primary concerns that IT organizations have with the use of the Internet are uptime, latency and cost.

IT organizations typically design their WAN based on the following criteria:

1. Minimize cost
2. Maximize availability
3. Ensure appropriate performance

As shown in **Table 4**, MPLS is regarded by the ***Survey Respondents*** as doing a good job at ensuring appropriate performance because it exhibits relatively small amounts of delay, jitter and packet loss. Unfortunately, MPLS is regarded poorly relative to the goal of minimizing cost. In contrast, the Internet is regarded relatively well on the goal of minimizing cost but is regarded relatively poorly on the goal of ensuring appropriate performance. In addition, the ***Survey Respondents*** expressed concerns about both MPLS and the Internet relative to the goal of maximizing availability.

One viable approach to WAN design is to use both the Internet and MPLS in ways that maximize the benefits of each while minimizing their deficiencies.

As was pointed out in the section of this report entitled *The Emergence of Cloud Computing and Cloud Networking*, the goal of cloud computing is to enable IT organizations to achieve a dramatic improvement in the cost effective, elastic provisioning of IT services that are *good enough*. As that section also pointed out, in order to support a small number of business critical services and applications, a cloud network that is *good enough* will have to provide the highest possible levels of availability and performance. However, in a growing number of instances, a cloud network is *good enough* if it provides a best effort level of service at a reduced price. Hence, independent of the concerns those IT organizations have about the Internet:

In a growing number of instances, Internet-based VPNs that use DSL for access are 'good enough' to be a cloud network.

Some of the concerns that IT organizations have with the use of the Internet such as uptime, stem from the fact that in many cases IT organizations access the Internet over a single DSL link. The availability of DSL is somewhat lower than the availability of access technologies such as

T1/E1 links. One impact of this reduced availability is that Internet VPNs based on DSL access are often used only as a backup connection to a primary private WAN circuit. This is unfortunate because the shortfall in quality is fairly small when compared to the dramatic cost savings and additional bandwidth that can be realized by using broadband connections such as DSL and cable. One technology that addresses this issue is referred to as an *aggregated virtual WAN*.

The key concept behind an aggregated virtual WAN is that it simultaneously utilizes multiple enterprise WAN services and/or Internet connections in order to optimize reliability and minimize packet loss, latency and jitter.

Aggregated virtual WANs and other types of alternate WAN services are discussed later in this section of the report. As that discussion highlights, aggregated virtual WANs have the potential to maximize the benefits of the Internet and possibly MPLS while minimizing the negative aspects of both.

Local Access to the Internet

The traditional approach to providing Internet access to branch office employees has been to carry their Internet traffic on the organization's enterprise network (e.g., their MPLS network) to a central site where the traffic was 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 security policy. One disadvantage of this approach is that it results in extra traffic transiting the enterprise's WAN, which adds to the cost of the WAN. Another disadvantage of this approach is that it usually adds additional delay to the Internet traffic. The fact that centralized Internet access exhibits these disadvantages is significant because as highlighted in **Table 5**, cost and delay are two of the primary concerns that IT organizations have relative to the use of the Internet.

Some of the concerns that IT organizations have about the use of the Internet are exacerbated by backhauling Internet traffic to a central site.

The ***Survey Respondents*** were asked to indicate how they currently route their Internet traffic and how that is likely to change over the next year. Their responses are shown in **Table 5**.

Table 5: Centralized Access to the Internet

Percentage of Internet Traffic	Currently Routed to a Central Site
100%	32.1%
76% to 99%	17.3%
51% to 75%	15.6%
26% to 50%	13.1%
1% to 25%	12.2%
0%	9.7%

The way to read the data in **Table 5** is that 32.1% of the ***Survey Respondents*** route 100% of their Internet traffic to a central site and that 17.3% of the ***Survey Respondents*** route between 76% and 99% of their Internet traffic to a central site

The ***Survey Respondents*** also indicated that driven in part to save money and in part to improve application performance that:

Over the next year, IT organizations will make an increased use of distributed access to the Internet from their branch offices.

Cloud Networking Without the Internet

There is a temptation to associate the WAN component of *cloud networking* either exclusively or primarily with the traditional Internet[8]. However, due to a variety of well-known issues, such as packet loss at peering points, BGP's inability to choose the path with the lowest delay, the TCP Slow start algorithm, the Internet often exhibits performance problems. As such, the Internet is not always the most appropriate WAN service to use to access cloud computing solutions. To put the use of the Internet into context, The ***Survey Respondents*** were asked to indicate which WAN service their users would most likely use when accessing public and private cloud computing services over the next year. Their responses are shown in **Table 6**.

Table 6: WAN Services to Access Cloud Computing Services

	The Internet	An Internet Overlay	A traditional WAN service such as MPLS	WAN Optimization combined with a traditional WAN service; e.g. MPLS
Public Cloud Computing Services	61.2%	4.9%	18.8%	15.1%
Private Cloud Computing Services	35.3%	1.0%	36.7%	27.0%

The data in **Table 6** indicates that IT organizations understand the limitations of the traditional Internet relative to supporting cloud computing. In particular:

In roughly forty percent of the instances that business users are accessing public cloud computing services, the primary WAN service that they intend to use is not the traditional Internet.

In almost two thirds of the instances that business users are accessing private cloud computing services, the primary WAN service that they intend to use is not the traditional Internet.

[8] Throughout this report, the phrase "traditional Internet" will refer to the use of the Internet without the use of any optimization functionality.

Techniques that IT organizations can use to mitigate their concerns about the use of the Internet are discussed later in this section of the report.

Service Level Agreements

As previously stated, the majority of IT organizations utilize MPLS. One of the reasons for the popularity of MPLS is that the major suppliers of MPLS services offer a number of different classes of service (CoS) designed to meet the QoS requirements of the varying types of applications that transit a WAN. For example, real-time applications are typically placed in what is often referred to as a Differentiated Services Code Point (DSCP) Expedited Forwarding class that offers minimal latency, jitter, and packet loss. Mission critical business applications are typically relegated to what is often referred to as a DSCP Assured Forwarding Class.

Each class of MPLS service is typically associated with a service level agreement (SLA) that specifies contracted ranges of availability, latency, packet loss and possibly jitter. Unfortunately, in many cases the SLAs are weak. In particular, it is customary to have the SLAs be reactive in focus; i.e., the computation of an outage begins when the customer opens a trouble ticket. In most cases, the carrier's SLA metrics are calculated as network-wide averages rather than for a specific customer site. As a result, it is possible for a company's data center to receive notably poor service in spite of the fact that the network-wide SLA metrics remain within agreed bounds. In addition, the typical level of compensation for violation of service level agreements is quite modest.

To gauge the effectiveness of SLAs that IT organizations receive from their network service providers (NSPs), the ***Survey Respondents*** were asked to indicate which of the following best describes the SLAs that they get from their NSPs for services such as MPLS.

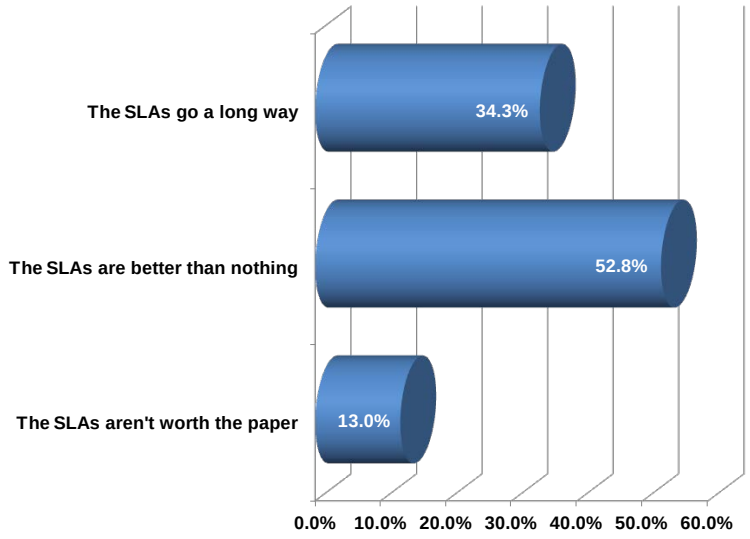
- The SLAs go a long way towards ensuring that we get a quality service from the network service provider.
- The SLAs are better than nothing, but not by much.
- The SLAs are not worth the paper they are written on.

Their responses are shown in **Figure 1**.

The fact that two thirds of the ***Survey Respondents*** indicated that the SLAs that they receive from network service providers are either not worth the paper they are written on, or that the SLAs they receive are not much better than nothing, demonstrates the weak nature of most SLAs.

The majority of IT organizations don't regard the SLAs that they receive from their network service providers as being effective.

Figure 1: The Effectiveness of SLAs



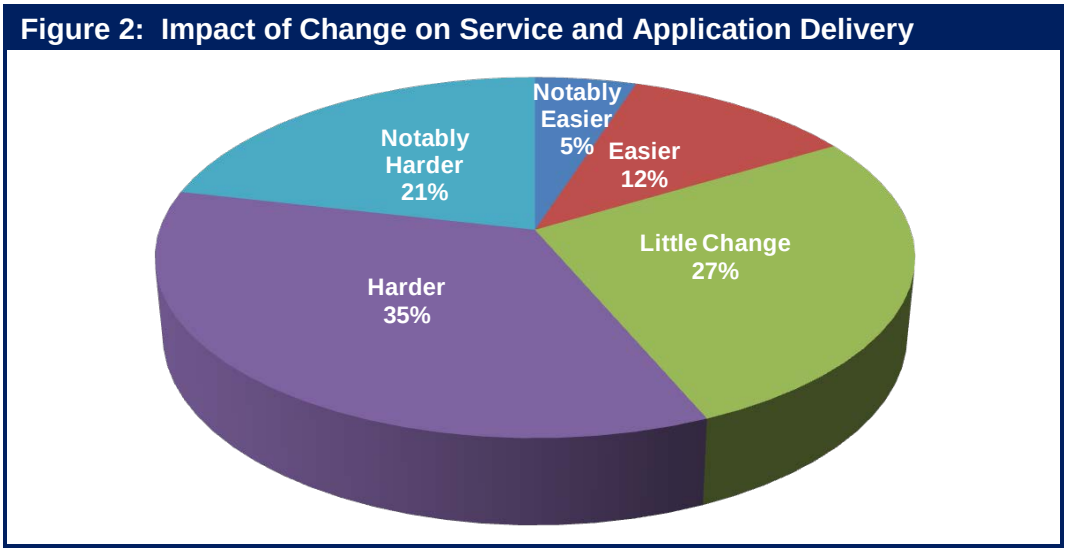
Optimizing the Performance of IT Resources

Background

This subsection of The Report will discuss techniques that IT organizations can implement to overcome the limitations of protocols and applications and to optimize the use of their servers. The focus of this subsection is on how these techniques enable IT organizations to ensure acceptable application and service delivery over a WAN. The discussion in this subsection will focus on two classes of products: WAN Optimization Controllers (WOCs) and Application Delivery Controllers (ADCs).

The introduction to this section of The Report discussed how the adoption of cloud computing in general is impacting the WAN and also discussed some of the specific factors that are driving change in the WAN. These factors included both the increasing number of mobile workers and the impact of multiple forms of virtualization. In order to gauge the effect that these factors have on the ability of an IT organizations to ensure acceptable application and service delivery, The ***Survey Respondents*** were asked "How will the ongoing adoption of mobile workers, virtualization and cloud computing impact the difficulty that your organization has with ensuring acceptable application performance?" Their responses are shown in **Figure 2**.

Figure 2: Impact of Change on Service and Application Delivery



One conclusion that can be drawn from **Figure 2** is that:

The majority of IT organizations believe that factors such as the growth in the number of mobile workers and the increase in the use of virtualization and cloud computing will make ensuring acceptable service and application delivery either harder or notably harder.

WAN Optimization Controllers (WOCs)

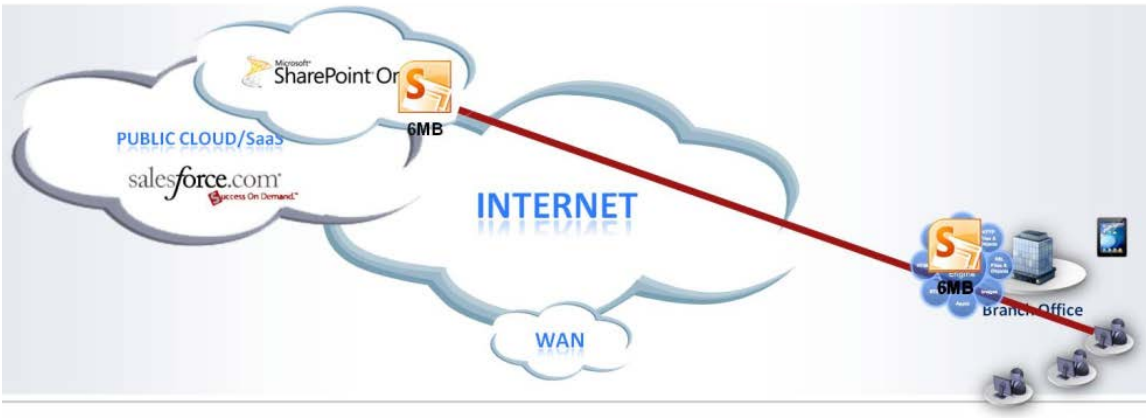
Goals of a WOC

The goal of a WOC is to improve the performance of applications and services that are delivered across a WAN from the data center either to a branch office, a home office or directly to a mobile user. In some cases the data center is owned and managed by the enterprise IT organization and in other cases it is owned and managed by a CCSP. The WOC accomplishes this goal by implementing techniques to overcome the limitations of the WAN such as constrained bandwidth, delay and packet loss.

WOCs are often referred to as *symmetric solutions* because they usually require complementary functionality at both ends of the connection; i.e., a WOC in the data center and another WOC at the branch office. However, the requirement to improve the performance of applications and services acquired from a CCSP has been the impetuous for the deployment of WOCs in an asymmetric fashion. One of the advantages of an asymmetric deployment of a WOC is shown in **Figure 3**. As shown in the figure, in an asymmetric deployment of a WOC content is downloaded from a CCSP to a WOC in a branch office. Once the content is stored in the WOC's cache for a single user, subsequent users who want to access the same content will experience accelerated application delivery. Caching can be optimized for a range of cloud content, including Web applications, streaming video (e.g., delivered via Flash/RTMP or RTSP) and dynamic Web 2.0 content.

As previously described, IT organizations are moving away from a WAN design in which they backhaul their Internet traffic from their branch offices to a central site prior to handing it off to the Internet. Also, as is described in the next section of this report, there are a variety of techniques that enable IT organizations to improve both the price-performance and the availability of distributed Internet access. As a result of these factors, asymmetric WOC deployment as described in the preceding paragraph will increasingly be utilized as part of a network design that features distributed Internet access. However, for this network design to be effective, IT organizations need to ensure that the design includes appropriate security functionality.

Figure 3: Asymmetric WOC Deployment



Modeling Application Response Time

A model is helpful to illustrate how the performance of a WAN can impact the performance of an application and it also serves to illustrate how a WOC can improve application performance. The following model (**Figure 4**) is a variation of the application response time model created by Sevcik and Wetzel[9]. Like all mathematical models, the following is only an approximation. For example, the model shown in **Figure 4** doesn't account for the impact of packet loss.

As shown below, the application response time (R) is impacted by amount of data being transmitted (Payload), the WAN bandwidth, the network round trip time (RTT), the number of application turns (AppTurns), the number of simultaneous TCP sessions (concurrent requests), the server side delay (Cs) and the client side delay (Cc).

Figure 4: Application Response Time Model

$$R \approx \frac{Payload}{Goodput} + \frac{(\#\ of\ AppsTurns * RTT)}{Concurrent\ Requests} + Cs + Cc$$

In order to improve the performance of applications that are delivered over the WAN, WOCs implement a variety of techniques. For example, to mitigate the impact of a large payload, WOCs implement techniques such as compression and de-duplication. These techniques are explained in detail in The 2012 Application Delivery Handbook. The handbook also details criteria that IT organizations can use to evaluate WOCs as well as specific techniques that WOCs need to support in order to optimize:

- The rapidly growing amount of traffic that goes between data centers
- Desktop virtualization
- Delay sensitive applications such as voice, video and telepresence

The 2012 Application Delivery Handbook also describes techniques that can optimize the delivery of applications to mobile workers. Many IT organizations, however, resist putting any additional software on the user's device. In addition, many users resent having multiple clients (e.g., WOC, SSL VPN, IPSec VPN, wireless/cellular access) on their access device that are not integrated. One option for IT organizations on a going forward basis is to implement WOC software on mobile devices that is integrated with the other clients used by mobile workers. As is explained below, an alternative way that IT organizations can improve the performance of applications and services delivered to mobile users is to utilize an optimization service from a CCSP.

Application Delivery Controllers (ADCs)

The current generation of ADCs evolved from the earlier generations of Server Load Balancers (SLBs) that were deployed in front of server farms. While an ADC still functions as a SLB, the ADC has assumed, and will most likely continue to assume, a wide range of sophisticated roles

[9] Why SAP Performance Needs Help, NetForecast Report 5084,http://www.netforecast.com/ReportsFrameset.htm

that enhance server efficiency and security and which provides asymmetrical functionality to accelerate the delivery of applications from the data center to individual remote users.

An ADC provides more sophisticated functionality than a SLB does.

Referring back to **Figure 4**, one of the factors that increase the application response time is server side delay. An ADC can reduce server side delay and hence can reduce the application response time. In particular, the ADC can allow a number of compute-intensive functions, such as SSL processing and TCP session processing, to be offloaded from the server. Server offload can increase the transaction capacity of each server, reducing the number of servers required for a given level of business activity.

The 2012 Application Delivery Handbook describes the primary techniques implemented by ADCs and identifies criteria that IT organizations can use to evaluate ADCs

Virtual Appliances

The section of this report entitled *The Emerging Data Center LAN* used the phrase *virtual switch* in two fundamentally different ways. One way referred to making two or more physical switches appear to be a single logical switch. The other way referred to the switching functionality that resides inside of a virtualized server.

In similar fashion, it is possible to look at a *virtual appliance* in a variety of fundamentally different ways. For example, two or more appliances, such as ADCs, can be combined to appear as a single logical ADC. Alternatively, a single physical ADC can be partitioned into a number of logical ADCs or ADC contexts. Each logical ADC can be configured individually to meet the server-load balancing, acceleration and security requirements of a single application or a cluster of applications.

However, the most common use of the phrase *Virtual Appliance* refers to what is typically appliance-based software, together with its operating system, running in a VM. Virtual appliances can include WOCs, ADCs, firewalls, routers, IDS, IPS and performance monitoring solutions. As explained in the next subsection of this report, virtual appliances make it easier for an IT organization to deploy network and application optimization functionality at a CCSP's data center. That, however, is not the only advantage of a virtualized appliance.

One of the compelling advantages of a virtualized appliance is that the acquisition cost of a software-based appliance can be notably less than the cost of a hardware-based appliance with same functionality.

In many cases the acquisition cost of a software-based appliance can be a third less than the cost of a hardware-based appliance[10]. A software-based solution can potentially leverage the functionality provided by the hypervisor management system to provide a highly available system without having to pay for a second appliance[11].

[10] The actual price difference between a hardware-based appliance and a software-based appliance will differ by vendor.
[11] This statement makes a number of assumptions, including the assumption that the vendor does not charge for the backup software-based appliance.

In addition, many IT organizations choose to implement a proof-of-concept (POC) trial prior to acquiring an appliance such as a WOC or an ADC. Using WOCs as an example, the purpose of these trials is to enable the IT organization to quantify the performance improvements provided by the WOCs and to understand related issues such as the manageability and transparency of the WOCs. While it is possible to conduct a POC using a hardware-based WOC, it is easier to do so with a virtual WOC. This follows in part because a virtual WOC can be downloaded in a matter of minutes, whereas it typically takes a few days to ship a hardware-based WOC. The value of the ease of downloading a virtual appliance is magnified in those cases in which the appliance is being delivered to a country where it takes a long time to get through customs.

Virtual appliances make is easier to conduct a proof of concept trial.

In addition to cost savings and making POCs easier, another advantage of a virtual appliance is that it offers the potential to alleviate some management burdens because most of the provisioning, software updates, configuration, and other management tasks can be automated and centralized at the data center. An example of this is that if virtualized appliances have been deployed, then it is notably easier than it is in a more traditional environment for various networking functions (WOC, ADC, firewall, router, etc.) to be migrated along with VMs in order to replicate the VMs's networking environment in its new location.

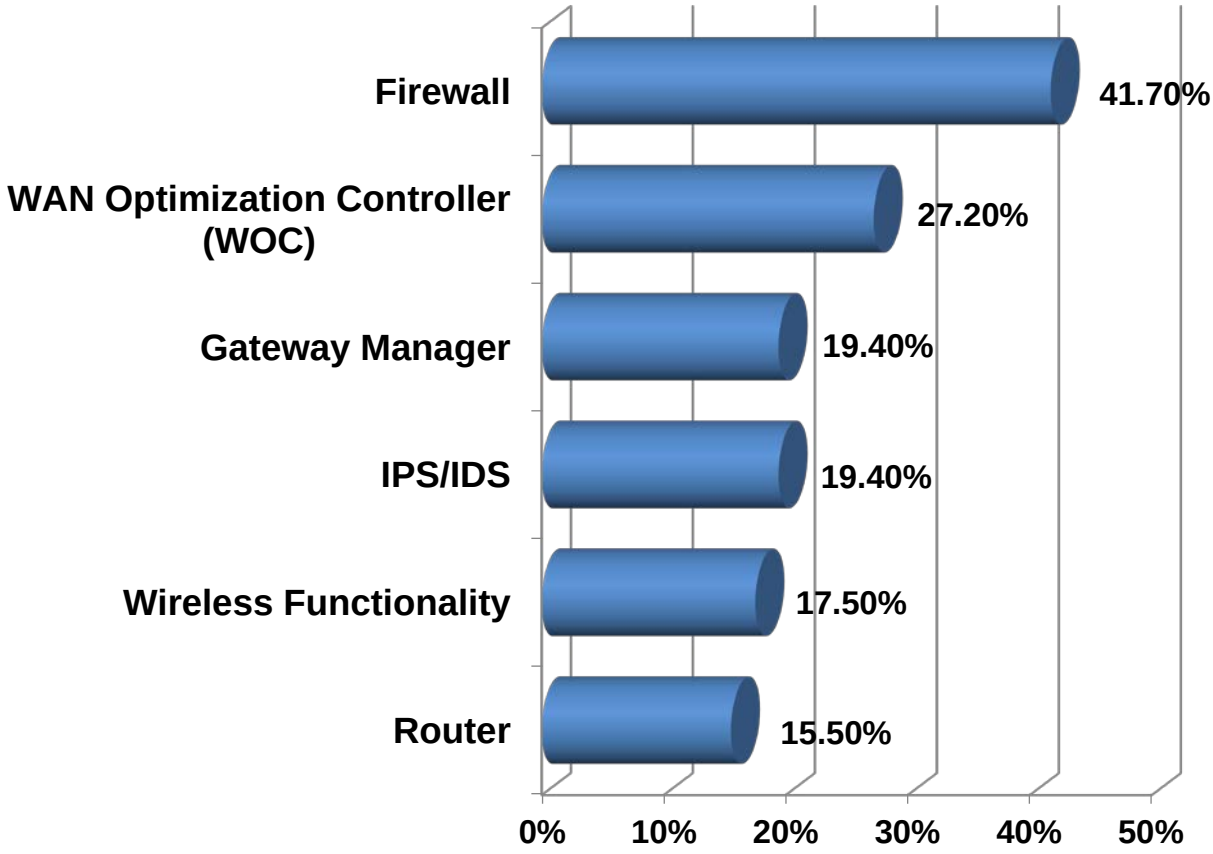
In many instances the benefits of the dynamic movement of a VM from one server to another are maximized if the supporting infrastructure is virtualized and can also be dynamically moved.

A virtualized ADC also makes it easy for an IT organization to package and deploy a complete application. One example of this packaging is the situation in which an entire application resides on VMs inside a physical server. The virtualized ADC that supports the application resides in the same physical server and it has been tuned for the particular application. This makes it easy to replicate or migrate that application as needed. In this case, a virtualized ADC also provides some organizational flexibility. For example, the virtual ADC might be under the control of a central IT group or it might be under the control of the group that supports that particular application. The later is a viable option from an organizational perspective because any actions taken by the application group relative to their virtual ADC will only impact their application.

A virtual firewall appliance can also help IT organizations meet some of the challenges associated with server virtualization. That follows because virtual firewall appliances can be leveraged to provide isolation between VMs on separate physical servers as well as between VMs running on the same physical server. Through tight integration with the virtual server management system, virtual firewall appliances can also be dynamically migrated in conjunction with VM migration where this is necessary to extend a trust zone to a new physical location. In addition, hypervisor APIs, such as VMware's Vsafe, can allow physical/virtual firewall consoles to monitor servers for abnormal CPU, memory, or disk activity without the installation of special agent software.

The Survey Respondents were asked whether or not their company had deployed virtual functionality in their branch office networks. Fifty-five percent indicated that they had and those respondents were then asked to indicate the type of virtual functionality their organization had implemented. Their responses are shown in **Figure 5**.

Figure 5: Virtual Functionality in Branch Offices



Optimizing Access to Public Cloud Computing Solutions

The conventional wisdom in the IT industry is that one of the key challenges facing IT organizations that use public cloud based solutions is improving the performance of those solutions. In order to understand how IT organizations intend to optimize the performance of services that they acquire from CCSPs, the ***Survey Respondents*** were given the following question:

If your company either currently acquires services from an Infrastructure-as-a-Service (IaaS) provider or you expect that it will within the next year, which of the following best describes the primary approach that your company will take to optimize the performance of those services?

Their responses are shown in **Table 7**.

Table 7: Optimizing IaaS Services	
Technique	**Percentage of Respondents**
Don't know	35.3%
We will leverage optimization functionality provided by the IaaS provider	27.8%
We will not do anything	18.2%
We will place a WAN optimization controller on the service provider's site and on our site	16.0%
We will use an optimization service from a company such as Akamai or Aryaka	2.7%

One conclusion that can be drawn from the data in **Table 7** is:

The majority of IT organizations are either undecided about how they will optimize the performance of IaaS services or they intend to do nothing.

The data in **Table 7** also shows some interest on the part of IT organizations to place a WOC on premise at an IaaS provider's data center. Referring back to the discussion in the previous subsection, IT organization will have a notably easier time placing an optimization device, whether that is a WOC or an ADC, at an IaaS provider's data center if the device is virtualized. That follows because if the device is virtualized, the IT organization can control the deployment of the functionality. If the device is physical, then the IT organization needs to get the IaaS provider to offer space for the device and to install it.

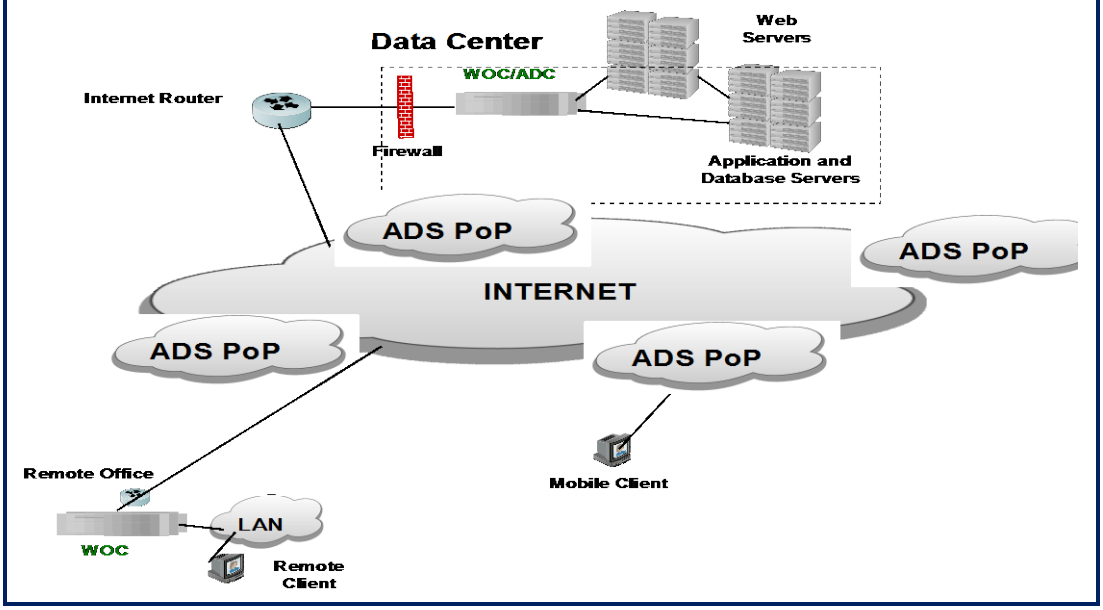
Alternative WAN Services

As noted, there isn't a new generation of fundamentally new technology focused on the WAN that is currently under development. However, as is described below, there are a number of WAN service alternatives that are variations on existing WAN technologies and services that better enable IT organizations to meet their WAN design goals. A number of these alternatives are either complementary to the WAN optimization technologies previously discussed or they depend partially on WAN optimization technologies to deliver acceptable levels of service quality.

An Internet Overlay

As described in the preceding subsection, IT organizations often implement WOCs and ADCs in order to improve network and application performance. However, these solutions make the assumption that the performance characteristics within the WAN itself can't be optimized because they are determined by the relatively static service parameters controlled by the WAN service provider. This assumption is reasonable in the case of WAN services such as MPLS. However, this assumption doesn't apply to enterprise application traffic that transits the Internet because there are significant opportunities to optimize performance within the Internet itself based on implementing an Internet overlay. An Internet overlay leverages service provider resources that are distributed throughout the Internet in order to optimize the performance, security, reliability, and visibility of the enterprise's Internet traffic. As shown in **Figure 6**, all client requests to the application's origin server in the data center are redirected via DNS to a server in a nearby point of presence (PoP) that is close to users of the application, typically within a single network hop. This edge server that is close to the users then optimizes the traffic flow to the server closest to the data center's origin server. Throughout this section, the Internet overlay that is depicted in **Figure 6** will be referred to as an Application Delivery Network (ADN).

Figure 6: An Application Delivery Service



An ADN provides a variety of optimization functions that generally complements the functionality provided by WOCs and ADCs. One such function that is often provided by an ADN is content offload. This calls for taking static content out of a data-center and placing it in caches in servers and in replicated in-cloud storage facilities that are close to the users. Because the content is close to the users, IT organizations that offload content and storage improve response time and simultaneously reduce both their server utilization as well as the bandwidth utilization of their data center access links.

Some of the other functionality that is often associated with an ADN includes:

- Route optimization
- Transport optimization
- HTTP protocol optimization
- Visibility

In addition to the functionality listed above, some ADNs incorporate Web application firewall functionality.

One use case for an ADN that is growing in importance stems from that fact that not all CCSPs will support virtual WOC instances in their data centers. This is particularly true of SaaS providers. Access to services provided by a CCSP can be accelerated via an ADN.

An Integrated Private-Public Solution

In almost all instances when a user accesses a CCSP-provided application or service they do that over the Internet and not over a private WAN service such as MPLS. That follows in large part because from the perspective of the CCSP, one or two high-speed Internet connections are much simpler and more economical to provision and manage than are connections to the varying private WAN services offered by multiple network service providers. In addition, the high fixed costs of these private WAN services can detract significantly from the overall cost-effectiveness of providing SaaS-based applications.

As previously discussed it is quite common for IT organizations to provide Internet access to branch office employees by carrying their Internet traffic on a private network (e.g., an MPLS network) to a central site where the traffic is handed off to the Internet. As was also previously discussed, many IT organizations have implemented WOCs in order to overcome the performance challenges that are associated with private WAN services. This means that the existing WOCs can utilize technology to overcome performance challenges such as TCP's retransmission timeout and the TCP slow start algorithm over the private WAN that connects a branch office to a central site. However, in the traditional scenario, once that traffic is handed off to the Internet, the performance of the application is negatively impacted by the limitations of TCP and by the transmission impairments (e.g., delay, jitter, packet loss) within the Internet.

It is possible to mitigate the impact of the performance impairments that are associated with employees in branch office using both a private network and the Internet to access a CCSP-provided applications and services by implementing an end-to-end approach to network and application optimization. A key component of the end-to-end approach is to integrate the optimization that is in place for private WANs with the performance gains that are provided by an ADN. As part of this integration, key functionality that is part of the ADN must be integrated into the WOC that sits in the enterprise data center. In addition, WOCs have to be distributed to the

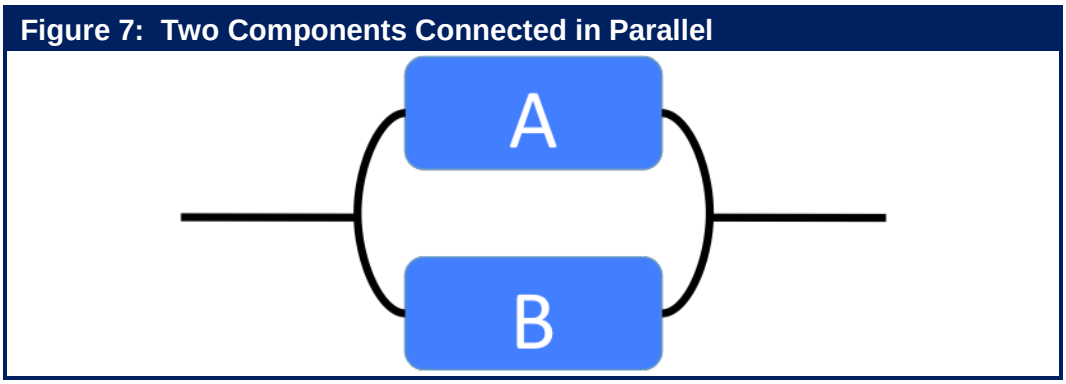
PoPs that support the ADN. The integration ensures a seamless handoff of functionality such as TCP optimization between the WOC in the data center and the ADN.

Dual ISP Internet VPN with Policy Based Routing

A preceding section of this report identified the concerns that IT organizations have with the use of the Internet. The two primary concerns are uptime and latency. Another approach to overcoming the limitations of the Internet is to connect each enterprise site to two ISPs. Having dual connections can enable IT organizations to add inexpensive WAN bandwidth and can dramatically improve the reliability and availability of the WAN.

For example, **Figure 7** depicts a system that is composed of two components that are connected in parallel.

Figure 7: Two Components Connected in Parallel



The system depicted in **Figure 7** is available unless both of the two components are unavailable. Assuming that each component is a diversely routed DSL or cable access line and that one of the access lines has an availability of 99% and the other has an availability of 98%, then the system has an availability of 99.98%. Alternatively, if both access lines have an availability of 99%, then the system is available 99.99% of the time[12]. This level of availability is equal to or exceeds the availability of most MPLS networks.

Traffic can be shared by the two connections by using Policy Based Routing (PBR). When a router receives a packet, it normally decides where to forward it based on the destination address in the packet, which is then used to look up an entry in a routing table. Instead of routing by the destination address, policy-based routing allows network administrators to create routing policies to select the path for each packet based on factors such as the identity of a particular end system, the protocol or the application.

Perhaps the biggest limitation of the PBR approach it that it creates a static allocation of traffic to multiple links and it doesn't have the ability to reallocate the traffic when the quality of one of the links degrades. The static nature of the policies means that, unless there is an outage of one of the links, a given class of traffic will always be allocated to the same network connection.

Dual ISPs and PBR can be used in conjunction with WOCs to further alleviate the shortcomings of Internet VPNs, bringing the service quality more in line with MPLS at a much lower cost point.

[12] If, as described later, 4G is added as a third access technique and if each access technique has an availability of 99%, then the system as a whole has an availability of 99.9999%.

For example, a WOC can classify the full range of enterprise applications, apply application acceleration and protocol optimization techniques, and shape available bandwidth in order to manage application performance in accordance with enterprise policies. As a result,

In many situations, a dual ISP-based Internet VPN with PBR can deliver a level of CoS and reliability that is comparable to that of MPLS at a significantly reduced price.

Part of the cultural challenge that IT organizations have relative to migrating traffic away from their MPLS network and onto an Internet based network is that Internet based networks don't provide a performance based SLA. However, as previously described, the majority of IT organizations don't place much value in the SLAs that they receive from their network service providers.

Hybrid WANs with Policy Based Routing

As noted, some IT organizations are reluctant to abandon traditional enterprise services such as MPLS. An alternative design that overcomes their concerns is a hybrid WAN that leverages multiple WAN services, such as traditional enterprise WAN services and the Internet, and which uses PBR for load sharing. The advantage of a hybrid WAN is that the CoS of MPLS can be leveraged for delay sensitive, business critical traffic with the Internet VPN used both for other traffic and as a backup for the MPLS network. As in the case of the dual ISP based Internet VPN, the major disadvantage of this approach is the static nature of the PBR forwarding policies. Since PBR cannot respond in real time to changing network conditions, it will consume more costly bandwidth than would a dynamic approach to traffic allocation. A second drawback of hybrid WANs based on PBR is that they can prove to be overly complex for some IT departments. As with many other types of WAN services, hybrid WANs can also be used in conjunction with WOCs and ADCs.

Aggregated Virtual WANs

A relatively new class of device has emerged to address the shortcomings of PBR-based hybrid WANs. WAN path controller (WPC) is one phrase that is often used to describe devices that work in conjunction with WAN routers to simplify PBR and to make the selection of the best WAN access link or the best end-to-end WAN path from a number of WAN service options.

Some members of this emerging class of products are single-ended solutions whereby a device at a site focuses on distributing traffic across the site's access links on a per-flow basis. Typical capabilities in single-ended solutions include traffic prioritization and bandwidth reservation for specific applications. These products, however, lack an end-to-end view of the available paths and are hence limited to relatively static path selections.

In contrast, symmetrical or dual-ended solutions are capable of establishing an end-to-end view of all paths throughout the network between originating and terminating devices and these solutions can distribute traffic across access links and specific network paths based on either a packet-by-packet basis or a flow basis. These capabilities make the multiple physical WAN services that comprise a hybrid WAN appear to be a single *aggregated virtual WAN*.

Aggregated virtual WANs (avWANs) represent another technique for implementing WANs based on multiple WAN services (e.g., MPLS, Frame Relay and the Internet) and/or WANs based on

just multiple Internet VPN connections. An aggregated virtual WAN transcends simple PBR by dynamically recognizing application traffic and allocating traffic across multiple paths through the WAN based on real-time traffic analytics, including:

- The instantaneous end-to-end performance of each available network: This allows the solution to choose the optimal network path for differing traffic types. One differentiator among virtual WAN solutions is whether the optimal path is chosen on a per packet basis or on a per flow basis. Per packet optimization has the advantage of being able to respond instantaneously to short term changes in network conditions.

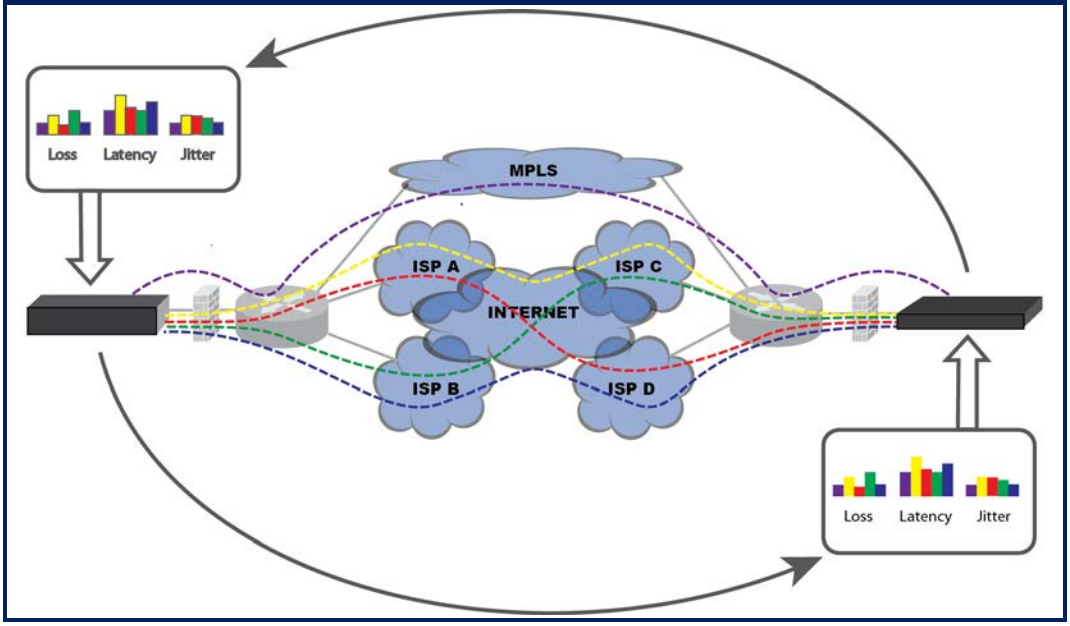
- The instantaneous load for each end-to-end path: The load is weighted based on the business criticality of the application flows. This enables the solution to maximize the business value of the information that is transmitted.

- The characteristics of each application: This includes the type of traffic (e.g., real time, file transfer); the performance objectives for delay, jitter and packet loss; as well as the business criticality and information sensitivity.

As previously noted, one of the primary reasons why IT organizations backhaul their Internet traffic to a central site over an enterprise WAN service is because of security concerns. In order to mitigate those concerns when using an avWAN for direct Internet access, the avWAN should support security functionality such as encryption.

Like other hybrid WANs, an avWAN (**Figure 8**) allows IT organizations to add significant amounts of additional bandwidth to an existing MPLS-based WAN at a relatively low incremental cost. In addition to enabling the augmentation of an MPLS WAN with inexpensive Internet connectivity, aggregated virtual WANs also give IT organizations the option to reduce its monthly ongoing expense by either eliminating or reducing its MPLS connections while simultaneously providing more bandwidth than the original network design provided.

Figure 8: Aggregated Virtual WANs



As shown in **Figure 8** because the two avWAN appliances work together to continuously measure loss, latency, jitter and bandwidth utilization across all of the various paths between any two locations, an aggregated virtual WAN can rapidly switch traffic away from a path that is exhibiting an unacceptable level of performance. This capability, combined with the availability advantages of parallel systems as depicted in **Figure 7**, means that all of the bandwidth in each of the paths can be used most of the time, and that most of the bandwidth can be used virtually all of the time. This combination of capabilities also underscores the ability of aggregated virtual WANs to deliver performance predictability that equals, and in many cases exceeds, that of a single MPLS network.

Because of the high availability and performance predictability of aggregated virtual WANs, IT organizations can now leverage a number of WAN services that are dramatically lower in cost than traditional MPLS services. This includes DSL and cable Internet access from branch offices and fiber access to the Internet from data centers. It also positions IT organizations to take advantage of the huge volumes of very inexpensive Internet access bandwidth that are typically available at co-location facilities.

While the preceding discussion focused on DSL and cable access to the Internet it is important to realize that there is an ongoing deployment of 4G services on the part of most wireless service providers. There will be some variability in the effective bandwidth of 4G services based in part on the fact that the wireless service providers will not all implement the same technologies. It should generally be possible, however, for users of these services to realize throughput in the range of three to four megabits per second, which is roughly equivalent to two T1 or E1 access lines. This will make 4G services a viable access service for some branch offices. For example, a 4G service could be combined with Internet access via DSL as part of a virtual WAN. In addition to providing cost savings, due to the inherent diverse routing associated with 4G and DSL, this design would provide a very high level of reliability.

Network-as-a-Service

As shown in **Table 4**, the two biggest concerns that IT organizations have with the use of MPLS are its cost and the amount of time it takes to implement new circuits. An emerging WAN service, referred to as Network-as-a-Service (NaaS), is intended to avoid those concerns. As shown in **Figure 8**, NaaS is built using a core network that interconnects a distributed set of Points of Presence (POPs). The phrase *NaaS* implies that unlike MPLS, the service can be deployed rapidly – typically within a day by leveraging Internet links for the first and last mile connections while providing a reliable private core network and additional network intelligence. The service also allows IT organizations to add capacity on demand, rather than provisioning and paying for bandwidth to support future requirements.

In order to meet enterprise requirements, the NaaS must deliver extremely high quality, predictable performance. It must also have enough POPs so that it is close to customers' sites. Some of the other specific characteristics of a NaaS that IT organizations should expect include:

- Centralized visibility across the WAN
- Low latency
- Diversity and redundancy
- Low packet loss
- Instant access to cloud based services or applications
- Support for multiple access methods

- Enterprise class security based on IPSec
- Low total cost of ownership

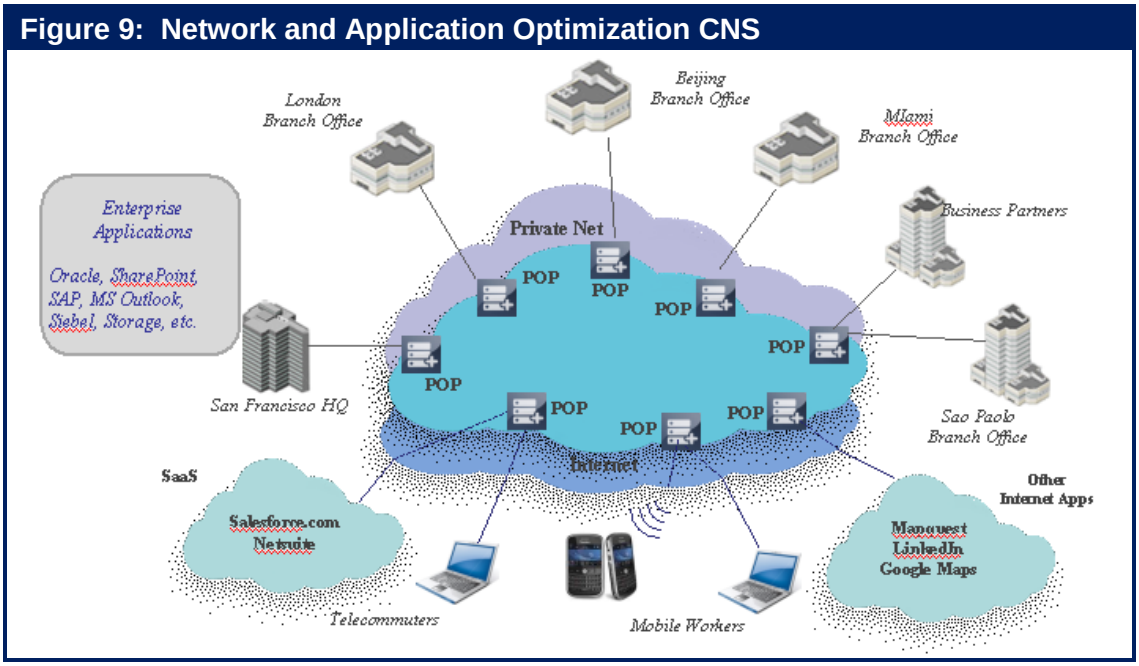
To mitigate the impact of packet loss on the first and last mile, the service should support a multi-segment TCP optimization architecture on those links as well as on the links that connect the POPs in order to ensure a rapid response to packet loss. The service should also honor industry standard QoS markings.

The next subsection of The Report discusses cloud-based network and application optimization that is based on a network similar to the one described above. Another key feature of a NaaS is that it should allow a customer to use the basic service if that is their choice, but it should also enable the customer to quickly upgrade to add the optimization capabilities discussed in the following subsection of The Report.

Cloud-Based Network and Application Optimization

As mentioned in the section of this report entitled *The Emergence of Cloud Computing and Cloud Networking*, network and application optimization has become available from CCSPs as a Cloud Networking Service (CNS). In this situation, instead of a physical or virtual WOC at each site, the WOC functionality is provided at the CCSP's cloud data centers or POPs, which ideally are in close proximity to the enterprise users, the data centers and the providers of other cloud services. As shown in **Figure 9**, the PoPs are interconnected by the CCSP's core network with customer access to each PoP provided via the Internet or via an enterprise WAN service. The CNS core network could be an Internet overlay, a private IP network or possibly a multi-carrier MPLS/IP network that uses intelligent routing capabilities similar to an aggregated virtual WAN or ADS in order to provide high levels of performance and reliability.

Figure 9: Network and Application Optimization CNS



In **Figure 9** a variety of types of users (e.g., mobile users, branch office users) access WAN optimization functionality at the service provider's points of presence (POPs). These POPs are inter-connected by a dedicated, secure and highly available network. To be effective, the

solution must have enough POPs so that there is a POP in close proximity to the users. In addition, the solution should support a wide variety of WAN access services.

There are at least three distinct use cases for the type of solution shown in **Figure 9**. One such use case is that this type of solution can be leveraged to solve the type of optimization challenges that an IT organization would normally solve by deploying WOCs; e.g., optimizing communications between branch office users and applications in a corporate data center or optimizing data center to data center communications. In this case, the factors that would cause an IT organization to use such a solution are the same factors that drive the use of any public cloud based services; e.g., cost savings, reduce the time it takes to deploy new functionality and provide functionality that the IT organization could not provide itself

The second use case is the ongoing requirement that IT organizations have to support mobile workers. Some IT organizations will resolve the performance challenges associated with supporting mobile users by loading optimization software onto all of the relevant mobile devices. There are two primary limitations of that approach. One limitation is that it can be very cumbersome. Consider the case in which a company has 10,000 mobile employees and each one uses a laptop, a smartphone and a tablet. Implementing and managing optimization software onto those 30,000 devices is very complex from an operational perspective. In addition, as previously discussed the typical smartphone and tablet doesn't support a very powerful processor. Hence, another limitation is that it is highly likely that network and application optimization software running on these devices would not be very effective.

The third use case for utilizing a solution such as the one shown in **Figure 9** is the expanding requirement that IT organizations have to support access to public cloud services. As previously mentioned, in some instances it is possible for an IT organization to host a soft WOC at an IaaS provider's site. However, that is generally not possible at a SaaS provider's site, and in any case a solution with a WOC at either end of a long distance Internet connection cannot address the congestion-based loss that occurs on the Internet. A Cloud-based optimization solution can improve users' access to cloud services by providing to the users the type of functionality typically provided in a WOC: reducing the amount of loss and high latency experienced between the end user's location and the location of the cloud service as well as and minimizing the impact of packet loss when it does occur.

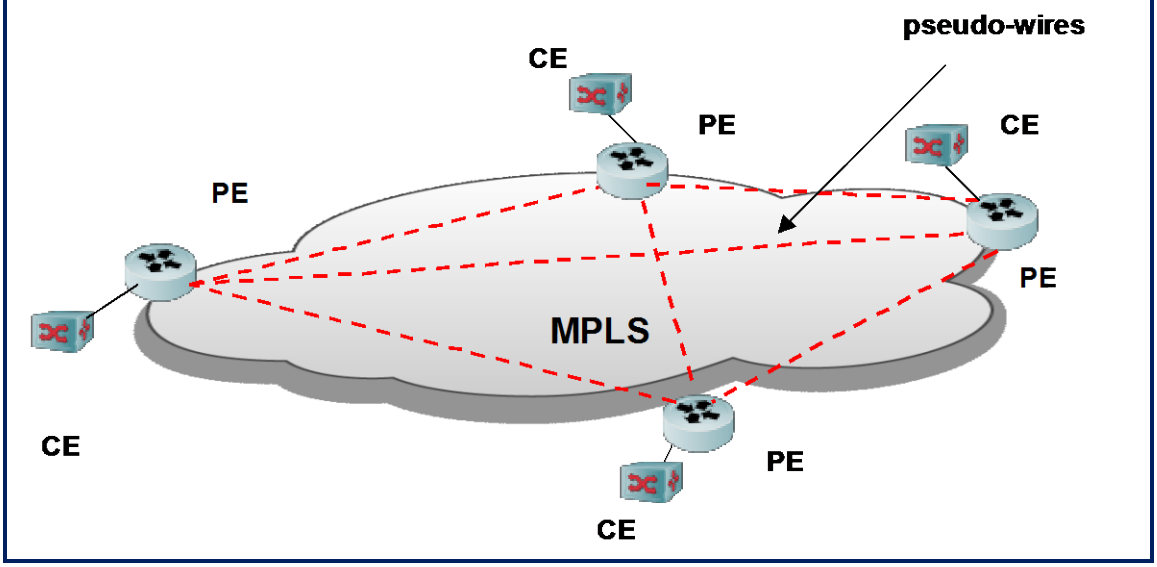
VPLS

As previously mentioned:

VPLS represents the combination of Ethernet and MPLS.

VPLS is a class of VPN that supports the connection of customer edge (CE) Layer 2 switches at multiple sites into a single bridged, multipoint-to-multipoint domain over a service provider's IP/MPLS network, as shown in **Figure 10**. VPLS presents an Ethernet interface to customers that simplifies the LAN/WAN boundary for Service Providers and customers, and enables rapid and flexible service provisioning. All sites in a VPLS appear to be on the same LAN, regardless of location. A companion technology, Virtual Private Wire Services (VPWS), provides point-to point services.

Figure 10: A VPLS Service Linking Four Customer Sites



With VPLS, either the Border Gateway Protocol (BGP) or the Label Distribution Protocol (LDP) is used to create the required pseudo-wires to fully mesh the provider edge (PE) devices serving the customer sites. Meshed pseudo-wires support the multipoint-to-multipoint nature of the virtual LAN and improve reliability. Reliability is enhanced because in case of failure in the MPLS network, traffic will automatically be routed along available backup paths, providing very short failover times.

VPLS MPLS packets have a two-label stack. The outer label is used for normal MPLS forwarding in the service provider's network. If BGP is used to establish the VPLS, the inner label is allocated by a PE as part of a label block. If LDP is used, the inner label is a virtual circuit ID assigned by LDP when it first establishes a mesh between the participating PEs. Every PE keeps track of assigned inner label, and associates these labels with the VPLS instance.

Table 8 provides a high level comparison of the different types of Ethernet WAN services available for LAN extension between data centers. It should be noted that there are other options for LAN extension, such as Ethernet over leased dark fiber and Ethernet over GRE tunneling through a private IP network.

Table 8: Ethernet WAN Service Types

Service Topology	Access Link	Provider Core	Service Type	Tunneling
Ethernet end-end	Ethernet	Ethernet	Pt-Pt or Mpt-Mpt	802.1Q or Q in Q
Ethernet/IP	Ethernet	IP	Pt-Pt or Mpt-Mpt	L2TPv3
VPLS/VPWS	Ethernet	MPLS	Pt-Pt or Mpt-Mpt	EoMPLS

Software Defined Networking (SDN)

As mentioned in the section of The Report entitled *Software Defined Networking* the most common discussion about implementing SDN focuses on the data center. However, as was also previously mentioned, Google has implemented SDN in their WAN, referred to as the G-Scale WAN, which interconnects their data centers. While SDN will not be a mainstream WAN technology for at least a couple of years, it does potentially represent a new approach to wide area networking.

As previously discussed, the G-Scale control plane is based on BGP and IS-to-IS and the OpenFlow-only switches are very simple 128 port 10 GbE switches that were built by Google using merchant silicon. It is important to note that when Google built these switches, 128 port 10 GbE switches had not yet been introduced in the commercial market. Google also built their own traffic engineering (TE) service. Their TE service collects both real-time utilization metrics and topology data from the underlying network as well as bandwidth demands from applications and services. The Google TE service uses this data to compute the best path for traffic flows and then programs those paths into their switches. The way that Google implemented the G-Scale WAN each site is comprised of multiple switch chassis to provide both scalability and fault tolerance. The sites are connected together and multiple controllers communicate with the switches using OpenFlow.

Google started this project in January 2010 and by early 2012 all of their data center backbone traffic was being carried on the G-Scale WAN. According to Google, some of the benefits of the G-Scale WAN include:

- Unified view of the network fabric: This simplifies configuration, management and provisioning.
- High Utilization: The centralized traffic engineering allows Google to achieve network utilization of up to 95%.
- Faster failure handling: In addition to handling failures faster, the systems converge more rapidly to target optimum and the behavior is predictable.
- Faster time to market/deployment: This comes in part from the fact that only features that are needed are developed.
- Hitless upgrades: The separation of the control plane from the forwarding plane enables hitless software upgrades without packet loss or capacity degradation.
- Elastic compute: The compute capacity of network devices is not longer a limiting factor. Large scale computation is done using the latest generation of servers.

According to Google, some of the challenges they faced were:

- OpenFlow: At the time they started the project, the OpenFlow protocol was just being developed and hence was not feature rich.
- Fault tolerant OpenFlow controllers: To provide high availability and scalability, multiple OpenFlow controllers must be provisioned which at the time that Google deployed the G-Scale WAN, required extra work on their part.
- Partitioning functionality: There is an ongoing lack of clarity in the industry as to what functionality should reside in network devices and what should reside in controllers.
- Flow programming: For large networks, programming of individual flows can take a long time.

Emerging Cloud Networking Specific Solutions

The preceding discussion of WAN services provided some insight into the interplay between the general requirements of cloud computing and the capabilities of WAN services to meet those requirements. One of the goals of this subsection of The Report is to describe the functionality that is required to support a particular form of hybrid cloud computing – cloud balancing. Another goal of this subsection of The Report is to describe some of the optimization functionality that is being developed specifically to support cloud computing.

Cloud Balancing

The phrase *hybrid cloud computing* refers to an IT organization providing IT services in such a way that each of the services is based in part on the private cloud that the IT organization operates and in part on the applications or services provided by one or more CCSPs. A hybrid cloud relies on a WAN to provide the connectivity between the enterprise's locations, including the enterprise's data center(s) and its remote sites, and the CCSP's data center. One of the goals of cloud balancing is to have the collection of individual data centers appear to both users and administrators as a single cloud data center, with the physical location of application resources as transparent as possible. The goal of having the location of application resources be transparent creates a number of requirements. This includes:

- **VLAN Extension**
  As is the case for private clouds, hybrid clouds depend heavily on VM migration among geographically dispersed servers connected by a WAN in order to ensure high availability and dynamic response to changes in user demand for services. The VLANs within which VMs are migrated must be extended over the WAN between and amongst the private and public data centers. This involves the creation of an overlay network that allows the Layer 2 VLAN traffic to be bridged or tunneled through the WAN.

- **Secure Tunnels**
  These tunnels must provide an adequate level of security for all the required data flows over the Internet. For the highest level of security, this would typically involve both authentication and encryption, such as that provided by IPsec tunnels.

- **Universal Access to Central Services**
  All application services, such as load balancing, DNS, and LDAP, should be available and function transparently throughout the hybrid cloud. This enhances security as well as transparency by allowing these application services to be provisioned from the private enterprise data center and by eliminating manual intervention to modify server configurations as the application and its VM are transferred from the private cloud to the public cloud.

- **Application Performance Optimization**
  Application performance must meet user expectations regardless of the location of the users or the IT resources that the users are accessing. This means that the public cloud data centers need to offer the same WAN optimization and application acceleration capabilities that are deployed within the enterprise. In addition, WOCs may well be needed between the enterprise's private cloud data center(s) and the public cloud data center(s) in order to accelerate VM migration, system backups, and other bulk data transfers between these data centers.

- **Interoperability Between Local and Global ADC Functions**
  Cloud balancing is based on making routing decisions based on a combination of local and global variables. This requires interoperability between local and global ADC functions.

- **Synchronizing Data between Cloud Sites**
  In order for an application to be executed at the data center that is selected by the cloud balancing system, the target server instance must have access to the relevant data. In some cases, the data can be accessed from a single central repository. In other cases, the data needs to co-located with the application. The co-location of data can be achieved by migrating the data to the appropriate data center, a task that typically requires highly effective optimization techniques. In addition, if the data is replicated for simultaneous use at multiple cloud locations, the data needs to be synchronized via active-active storage replication, which is highly sensitive to WAN latency.

WAN Optimization and Application Delivery for Cloud Sites

One of the most significant trends in the WAN optimization market is the development of new products and new product features that are designed to enable IT organizations to leverage public and hybrid clouds as extensions of their enterprise data centers. Some recent and anticipated developments include:

- **Cloud Optimized WOCs**
  These are purpose-built virtual WOC appliances for deployment in public cloud environments. Cloud optimized features include compatibility with cloud virtualization environments, SSL encryption and acceleration, and automated migration or reconfiguration of virtual WOCs in conjunction with VM provisioning or migration. As previously mentioned, WOCs can either be deployed in a symmetric fashion, with a WOC on each end of the WAN link; or in an asymmetric fashion, with a WOC deployed just in a branch office.

- **Cloud-based WAN Optimization Service**
  As mentioned in the Cloud-based Network and Application Optimization section above, this solution both leverages the Internet ecosystem and is a solution that provides accelerated, reliable access to public cloud services. It combines cloud-based WAN Optimization technology with a reliable core network, using globally distributed POPs and centralized WAN and application-layer visibility. Optionally an appliance can be deployed on premise for last mile bandwidth scaling. The service is intended to deliver the performance of WOC solutions without the high cost of MPLS or the cost and management overhead of traditional WAN Optimization appliance solutions, in a single combined, fully-managed service with no capital expenditures.

- **Cloud Storage Optimized WOCs**
  These are purpose-built virtual or physical WOC appliances for deployment in the enterprise's data center(s) and also at public cloud Storage as a Service environments that are used for backup and archival storage. Cloud optimized features can include support for major backup and archiving tools, de-duplication to minimize the required data transfer bandwidth and the storage capacity that is required, and support for SSL and AES encryption.

- **Cloud Optimized Application Delivery Controllers**
  One trend in the evolution of ADCs is increasing functional integration with more data center service delivery functions. As organizations embrace cloud computing models, service levels need to be assured irrespective of where the applications are hosted. As is the situation with WOCs, ADC vendors are in the process of adding enhancements that support the various forms of cloud computing, including:

- **Hypervisor–based Multi-tenant ADC Appliances**
  Partitioned ADC hardware appliances have for some time allowed service providers to support a multi-tenant server infrastructure by dedicating a single partition to each tenant. Enhanced tenant isolation in cloud environments can be achieved by adding hypervisor functionality to the ADC appliance and by dedicating an ADC instance to each tenant. Each ADC instance is then afforded the same type of isolation as a virtualized server instance, with protected system resources and address space. A combination of hardware appliances, virtualized hardware appliances and virtual appliances provides the flexibility for a cloud service provider to offer highly customized ADC services that are a seamless extension of an enterprise customer's IT environment.

- **Cloud Bursting and Cloud Balancing ADCs**
  Cloud bursting refers to directing user requests to an external cloud when the enterprise private cloud is at or near capacity. Cloud balancing refers to routing user requests to application instances deployed in the various different clouds within a hybrid cloud. Cloud balancing requires a context-aware load balancing decision based on a wide range of business metrics and technical metrics characterizing the state of the extended infrastructure. By comparison, cloud bursting can involves smaller set of variables and may be configured with a pre-determined routing decision. However, cloud bursting may require rapid activation of instances at the remote cloud site or possibly the transfer of instances among cloud sites. Cloud bursting and balancing can work well where there is consistent application delivery architecture that spans all of the clouds in question. This basically means that the enterprise's application delivery solution is replicated in the public cloud. One way to achieve this is with virtual appliance implementations of GSLBs and ADCs that support the range of variables needed for cloud balancing or bursting. If these virtual appliances support the IaaS cloud hypervisors, they can be deployed as VMs at each cloud site. The architectural consistency insures that each cloud site will be able to provide the information needed to make global cloud balancing routing decisions. When architectural consistency extends to the hypervisors across the cloud, integration of cloud balancing/bursting ADCs with the hypervisors management systems can help the routing of application traffic synchronized with private and public cloud resource availability and performance. Access control systems integrated within the GSLB and ADC make it possible to maintain control of applications wherever they reside in the hybrid cloud.

Planning for WAN Evolution

The ***Survey Respondents*** were asked "As your organization evolves its WAN over the next two years, which of the following describes the expectations that your organization will have for the functionality that the WAN will provide. The question had seven classes of WAN functionality and the **Survey Respondents** were asked to indicate all of the classes that applied in their environment. The responses of all of the ***Survey Respondents*** as well as just the ***Survey Respondents*** who work in large companies[13] are shown in **Table 9**.

Table 9: WAN Expectations

	All of The Survey Respondents	The Survey Respondents who work for Large Companies
Provide high level functionality such as security or optimization	50%	62%
Provide basic connectivity between users and business critical IT resources	57%	59%
Utilize basic QoS functionality to support voice, video and telepresence	49%	51%
Be aware of the applications and end points that It supports and adjust accordingly	41%	48%
Utilize not only basic QoS functionality, but also media-aware controls to support enhance voice	37%	44%
Provide integrated security	36%	43%

The data in **Table 9** indicates that the majority of IT organizations continue to see that one role of their WAN is to provide basic connectivity. However, the data also indicates that the majority of all IT organizations and an even bigger majority of large IT organizations also see that on a going forward basis, that their WAN must provide a range of higher value services that correspond closely to the functionality previously discussed in this section of The Report.

The Survey Respondents were also asked two additional questions. Those questions were:

1. *Does your organization have an architecture or strategy document that outlines the current state and likely evolution of your WAN?*

2. *Does the document have a significant influence on decision making around issues such as the choice of technologies, services and vendors (a.k.a., is it effective)?*

Their responses are shown in **Table 10** and **Table 11**.

[13] Throughout this section of The Report, the phrase *large companies* refers to companies with 10,000 or more employees.

Table 10: Does Your Organization have a Documented WAN Strategy?		
	Yes	**No**
All Companies	50%	50%
Large Companies Only	76%	24%

Table 11: Is Your WAN Strategy Effective?		
	Yes	**No**
All Companies	76%	24%
Large Companies Only	77%	23%

One conclusion that can be drawn from the data in **Table 10** and **Table 11** is that:

Slightly over a third of all companies, and slightly over a half of large companies have an effective WAN strategy.

In order to successfully respond to the challenges described in this report, IT organizations must create an effective strategy for how they will evolve their WAN. As described in this report, a key component of the WAN strategy that IT organizations must develop is to identify how the organization will continue to provide the same functionality as it does today, as companies make increasing use of public cloud computing services, independent of whether or not traffic is backhauled to a corporate data center prior to being handed off to the Internet. This functionality includes the ability to:

- Optimize application performance
- Provide intelligent QoS that reflects business priorities, not network priorities
- Provide end-to-end visibility of application performance over all segments of the network
- Dynamically route network traffic according to changing conditions
- Enable the growing adoption of all forms of Cloud Computing; e.g., Public, Private, Hybrid.
- Support a variety of end user devices and mobile workers
- Provide integrated network security regardless of the end user device and whether or not they are mobile
- Provide the ability to manage network performance and security policies centrally no matter where and who owns the hardware of the IT infrastructure

About the Webtorials® Editorial/Analyst Division

The Webtorials® Editorial/Analyst Division, a joint venture of industry veterans Steven Taylor and Jim Metzler, is devoted to performing in-depth analysis and research in focused areas such as Metro Ethernet and MPLS, as well as in areas that cross the traditional functional boundaries of IT, such as Unified Communications and Application Delivery. The Editorial/Analyst Division's focus is on providing actionable insight through custom research with a forward looking viewpoint. Through reports that examine industry dynamics from both a demand and a supply perspective, the firm educates the marketplace both on emerging trends and the role that IT products, services and processes play in responding to those trends.

For more information and for additional Webtorials® Editorial/Analyst Division products, please contact Jim Metzler at jim@webtorials.com or Steven Taylor at taylor@webtorials.com.

Published by Webtorials Editorial/Analyst Division
www.Webtorials.com

Division Cofounders:
Jim Metzler
jim@webtorials.com
Steven Taylor
taylor@webtorials.com

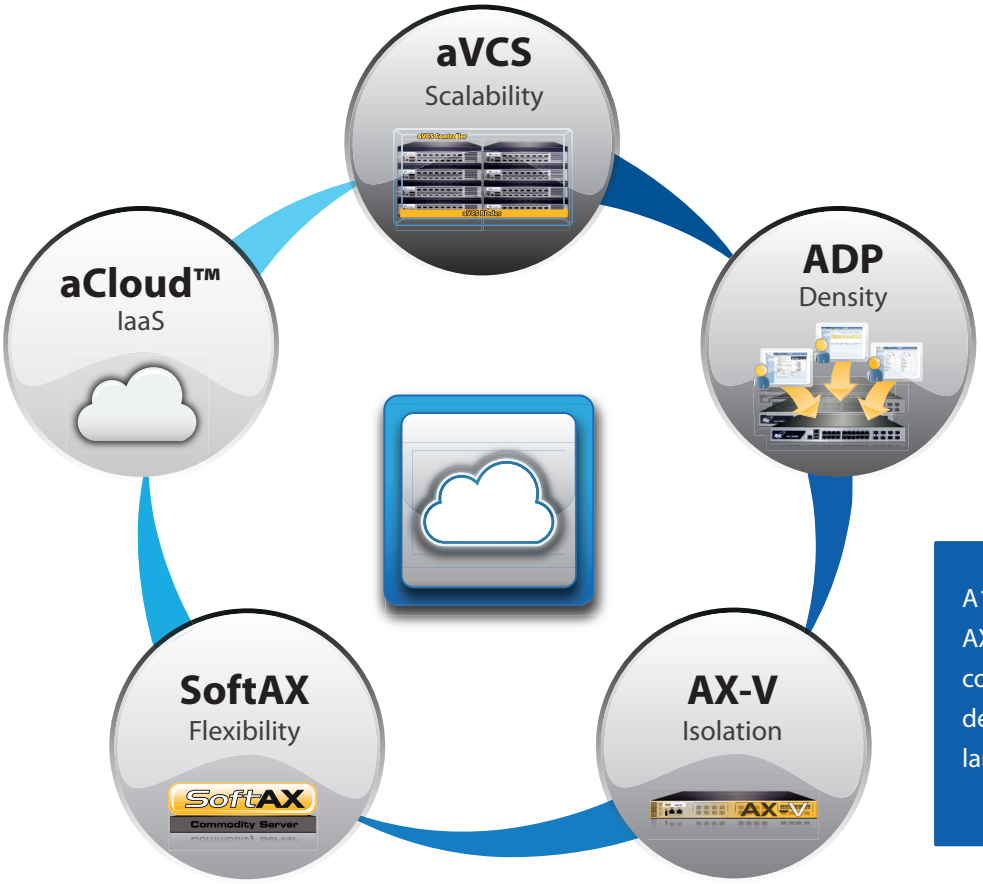
Professional Opinions Disclaimer

All information presented and opinions expressed in this publication represent the current opinions of the author(s) based on professional judgment and best available information at the time of the presentation. Consequently, the information is subject to change, and no liability for advice presented is assumed. Ultimate responsibility for choice of appropriate solutions remains with the reader.

For editorial and sponsorship information, contact Jim Metzler or Steven Taylor. The Webtorials Editorial/Analyst Division is an analyst and consulting joint venture of Steven Taylor and Jim Metzler.

Customer Driven Innovation

Cost Effective Cloud Networking | Virtualization as the Enabler



A10 also offers a powerful choice of AX Series ADC form factors with comprehensive management options, delivering flexibility and efficiency for large scale deployments.

AX Series Virtualization Products & Solutions

Based on A10's award-winning AX Series Application Delivery Controllers (ADC) and Advanced Core Operating System (ACOS™) architecture, enterprises and service providers will have the flexibility to choose the following scale-as-you-grow virtualization options.

SoftAX™

- SoftADC: AX virtual machine (VM) on commodity hardware
- Rapidly scale with commodity hardware
- Reduce hardware costs and upload to compatible cloud providers

AX-V Appliance

- SoftADC: AX virtual machine (VM) on AX Series hardware
- SoftAX flexibility with AX hardware performance and reliability
- Guaranteed performance, certifications, support and optimized hardware

AX Virtual Chassis System (aVCS™)

- Cluster multiple AX devices to operate as a unified single device
- Scale while maintaining single IP management
- Reduce cost and simplify management while adding devices as you grow

Application Delivery Partitions (ADPs)

- Divide the AX platform resources for individual applications
- Enables quality multi-tenancy with granular resource allocation
- Reduce the number of appliances to host multiple applications

The Application Fluent Data Center Fabric

Introduction

The rise of virtualization and cloud computing requires the selection of a best-of-breed data center switching solution as part of an enterprise's overall data center strategy. And at the heart of this strategy is the need to deliver a high quality user experience with new virtualized applications, including video, on new devices such as smart phones and tablets. However, the traditional 3-layer networks designed for a client/server communication model cannot meet the requirements of these new applications and devices, nor can it address the new requirements of virtualized servers and desktops.

Application Fluency for the Data Center

Resilient Architecture

- Simplified 10 & 40 GigE network with low latency and ready for 100 GigE
- Multi-path data center network extends between data center sites and to public cloud
- Supports definition of virtual data centers
- Ready for storage convergence with lossless Ethernet

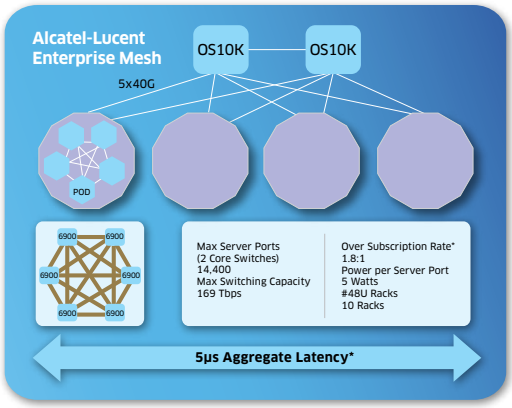
Automatic Controls

- Application profiles ensure that the network is aware of application provisioning, security and QoS requirements
- The network will automatically sense virtual machine location and movement
- The network will automatically adjust to VM motion within and between data center sites

Streamlined Operations

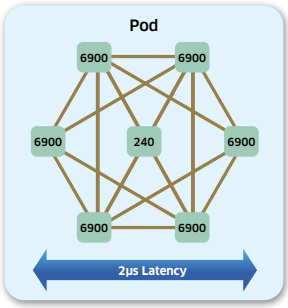
- Applications are automatically provisioned
- Core switches automatically configure top of rack switches
- Converged management for data center network and virtual machine mobility
- Low power consumption

The Alcatel-Lucent Mesh



Alcatel-Lucent provides a unique Application Fluent approach to maximize the benefit from virtualization technologies for servers, the desktop, as well as the network. Alcatel-Lucent's application fluent data center fabric can scale from several hundred to over 14,000 server facing ports while keeping aggregate latency at 5ms, and can automatically adapt to virtual machine movement no matter which server virtualization platform is used.

The Alcatel-Lucent Virtual Network Profile (vNP), embedded in the Alcatel-Lucent Mesh, includes the critical information the fabric needs to understand each application, including provisioning requirements, security profiles, and expected quality of service levels. With this knowledge, the network can manage applications as services, including automatically discovering the location of each virtual machine, modifying the network configuration to follow virtual machine moves and providing an integrated view on visibility on VM movement and current location from a network perspective.



Application fluency in the corporate data center includes its transformation into a multi-site private cloud by extending layer 2 connectivity between data center sites and allowing for seamless delivery of public cloud-based services on the corporate network.

The Alcatel-Lucent Mesh enables enterprises to provide a high quality user experience with mission critical, real-time applications, and to improve agility in deploying new applications while significantly reducing data center costs.

Open Ecosystems and Market Success

Alcatel-Lucent Enterprise is committed to open standards, allowing enterprises to select best-of-breed suppliers for their complete data center solution: servers, storage, data center fabric, and data center interconnect.

- Winner: Best of Interop 2011 for Data Center Switching and Storage
- Data center ecosystem partners include Emulex, NetApp, VMware, Citrix, and QLogic
- Participant in IEEE sponsored Shortest Path Bridging interoperability test with Avaya, Huawei, Solana and Spirent
- Over 20 million Ethernet ports shipped

For More Information

Alcatel-Lucent Data Center Switching Solution
Alcatel-Lucent Application Fluent Networks
Alcatel-Lucent Enterprise

**Assuming Server to Server Traffic 70% within a Pod, 20% between Pods and 10% Via Core*

Alcatel·Lucent
Enterprise

Visibility. Control. Optimize SaaS, BYOD, and Social Media

How to Lower Networking Costs and Safely Improve Performance

So many of the dominant trends in applications and networking are driven from outside the organization, including software-as-a-service (SaaS), bring-your-own-device (BYOD), Internet streaming video, and social networking. These technologies of an Internet connected world are fundamentally changing how we live and work every day. Yet, Network Administrators struggle to see and control these traffic streams from the Internet.

As businesses have opened their networks to SaaS applications, users are quickly starting using business bandwidth to access recreational websites and download BYOD updates, applications, and upload photos, videos and backups. This has created overburdened networks and slows the response of both cloud-based and internally delivered applications.

But with Visibility and Control from Blue Coat, Network Administrators can see all traffic on their networks and apply policies that can separate and control application traffic, and ensure internal and SaaS application performance.

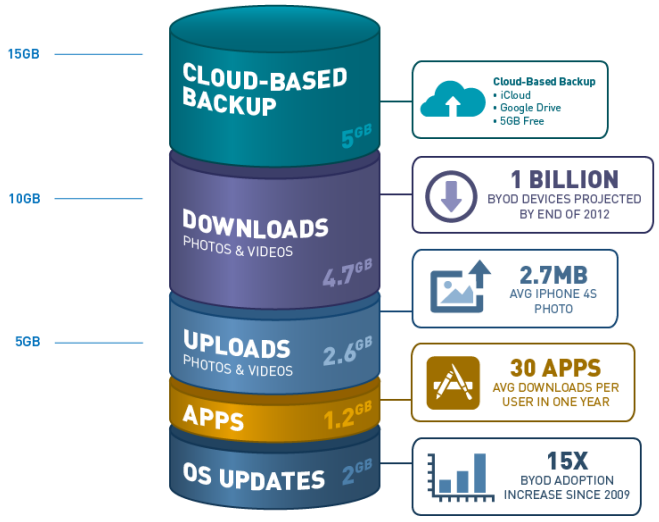
First: Visibility of all traffic on all ports – Understand what is on your network

Blue Coat PacketShaper leverages Blue Coat WebPulse™, an Internet Intelligence Service powered by a global community of 75 million users, the Cloud Service is able to deliver real-time categorization of Internet applications and web traffic.

WebPulse is based on sound analysis-system design principles:

- Massive input: WebPulse analyzes up to 1 billion web requests per day.
- In-depth analysis: 16 layers of analysis support over 80 categories in 55 languages.
- Granular policy: Up to 4 categories can be applied to each web request for multi-dimensional ratings.
- Speed: Automated systems process inputs – in most cases, in real time.
- Results: This collective intelligence allows WebPulse to categorize new Internet applications and websites quickly to PacketShaper without software updates/upgrades.

BYOD BANDWIDTH CONSUMPTION - JUNE 2011 TO JUNE 2012



The graphic details the impact of BYOD and Recreational video traffic can have on a network if left unchecked.

Second: Optimize Performance

SaaS, BYOD, Video and Social Media present challenges to network capacity and user patience. Blue Coat WAN Optimization helps overcome these challenges.

Chatty protocols and multi-megabyte files can hurt SaaS performance. Video requirements destroy capacity plans. Blue Coat's asymmetric, on-demand video caching and live stream splitting boost video capacity up to 500x – whether it's corporate or recreational video. For SaaS, our CloudCaching Engine improves performance by 3-93x, dramatically raising productivity for SaaS users at branch locations.

And now Blue Coat ProxySG/MACH5 technology secures SaaS applications as it accelerates their performance. ProxySG/MACH5 connects directly to the Blue Coat Cloud Service, enforcing SaaS user policies and leveraging WebPulse to scan and filter cloud traffic. Branch users can access applications like SAP, Salesforce, and RightNow without the burden of bandwidth slowdowns or risk of malware threats.

On The Road To The Cloud?

agility made possible™

ca technologies

With Converged Infrastructure Management and Network Automation, CA Technologies' allows you to transform your IT management functionality…reduce complexity and proactively optimize infrastructure while reducing costs…for a superior customer experience.

The Cloud Challenge…Increasingly CIO'-s and CEO'-s are looking to the IT organization to help deliver differentiation to the marketplace through innovation. As well, some organizations are looking to the Cloud to help them become more agile. Today "Cloud is synonymous with "Agility" but can you ensure your business services and guarantee application performance and availability in the cloud? How can you be proactive and optimize your infrastructure for lower costs while still delivering the highest quality user experience?

Cloud-Enable Your Network…CA Technologies Converged Infrastructure Management delivers ease of use and simple deployment while getting you up and running quickly with prescriptive OOTB capabilities- the benefit of IT organizations that say "It works as advertised." As well as functionality that can go deeper for dedicated IT teams showing them visibility into the infrastructure they specifically manage.

Access a single user interface for actionable performance, availability, flow capacity and application response information for all Layer 2 and Layer 3 technologies.

CA Technologies Converged Infrastructure Management delivers up to 25X Faster Problem Resolution While Reducing Total Cost by as Much as 50%. It helps you deliver a superior, differentiated customer experience – quickly and economically while -

Speeding proactive triage and remediation with less effort

- Analytics translate disparate data into intelligent views for up to 25x faster problem resolution

Meeting massive scalability demands cost-effectively

- Monitoring leading nationwide voice and video network with only two management servers

Shifting operations costs to innovation

- Converged infrastructure management reduces total costs by as much as 50%

Improving revenue streams

- Generate differentiated new sources of revenue and onboard new clients faster

The Cloud and Network Automation…CA Technologies Network Automation enables cloud-readiness all across your network, making your operation more efficient, more cost-effective and safer. Automation allows your workers to be more productive, improves your compliance and security issues, diminishes the risk of failure and ensures safe and immediate disaster recovery.

Automated dashboard for data collection and analysis to improve remediation options like manual time and level of effort.

Just some of the ways Network Automation helps enable Cloud is:

- Tasking over manual, error-prone processes of provisioning network devices.
- Detecting network changes and addressing their impact with troubleshooting and notifying in real time when issues are detected.
- Knowing and showing who is on the network, where and when at any given time, as well as archiving historical configurations.
- Updating network configuration changes on a wide number of devices from a central location automatically.
- Obtaining a current inventory of all components on the network and detecting policy and compliances failures in real time.
- Backing up all network configuration son a near real time basis, allowing restoration to take place in a matter of minutes.

Whether you are looking for ease-of-use, enterprise scalability or automation on your journey to the cloud, CA Technologies will help you deliver the innovation and agility that today's business services demand.

Visit us at http://www.ca.com/converge or http://www.ca.com/us/it-automation.aspx

Simplify and Accelerate Private Cloud Deployments with Cisco's Virtual Networking Portfolio

Cisco and a Multi-Vendor Ecosystem Provide Cloud-ready Network Solutions

ROLE OF THE NETWORK PLATFORM IN CLOUD
Access to Critical Data, Services, Resources and People • Core fabric connects resources within the data center and data centers to each other • Pervasive connectivity links users and devices to resources and each other • Network provides identity- and context-based access to data, services, resources and people
Granular Control of Risk, Performance and Cost • Manages and enforces policies to help ensure security, control, reliability, and compliance • Manages and enforces SLAs and consistent QoS within and between clouds, enabling hybrid models and workload portability • Meters resources and utilization to provide transparency for cost and performance
Robustness and Resilience • Supports self-healing, automatic redirection of workload and transparent rollover • Provides scalability, enabling on-demand, elastic computing power through dynamic configuration
Innovation in Cloud-specific Services • Context-aware services understand identity, location, proximity, presence, and device • Resource-aware services discover, allocate, and pre-position services and resources • Comprehensive insight accesses and reports on all data that flows in the cloud

The Power of Cloud for the Enterprise

Business and IT executives are confronted daily by conflicting and exaggerated claims of how cloud will transform their industries, but the lure of transformative efficiency and agility is hard to ignore. Understanding the objectives and obstacles to cloud, as well as the solutions to overcome those obstacles is the key to achieving cloud-readiness.

Defining Cloud

In the simplest terms, cloud is IT delivered as a service over the network. Going a level deeper, cloud is a model in which IT resources and services are abstracted from the underlying infrastructure and provided on demand and at scale in a multi-tenant environment.

- *On demand* means that resources can be provisioned immediately when needed, released when no longer required, and billed only when used.
- *At scale* means the service provides the experience of infinite resource availability to meet whatever demands are made on it.
- *Multi-tenant environment* means that the resources are provided to many consumers - for example, business units -from a single physical infrastructure.

Note that the physical location of resources (on or off premises) is not a part of this statement. From the perspective here, that aspect has more to do with the way the cloud is sourced than with what the cloud does.

CISCO VIRTUAL NETWORK PORTFOLIO

Routing and Switching

- Cisco Nexus 1000V virtual switch
- Cisco Cloud Services Router (CSR) 1000V

Security and VPN

- Cisco Virtual Security Gateway for Nexus 1000V (included in Nexus 1000V Advanced Edition)
- Cisco Adaptive Security Appliance (ASA) 1000V Cloud Firewall

WAN Optimization

- Cisco Virtual Wide Area Application Services (vWAAS)

Network Analysis and Monitoring

- Cisco Prime Virtual Network Analysis Module (NAM)

Application Delivery Controllers

- Citrix NetScaler VPX virtual application delivery controller

Virtual Services Deployment Platform

- Cisco Nexus 1100 Series Virtual Services Appliance

Cloud Orchestration and Management

- Cisco Intelligent Automation for Cloud
- Cisco Virtual Network Management Center (VNMC)

To learn more about Cisco's complete virtual networking portfolio: http://cisco.com/go/1000v

Barriers to Adoption

While most enterprises have recognized the potential benefits of cloud, practical concerns and perceived challenges have hampered the widespread adoption of cloud technologies and services. Many of these barriers can be understood as questions of trust: Can the cloud be trusted to deliver the same capabilities at the same service levels in the same controlled way as traditional IT?

- **Security**: Can the same security available to applications be applied in the cloud?
- **Compliance**: Can applications in the cloud meet the same regulatory compliance requirements?
- **Reliability and quality of service (QoS)**: Can the same service-level agreements (SLAs) for reliability and QoS be met in the cloud, especially given the multi-tenant use of the underlying IT infrastructure?
- **Control**: Can application owners still have the same amount of control over their applications and the infrastructure supporting them in the cloud?
- **Fear of vendor lock-in**: Will use of a particular vendor for cloud services or infrastructure prevent use of a different one in the future, or will the enterprise's data and applications be tightly locked into a particular model?

These concerns represent questions of technology and governance, but do not address any potential organizational friction that might arise from adopting cloud. For example, who will manage which part of the cloud or who will determine which applications to migrate to the cloud. Cisco believes that all these concerns can be met with the right technology, architecture, and approach.

Practical Solutions for Cloud-ready Virtual Networks and Infrastructure

The Cisco Virtualized Multi-Tenant Data Center (VMDC) architecture provides an end-to-end architecture and design for a complete private cloud providing IaaS capabilities. VMDC consists of several components of a cloud design, from the IT infrastructure building blocks to all the components that complete the solution, including orchestration for automation and configuration management. The building blocks are based on stacks of integrated infrastructure components that can be combined and scaled: Vblock™ Infrastructure Packages from the VCE coalition developed in partnership with EMC and VMware and the Secure Multi-Tenancy (SMT) stack developed in partnership with NetApp and VMware. Workload management and infrastructure automation is achieved using BMC Cloud Lifecycle Management (CLM). Clouds built on VMDC can also be interconnected or connected to service provider clouds with Cisco DCI technologies. This solution is built on a service delivery framework that can

be used to host other services besides IaaS on the same infrastructure: for example, a virtual desktop infrastructure VDI).

These solutions for building private clouds are also being used by service providers to build cloud infrastructures on which to provide public, hybrid, and virtual private clouds to their enterprise customers. With service providers and enterprises, Cisco is developing an ecosystem of cloud providers, builders, and consumers. This ecosystem will be able to take advantage of common approaches to cloud technology, management, interconnection, and operation.

Where to Begin Your Cloud Journey

Cisco is working with its broad ecosystem of partners to assist some of the world's leading institutions in their initial cloud deployments. Cisco will have a central role in the unique journeys of enterprises, small and medium-sized businesses (SMBs), public-sector organizations, and service providers as they move to cloud.

When the topic of cloud comes up, the conversation often focuses on the newest technologies and the latest service provider offerings. However, Cisco believes that every conversation needs to begin with an understanding of the expected business outcomes. Is the goal lower total cost of ownership (TCO) or greater agility and innovation, or some blend of the two? The journey to cloud has many paths; starting the journey without a clear understanding of the destination can lead to disappointing results.

Enterprises should start the journey to cloud by answering some basic questions:

- What is the expected impact of cloud on my business?
- Which applications can and should I move to the cloud?
- What cloud deployment model is best suited for each of my applications?
- How do I maintain security and policy compliance in the cloud?
- How do I transition my organization to best take advantage of cloud?

The answers to these questions will fundamentally shape your cloud strategy. We are helping customers define and implement a pragmatic approach to cloud. We deliver solutions that address our customers' unique business architecture and needs, align with regulatory constraints, and are optimized according to the customer's individual preferences for performance, cost, and risk.

For More Information

As you begin your own journey to the cloud, we invite you to discuss the right approach for your organization with your Cisco account manager, channel partners, and other IT advisors. For additional information about cloud, please visit: http://www.cisco.com/go/cloud.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Application Performance for Business Efficiency

The unique way to guarantee business application performance over the WAN, increase IT productivity and save on IT costs.

82%* of organizations suffer application performance problems.

63%* of organizations don't know the number of apps using the network.

72%* of organizations use very occasionally their network to its full data transmission capacity.

Business and IT performance are tightly coupled…

Losing 5 minutes per day for poor application performance means 1% of productivity drop which can turn down profitability by 10%.

**Ipanema Killer Apps survey 2012*

IT departments are witnessing change at a pace never seen before

Transformation is occurring as CIOs seek to access the benefits offered by Unified Communications, cloud computing, internet-based applications and consolidation, amongst many other strategic projects.

These initiatives are aimed at increasing enterprise's business efficiency. While they simplify the way IT is delivered to users, they increase the complexity and the criticality of corporate networking as applications and users rely more than ever on the continuous, reliable and consistent flow of data traffic.

In order to protect the business and the significant investments made in transformative applications such as Unified Communications and SaaS the network must be more intelligent, more responsive and more transparent. Ipanema's revolutionary self-learning, self-managing and self-optimizing Autonomic Networking System™ (ANS) automatically manages all its tightly integrated features to guarantee the application performance your business requires over the global network:

- Global Application Visibility
- Per connection QoS and Control
- WAN Optimization
- Dynamic WAN Selection
- SLA-based Network Rightsizing

Business efficiency requires guaranteed application performance

- Know which applications make use of your network…
- Guarantee the application performance you deliver to users…
- Manage cloud applications, Unified Communications and Internet growth at the same time…
- Do more with a smaller budget in a changing business environment, and to prove it…

With Ipanema, control all your IT transformations!

What our customers say about us:

Do more with less

"Whilst data volume across the Global WAN has increased by 53%, network bandwidth upgrades have only grown by 6.3%. With Ipanema in place we have saved $987k this year alone."

Guarantee Unified Communications and increase network capacity

"Ipanema is protecting the performance our Unified Communication and Digital Signage applications, improving our efficiency as well as our customers' satisfaction. Moreover, we have been able to multiply our available capacity by 8 while preserving our budget at the same time."

Reduce costs in a cloud environment

"With Ipanema, we guaranteed the success of our cloud messaging and collaboration deployment in a hybrid network environment, while dividing per 3 the transfer cost of each gigabyte over our global network."

For $3/employee/month, you guarantee the performance of your business applications… and can save 10 times more!

Ipanema's global and integrated approach allows enterprises to align the application performance to their business requirements. With an average TCO of $3/employee/month, Ipanema directly saves x10 times more and protects investments that cost x100 times more:

- **Application performance assurance**: Companies invest an average of $300/employee/month to implement the applications that support their business. At a mere 1% of this cost, Ipanema can ensure they perform according their application SLAs in every circumstance, maximizing the users' productivity and customers' satisfaction. While they can be seen as "soft money", business efficiency and investment protection are real value to the enterprise.
- **Optimized IT efficiency**: Ipanema proactively prevents most of the application delivery performances problems that load the service desk. It automates change management and shortens the analysis of the remaining performance issues. Global KPIs simplify the implementation of WAN Governance and allow better decision making. This provides a very conservative direct saving of $15/employee/month.
- **Maximized network efficiency**: Ipanema's QoS & Control allows to at least doubling the actual capacity of networks, deferring upgrades for several years and saving an average of $15/employee/month. Moreover, Ipanema enables hybrid networks to get access to large and inexpensive Internet resources without compromising the business, typically reducing the cost per Mbps by a factor of 3 to 5.

Enabling the cloud:

Award-winning NEC ProgrammableFlow® Open Software Defined Networking…

…delivering automated, efficient, and agile networks for the cloud

NEC's ProgrammableFlow network suite was the first commercially available SDN solution to leverage the OpenFlow protocol—enabling network-wide virtualization, allowing customers to easily deploy, control, monitor, and manage multi-tenant network infrastructure in a cloud environment. This architecture delivers better utilization of all IT assets, and helps provide ongoing investment protection as customers add functionality or upgrade their networks. NEC's approach simplifies network administration and provides a programmable interface for unifying the deployment and management of network services with the rest of IT infrastructure.

Specific functions customers prize include:

- **Drag and drop network design:** The GUI interface to the ProgrammableFlow Controller includes the familiar CLI found on most routers and switches today, so with minimal training a network admin can easily point and click to design an entire network from the single pane provided by the PF6800. This can radically reduce network programming and design time and errors caused previously by human intervention.
- **VM mobility:** With the ability to readily direct traffic throughout the data center—or throughout multiple data centers, it is possible to better manage all of the resources in a data center. For example, in NEC's own data centers in Japan, where they have recently implemented the ProgrammableFlow Fabric, it has enabled them to spread traffic between East and West Japan, offloading servers in East Japan that were nearing capacity, and postponing purchase of new servers, for a substantial saving. VM Mobility also enabled Nippon Express to complete a data center consolidation move that normally would have taken 2 months down to 10 days.
- **Bandwidth monitoring and traffic flow visualization:** This feature of the PF6800 provides performance monitoring of network flows and centralized management of network traffic, reducing bottlenecks and enabling smooth, streamlined network operations with substantially improved network admin productivity.
- **Secure, multi-tenant networks:** Secure, multi-tenant networks from the PF6800 enables customers like Genesis Hosting to expand their service offering with new sources of revenue potential. Genesis also reports software engineering investments were reduced by 100 hours each month with the advancements provided by ProgrammableFlow multi-tenancy.
- **Automation and administration of business policy to network management:** With network services aligned with business policy, automation such as prioritizing classes of applications or specific applications over other enterprise activity during peak loads is now possible with the ProgrammableFlow Network Suite, with multiple paths provided automatically. These capabilities offer significant value, particularly to enterprises engaged in heavy transaction loads.
- **Load balancing:** Traditional networking protocols often lead to performance-reducing bottlenecks. ProgrammableFlow uses path selection algorithms to analyze traffic flow across the network, check all available paths, and customize traffic flows to maintain performance and fully utilize network capacity. This increases the utilization of the network and improves application performance.

Backed by a 100-year history of technology innovation, NEC helps customers improve performance and solve their toughest IT challenges.

To learn more about how NEC can help you optimize your network for the cloud, visit necam.com/pflow or call your NEC Account Manager today.

Empowered by Innovation NEC

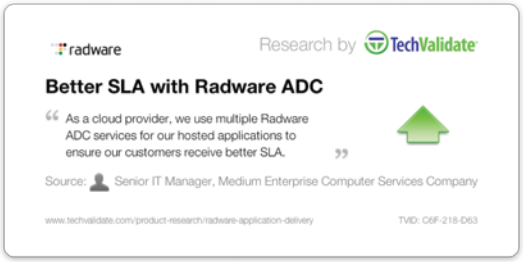
® ProgrammableFlow is a registered trademark of NEC Corporation

Expand Your Cloud Offering with Advanced Cloud ADC Solutions

Challenges in the Cloud Provider Business

The broad adoption of cloud based services by enterprise organizations and the multiple entrants into the cloud and hosting business challenges cloud providers to differentiate their service offerings and attract customers. Cloud providers face multiple challenges in establishing their business.

The first challenge is the infrastructure availability challenge. In an effort to provide uptime assurance at the base service level, or as a value added service offering, cloud providers must provide continuous availability of customer resources. One threat impacting the business availability is general connectivity: infrastructure outages and disruption events in which providers are dependent on external utilities and their running equipment. Failure to these can have significant adverse affect on the providers' business. Furthermore, part of the scalability value proposition of a cloud provider is the ability to scale-out application infrastructures – without load balancers, application scale-out is virtually impossible.



Above all, cloud providers are pressed to build solutions with minimal capital expenditure, maintain low operational costs and rapidly meet spikes in customer demand. Flexible procurement models by vendors and platforms that are easily scalable and centrally managed support the overall operational constraints faced by cloud providers.

Radware Solutions for Cloud Service Providers

Radware offers a set of fully integrated infrastructure availability and security solutions to meet the demands of cloud providers worldwide. Radware's solutions are comprised of the following components as illustrated in the figure below:

- **Radware ADC-VX™** – highly scalable ADC virtualization and consolidation solution offering high speed global and local load balancing, application acceleration and SSL offloading that supports dynamic availability requirements of cloud customers. ADC-VX can host multiple fully isolated, fully featured vADC instances.

- **Radware Alteon VA®** – flexible virtual ADC instance running atop most commercial, general purpose x86 server hypervisors.

- **Radware VADI®** – comprehensive virtual application delivery infrastructure solution including Alteon VA and ADC-VX-based virtual ADCs (vADC) and vDirect, an ADC service automation plug-in that simplifies ADC service deployment in cloud environments.

Radware's solutions enable cloud providers and hosts to offer more reliable and scalable infrastructure services to their customers. Resilience and scalability are key attributes of a cloud service as enterprises are contemplating the extent of cloud service adoption.

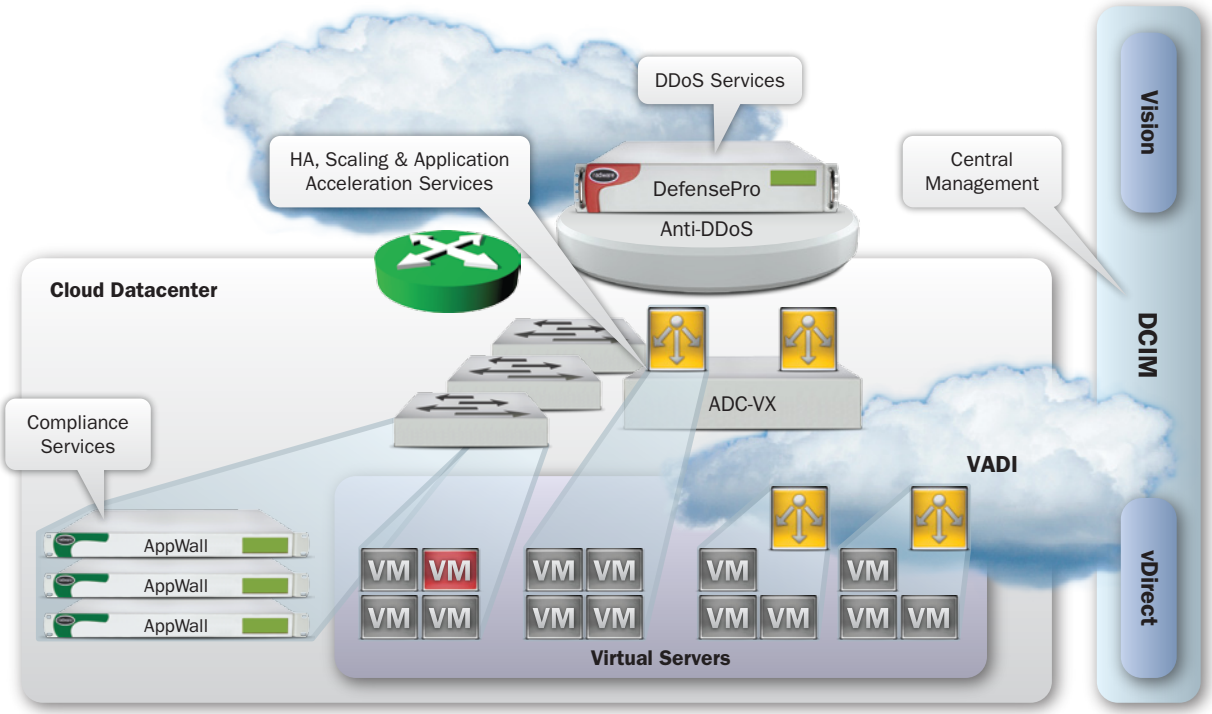


Figure 1 - Radware Service Architecture for Cloud

Benefits of Radware Solutions for Cloud Service Providers

1. Offer increased level of availability to cloud customers through highly available deployments of load balancing and application delivery services. High availability can be offered across any hardware form factor and location.
2. Seamlessly offer scale-out services to cloud customers inside cloud datacenters and across cloud datacenters by leveraging advanced health monitoring and KPI based global server load balancing.
3. Host a large scale of diverse services over a shared, purpose-built ADC infrastructure while fully isolating ADC instances associated with the different services.
4. Easily integrate application delivery and load balancing services into existing cloud service orchestration frameworks, home grown management tools and applications.
5. Simplify operations with a single management system controlling the entire set of Radware products in the cloud datacenter.
6. Cloud providers can offer additional value-add services such as application acceleration and application performance monitoring to their customers. All this while easily bundling the services into service packages and increasing customer confidence of rolling out applications in the cloud.

Summary

Radware application delivery and security solutions for cloud and hosting providers offer exceptional capabilities that greatly enhance the resilience, scalability and breadth of services offered by cloud and hosting providers. The value of the Radware is derived from 3 main benefits: (1) ability to enhance stability and scalability of cloud provider infrastructure (2) capability to help cloud providers build value added network services and offer these to their customers and (3) enabling these capabilities with minimal integration efforts and enhanced control.

Radware works with cloud providers globally addressing the key application delivery requirements presented in a cloud infrastructure through innovative cloud specific solutions.

For more information please visit http://www.radware.com