Acceleration Systems
WAN Optimization:
Hardware vs. Cloud-Based Solutions

September 2016, v1.2

Table of Contents

Introduction

Wide area network optimization involves a number of technologies, all of which either reduce or prioritize the traffic going over a data circuit. A reduction in data traffic yields several important benefits to companies and organizations: 1) Better utilization of data circuits in wide area networks (WAN), can either delay the acquisition of larger, more expensive data circuits, or reduce traffic to the point where smaller, less expensive circuits can be deployed; and 2) the reduced latency of a completed transaction creates a better end-user experience. This white paper examines some of the different optimization technologies that are available, and the benefits of cloud-based WAN optimization services.

Wide area networks are the central nervous system for any company or organization. Just as a human's nervous system can get over stimulated, WANs can become overloaded and out of balance. The result: poor performance.

Network congestion manifests itself in a number of ways, ranging from increased transaction times, or latency, all the way to a complete network failure. Network planners typically plot bandwidth usage, then implement larger data circuits in hopes of staying ahead of the demand; however, adding ever-increasing amounts of bandwidth come at a high cost.

Bandwidth demand is growing at a rapid rate. Global consumer Internet traffic grew 21% over the previous year in 2014.[1] Further, global IP traffic has increased more than fivefold over the last five years and will increase nearly threefold over the next five years.[2]

IT organizations have two options to deal with this challenge: 1) Add more bandwidth to their wide area networks; or 2) reduce the traffic going over the existing network.

Due to limited financial resources, IT organizations are turning to network tools that reduce the amount of data travelling across their WANs. Different technologies are available to achieve this task. Collectively, these tools are known as WAN optimization because they speed up the data flow.

Optimization Technologies

There are a number of WAN optimization technologies that reduce traffic and/or enhance the end-user experience. Some of the more well-known include: caching, traffic shaping, and data de-duplication.

Caching is a particularly useful technique to enhance the web browsing experience. It is based on the premise that users routinely visit their favorite websites for content. Think of the needless de-duplication when 1,000 users all retrieve data from a popular website over the same circuit. A caching engine recognizes and chronicles frequently visited web sites, and then stores the content so it can be accessed on a local basis. As a

[1] Mary Meeker's 2015 Internet Trend Report
[2] Cisco Visual Networking Index: Forecast & Methodology 2014-2019

result, web pages load faster and the web surfing experience is greatly enhanced. In the example above, caching would eliminate duplicate data traversing a data circuit by a thousand-fold.

Traffic shaping is a technique that allows network engineers to prioritize specific types of data packets. Certain types of applications, such as voice and videoconferencing, are real time applications and perform poorly when packets are missing or received in the wrong order. The result is a garbled conversation on a voice call, or poor audio or tiling during a videoconference.

Data de-duplication is a technique that looks for a repetitive string of packets and then eliminates the duplicates. To accomplish this task, the data stream is broken into specific sized blocks and inspected. A compression engine memorizes the data pattern of each block. As new blocks of data are analyzed, new patterns are recognized and stored in the memory of the compression engines on either end of the data circuit. Subsequent blocks of data are analyzed against the patterns stored in memory. Should a new block of incoming data match a pattern previously seen before, the compression engine replaces the block of data with a single, unique character. At the other end of the network, the process is reversed. When the unique character is received at its final destination, the duplicate block of data is reinserted into the data stream as originally transmitted. Transmitting a single character instead of thousands, or tens of thousands, can dramatically lower the amount of data transferred over the connection.

Traditional WAN optimization solutions utilize an appliance at either end of each data circuit, working in a point-to-point fashion. Appliances are available in different sizes, and as the speed of the circuit increases, the processing power and the associated cost of the appliance also increases.

Total Cost of Ownership - Hardware Based Solutions

Before an IT organization invests in an optimization solution, a return on investment (ROI) calculation must be made to determine if the solution makes financial sense.

The following facets should be considered when calculating ROI:

- Capital cost of hardware
- Ongoing software maintenance
- Complexity of the installation
- Ongoing technical support
- Network interruptions during the rollout

The capital cost of hardware and software is fairly straightforward but growth needs to be considered. Is the hardware appliance scalable and will it need an upgrade? What if bandwidth needs exceed the anticipated requirements? If so, what will this cost? Hardware costs can easily exceed six figures for large installations.

Annual software maintenance contracts are typically calculated at 19% of the cost of the hardware. Therefore, the software maintenance fee for an appliance supporting a single campus costing $100,000.00 would be $19,000.00 each year.

The cost of deploying an optimization solution needs to be factored into the ROI as well. Rolling out an optimization solution on a network-wide basis is challenging on several fronts. In today's business climate, IT staffs are very lean and, in most organizations, there is a lack of manpower available to actually implement the solution. Secondly, WAN optimization appliances are complex and manufacturers require certification to work on the equipment. Both of these challenges point to the need for outside consultants possessing specialized knowledge to perform installations, and in many cases, ongoing technical support.

The last important point to include in an ROI calculation is the cost of network downtime during the installation period. Any additional cost, such as redundant network equipment or additional manpower, should be included in the ROI calculation.

Although the benefits of optimization are clear, adoption rates are dramatically skewed towards offices supported by high-speed circuits. In short, existing optimization solutions are too expensive to effectively deploy for medium and small offices.

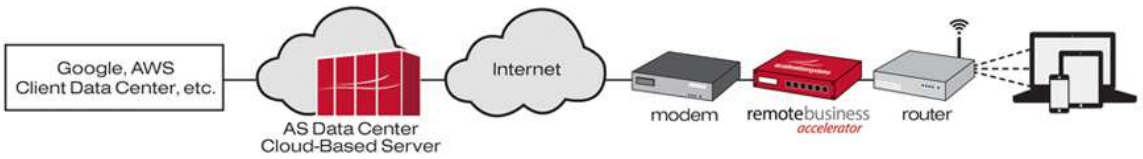
A Better Deployment Model: Cloud-based Solution

What is needed is a better delivery model, allowing the benefits of optimization technology to be deployed at every network location, regardless of size.

By moving the optimization hardware normally residing at a customer's office into the cloud, capital and support costs can be spread over thousands of locations instead of just one. Acceleration Systems has 4 patents for this disruptive technological design. WAN and Web acceleration solutions can now be deployed in a cost effective manner to all locations in a network, no matter how large or small the office.

Acceleration Systems has incorporated 26 different data de-duplication, data caching, and traffic shaping techniques into its cloud-based solution to minimize the traffic traversing a data circuit. The different technologies employed are complimentary to one another, each improving the overall data reduction and effectiveness of the system.

The Scalable Proxy Optimization Cloud (SPOC) is at the center of our bandwidth optimization solution. The SPOC is a multi-user cloud server environment that Acceleration Systems uses to coordinate data flow between the Remote Business Accelerator (RBA) and the Internet. SPOCs are deployed globally at strategic locations around the world and are directly connected to the backbone of the Internet. The SPOC can be deployed as either a public or private cloud to meet the needs of individual customers. The RBA is an inexpensive piece of hardware, located at each remote location. Typically, the RBA is plugged in behind the main router at the remote office.



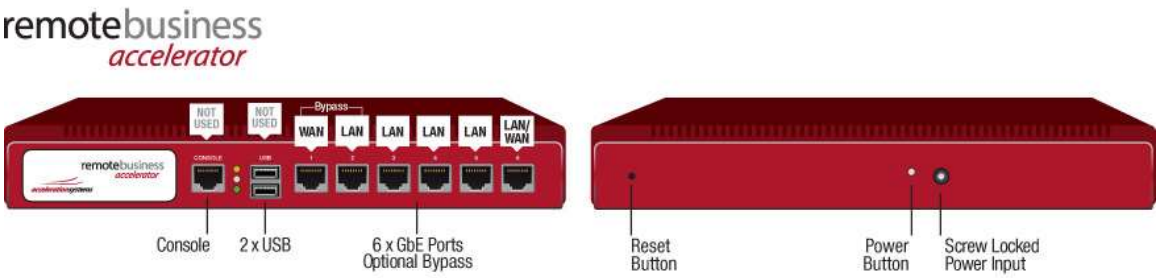
Each RBA works in conjunction with the cloud, reducing and optimizing traffic. Data flowing through the RBA is inspected in blocks; each block consists of 4,000 bytes. As each 4K block passes through the cloud or the RBA, its pattern is recognized and committed to memory. A unique data character is sent in lieu of sending data blocks that have been previously sent, thereby reducing data traffic.

Comparing Cloud-based to Hardware-based Solutions

The capital cost for a Remote Business Accelerator is less than $1,000.00. Organizations can equip 20-30 offices for the same capital cost as two high-end data compression engines.

Cloud-based optimization services are based on the number of concurrent users, and contract duration, such as month-to-month, 1-year, or 2-year subscriptions. It is not uncommon for cloud-based subscription fees to be an order of magnitude lower than the software maintenance fees on licensed products.

To deploy an RBA, unplug the local area network (LAN) connection going into the router at the remote location. Plug an Ethernet cable into the vacated port of the router and connect it to the RBA. Then plug the LAN connection into another port on the RBA and power the RBA up. Total investment of time: about 1 minute.



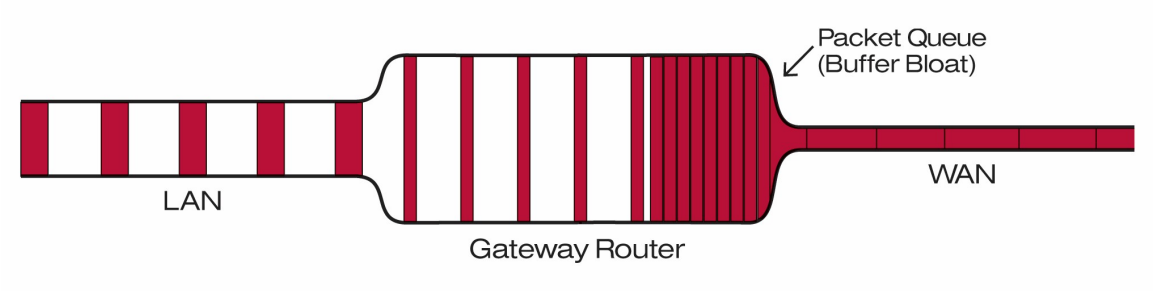
The RBA is preprogrammed by Acceleration Systems engineers to connect to the cloud upon restoration of power. It will then check for software updates to insure it is running the most current software package. The RBA will download any updates automatically, if needed, and begin active service. This eliminates complexity and the only training required is the ability to plug in an Ethernet cable.

The lack of complexity also eliminates the need for highly trained technical support personnel. Acceleration Systems provides both pre- and post-sales support. Should technical support ever be required after network deployment, Acceleration Systems provides help desk support on 24x7x365 basis.

Network interruptions and network downtime are significant issues, no matter how short the duration. An employee with an IT background can install an RBA in less than 30 seconds if all equipment and cables are unboxed and prepared ahead of time.

Adaptive Traffic Management

When two networks running at different rates interface with one another, the flow of packets from the faster network must be managed. To deal with the difference in speeds, routers contain buffers that hold a specific number of data packets.



When bursts of data arrive at a router, the excess data queues up in the buffer. If the queue becomes too great, time sensitive packets, such as Voice over IP (VoIP) and video, get buried in the buffer, thereby causing garbled voice conversations and tiling in a video presentation.

Ideally, the queue quickly drops to a manageable level or dissipates altogether. But if the high-speed side continues to send data at a rate above the capacity of the second leg of the connection, the queue will not dissipate. It will increase. This increase is referred to as buffer bloat.

Configuring Quality of Service (QoS) for a network has long been a challenge for network administrators due to changing throughputs on data links. Establishing priorities and bandwidth allocations for various types of traffic can be a daunting task due to constantly shifting data speeds and surges in bandwidth demand.

Acceleration Systems employs an advanced traffic management algorithm called Adaptive Queue Management (AQM). AQM senses changes in the upstream capacity of an Internet connection in real-time and actively regulates the rate that LAN-side applications send data to the Internet gateway device. AQM responds directly and dynamically to the available bandwidth at the gateway router. Coupled with an advanced Quality of Service prioritization system, AQM greatly improves the user experience. This is the first commercial implementation of AQM and the technology is currently not available from any other vendor.

Conventional QoS assumes that connections run at a constant speed. But most connections are shared, so the capacity changes constantly. Even on dedicated links the amount of available bandwidth fluctuates as users contend for a finite resource. So, a key underlying assumption to conventional QoS is seldom, if ever, met.

Conventional TCP flow control (and subsequent refinements such as RED and explicit congestion notification) do not prevent buffer bloat. Once the gateway buffer fills or overflows, time sensitive packets such as VoIP may get delayed due to their position in

the queue. As a result, the queue manager may not prioritize VoIP packets quickly enough to maintain call quality – even when VoIP traffic is given a top priority in the QoS configuration.

Acceleration Systems developed a prioritization hierarchy for protocols and data types. Their priority stack has proven successful in over 95% of the environments in which it has been deployed. This virtually eliminates the need for custom configuration. Administrators do have the ability to adjust QoS settings to meet local requirements in the rare instances where doing so would be beneficial.

With Acceleration Systems' Adaptive Queue Management efficiently controlling the flow of data across the network, the queue manager is able to prioritize traffic in a timely manner based upon a predefined hierarchy of protocols and data types. This combination of flow control and prioritization produce a more efficient use of network capacity than has previously been achieved and does so with little or no effort on the part of network administrators.

Case Study

Sigma Cubed is a service company that provides seismic mapping and down-hole seismic services for energy companies. The company begins a project by deploying a grid of geophones in the earth around a well, or formation, that is about to be fractured or stimulated. Sound waves reflecting off different layers of rock below the Earth's surface are captured by the geophones and the data is stored for processing. Sigma Cubed then creates images of the down-hole reservoir holding the crude oil or natural gas.

The data is time sensitive and is sent via a satellite link to Sigma Cube's data processing center. In addition to sending several types of seismic data, the company also supports VoIP traffic and remote desktop support programs over the same satellite connection.

Tommy Remmert, Real Time Operations Engineer, with Sigma Cubed, explained his company's need, saying "We need to send data from the well location back to Houston for processing. Depending on the number of geophones we deploy at a site, the data file size can range from 15 Megabytes 2 Gigabytes."

Remmert explained that the company tested multiple vendor's WAN optimization products, including Riverbed and Replify. Each vendor's product was tested in a lab setting using the same data files. Multiple data transfer techniques were tested on each of the different platforms, including: FTP, UDP Push, AMQP, and several others. In addition, VoIP traffic was sent during the process to insure there was no degradation of voice quality.

Remmert continued, stating, "We were interested to see if we could reduce the amount of data traffic going over our satellite link and were somewhat skeptical we would see any improvement because of the random nature of the data in a seismic file. After implementing Acceleration Systems bandwidth optimization solution, we were shocked to see the dramatic increase in transfer rates across the satellite. We were also pleased

to see an improvement in the performance of the different desktop support applications we use on remote computers."

All three solutions tested reduced the amount of traffic going over the Sigma Cube's satellite link, but Remmert noted several key differentiators, stating, "The Riverbed hardware was extremely complex and difficult to configure. You have to be a rocket scientist to configure their gear. Plus, equipping every one of our seismic trucks with a $15,000 appliance was going to be very expensive. I was able to get technical support from Riverbed but not to the same level that I received from Acceleration Systems. Acceleration Systems' hardware was extremely easy to install and any technical questions I had were resolved in a couple of minutes on the phone."

Sigma Cubed also bench tested Replify software. "We saw a reduction in data traffic, but Replify's software had to be installed on every computer at a remote site," Remmert explained. "With Sigma Cubed personnel, as well as our contractors, coming and going at a well site, it just too difficult to make sure all devices are configured the same. With Acceleration Systems, all we have to do is plug in a single RBA into our satellite modem and all devices connected to the modem are accelerated."

Conclusion

WAN optimization offers significant benefits. Reducing the data traffic on a WAN will extend the life of existing circuits, thereby delaying the need to buy more bandwidth. If bandwidth is expensive, such as a satellite circuit, the reduction of data traffic may allow IT organizations to provision smaller, less expensive satellite circuits, allowing the company to drive costs out of their business model immediately.

Traffic reduction also provides a better end-user experience. Regardless whether the end-user is conducting an Internet search or receiving an email with an Excel spreadsheet attached, the transaction is completed in less time, thereby increasing productivity.

Prior to a capital expenditure, an ROI calculation must be prepared by a company to verify the investment is warranted. The total cost of ownership of a hardware-based WAN optimization should include the following: capital cost of optimization hardware, consulting fees to configure and install all hardware, annual software maintenance fees for every piece of hardware, and the cost of network interruptions or network downtime during the installation period.

Moving the complexity and cost of a WAN optimization system into the cloud allows the cost to be spread over thousands of remote sites. A cloud-based WAN optimization system from Acceleration Systems incorporates a pre-configured Remote Business Accelerator costing less than $1,000.00. The installation of this device at a remote office involves the connection of two Ethernet cables. Upon power-up, the Remote Business Accelerator automatically connects to the cloud, checks for any updates, and then goes into service.

A cloud-based WAN optimization solution from Acceleration Systems requires significantly less capital investment than with a hardware-based solution. A cloud-based solution can reduce capital expenditure as much as $15,000 - $100,000 per location.

But capital expenditures aren't the only savings. It important to note that the ongoing subscription fees for a cloud-based WAN optimization solution are typically 5 -10 times lower than the equivalent software maintenance fees charged by hardware vendors.

Acceleration Systems offers the same great benefits that hardware-based solutions provide, but without the cost or complexity, allowing IT organizations to quickly deploy a cost-effective, cloud-based WAN optimization solution.

You can learn more about Acceleration Systems and their cloud-based WAN optimization solutions at http://www.accelerationsystems.com.