2114 West 7th Street
Tempe, AZ 85281 USA
Voice +1.480.333.2200
E-mail sales@comtechefdata.com
Web www.comtechefdata.com

How ISPs Can Maximize IP-Based Satellite Bandwidth

June 2011

Overview

The advent of new communication devices and IP-network convergence are enabling many innovative applications, which are in turn, placing greater demands on hub infrastructure and network bandwidth. As network and application infrastructure becomes more complicated, IT personnel are challenged to keep up with the growing demands, and they are looking to application delivery and network optimization solutions to improve VSAT availability and performance.

For virtually all IT organizations, there are business imperatives to ensure economical utilization of Internet connectivity and overall IT infrastructure. When end users are located in places where terrestrial network links are either too expensive, or have limited or no availability, organizations look to satellite Internet broadband communications to bridge the digital divide.

Latency is the result of the delay caused by the physical distance that a signal travels from an earth-based station to the orbiting satellite, and back to earth to complete a round-trip data transmission. Because of the physics involved in the round-trip-time (RTT), end-to-end latency can take 500 milliseconds or more, due to geosynchronous satellites orbiting 23,000 miles above the equator. This physical limitation is common to all satellite network service providers. But, this is just the tip of the iceberg. VSAT systems that deliver IP-based broadband services incur much greater levels of latency due to the extreme number of TCP and application protocol handshakes and redundant data. However, Internet Service Providers (ISPs) can gain a competitive advantage by using technologies to optimize satellite links in order to avoid additional “network induced latency” caused by excessive TCP handshakes, web application turns and content delivery inefficiencies that factor into the total latency. This is precisely what our Stampede FX Series solves with advanced content reduction, application acceleration and traffic shaping technologies.

Indeed, there are huge opportunities for ISPs to take advantage of a host of applications delivered to devices such as smart phones, desktop computers, tablet and notebook computers, all consuming huge amounts of data over limited amounts of bandwidth – at unprecedented rates. And, the amount of bandwidth consumed by these devices is only going to grow. Not only are these devices consuming huge amounts of data (and therefore, bandwidth), they are enabling new capabilities that are delivered over the Internet. Today’s new TVs are Internet-enabled. Movies are downloaded from NetFlix, Hulu and other sites. Videos are streamed from YouTube. Music is available from many sources such as iTunes, Pandora and others. Social networking sites such as Facebook, Twitter and LinkedIn are placing even greater demands on network bandwidth. The rise in the amount of data consumed by individual users and enterprises is increasing exponentially. The challenge for ISPs with satellite links is to cost-effectively keep up with the enormous demand for limited bandwidth, while conserving it.

WAN optimization vendors provide WAN Optimization Controller (WOC) appliances that are designed to speed application delivery over IP-based networks. However, traditional WOC vendors have focused on enterprise customers and ISPs with terrestrial networks, leaving ISPs within the VSAT market restricted in their ability to fully optimize the limited bandwidth within their satellite links. This created a large void for ISPs that were not able to find adequate, high-value network optimization products that met their performance needs, and were priced to meet their financial requirements. However, that has all changed with our Stampede FX Series.

Stampede FX Series Architecture

We are a global leader in satellite bandwidth efficiency and link optimization, providing advanced communication solutions that include Advanced VSAT Solutions, Modems, RAN & WAN Optimization, Managed Bandwidth and RF products.

Our FX Series appliances are WAN optimization devices based on a single platform and architecture that supports essential stand-alone hub optimization, and advanced hub and remote site optimization. For unsurpassed WAN optimization, the FX Series delivers content reduction, network and application acceleration, traffic management, Quality of Service (QoS) and flexible deployment and scalability that ISPs need to be successful. For example, with the FX Series in the hub, ISPs have the flexibility to use the essential content reduction features to support subscriber customers. Additionally, you can use the same FX Series in the hub, and provide remote FX Series appliances (or remote site software) to users who wish to have more advanced

content reduction and acceleration capabilities within a hub-spoke and even within a point-to-point trunking environment.

This paper discusses how our FX Series optimizes WAN infrastructure to improve bandwidth utilization and performance by minimizing the amount of bandwidth that TCP and applications consume. The products integrate compression, caching, acceleration, and QoS into a single, stand-alone appliance that delivers unmatched price/performance value.

Three primary Characteristics that Challenge Satellite Channels

High cost of bandwidth – ISPs are under continuous pressure to reduce operating expenses (OPEX) and improve profitability. Due to the high cost per megabit/s (in many cases an order of magnitude or more than the cost of terrestrial links), ISPs are looking for technologies that allow them to squeeze more traffic into satellite links. By leveraging bandwidth-efficient technologies, OPEX can be reduced and satellite link throughput can be significantly increased.

Limited data rates – Because of the high cost for bandwidth, satellite links are data rate limited, which leads to links becoming filled to capacity.

High latency – Latency can occur due to network congestion, distance, delay, packet loss and jitter. Latency reduces the amount of traffic that can be transmitted through a satellite link.

Network Traffic Causes Latency and Congestion

Below are some of the key aspects of network traffic that challenge satellite networks:

Chatty protocols – Transport and application protocols require many handshakes (or turns); often hundreds of turns are required in order to complete a transaction.

Redundant data – Very often data (text, images, attachments, HTTP, etc.) is sent over satellite links with non-essential elements. Satellite links can get filled up with unnecessary data, thus lowering performance, and limiting traffic flows. For example, applications that display images such as JPEGs use high resolutions that are typically not required for "normal" displays. This results in large amounts of bandwidth being utilized, and much longer load times for those images.

Repetitively accessed data – When repetitively accessed data are routinely accessed by end users, satellite links become saturated with requests for the same files. This adversely impacts users with long response times.

Fortunately, there are solutions to address these problems, alleviating the adverse effects that satellite links have on application delivery. Application delivery and network optimization solutions play an important role in improving performance, simplifying complexity and providing flexibility in scaling IT operations. Below are three technology areas that address the satellite application delivery problems.

Technologies that Optimize Satellite Bandwidth

Content Reduction – Content reduction technologies shrink the amount of data delivered over satellite networks. Using less data bits to provide the same information is equivalent to making the pipe bigger. This enables ISPs to support more users and applications without increasing bandwidth. By removing non-essential data from the network and the inefficiencies caused by web applications, ISPs are able to utilize less bandwidth to support traffic, thus, increasing the efficiency of the total amount of bandwidth.

Application Acceleration – The proliferation of chatty and reliable protocols can wreak havoc with long latencies, preventing a pipe from filling and greatly increasing the user response time. Acceleration technologies address both of these issues by allowing the pipe to fill quickly and reducing the number of round trips required to complete a connection transaction. A rapidly growing number of applications are placing greater demands on satellite links. To address this problem, advanced network and application acceleration technologies are providing ISPs with bandwidth optimized links able to handle more traffic, allowing them to remain competitive**.**

Traffic Management –While acceleration technologies are effective in keeping the pipe full, it is equally important to keep it filled with the right traffic. Traffic management allows the operator to allocate data rates (bandwidth) appropriately, classify the traffic, and set the appropriate priority on the types of traffic. Lower priority traffic can be blocked, or throttled back, when the link is full. Traffic management technologies such as QoS/Traffic Shaping enable ISPs to measure and control network traffic and packets going over satellite links. This allows the channel to be filled with the right traffic, while ensuring time sensitive traffic (e.g. /Voice over IP) is delivered with the optimal bandwidth and priority to achieve a high quality user experience.

Bandwidth optimization consisting of content reduction, application acceleration, and traffic management technologies provide network administrators with the ability to improve head-end and satellite link performance – the ability to fully utilize the link by optimizing bandwidth and improving traffic throughput. Some of the key benefits include:

- Reducing the amount of non-essential data sent over satellite links
- Reducing the number of TCP and application turns (handshakes) required to complete transactions
- Offloading computationally intensive tasks from servers and clients
- Measuring and controlling network traffic and packets going over satellite links

FX Series Hub and Remote Site Optimization Capabilities

Hub WAN optimization solutions enable ISPs to reduce critical bandwidth utilization, accelerate application delivery, support more user traffic and applications, and provide a quality user experience. The FX Series hub optimization solution is ideal for ISPs that need to improve their satellite network performance. The FX Series can be deployed as a hub-only Application Delivery Controller (ADC) or as a two-sided WAN Optimization Controller (WOC) by adding a Remote appliance (REM). The same ADC can optimize single-sided links, as well as two-sided links as shown in figure 1.

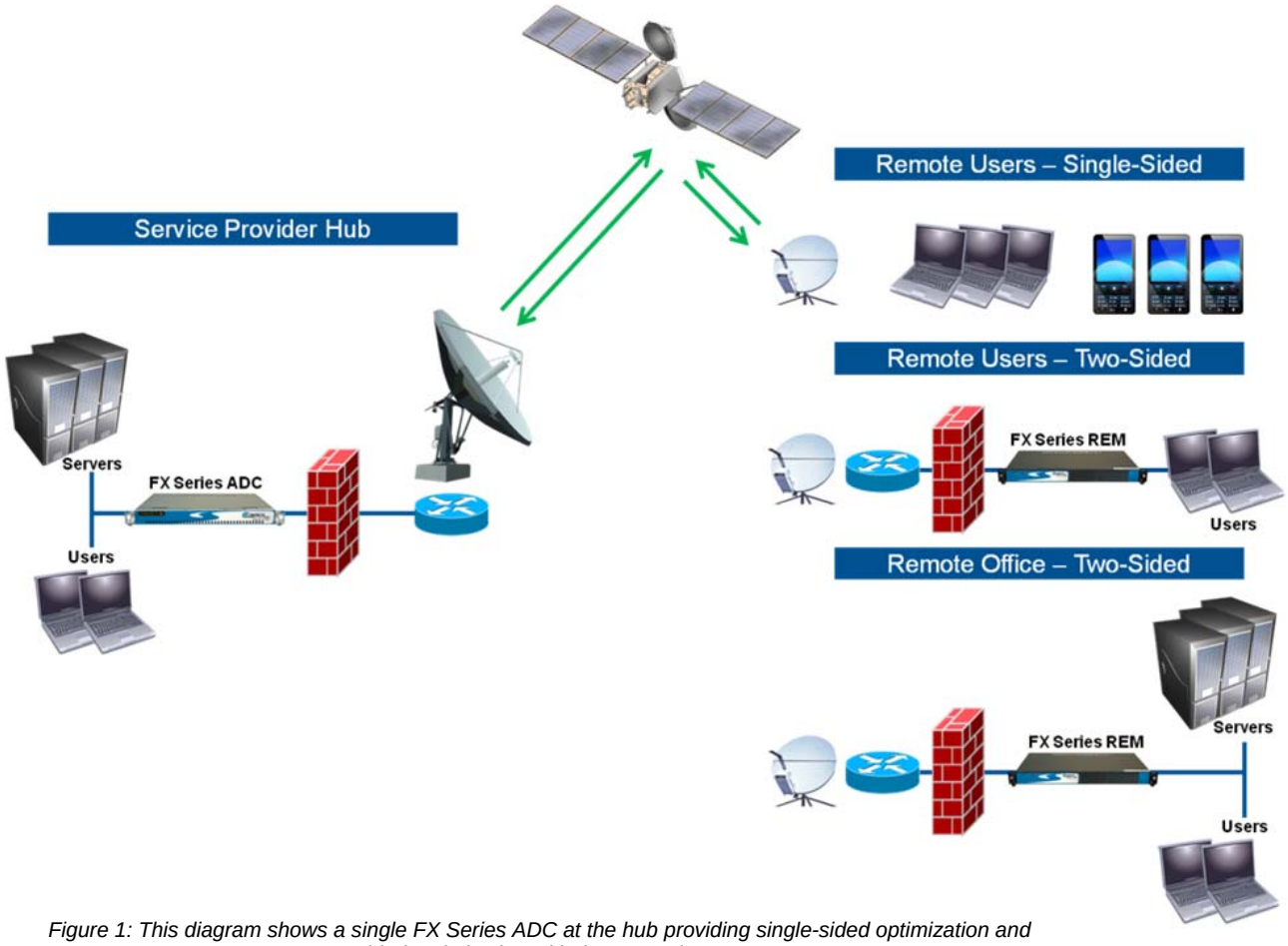

Figure 1: This diagram shows a single FX Series ADC at the hub providing single-sided optimization and two-sided optimization with the FX Series REM.

Data Compression

Compression – In single-sided deployments, the FX Series ADC will compress all HTTP traffic that is not already compressed. This compression is done in a manner that the web browser can decompress.

Bi-directional data compression – We utilize various compression techniques to reduce the amount of data that must be sent across the network. When the FX Series is deployed in a two-sided environment, our bi-directional compression provides for:

- All HTTP Headers
- Application Cookies
- All Text and Data Objects
- JPEG files with Image Reduction, yielding very acceptable quality for all attachments, file uploads and downloads

Image Reduction and Smoothing – This process reduces the amount of data required to represent an image without significantly altering the visual perception of the image by reducing the high frequency components or the sharpness of an image. A moderate amount of smoothing can significantly reduce the amount of data without visible detection. The FX Series provides advanced Image Reduction and Smoothing. Depending on the JPEG, the compression is often in the range 9:1.

Dynamic Data Suppression (DDS) – Using block level de-duplication, DDS is a technique for recognizing and replacing repetitive streams of payload data with signatures prior to transmission over satellite links. DDS is not application protocol specific and can be applied to most TCP application traffic. The FX Series intelligently monitors the data stream and distinguishes protocol headers which change frequently from payload data, which is often static. The FX Series extracts this payload data and segments it into blocks, storing each block into persistent memory known as a "byte cache". Blocks of data are replaced with a signature for that data, achieving significant data reduction.

Content Caching

Static Caching – The FX Series utilizes caching to accelerate content requests by maintaining copies of routinely accessed content to eliminate unnecessary requests. By keeping local copies of frequently requested content, the FX Series allows organizations to significantly reduce upstream bandwidth usage and cost, while improving performance. The FX Series acts as an intermediary from end users requesting content (e.g. files, web pages, other resources) from servers.

For static caching in the FX Series ADC, end users will see improved response time. In two-sided applications, content that is cached by the FX Series REM is not resent on the satellite link, which results in higher effective throughput, as well as significantly improved response times.

Some of the key benefits include:

- Reducing bandwidth consumption
- Improving user response times by delivering content quicker

Cache Differencing - Cache Differencing is only available in a two-sided implementation. It maintains identical copies of the browser's cache at the FX Series REM and on the FX Series ADC appliance. The FX Series uses intelligent differencing technology to understand what data has actually changed, and then transfers only the changed data. Serving content requests from the remote cache conserves satellite bandwidth and improves the user response time.

Traditionally, pages can be marked as cacheable and will have expiration dates. When they expire they must be retrieved from the original server, resulting in additional data being transmitted across the satellite network. Within a hub and remote site deployment, the FX Series REM caches all pages returned to the browser (even pages that are marked as non-cacheable) and performs validation when needed to ensure that no stale data is returned to the browser. When the browser asks for a page or an item that has expired or been marked as non-

cacheable, the FX Series REM sends a validation request to the FX Series ADC appliance. If the FX Series is aware of the last page the client cache contains and can compute differences in the page, it sends just the differences to an expired page or non-cached page. If the differences are too big, or if the FX Series appliance no longer has retained the last version that the client has, then the entire page is returned and subsequently cached for future possible differencing. The client in turn reconstructs the requested page, caches it, and returns it to the browser. Checksums are calculated by the FX Series ADC and verified by the FX Series REM. While this technique adds value on expired pages, it is extremely effective for dynamic page generation.

Microsoft Update Caching – The FX Series intelligently caches Microsoft® updates on the client side saving significant bandwidth attributed to “Patch Tuesday”. The FX Series caching methodology handles the complicated procedures employed by Microsoft and other vendors to request updates by requesting "partial objects". This reduces the amount of data sent over satellite links. The FX Series dramatically curbs bandwidth consumption by caching software updates published by Microsoft, Symantec, Adobe, Apple and other leading software vendors.

Most satellite service providers are aware of the bandwidth impact of “Patch Tuesday” – the day that Windows updates are distributed. The delivery of these updates is performed when software that resides on client devices downloads the new content in the background by requesting "partial content" over HTTP. The complex nature of "partial-content" HTTP requests thwarts the capabilities of most caching devices; however the FX Series REM appliance’s caching engine can handle these requests. Once the content is cached by the FX Series REM, subsequent retrievals by the updating agents that request "partial-content" will be satisfied by the FX Series, eliminating the need to repetitively transfer the same updates over satellite links.

Content distribution - Pre-Caching – When large amounts of repetitively accessed data are routinely accessed by end-users, the satellite link can become saturated with requests for the same file. In addition, users can be adversely impacted with long response times for critical data. The FX Series Pre-Caching feature helps to eliminate this bottleneck and improve end user response time.

TCP Optimization

The FX Series reduces TCP overhead from satellite links by controlling traffic congestion. For satellite-based Internet communications, TCP is used each time a user requests and receives data from a server. Unfortunately, the ability to effectively control the flow of data packets over satellite links to ensure reliable packet delivery has been a continual challenge for network administrators since the inception of TCP/IP. The real challenge is to understand how to manage the inherent bandwidth limitations within a given satellite link, and the inconsistencies, or ebbs and flows of real-time network traffic. These factors directly impact the ability to reliably and efficiently deliver traffic. To address the inefficiencies of TCP, the FX Series manages satellite links in several ways to intelligently tune TCP handshaking to reduce bandwidth consumption, ensure the reliable flow of data, and provide full utilization of the satellite link.

The FX Series provides multiple TCP congestion avoidance algorithms. These congestion avoidance algorithms deliver multi-faceted and highly effective TCP connection management to improve upon the standard TCP default setting within bandwidth delayed environments such as VSAT.

The FX Series implements Fast Start strategies to initiate connections with servers, as well as between the FX Series ADC and clients independently, so that both the server and the clients are optimized. This ability to increase data transfers without requiring additional bandwidth provides satellite service providers with a highly cost-effective solution.

Located within the satellite operations center, the FX Series ADC appliance maintains a consistent pool of TCP connections between end user clients and the servers. No changes are needed on end user devices to accomplish this and no upgrading or configuring of the servers is required. The FX Series monitors outbound packets and buffers packets from the servers by sending its own acknowledgements in place of the client’s. This dramatically speeds up file transfers, and provides an environment for fully utilizing bandwidth by minimizing the number of TCP retransmissions. When TCP acknowledgements arrive from a client, the FX Series establishes the quickest transfer rate that can be used to deliver packets to the client without overflowing the satellite link and sends packets from its buffer to the client.

Within a two-sided implementation, the FX Series uses TCP ACK Spoofing which operates on accelerated TCP streams and provides intermediate ACK's at the FX Series REM device, which relies on the RTT calculation to set TCP windows. This allows the application window to grow, taking advantage of available bandwidth and delivering data faster.

Multiplexing of Accelerated Data Streams – To remove the burden of establishing and terminating TCP connections from web and application servers, within a two-sided implementation, the FX Series utilizes Connection Multiplexing to allow satellite links and servers to handle more traffic. The FX Series connection management services provide TCP connection offloading for application-aware modules such as HTTP, HTTPS, CIFS, MAPI, POP3, SMTP, FTP and other protocols. Within a stand-alone hub deployment, the FX Series ADC reuses already established connections with back-end servers (rather than establishing new server connections), and multiplexess the connections on behalf of servers to eliminate TCP connection overhead. This reduces costly handshakes and dramatically lowers bandwidth consumption.

Traffic Prioritization

Traffic Prioritization shapes bandwidth utilization so that capacity requirements for applications are always met. By controlling the transmission of specific applications, administrators are able to keep within bandwidth limits and prioritize applications over satellite links.

QoS is used to classify and prioritize network applications based on business objectives to guarantee optimal application performance by assigning priority or bandwidth guarantees and limits for each application. Network administrators use QoS traffic management to define preferential treatment for certain classes of traffic. Hub deployments utilize QoS to provide TCP rate controls and priority levels capable of supporting many remote sites, while, remote sites use QoS to prioritize applications and user types from the remote locations.

Even with proper QoS priorities in place, applications with large file transfers, such as CIFS, FTP, and backup systems can starve real-time applications such as Voice over IP (VoIP) and streaming media (e.g. video). Even if a real-time application has priority, the bulky nature of large transfers takes too long to clear the satellite link – even when queuing and traffic shaping are enabled. The added latency that results can make a critical application such as VoIP impossible for many remote sites. To address this problem, the FX Series reduces the size of data packets, and intelligently manages packets based on satellite link capability and application profiles.

The FX Series Traffic Prioritization is a powerful feature which ensures on-time delivery of time-critical information. Through specific port assignments, priorities and policies can be assigned at the database-level, guaranteeing QoS for critical applications. The FX Series Traffic Prioritization feature allows different TCP ports to be assigned to individual applications, such as email and VoIP. These TCP ports are used in conjunction with Traffic Prioritization hardware to provide different levels of service quality. This allows administrators to define the QoS attributes used for traffic control based on either "Traffic Classes" or "Bandwidth Pools".

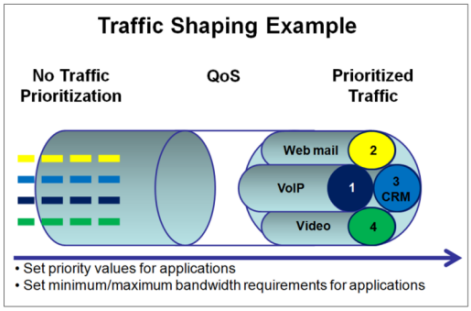

Figure 2: Traffic Shaping allows administrators to classify and prioritize traffic to improve application performance over satellite links.

Flexible Deployment and Scalability

Bandwidth Pooling - ISPs use Bandwidth Pooling to enforce a rate-limit on traffic coming from their hub out to different subnets. Bandwidth pooling reduces packet loss and addresses issues associated with carrier overload, when for example; an ISP that acquires a 50 Mbps connection from a carrier, and the ISP customer

uses 55 Mbps. While the extra 5 Mbps that the customer uses may not seem like much, that connection will incur additional packet loss. If that extra packet loss represents an additional 10%, eventually those lost packets will generate more packets on the network. Before you know it, the extra 10% packet loss will represent a 30% impact of bandwidth packet loss on the network.

The FX Series uses Bandwidth Pooling to enforce traffic limits based on the network administrator-defined throttle of the bandwidth pool so that the data going through the pool does not exceed the specified parameters. The FX Series buffers and throttles traffic to prevent packet loss and to deliver smooth traffic flow across the network connection. We provide an easy-to-use Web GUI around the complex rate shaping technology that is built into the FX Series. With the Web GUI, network personnel are able to easily configure bandwidth pools by defining a bandwidth pool and giving it a rate, adding separate subnets to that pool, and then dynamically throttle the traffic according to the specifications of the pool.

IP Source Preservation – IP Source Preservation is used to support security policies that require a specific source IP address, or range of IP addresses. IP addressing problems can occur when, for example, an end user is involved with illegal online activity and the IP address of the acceleration device is recorded in the Web server's logs. If the IP address of the acceleration device is used to make the client request to the server, it will likely be placed on a blacklist, and therefore cause considerable network problems. By spoofing the IP address of the client, the FX Series is able to avoid this problem. For example, in the event where a situation is deemed inappropriate, such as a SPAM event, the sending device source IP address will be blacklisted. To avoid this problem, the FX Series uses the organization's source IP address.

FX Series Deployment Examples

The FX Series addresses the need for ISPs to offer advanced hub and remote site optimization. In addition to the technologies described in the hub deployment, the FX Series adds additional advanced optimization and acceleration capabilities to support remote sites. Below are some examples of how the FX Series can be utilized.

- Social networking can significantly impact an ISP's satellite bandwidth within discrete user populations. To address this, an ISP can take advantage of the FX Series' advanced content reduction capabilities. For example, within a small, remote town or village, one of their subscribers downloads a video, posts the link on his Facebook page, shares it with his friends, who then share it with their friends. That activity can consume a lot of the ISPs bandwidth. Multiply that type of activity by hundreds, and even thousands of similar situations within that town, and you can see the bandwidth being quickly consumed. However, if the ISP has a FX Series REM appliance located in that village, the FX Series REM will cache the video, and subsequently serve the video locally. Additional advanced content reduction technologies will further free-up the ISPs bandwidth.
- Another example is where a small village being served by a teleport (with an FX Series ADC), has a FX Series REM appliance located in a local school system to serve the faculty and students. Additionally, the school system can set up a wireless network to allow the surrounding schools, and if they choose to, the rest of the village population, to connect to the network. Not only does this provide a reliable network service to that village, but, similar to the example above, much of the content can be served locally, and advanced content reduction will provide the ISP with significant bandwidth savings.
- For ISPs that need a WAN optimization solution to resell to their customers for deployment within their networks, the FX Series addresses a corporate customer's need to accelerate user traffic and applications between the ISP and the customer's remote sites.

Choosing the Right Solution

The solution that ensures significant bandwidth utilization, optimal throughput and flexible deployment options is one that is developed for the network environment that meets the specific technology and business needs of the organization.

Many WOC vendors are focused on selling to enterprise customers, offering products that include features designed for enterprise environments, with functionality that ISPs don't require. In contrast, the FX Series provides a single platform that supports the VSAT market with a variety of deployment options that ISPs need for advanced content reduction, network and application acceleration, traffic management, and flexible and scalable deployment capabilities.

IT managers can adapt hub technologies to meet their business needs by eliminating satellite link latency and congestion. Applying application acceleration and link optimization policies will efficiently manage bandwidth, control traffic rates and ensure reliable network service. The FX Series brings the superior advantages of efficiently optimizing hub operations and satellite links, and eliminating the adverse effects from chatty applications that add latency and slow network performance. By using the FX Series, ISPs are able to maximize satellite link bandwidth, accelerate application delivery and provide a quality user experience. The FX Series products enable satellite ISPs to:

- **Architecture** – Deploy the right WAN optimization solutions to best support the different business models within which they are deployed (the right product for the right network environment)
- **Content reduction** – Reduce the amount of data required to be stored and transmitted (bandwidth reduction)
- **Network optimization** – Reduce the number of TCP round trip requests and response turns required for transactions (network offloading)
- **Application acceleration** – Mitigate the inefficiencies of chatty protocols such as CIFS, MAPI, FTP, POP3, HTTP(S) (server offloading)
- **Traffic management** – Avoid filling satellite links with non-essential and redundant data, or oversubscribing the links (eliminate network congestion and poor network performance)
- **Deployment flexibility** – Reliably scale network infrastructure to maintain control over the network (easily make network changes, stay competitive and reduce deployment costs)

Summary

The most important aspect of acquiring the appropriate WAN optimization solution is finding one that is flexible, supports a hybrid deployment model and delivers the optimum bandwidth efficiencies and network performance to meet ISP requirements.

The FX Series has a powerful architecture and operating system built for high-performance, and delivers flexible scalability to support an ISP's need to handle their ever-increasing demands for more bandwidth and faster application delivery. In addition to FX Series' robust architecture, it utilizes full-featured advanced content reduction, bandwidth optimization, network and application acceleration and scalable deployment capabilities that ISPs need to be competitive and maintain satisfied customers.

The FX Series supports ISPs with a platform that combines flexibility, brawn and intelligence for significantly better hub and satellite link performance, scalability and OPEX savings. With power and versatility, the FX Series architecture provides the highest overall value. It enables ISPs to optimize their businesses that rely on satellite infrastructure. Their networks can run more efficiently and reliably, while lowering operational costs. The benefits of content reduction over satellite links are very tangible – bandwidth is more efficiently used and data transfer times are significantly improved.

For additional information on our FX Series products, please contact us:
U.S. & Canada: +1.800.763.3423
Outside U.S. & Canada: +1.937.291.5035