Feature Brief

Packet Flow Acceleration (PFA)

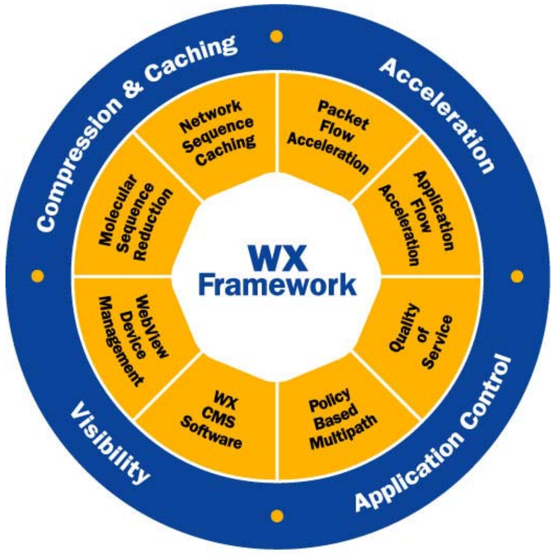
Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089 USA
408 745 2000 or 888 JUNIPER
www.juniper.net

Part Number: 210004-001 Aug. 2005

Overview

The Juniper Networks WX™ and WXC™ application acceleration platforms, two members of a larger family of solutions that improve application response times, provide IT with a range of technologies to overcome the technical limitations of WANs and speed application delivery across wide-area links. These features are implemented as part of the integrated WX Framework, which defines specific attributes that an application acceleration platform must have to overcome the bandwidth, latency, congestion and manageability issues that impede application performance over the WAN. Each element of the WX Framework addresses a specific challenge that prevents applications from running efficiently over the WAN. Those elements are organized into the following four categories:

The Integrated WX Framework



Compression and Caching:

The WX Framework includes Molecular Sequence Reduction™ (MSR™) technology, next-generation, memory-based compression that frees up WAN capacity by eliminating repeated data patterns. The MSR feature is complemented by the Network Sequence Caching technology, which uses hard disks to store and recognize large repeated patterns, even if they were sent days or weeks earlier.

Acceleration:

The acceleration component of the WX Framework includes Packet Flow Acceleration™ (PFA™) techniques, which combat the effects of latency on the TCP protocol. The Application Flow Acceleration™ (AppFlow™) technology augments that TCP acceleration with protocol-specific acceleration for applications such as Exchange, Microsoft file services, and web. The AppFlow feature pipelines multiple data blocks and web objects across the WAN, improving user productivity by reducing their wait times as multiple round trips complete.

Application Control:

The WX Framework includes Policy-based Multipath™ (Multipath™) optimization technology, which allows IT to direct specific application flows to a specific WAN link when multiple links are available. Application control features also include Quality of Service (QoS) and bandwidth-management tools for prioritizing critical application traffic and ensuring bandwidth availability.

Visibility:

The visibility components of the WX Framework include WebView device management for configuring and managing individual WX and WXC application acceleration platforms, and the WX Central Management System™ (CMS™) software for gaining visibility into and centralized control over WX and WXC platforms distributed throughout an organization.

Each of these elements interacts with one another to dynamically adjust and improve their capabilities. Working in concert, these features provide IT with the greatest degree of application acceleration and WAN optimization.

Packet Flow Acceleration (PFA)

One of the most critical issues to address in WAN transmissions is the impact that latency has on TCP-based applications. Often, when WAN bandwidth is no longer an issue, latency imposes restrictions that impact application throughput.

The WX Framework's Packet Flow Acceleration (PFA) technology features a number of techniques that accelerate the performance of TCP-based applications. These include:

- Active Flow Pipelining
- Forward Error Correction
- Fast Connection Setup

Active Flow Pipelining (AFP) is intended for applications that use high-latency links such as high- bandwidth/long-haul connections or satellite connections.

Forward Error Correction (FEC) enables the transmitting WX or WXC platform to send recovery packets along with all data packets so that if packets are lost, the receiving WX or WXC device can reconstruct them without requesting a retransmission. FEC is intended for use in high-loss environments, such as satellite connections, that do not have a link-layer error-correction scheme.

Fast Connection Setup (FCS) eliminates a round trip time for short-lived TCP connections across high-latency links. Flow Pipelining makes it possible for TCP sessions with small window sizes, such as Windows 98 or many versions of Unix, to have the same performance as TCP sessions with window sizes of 64 KB.

This feature brief provides additional information about the benefits of these specific features, the way they work, and how to interpret PFA profile results.

Active Flow Pipelining (AFP)

Active Flow Pipelining (AFP) is intended for high-latency environments, such as satellite connections, and long-haul, high-bandwidth links. In these environments, TCP slows down the transmission of data because it allows at most one receive window's worth of data to be in flight per round-trip time (RTT). For high-latency and high-bandwidth links, this window size may be significantly smaller than the capacity of the WAN pipe (in particular the capacity of that pipe enhanced by high compression).

AFP solves this problem by logically separating each TCP session crossing the WAN into three independent reliable sessions: the first between the TCP source and the WX or WXC platform near the source; the second between two WX/WXC devices; and the third between the WX or WXC platform near the destination and the destination itself. The WX/WXC platform near the source sends ACKs to the source at exactly the rate needed to continuously fill the WAN pipe with compressed data. This rate is governed by several factors, including the speed of the link, the current number of active TCP sessions using the link, and the current error rate of the link.

The data is then transferred from the source WX/WXC device across the WAN to the destination device using the WX Framework's reliable transport protocol. Congestion is managed by outbound QoS, and packet order is preserved.

Once the data arrives at the destination device, it is delivered to the targeted recipient using TCP. The entire operation of terminating TCP at the source, using the WX Framework's reliable transport protocol for transmission across the WAN, and delivering the data via TCP to the destination is completely transparent to the TCP hosts.

AFP Technical Overview

Figure 1 illustrates the operation of AFP. In this figure, all TCP connections from the TCP sender and receiver are locally terminated on WX/WXC platforms P1 and P2. A reliable IP tunnel is built between the two devices. The WX/WXC platforms know the bandwidth of the WAN circuit and send data across the link at exactly this speed. In turn, P1 accepts data from the sender at exactly the speed necessary to continuously fill the WAN pipe at its current effective capacity.

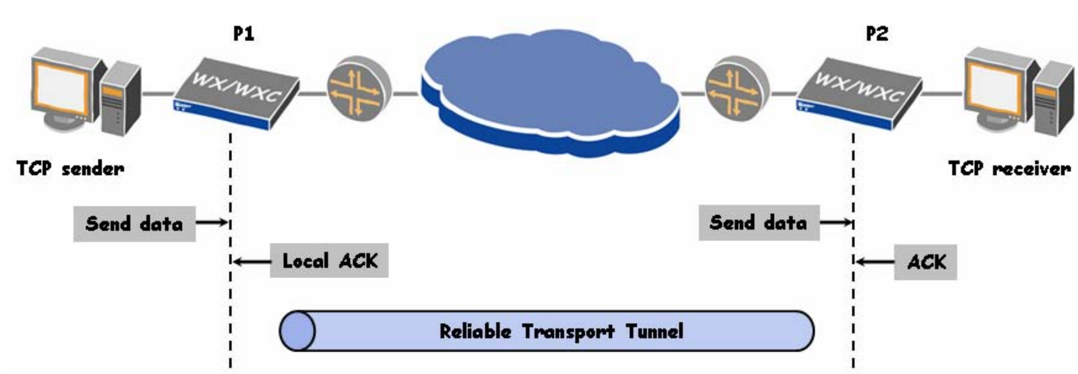


Figure 1: Active Flow Pipelining

Since the TCP connections are now locally terminated, all data ACKs are sent at LAN rates, enabling the TCP senders to very quickly use all available bandwidth. And since ACKs are sent at the rate governed by the WAN link speed, the full bandwidth of the WAN circuit is utilized, even in the case of very high latency. This mechanism enables the WX/WXC platforms to ensure that the TCP hosts send at the maximum possible rate given the WAN circuit size.

From the moment the data is acknowledged by the WX/WXC device at the sender – P1 – a system of guarantees ensures that the data is delivered reliably and in order through the reliable tunnel and to the destination. The guarantees depend on buffers at the source WX/WXC platform, a fast loss-detection and retransmission request mechanism at the destination platform, and the use of TCP between the destination WX/WXC device and the destination itself.

In the case of an error-prone link, Forward Error Correction (FEC) may optionally be employed to protect against data loss. This mechanism can improve performance by significantly reducing the number of times the receiving device – P2 – detects a lost packet and requests a retransmission (see the FEC section below).

The net effect of these features is that the TCP sender and receiver are shielded from loss and latency issues on the WAN, and they are reliably and transparently given the maximum available bandwidth the WAN circuit can provide.

Applications That Can Benefit from AFP

AFP is intended for non-chatty applications running over high-latency links. High latency exists when the product of bandwidth*delay of the WAN exceeds the TCP receive window size of the host. Put more simply, if TCP hosts don't have a large enough window to fill the WAN pipe, the link has high latency. Examples of applications that can benefit from AFP include CAD/CAM file sharing, source control replication, FTP, and any other non-chatty application.

Many WANs suffer from the "long, fat pipe" problem. As the distance or WAN latency increases and the bandwidth increases, TCP applications become throughput-bound based on the overall bandwidth*delay product of the WAN (see RFC1323). AFP helps minimize this problem by putting more data in flight than the TCP senders otherwise would. As a result, TCP receivers do not have to wait as long for packets to traverse the WAN and application performance is increased. The higher the WAN link's latency, the higher the potential AFP performance gain.

In general, AFP improves performance if the product of the effective bandwidth and latency (the bandwidth*delay product) exceeds the TCP window size. Latency is measured by the circuit's round-trip time. Note that 64 KB is the typical TCP window size for Windows 2000 and later. However, for Windows 98, the TCP window size is 16 KB.

Let's look at an example for calculating the bandwidth*delay product. A 1.5 Mbps T-1 link, with 50 percent data reduction due to compression and latency of 200 ms, has a bandwidth*delay product of:

```
(3,088,000 bps * 0.200 seconds)= 617,600 bits or 617,600/8 =
77,200 bytes
```


In this case, AFP will improve performance since the bandwidth*delay product is greater than 64 KB, the maximum TCP window size.

Another example of the dramatic impact AFP can have on high-bandwidth links is a long-haul, 45 Mbps T-3 link. On that link, if the latency is 50 ms and a 50 percent data reduction doubles the effective bandwidth, the bandwidth*delay product is:

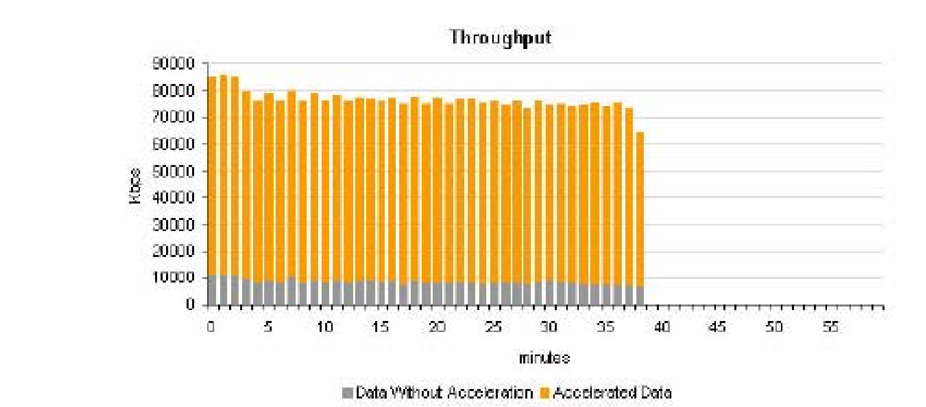
```
(90,000,000 bps * 0.050 seconds)/8 = 562,500 bytes
```

In this case, AFP will dramatically improve bandwidth available for the TCP host. If the TCP host has a 64 KB window size, using AFP will increase the available WAN bandwidth more than 8x. Juniper Networks has developed a bandwidth calculator to help determine the benefit of AFP and MSR on high-latency links.

Analyzing Results

Figure 2 shows an example of an Active Flow Pipelining traffic analysis. Traffic is grouped by application (this case shows just one application – FTP). For each application, the total number of TCP sessions is displayed. The total TCP sessions are further broken down into the number of sessions eligible for acceleration and the total amount of traffic sent for those sessions. These are sessions which have the potential to benefit from AFP.

In inline mode, the WX/WXC device takes a measurement of the actual session throughput achieved by AFP. Analytical tools then estimate what the session throughput would have been without AFP enabled. (Note that this estimate is just that; the analytical tools cannot determine exactly what the throughput would have been with 100 percent accuracy.) From these two numbers, an estimated acceleration factor can be derived. In this example, AFP increases the available bandwidth for the FTP session from just under 10 Mbps to faster than 75 Mbps.



Application	Total TCP Sessions (count)	Accelerated Sessions (count)	Traffic (MB)	Average Accelerated Session Throughput (Mbps) Actual	w/o Accel*	Estimated Acceleration Factor
FTP	1	1	21,344.45	76.52	8.30	9.2 X

Figure 2: Active Flow Pipelining sample results

Forward Error Correction (FEC)

Forward Error Correction (FEC) enables the sending WX or WXC platform to send recovery packets along with all data packets so that, in case packets are lost, the receiving WX/WXC device can reconstruct them without requesting a retransmission. IT can specify the number of recovery packets to be sent per block of data packets.

The idea is that for every "n" packets, a recovery packet is created that contains material that can be used to reconstruct any of the "n" packets that might be lost. All that is needed for reconstruction is for the packets around the lost packet to be available. The recovery packet can be used to reconstruct the lost packet(s) based on information contained in the surrounding packets.

Unlike some FEC schemes, the Packet Flow Acceleration technique's FEC has minimal impact on forwarding performance. Some FEC schemes operate on blocks of data, creating additional latency because they require data packets to be held for as long as it takes for the entire block of packets — including the recovery packets — to arrive. The PFA FEC scheme introduces no delay when there is no loss occurring, and lost packets are recovered as soon as the recovery packets for the given block of data arrive.

FEC is intended for use in high-loss environments, such as satellite connections.

Analyzing Results

Figure 3 shows an example of the FEC recovery statistics. For each WX/WXC endpoint, the monitoring tools display a count of the total number of FEC-enabled packets. The tools also show a count of the total number of lost packets.

The tool then displays how many of the lost packets were recovered without re-transmission and how many were not recovered and therefore required retransmission. IT can look at these results on a per-device basis or aggregated across multiple WX or WXC platforms.

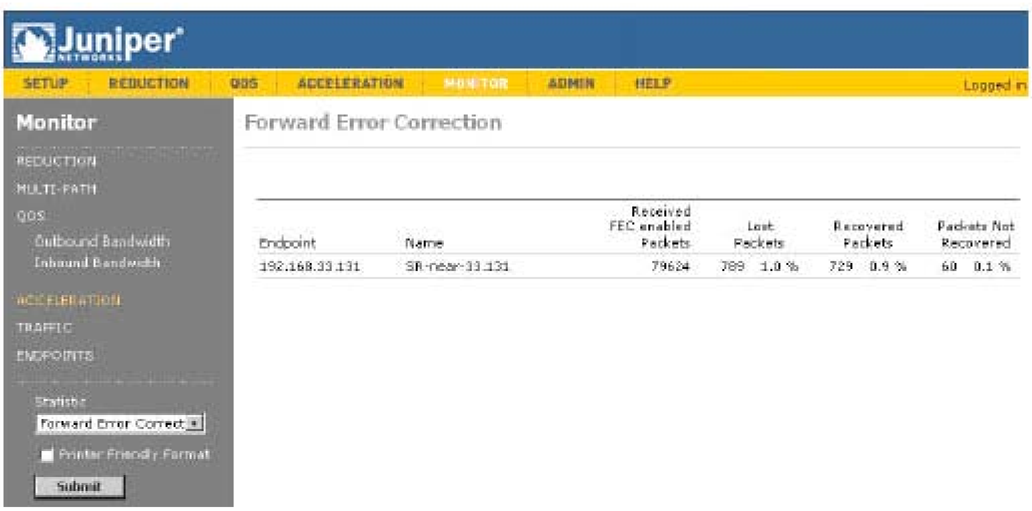


Figure 3:   FEC statistics

Fast Connection Setup (FCS)

Fast Connection Setup (FCS) is for applications that use TCP connections requiring 10 or fewer round trips. For each TCP connection, one round-trip time (RTT) is eliminated. FCS should be enabled *only* for applications that use short-lived TCP connections. FCS Technical Overview

Figure 4 illustrates how FCS works. In this figure, the TCP receiver initiates the three-way handshake that starts the TCP connection by sending a SYN packet to the TCP sender. As soon as the SYN packet is seen by the P1 WX/WXC device, P1 sends a local SYN/ACK to the TCP receiver. This SYN/ACK stimulates the TCP receiver to send the final ACK along with the initial request for data to the TCP sender. At the same time, the original SYN packet is in flight across the WAN. As a result, the three-way handshake process is accelerated by one RTT.

The TCP sender receives the SYN and immediately sends a SYN/ACK back to the TCP receiver. Shortly after this transmission, the TCP sender receives the three-way handshake ACK and the request for data. The TCP sender now puts its data in flight. When the SYN/ACK sent by the TCP sender arrives at P1, instead of forwarding the SYN/ACK on to the TCP Receiver, P1 consumes that SYN/ACK. Remember — the TCP receiver already received a SYN/ACK from P1. Finally, the data that the TCP sender has put in flight is received by the TCP receiver.

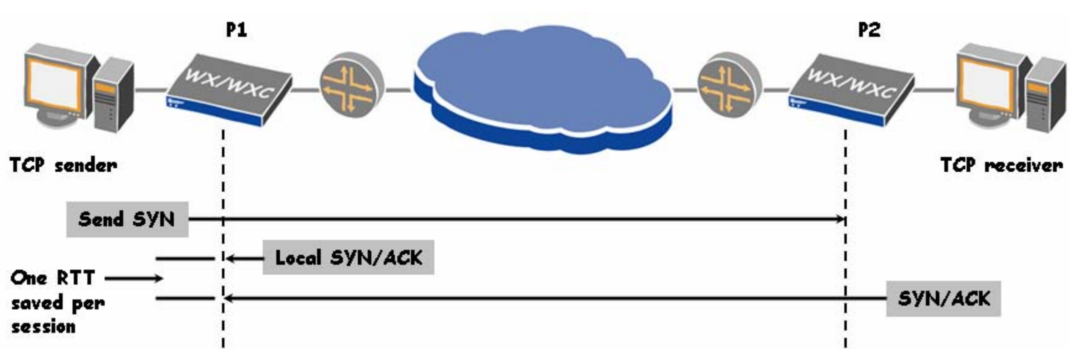


Figure 4: Fast Connection Setup (FCS)

The entire transmission has been sped up by one round-trip time (RTT), resulting in an acceleration of 10 percent to 50 percent for short-lived (fewer than 10 RTTs) TCP sessions.

Applications That Can Benefit from FCS

FCS is particularly useful in pre-Windows 2000 environments. Pre-Windows 2000 clients use NetBIOS, rather than CIFS, for file transfer. NetBIOS is a much chattier protocol than CIFS, so it establishes many short-lived TCP connections.

FCS is also beneficial for HTTP traffic, providing the greatest benefit to HTTP 1.0 traffic (pre-Windows 2000) since that protocol creates many more short-lived TCP connections than HTTP 1.1 (although even HTTP 1.1 should see some benefit).

FCS can also be useful for custom, home-grown enterprise WAN applications that may not have been designed with the WAN in mind. Many custom applications are developed on the corporate LAN and tend to perform poorly when moved to the WAN. If the home-grown application creates many short-lived TCP connections, it can benefit from FCS.

Fast Connection Setup's advantages are best seen in high-latency environments such as satellite networks, because the one saved RTT results in the greatest time savings when latency is high. If latency is low, FCS will not provide much benefit since the time savings are so minimal.

Summary

Packet Flow Acceleration features can significantly improve application performance over the WAN, especially on high-latency connections. The key to understanding the degree of benefit is to understand the way the features work and the type of applications that can benefit from each feature.