



Optimize WAN and LAN Application Performance with TCP Express

For enterprises delivering Internet and extranet applications, TCP/IP inefficiencies, coupled the effects of WAN latency and packet loss, all conspire to adversely affect application performance. The result of these inefficiencies has inflated the response times for applications, and significantly reduced bandwidth utilization efficiency (ability to “fill the pipe”).

White Paper
by F5



WHITE PAPER

Optimize WAN and LAN Application Performance with TCP Express

Overview

For enterprises delivering Internet and extranet applications, TCP/IP inefficiencies, coupled with the effects of WAN latency and packet loss, conspire to adversely affect app performance. The result? Inflated response times for applications, and significantly reduced bandwidth utilization efficiency (the ability to "fill the pipe").

F5's BIG-IP Local Traffic Manager provides a state-of-the-art TCP/IP stack that delivers dramatic WAN and LAN application performance improvements for your real-world network—not packet-blasting test harnesses that don't accurately model actual client and Internet conditions.

This highly optimized TCP/IP stack, called TCP Express, combines several cutting-edge TCP/IP techniques and improvements in the latest RFCs. Numerous improvements and extensions developed by F5 minimize the effect of congestion and packet loss and recovery. Independent testing tools and customer experiences show TCP Express delivers up to a 2x performance gain for end users, and a 4x improvement in bandwidth efficiency with no change to servers, applications, or the client desktops.

Solution

TCP Express: F5's Optimized TCP Stack

F5's TCP Express is a standards-based, state of the art TCP/IP stack that leverages optimizations natively supported in various client and server operating systems, and optimizations that are not operating-system specific. F5's TCP/IP stack contains hundreds of improvements that affect both WAN and LAN efficiencies, including:

- **For high-speed LANs:** F5's TCP stack quickly expands buffer sizes and detects low latency to manage congestion.
- **For low-speed WANs:** F5's TCP stack detects client speed and estimates bandwidth to limit packet loss and recovery in the case of dropped packets.

At the heart of the BIG-IP Local Traffic Manager is the TMOS architecture that provides F5's optimized TCP/IP stack to all BIG-IP platforms and software add-on modules. These unique optimizations, which extend to clients and servers in both LAN and WAN communications, place F5's solution ahead of packet-by-packet systems that can't provide comparable functionality—nor can they approach these levels of optimization, packet loss recovery, or intermediation between suboptimal clients and servers.



WHITE PAPER

Optimize WAN and LAN Application Performance with TCP Express

The combination of F5's TMOS full proxy architecture and TCP Express dramatically improves performance for all TCP-based applications. Using these technologies, BIG-IP has been shown to:

- Improve transfer rates for all connecting client types
 - 79% performance boost on average for broadband users
 - 35% performance boost on average for dial-up clients
- Improve connection reliability for dial-up clients
 - 56% reduction on average in TCP/IP errors (mostly TCP timeouts)
- Increase bandwidth efficiency across existing ISP providers
 - 224% increase of data placed on the wire (3.2x improvement)
 - 50% packet reduction on the wire (2x improvement)
 - Eliminated 63% of "empty" TCP packets (2.7x improvement)

The following sections describe the TMOS enabling architecture, as well as a subset of the standard TCP RFCs and optimizations that TCP Express uses to optimize traffic flows. Because there is no one-size-fits-all solution, this paper also describes how to customize TCP Profiles and handle communications with legacy systems.

TMOS Architecture and TCP/IP Stack Brokering

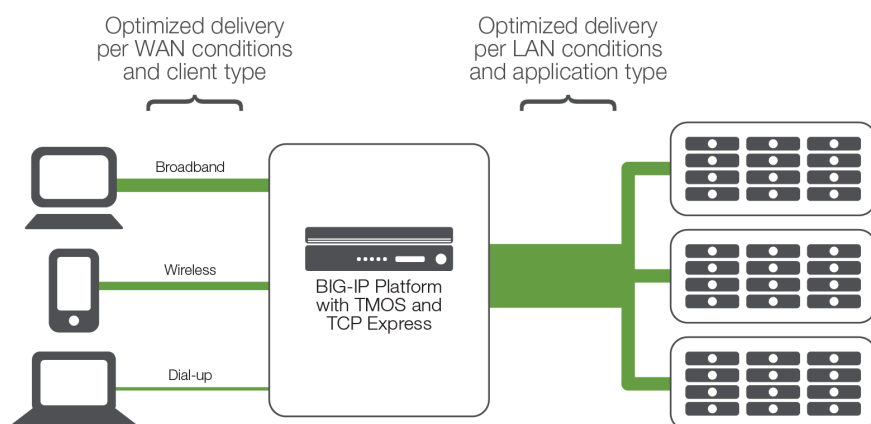
Most organizations don't update server operating systems often, and some applications continue to run on very old systems. This legacy infrastructure can be the source of significant delays for applications as they are delivered over the WAN. The BIG-IP Local Traffic Manager with TCP Express can shield and transparently optimize older or non-compliant TCP stacks that may be running on servers within a corporate data center. This is achieved by maintaining compatibility with those devices, while independently leveraging F5's TCP/IP stack optimizations on the client side of a connection—providing fully independent and optimized TCP behavior to every connected device and network condition.

As a full proxy that bridges various TCP/IP stacks, TMOS is a key enabler to many of the WAN optimizations included in F5's unique TCP Express feature set. Client and server connections are isolated, controlled, and independently optimized to provide the best performance for every connecting device.

The BIG-IP Local Traffic Manager eliminates the need for clients and servers to negotiate the lowest common denominator for communications. It intermediates on behalf of the client (called Stack Brokering) and uses TCP Express to optimize client-side delivery while maintaining server-optimized connections on the inside of the network as shown in the following figure.

WHITE PAPER

Optimize WAN and LAN Application Performance with TCP Express



Often, organizations don't have the resources—or don't need—to remove or replace their legacy servers and applications. To accommodate these systems, the BIG-IP Local Traffic Manager provides mediation to translate between non-optimized or even incompatible devices, including:

- Maintaining separate MSSs for clients and servers to ensure both are transmitting data at an optimal rate. Clients and servers communicate the MSS to determine the maximum amount of data that may be put in a segment of TCP communication. The two parties negotiate the MSS in an attempt to create the most compatible communication, but this often leaves either the client or server not optimized by forcing them to negotiate the lowest common MSS value.
- Maintaining optimizations like TCP Selective Acknowledgements (SACK) or TCP Timestamps (and much more) to clients when they connect to non-supported servers.
- Dynamically and automatically optimizing TCP window sizes and TCP congestion information for each connected device (every client and every server).
- Maintaining interoperability for stacks such as Windows to older Solaris systems that will not interoperate with a TCP FIN-PUSHs. This is just one example of the various stack interoperability issues that often challenge business trying to server a broad user population.

In addition to improving WAN communications, the BIG-IP Local Traffic Manager translates these capabilities across the entire infrastructure by acting as a bridge or translation device between all clients and back-end servers. The net result is that the BIG-IP Local Traffic Manager improves performance while masking inefficiencies in the network. This reduces cost and complexity by eliminating the need to update and tune every client and every server.



F5 Improvements on TCP/IP RFCs

Some of the most important F5 TCP/IP improvements include:

- Client acceleration and error avoidance
- Link utilization improvements
- Customizable TCP controls

These improvements were made to industry-standard RFCs. The following sections highlight some of the key RFCs in TCP Express.

Client Acceleration and Error Avoidance RFCs

- **Delayed Acknowledgements, Nagle Algorithm** (RFC 896, 1122) Enables the BIG-IP Local Traffic Manager to accelerate data delivery by reducing the number of packets that must be transmitted. Delayed Acknowledgements provides a standard mechanism for deciding when acknowledgement packets need to be sent to help reduce redundant acknowledgement packets. In addition, the Nagle Algorithm provides a standard procedure for coalescing many smaller packets into fewer, larger packets.
- **Selective Acknowledgements** (RFC 2018, RFC 2883) Enables the BIG-IP Local Traffic Manager to more effectively and quickly deal with lost and reordered packets on WANs and lossy networks. This is enabled by default on Windows XP and later for Internet communications, as well as all other modern TCP stacks. Extensions enable specifying the use of a SACK option or acknowledging duplicate packets.
- **Explicit Congestion Notification ECN** (RFC 3168, 2481) Enables the BIG-IP Local Traffic Manager to proactively signal peers that intermediate routers are being overloaded so that they can back off and avoid packet loss. The reserved flags in the TCP header (ECE and CWR) can be used to communicate congestion back to the peer.
- **Limited and Fast Retransmits** (RFC 3042, RFC2582) Enables the efficient retransmission of lost data, which can eliminate the effects of timeouts from packet loss.
- **Adaptive Initial Congestion Windows** (RFC 3390) Mitigates the cost of TCP slow start congestion avoidance. Studies on larger initial congestion windows have shown a 30% gain for HTTP transfers over satellite links and a 10% improvement for 28.8 bps dialup users with no accompanying increase in the drop rate. With TCP connections that are sharing a path (16 KB transfers to 100 Internet hosts), an increase in window segment resulted in roughly a 25% improvement for transfers using the four segments (512 byte MSS) initial window when compared to an initial window of one segment.

Link Utilization Improvements RFCs

- **TCP Slow Start with Congestion Avoidance** (RFC 3390, 2581) This is a method of converging on the right amount of data to put on the link without clogging it up so that packets don't get dropped. This capability helps organizations increase bandwidth utilization, realizing higher throughput rates



WHITE PAPER

Optimize WAN and LAN Application Performance with TCP Express

on their existing public Internet connections and leased lines.

- **Bandwidth Delay Control** (RFC 793, RFC 2914, RFC 1257) F5 has improved and expanded on Bandwidth Delay Calculation to more accurately estimate the optimal load to put on the network without exceeding it.

TCP Extension RFCs

- **TimeStamps** (RFC 1323) The BIG-IP Local Traffic Manager allows for selective use of timestamps that add data to the TCP segment to aid with other optimizations. While the benefit of Timestamps is great over a modern network, some legacy routers and NATing devices zero out or do not update the timestamps, negating the benefit. As such, these and other capabilities can be tuned on a per profile basis.
- **Improve TCP TIME-WAIT Assassination Hazards** (RFC 1337) There are a few possible communication errors that can be avoided by optimizing TIME-WAIT behavior, especially avoiding action when receiving reset segments while in the TIME-WAIT state.
- **Defending Against Sequence Number Attacks** (RFC1948) TCP Express blocks most sequence number guessing attacks, using secure ISN generation.
- **Improve TCP Congestion Management** (RFC 3168) TCP Express implements all the latest TCP congestion avoidance and congestion recovery methods available on the Internet today to increase usable bandwidth and speed recovery in case of congestion.
- **Improve TCP Slow-Start for TCP with Large Congestion Windows** (RFC 3742) Uses a more conservative slow-start behavior to prevent large amounts of loss when connections use very large TCP windows.
- **Appropriate Byte Counting** (RFC 3465) Uses the number of previously unacknowledged bytes each ACK used to provide more intelligent window scaling and increase TCP performance.
- **Improve TCP Fast Recovery Algorithm (NewReno)** (RFC 3782) The NewReno modification to TCP's Fast Recovery Algorithm specifies a slight modification, whereby a TCP sender can use partial acknowledgements to make inferences that determine the next segment to send in situations where SACK would be helpful, but isn't available.

Collective Performance Improvement

Because TCP Express implements literally hundreds of real-world TCP interoperability improvements, and fixes or provides a workaround to commercially-available product stacks (Windows 7 and up, IBM AIX, Sun Solaris, and more), no single optimization technique accounts for the majority of the performance improvements. These optimizations are dependent on specific client/server type and traffic characteristics. For example:

- With broadband where there is much more bandwidth, most TCP is naturally less efficient at utilizing the full link capacity, so the BIG-IP offers additional optimizations.



WHITE PAPER

Optimize WAN and LAN Application Performance with TCP Express

- With a client on dial-up, the key advantages of TCP Express is that BIG-IP can reduce the total number of packets transmitted for a given transaction, as well as providing for faster retransmissions.

The BIG-IP still reduces packet round trips and accelerates retransmits just like dial-up, but with faster connections. The BIG-IP Local Traffic Manager and TCP Express also optimizes congestion control and window scaling to improve peak bandwidth. Although improvements for dial-up users may be the most noticeable, improvements for broadband users are the most statistically obvious because of how dramatically some enhancements improve top-end performance on faster links.

As a general rule, the more data that is exchanged, the more bandwidth optimizations apply. The less data that is exchanged, the more Round-Trip delay Time (RTT) optimizations apply. Therefore, traffic profiles that don't exchange a lot of data, such as dial-up, would see more optimization than broadband. For traffic profiles that do exchange a lot of data, broadband would see the most optimization. In both cases, significant gains can be realized using TCP Express.

Customizable TCP Controls

While TCP Express is automatic and requires no modifications, the BIG-IP Local Traffic Manager gives users advanced control of the TCP stack to tune TCP communications according to specific business needs. This includes the ability to select optimizations and settings at the virtual server level, per application being fronted on the device. Administrators can use a TCP profile to tune each of the following TCP variables:

- time wait recycle
- delayed acks
- proxy mss
- proxy options
- deferred accept
- selective acks
- ecn
- limited transmit
- rfc1323
- slow start
- bandwidth delay
- nagle
- proxy buffer



Administrators can also use these controls to tune TCP communication for specialized network conditions or application requirements. Customers in the mobile and service provider industries find that this flexibility gives them a way to further enhance their performance, reliability, and bandwidth utilization by tailoring communication for known devices (like mobile handsets) and network conditions.

Tuning Stack Settings for Applications

TCP Express provides flexible stack settings to optimize custom services—for example, you can adjust these settings to optimize an ASP application delivered to mobile users. The following table describes the BIG-IP Local Traffic Manager modifiable stack settings.

Setting	Description
Recv window 65535	The BIG-IP Local Traffic Manager's default receive window is 16384. This can cause certain TCP stacks to 'throttle' (slow down) when communicating with the BIG-IP LTM. Setting it to 65535 results in reduced time to last byte (TTLB) at the potential expense of more memory utilization.
Send buffer 65535	Increasing the BIG-IP Local Traffic Manager's default send buffer to 64K enables more data to be put on the network at a time if the congestion window allows it, at the potential expense of greater memory utilization.
Proxy buffer high and low 128K	F5 has empirically found these modified defaults offer better real-world performance for most sites based on average page sizes. These values control the amount of data the BIG-IP Local Traffic Manager receives from the server for Content Spooling. This comes at the expense of potentially increased memory utilization.

Optimizing Highly Interactive Traffic over the LAN

If the traffic on a LAN is highly interactive, F5 recommends a different set of TCP settings for peak performance. F5 has found that Nagle's algorithm works very well for packet reduction and general compression/RAM caching over a WAN. In addition, tweaks to various buffer sizes can positively impact highly-interactive communications on low-latency LANs, with the only possible cost being increased memory utilization on the BIG-IP Local Traffic Manager.

The following table describes modifiable TCP profile settings.

Setting	Description
---------	-------------



WHITE PAPER

Optimize WAN and LAN Application Performance with TCP Express

bandwidth delay disable	Disables bandwidth limiting. On real-world networks, TCP stacks often push so much data onto the network that drops occur. Bandwidth delay product limiting causes the BIG-IP Local Traffic Manager to determine the optimal amount of data to inject into the network per RTT, and does not exceed this. In case the traffic profile has lots of tiny objects, the client exhibits a "stretch ACK violation bug" (older Linux kernels do this, for example), and this parameter should be disabled.
Nagle disable	Nagle's algorithm holds data until the peer ACKs up to prevent putting tiny packets on the network. Enabling it results in better real-world performance over the WAN, but can make it look as though the BIG-IP Local Traffic Manager is imposing latency on a LAN due to holding packets less than the TCP MSS until the peer ACKs outstanding data.
Ack on push enable	<p>Causes the BIG-IP Local Traffic Manager to immediately send a TCP acknowledgement when a TCP packet with the PSH flag is received. Increases peak bandwidth when transferring large files to/from Windows machines on a LAN.</p> <p>BIG-IP administrators can flexibly tune the number of ingress/egress TCP ACK packets to reduces ingress ACKs by sending fewer PUSH flags. This address TCP protocol designed-in operation to send an immediate ACK to a PUSH segment even if delayed ACKs or stretch ACKs are enabled. Four different modes for sending PUSH flags are enabled: Default, None, One and Auto. This flexibility provides administrators with control over the frequency of PUSH segments sent</p>
Recv window 65535	The BIG-IP Local Traffic Manager's default receive window is 16384. This can cause certain TCP stacks to 'throttle' (slow down) when communicating with the BIG-IP LTM. Setting it to 65535 results in reduced time to last byte (TTLB) at the potential expense of more memory utilization.
Send buffer 65536	Increases the BIG-IP Local Traffic Manager's default send buffer to 64K, which enables more data to be put on the network at a time, if the congestion window allows it at the potential expense of greater memory utilization.
Proxy buffer high and low 128K and 96K respectively	F5 has found these modified defaults offer better real-world performance for most sites based on average page sizes. These values control the amount of data the BIG-IP Local Traffic Manager receives from the server for content spooling. This comes at the expense of potentially increased memory utilization.
Slow start disable	Usually not required, but if measuring time to last byte (TTLB) on a LAN, disabling slow start can have a small, but positive impact on reducing latency.



Other F5 Acceleration Technologies

TCP Express is complemented by other F5 acceleration features and products that work to further reduce user download times and optimize infrastructure resources.

Other acceleration features that are integrated with the BIG-IP Local Traffic Manager include:

- **HTTP Compression** uses highly configurable GZIP compression capabilities to reduce the size of bytes transferred over a line.
- **Fast Cache** offloads servers and saves server CPU by caching priority applications and extending control for hosting multiple applications on a shared system. Cache storage of compressed data generates even faster content delivery and improves BIG-IP scalability.
- **OneConnect** increases server capacity up to 60% by offloading TCP connections from servers.
- **Content Spooling** reduces the TCP overhead on servers to increase server capacity up to 15% by lowering the amount TCP segmentation that must be performed on servers.

Conclusion

For organizations looking to improve the capacity and performance of their infrastructure, the BIG-IP Local Traffic Manager provides a unique solution that transparently makes every connecting client and server work more efficiently. F5's unique TCP Express delivers unmatched, real-world network and application performance improvements, and offers organizations an unprecedented level of control to optimize TCP communications for mission-critical applications.

F5 Networks, Inc.
401 Elliott Avenue West, Seattle, WA 98119
888-882-4447 f5.com

Americas
info@f5.com

Asia-Pacific
apacinfo@f5.com

Europe/Middle-East/Africa
emeainfo@f5.com

Japan
f5j-info@f5.com