Experience the Network as Your Customers Do—Closing the Turn-up Gap

Introduction

Traditionally, Layer 2/3 turn-up tests such as RFC 2544 have been conducted when installing Ethernet services. After providers "certify" their networks with either an RFC 2544 test (or even the new Y.1564 test), they can still receive complaints of poor application performance from business-end customers using video conferencing, YouTube, Facebook, or cloud-based applications.

The gap in installation testing, namely the omission of transmission control protocol (TCP)-layer testing, which is key to optimal end-customer application layer performance, is the cause for this disconnect. Figure 1 portrays a simplified view of the protocol stack and the gap between current turn-up testing methodologies and the end-user experience.

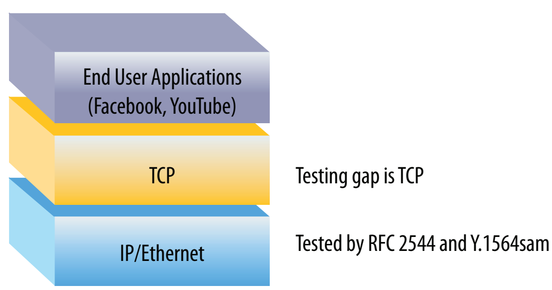

Figure 1. Simplified protocol stack and the gap between turn-up testing and the end-user experience

This testing gap does not let network providers experience network performance like their customers, so they need a solution that can verify TCP-layer performance before end-customer activation. Testing at the TCP layer can eliminate additional truck rolls, tech support calls, and customer churn which provides substantially positive implications to providers' operating expenses (OpEx).

This white paper:

- briefly introduces the TCP protocol
- summarizes some common customer-premises equipment (CPE) and network issues that can adversely affect TCP and application performance
- introduces the new IETF RFC 6349 TCP test methodology
- demonstrates the benefits to network providers who conduct RFC 6349-based TCP-layer installation testing.

Network and CPE Issues that Adversely Affect TCP

TCP operates at open system interconnection (OSI) Layer 4 and resides on top of the IP Layer 3. One of the most important aspects of TCP is that it is reliable; and, if packets are dropped, TCP will ensure that they are retransmitted to the receiver.

Additionally, on a wide-area network (WAN) link TCP must be properly configured to adjust the number of bytes the sender can transmit before receiving an acknowledgment (ACK) from the receiver. This number of bytes "in-flight" is commonly referred to as the TCP Window; although, in reality, there are several TCP-window mechanisms at work.

Figure 2 depicts the concept of TCP in-flight data bytes on a 45 Mbps WAN link with 25 ms round-trip delay (RTD), or latency.

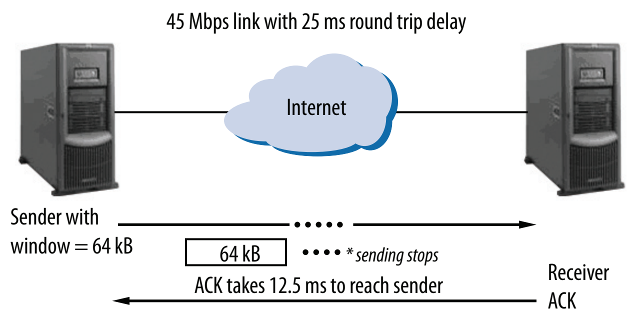

Figure 2. Illustration of TCP in-flight data bytes on a 45 Mbps WAN link with 25 ms RTD

In Figure 2, the TCP window is improperly tuned and only 64 kB are transmitted from the sender before requiring an ACK.

As RFC 6349 describes, the bandwidth delay product (BDP) is the optimum TCP window, calculated as:

$$\text{BDP} = \frac{\text{link bottleneck bandwidth x round-trip time}}{8}$$

In this example, the BDP would be 140 kB, which is more than twice the size of the sender's 64 kB window and the sender would only achieve about 20 Mbps throughput.

Another key attribute of TCP is that it is bursty rather than a constant bit rate. So a Gigabit Ethernet (GigE) local area network (LAN) on a 100 Mbps WAN will result in several instances where the WAN network improperly handles GigE "bursts," causing dropped packets and TCP retransmissions. Higher network latency dramatically affects TCP throughput, because the TCP state machine must rise up to the optimal transmission rate (without packet loss).

The primary means to "downshift" TCP from a LAN to a WAN are buffering and traffic shaping. Figure 3 illustrates network buffering and the output queue of a network device. The output queue prioritizes traffic based upon various quality of service (QoS) mechanisms, such as differentiated services code point (DSCP), virtual LAN (VLAN) tags, and others, and also allocates queue depths per traffic class. Using default queue depths can drop bursty TCP traffic. The packets drops cause TCP retransmissions which can seriously degrade end-user experience.

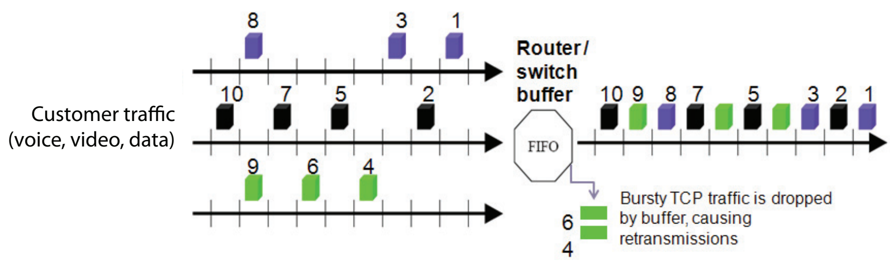

Figure 3. Network buffering and output queue of a network device

The second means to downshift from a LAN to a WAN is traffic shaping or "intelligent" network buffering, where the network device shapes the traffic according to the committed information rate (CIR). Traffic shaping should be performed at the CPE edge device, but network providers also can shape traffic to substantially benefit TCP performance and the end-customer experience.

By not shaping TCP traffic as it downshifts from a higher speed interface to a lower speed, network policers can detrimentally affect TCP performance. Contrary to shaping, policing chops excess traffic above the CIR, causing TCP retransmissions and seriously degrading end-user performance. Figure 4 contrasts the function of a traffic shaper versus a policer.

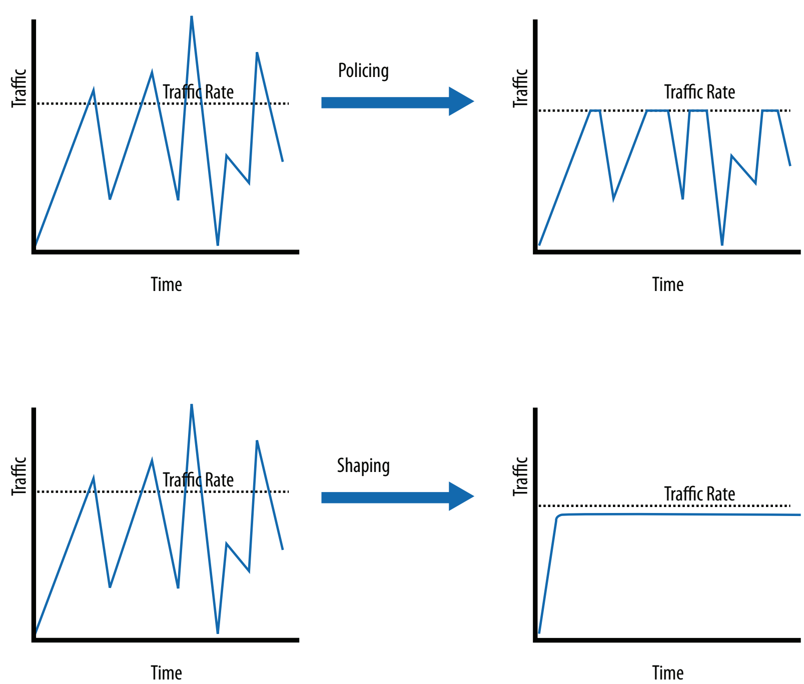

Figure 4. Contrasting traffic shaping versus policing

Using RFC 6349 test methods before customer activation to tune buffer queuing and traffic shaping will optimize TCP throughput over the WAN.

RFC 6349 TCP Test Methodology

RFC 6349 is the new, innovative TCP throughput test methodology that JDSU co-authored along with representatives from Bell Canada and Deutsche Telecom. Recently issued by the Internet Engineering Task Force (IETF) organization, RFC 6349 provides a repeatable test method for TCP throughput analysis with systematic processes, metrics, and guidelines to optimize the network and service performance.

RFC 6349 recommends always conducting a Layer 2/3 turn-up test before TCP testing. After verifying the network at Layer 2/3, RFC 6349 specifies conducting the following three test steps.

- **Path MTU detection (per RFC 4821)** to verify the network maximum transmission unit (MTU) with active TCP segment size testing to ensure that the TCP payload remains unfragmented.
- **Baseline round-trip delay and bandwidth** to predict the optimal TCP Window size for automatically calculating the TCP BDP.
- **Single and multiple TCP connection throughput tests** to verify TCP Window size predictions that enable automated "full pipe" TCP testing.

TCP retransmissions are normal phenomena in any TCP/IP network communication. Determining the number of retransmissions that will impact performance is difficult when simply using the number itself. RFC 6349 defines a new metric to gain insight into the relative percentage of a network transfer that was used due to the retransmission of a payload.

This metric is the TCP Efficiency metric, or the percentage of bytes not retransmitted, and is defined as:

$$\frac{\text{transmitted bytes} - \text{retransmitted bytes}}{\text{transmitted bytes}} \times 100$$

Transmitted bytes are the total number of TCP payload bytes transmitted including the original and retransmitted bytes. This metric provides a comparison between various QoS mechanisms such as traffic management, congestion avoidance, and various TCP implementations such as Windows XP and Linux to name a few.

For example, if 100,000 bytes were sent and 2,000 had to be retransmitted, the TCP Efficiency would be calculated as:

$$\frac{102{,}000 - 2{,}000}{102{,}000} \times 100 = 98.03\%$$

Note that packet loss percentages at Layer 2/3 do not directly correlate to retransmission percentages of bytes because the distribution of the packet loss can widely affect the manner in which TCP retransmits. The TCP Efficiency metric allows network providers to establish a TCP loss threshold for various class-of-service (CoS) levels.

RFC 6349 also defines the Buffer Delay Percentage, which represents the increase in round-trip time (RTT) during a TCP throughput test from the baseline RTT, which is the RTT inherent to the network path without congestion.

The Buffer Delay Percentage is defined as:

$$\frac{\text{average RTT during transfer} - \text{baseline RTT}}{\text{baseline RTT}} \text{ x } 100$$

For example, use the following formula to calculate the Buffer Delay Percentage of a network with a baseline RTT path of 25 ms that increases to 32 ms during an average RTT TCP transfer:

$$\frac{32 - 25}{25} \text{ x } 100 = 28\%$$

In other words, the TCP transfer experienced 28-percent additional RTD (congestion) which may have caused a proportional decrease in overall TCP throughput leading to longer delays for the end user.

RFC 6349 complements RFC 2544 testing filling the gap between the end user-experience and the manner in which the provider tests the network. The following table illustrates the network-test applicability of RFC 2544 and Y.1564sam along with the additional benefits of conducting an RFC 6349-based TCP test.

Turn-up Related Problem	RFC 2544	Y.1564sam	RFC 6349
Single-service, Layer 2/3 SLA issues, such as loss and jitter	X	X	N/A
Multi-service, Layer 2/3 SLA issues, such as service prioritization, loss, and jitter		X	N/A
Demonstrate the effect of end-customer TCP window size on throughput (CPE issue)			X
Inadequate device buffers to handle bursty customer applications			X
Policing effects to TCP performance			X

A companion application note, *RFC 6349 Testing with JDSU TrueSpeed™* provides details behind this important new methodology and also the automated JDSU test implementation of this new IETF RFC.

Two Scenarios of Unsatisfied Business Customers

Typically two reasons account for business-customer complaints of poor application performance over their network:

1. Business customers misconfiguring the customer equipment or running a flawed "speed test" procedure. Running over a high-bandwidth WAN network using default TCP window sizes can significantly reduce achievable TCP performance. An example of a flawed speed test is when users run a TCP test using an open-source tool, such as iperf, on a poorly performing PC and being unable to achieve the throughput of the SLA.
2. Network providers may experience network issues that require tuning such as inadequate port buffer sizes for TCP bursts. Traditional Layer 2/3 testing will not stress the network similar to the way bursty TCP traffic will. Bursty business-customer traffic will be policed, causing retransmissions and poor performance. Increasing port-buffer sizes (a form of traffic shaping) can greatly reduce packet loss and improve TCP performance.

A summary of two real-world case studies illustrate these two outcomes in the subsection that follows. Each case study uses traditional Layer 2/3 installation procedures without conducting TCP-layer testing prior to customer activation. These two case studies highlight the additional troubleshooting costs and customer churn that providers can avoid when conducting TCP testing prior to customer activation.

CPE Issue: Non-Optimal TCP Configuration

For this scenario, a business customer with two locations purchased a 100 Mbps transparent LAN service from a business provider with the network configuration shown in Figure 1.

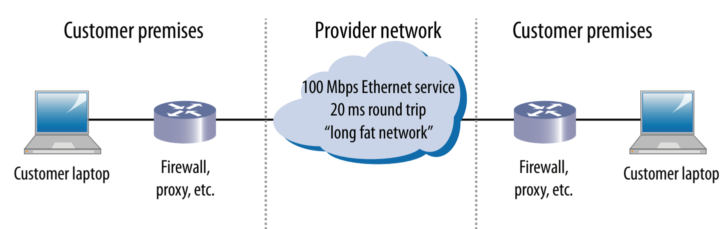

Figure 5. The business customer's 100 Mbps transparent LAN configuration

Table 1 shows the results of the RFC 2544 test and the customer's speed test that the network provider ran using FTP downloads and uploads.

RFC 2544 result	FTP result
100 Mbps throughput	25 Mbps throughput

Naturally, the business customer was unhappy with the FTP performance and contested the service provided. Much finger-pointing ensued along with several truck rolls in which the provider technicians re-ran the RFC 2544 test.

The problem in this scenario is that a 100 Mbps, 20 ms latency network, commonly referred to as a long fat network (LFN), requires much larger TCP-window settings in end-host computers. The ideal TCP window setting (as specified in RFC 6349) for this scenario can be calculated using the following formula:

$$\frac{100 \text{ Mbps} \times 20 \text{ ms}}{8} = 250{,}000 \text{ bytes}$$

Even though traditional network providers are not responsible for CPE devices, they should conduct RFC 6349 TCP tests between demarcation points. RFC 6349 provides guidelines for various network conditions as well as the associated optimal TCP window sizes which helps network providers optimize the network experience for its end customers, creating greater satisfaction and loyalty.

For this white paper, the network provider ran four different TCP window size tests and Table 2 reflects the actual results achieved.

TCP window size	Throughput
32 kB	12.5 Mbps
64 kB	25 Mbps
128 kB	50 Mbps
256 kB	100 Mbps

The tests clearly showed that the network was operating properly and that the end customer needed to enable the TCP window-scaling option in their CPE hosts. It turned out that the end customer was not using window scaling on one of the servers and the maximum window was only 64 kB (achieving only 25 Mbps of the 100 Mbps CIR).

In the end, we estimated that the additional OpEx costs to the provider were in the range of $7,000. This provider now uses the RFC 6349 test methodology in their standard service-activation procedure.

Network Provider Issue: Inadequate Network Buffers

For this scenario, a business customer with two locations purchased a 300 Mbps transparent LAN service from a business provider with the network configuration shown in Figure 2. Note that the latency between locations was approximately 50 ms.

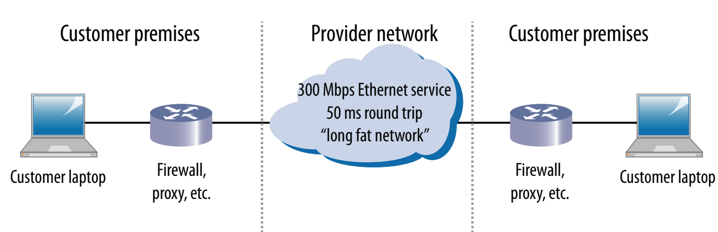

Figure 6. The business customer's 300 Mbps transparent LAN configuration

Table 3 shows the results of the RFC 2544 test and the customer's speed test that the network provider ran using the iperf program.

RFC 2544 result	iperf result
300 Mbps throughput	95 Mbps throughput

Again, much finger-pointing ensued along with several truck rolls in which the provider technicians re-ran the RFC 2544 test. The end customer demanded that the provider conduct some form of TCP-layer speed test because they lacked confidence in the results that the provider had presented.

In this test scenario, the BDP was very large and is calculated as follows:

$$\frac{300\text{ Mbps x } 50\text{ ms}}{8} = 1{,}875{,}000\text{ bytes}$$

Indeed this is a very large BDP, but the TCP-savvy end customer was running parallel TCP sessions (30 in this case) to properly attempt to "fill the pipe."

RFC 6349 includes parallel-TCP connection testing which is an important test method for extremely large BDPs. The provider used the JDSU TrueSpeed solution to replicate this test scenario along with the JDSU hardware TCP solution to remove any uncertainty about whether the processing performance of the end customer's workstation running the iperf program had affected the test results.

For this scenario, the provider actually replicated the problem and could only achieve the same 95 Mbps throughput as the end customer had claimed.

The provider identified the problem as inadequate default buffering in an edge-routing device. Increasing the buffer size in the device enabled the provider to retest and verify proper 300 Mbps TCP performance.

In the end, we estimated the additional OpEx costs to the provider were in the range of $15,000. This provider now uses the RFC 6349 test methodology in their standard service-activation procedure.

Conclusion

This white paper highlights the gap between current Ethernet service-activation methods RFC 2544 and Y.1564 and the end-customer experience. Customer applications ride on top of the TCP layer and its performance is greatly affected when moving from a LAN to WAN environment.

Augmenting RFC 2544 and/or Y.1564 service activation with an RFC 6349-based test enables the provider to share the customer's network experience, substantially save on OpEx, and dramatically increase first-time customer satisfaction, leading to higher customer loyalty and increased profitability.

JDSU adopted RFC 6349 TCP testing and has innovatively automated it. The JDSU TrueSpeed test capability is the industry's first RFC 6349-compliant test solution. TrueSpeed enables a network provider to run traditional RFC 2544 installation and TCP tests during the same truck roll and with the same skill-level technician and in less than five minutes.

The companion application note, *RFC 6349 Testing with JDSU TrueSpeed*, explains in detail the new RFC and highlights the TrueSpeed test feature.

Test & Measurement Regional Sales

NORTH AMERICA	LATIN AMERICA	ASIA PACIFIC	EMEA	WEBSITE: www.jdsu.com/test
TEL: 1 866 228 3762	TEL: +1 954 688 5660	TEL: +852 2892 0990	TEL: +49 7121 86 2222	
FAX: +1 301 353 9216	FAX: +1 954 345 4668	FAX: +852 2892 0770	FAX: +49 7121 86 1222	

30173004 000 0911 RFC6349.WP.TFS.TM.AE September 2011