
ExpertTCP™ - TCP Throughput Testing

(per RFC-6349)



818 West Diamond Avenue - Third Floor, Gaithersburg, MD 20878
Phone: (301) 670-4784 Fax: (301) 670-9187 Email: info@gl.com
Website: <https://www.gl.com>

Outline

- Background
 - RFC-2544, Y.1564 (SAM), RFC-6349, SLA
- TCP Principles
 - TCP Throughput Inter-Relationships
 - Bandwidth * Delay Product
 - Bottleneck Bandwidth (BB)
 - TCP Congestion Window (TCP CWND) and TCP Receive Window (RWND)
 - Packet Loss Rate
 - Retransmission Schemes (Go Back N, Selective Repeat)
- GL Hardware Platforms
- TCP Throughput Measurement
 - Path MTU Discovery
 - Round Trip Time Measurement
 - Measure TCP Throughput
- Screenshot

Performance Testing of Packet / Ethernet Connections and Networks

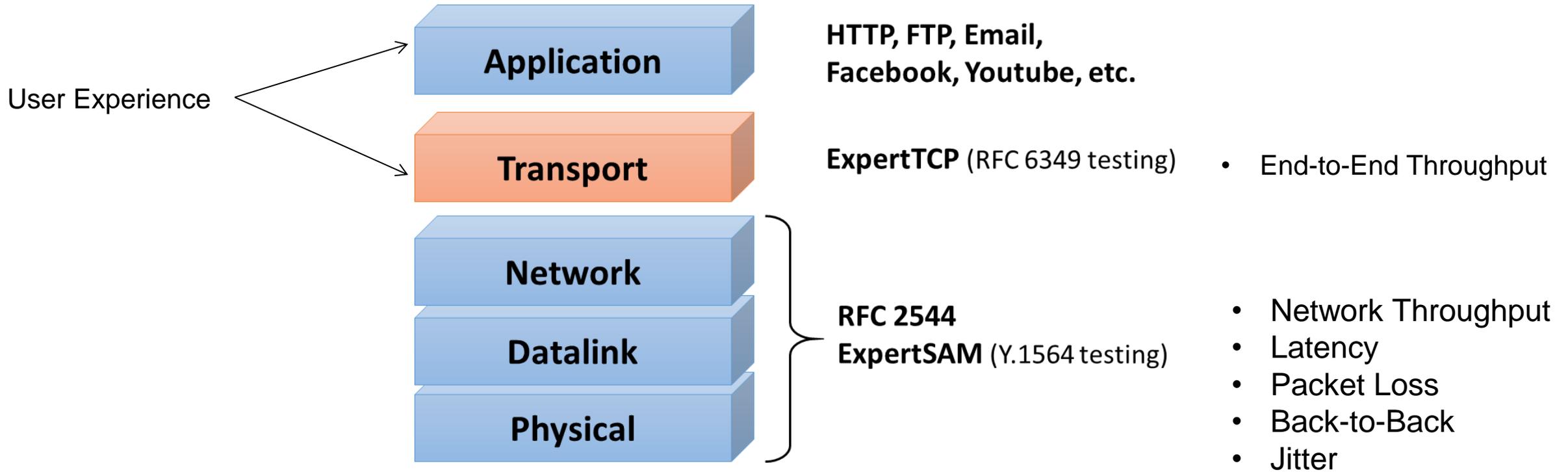
For Predictable Managed Networks

- RFC-2544 ← Service Level Agreements from Network Providers, a must
- ITU Y.1564 (SAM) ←
- RFC-6349 (TCP) ← User Experience, Application-Network Sensitive, TCP Tuning

SAM – Service Activation Methodology

TCP – Transmission Control Protocol

Packet / Ethernet Testing



Typical SLA

EXHIBIT D – Service Level Agreements

1. Service Level Agreement Matrix

Category/Service	Service Level Agreement Metrics				
	Mean Time To Repair	Availability	Packet Delivery or Loss	Jitter	Latency
Internet Services					
Internet Dedicated (North American IP Network Only)	4 hrs to 8 hrs depending on access	99.90%	≥ 99.50%	≤ 1 ms	≤ 45 ms
SOHO Services					
Internet Cable	24 hrs (Excludes Weekends and Holidays)	99.00%	99.00%	≤ 4 ms	≤ 75 ms
Internet DSL – Office & Solo					
Internet Satellite Enterprise & Office	N/A	99.90%	≤ 1 %	N/A	N/A
Managed PBX and VoIP Services					
Hosted IP Centrex	≤ 4 hrs	99.90%	EF- ≥ 99.995%, AF4x - ≥ 99.99% depending on access	≤ 1 ms	≤ 36 ms
IP Flexible T1, IP Integrated Access, IP Trunking					

Typically

Packet Loss

0.0005 % to 1%

Latency

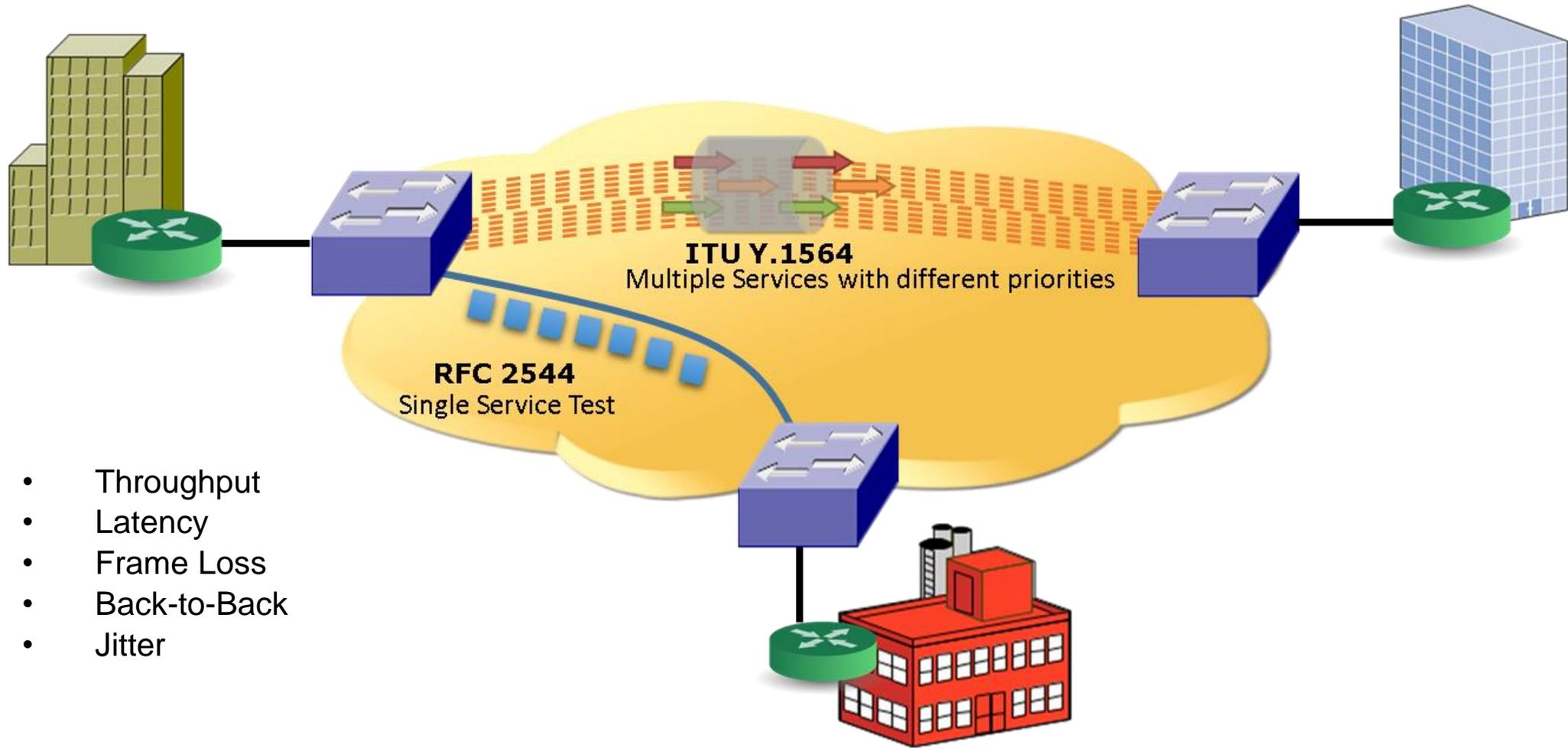
36 to 75 ms

Availability

99% to 99.9%

RFC-2544 vs. ITU Y.1564 (ExpertSAM™)

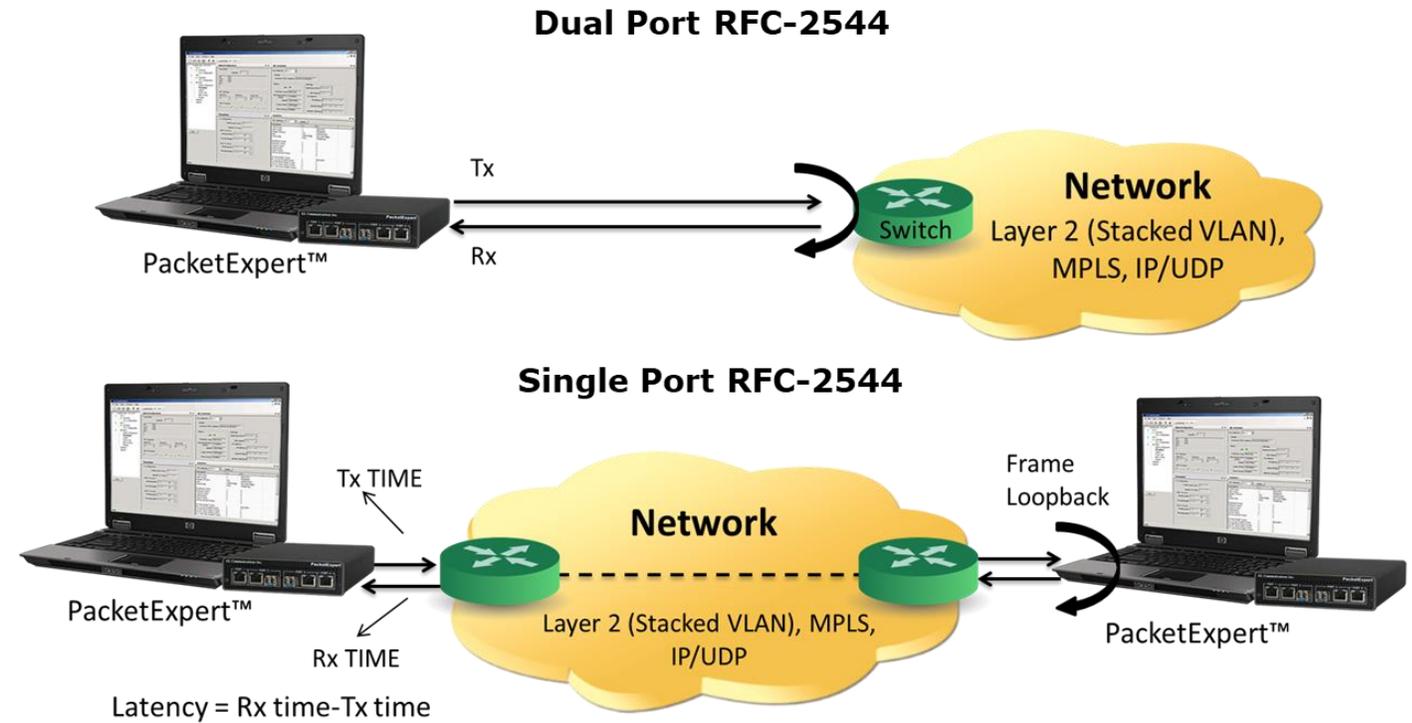
Both are Connection-less



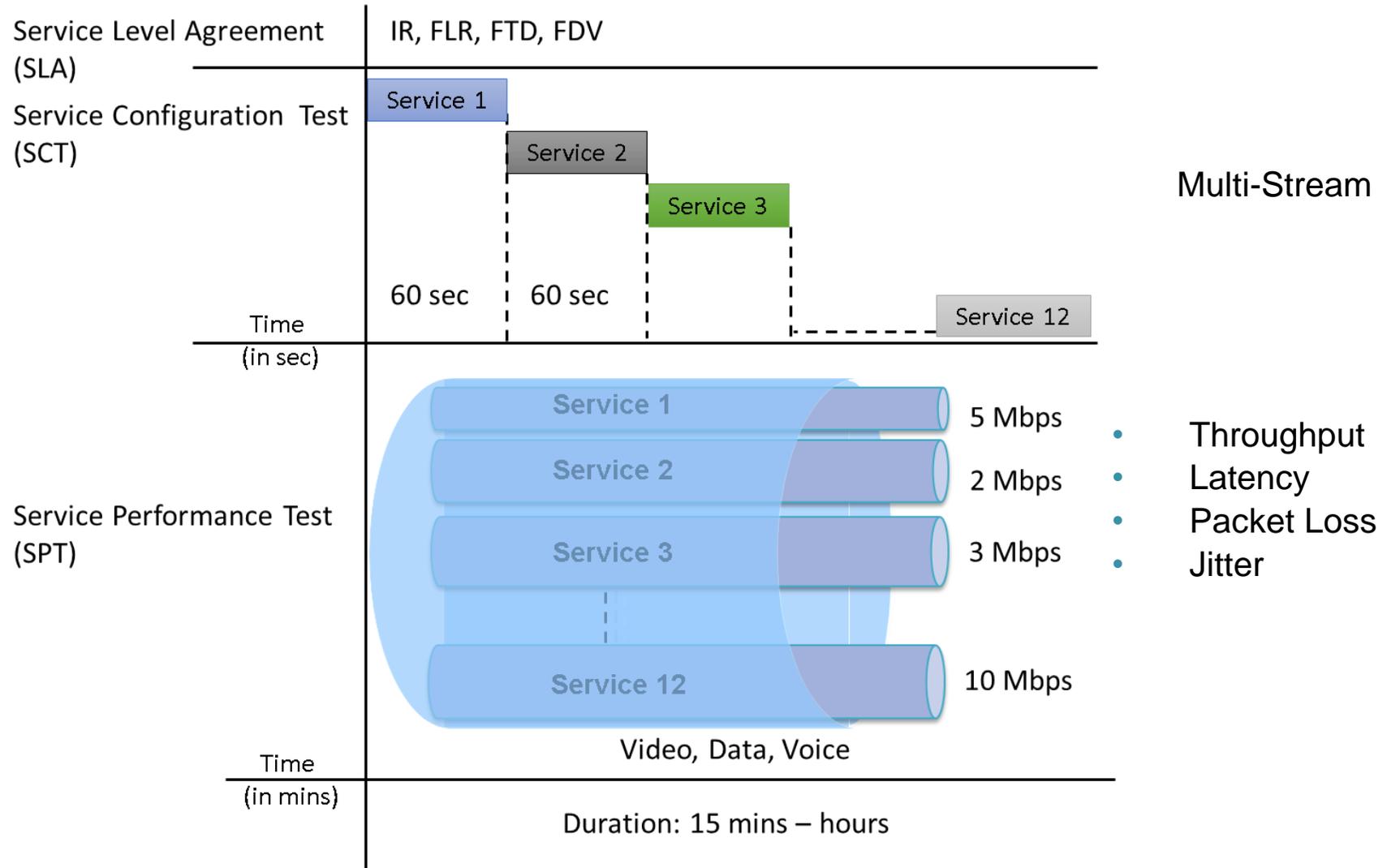
- Throughput
- Latency
- Frame Loss
- Back-to-Back
- Jitter

RFC-2544 Testing

- ExpertTCP™ testing is performed using the RFC 6349 standard
- To conduct this test, users need two PacketExpert™ devices — one as the client and the other as the server
- The ExpertTCP™ test covers both upload (Client to Server) and download (Server to Client), measuring TCP throughput and efficiency
- RFC-2544 test application includes the following tests:
 - Throughput - Maximum number of frames per second that can be transmitted without any error
 - Latency - Measures the time required for a frame to travel from the originating device through the network to the destination device
 - Frame Loss - Measures the network's response in overload conditions
 - Back-to-Back - It measures the maximum number of frames received at full line rate before a frame is lost



ITU Y.1564 (ExpertSAM™)



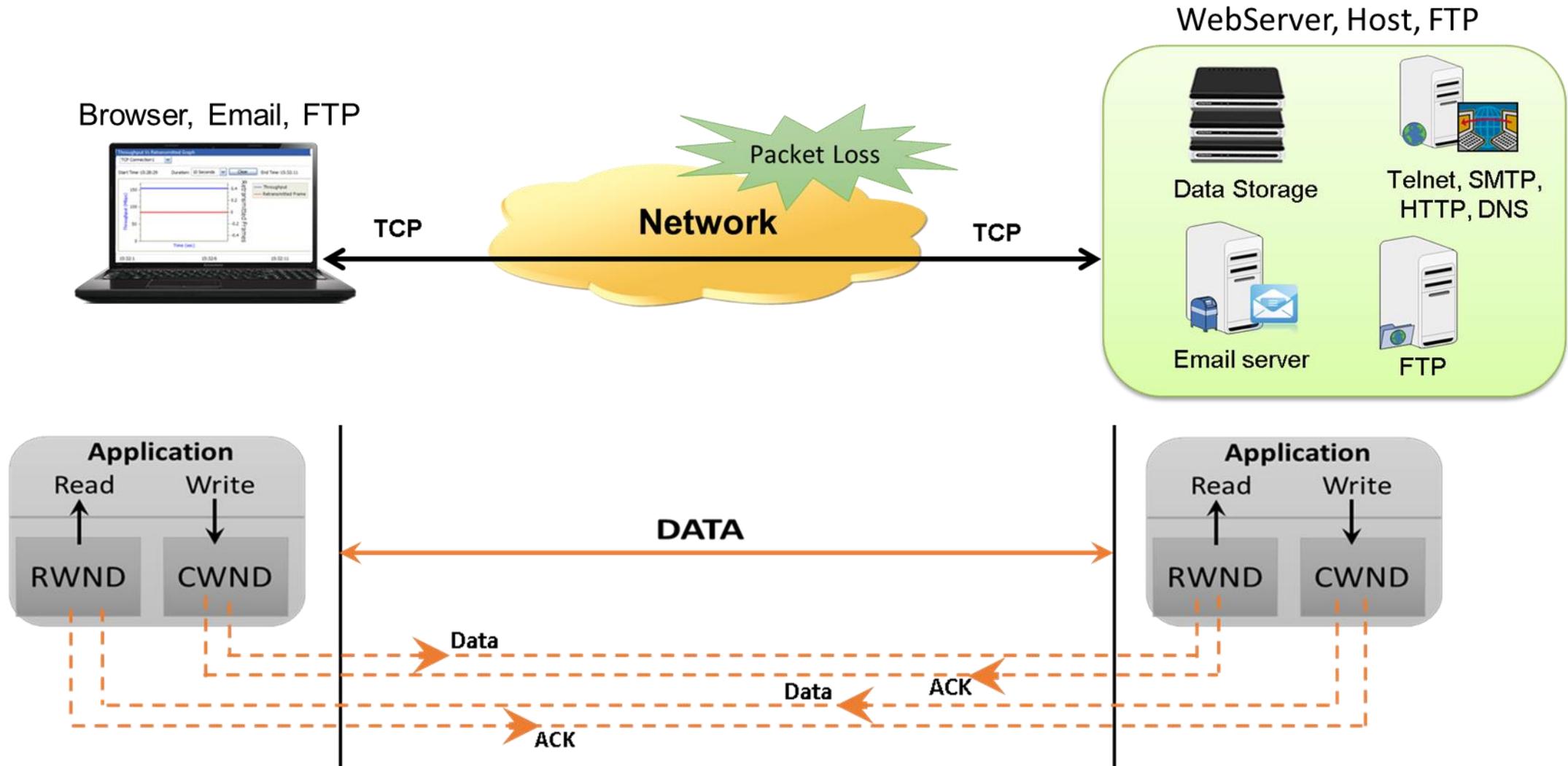
Testing Relevance

Problems	RFC-2544	Y.1564	RFC-6349
Single-service Layer 2/ 3/ 4 SLA Issues like loss, jitter	Yes	Yes	N/A
Multi-service Layer 2/ 3/ 4 SLA Issues like loss, jitter	No	Yes	N/A
TCP window sizes (CPE issues)	No	No	Yes
Excessive retransmissions due to policing	No	No	Yes

- Running RFC-2544, Y.1564 or another L2/L3 layer test is always first step
- However, even after these performance tests are passed with good results, end-customers can still complain that the “network is slow” and the cause of poor application performance (i.e., FTP, web browsing, etc.)
- Lack of TCP testing is a turn-up gap because end-customer applications are transported using TCP
- Save operating expense costs by eliminating or quickly resolving painful end-customer finger pointing scenarios

TCP Principle

(Packet Loss and Waiting for ACK Reduces Throughput)

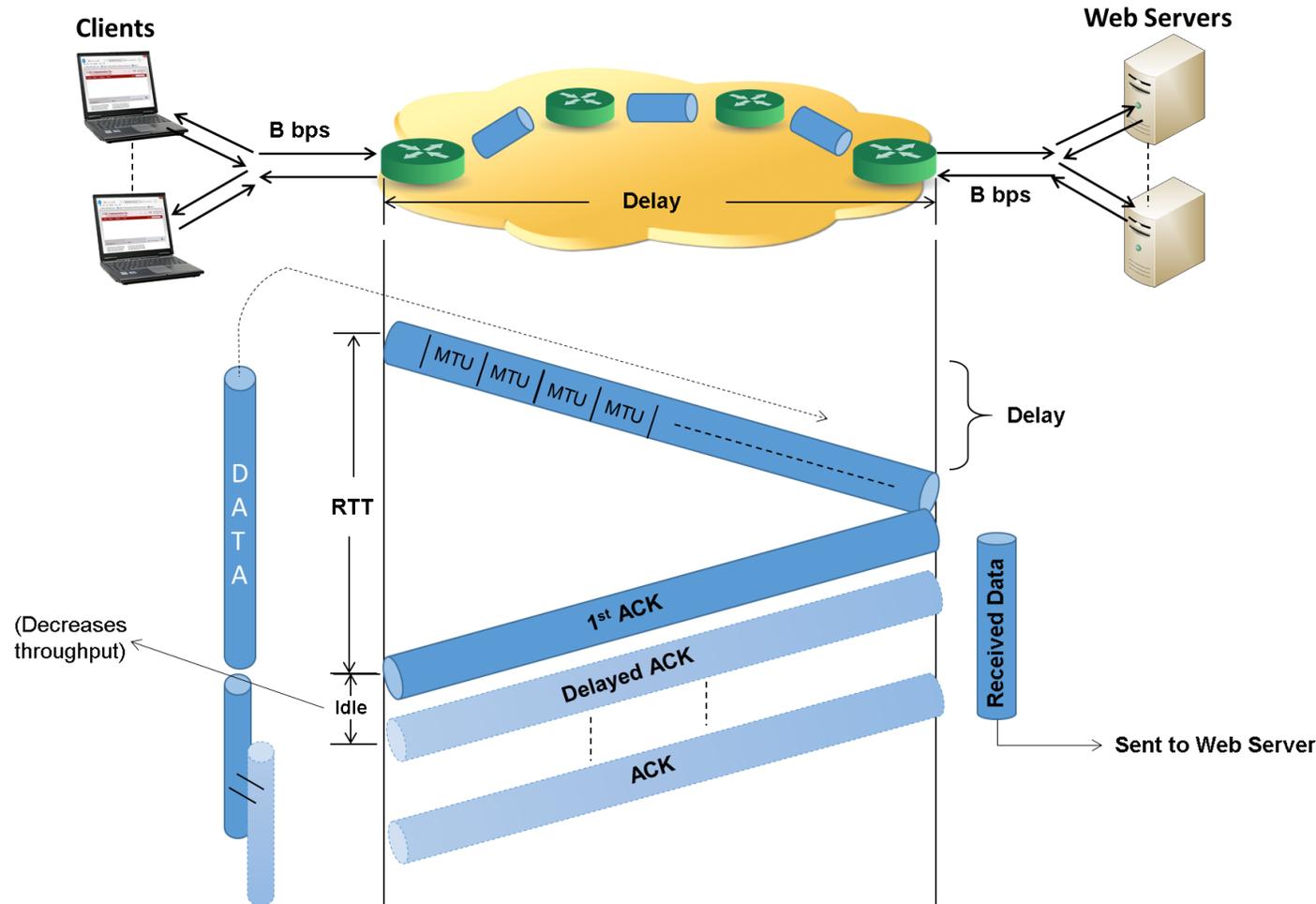


Major TCP Throughput Inter-Relationships

- Bandwidth of Applications
- Latency/Delay of Networks
- Packet Loss Networks
- TCP Retransmission Scheme
- Maximum Transmit Unit of Network
- Transmit/Receive Windows of TCP
- # (number) of TCP Simultaneous Connections

Bandwidth Delay Product (Bits or Bytes)

Application and Network are Matched, TCP is Tuned



$B = 10 \text{ Mbps}$
 $RTT = 50 \text{ ms}$

$B * 50 = 500,000 \text{ bits}$
or $62,500 \text{ Bytes}$

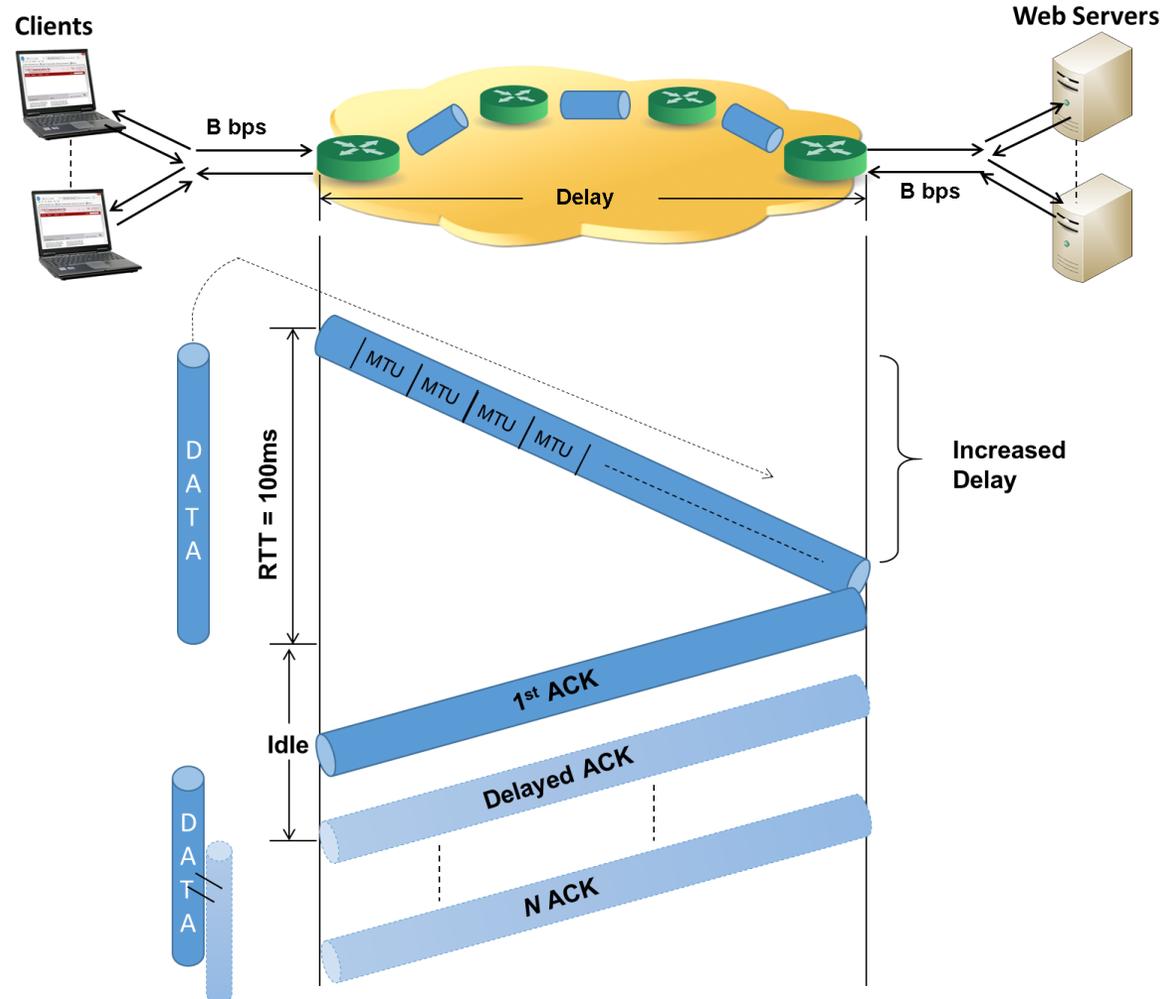
$65,535 \text{ Bytes}$ is max window

Achieving max throughput

Bandwidth (B) - Bandwidth (bps), Mbps, the maximum rate at which an application can transmit or receive data (the smaller of the two). Line rate may be shared among applications.

Bandwidth Delay Product (BDP) - measured in bits or bytes (divided by 8), the number of bits (or bytes) in the network that are unacknowledged (in transit), $B \text{ (bps)} * RTT \text{ (secs)} = \text{BDP bits}$.

Effect of Increased Network Delay or Smaller Tx or Rx Buffers



B = 10 Mbps
RTT = 100 ms

$B * 100 = 1,000,000$ bits
or 125,000 Bytes

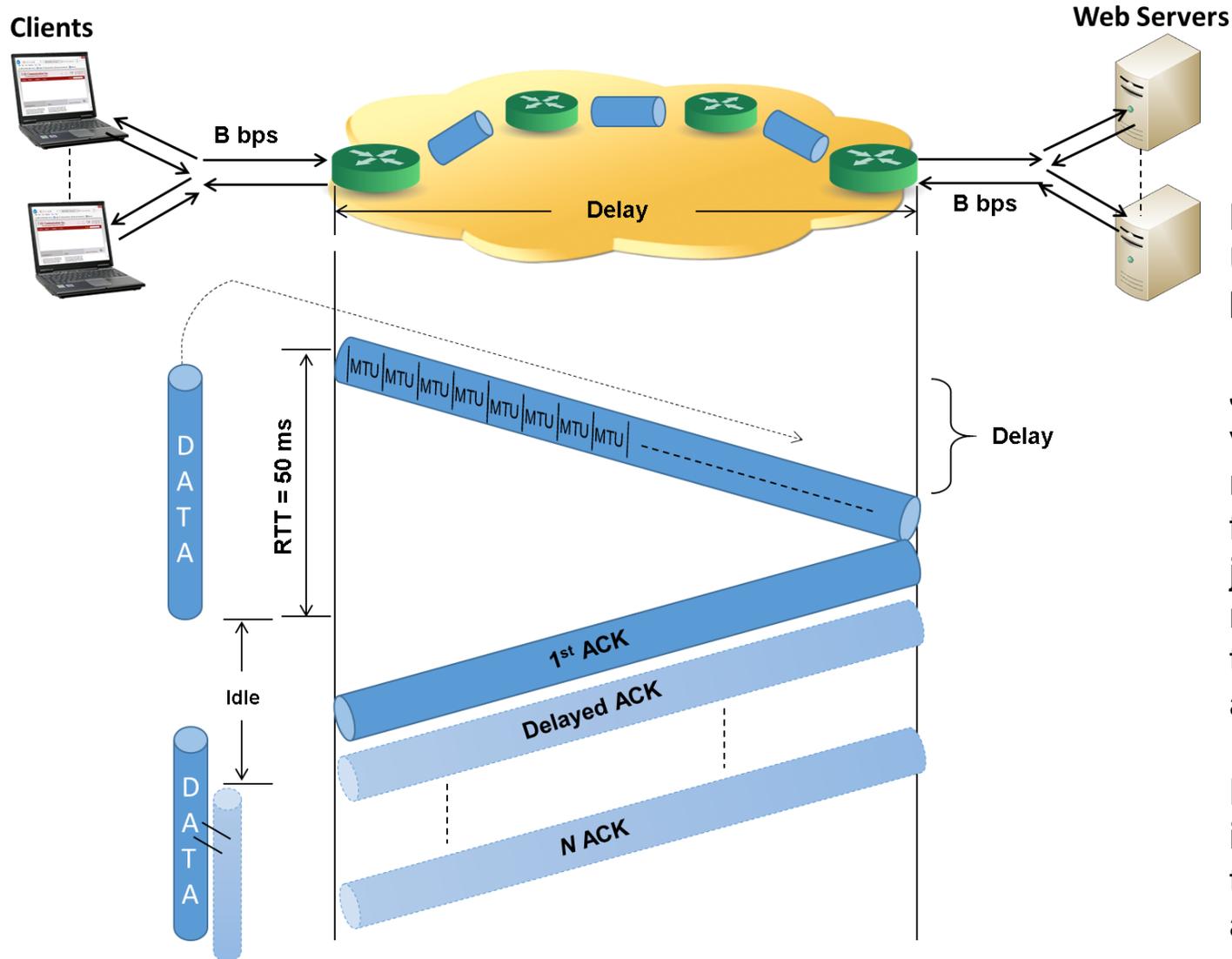
But 65,535 Bytes is max window

NOT Achieving max throughput, 50% or less

Latency, Delay, Round Trip Time (RTT) - in seconds (secs), or milliseconds (ms), round trip time includes acknowledgement delay.

TCP Throughput - bits/second (bps), million bits/second (Mbps), One way throughput (RFC2544, Y.1564), Round-trip throughput (RFC-6349) is a different story since retransmissions and acknowledgements are involved.

Effect of Increased Application Bandwidth



$B = 20$ Mbps
 $RTT = 50$ ms

$B * 50 = 100,000$ bits
 or 125,000 Bytes

But 65,536 Bytes is
 max window

NOT Achieving max
 throughput, 50% or
 less

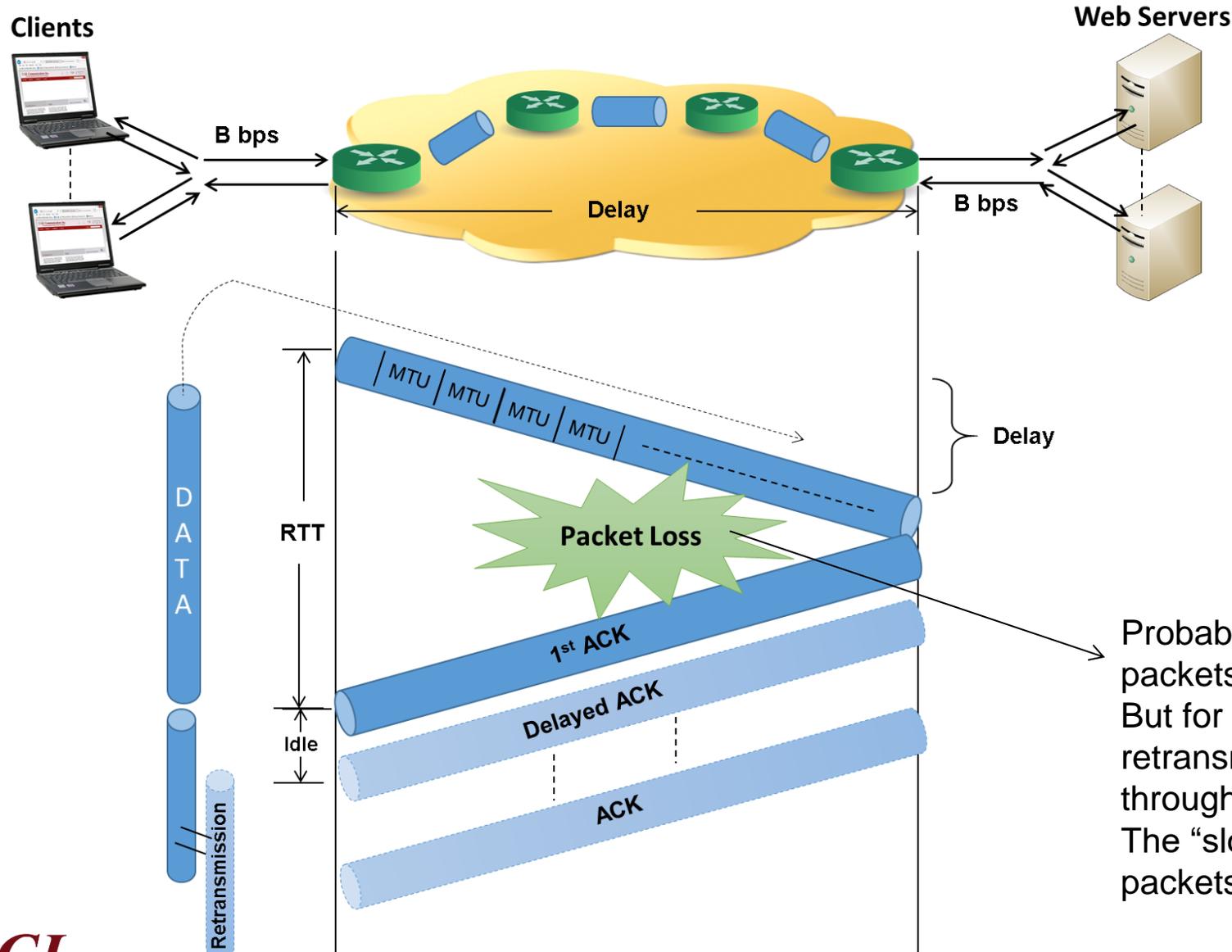
Maximum Transmission Unit (MTU) - Approx. 1500 bytes, max packet size.

Jitter - Instantaneous variation in RTT, e.g. if RTT is nominally 100 ms, but varies from 80 ms to 120 ms, then jitter is +/- 20ms, or 40 ms. Since jitter affects ACK time, TCP throughput is affected.

Packet Loss Rate - Very important factor affecting TCP throughput, could be as high as 2%.

Excess Bandwidth may be used for additional TCP Connections

Effect of Packet Loss Rate and Retransmission Scheme



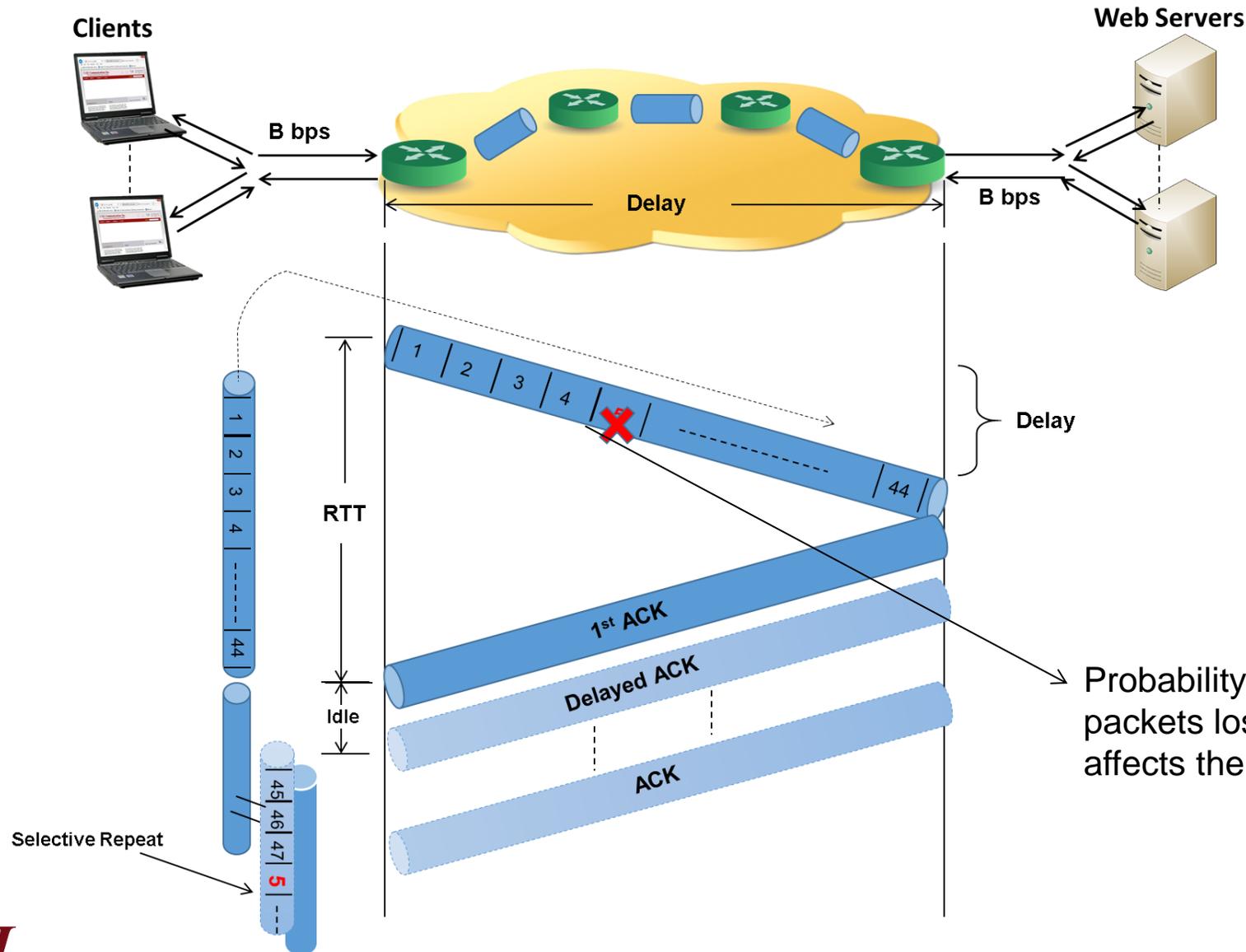
For **Go Back N** retransmission scheme, and if unacked packets is maximum ~ 43 or 44, then Packet Loss effects are very serious!

Packet Loss TCP Throughput

0 %	100%
0.1 %	< 50%
1 %	< 10%
2 %	0 %

Probability that one or more MTU packets or ACK packets is lost is very high!! Can be 1 !!!
 But for every lost MTU packet or ACK packet, 43 retransmissions occur. This results in near zero throughput.
 The "slow start phase" results in very few "in flight" packets.

Effect of Packet Loss Rate and Retransmission Scheme (Contd.)



For **Selective Repeat** retransmission scheme, and if unacked packets is maximum ~ 43 or 44, then Packet Loss affects TCP Throughput linearly for “low” Packet Loss rates.

Packet Loss	TCP Throughput
0 %	100%
0.1 %	> 99 %
1 %	> 95 %
2 %	? %

Probability that one or more MTU packets or ACK packets lost is very high! But the retransmission only affects the lost packets, not other packets.

ExpertTCP™ (RFC-6349 Testing)

The TCP Throughput Testing is conducted in 3 steps simultaneously on up to 16 application streams:

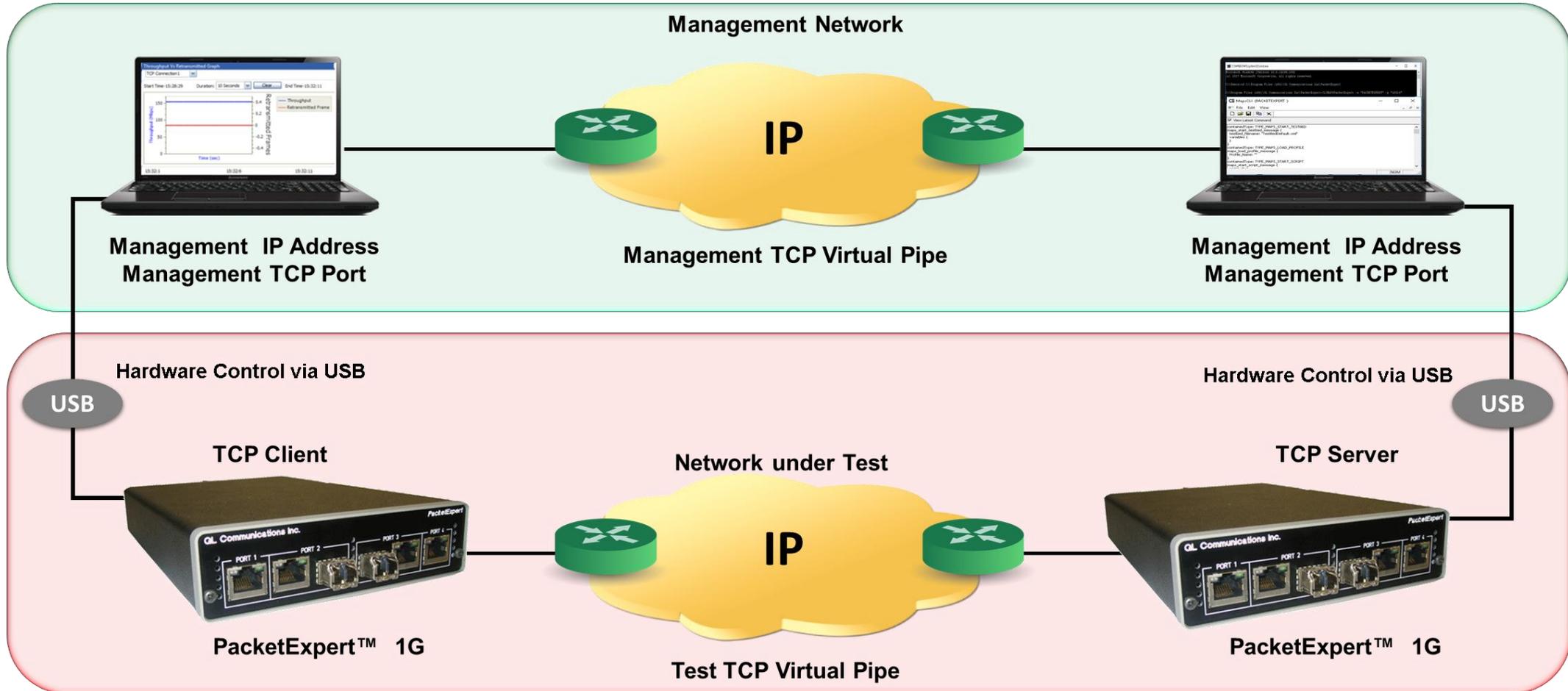
- **Path MTU Discovery** - What is the maximum packet size that can successfully traverse the network?
- **Round Trip Time (RTT) Measurement** - Timestamp based RTT discovery of transmitted packet until acknowledgement packet arrives from far end
- **Measure TCP Throughput** - Complete measurements per RFC-6349 definitions to provide TCP Throughput results

GL's ExpertTCP™ Provides Reports and Graphs of all Results

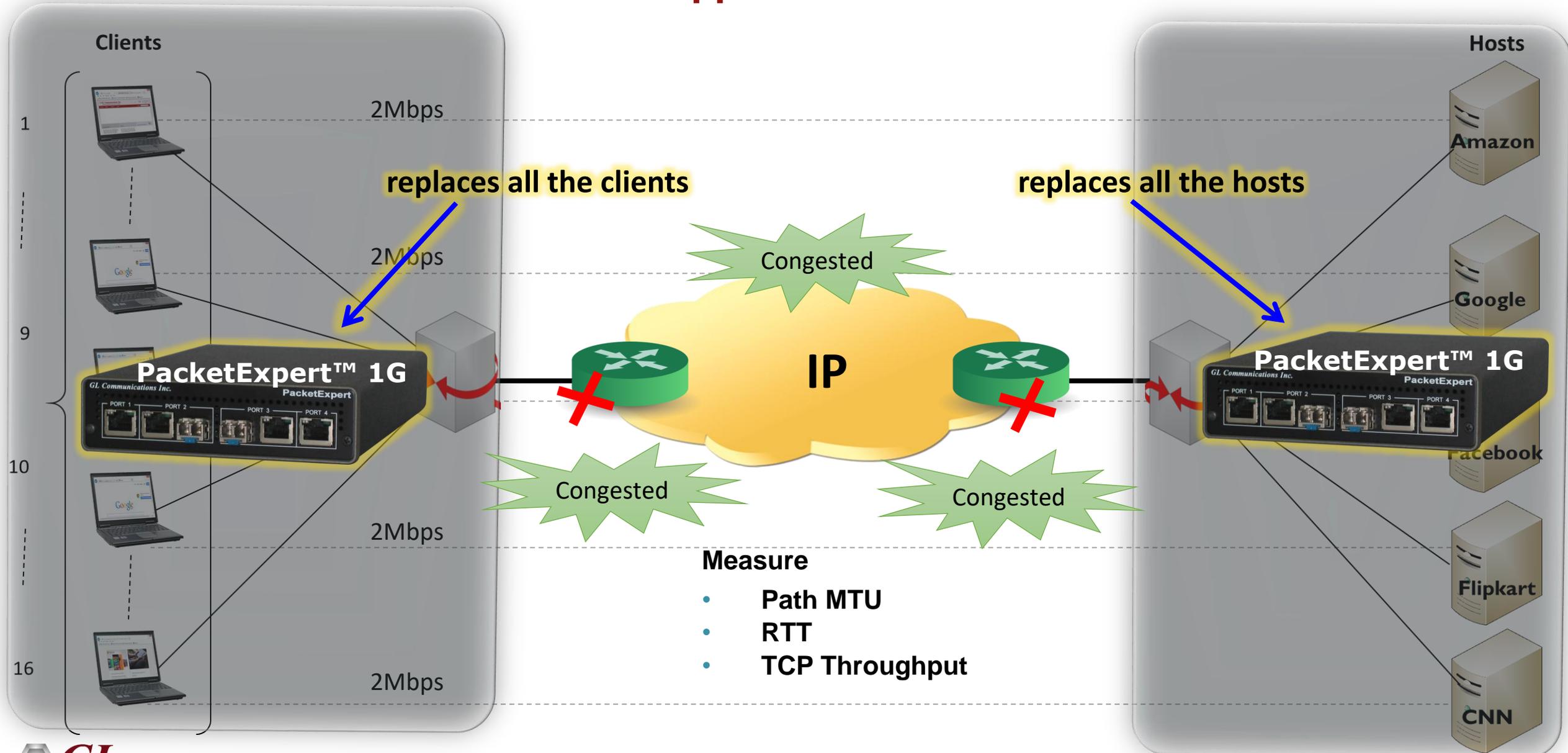
GL Hardware / Software ExpertTCP™

Basic Setup

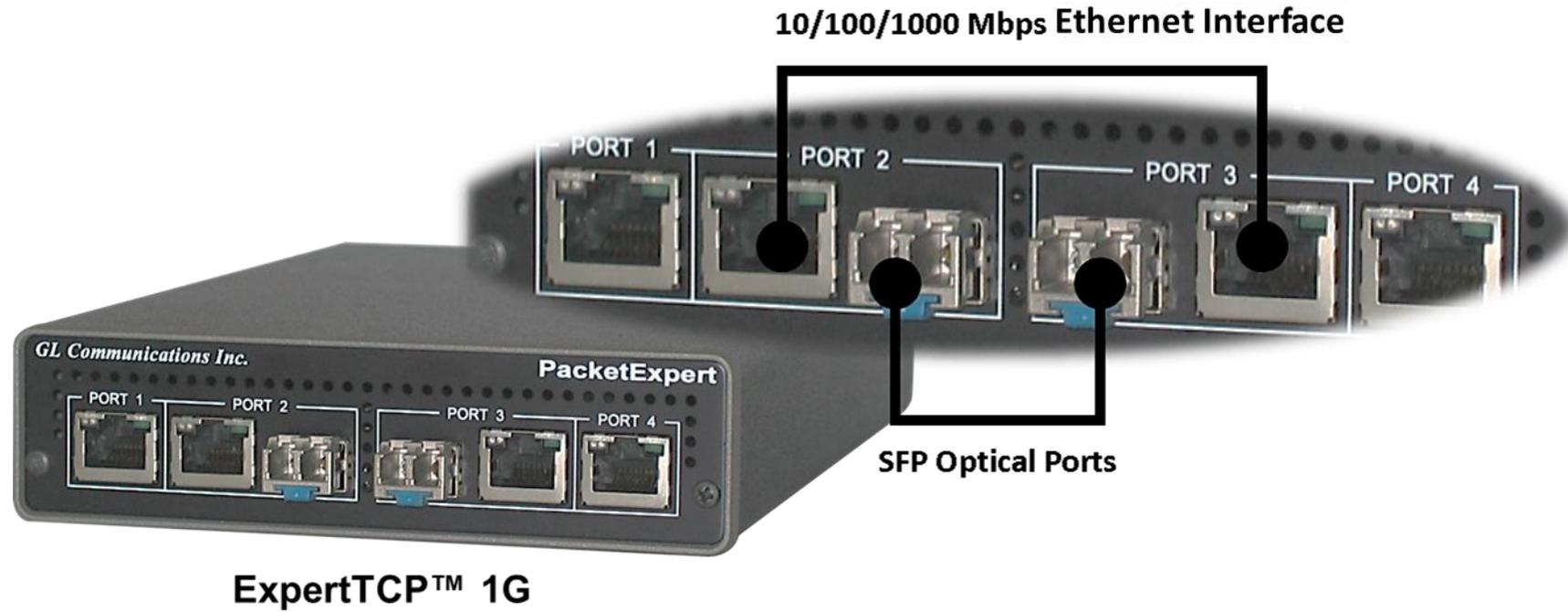
Test Configuration of Client and Server
Measurement Results from Server to Client



End-to-End Application Performance



ExpertTCP™ 1G Ports

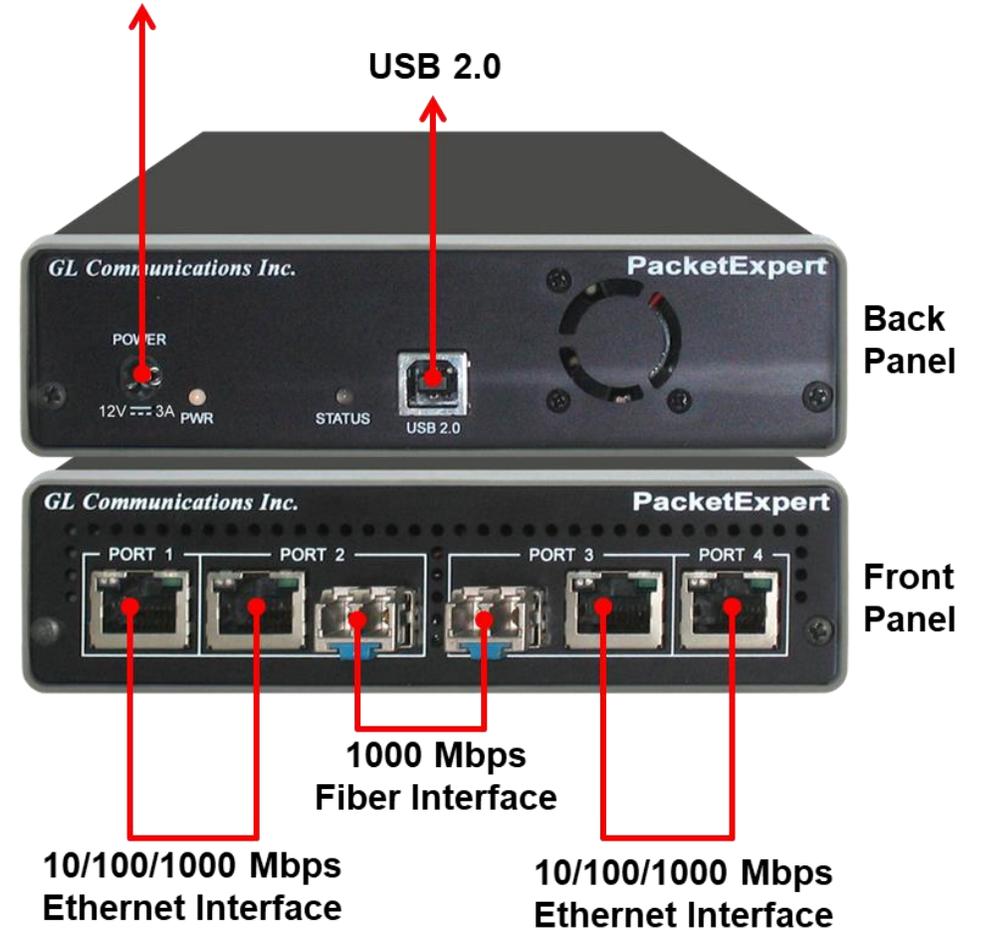


- **TCP Client and Server** will be supported in two different applications.
- In 1G, **Port 2** is used.

PacketExpert™ 1G Portable Unit

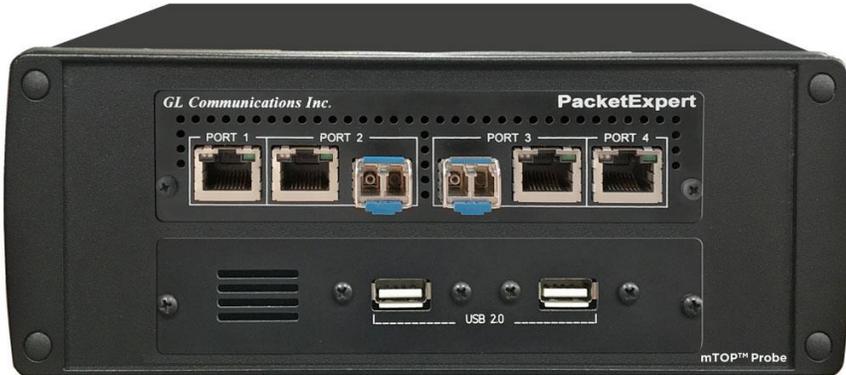
- Interfaces
 - 2 x 10/100/1000 Base-T Electrical only
 - 2 x 1000 Base-X Optical OR 10/100/1000 Base-T Electrical
 - Single Mode or Multi Mode Fiber SFP support with LC connector
 - Optional 4-Port SMA Jack Trigger Board (TTL Input/Output)
- Protocols:
 - RFC 2544 compliance
 - ITU-T Y.1564 (ExpertSAM)
- Power:
 - +12 Volts (Medical Grade), 3 Amps
- Bus Interface:
 - USB 2.0

Power: 12V (Medical Grade), 3A



PacketExpert™ mTOP™ Probe

Front Panel View



Rear Panel View



- Portable Quad Port Ethernet/VLAN/MPLS/IP/UDP Tester with 4 Electrical Ethernet Ports (10/100/1000 Mbps) and 2 Optical Ports (100/1000 Mbps). Embedded with Single Board Computer (SBC)
- **SBC Specs:** Intel Core i3 or optional i7 NUC Equivalent, Windows® 11 64-bit Pro Operating System, USB 3.0 and USB 2.0 Ports, 12V/3A Power Supply, USB Type C Ports, Ethernet 2.5GigE port, 256 GB Hard drive, 8G Memory (Min), Two HDMI ports
- Each GigE port provides independent Ethernet/VLAN/MPLS/IP/UDP testing at wire speed for applications such as BERT, RFC 2544, and Loopback. BERT is implemented for all layers
- RFC 2544 is applicable for Layers 2, 2.5, and 3, and Loopback is applicable for Layers 2, 3, and 4

PacketExpert™ High-Density 12/24 GigE Ports mTOP™ Rack

PacketExpert™ SA (PXE112) is a 12-Port PacketExpert™ w/ Embedded Single Board Computer (SBC).

SBC Specs: Intel Core i3 or optional i7 NUC Equivalent, Windows® 11 64-bit Pro Operating System, USB 3.0 and USB 2.0 Ports, ATX Power Supply, USB Type C Ports, Ethernet 2.5GigE port, 256 GB Hard drive, 8G Memory (Min), Two HDMI ports.

19" 1U Rackmount Enclosure (If options, then x 3).

PacketExpert™ SA (PXE124) is a 24-Port PacketExpert™ w/ Embedded Single Board Computer (SBC).

SBC Specs: Intel Core i3 or optional i7 NUC Equivalent, Windows® 11 64-bit Pro Operating System, USB 3.0 and USB 2.0 Ports, ATX Power Supply, USB Type C Ports, Ethernet 2.5 GigE port, 256 GB Hard drive, 8G Memory (Min), Two HDMI ports.

19" stacked 1U Rackmount Enclosure (If options, then x 6).

PacketExpert™ SA (PXE112)

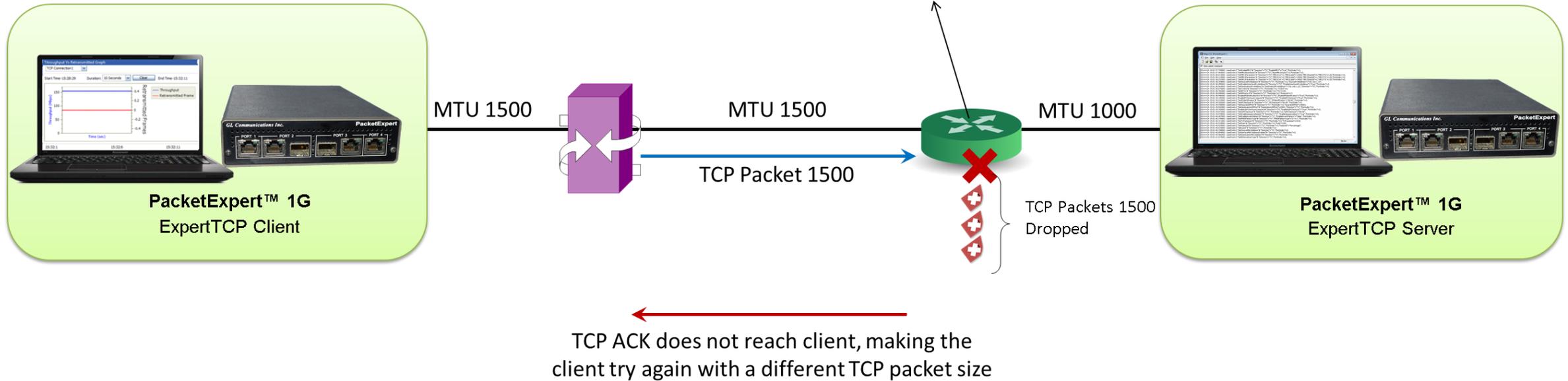


PacketExpert™ SA (PXE124)

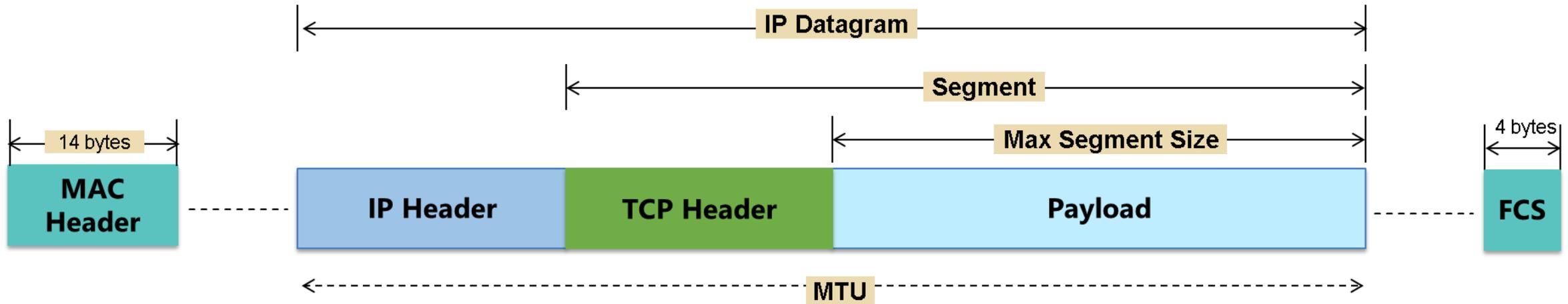


Step 1. Path MTU Discovery

Client sends packet with Don't Fragment (DF) bit set

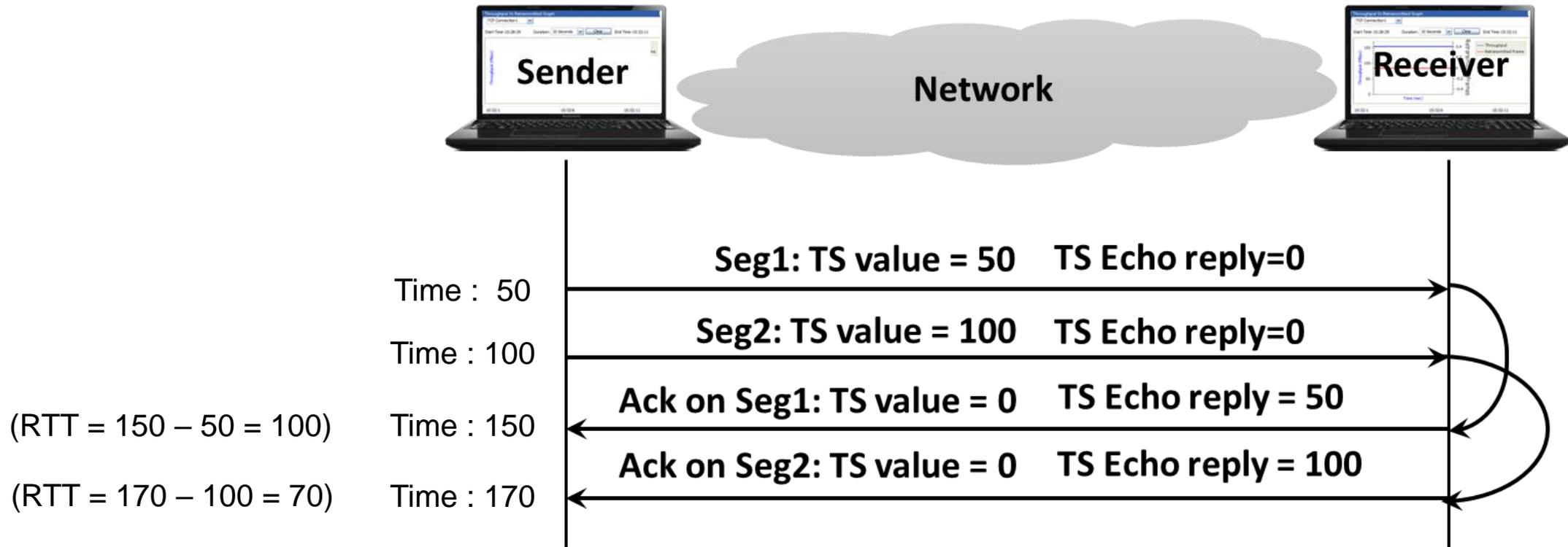


Step 1. Path MTU Discovery



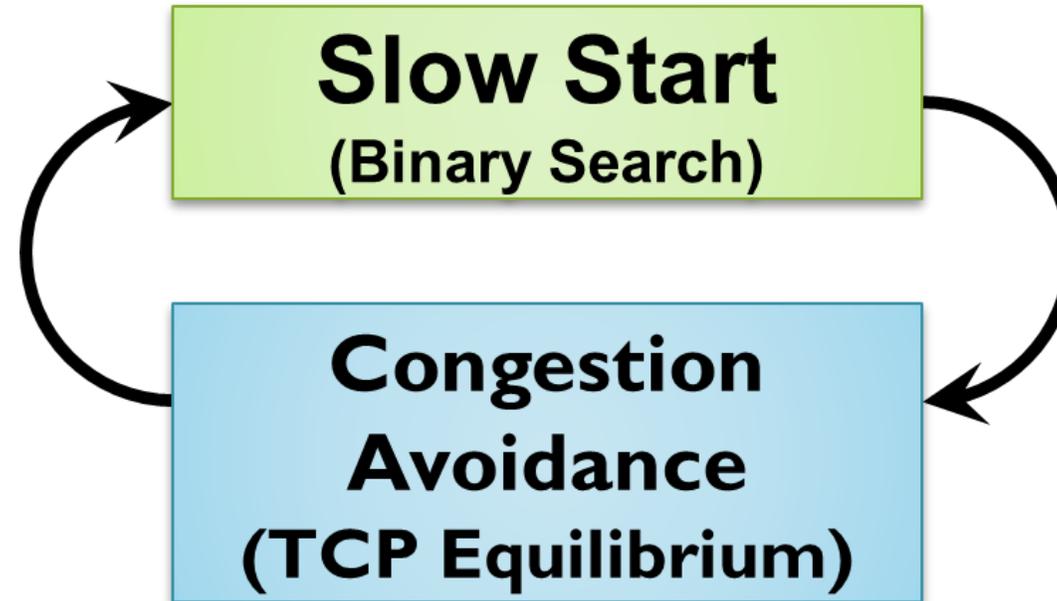
- Path MTU discovery as per RFC 4821 - PLPMTUD - Packetization Layer Path MTU Discovery.
- DF (Do Not Fragment) bit is set to avoid fragmentation when traversing through network.
- The algorithm uses TCP retransmit conditions to search for the MTU.
- Each conclusive probe narrows the MTU search range, either by raising the lower limit on a successful probe or lowering the upper limit on a failed probe.
- Path MTU is discovered for both directions in case of bi-directional test.

Step 2. Timestamp based RTT Measurement

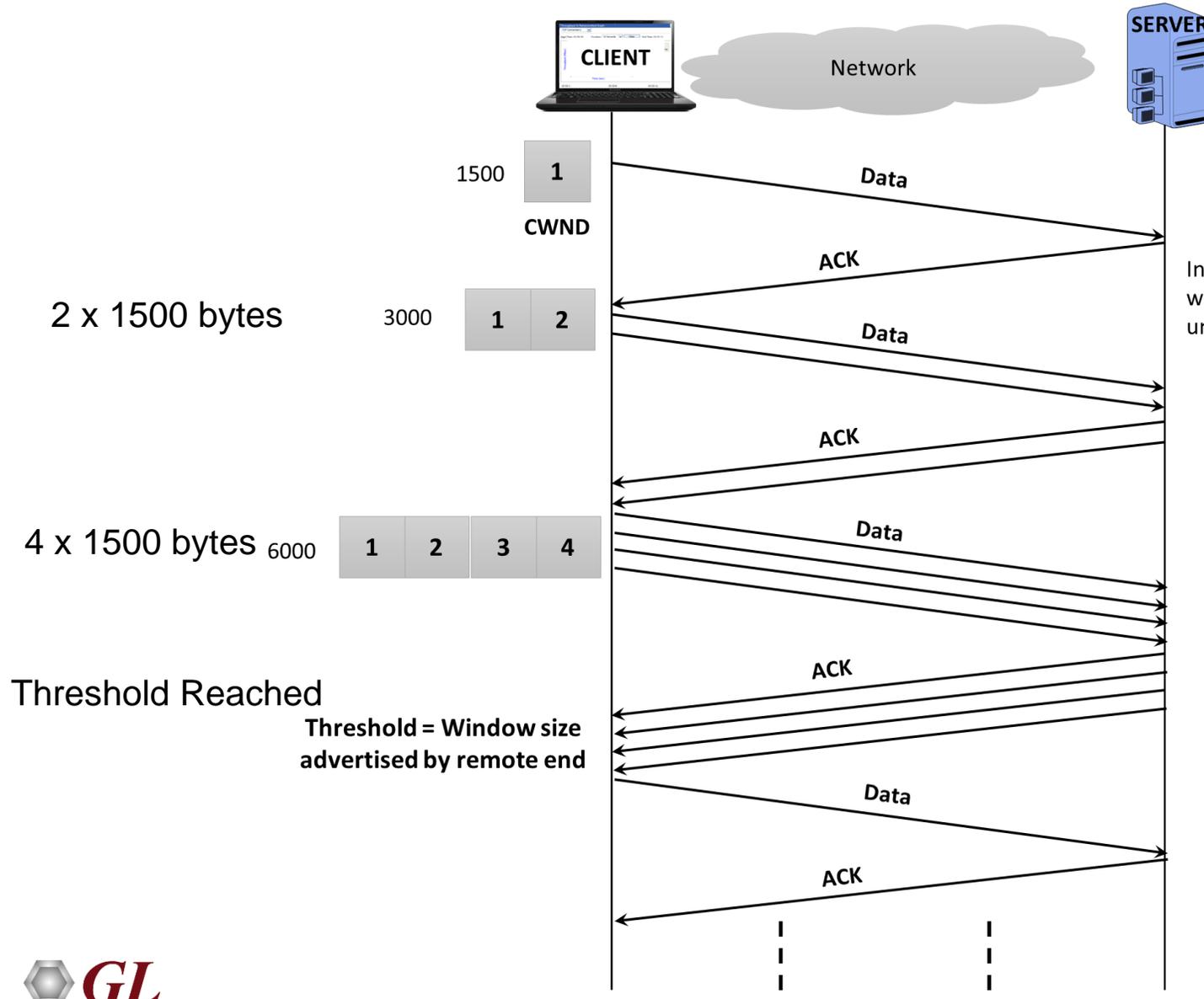


- Timestamp based RTT Measurement (RFC1323).
- Tx segment includes current time in option field, Receiver echoes timestamp in ACK.

Step 3. Now Ready to Measure TCP Throughput



Step 3. Slow Start TCP Throughput Measurement



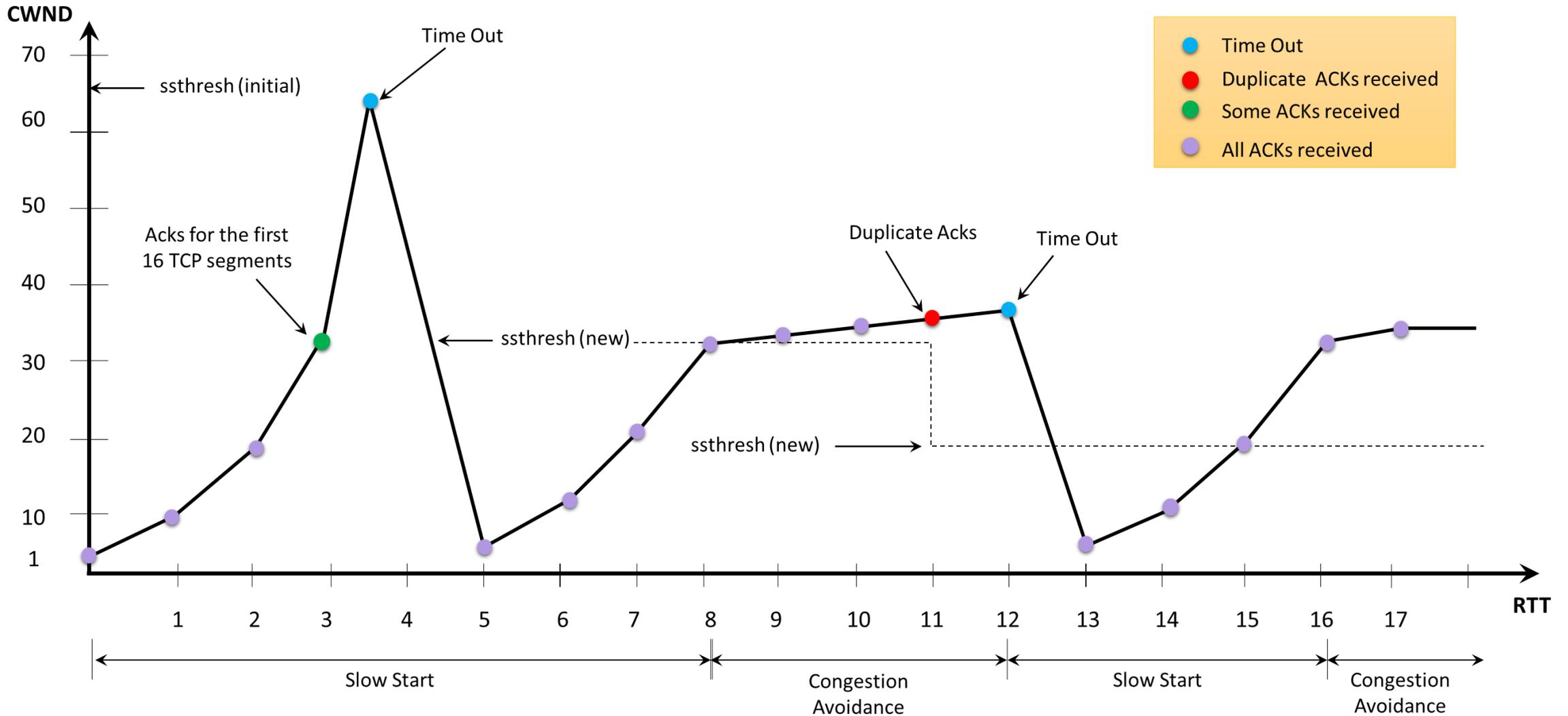
In Slow start, the congestion window increases exponentially until it reaches threshold

Slow Start - Initially send two TCP Segments
If Acks received, then send double the number of TCP Segments.

Continue doubling until the Receiver "ssthreshold" # is reached, or Acks are not received and Timeout is reached, then halve the send TCP segments.

If Acks are received send TCP segments are incremented by one, until again Timeout is reached, then number of send TCP segments is halved and the process continues.

Step 3. TCP Throughput Equilibrium



- Time Out
- Duplicate ACKs received
- Some ACKs received
- All ACKs received

Screenshots of Software Operation

ExpertTCP™ Main Screen

The screenshot displays the ExpertTCP main interface with the following components:

- Application:** ExpertTCP (Beta)
- Left Navigation Tree:**
 - ExpertTCP
 - Config
 - Remote
 - Interface (Local)
 - Interface (Remote)
 - Network Setup
 - TCP Setup
 - Test Setup
 - Results
 - Overall Status
 - Path MTU Results
 - Baseline RTT Results
 - Test Parameter Summary
 - Overall Results
 - RTT Results
 - Throughput Results
 - Statistics
 - Final Results
 - Graph
 - Throughput
 - ThroughputVsRTT
 - ThroughputVsRetransmitted
 - Port Statistics (Local)
 - Port Statistics (Remote)
 - Reports

- Interface (Local) - Top Left Panel:**
- Details:** Hardware MAC address: 00-21-C2-00-04-A9
- Settings:** Interface Type: Electrical, Link Speed: 1000Mbps, Disable Auto Negotiation, Apply
- Status:** Link: ●, Interface Type: Electrical, Auto-Negotiation Status: Complete, Speed: 1000 Mbps, Duplex Mode: Full Duplex, Flow Control: Enabled
- Interface (Local) - Top Right Panel:** (Identical to the top left panel)
- Network Setup - Middle Panel:**
- Diagram: Client (Local) — Network Under Test — Server (Remote)
- Client (Local):** MAC Address: User Defined (00-21-c2-00-04-a9), IP Address: 192.168.1.111, Subnet Mask: 255.255.255.0, Default Gateway: 192.168.1.1
- Network Under Test:** Link Type: Symmetrical Asymmetrical, Upstream CIR: 100 Mbps, Downstream CIR: 1000 Mbps
- Server (Remote):** MAC Address: 00-21-c2-00-04-a9, IP Address, Subnet Mask, Default Gateway
- Overall Status - Bottom Right Panel:**
- Test Status: Idle
- Current Direction: -
- Current Test Table:

Test	Status	Result
- TCP Connection Status: (Empty list)

Network Setup

All settings configured locally on the client side

Remote

Remote Server IP Address: 192 . 168 . 1 . 232

Status: Connected 

Interface - Local (Port2)

Details

Hardware MAC address: 00-21-C2-00-09-B1

Status

Link: 

Interface Type: Electrical

Auto-Negotiation Status: Complete

Speed: 1000 Mbps

Duplex Mode: Full Duplex

Flow Control: Enabled

Settings

Interface Type: Electrical

Link Speed: 1000Mbps

Disable Auto Negotiation

Interface - Remote (Port2)

Details

Hardware MAC address: 00-21-C2-00-04-CE

Status

Link: 

Interface Type: Electrical

Auto-Negotiation Status: Complete

Speed: 1000 Mbps

Duplex Mode: Full Duplex

Flow Control: Enabled

Settings

Interface Type: Electrical

Link Speed: 1000Mbps

Disable Auto Negotiation

Network Setup (Contd.)

Separate Upstream and Downstream bandwidths configurable for asymmetrical path

Network Setup

The diagram illustrates a network setup with three main components: Client (Local), Network Under Test, and Server (Remote). The Client (Local) is connected to the Network Under Test, which is connected to the Server (Remote). The Client (Local) configuration includes a MAC Address of 00-21-c2-00-05-02 (User Defined checkbox is unchecked), IP Address 192.168.1.111, Subnet Mask 255.255.255.0, and Default Gateway 192.168.1.1. The Network Under Test configuration includes a Link Type of Symmetrical (selected) and Asymmetrical (unselected), Upstream CIR 10 Mbps, and Downstream CIR 10 Mbps. The Server (Remote) configuration includes a MAC Address of 00-21-c2-00-06-1e (User Defined checkbox is unchecked), IP Address 192.168.1.222, Subnet Mask 255.255.255.0, and Default Gateway 192.168.1.1.

Component	MAC Address	IP Address	Subnet Mask	Default Gateway	Link Type	Upstream CIR	Downstream CIR
Client (Local)	00-21-c2-00-05-02	192.168.1.111	255.255.255.0	192.168.1.1	Symmetrical	10 Mbps	10 Mbps
Network Under Test	-	-	-	-	Symmetrical	10 Mbps	10 Mbps
Server (Remote)	00-21-c2-00-06-1e	192.168.1.222	255.255.255.0	192.168.1.1	-	-	-

TCP Setup

Single TCP connection

TCP Setup

No of TCP Connection ▼

TCP Port Configuration Automatic Manual

TCP Connection No.	Client Port	Server Port
1	5000	6000

Multiple TCP connections

TCP Setup

No of TCP Connection ▼

TCP Port Configuration Automatic Manual

TCP Connection No.	Client Port	Server Port
1	5000	6000
2	5001	6001
3	5002	6002
4	5003	6003
5	5004	6004
6	5005	6005
7	5006	6006
8	5007	6007

TCP Setup (Contd.)

The screenshot shows a 'Test Setup' window with the following configuration:

- Direction:** Upstream (indicated by a green up arrow), Downstream (indicated by a green down arrow), Upstream and Downstream (indicated by a green double-headed arrow).
- Transfer Size:** 100.000 MBytes.
- Test Selection:**
 - Run Throughput Test
 - Run Path MTU Test
 - Upstream MTU: 1500 Bytes
 - Downstream MTU: 1500 Bytes
 - Run Baseline RTT Test
 - Upstream RTT: 250.049 msec
 - Downstream RTT: 250.030 msec

Upstream/Downstream/Bidirectional

- Path MTU - run test and discover or user can enter manually
- Baseline RTT - run test and find out or user can enter manually
- Separate Path MTU/Baseline RTT configuration for Upstream/Downstream directions for asymmetrical paths

Status and Results

Overall Status

Test Status: Done

Current Direction: -

Current Test

Test	Status	Result
Path MTU (Upstream)	↑	✓
Baseline RTT (Upstream)	↑	✓
Throughput (Upstream)	↑	✓

TCP Connection Status:

Connection No.	Source Port	Destination Port	Status
0	5000	6000	Connection Closed

Path MTU results

Upstream Downstream

Path MTU: 1500 Bytes

Baseline RTT Results

Upstream Downstream

Trial Duration: 91

Average RTT: 50.018 msec

Minimum RTT: 50.015 msec

Maximum RTT: 50.040 msec

Baseline RTT Value Selected: 50.015 msec

Test Parameter Summary

Upstream Downstream

Baseline RTT: 50.015 msec

Calculated BDP: 625.190 KBytes

TCP Window: 65535 Bytes

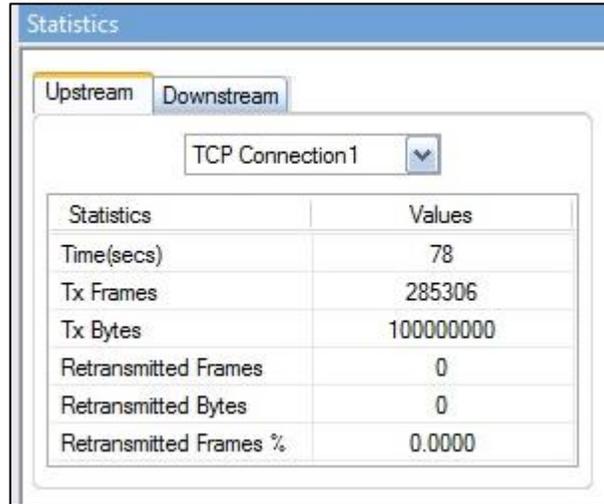
Path MTU: 1500 Bytes

MSS Used: 1448 Bytes

No of TCP Connection: 1

Transfer Size: 100.000 MBytes

Statistics and Periodic Results

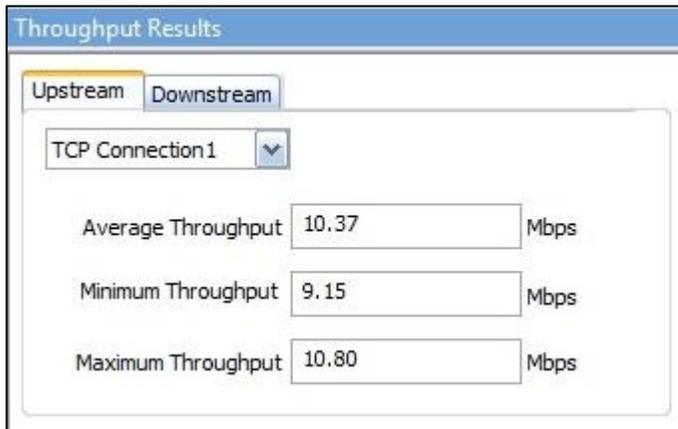


The screenshot shows a window titled "Statistics" with tabs for "Upstream" and "Downstream". A dropdown menu is set to "TCP Connection 1". Below is a table with two columns: "Statistics" and "Values".

Statistics	Values
Time(secs)	78
Tx Frames	285306
Tx Bytes	100000000
Retransmitted Frames	0
Retransmitted Bytes	0
Retransmitted Frames %	0.0000

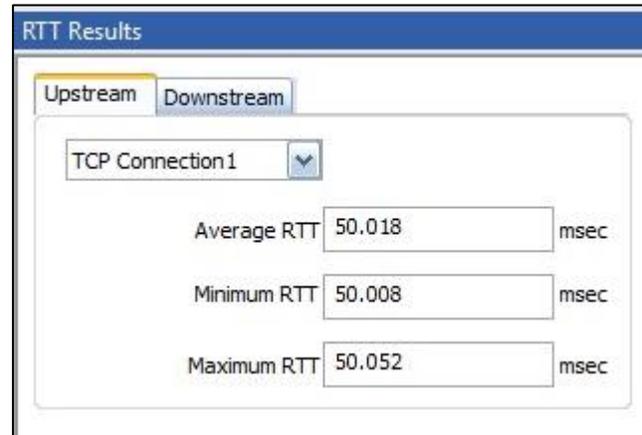
Statistics are updated every second and includes -

- TCP Transmitted Frames/Bytes
- TCP Retransmitted Frames/Bytes
- Retransmitted Bytes Percentage



The screenshot shows a window titled "Throughput Results" with tabs for "Upstream" and "Downstream". A dropdown menu is set to "TCP Connection 1". Below are three rows of throughput data, each with a text input field and the unit "Mbps".

Average Throughput	10.37	Mbps
Minimum Throughput	9.15	Mbps
Maximum Throughput	10.80	Mbps



The screenshot shows a window titled "RTT Results" with tabs for "Upstream" and "Downstream". A dropdown menu is set to "TCP Connection 1". Below are three rows of RTT data, each with a text input field and the unit "msec".

Average RTT	50.018	msec
Minimum RTT	50.008	msec
Maximum RTT	50.052	msec

Throughput and RTT values are calculated every second and displayed. Minimum, Maximum and Average Values are displayed.

Final Results

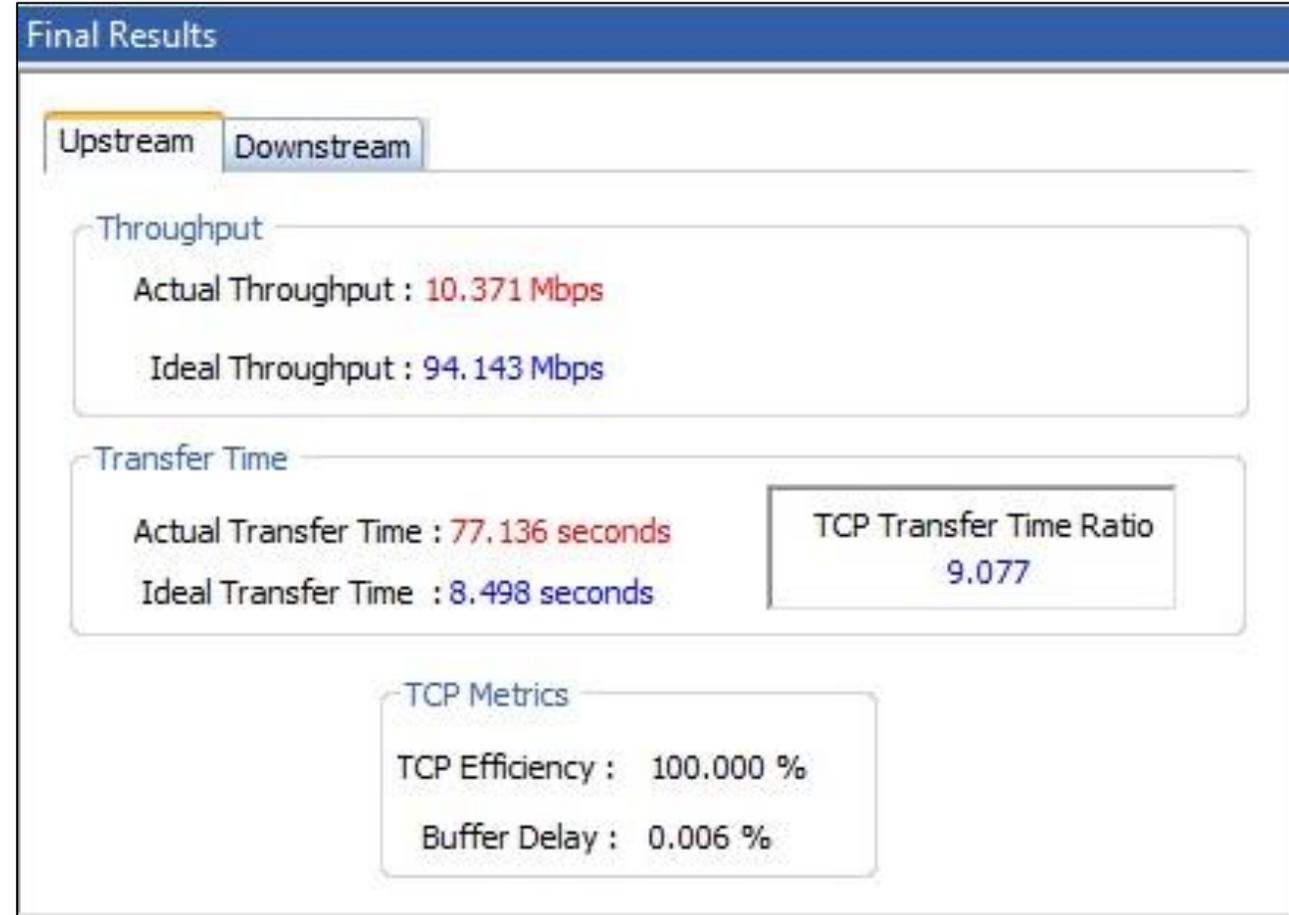
Ideal Throughput - the maximum possible TCP throughput for the given CIR.

Ideal Transfer Time - the time taken to transfer the test data size at the ideal throughput.

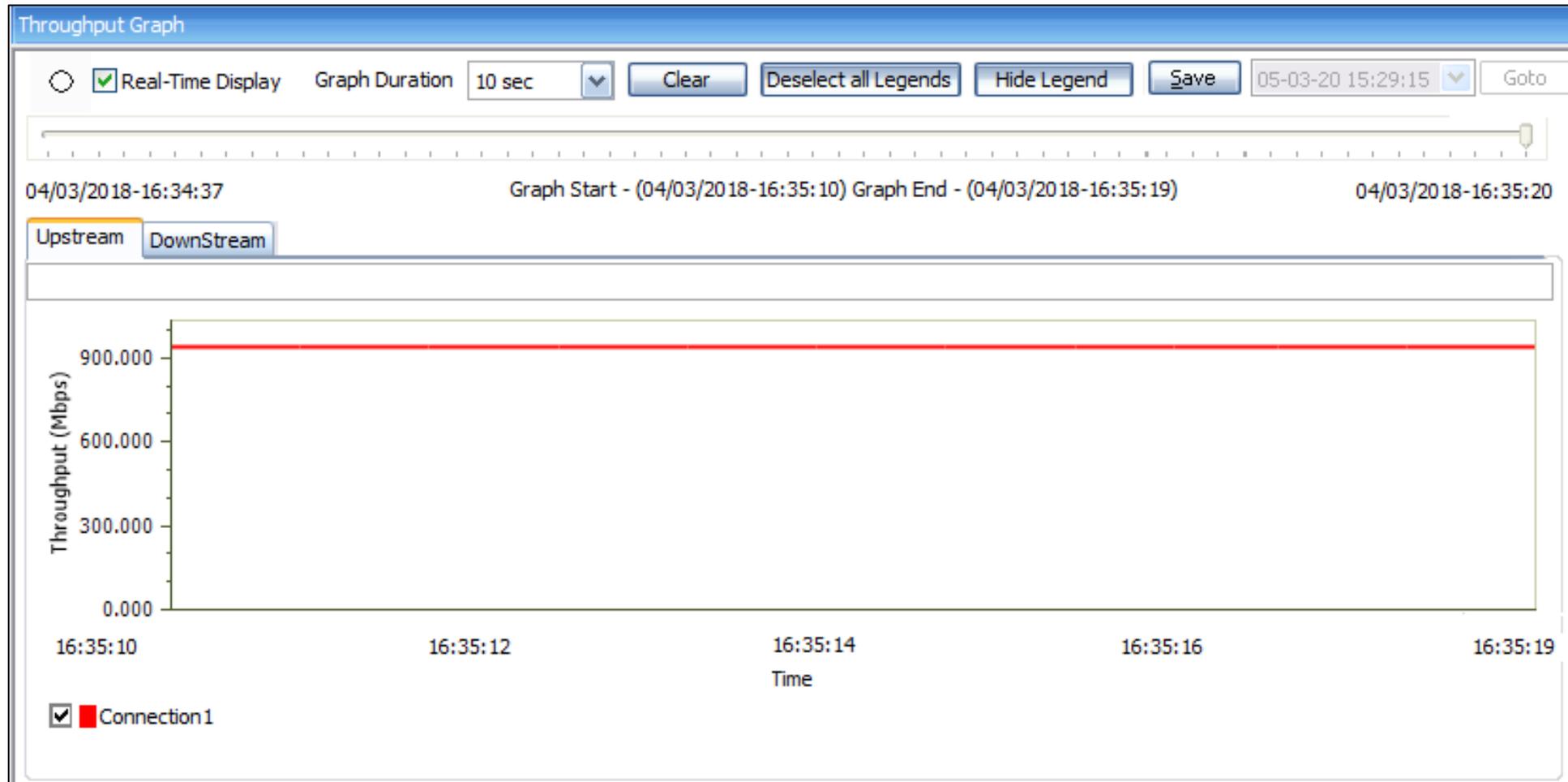
TCP Transfer Time Ratio - Measure of how much Actual transfer time is greater than the Ideal transfer time.

TCP Efficiency - measure of the number of Transmitted bytes compared to the retransmitted bytes.

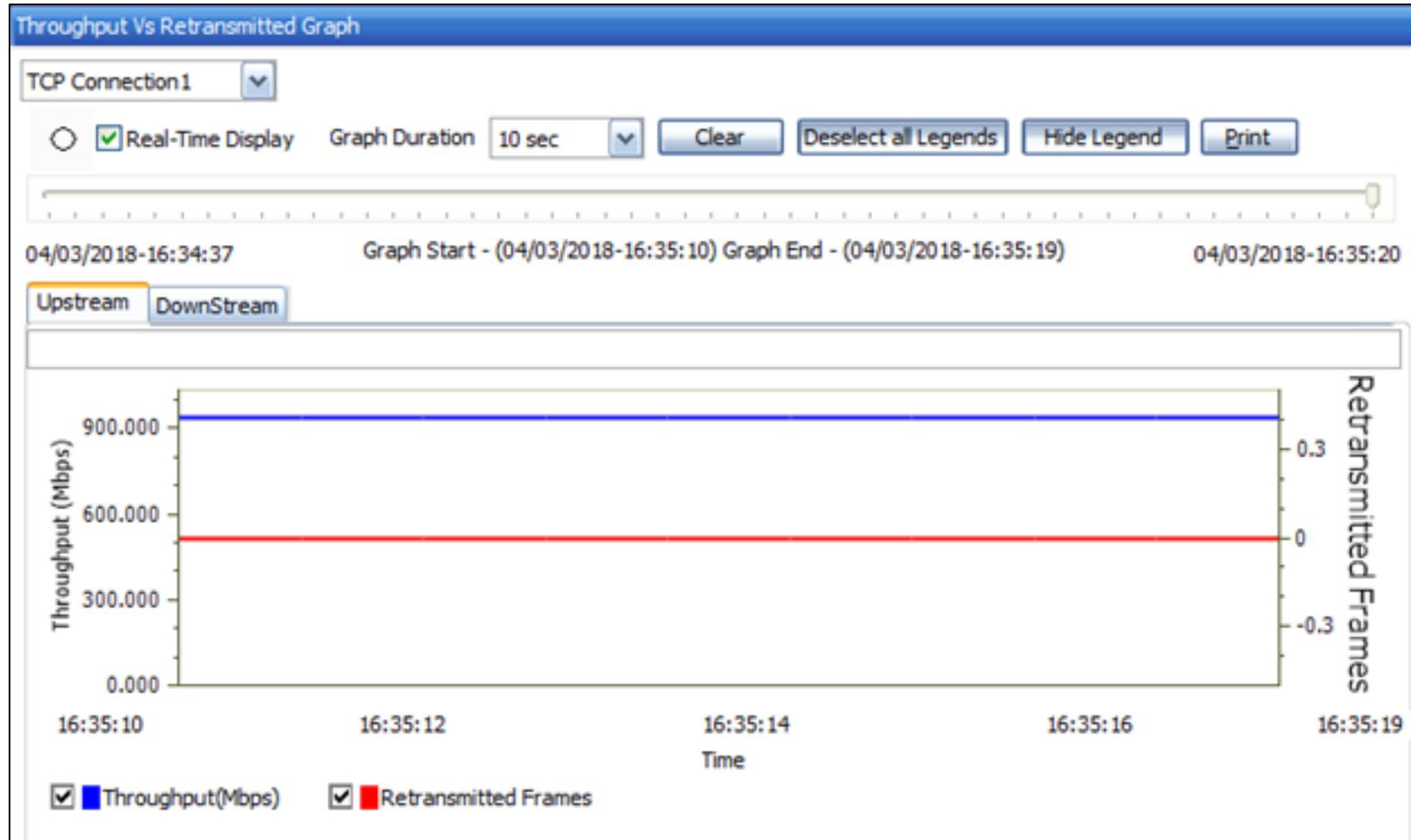
Buffer Delay - measure of how much the RTT increases during the actual TCP Throughput test compared to the Baseline RTT.



Throughput Graph



Throughput vs. Retransmitted Frames Graph



Multiple TCP connections

With 8 TCP connections

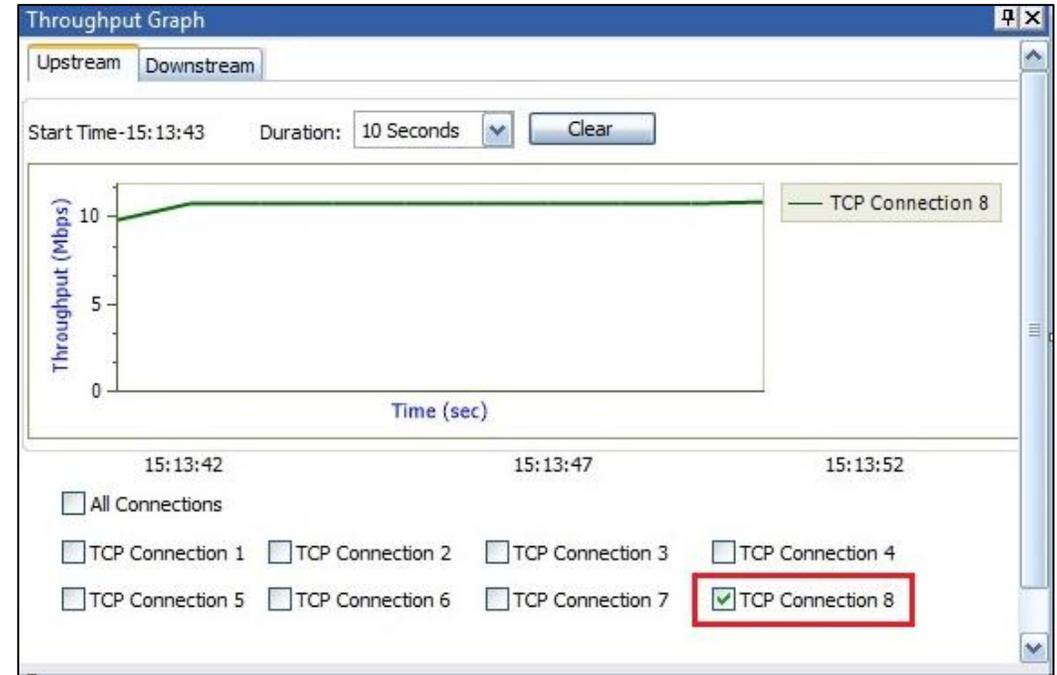
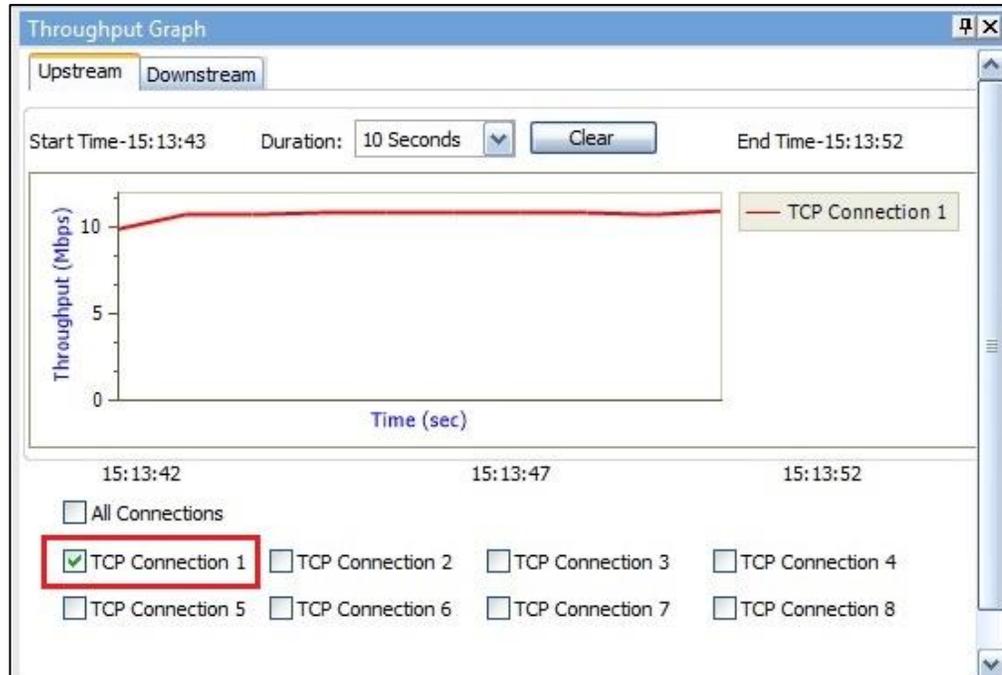
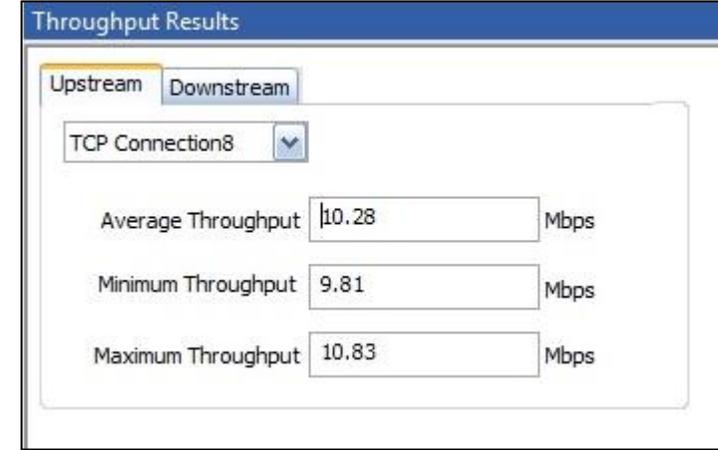
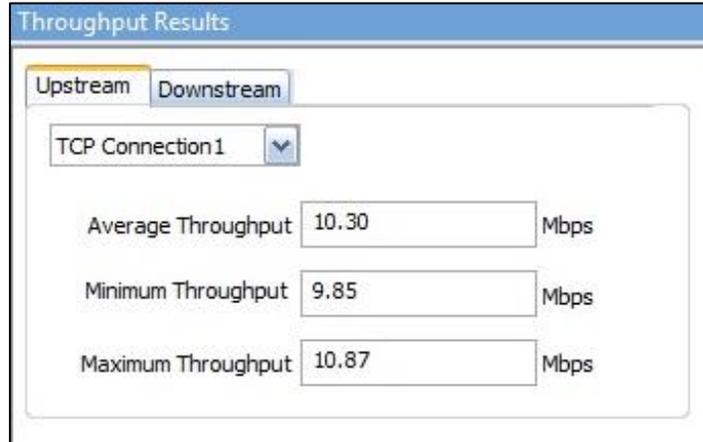
The screenshot shows a window titled "Test Parameter Summary" with two tabs: "Upstream" and "Downstream". The "Downstream" tab is active. The window displays several network parameters in a list format:

- Baseline RTT: 50.022 msec
- Calculated BDP: 625.274 KBytes
- TCP Window: 524280 Bytes** (highlighted with a red box)
- Path MTU: 1500 Bytes
- MSS Used: 1448 Bytes
- No of TCP Connection: 8
- Transfer Size: 100.000 MBytes

A red arrow points from the "TCP Window" value to the text: "TCP window of 5,24,280 bytes shared among 8 connections".

Multiple TCP Connections - Throughput

Individual Throughput for each connection



Multiple TCP Connections - Result

Improved Overall Throughput



Thank you