ExpertTCP™ - TCP Throughput Testing (per RFC-6349)
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**
Outline..

• TCP Throughput Measurement
  • Path MTU Discovery
  • Round Trip Time Measurement
  • Measure TCP Throughput

• Screenshot

• Video
Performance Testing of Packet / Ethernet Connections and Networks

For Predictable Managed Networks

- RFC-2544
- ITU Y.1564 (SAM)
- RFC-6349 (TCP)

Service Level Agreements from Network Providers, a must
User Experience, Application-Network Sensitive, TCP Tuning

SAM – Service Activation Methodology
TCP – Transmission Control Protocol
Packet / Ethernet Testing

User Experience

Application
- HTTP, FTP, Email, Facebook, Youtube, etc.

Transport
- ExpertTCP (RFC 6349 testing)

Network
- RFC 2544
- ExpertSAM (Y.1564 testing)

Datalink
- Network Throughput
- Latency
- Packet Loss
- Back-to-Back
- Jitter

Physical
- End-to-End Throughput
## Typical SLA

### EXHIBIT D – Service Level Agreements

**1. Service Level Agreement Matrix**

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

**Typically**

- **Packet Loss**: 0.0005% to 1%
- **Latency**: 36 to 75 ms
- **Availability**: 99% to 99.9%
RFC-2544 vs. ITU Y.1564 (SAM)

Both are Connection-less

- Throughput
- Latency
- Frame Loss
- Back-to-Back
- Jitter
RFC-2544 Testing

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.

**Background**

- **Dual Port RFC-2544**
  - Tx (Transmission)
  - Rx (Reception)

- **Single Port RFC-2544**
  - Tx TIME (Transmission Time)
  - Rx TIME (Reception Time)

**Network**

- Layer 2 (Stacked VLAN), MPLS, IP/UDP

**Frame Loopback**
ITU Y.1564 (SAM)

- Throughput
- Latency
- Packet Loss
- Jitter
### Testing Relevance

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

- Running RFC-2544, Y.1564 or other 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.
Background

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
- # of TCP Simultaneous Connections
Bandwidth * Delay Product (Bits or Bytes)

Application and Network are Matched, TCP is Tuned

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), $B \text{ (bps)} \times RTT \text{ (secs)} = \text{BDP bits}$

B = 10 Mbps
RTT = 50 ms

B*50 = 500,000 bits or 62,500 Bytes

65,535 Bytes is max window

Achieving max throughput:

GL Communications Inc.
Effect of Increased Network Delay or Smaller Tx or Rx Buffers

**Background**

- **B = 10 Mbps**
- **RTT = 100 ms**

- $B \times 100 = 1,000,000$ 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

**Background**

- \( B = 20 \text{ Mbps} \)
- \( \text{RTT} = 50 \text{ ms} \)

\[ B \times 50 = 100,000 \text{ bits} \]

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 +/- 20 ms, 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 & Retransmission Scheme

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

<table>
<thead>
<tr>
<th>Packet Loss</th>
<th>TCP Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 %</td>
<td>100%</td>
</tr>
<tr>
<td>0.1 %</td>
<td>&lt; 50%</td>
</tr>
<tr>
<td>1 %</td>
<td>&lt; 10%</td>
</tr>
<tr>
<td>2 %</td>
<td>0 %</td>
</tr>
</tbody>
</table>

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 & Retransmission Scheme (cont.)

Background

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

<table>
<thead>
<tr>
<th>Packet Loss</th>
<th>TCP Throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 %</td>
<td>100%</td>
</tr>
<tr>
<td>0.1 %</td>
<td>&gt; 99 %</td>
</tr>
<tr>
<td>1 %</td>
<td>&gt; 95 %</td>
</tr>
<tr>
<td>2 %</td>
<td>? %</td>
</tr>
</tbody>
</table>

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:

1. **Path MTU Discovery** - What is the maximum packet size that can successfully traverse the network?

2. **Round Trip Time (RTT) Measurement** - Timestamp based RTT discovery of transmitted packet until acknowledgement packet arrives from far end.

3. **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

Measure
- Path MTU
- RTT
- TCP Throughput
PacketExpert™ 10G Standalone
- 2 x 1 Gbps Optical OR Electrical
- 2 x 10 Gbps Optical only

PacketExpert™ 10GX Standalone
- 4 x 1 Gbps Optical OR Electrical
- 2 x 10 Gbps Optical only

PacketExpert™ 1G (4 Port)

PacketExpert™ 10G Tablet Inspired
(Coming soon)
mTOP™ 1U/2U Rack Option

- 19” rack option, w/ Embedded Single Board Computer (SBC)
- SBC Specs: Intel Core i3 Equivalent, Windows® 10 64-bit Pro, USB 3.0 Hub, ATX Power Supply, 240GB Hard drive, 8G Memory (Min), Two HDMI ports for display
ExpertTCP™ 1G Ports

- TCP Client and Server will be supported in different applications.
- In 1G, Port 1, Port 2, Port 3, Port 4 are used.
ExpertTCP™ 10G Ports

- TCP Client and Server will be supported in different applications.
- In 10G, Port 3 and Port 4 are used.
Step 1. Path MTU Discovery

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

Network device rejects packets that are bigger than the supported MTU, and will not forward the packet towards server

TCP ACK does not reach client, making the client try again with a different TCP packet size
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

Sender  

Network  

Receiver

<table>
<thead>
<tr>
<th>Time : 50</th>
<th>Seg1: TS value = 50</th>
<th>TS Echo reply=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time : 100</td>
<td>Seg2: TS value = 100</td>
<td>TS Echo reply=0</td>
</tr>
<tr>
<td></td>
<td>Ack on Seg1: TS value = 0</td>
<td>TS Echo reply = 50</td>
</tr>
<tr>
<td></td>
<td>Ack on Seg2: TS value = 0</td>
<td>TS Echo reply = 100</td>
</tr>
</tbody>
</table>

(RTT = 150 – 50 = 100)

(RTT = 170 – 100 = 70)
Step 3. Now Ready to Measure TCP Throughput

Slow Start
(Binary Search)

Congestion Avoidance
(TCP Equilibrium)
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

- CWND
- Time Out
- Duplication of ACKs received
- Some ACKs received
- All ACKs received

- sssthresh (initial)
- sssthresh (new)

- Acks for the first 16 TCP segments

- Slow Start
- Congestion Avoidance
- Slow Start
- Congestion Avoidance
Screenshots of Software Operation
ExpertTCP™ Main Screen
Test Setup with Impairments

Client

ExpertTCP™

IPLinkSim

0% Errors
25ms

USB

25ms
0% Errors

Server

ExpertTCP™
Network Setup

All settings configured locally on the client side
Network Setup (cont...) 

Separate Upstream and Downstream bandwidths configurable for asymmetrical path
TCP Setup

Single TCP connection

Multiple TCP connections
TCP Setup (cont...)

- 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:
  - Path MTU (Upstream)
  - Baseline RTT (Upstream)
  - Throughput (Upstream)

TCP Connection Status:
- Connection No.: 0
- Source Port: 5000
- Destination Port: 6000
- Status: Connection Closed

Path MTU results:
- Path MTU: 1500 Bytes

Baseline RTT Results:
- Trial Duration: 91 msecs
- Average RTT: 50.018 msecs
- Minimum RTT: 50.015 msecs
- Maximum RTT: 50.040 msecs
- Baseline RTT Value Selected: 50.015 msecs

Test Parameter Summary:
- Baseline RTT: 50.015 msecs
- 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

Statistics are updated every second and includes:

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

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

With 0.1% Packet Loss
Throughput vs. Retransmitted Frames Graph

With 0.1% Packet Loss
Multiple TCP connections

With 8 TCP connections

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline RTT</td>
<td>50.022 msec</td>
</tr>
<tr>
<td>Calculated BDP</td>
<td>625.274 KBytes</td>
</tr>
<tr>
<td>TCP Window</td>
<td>524280 Bytes</td>
</tr>
<tr>
<td>Path MTU</td>
<td>1500 Bytes</td>
</tr>
<tr>
<td>MSS Used</td>
<td>1448 Bytes</td>
</tr>
<tr>
<td>No of TCP Connection</td>
<td>8</td>
</tr>
<tr>
<td>Transfer Size</td>
<td>100.000 MBytes</td>
</tr>
</tbody>
</table>

TCP window of 5,242,800 bytes shared among 8 connections
Multiple TCP Connections - Throughput

Individual Throughput for each connection
Multiple TCP Connections - Final Result

Improved Overall Throughput
THANK YOU
Questions