Understanding Carrier Ethernet Throughput

Am I getting the throughput I should be getting?

2010 July Version 2

Table of Contents

Table of Contents

Introduction 3
Identifying Throughput Problems 4
Basic MEF Service Concepts 5
  Protocol Basics for Carrier Ethernet 5
  The Basic Elements of Service 7
Factors Affecting Throughput 9
  Customer Challenges in Conforming to the SLA 9
  The Impacts of Frame Size and Protocol on Throughput Efficiency 10
  Standard TCP Window Size and Effect on Throughput 12
  Using Window Scaling to Increase TCP Throughput 14
  Dropped/Errored Frames and the Retransmission Effect 15
Monitoring Throughput 16
  General Network Management Tools 16
Verifying Carrier Ethernet Performance 18
  Troubleshooting with PC Based Test Tools 18
  Software Based Test Tools 18
  Hardware Based Test Tools 19
Summary and Recommendations 19
Glossary Of Abbreviations 20
References 21
Acknowledgements 22

Introduction

Globalization, virtualization, and mobile computing drive a seemingly insatiable demand for bandwidth, and only Carrier Ethernet efficiently scales up to meet this demand. Customers seeking high performance business Ethernet services can now easily purchase faster Ethernet connections at 10 Mbit/s to 1 Gbit/s and beyond. But sometimes users believe they are receiving lower throughput than they expected. This perception can be due to poor application performance which is caused by factors un-related to Ethernet service throughput. Many IP and application layer factors affect a user's application experience when utilizing an Ethernet service, most of which are under their own direct control.

First and foremost, obtaining a good service requires selecting an Ethernet service provider that is MEF certified to deliver a high quality Carrier Ethernet service. Secondly, Enterprise users must ensure that they are shaping the bandwidth offered to the network to match the bandwidth profile of the service level agreement (SLA). For example, driving a 50Mbit/s Ethernet service with 100 Mbit/s for a time period larger than the contracted Committed Burst Size (CBS) will provide a poor user experience, with dropped traffic, re-transmissions, and net throughputs which are much lower than expected. Other key application issues include optimally setting Transport Control Protocol (TCP) window size on applications which require higher speed or are delivered over services with longer delay. For instance, TCP window limitations can be seen on Ethernet services transmitting at information rates as low as 13 Mbit/s when combined with transmission delays in the range of 40 ms. In addition, the Ethernet frame size, the selection of higher layer protocols and error rates can all affect both delay and throughput of an application being delivered over an Ethernet service.

This MEF white paper presents an overview of the more common factors affecting Carrier Ethernet throughput, provides some pointers for getting more performance from higher layer protocols, and shows how to measure bandwidth throughput of a Carrier Ethernet service. Note that to simplify the scope of this discussion; the focus of the whitepaper is on E-Line services that are point to point. More sophisticated E-LAN and E-Tree services are influenced by the same factors plus some additional factors such as multicast/broadcast storms which are specific to their multi-point topologies.

Identifying Throughput Problems

Ethernet service providers report that a significant portion of customer trouble tickets are opened due to poor application performance. Typically, enterprise customers call their service provider when the transfer rate between the host and servers across the service provider's network appears to be slow (below the contracted throughput rate). When they think there might be a problem, some end users "test" the rate by running a file transfer between two sites (as measured by their operating system). By looking at the transfer rate shown in Figure 1, an end user would presume that the maximum rate of their link is 339 KByte/s, or 2.78 Mbit/s. If the contracted service is supposed to be at 50 Mbit/s committed information rate (CIR) as in Figure 2, the user will be frustrated.

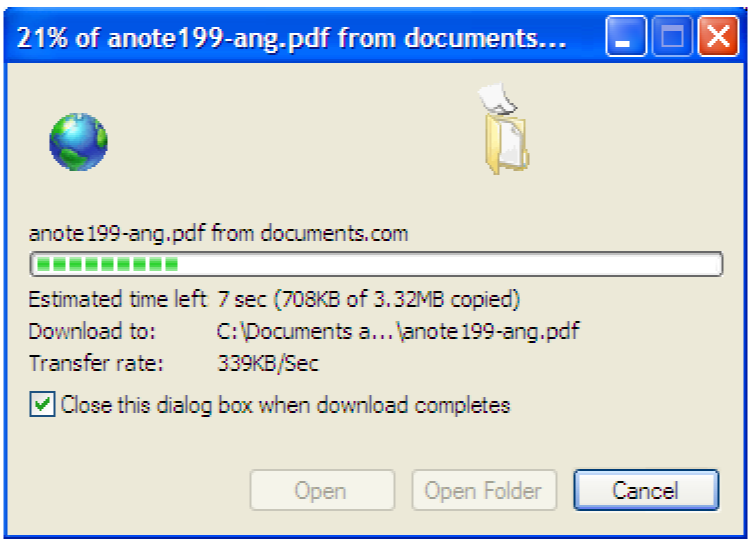

Figure 1. Windows XP file transfer dialog box showing transfer rate

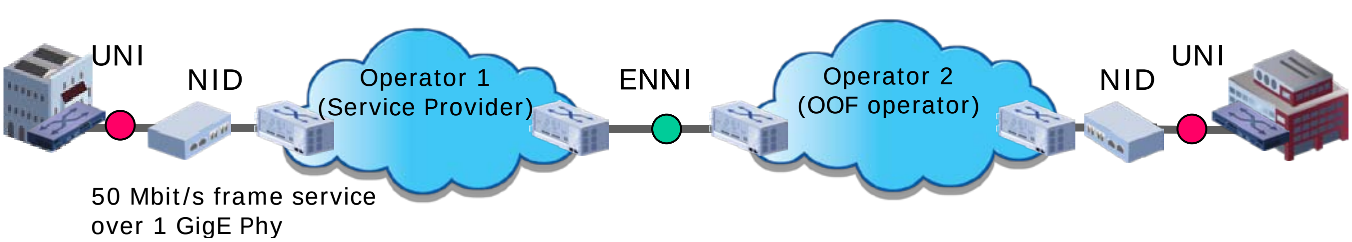

Figure 2. Customer's high performance 50 Mbit/s Ethernet Virtual Connection

Note that in Figure 2, the customer has purchased a 50 Mb/s E-Line service, but has a 1 Gbit/s Ethernet physical interface User Network Interface (UNI). A Network Interface

Device, (NID), may be installed on his premises by the Service Provider to provide various network terminating functions. The Service Provider (Operator1) partners with another OOF (Out of Franchise) Operator 2, interconnecting their networks at the External Network to Network Interface (ENNI). The two carriers have constructed an Ethernet Virtual Connection (EVC) to implement the E-Line service.

The end user with the throughput problem may react by blaming the service providers for the lower achieved data rate because he or she thinks the network does not perform as detailed in their service level agreements (SLAs). As these users complain to their IT departments, the IT personnel will often test the links with more advanced software-based tools to validate these claims. However, because their tools are PC-based, they also come with all of the limitations found in PC operating systems. Should a less seasoned IT professional use those tools, he might come to the same conclusion as the end user and open a trouble ticket with the service provider.

Basic MEF Service Concepts

Protocol Basics for Carrier Ethernet

Before discussing the issues and behaviors of Ethernet Services and test tools in greater detail, it is important to understand the protocols found on Ethernet links.

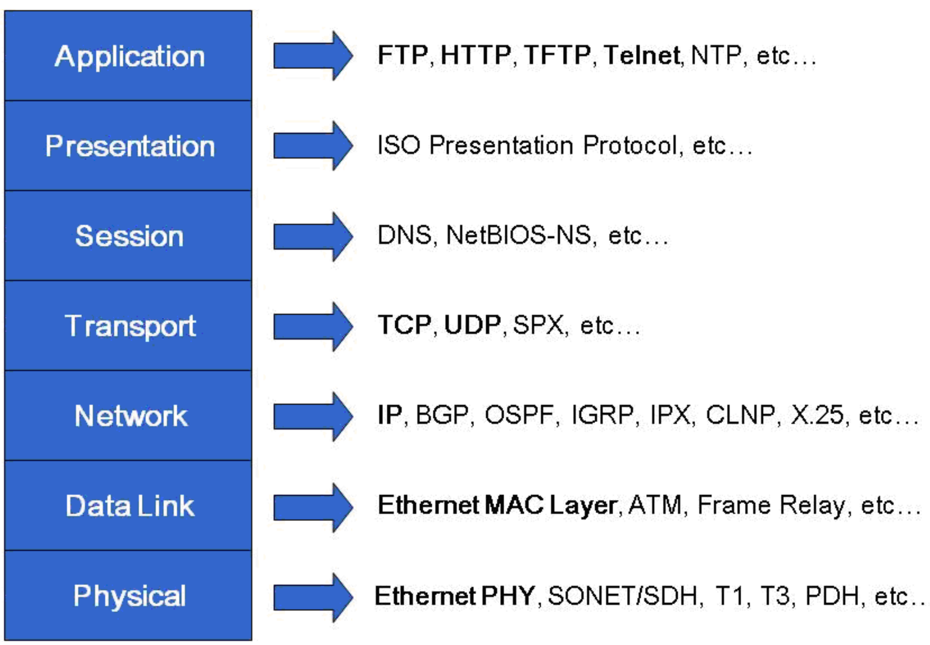

Figure 3. OSI Reference Model with examples of protocols for each layer

The Open Systems Interconnect (OSI) model is shown above and is used as a model for developing data network protocols. Each layer works with the layers above and below them to enable communication with the same layer of another stack instance. Carrier

Ethernet primarily is defined in Layers 1 and 2. Other familiar layers would be the IP layer (Layer 3) as well as the TCP or UDP protocols found in layer 4. Note that the FTP session in Figure 1 generally rides over the TCP protocol in layer 4.

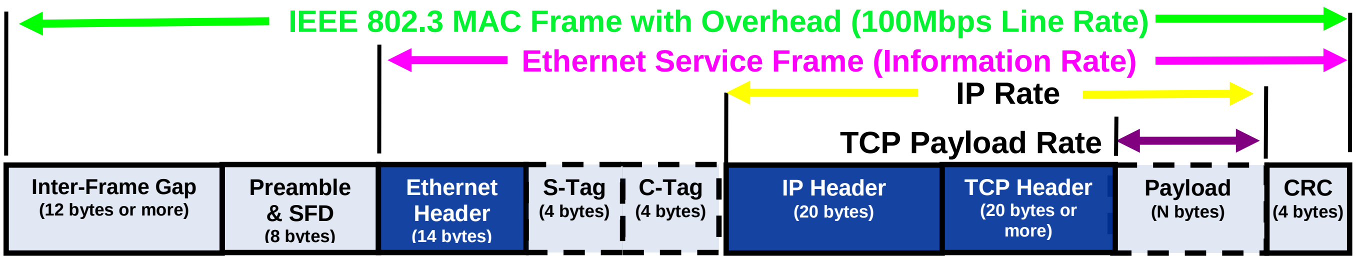

Figure 4. Sample TCP/IP Overhead

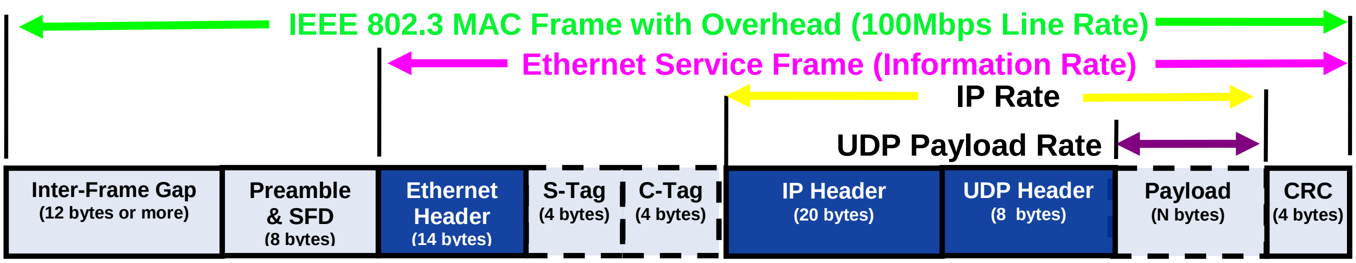

Figure 5. Sample UDP Overhead

Carrier Ethernet consists of frames transported across the network. We can see how the various layers occupy the frame in Figure 4 and Figure 5. At the physical layer, frames are separated by an Inter-Frame Gap, and then a Preamble and Start of Frame Delimiter (SFD) serves to align the receiver on the frame to come. The Ethernet header provides most of the Layer 2 local area network addressing information. S-Tags and C-Tags may be applied by the service provider and customer, respectively, to identify specific virtual circuits. Further bytes located farther into the packet handle the layers above Layer 2. Note that the Committed Information Rate (CIR) is defined in MEF 10.2 the rate in bits per second of all the bytes designated as Ethernet Frame in Figure 4 and Figure 5. The Inter-Frame Gap, Preamble & SFD are not counted towards the Committed Information Rate of an Ethernet Virtual Connection.

The Transport layer handles end-to-end connections and reliability of applications. There are two main protocols at the transport layer for the IP protocol suite — the transport control protocol (TCP) and the user datagram protocol (UDP). These two protocols are the basis of all modern data communications. Depending on the application, it will either use TCP or UDP.

UDP, a more basic transport layer protocol, was created to be as simple as possible so basic information could be transported between two hosts without requiring set up of special transmission channels or data paths. UDP's simplified structure makes it connectionless and less reliable, but very fast with no upward limits on throughput. UDP messages can be lost, duplicated or arrive out-of-sequence. With no flow control, messages can arrive faster than they can be processed. UDP relies on the application layer for everything related to error validation, flow control and retransmission.

UDP is stateless by nature (meaning that each message is seen as an independent transaction). This behavior is used by real-time applications such as IPTV, VoIP, Trivial File Transfer Protocol (TFTP) and network gaming, which benefit from the protocol's low latency (total delay of getting data from one application to another through the network) and lack of speed limitations.

TCP, on the other hand, is a connection-oriented protocol that requires handshaking to set up end-to-end communications. TCP provides the following functions and benefits:

- Reliable, transparent transfer of data between networked end points
- End-to-end error detection, recovery and data flow control
- Sequential numbering of frames to keep track of them if they arrive out of order
- Segmentation and reassembly of user data and higher-layer protocols
- Adaptive transmission data rate control to utilize the available speed of the link

TCP is not perfect however. These details will be covered in more depth in the following sections, but here is a short summary. TCP has more overhead bytes than UDP, with 20 or more bytes in TCP's overhead compared with only 8 bytes in UDP. In TCP, extra bytes are used for things like window scaling and selective acknowledgment. With TCP there is some delay in setting up a reliable link, because it goes through a significant handshaking procedure to start the session. TCP's continuous acknowledgments and need to buffer data for potential retransmission can bog down processor performance. Finally, TCP's performance can be significantly compromised by transmission latency, or Frame Delay. This latency is increased by many factors:

- Longer distance of transmission
- Slower propagation velocity of transmission
- A large quantity of network elements, and longer delay introduced by each of the network elements
- Use of large Ethernet frames

Having taken a look at some basic elements of the protocols, now let's delve further into why we may not be getting the throughput we think we should.

The Basic Elements of Service

The user should understand the basic elements of the Service Level Agreement (SLA). Users purchase bandwidth, called Committed Information Rate (CIR), from the Service Provider. Because the user's traffic may be bursty in nature, a Committed Burst Size (CBS) may also be specified to guarantee SLA performance levels on bursts of frames which exceed the CIR by the number of bytes in the Committed Burst Size. Note that the CIR will typically be far less than the speed of the physical interface to which the customer is attached. For instance in Figure 2, the CIR is 50 Mbit/s and the physical interface is Gigabit Ethernet. If the customer inputs data to the network in accordance with these limits, the Service Provider SLA guarantees to deliver the traffic at the CIR meeting certain performance objectives for Frame Delay, Frame Delay Variation, and Frame Loss Ratio.

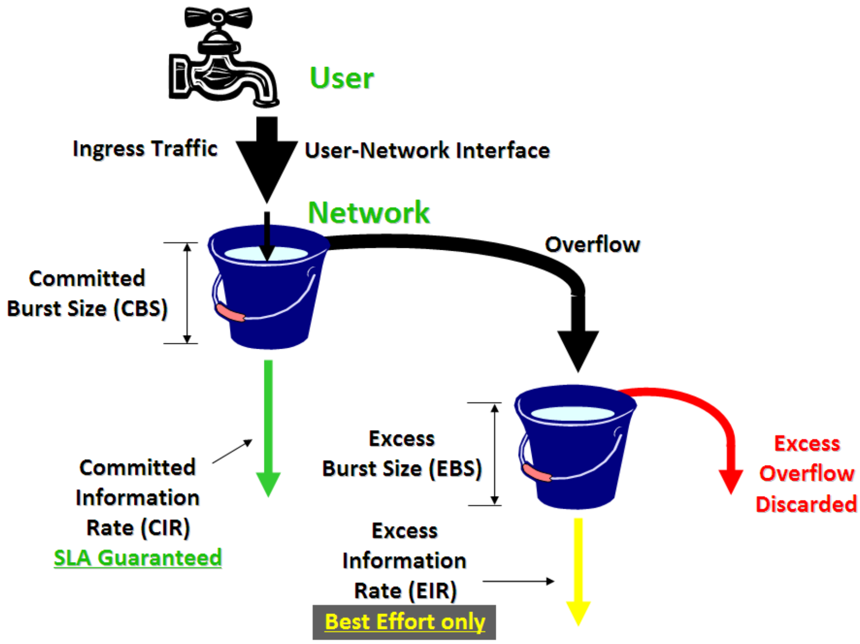

Figure 6. Leaking Buckets Model of Carrier Ethernet Service

The Service Provider may also offer an Excess Information Rate (EIR) and Excess Burst Size (EBS) in which it agrees to carry the traffic if there are no congestion problems in the network. Traffic which conforms to the CIR/CBS criteria is called green traffic and treated according to the SLA. Traffic which is above the CIR/CBS rate/size but within the EIR/EBS is labeled as yellow traffic in order that it can be dropped anywhere within the network that congestion is a problem. Yellow traffic is delivered on a best efforts basis and is not included in the SLA calculation. Traffic that exceeds the EIR/EBS is called red traffic and should be dropped immediately upon entry into the service provider's network. This traffic management procedure is loosely described in Figure 6.

In this model, the customer's data flows out of the "tap" and into the network in bursts of various size and flow rate. A "bucket" catches the flow and provides for the managed entry of the data into the network out the hole in the bottom of the bucket. If too much data is put into the network, the bucket overflows. A second bucket may capture this data in the same fashion, giving it additional protection with additional managed data entry into the network at the EIR. If the second bucket overflows or there is no second bucket, the overflowing traffic should be discarded.

Unlike real buckets leaking water, however, frames entering the network are immediately classified as green, yellow or red and transmitted onward without delay if network congestion allows.

Another way to think of this process is that the customer's traffic has to go through a combination bandwidth profiler and policer as soon as it enters the service provider's network. The bandwidth profiler characterizes each frame as green, yellow, or red, depending on how it matches the purchased bandwidth profile for the service. Any red traffic is immediately discarded by the policer. This behavior is shown in Figure 7

Figure 7. Bandwidth Profiling with Policing

Factors Affecting Throughput

Customer Challenges in Conforming to the SLA

The transition from legacy services such as T1, T3, Frame Relay and ATM to Carrier Ethernet has created some unintended consequences. Not all customers have conforming equipment facing the network which properly limits/shapes the traffic outbound to the network, with deleterious results. For instance, on the 1 GigE interface of Figure 2, if the customer's equipment accidentally transmits long bursts of data at 150 Mbit/s instead of the SLA's Committed Information Rate of 50 Mbit/s, 67% of the data may be lost and network breakdown will likely result. If the committed burst size is 30 KBytes, the committed burst will be used up and the traffic policer will start discarding service frames in about 2.4 milliseconds. Another Carrier Ethernet implementation issue could be that customer equipment may expect the network to react to layer 3 Pause Frame flow control protocol, but Pause Frames are not appropriate for Carrier Ethernet because they don't properly differentiate between individual layer 4 traffic streams that can be paused, like file download, and layer 4 streams that can't be paused, like VoIP. Carrier Ethernet instead requires the customer's equipment to self-limit to the SLA CIR. Finally, should CPE switches or routers not support shaping or CIR conventions at all, this limitation may be accommodated by carefully engineering the CBS and the CIR on the Ethernet service to efficiently handle whatever natural Ethernet traffic the network generates.

The Impacts of Frame Size and Protocol on Throughput Efficiency

Layer 4 protocols such as TCP and UDP allow the user to select different frame sizes for transmission. On applications that need a lot of bandwidth, larger frames have the effect of generating better payload utilization of that bandwidth, because the substantial overhead of each frame is spread out over many more payload or application bytes. But in some cases, end users may want to use small frames with inefficient overhead in order to reduce latency for applications such as VoIP. Figure 8 shows this behavior by plotting bit rates for the line, Ethernet service frames (called Information Rate), IP frame, UDP payload, and TCP payload transmission against the corresponding frame sizes. These four layered protocols are shown to illustrate how the bandwidth of the line gets eaten up by each additional layer of protocol operating over the line. Given an underlying line rate of 100 Mbit/s, the plot shows the maximum theoretical throughput per protocol type and does not take into account other throughput factors such as TCP protocol handshaking. Note that the Information Rate referred to in Committed Information Rate is the bit rate of the Ethernet frames, the top curved line in the Figure 8. Information Rate includes all the bits in the Ethernet frame starting with the MAC address and ending with the Frame Check Sequence, and excludes the Inter Frame Gap, the Preamble, and the Start of Frame Delimiter. Figure 4 and Figure 5 provide the basic frame structure, overhead size and some terminology used in Figure 8, Maximum Throughput at varying Frame Sizes at 100 Mbit/s Line Rate.

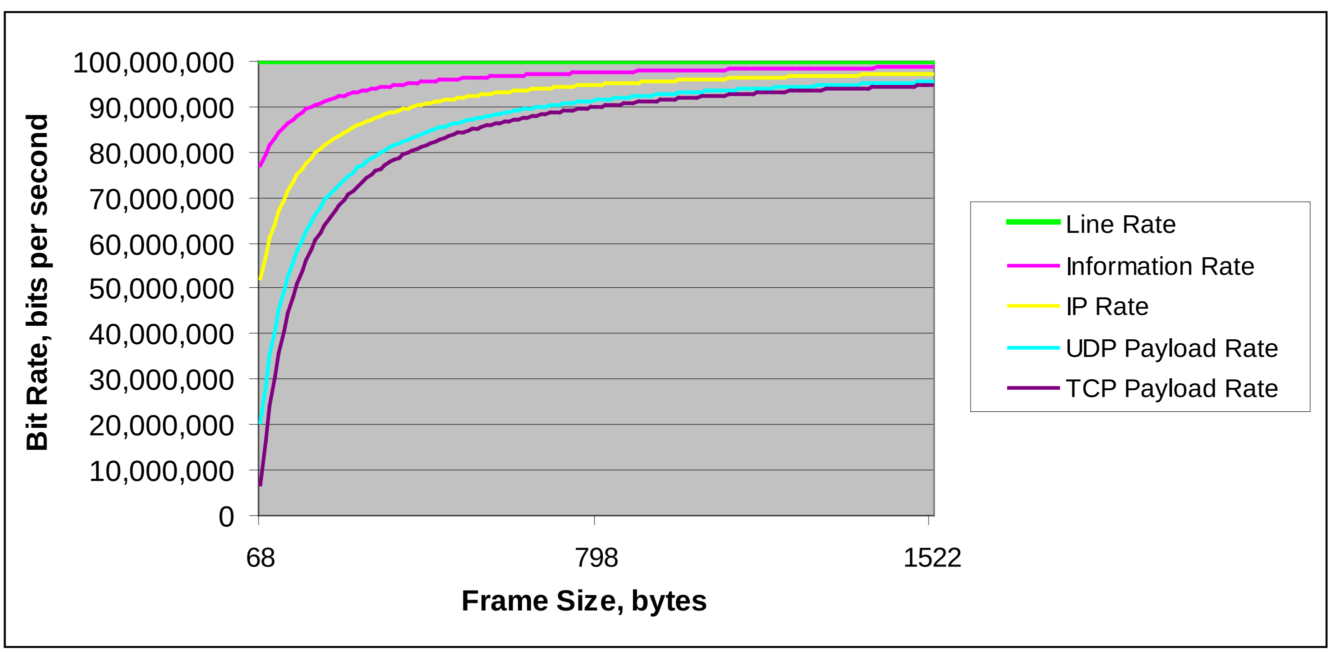

Figure 8. Maximum Throughput at varying Frame Sizes at 100Mbit/s Line Rate

Another subtlety in addressing Carrier Ethernet is the difference between line rate, the utilized line rate and the information rate. The line rate is the commonly known rate such as 100 Mbit/s. The line rate is the number of bits per second transmitted including all the bits of the Ethernet frame, plus the interframe gap, the preamble, and the start of frame delimiter. The utilized line rate is the same as the line rate, except that it counts only the minimum number of bits for each interframe gap. If the customer is not transmitting Ethernet frames, then the line is not being utilized, and there will be a very long interframe gap much longer than the 12 byte minimum. The information rate is the number of bits that get transmitted in a second, when counting just the Ethernet frame bits themselves. By definition, the utilized line rate will always be higher than the information rate. So for example, you would never provision 2 each 50 Mbit/s Committed Information Rate EVCs at a 100 Mbit/s physical interface because there is no capacity left over to handle the overhead that goes on top of the 50 Mbit/s information rates. Figure 9 illustrates for us what goes on in the Ethernet line when traffic is generated at a 50 Mbit/s Information Rate at frame sizes varying from the minimum VLAN size of 68 bytes to the maximum standard size of 1522 bytes. From Figure 9 we see that a Carrier Ethernet EVC running at the full 50 Mbit/s CIR loads up the physical interface with about a 52 to 65 Mbit/s utilized line rate depending on the frame size. The physical interface would need a line rate greater than 100 Mbit/s to transmit two EVCs running at 50 Mbit/s information rate. Because the physical interface can only run at 100 Mbit/s, frames would be dropped and the customer would be unhappy. Also, Figure 9 shows that the payload of the Layer 4 UDP flow will necessarily have a smaller data rate than the EVC itself. The exact differences in rates are a function of the Ethernet frame size and size of Ethernet, IP, and UDP headers.

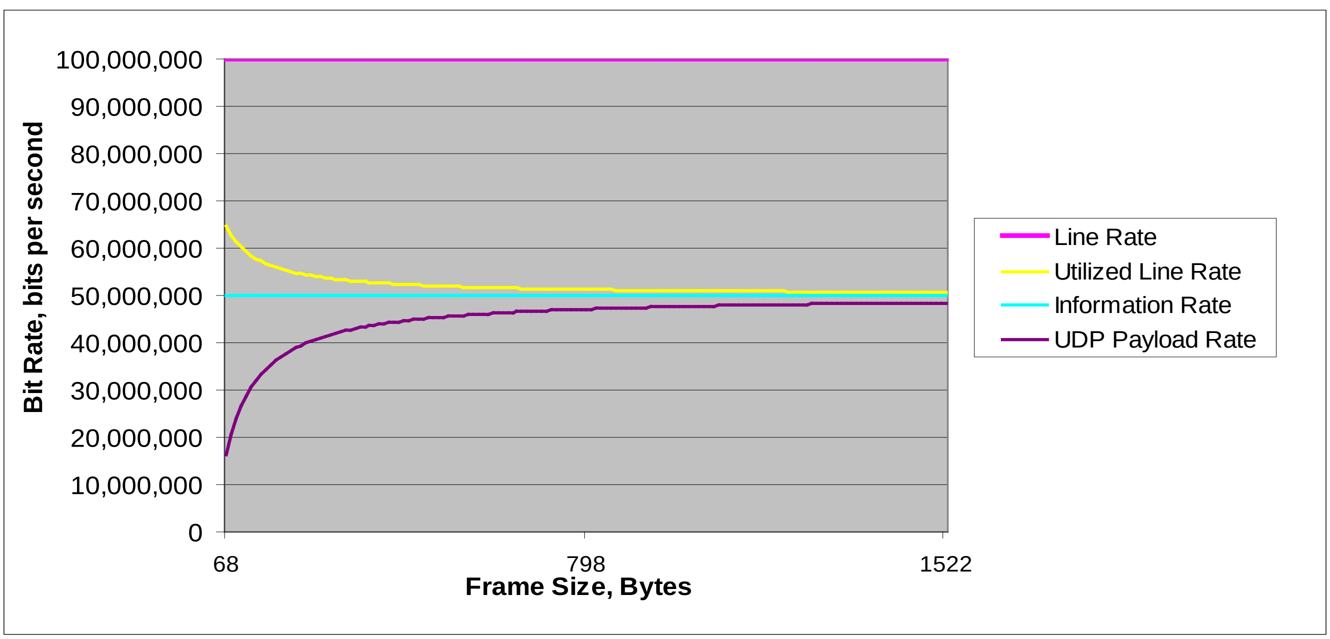

Figure 9. Impact of Frame Size on Line Rate and UDP Payload Rate at Fixed Information Rate

Standard TCP Window Size and Effect on Throughput

In TCP, each TCP segment (name of a data packet at Layer 4) is accounted for. This means that for each block of information sent across a data path, an acknowledgement must be received before sending an additional block of data. As it would be very ineffective to send only one segment at a time and then wait for each acknowledgment, TCP has a built-in capability to send multiple segments into the network at the same time; this capability also serves as a flow control mechanism. Should a receiving host have trouble processing all of the received data; it will delay the acknowledgment to the sending host.

A graphical view of TCP flow control is shown in Figure 10. The graph shows the total amount of memory available to the issuance of sequence numbers. The frames that are transmitted are assigned a limited set of numbers corresponding to the cumulative number of bytes transmitted since the start of the session. When the total number of transmitted bytes exceeds 2^{32}, the numbering goes back to the starting number and repeats. It is important that all the frames that are currently in transmission are received with unique numbers so that they can be reassembled in the right order. In this diagram, the blue arc shows these active frames and the range of numbers that have been allocated to them. As long as the blue arc is much smaller than the circumference of the circle, there is no problem. To assist with transmission rate optimization, the receiver advertises (rwnd [receiver window] advertisement) how much window it has available for the transmitter to fill up – the transmitter is free to continue boosting its transmitted rate if the advertisement shows buffer space available and if it hasn't received a pause frame from the receiver.

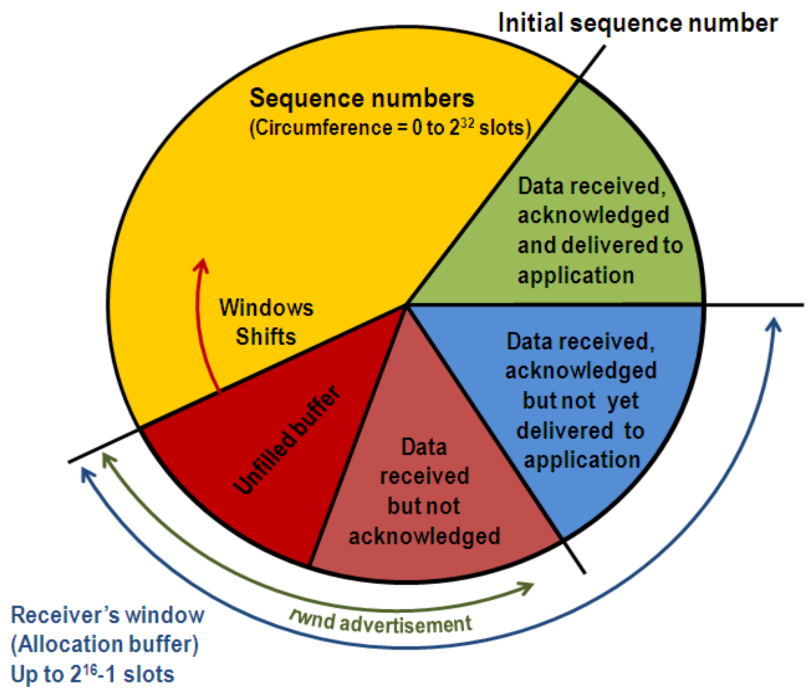

Figure 10. TCP sliding window flow control protocol (source: Wikipedia)

Another notion that must be addressed before moving forward is the effect of high bandwidth and/or high latency on the TCP protocol. As the bit rate increases, the amount of data required to fill the pipe until an acknowledgement is received grows linearly. It also grows linearly with latency (or Frame Delay). This behavior is defined as the bandwidth-delay product. Its formula is:

Pipe Capacity (bits) = bandwidth delay product = bandwidth (bits/s) x round-trip time (s)

Bandwidth is the transmission rate of the link. Round-trip time is the amount of time for bits to travel across the network and back. Pipe capacity refers to the number of bits "in flight." In other words, capacity is the maximum number of bits in transit over the link at a given time. Since TCP transmission requires acknowledgement, a sender needs to buffer at least enough bits to continue sending until an acknowledgement is received. This sender buffer for TCP is the same as the TCP receive window, and is a maximum of 64 kB for standard TCP. So for TCP, the bandwidth delay product can be rewritten as follows:

$$\text{Bandwidth} \leq \frac{\text{Capacity}}{\text{Round-Trip Time}}$$

Thus, the maximum bandwidth for a TCP circuit is limited by the size of the TCP receive window and as well as the round-trip time. The following table provides an example of the bandwidth-delay product for a link with 40 ms round-trip latency.

Circuit Rate	Payload rate (Mbit/s)	Capacity (Kbits)	Capacity (KBytes)
DS1 (1.5M)	1.536	61	7.5
E1 (2M)	1.984	79	9.68
64KB Windows TCP max	13.1	524	64
DS3 (45M)	44.21	1,768	215
100BASE-T	100	4,000	488
OC-3/STM-1 (155M)	150	5,990	731
OC-12/STM-4 (622M)	599	23,962	2,925
1000BASE-T	1,000	40,000	4,882
OC-48/STM-16 (2.5G)	2,396	95,846	11,700
OC-192/STM-64 (10G)	9,585	383,386	46,800
10GBASE-SW (WAN)	9,585	383,386	46,800
10GBASE-SR (LAN)	10,000	400,000	48,828

Table 1. Bandwidth-delay product for different circuit rates with 40ms round-trip time

The column of interest is Capacity (in KBytes). This theoretical value provides the maximum number of bytes in the system at any time so that the link is filled to the maximum and that TCP can resend any dropped or errored segments. In a standard TCP implementation, the maximum allowable TCP window is 65,535 bytes; this means that at a rate of 13.1 Mbit/s and more, with a round-trip time of 40 ms, a server running normal

TCP cannot fill the circuit at 100%. This is a theoretical maximum; unfortunately, the network might drop frames along the way, making a lower payload rate more likely.

Companies who buy higher speed Carrier Ethernet circuits often have hundreds of users or processes at a location that are sharing the circuit. In this case, the circuit will have many TCP/IP streams sharing a single circuit. The effective bandwidth that any one of the users is using will likely be not that high, and the overall Ethernet Virtual Connection bandwidth may be used efficiently without any TCP/IP extensions as discussed in the next paragraph.

Using Window Scaling to Increase TCP Throughput

Window scaling, RFC 1323, is a technique used to extend TCP's throughput. The 16-bit counter limitation for unacknowledged frames is expanded to 32 bits, which greatly expands the bandwidth delay product through which TCP can be transmitted by a factor of roughly 65,000. Although developed many years ago, window scaling is not easily available to most computer end users today. Techniques exist to manually modify operating systems like the Microsoft Windows system registry to invoke window scaling, but that is simply beyond the capability of most users. Nonetheless, users who need to get high bandwidth performance from a single TCP/IP stream on a long Ethernet circuit should get some help to investigate this option. Figure 11 shows the potential speed boosts available for a single TCP session. The 13 Mbit/s limit of standard TCP expands to well over 10 Gbit/s with TCP Window Scaling.

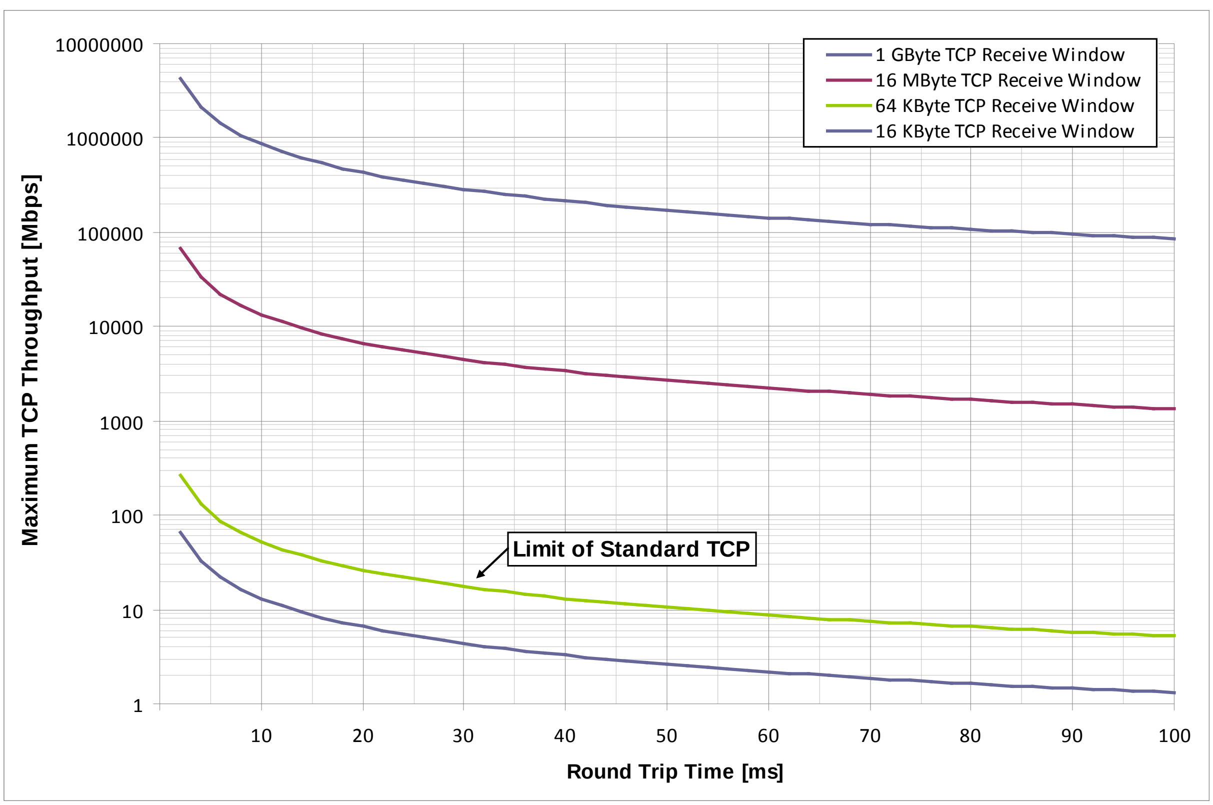

Figure 11. TCP Window Scaling Greatly Boosts Single Session Throughput

Dropped/Errored Frames and the Retransmission Effect

All transport circuits have some underlying error rate. When the bit error rate is very low on the order of 10^{-6} or much less depending on the application, users will generally not see much service degradation. Some protocols such as TCP require a retransmission any time an error occurs anywhere in the frame, or if the frame is dropped during transmission. The frame size has a magnifying effect which multiplies the impact of a bit error rate. For instance if a frame is about 120 bytes long, that amounts to about 1,000 bits. A bit error rate of 10^{-12} is magnified into a frame error rate of 10^{-9}. The effects of frame loss (whether by bit error or dropped traffic) on overall network throughput can be modeled as follows (Mathis, et. al.).

$$\text{Throughput} \leq \frac{\text{Maximum Segment Size}}{\text{Round-Trip Time} * \sqrt{\text{Probability of Frame Error}}}$$

From this formula, we can see that we can increase the throughput by reducing the round trip time, reducing the frame error rate, and/or increasing the segment size. Maximum Segment Size is the size of user data. The effects of frame loss are plotted as a function of round-trip time in Figure 12 below. Note that for very high frame error rates like 10^{-2}, throughput can be limited to less than 2 Mbit/s by frame error rate alone (for delays over 60ms). Note also that the range of round trip times correspond to Carrier Ethernet services running over longer distances (with round trip times 2 ms and larger). For delays much shorter than 2ms, the store and forward time of large packets can come into play. Recall also that traffic offered to the network above the CIR is carried only on a best efforts basis. Excess traffic may be dropped all or in part. In the case of serious defect in the implementation of service where offered traffic greatly exceeds the CIR, half or more of the frames may be dropped (eg policed) and protocols like TCP may completely break down and carry virtually no data at all.

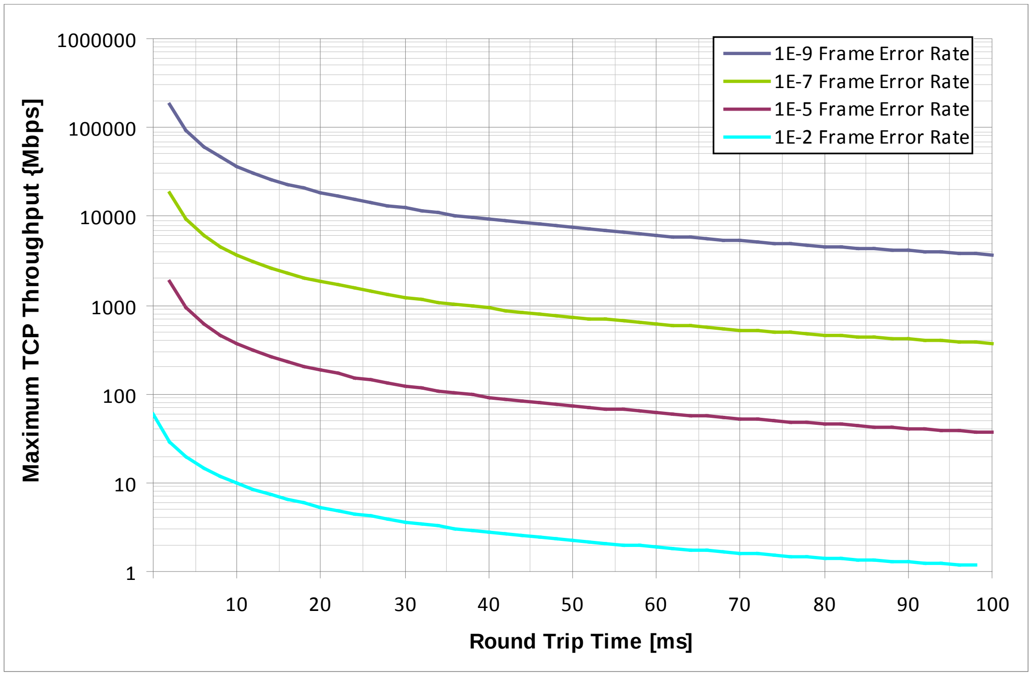

Figure 12. Frame Error/Loss Ratio Limits Maximum Throughput

Monitoring Throughput

General Network Management Tools

A wealth of literature and tools guide end users and network operators in the successful management of their Ethernet networks. One good tool is flow analysis applied in the service provider network or in the end user network to identify top talkers and see why they are utilizing so much bandwidth. It could be that malware has entered the computer of an end user, generating excessive spam email traffic which saps private and public networks alike. Or, peer to peer file exchange may be occurring in violation of copyright laws at the same time as absorbing great network capacity. These sorts of problems can make users think something is wrong with their Carrier Ethernet service, when in reality, the service is working fine and provides plenty of bandwidth for proper usages. Service Providers can gain access to special flow analysis software and systems available in their router/switch management systems that provide excellent insight into the exact real-time sources of loads on the network. End users and service providers alike are advised to consult the rich literature on these general network management subjects. Sample graphics from these systems are shown in Figure 13 and Figure 14.

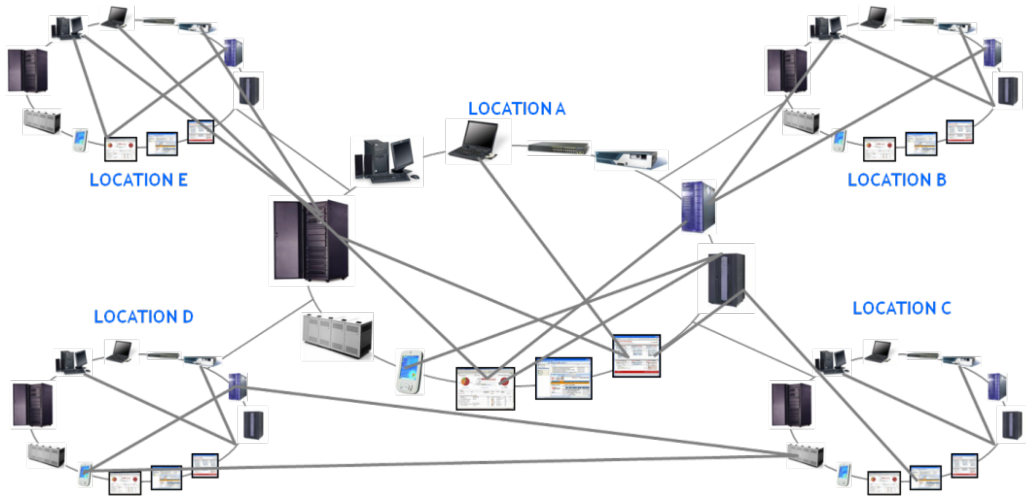

Figure 13. Flow Analysis Diagramed Through the Network

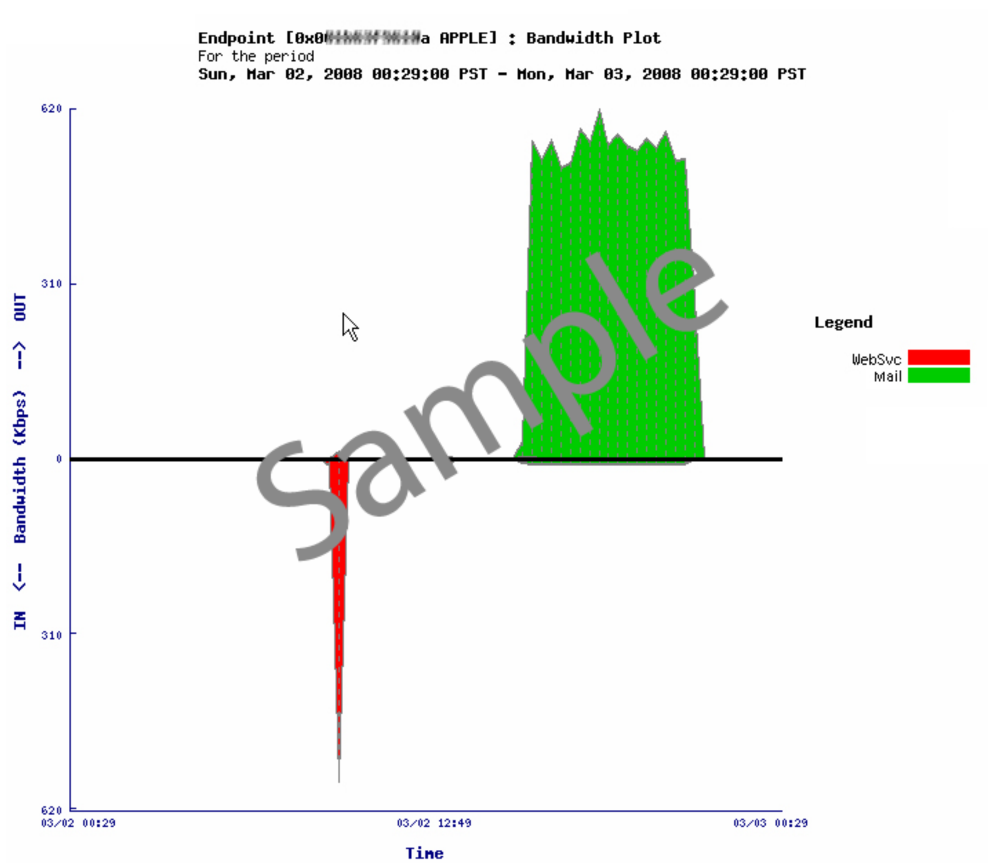

Figure 14. Single User Spam Malware Load on Network

Verifying Carrier Ethernet Performance

Troubleshooting with PC Based Test Tools

So now let's turn to the subject of how an end user can independently verify the performance of and troubleshoot the Carrier Ethernet service. For our new Carrier Ethernet User from Figure 1 and Figure 2, how relevant was a File Transfer Protocol (FTP) download rate when trying to validate the performance of an Ethernet service? We have just seen issues that could have caused that problem, but there are more. One must first understand the FTP environment. Like any other application, an FTP session relies on the underlying hardware, software and communication protocols. The performance of a PC is very much aligned with its hardware and the characteristics of the CPU (speed, its cache memory, RAM). The operating system and the different background programs loaded are additional factors. Firewalls, anti-virus and spy-ware can further limit the performance of a PC. From an operating system perspective, this is where the OSI stack resides. The network performance of a PC is directly related to the OS it is using. By default, the TcpWindowSize registry key value is set to 65,535 bytes, which affects the TCP performance in high-bandwidth networks. Although there are utilities such as window scaling to increase this value, some applications, like FTP, may possibly override the TcpWindowSize registry and use the 65,535 value, thereby reducing performance.

Software Based Test Tools

Freeware software tools and online bandwidth test sites receive a lot of publicity from different sources as they can help test and benchmark networks. These tools use the same PC architecture as an FTP download test. Although the TcpWindowSize registry could be bypassed with these tools, their performance is also directly related to the PC performance. A PC that does not have enough RAM memory or has too many background programs loaded will perform differently than another PC that is more recent and has more memory. Although the measurement can provide some insight on the problems on a network, the measurement will not be as repeatable and reliable as others will with dedicated hardware.

Again, the bandwidth-delay product will influence performance. If one doesn't have the capability to extend the TCP window size, the only way to prove that a link can support 100% load of TCP traffic is to start multiple test sessions. Having multiple TCP streams will fill the link under test, but multiple TCP streams will be "fighting" for the bandwidth and may degrade the PC performance they are running on. The peak rate of all test streams might come close to the configured throughput of the link, but looking at the average may show that it is way off.

Finally, when using these software-based tools, bear in mind that they are operating generally at layer 4 and higher, and so they will generally not provide a direct measurement of the conformance of the Layer 2 Carrier Ethernet service to the Service Level Agreement.

Hardware Based Test Tools

Hardware based test equipment for testing Ethernet services is also available and provides definitive confirmation of whether the Carrier Ethernet service is performing properly. This equipment may be portable/hand-held or integrated into other CPE or network elements. These basic Ethernet test instruments have the capability to format test traffic up to wire speed (the maximum possible line rate) for the service being tested – even for GbE and 10 GbE services. In addition, they look at traffic at layer 2, 3 and even 4 or higher in some cases. The test sets have a dedicated OSI stack which ensures that higher level protocol layers or applications can utilize all the measured bandwidth. With this equipment end users or service provider field technicians can reliably verify that they are getting the committed information rate, frame delay, and dropped frames per the Service Level Agreement (SLA) on the layer 2 Ethernet service. Technicians can make long term tests to see if the network has certain times of the day where it underperforms. The test set's layer-2 round-trip-time measurement is the value of circuit transmission delay used for calculating the bandwidth delay product. If a deeper analysis of circuit performance is needed, the test set can be used to invoke portions of the RFC 2544 or MEF test suites relevant to the service under test. There are many whitepapers and articles available online that describe these test suites in much more detail. Dedicated Ethernet test sets are quite affordable for enterprise users who really want to understand the performance of their Ethernet service. Figure 15 shows how you can plug in these test sets to make the needed Carrier Ethernet measurements.

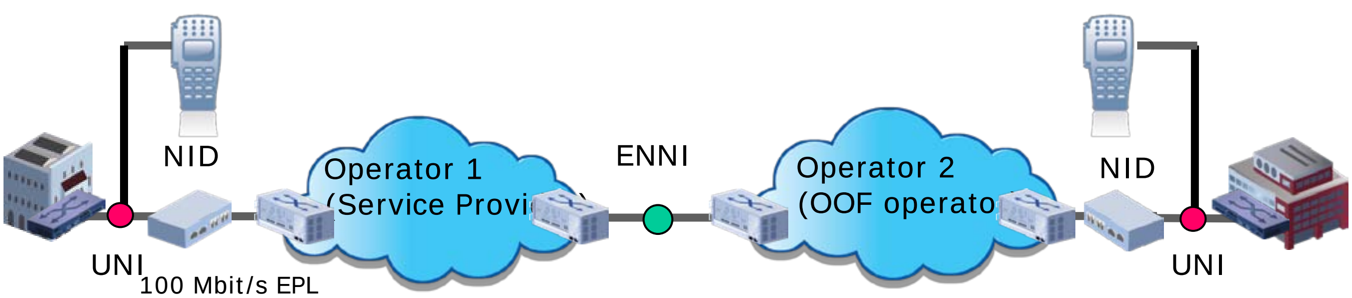

Figure 15. Testing an EVC

Summary and Recommendations

Carrier Ethernet customers can optimize their service performance as follows. First, they should make sure their Ethernet carrier is MEF certified to ensure delivery of a high quality Ethernet service. Then, they should make sure that they are using routers that properly limit or shape the traffic they send to the network so that problems with discarded overflow traffic will be avoided. The customer should ensure that their router does not burst a greater number of frames than can be safely captured according to the service's committed information rate and committed burst size. If they have any question about the performance of their layer 2 Ethernet service, they or their service provider should do a simple point-to-point test with Ethernet test sets to verify that the circuit they have purchased is living up to its Service Level Agreement. Users should be aware that the throughput indicator on a PC file transfer or online bandwidth test site is likely showing the limitation of the TCP/IP or FTP protocol rather than the high-speed layer 2 Carrier Ethernet service.

End users may also want to measure the Carrier Ethernet circuit's round trip delay (or have the service provider measure it), so that they can calculate the upper limit bandwidth achievable per standard TCP/IP stream from Figure11. If the delay supports the necessary throughput, no further change is required. If not, end users may want to explore getting help to utilize RFC 1323 window scaling or other technique to get more performance out of their TCP/IP and FTP protocols on their high speed circuits, especially those with higher latency. If end users chose to use UDP to improve throughput, they should bear in mind that with UDP they will get some errors on the received data, and they will either need to accept those errors or use a higher layer in the OSI stack to ensure received data integrity. End users can further tune their Carrier Ethernet Circuit by using large frames to get the most efficient use of bandwidth, or by using short frames to get the lowest possible latency.

Glossary Of Abbreviations

ARP	Address Resolution Protocol
ATM	Asynchronous Transfer Mode
BCP	Bridging Control Protocol
BPDU	Bridge Protocol Data Unit
BWA	Broadband Wireless Access
CBS	Committed Burst Size
CIR	Committed Information Rate
CFM	Connectivity Fault Management
CLEC	Competitive Local Exchange Carrier
CPE	Customer Premise Equipment
C-Tag	Customer Tag (VLAN Id)
EFM	Ethernet in the First Mile
EBS	Excess Burst Size
EIR	Excess Information Rate
E-LAN	Ethernet-LAN Service
E-Line	Ethernet Point-to-Point
ENNI	External Network to Network Interface
EPL	Ethernet Private Line
E-Tree	Ethernet Tree service (1 to many)

EVC	Ethernet Virtual Connection
EVPL	Ethernet Virtual Private Line
FTP	File Transfer Protocol
GbE	Giga Bit Ethernet
IEEE	Institute of Electrical & Electronics Engineers
IETF	Internet Engineering Task Force
ILEC	Incumbent Local Exchange Carrier
IP	Internet Protocol (Layer 3)
IPTV	Internet Protocol Television
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
LAN	Local Area Network
MAC	Media Access Control (layer 2 protocol)
MEF	New name for entity formerly known as Metro Ethernet Forum
MSO	Multiple Service Operator (Comcast, COX, Time Warner Cable, etc)

NID	Network Interface Device
OAM	Operations, Administration and Maintenance
OOF	Out of Franchise
OSI	Open Systems Interconnect
PC	Personal Computer
QoS	Quality of service
RFC	Request For Comment (an IETF tool for organizing/communicating comments)
SLA	Service Level Agreement
SLO	Service Level Objectives
SOF	Start Of Frame
S-Tag	Service Tag (VLAN Id)

T1	T-Carrier 1 (1.544Mb/s)
T3	T-Carrier 3 (44.736 Mb/s)
TCP	Transport Control Protocol (Layer 4)
TCP/IP	Transport Control Protocol/Internet Protocol
TDM	Time Division Multiplexing
TFTP	Trivial File Transfer Protocol
UDP	User Datagram Protocol
UNI	User to Network Interface
VoIP	Voice over Internet Protocol
VLAN	Virtual LAN

References

Matthew Mathis, Jeffrey Semke, Jamshid Mahdavi. *The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm*. Computer Communication Review, 27(3), July 1997. (http://www.psc.edu/networking/papers/model_ccr97.ps)

MEF Technical Specification, MEF 11. User Network Interface (UNI) Requirements and Framework. November 2004.

MEF Technical Specification, MEF 10.2, “Ethernet Services Attributes – Phase 2”, October 2009.

IETF RFC 2544, Benchmarking Methodology for Network Interconnect Devices, March 1999.

IETF RFC 1323, TCP Extensions for High Performance, May 1992

Acknowledgements

The MEF thanks the following member companies for their contribution to this document

Contributor	Company
Mike Bugenhagen	CenturyLink
Fred Ellefson	ADVA Optical Networking
Craig Fanti	Canoga Perkins
Phil Fine	Calix
Brooke Frischemeier	Cisco
Bruno Giguere	EXFO
Steve Holmgren	att
Roman Krzanowski	Verizon Business
Ayal Lior	Independent Consultant
Paul Marshall	Sunrise Telecom
Steve Olen	Omnitron Systems
Brian Rose	Cox Cable
Abel Tong	Omnitron Systems

More information and updates on
Carrier Ethernet Services
can be found at
www.metroethernetforum.org