Fundamentals of Ethernet
10 Megabit Ethernet to 10 Gigabit Ethernet

Ethernet History

The University of Hawaii's ALOHA network is considered to be the ancestor of all shared media networks. In 1968, Norman Abramson pioneered the precepts of Ethernet by developing this packet radio networking system that ran at 4800 bps and 9600 bps.

A few years later in 1973, Robert Metcalfe and David Boggs at Xerox Corporation in Palo Alto, CA applied the ALHOA network principles and created the world's first Local Area Network (LAN). Initially named ALTO ALOHA, the name was later changed to Ethernet. This first version of Ethernet ran at speeds up to 2.94 Mbps. One of the first customers of Ethernet was the White House, where it was used for word processing. Beyond this high-profile customer, though, this version of Ethernet was not successfully commercialized.

The first commercial release of Ethernet was by DEC, Intel, and Xerox (DIX) in 1980 as Ethernet, Version 1. It was commonly referred to as Ethernet DIX80. The second revision release, Ethernet, Version 2, was released in 1982. It was commonly referred to as Ethernet DIX82. Ethernet, Version 2 is the standard of Ethernet technology that is in use today.

In 1980, the IEEE formed Project 802 to provide a framework for the standardization of LAN technology. Novell released Novell Netware '86 in 1983, which used a proprietary frame format based on a preliminary specification of the IEEE 802.3 standard. This is the same Novell software that is used today to manage printers and servers.

In 1983, the IEEE approved the

IEEE 802.3 standard, which included IEEE 802.2 Logical Link Control (LLC). This made Novell Netware's proprietary format incompatible with the latest technology. In order to resolve this incompatibility, Sub-Network Access Protocol (SNAP) was created for the new IEEE 802.3 standard.

Once the overall packet specifications were finalized, the transmission medium needed to be agreed upon. In the late 1980s, SynOptics Communications developed a mechanism for transmitting 10 Mbps Ethernet signals over twisted-pair cables.

It was this combination of a low cost transmission medium with agreed packet technology specifications that led to the wide deployment of Ethernet. The Ethernet-over-twisted-pair specification (10 BASE-T) was approved by the IEEE in 1990 as the IEEE 802.3i standard. It quickly became the preferred Ethernet media type.

Section 1: OSI Model Overview

The International Standards Organization (ISO) designed the Open System Interconnect (OSI) model for data communications. This model, in some form, is followed by ALL data communications. Any time two or more computers transmit information, they follow the OSI model.

Layer 7	Application Layer
Layer 6	Presentation Layer
Layer 5	Session Layer
Layer 4	Transport Layer
Layer 3	Network Layer
Layer 2	Data Link Layer
Layer 1	Physical Layer

Table 1: OSI model

The OSI model (Table 1) is a series of basic building blocks. Each block has its own function and role in getting data from one point to another.

– **Physical Layer: Layer 1**
 The physical layer represents the "pipe". This is normally what major service providers offer to their customers. There are two parts to the physical layer – the physical media and the bandwidth. The physical media can be twisted pair copper, coax, or fiber. The bandwidth is a combination of signal characteristics and rates (T1 B8ZS, for example).

– **Data Link Layer: Layer 2**
 The data link layer represents the beginning of the data and offers the basic data framing. This layer can be provided by the carrier (frame relay, for example), or it can be provided by the customer in a point-to-point leased line environment (HDLC or PPP, for example).

– **Network Layer: Layer 3**
 The network layer represents the beginning of customer traffic. This is the layer where computers speak to each other and data is addressed for end-to-end communications. The addressing type for this layer is often a router or a computer. Examples of network layer protocols include IP and IPX.

– **Transport Layer: Layer 4**
 At this layer, the lower layers and the application meet. The transport layer identifies the application that rides within the data packet and makes sure that all of the packets get from the source to the destination. This layer also specifies the connection type, or transfer protocol, of the application.

 There are two types of transport protocols – a connection-oriented transport protocol (TCP) and a connectionless transport protocol (UDP). Connection-oriented applications require all packets to get from the source to the destination. An example of this transport protocol is e-mail. If all of the packets making up the e-mail don't get transmitted, the e-mail is unreadable. Connectionless applications are those that do not require all packets to get from the source to the destination. An example of this transport protocol is streaming audio. If a packet or two is missed, the computer will skip a few bars of music and keep playing.

– **Session, Presentation, and Application Layers: Layers 5, 6, and 7**
 In most data communication networks today, these layers merge together into the application layer. Aspects of Lotus Notes, POP3 mail, SMTP mail, and Web surfing all have layers 5, 6, and 7 built into them. For purposes of this document, these layers will be combined into layer 7 and will be referred to as the application layer.

OSI Model Analogy

In order to better understand the now five-layer model (physical layer, data link layer, network layer, transport layer, and application layer), a parallel model built around a familiar process – mail delivery – is offered.

OSI Model Importance

The most important aspect of the OSI model is apparent when dealing with a service issue or performing service turn-up. Each layer builds upon the next layer – the bottom layer being the physical layer. Looking at the mail model analogy (Table 2), a closed road or broken down mail truck means no mail delivery. No matter how perfect the envelope or how well it is addressed, if the truck can't carry the mail, it won't get to its destination. Similarly, if the copper that carries the customer data is bad or if the T1 is incorrectly optioned, the data at layers 2 and up will not properly get from the source to the destination.

Application Layer	The application is the letter itself. It is the actual piece of information sent from one location destined for another location to be read by a specific person.
Transport Layer	A final distinguishing factor is the name that appears on the envelope. Just as a household can have multiple residents, a computer can have multiple applications. The name on the envelope identifies who should read the letter.
Network Layer	The address on the envelope represents the network layer. In this case, the address is the street address, city, state, and zip code.
Data Link Layer	The data link, being the basic data format, would be represented by the envelope that the letter is put into. The envelope distinguishes one letter from another.
Physical Layer	The roads and trucks that carry mail are analogous to the physical layer. The roads represent the copper or fiber, while the truck represents the technology (T1, for example).

Table 2: OSI five-layer model vs. mail model analogy

OSI Layer	Technology
Application	Lotus Notes, WWW, SMTP, POP3
Transport	TCP, UDP, SPX
Network	IP, IPX
Data Link	HDLC, PPP, Frame Relay, ATM, MAC
Physical	QFSK (modem), T1, T3, SONET, DMT (DSL), 802.3 (Ethernet)

Table 3: Technologies associated with the OSI layers

When looking at problem circuits, some problems are easier to identify than others. If a farmer in Georgia cuts through a pair of copper cables, none of the data will transmit across the copper. Since the failure is total and usually very obvious, it is easy to identify and resolve. The more difficult problems to solve are on the marginal circuits. These types of problems tend to come and go and are difficult to identify. Many of these marginal problems will present themselves at higher layers, even though the problem is at a lower layer of the OSI model.

Examples of OSI Layer Technologies

Table 3 provides examples of technologies and the layers that they reside in within the OSI model. It is important to remember that almost any layer 3 technology can reside on any layer 2 technology, which in turn can reside on any layer 1 technology. The layers can be switched around depending on the network architecture (dial-up or DSL). Each layer is independent of the layers above and below it.

Using the OSI layer and technology information from Table 3, Table 4 shows three specific methods of how end users can access the Internet.

Notice that the end-to-end addressing scheme (layer 3 – IP) is constant no matter how the Internet is accessed.

OSI Layer	From Home - Modem	From Home - DSL	From Office
Application	HTTP (WWW)	HTTP (WWW)	HTTP (WWW)
Transport	TCP	TCP	TCP
Network	IP	IP	IP
Data Link	PPP	ATM	Frame Relay
Physical	QFSK	DMT	T1

Table 4: Three different end user methods

Section 2: Ethernet Technology Fundamentals

Physical Layer – 10 Mbps, 100 Mbps, and 1 Gbps Ethernet

The physical layer for Ethernet is defined by certain electrical and bit rate specifications. The electrical specifications are based on the IEEE 802.3 Ethernet standard.

Line Coding

10 Mbps, 100 Mbps, and 1 Gbps Ethernet technologies utilize a specific type of line coding referred to as 8B/10B. 8B/10B conversion is a simple algorithm that converts 8 bits into 10 bits, producing a 25% overhead. This effectively increases the line rate without increasing the actual throughput. There are two main reasons for the use of this particular type of line coding.

(1) DC Current Elimination: In order to keep the average DC current on the wire equal to zero, there must be an equal number of positive pulses and negative pulses. The algorithm for 8B/10B converts the 8-bit sequence, where a positive or negative DC voltage would present itself. Those 8 bits are converted to a 10-bit sequence, eliminating the DC current artifact.

(2) Clocking: There is no common clock within an Ethernet circuit. The bit stream from one device drives the clock on the receiving device. If there are two many zeros (absence of pulses) in the bit stream, the receiving device will lose synchronization. The 8B/10B conversion takes the 8 bits without enough clocking pulses and converts it to 10 bits without enough positive pulses.

In terms of testing, there are no options or settings for line coding. All test sets for Ethernet support this type of line coding.

Duplex Options

Another important physical layer characteristic of Ethernet is its full-duplex or half-duplex operation. A full-duplex circuit is able to transmit and receive at the same time, similar to a telephone where a person can speak and hear at the same time. A half-duplex circuit is either speaking or listening; it is incapable of both operations simultaneously. Since it can only operate in one direction, a half-duplex circuit only offers the user about half of its actual bandwidth. In other words, a 100 Mbps circuit running in a half-duplex environment offers the user only about 50 Mbps worth of actual data throughput.

10 Megabit Ethernet (10 BASE-X) is most commonly deployed in a half-duplex environment. Users requiring lower bandwidth applications are able to utilize 10 BASE-X and save the cost of higher bandwidth equipment.

100 Megabit Ethernet (100 BASE-X) can be deployed in either a full-duplex or half-duplex environment, depending on bandwidth requirements. Most switches deployed in a LAN today are capable of 100 BASE-X.

Gigabit Ethernet (1000 BASE-X) is almost always deployed in a full-duplex environment, allowing

network routers and switches to take full advantage of the bandwidth. The IEEE 802.3 Ethernet standard, however, does allow for 1000 BASE-X deployment in a half-duplex environment.

Physical Layer – 10 Gbps Ethernet

LAN Versus WAN

There are two types of 10 Gigabit Ethernet (10 GigE) – WAN and LAN. The LAN specification (10G BASE-R) is very much like traditional Ethernet. It has no common clock, and it is very similar in the physical layer. The WAN specification (10G BASE-W) is based on the SONET/SDH signal architecture at the physical layer.

Line Coding

10 Gigabit Ethernet LAN

10 GigE LAN has a line coding of 64B/66B instead of the 8B/10B coding used for lower speed Ethernet services. The main reason for the switch from 8B/10B is the bit rate.

At 10 Gigabit, an 8B/10B formatted signal has a line rate of 12.5 Gbps. The 64B/66B signal, with significantly less overhead (3% instead of 25%), has an effective line rate of 10.3 Gbps.

10 Gigabit Ethernet WAN

The 10 GigE WAN specification utilizes SONET physical layer standards. Most often, an NRZ optical pulse is used. There is also a common clock; there-fore, no 64B/66B conversion is required.

The most critical physical layer com-ponent of a 10 GigE WAN signal is the SONET framing structure. In a 10 GigE WAN signal, the Ethernet packet is placed into a SONET frame. This frame also includes the section, line, and path overhead. In this type of structure, there are several over-head bytes that are slightly different than the traditional specification (Table 5).

	Byte	Function	Usage
Section	A1	Frame Alignment	Supported
	A2	Frame Alignment	Supported
	B1	Section Error Monitoring	Supported
	D1-D3	DCC	Not Supported
	E1	Orderwire	Not Supported
	F1	Section User Channel	Not Supported
	J0	Section Trace	Specific Value
	Z0	Growth	Not Supported
Line	B2	Line Error Monitoring	Supported
	D4-D12	DCC	Not Supported
	E2	Orderwire	Not Supported
	H1/H2	Pointer	Specific Value
	H3	Pointer Action	Specific Value
	K1/K2	APS	Specific Value
	M0	REI-L	Not Supported
	M1	STS-N REI-L	Supported
	S1	Synchronization	Not Supported
	Z1/Z2	Growth	Not Supported
Path	B3	Path Error Monitoring	Supported
	C2	STS Path Signal Label	Specific Value
	F2	Path User Channel	Not Supported
	G1	Path Status	Supported
	H4	Multiframe Indicator	Not Supported
	J1	STS Path Trace	Specific Value
	N1	TCM	Not Supported
	Z3/Z4	Growth	Not Supported

Table 5: 10 Gbps Ethernet WAN physical layer overhead

Duplex Options

10 GigE Ethernet (10G BASE-X) must be deployed as a full-duplex service. There are no specifications in the standard allowing for a half-duplex option.

WAN Interface Sublayer (WIS)

One of the main differences between 10 GigE WAN and 10 GigE LAN is the SONET portion. The WIS is responsible for mapping Ethernet frames into a SONET frame and demapping Ethernet frames from a SONET frame.

Generally, the WIS is designed to:

– Map Ethernet MAC frames into a SONET frame.

– Implement framing, scrambling, and defect/anomaly (error/alarm) detection to allow for compatibility with the SONET requirements.

– Provide a 9.95328 Gbps effective data rate at the service interface, conforming to the requirements of a SONET STS-192c signal.

Data Link Layer

The data link layer for Ethernet is the same for 10 BASE-X, 100 BASE-X, 1000 BASE-X (GigE), or 10G BASE-X (10 GigE). This layer is referred to as the Media Access Control (MAC) layer. It is the beginning of the basic data format for Ethernet.

Table 6 shows a basic Ethernet frame. There are four main parts to the frame – destination address, source ad-dress, control information, and the FCS.

Destination Address	Source Address	Frame Type	Information (Data)	FCS

Table 6: Basic Ethernet frame

– Destination and Source Address Fields: The destination and source address fields, as their names suggest, are the fields in the data frame that identify the destination and source MAC addresses for the frame. The source address is the device that transmitted the frame, and the destination address is the device destined to receive the frame.

– Frame Type: This field contains information that determines the format of the frame, either an Ethertype field for Ethernet, Version 2 or a length field for IEEE 802.3.

– Data Field: This field contains the bulk of the frame. This is where the upper layer information is encapsulated.

– FCS Field: This is the frame check sequence. The FCS is a calculation performed by the equipment generating the frame on the total bits in the frame. If any of the bits change while the packet traverses the network, the FCS value will no longer be valid. The device receiving the frame at the far end will see that the frame has been corrupted during transmission, and it will discard the frame.

All of the previous information applies to 10 BASE-X, 100 BASE-X, and 1000 BASE-X (GigE) services. There is, however, one specification of the Ethernet standard that is more applicable for GigE services than for 10 BASE-X and 100 BASE-X services – Pause Control.

Pause control frames allow Ethernet elements to throttle the actual throughput of the link in real time. Most elements can support full 10 BASE-X and 100 BASE-X rates. When GigE was first released, many elements could not support long durations of full bandwidth routing. Because of this, the pause control specification allowed a local element to tell the far end element to slow down until the local element caught up. Although not as prevalent as a few years ago, this is still part of the Ethernet standard and can be seen in deployed networks.

If the physical layer is bad, all of the information above it will be corrupted. For an Ethernet deployment, there are several danger zones that can cause the physical layer to be bad. Dirty fiber connections or bad media converters (electrical-to-optical or short range optical to long range optical) are two examples of physical layer problems that will generate customer traffic errors. If a customer or a carrier element is registering bad FCS frames, the cause is often a bad physical layer.

Table 7 shows the OSI model for Ethernet that has been discussed thus far.

Layers 5/6/7	Application
Layer 4	Transport
Layer 3	Network
Layer 2	**MAC**
Layer 1	**802.3**

Table 7: OSI model showing 802.3 and MAC layers

Network Layer

The network layer resides within the information field of the data link layer. This layer contains individual computer addresses or Web site addresses. Commonly used network layer protocols include IP (most common) and IPX (Novell). This document will focus on IP for this section, since IP is the technology that almost all carriers are moving forward with in order to provide next-generation services. Almost all potential customers are standardized on IP-based networks as well.

The overall role of IP is the routing of the packet from the source to the destination. It is not responsible for quality of service (QoS). It does not keep track of numbers of packets or lost packets throughout the network. These functions are the responsibility of higher layers of the OSI model.

As with the MAC layer, the IP layer contains a source address, a destination address, and an FCS. Table 8 shows the IP portion of an Ethernet frame. It is more complicated than a MAC frame.

Version	IHL	TOS	Total Length	
Identifier			Flags	Fragment Offset
TTL		Protocol	Header Checksum (FCS)	
Source Address				
Destination Address				
Information (Data)				
Options and Padding				

Table 8: IP portion of an Ethernet frame

– Like the MAC frame, the IP frame includes a destination address, a source address, and an FCS. There is a difference, however. The destination and source addresses are the final end point addresses and are not the next addressable ports. See the insert to the right for more information regarding what an IP address looks like and how it relates to the Internet.

– The total length field identifies the overall length of the information field. The overall length of the information field can range from 46 bytes to 1500 bytes.

It is important to note that the information field can have a wide range in length. This allows for various sized packets to be put into one IP frame. For example, an Internet URL request is a short connection request. The response is often a large Web page. The requesting packet would be small, but the response packets would be larger to accommodate the larger amounts of data in the Web page.

– The information field represents the data placed into the IP packet. This includes all of the upper layer information at the transport and application layers.

– Finally, the FCS completes the frame. The FCS is a layer 3 frame check sequence. The IP FCS allows a technician to differentiate layer 2 versus layer 3 FCS issues.

Another layer can now be added to the OSI model for Ethernet carrier-based deployments. The physical layer (layer 1) and the data link layer (layer 2) have been defined. Layer 3 can also be referred to as the IP layer (Table 9).

Now that the IP layer has been added, there is another layer for the occurrence of events. It is obvious that the main goal of IP is transmitting packets from a beginning point (source) to the end point (destination). This is apparent from the basic IP frame format.

If a user incorrectly addresses a packet, the packet will not arrive at the proper destination (like incorrectly addressing an e-mail). If the addressing scheme is flawed (the DNS server is not operating properly, for example), users will not be able to transmit their data to the destination. As discussed earlier, any errors at lower layers will corrupt the traffic in the layers above it. Therefore, it is critical that the physical and MAC layers are clean for IP to run properly.

Transport Layer

The final layer of the OSI model prior to the actual desired data is the transport layer. There are two main protocols that reside over IP and are common transport protocols in an IP network. These protocols are Transmission Control Protocol (TCP) and User Datagram Protocol (UDP).

This section is divided into two separate sections – TCP and UDP. These two protocols are very different in their roles and responsibilities, and it is important to distinguish between the two protocols.

Layers 5/6/7	Application
Layer 4	Transport
Layer 3	**IP**
Layer 2	**MAC**
Layer 1	**802.3**

Table 9: OSI model showing the IP layer

TCP – A Connection-Oriented Transport Protocol

TCP has six main responsibilities. They include:

– Basic data transfer

– Reliability

– Flow control

– Multiplexing

– Connection management

– Security

IP and the Internet

IP addresses have four different value locations, each ranging from 0 to 255. An IP address may look like 212.43.52.123.

The mechanism that allows a user to get from his local computer to an Internet site is through the IP addressing scheme built into the Internet. However, you don't enter an IP address into the URL field of your browser. You enter a Web address, such as JDSU's www.jdsu.com.

When you enter a Web address into the URL field, your computer sends that URL to a Domain Name Server (DNS). This server converts the Web address to its IP address. JDSU's IP address on the Internet is 157.234.223.80.

In fact, you will get to the same location on the Internet if you enter the IP address above in the URL field instead of entering www.jdsu.com.

– Basic data transfer and reliability: TCP, being a connection-oriented transport protocol, makes sure that all data gets from the beginning to the end of the network. Sequence numbers are built into the TCP overhead information of the TCP frame. These sequence numbers keep track of all of the information sent and the order in which it arrives. If any packets do not arrive, the TCP layer knows which packets didn't arrive and requests those lost packets. It is this layer that allows for small blips in local area networks and wide area networks to go unnoticed by the user.

– Flow Control: In addition to the sequence numbers discussed previously, there is another portion of the header that contains a value known as a "window size". A window size is effectively the amount of data each end point will receive prior to acknowledging its receipt of the data. A larger window size is more efficient than a smaller window size. The problem with a large window size occurs when there are lost packets. Because a larger window size indicates more data between acknowledgements, more data will need to be retransmitted for each error if there is an error in that time period.

– Multiplexing: Users are used to running multiple applications on their PCs simultaneously. Oftentimes they are checking e-mail and accessing one or more Web sites at the same time. TCP not only connects them to the other end point (Web site or e-mail), but it also manages which packets entering their computer are from the Internet or e-mail and makes sure that outbound packets are properly identified by the far end.

– Connection Management and Security: When two end points begin a conversation, the requesting end point requests a connection to the receiving end point. The receiving end point manages the connection and, if implemented, will attempt to confirm that the requesting end has the right to access the information.

Now that TCP's responsibilities have been identified, the frame and its parts can be discussed. Table 10 shows a typical TCP frame.

Source Port					Destination Port			
Sequence Number								
Acknowledgement Number								
Offset	Reserved	URG	ACK	PSH	RST	SYN	FIN	Window
Checksum					Urgent Pointer			
Options and Padding								
Information (Data)								

Table 10: Typical TCP frame

Compared to a less featured protocol, such as MAC, the TCP frame is extremely complicated and has many different fields that are responsible for the various tasks described above. It is outside the scope of this document to define each portion of the TCP header. However, a few of the fields are worth discussing.

– Source and Destination Ports: These are the address fields that identify the application type.

– Sequence and Acknowledgement Numbers: These are the fields that keep track of the packet sequences and which packets have and have not arrived from the far end sender.

– Checksum: This field, like other checksum fields, represents the basic FCS for the frame.

UDP – A Connectionless Transport Protocol
UDP is a simpler protocol than TCP. UDP is designed with the following features:

– Basic data transfer

– Connection management

Based on this reduced feature set, the make up of the UDP frame provides a good understanding of how it works and why it has a reduced feature set (Table 11).

Source Port	Destination Port
Length	Checksum
Information (Data)	

Table 11: Typical UDP frame

The main use for UDP is for those applications that do not require the arrival of all of the data in order to work. More importantly, UDP is for those applications that cannot utilize the information unless it arrives in the sequence in which it was sent. Examples of UDP applications include VoIP and streaming video.

When a user logs into a streaming Web site, such as www.cnn.com, to watch the latest news report, he uses UDP. CNN has many viewers with different computers, connection speeds, and link qualities. CNN can't stop transmitting video and retransmit the packets that one user did not receive. CNN doesn't want to know anything about what their users are getting in real time. UDP allows users to connect to CNN without all of the management and security that would cause the video service to fail.

Layers 5/6/7	Application
Layer 4	TCP/UDP
Layer 3	IP
Layer 2	MAC
Layer 1	802.3

Table 12: OSI model showing the TCP/UDP layer

Transport Layer Summary
With a solid understanding of the transport layer, the OSI model can be re-defined as a complete set of building blocks shown in Table 12.

VLAN Tagging
A Virtual Local Area Network (VLAN) tag is made up of two parts, a tag value (0 to 4095) and a priority value (0 to 7). A VLAN tag is a way to separate traffic on a LAN into different sub-groups. There are two main reasons for adding VLAN tags to a network – traffic routing and traffic prioritization.

Layers 5/6/7	Application
Layer 4	TCP/UDP
Layer 3	IP
Layer 2.5	VLAN
Layer 2	MAC
Layer 1	802.3

Table 13: OSI model showing VLAN tagging

– Traffic Routing: VLAN tags allow for quicker routing on higher-speed networks. With VLAN tags enabled, the element performing the routing does not have to look to the IP layer to find the ultimate destination, saving processing power and reducing the time to route each packet. Any testing traffic placed on a network routed with VLANs must have a valid VLAN tag.

– Traffic Prioritization: If a network has both voice and data traffic on it, the voice traffic should have priority over the data traffic. Most networks rely on VLAN tags to differentiate the voice from the data traffic. If the network gets busy, the routers will drop the data traffic and route the voice traffic. Although there is a priority value that can be set with each VLAN tag, most networks utilize different tag values for different levels of prioritization.

A VLAN tag effectively sits between the MAC layer and the IP layer. It is sometimes referred to as layer 2.5 (Table 13).

Section 3:10 BASE-X, 100 BASE-X, 1G BASE-X, and 10G BASE-X Deployment

Currently, most carriers are offering three versions of the Ethernet services. They are referred to as transparent LAN, switched Ethernet, and routed Ethernet.

Deployment Options

The three deployment options are based on the OSI layer that is required for the traffic to transverse the network.

Transparent LAN (OSI Layer 1 Service)

The term "transparent" is used to indicate that the LAN sites on either end of the service are connected by a pipe and have no idea if the Ethernet traffic is traveling 5 feet or 500 miles. Transparent service only requires that the first layer of the OSI model (the physical layer) is properly formatted in order to transmit traffic. If the electrical or optical characteristics of the signal are correct, the service will transmit the data. This service is ONLY point-to-point.

Typically, this type of service is offered via a Dense Wave Division Multiplexing (DWDM) system. Instead of a SONET card as the customer interface, the card is an Ethernet card (typically gigabit). The signal that the card receives is transmitted over the long haul and delivered to the far end. There, it is converted back to the appropriate wavelength for the Ethernet service.

Another way for transparent LANs to work is through media converters. A media converter takes a signal, electrical or optical, and converts it into a long haul optical signal. This allows 10 BASE-X or 100 BASE-X services to travel for miles. Gigabit Ethernet, although already optical, can also benefit from media converters. Normally gigabit Ethernet is available at 850 nm or 1310 nm, both considered short haul wavelengths. A media converter can receive the 850 nm or 1310 nm signal and convert it to 1550 nm.

A final option for transparent services is using a standard Add/Drop Multiplexer (ADM). Like the DWDM-based option, the Ethernet signal is placed directly into the transport system. Unlike the DWDM option, the signal is limited by the SONET signal structure. If the SONET pipe available to the Ethernet is only an STS-12 (622 Mbps), the Ethernet is limited to the bandwidth of an OC-12 service. For 10 BASE-X or 100 BASE-X, this is more than enough bandwidth. For gigabit Ethernet (1.25 Gbps), however, an STS-12 mapped Ethernet circuit only offers about 60% of the bandwidth required. The customer only has access to 60% of the total possible bandwidth. To overcome this, some providers are offering gigabit Ethernet encapsulated in an STS-24c (1.25 Gbps) or an STS-48c (2.5 Gbps).

Switched Ethernet (OSI Layer 2 Service)

This type of offering has more flexibility, but it is slightly more complicated to turn-up and troubleshoot. In order for this service to operate, the customer must provide some type of addressing. In OSI terms, this requires the customer to correctly address at layer 2 (MAC) or layer 2.5 (VLAN). By using addressing schemes, the carrier can sell the service as a point-to-multipoint service instead of just a point-to-point service and can offer prioritization of the traffic.

A layer 2 (MAC) service is typically offered using a Multi-Service Platform (MSP) with a switch card. The switch card looks at an Ethernet packet and switches it based on its destination MAC address. An incorrect destination MAC address causes the switch card to ignore the packet.

A layer 2.5 (VLAN) service is also available with carriers today. VLAN tagging allows the user to easily set up a point-to-multipoint network using a very simple addressing scheme. The VLAN tags are very easy for the carrier to read, allowing for quick and efficient routing and prioritization.

Routed Ethernet (OSI Layer 3 Service)

A routed Ethernet service is a layer 3 (IP) service. In order for a fully IP-addressable service to be offered, a large IP network is required at the core. This is typically a Packet Over SONET (POS) network. Turning up a routed service requires a test set capable of generating traffic with an IP address.

Another type of routed service is generically referred to as managed Ethernet. When the term managed is added to the service, the carrier owns that portion of that service. An example of a managed Ethernet service is a VPN. Many enterprise networks currently utilize VPNs, but the VPN server is owned and operated by the enterprise customer.

Network Architecture

For carriers, network architecture consists of the different types of equipment and architecture that is used to deploy Ethernet. Currently, both point-to-point services as well as point-to-multipoint services are available.

WDM Deployment

This deployment is based on a transparent service using WDM elements only. Effectively, the carrier takes the customer LAN signal and converts the signal into a WDM wavelength. This method of deployment works for any Ethernet rate.

Switched or Routed Deployment

A switched or routed deployment offers carriers more flexibility and growth than a WDM point-to-point Ethernet deployment. The service uses addressing schemes either at layer 2, layer 2.5, or layer 3.

Because these services require addressing to route the packets, any turn-up of the service will require addressing on the test set. If either the technician or the customer incorrectly addresses packets, the pipe will not transmit any of the traffic, making the circuit appear down at the physical layer. The problem, however, is at higher layers.

Section 4: Turn-up and Troubleshooting of Ethernet Networks

For the purposes of this section, the discussion will focus on the turn-up and troubleshooting of a basic Ethernet service. The screen shots shown throughout this section are from an Acterna FST-2802 TestPad Ethernet services test instrument. Other Acterna products, such as the Acterna DA-3400, T-BERD 8000, or MTS-8000, offer similar test feature sets and can be used instead of the FST-2802.

Overview of Turn-up and Troubleshooting

As with any service, turn-up and troubleshooting of Ethernet networks is critical to confirm that the service works prior to the hand-off to the end user. To confirm Ethernet services, the technician generates traffic and measures the traffic for various parameters. This section covers the types of traffic that need to be generated as well as the measurements that need to be performed.

Turn-up Testing

Traffic Generation

An Ethernet service is a pipe offered to the customer to transmit traffic from one point to another. In order to confirm that the pipe is clean and will transmit the customer's traffic, the technician must generate traffic and confirm that all of the traffic traverses the network without being corrupted.

When setting up a test set to generate traffic, there are three main parameters that must be specified: utilization, frame size, and traffic profile.

– Utilization: This is the most critical parameter. Depending on the service, the Ethernet pipe may transmit at a line rate of 1 Gbps or less. The carrier and type of network determines the maximum throughput. Therefore, generating traffic at the maximum line rate and confirming that the traffic is not corrupted is critical.

– Frame Size: Different frame sizes can affect Ethernet elements. Smaller frames cause elements to work harder than larger frames. The reason for this is that small frames have a smaller payload and less time for the element to process a frame before the next frame arrives. At high utilizations, the element may drop or corrupt frames.

– Payload: The payload is the PDU portion of the frame. For the most part, this portion is irrelevant to the Ethernet service. From a customer standpoint, this is the most critical portion of the service. Therefore, the ability to edit the payload may be a requirement for some turn-ups.

Traffic Rate – Constant Bandwidth

When setting utilization, there are several different units of measure. The two main units of measure are an actual bit rate (in Mbps or Gbps) or a percentage of the total available bandwidth. Stating bandwidth in terms of a percentage of the total available bandwidth is the most common.

When turning up a circuit, generating traffic at the maximum rate is the only way to confirm that the circuit can transmit the customer data at the guaranteed rate without errors. Depending on how the carrier offers the Ethernet service, the maximum bandwidth available to the end user may vary.

How Long to Generate Traffic?

Ethernet/SONET networks are often deployed with GBICs, SFPs, or XFPs. Depending on the quality of the GBICs, SFPs, and XFPs, the network may have a guaranteed error rate of 10-12 to 10-15. For a standard GigE network at an error rate of 10-12, the user will experience a few errors a day. At an error rate of 10-15, the user will experience about one error a week.

A 15-minute error-free test should be run to confirm that the network is properly provisioned and that the circuit is capable of handling the traffic. Longer tests offer more statistical information into the quality of the circuit and shorter tests offer less information.

The maximum bandwidth test should run error free and offer the customer proof that the circuit will transmit traffic appropriately.

Traffic Rate – Ramp

Another option for generating traffic is to step up (or ramp up) the traffic rate over time. The test involves setting a constant bandwidth, waiting for a short time, and then restarting the test at a higher bandwidth. The easiest way to accomplish this test is to have a test set do it for you.

By ramping up the traffic at specific intervals, the service can be proven to be error free at all line rates, not just at the maximum bandwidth being offered. If there are errors on the circuit, the step function will identify the rate at which the errors are being caused.

When setting up a ramp test, there are a couple of extra parameters that are required over the constant rate test. To begin generating any traffic, the technician must enter the step rate (2%, 5%, 10%, etc). Then, the technician sets the time interval at each step (20 seconds, 1 minute, 5 minutes, etc).

The ramp test, like the constant rate test, confirms that the service works and will transmit all of the customer's traffic without errors.

Traffic Rate – Bursty

Bursty traffic is a way to simulate real customer data, similar to the QRSS test pattern for a standard T1 Bit Error Rate (BER) test. When the test set is set to bursty, the test set varies the traffic in two important ways. First, traffic utilization is adjusted around a particular rate. If the technician sets the average at 50% traffic utilization, the utilization will fluctuate around 50%, which is similar to actual customer traffic.

The frame size is also varied by the test set. Customer traffic has a wide variety of frame sizes due to different applications and different requirements. By generating different sized frames in real time, the test set is able to emulate customer data more effectively.

Turn-Up Results

After setting up and generating traffic, the results of the test are analyzed in order to confirm that the service will or will not work per the service level agreement (SLA).

Interpreting Errors

When generating traffic, any received errors are an indication of a problem. Errors include runts, jabbers, and bad FCS frames. Either the customer network or the carrier network will drop any errored frames.

The screen shot below of the FST-2802 TestPad offers visibility into the types of errors that are tracked. Any errors will be displayed for the user to see. Errors will be displayed in two different categories – Error Stats and Summary. The summary view scans all of the results and displays any results that are out of specification as measured by the test instrument.

One important error result that is displayed in the summary view is the lost frames result. The FST-2802, using one of its packet generation options, can generate an Acterna test packet. This packet has a sequence number (similar to the TCP sequence number) and a timestamp. This allows the FST-2802 to perform real time QoS and SLA analysis including lost packet rate and round trip delay.
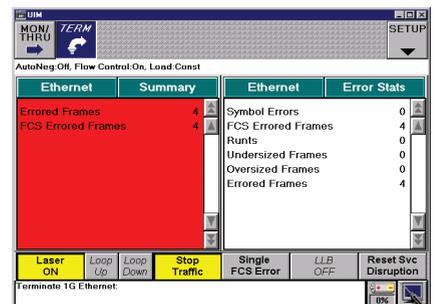
Interpreting Link Statistics

Whether the link has errors or is operating nominally, there are several link statistics the technician can use to confirm that the traffic he is transmitting is getting properly received by the test set.

The screen shot below displays the link statistics that the FST-2802 collects. All of the statistics captured on the screen gives the technician a complete view of how the circuit is behaving.

The first six results in the window are the ones that are used most frequently. The first three results show utilization as a percentage of total bandwidth, while the second three results show utilization as a frame rate.

On the right side of the window, notice the PAUSE Frames result. The pause control frames are those frames that tell the elements to slow down or speed up their transmission rate in order to most efficiently transmit the data.

RFC 2544 Testing

One of the most commonly used tests to turn-up Ethernet circuits is the

RFC 2544 test. RFC 2544 is a specification that was initially developed to qualify an Ethernet switch and define its capabilities. This test has been adopted to characterize Ethernet circuits prior to being handed off to the end user.

There are four tests that are part of the RFC 2544 specification and are relevant to circuit turn-up. They include:

– Throughput

– Latency

– Lost frames

– Back-to-back frames

– Throughput: The throughput test identifies the maximum bandwidth that the circuit will operate at. This value should match or be slightly greater than the provisioned bandwidth rate.

– Latency: The latency result provided by the RFC 2544 test is a round trip latency and is typically offered to the millisecond.

The latency of the network can be a somewhat tricky result. There are two main factors that create latency – network architecture and network traffic.

The network architecture adds a   fixed amount of latency into the network. The only way to reduce this component of latency is to physically re-route the circuit through fewer elements.

The network traffic component is only a factor for switched or routed Ethernet circuits. Transparent LAN services have reserved bandwidth in the elements, and therefore, traffic does not affect the circuit under test. On the other hand, the total traffic going through the router or switch can add latency to the circuit under test. Proper network capacity management, though, will limit this factor.

– Lost Frames: The RFC 2544 test will run a lost frames analysis at the rate determined by the throughput test. The appropriate result should be zero lost frames at the guaranteed throughput rate.

– Back-to-back frames: This test is mainly used in the manufacturing arena to determine the buffering capability of a single element. For circuit turn-up, this test is not required and doesn't offer useful information for the technician.

This test generates a number of frames at full line rate (10 Mbps, 100 Mbps, 1 Gbps, or 10 Gbps). The test set continues to drop the number of frames until the element under test drops none of the frames.

Troubleshooting

Troubleshooting Ethernet circuits presents a difficult problem. Many of the problems that occur on a data network often occur within the CPE LAN environment, outside of the carrier's domain of ownership. Troubleshooting and rectifying these issues is difficult without access to the customer network, which is not always easily obtained.

The carrier network can have its own set of problems, including traffic loading, capacity management, and fiber/cabling issues. One of the most common problems is fiber and cabling issues. Dirty fiber ends and improperly installed cable can affect a circuit's quality.

Using a test set to troubleshoot parameters such as utilization, errored frames, and the type of traffic can help the carrier determine where a problem originates.

Section 5: Common Terms and Definitions

There are a variety of terms that are applied to Ethernet technology. The following section explains many of these terms within the context of an Ethernet deployment.

(1) Utilization:
The utilization on a link is determined by comparing the packet rate to the overall bandwidth of a link. For example, a 100 Megabit Ethernet link (100 BASE-X) has an available bandwidth of 100 Mbps. If the packet rate is 40 Mbps, then the overall utilization is 40%. Utilization for data can fluctuate widely throughout the day on an operational network. Depending on user activity at any given time, utilization can jump from 0% to 100% and back to 0% within a few moments.

(2) Throughput:
Throughput is a measurement similar to utilization. Throughput, however, measures the number of packets sent versus the number of packets received. It is very much a quality of service metric. If 100 packets are sent and 90 of them arrive at the far end, then the throughput of the system is 90%.

(3) Round Trip Delay/Latency:
Round trip delay and latency are time measurements for a network. Round trip delay specifically addresses the time it takes for a packet to go from one point on the network to another point and back again. Latency is the time from one point to another. Latency can be measured for a single element (router) or for an entire network path.

(4) Frame Counts:
Frame counts are packet counts. Each packet, as it is received, is counted. It is also labeled as a good packet (good FCS) or a bad packet (FCS error). Occasionally, packets are also counted by their size.

(5) Payload:
The payload of a packet is often viewed as layer 4 and up, thus effectively the application. Payload can also be the line between the carrier service and the customer data. For frame relay, the line is between layer 2 and layer 3. For a point-to-point data T1, the payload line is between layer 1 and layer 2.

(6) Frame Size:
Frame size is measured from the beginning of an Ethernet packet to the end of the packet. Frame sizes range from 64 bytes to 1518 bytes. If VLAN tagging is added (layer 2.5), the maximum frame size increases to 1522 bytes.

(7) PDU:
A Protocol Data Unit (PDU) is typically the data contained at layer 3 and above within the data frame.

(8) CPE:
Customer Premises Equipment (CPE) usually refers to the customer equipment or the overall customer site.

(9) Retransmission:
Retransmissions occur when the far end does not receive all of the data that it was sent. TCP sequences packets so that lost packets are identified and retransmitted. The customer's router can often count these retransmissions and express them as an error condition. It is important to remember that errors in the lower layers can corrupt the higher layers, so retransmissions can be a sign of a problem with the physical layer, the data link layer, or even the network layer.

(10) Ping:
A ping is a packet that is sent from a source address to a destination address and back again. This allows the user to determine if the network will allow traffic to go from one point to another. If a customer can't ping the far end device, he can't send traffic there.

(11) Trace Route:
A trace route is a means for a user to trace all IP addres-sable devices in the network from one point to another. This allows the user to see all of the points along the way of a packet's journey.   An analogy to trace route is when FedEX scans a package as it goes through each one of its distribution centers. The user can see every point that the package touched on its journey.

(12) Runts/Undersize:
These errors are generically defined as any packet less than the minimum 64-byte length. In addition, the packet does not have an FCS value.

(13) Jabbers/Oversize:
A jabber is the opposite of a runt. These are overly long packets (>1518 bytes). Broken NIC cards/ports often cause jabbers.

(14) Bad FCS:
Bad FCS frames are those frames with an incorrect FCS value. These errors are counted when one or more of the bits in a packet have been switched (from a 1to a 0, for example).

(15) Collisions:
On a half-duplex Ethernet link, each computer must share the bandwidth with the rest of the computers. If two or more computers on a network broadcast at the same time, the packets "collide" and are unreadable. This event is known as a collision. On full-duplex links, this is a non-event because there are separate transmit and receive paths.

(16) Symbol Errors:
Symbol errors represent line coding issues at the physical layer.

MTU versus Frame Size

There is some confusion between the maximum transmission unit (MTU) settings in routers and switches and the frame size generation setting when dealing with Ethernet turn-up and troubleshooting. The frame size is the overall length of the Ethernet frame at layer 2 in the OSI model. Standard Ethernet frames range from 64 bytes to 1518 bytes without a VLAN tag and from 64 bytes to 1522 bytes with a VLAN tag. Larger Ethernet frames, typically called jumbo frames, can have a frame size as large as 10,000 bytes.

The MTU setting is a layer 3 packet size setting, often available in routers and switches. As its name implies, the user would set the MAXIMUM transmission unit as the largest packet that the router or switch can pass. Because this is a layer 3 measurement, the user must subtract the layer 2 (MAC) size and the layer 2.5 (VLAN tag) size of the overall frame to correctly set the MTU.

For most Ethernet networks, the MTU size should be set to 1500. This allows 1518 byte non-tagged Ethernet frames and 1522 byte VLAN-tagged Ethernet frames to pass the entire network untouched. Any frames larger than 1518 or 1522 bytes will be broken up into smaller frames by the first router and forwarded as multiple frames. MTU sizes can be set as high as 9000+ bytes to support jumbo frames.

Test & Measurement Regional Sales

NORTH AMERICA	LATIN AMERICA	ASIA PACIFIC	EMEA	WEBSITE: www.jdsu.com
TEL: 1 866 228 3762	TEL: +55 11 5503 3800	TEL: +852 2892 0990	TEL: +49 7121 86 2222	
FAX: +1 301 353 9216	FAX: +55 11 5505 1598	FAX: +852 2892 0770	FAX: +49 7121 86 1222	