CHAPTER

6

Ethernet Concepts

The system of nature, of which man is a part, tends to be self-balancing, self-adjusting, self-cleansing. Not so with technology.

— E.F. Schumacher

The term *Ethernet* is a catchall word used to describe the most common network architecture used in a majority of today's networks worldwide. If you were to say to someone, ''Describe an Ethernet cable,'' 99 out of 100 would probably respond that it consists of unshielded twisted pair (UTP) cable that is terminated on each end with RJ45 plugs. That is mostly true in today's network, but Ethernet technology has evolved from its early coaxial cable days to what it is today.

All Ethernet networks, no matter the type of cable that is in use, are Carrier Sense Multiple Access with Collision Detection (CSMA/CD) networks that adhere to the standards described in IEEE 802.3. This is true for either coaxial or UTP cable Ethernet networks. Let's review how Ethernet came about and how it evolved to its current emanation of Ethernet cable technology.

NOTE The term *Ethernet* is derived from two words: *ether* and *net*. Ether is a medium that can be made from pretty much anything. This is evident in today's network environment, where network signals can be carried over wire, fiber (fiber optic), or air (wireless). The word *net* may be short for *network*, but one of the authors likes the idea of visualizing a fishing net, where each node is tied to adjoining nodes, and there are multiple paths from one to the other.

6.1 The Beginning of Ethernet Technology

From 1973 to 1975, Ethernet had its start at the Xerox Palo Alto Research Center (PARC). Xerox filed a patent application in 1975 with the U.S. Patent Office for a Multipoint Data Communication System with Collision Detection. Patent 4,063,220 described how multiple data processing stations distributed along a branched cable segment would be able to communicate with each other. It included descriptions of the cable the devices needed to send and receive data on that cable. It also included a packet description outlining both source and destination addresses along with data and error fields.

In the experimental implementation of Ethernet, data rates were 3 Mbps, and the source and destination address fields were only provided 8 bits for addressing, which limited the number of devices that could be addressed on the network. There were 16 bits allocated for the packet type, which would be used to define a packet type that would be used within a particular protocol.

NOTE **Mbps means "megabits per second," where mega is the value of a million. So 100 Mbps is 100 million bits per second. Remember that a bit is a single binary digit of either zero or one. Even if only one stream of zeros was being generated, there are still 100 million of them in a second. It may represent a whole lot of nothing, but in the network world they truly have value.**

One of the original inventors on the Xerox patent, Robert Metcalfe, left Xerox in 1979 to form 3Com to promote LAN development and the use of PCs as nodes on the Ethernet network. He was instrumental in convincing Digital Equipment Corporation (DEC), Intel, and Xerox to work together to promote Ethernet as a LAN standard. This standard came to be known as the *DIX standard*, after the companies (DEC, Intel, Xerox) who came together to create the standard.

The DIX or Ethernet II standard describes a frame format that provides 48 bits each for destination and source addresses, along with 16 bits for the packet type. The standard also set the data rate at 10 Mbps. Figure 6-1 illustrates a DIX/Ethernet II frame.

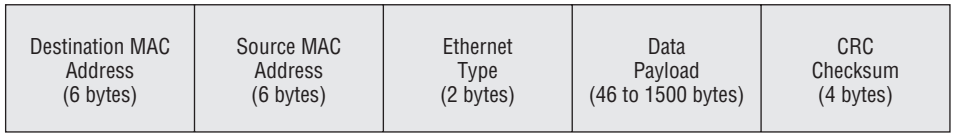


Figure 6-1 A DIX/Ethernet II frame

The Destination and Source Address fields are 6 bytes in length and are usually presented as a group of 12 hexadecimal numbers. These addresses are

called the *Media Access Control (MAC) addresses* and are a unique Ethernet hardware address assigned to a network interface card (NIC). The DIX/Ethernet II standard has been superseded by IEEE 802.3.

NOTE **Hexadecimal number system is an easy way of illustrating 4 binary bits, which can have values from 0 to 15. The values 0 through 9 are presented as their actual value, while the units 10 through 15 are represented by the alpha characters A through F, respectively. The 16 (the root *hexadeca* means 16) values that can be contained in a hexadecimal number are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E and F.**

Although Ethernet was originally designed to allow computers to communicate with each other over a coaxial cable as the broadcast transmission medium, twisted pair Ethernet cable systems have been under development since as early as the mid-1980s. The first network topology using UTP cable was StarLAN, and it was introduced with a data rate of 1 Mbps. However, StarLAN would eventually evolve into what became known as *10BASE-T*, which is the predominant UTP cable in use today.

Since the publication of IEEE 802.3 in 1985, there have been several amendments that provide for increased Ethernet rates. Table 6-1 lists the data rates that can be found in use today.

Table 6-1 Ethernet Types and Speeds

ETHERNET TYPE	SPEED
10BASE-T	10 Mbps
Fast	100 Mbps
Gigabit	1000 Mbps

POP QUIZ

What was the first type of cable used to form an Ethernet network?

Ethernet has emerged as the de facto network standard worldwide. It has withstood challenges from other networking protocols over time, and as a result, large numbers of products from a wide range of manufacturers are readily available and are able to successfully interconnect based on this standard. Due to the economies of scale, networking products have decreased in price while performance has increased. Ethernet allows for flexibility in network implementation that is easy to maintain and manage. The installed base for Ethernet networks

is huge, guaranteeing that Ethernet will be around for some time to come. There will always be improvements inserted into existing networks, but they will not cause a total dumping of the current Ethernet network.

6.2 Ethernet Components

We discussed how UTP cable evolved from UTP telephone wire used to create the StarLAN networks. It would stand to reason that some of the concepts would be carried over from the Telco influence in setting Ethernet standards. Ethernet components using UTP cabling fall into two categories:

- Data terminal equipment (DTE)
- Data communications equipment (DCE)

This nomenclature is part of the long-standing serial communications standard EIA RS-232. Much like that standard, the Ethernet standard uses this framework as the basis in developing standards for the electrical signal characteristics for Ethernet cabling and signals. Figure 6-2 illustrates a DCE and DTE device connected with UTP cable.

RANDOM BONUS DEFINITION

bridge port — A network interface on a bridge.

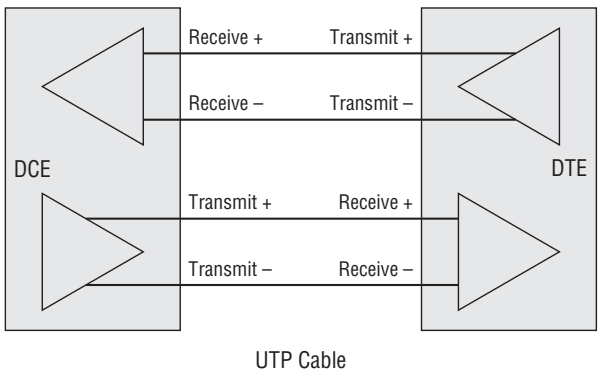


Figure 6-2 Interconnection of DCE and DTE Ethernet devices

This figure represents the conceptual interconnection of two Ethernet devices using UTP Ethernet cable. You will notice that the cable appears to be straight across, although physically the + and − wires are twisted together within the jacket of the cable. This type of Ethernet cable is often referred to as a *patch*

cable or a *straight-through cable* because there is no crossover from receive to transmit circuits.

NOTE **Twisted pair wire does have a purpose. The pair of wires are twisted together in a uniform manner with a fixed number of twists per foot. Why should the cable be twisted in the first place? To look pretty? To keep the wires from drifting apart? Okay, the answer is: to combat the effects of electromagnetic interference (EMI). Electrical waves are all about us, now more so than ever with the plethora of cell phones and other mobile devices. When these waves intersect wire, they can induce minute fluctuations in voltage. No big deal, right? Just a little static on the line. Wrong! These signals could cause erroneous data to be read, so signal integrity is an absolute necessity. (How would you like it if your ATM card was swallowed before you could get your money out?) Now, do not go adding extra twists to your Ethernet cable thinking this is going to increase your immunity. In reality, you will alter the electrical characteristics of the wire and cause reflections within the cable, which is bad as EMI. Leave the cables alone and go pop some bubble wrap if you need to keep those idle hands busy.**

So, we have DCE and DTE Ethernet devices, but which is which? A good way to remember this is by recalling the early days of RS-232. The term *data terminal equipment* often referred to teletypewriters, whereas *data communication equipment* most often referred to modems. When PCs were introduced, the majority of telecommunications was accomplished via a modem. (Yes, we recall those days — the 300 baud handset devices where you squeezed your phone's handset into the foam cuffs so it could receive the actual audio signals through the telephone.)

NOTE **A *handset* refers to a standard telephone like we had back in the olden days. The telephone wire was connected to the base, and the handset portion had a spiraled wire, which always managed to get so twisted that you found you could not talk on the phone unless your head was about a foot off the table where the base rested. The base contained the actual dialing mechanism, which allowed you to dial the number you wished to connect to. Yes, "dial" — where do you think the word originated? Surely, not from punching those minute buttons on the latest whiz-bang cell phone, which has given us a new set of human ailments such as "texting thumb."**

RANDOM BONUS DEFINITION

bandwidth — The data-carrying capacity of a device or communications channel.

As telephone technology evolved from mechanical dialing mechanisms to touch-tone dialing, modems also implemented those technologies. Even today's modems — whether external or internal modems embedded in a laptop, PCMCIA modem card, or PCI

modem card in a desktop computer — all support both dial and touch-tone dialing methodologies in their designs.

NOTE **What is meant by mechanical dialing? The old rotary phones had a dial with numbers and letters assigned around a dial mechanism shaped like a wheel with finger holes assigned the numbers 1 through 9 and 0 for either the number zero or Operator if that number was dialed first and by itself. A number was dialed by placing one's index finger in the hole with the corresponding number that was desired and then in a circular motion moving the dial to the stationery finger-stop and releasing the dial to allow it to step back. As it stepped back, it sent a pulse on the wire to the home office, where stepping relays would increment to set up the circuit corresponding to that number. Switching theory was developed and used by the telephone companies in order to eliminate human operators who would actually make the circuit connection for the caller. The number selected would determine the number of pulses, which stepped the home office stepping relay to that number. You can just imagine how many relays were required to set up those switching offices. Today's modems use a relay to pulse line the number of times required for the number to be dialed, and that is what is meant by the pulse setting on the modem.**

Touch-tone dialing was devised by the telephone companies to accomplish pretty much the same thing as pulse dialing. However, it uses a more modern technique of using distinct audio tones for each discrete number. If you ever listened to a modem dial with tone dialing, you know it sounds like automatons in sci-fi movies.

PCs pretty much replaced teletypewriters as the device to use for telecommunications. They were supplied with RS-232 serial ports. With a terminal emulation program, these PCs became the modern-day teletypewriter. We said that teletypewriters were DTE devices, so the PC with an Ethernet NIC is an Ethernet DTE device. Modems are DCE devices, and since they pass data along the network, devices like Ethernet hubs, routers, and switches are also considered to be DCE Ethernet devices.

POP QUIZ

An Ethernet network device that forwards data on the network would be considered what type of Ethernet device?

6.2.1 DCE and DTE Cabling Considerations

We mentioned that a straight-through cable was one where the wire from pin 1 would be connected to pin 1 on the other connector. Let's discuss the RJ-45 modular plug that is used on any UTP Ethernet cable. Figure 6-3 represents how an RJ-45 plug would look if you held the plug with its gold contacts facing you. Pin 1 of the plug will be on your left, with pin numbers incrementing until pin 8 on your right is reached. The pin numbering is sequential.[1]

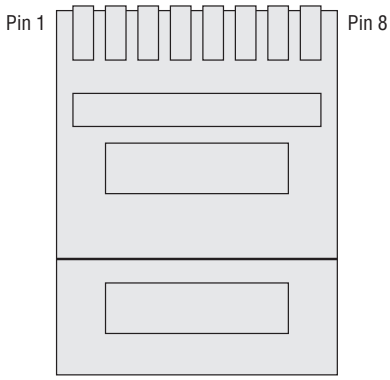


Figure 6-3 An RJ-45 modular plug

UTP Ethernet cable consists of four twisted pairs,[2] for a total of eight wires contained within an unshielded jacket. The wires are colored with four solid colored wires, each of which is twisted together with its mate, which is mostly white with a colored stripe that matches the color of its solid colored mate. How and to what pin these wires connect to on the RJ-45 plug adhere to old telephone company standards and are contained within the TIA/EIA-568-A and TIA/EIA-568-B standards. Table 6-2 lists the wiring scheme for T568A wiring, and Table 6-3 lists the wiring scheme for T568B wiring.

[1]Sequential is derived from the word *sequence*, which means one after the other. For those in the reading audience who find it difficult to grasp this concept, we shall be more precise in the pin numbering definition. Starting on the left with pin 1, the pin numbers increment in sequence: 2, 3, 4, 5, 6, 7, and pin 8, which is the last pin on the right. Now, if you tell us you can't count, then we have a major problem here, and you need additional help, which is beyond the scope of this book.

[2]Pair refers to the number two. So a twisted pair of wire would consist of two discrete wires which have been twisted together for . . . what? Noise immunity, good answer.

Table 6-2 T568A Wiring Pin-out

PIN	PAIR	WIRE	COLOR	ETHERNET SIGNAL
1	3	Tip	White/green	Transmit +
2	3	Ring	Green	Transmit –
3	2	Tip	White/orange	Receive +
4	1	Ring	Blue	
5	1	Tip	White/blue	
6	2	Ring	Orange	Receive –
7	4	Tip	White/brown	
8	4	Ring	Brown	

Table 6-3 T568B Wiring Pin-out

PIN	PAIR	WIRE	COLOR	ETHERNET SIGNAL
1	2	Tip	White/orange	Transmit +
2	2	Ring	Orange	Transmit –
3	3	Tip	White/green	Receive +
4	1	Ring	Blue	
5	1	Tip	White/blue	
6	3	Ring	Green	Receive –
7	4	Tip	White/brown	
8	4	Ring	Brown	

A straight-through cable can be wired with either the T568A or T568B wiring scheme as long as both ends of the cable are wired exactly the same using the same wiring pin-out.

A crossover Ethernet cable must have one plug wired with the T568A wiring scheme and the other plug wired following the T568B wiring pin-out. The purpose of a crossover cable is to interconnect to like devices, regardless of whether they are

RANDOM BONUS DEFINITION

Application layer — The highest layer of the seven-layer OSI model.

DCE or DTE devices. The crossover is to have the transmit signals from one device terminate on the receive signals of the other device so they can pass data between them. A quick analogy is connecting two microphones together; the two parties could scream into them but neither could hear the other. Now, if we take one microphone and crossed over to a speaker and did the same for the other microphone, then parties would be able communicate without a problem.[3] The same goes for Ethernet devices — just because there is some sort of Ethernet UTP cable strung between them does not mean they are ''supposed'' to communicate.

So, when you are having problems getting two Ethernet devices to communicate, the first place to look is at the Physical layer (such as the cable being used).

HELPFUL HINT

Since for the most part Ethernet cables use RJ-45 jacks, which are mostly clear plastic, it is fairly easy to determine if a Ethernet UTP cable is either a straight-through or crossover cable. Take the two connectors on the ends of the cable and hold them against each other with both plugs oriented in the same direction. Scan the colors of each. They should look exactly alike on a straight-through cable. If it is a crossover cable, you will notice that the colored wires on pins 1 and 2 of one plug have moved to pins 3 and 6 of the other, with the reverse also being true.

If for any reason the cables do match as described in this note, there is a likelihood it is a cable used for another purpose or it is supposed to be an Ethernet UTP cable but has been manufactured incorrectly.

Do yourself a favor: if you find cables in your box of goodies that appear different from what has been described in this note, discard them in the nearest wastebasket. Many countless hours have been wasted fighting problems with bad cables, not only by people in general but by network administrators who should know better.

For the frugally minded who cannot bear to toss anything away, our recommendation is to cut the ends off the cables so you will not be tempted to use them in your network. You may want to use them to tie up all those newspapers that have been collecting in the corner and bring them to a recycling drop-off in your community.

6.2.1.1 Interconnecting Like Ethernet Devices

We have already discussed that Ethernet devices fall into two categories, DCE or DTE type devices. It has also been stated that interconnecting to like Ethernet

[3]We fully acknowledge that his simple-minded analogy has very little likelihood of succeeding in the real world because there is a whole lot of electronics that needs to be added for it to actually work. The purpose of any analogy is to demonstrate in the simplest terms how something works.

devices requires the use of a crossover cable. For example, two PCs with NIC cards can be directly interconnected with a crossover cable, as illustrated in Figure 6-4.

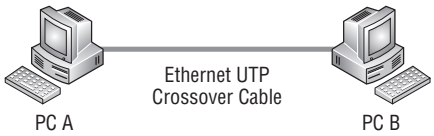


Figure 6-4 Two PCs interconnected via Ethernet

In this simple figure, the two computers are able to communicate with each other over the crossover cable. There must be some sort of networking protocol running on the PCs, such as TCP/IP, and some sort of application that will allow the sharing of data or devices (which may be locally connected to either or both of them). Some operating systems, such as Microsoft Windows and Apple Macintosh, are ''network-able'' and include tools and utilities to facilitate data and device sharing over the network.

POP QUIZ

If a cable is wired such that one plug is a T568A and the other is a T568B, it would commonly be referred to as _____ cable.

The last example showed two Ethernet DTE devices interconnected, but how about DCE devices? We already mentioned that DCE devices are in the form of hubs, switches, and routers, so we know we are dealing with that kind of device. Why would anyone want to connect those types of devices? To illustrate this, we will consider a few simple examples.

ACRONYM ALERT

BER — Bit error rate

The first example is a case where we have a stack of dumb,[4] eight-port, passive hubs and there is a small office with 15 workers who need to be interconnected to a local server to share the resources available on that server. Figure 6-5 illustrates one method of how these passive eight-port hubs may be used to accomplish this.

The three hubs are placed about the office for the ease of cabling between each other and the workstations connected to them. Since these hubs have eight ports, with one

ACRONYM ALERT

TTL — Time to live

[4]Dumb means exactly that: dumb. There is no internal intelligence contained within the unit.

port dedicated for linking to the other hub, this leaves seven available ports for workstation connections. As you can see in Figure 6-5, two of the hubs have seven workstations each connected to them. That leaves one workstation and the server to be connected to the LAN. The hub that is used to connect these devices and the other two hubs has only used four of the eight available ports, so if needed there are four ports remaining for future expansion. You can see from the cabling legend that the workstations and the server are connected to the LAN with a patch or straight-through Ethernet UTP cable. The hubs are connected to each other using crossover cables since we are interconnecting like DCE Ethernet devices.

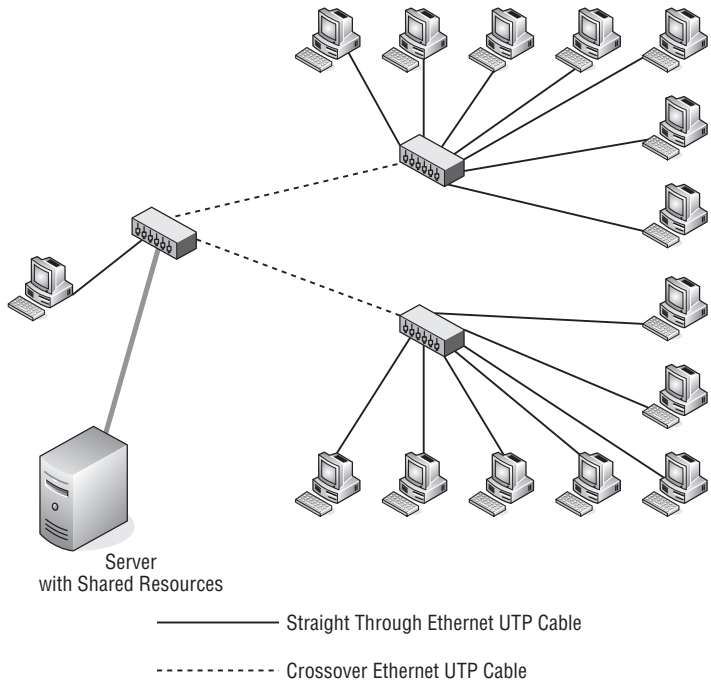


Figure 6-5 A LAN created with passive hubs

This scenario is not uncommon, and a few of you who may be familiar with cabling hubs today may be scratching your head. We remember the day when this was standard operating procedure for interconnecting passive hubs, so go with us on this one. Yes, there have been improvements in hub technology. One was actually adding what was called an *uplink port*, where a DTE port was added to the device to facilitate it being connected to another hub, with a patch cable eliminating the need to find a crossover cable, in case you forgot to purchase one when you purchased the hub. Another improvement is an uplink port with a switch dedicated to it that switches its receive and transmit

circuits to match the cable and the port it was connected to at the other end. The most recent innovation in hub and switch design is that all ports on the hub are now auto-sensing and auto-switching.

NOTE **Auto-sensing is accomplished by electronic circuits that determine if the incoming wires to a signal pair of pins are connected to a transmitter or a receiver. Once the "sense" of the wire is determined, this information is passed to the circuits responsible for auto-switching.**

Auto-switching is circuitry added to a port to configure the port to which pins receive and transmit circuits should be connected to. If one set of pins is determined to be a receive pair, then the other set of pins must be the transmit pair. Receive and transmit are mutually exclusive in that one set of pins must be the receive circuit and the other must be the transmit circuit. If both sets of pins are the same, either receive or transmit, the device is defective.

HELPFUL HINT

Most Ethernet devices with RJ-45 jacks to accommodate Ethernet UTP cables have LED[5] lights showing the link status. If there is no link indication, the first place to check is the cable. Both devices connected with the same cable should indicate link while connected. If you pull one end of the cable and the other device's link light is still illuminated, you may not be connected to the correct device. In large LAN implementations, many times a cable is pulled to ensure that it loses link so one knows the port assignment is correct on both ends of the cable.

We can see that look on your face. You are thinking that if devices can do auto-sensing and auto-switching, why do you have to learn the differences in cable types? The answer is, you may be correct if you are only doing new implementations and using stock cables you buy already assembled. However, there is a large installed base of legacy systems that have dedicated ports wired as either a DTE or DCE, so cable knowledge is essential.

Let's continue with another example. Remember, it still is not yet an auto-sensing/auto-switching world. Figure 6-6 shows a part of a larger installation at a corporate office. There are many user workstations, but for sake of illustration there

[5]LED is the acronym for light emitting diode. It is actually a semiconductor device that will illuminate when a current is passed through it. Some are single colored while others are able to change color depending on how the device is electrically driven.

are only a few in the figure drawing. This figure may represent a floor or department location within a building.

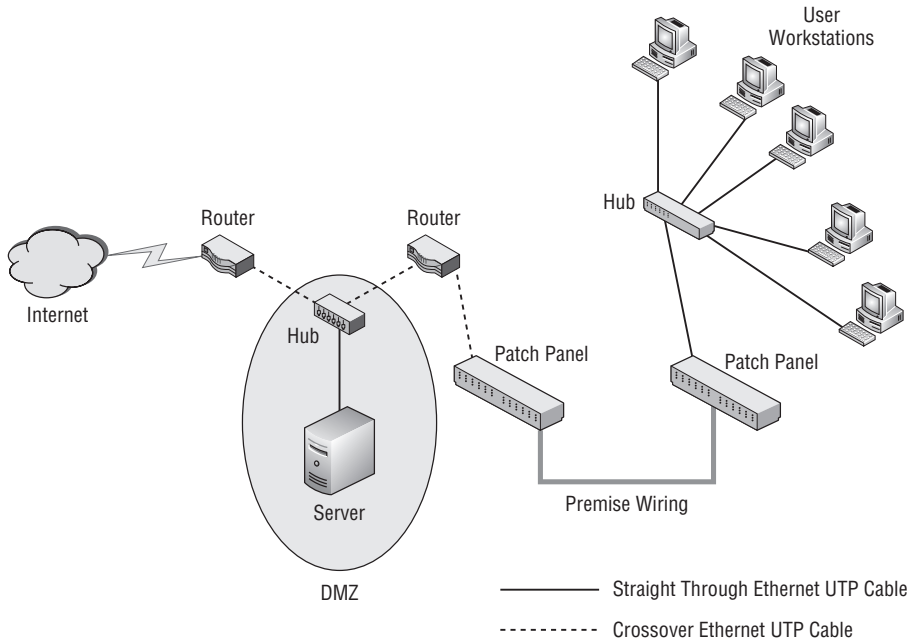


Figure 6-6 A larger LAN implementation

RANDOM BONUS DEFINITION

multimode fiber — An optical fiber that allows signals to propagate in multiple transmission modes simultaneously.

There are three DCE devices in this drawing, two routers and a hub that are interconnected using crossover cables. Off the hub there is a server connected with a patch cable/straight-through Ethernet UTP cable. The placement of the hub and server is considered a DMZ (demilitarized zone). The purpose of a DMZ is to regulate access to the networks it is connected to. In this scenario, there is a network of corporate user workstations that have access to a corporate server and the Internet. The routers within the DMZ have been programmed with policies that allow approved users from the Internet to have access to the corporate server but not to pass to any other networks connected to the DMZ. These routers and other equipment may be located in a data center on another floor from the users who need access to the server and the Internet. This is where premise[6] wiring comes in.

[6]*Premise* is the term used to represent a given locale like a home or building. Thus, premise wiring is the wiring contained within the building.

Cable needs to be run from the data center to the floor where the user workstations are located. This is done by running Ethernet-grade[7] cable, which is terminated on patch panels[8] located in the data center and the wiring closet on the floor where the user workstations are located.

HELPFUL HINT

We have seen wiring closets that are neat and orderly, and others with wire strung everywhere and piled on the floor like a large bowl of my mother's spaghetti and meatballs. (For more information on my mother's secret recipe, read the note on it.)

If you are a network administrator and want to do yourself a favor, please try to keep your wiring closets orderly and well labeled. You do not want to be called at all hours of the night or on vacation or even on your weekends off, and that will be the case each time someone is troubleshooting a problem and has no clue as to which cables go where. Do it right up front and you can truly have peace of mind. If not, your ears will be burning each time someone curses you for making their job harder.

The patch panels are wired with Category 5e or Category 6 cable from panel to panel as straight-through cables. There is no crossover taking place within the long-run cables. If a crossover is needed, it will be taken care of from the patch panel to the device using an Ethernet UTP crossover cable. This is illustrated in Figure 6-6 with the router that is connected to the patch panel. Notice on the other patch panel that although the switch is a DCE Ethernet device, it is connected with a patch cable. This is because it connects to the router at the other end, which is connected to the patch panel with a crossover cable, so that only a single crossover is required. Double crossover[9] cables will basically negate the crossover function, and the device link lights will not illuminate.

[7]Ethernet using UTP cable was initially designed on the idea of using existing premise wiring that was in place for telephone communications. With improvements in speed on Ethernet circuits, a higher quality cable was necessary to support these new requirements. Today's new cable installations should be using Category 5e or Category 6 cable, especially if Gigabit Ethernet is to be used.

[8]Patch panels are an old holdover from the telephone company days. However, remember the basis of Ethernet over UTP was to use existing premise wiring, which was telephone UTP cable. It stood to reason if those cables are attached to patch panels, then patch panels would become part of the Ethernet UTP connectivity equation.

[9]Double crossover is like a double negative: two negatives make a positive, so you don't have the crossover. It may come in handy sometime when you find yourself up to the armpits in crossover cables but are unable to find that one badly needed patch cable. Now, how would you connect them?

POP QUIZ

You are interconnecting two Ethernet devices, but neither device is showing a link light on the assigned port. List in order of likelihood where the problem might be.

The server and all the user workstations are DTE devices connecting to other DTE devices, so the cables used are straight-through (patch) Ethernet UTP cables. With the right routing protocols and security policies in place, users at the user workstations are able to access the local corporate server as well as the Internet, while the corporate LAN is protected from unauthorized users from the Internet.

AN UNRELATED MOMENT OF PAUSE – MAMA BRAMANTE'S SECRET SPAGHETTI AND MEATBALLS RECIPE

The thought of all of the cables in a wiring closet made Rich think of his mother's spaghetti and meatballs. Rich decided to share the recipe with you all: Well, the recipe is not under lock and key like you see in some of those commercials on TV, and no, the dog doesn't know it either. The reason it is so secret is that my mother had the knack of making it without measuring ingredients other than with her watchful eye. I always said she could cook for five or fifty and it would always be the same, and it was. There is nothing like a mother's cooking, eh?

So, I am going to give you a list of ingredients, and you can mix up a batch. You may surprise yourself and it could be almost as good as my mom's. My mother always started the sauce before the meatballs. (For you Italian readers out there, "sauce" is "gravy.")

- **Sauce Steps:**
  1. **Using a large pot, pour a liberal amount of olive oil to a depth of about a quarter of an inch and heat to a temperature that would fry whatever you place in it.**
  2. **Slice up (slice, not dice) a medium-sized onion. Add the onion to the oil and brown to a dark crisp. Remove the onion from the oil and set aside.**
  3. **Take some garlic cloves and slice them so you have these tiny garlic slabs. Add them to the oil and just brown (do not cook as long as the onions).**
  4. **Once the garlic is brown, add two cans of peeled Roma tomatoes into the olive oil/garlic mix. Be careful that the oil does not splatter back.**

(continued)

AN UNRELATED MOMENT OF PAUSE — MAMA BRAMANTE'S SECRET SPAGHETTI AND MEATBALLS RECIPE *(continued)*

5. Stir in one can of tomato paste and the fried onions. Stir for consistency and let simmer while making the meatballs.

◆ **Meatball Steps:**

1. Put about a pound or pound and a half of fresh ground beef into a large mixing bowl.
2. Grate in an amount of bread crumbs that is about a third of the hamburger volume. (Stale Italian bread allowed to thoroughly dry to a rock was used to make the bread crumbs. Not much was wasted when feeding six kids.)
3. Finely dice two garlic cloves and add to the mix.
4. Finely chop two or three sprigs of fresh parsley and add to the mix.
5. Grate in some fresh Parmesan or Romano cheese — about half a cup or slightly more.
6. Add salt (not too much, as the cheese is salty) and some ground black pepper.
7. Create a cavity in the mix and add three whole eggs into the mix.
8. Mix all the ingredients thoroughly so that the whole batch is consistent throughout.
9. In a large skillet, preheat olive oil to fry the meatballs in. Scoop up enough of the beef mixture to make a golf ball size meatball. Roll the meatball in the palm of your hand (wash your hands before and after this process) to form a firm ball that can withstand frying without falling apart.
10. Fry the meatballs to a deep brown crust on all sides before dropping them into the sauce.

◆ **Spaghetti Steps:**

1. Once everything is simmering in the large sauce pot, it is time to boil the water for the spaghetti.
2. Add a half teaspoon of salt to the spaghetti water and bring to a rapid boil.
3. Add a pound of spaghetti (smaller amount for a smaller gathering) to the water and stir in.
4. Keep an eye on the pot since rapidly boiling spaghetti has a tendency to foam up and overflow the pot.

(continued)

AN UNRELATED MOMENT OF PAUSE – MAMA BRAMANTE'S SECRET SPAGHETTI AND MEATBALLS RECIPE *(continued)*

5. **Once the spaghetti is cooked and is soft but firm to the bite (al dente), strain it in a colander.[10] Make sure the spaghetti is well drained.**

◆ **Serving Steps:**

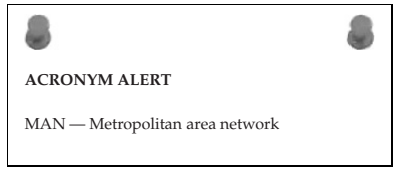
1. **Dump the colander of spaghetti into a large serving bowl.**
2. **Add some of the sauce (no meatballs) to the spaghetti and mix thoroughly to where the spaghetti has sauce on it but not is swimming in the sauce. I know there is a fine line to this, so add sauce slowly.**
3. **Once you are satisfied the spaghetti has sufficient sauce on it, fish out two meatballs for each diner and place in the bowl on top of the spaghetti.**
4. **Serve with freshly grated cheese on the side, a little vino, good company, and conversation.**

Congratulations! You have just served up Mama Bramante's favorite dish to *la famiglia*.

6.3 Ethernet and IEEE 802.3's Relationship to the OSI Model

There is a close similarity between the ISO OSI model and IEEE 802.3 model, with the difference being at the Data Link layer of the OSI model, as illustrated in Figure 6-7.

The Physical layer is the same in both models and is dependent upon the media[11] being used. This layer deals with parameters such as cable pin-out, signal electrical characteristics, modulation encoding of the data being modulated on carrier signals, and data synchronization.[12] Once it has been determined that the receive buffer has received a complete frame, the Data Link layer is signaled and the frame is passed up to that layer.



[10]For those of you who are uninformed about cooking utensils, a *colander* looks kind of like a leaky bucket or a hemispherical pot shot full of buckshot holes. Not useful for holding water, but it sure comes in handy when draining spaghetti.

[11]Media is in reference to the method of delivery of the data. Obviously in a wired network it depends on the type of cable and the NIC cards being used. However, other methods of delivery such as wireless and optical can be used. So media for the most part is how the data moves between data points.

[12]Data synchronization refers to the capability to detect the start of a data frame from a stream of data bits and the fact that the binary pattern is a complete frame.

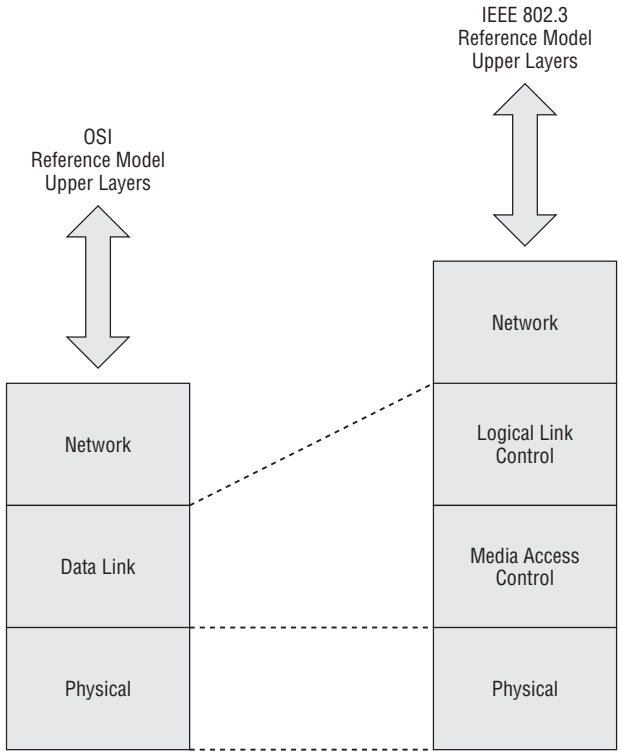


Figure 6-7 OSI's relationship to IEEE 802.3

In the OSI reference model, the Data Link layer accepts service requests from the Network layer and sends service to the Physical layer. It is the layer responsible for data transfer between adjacent network nodes and has the capability to detect and correct errors that may occur on the Physical layer. Although the Data Link layer is responsible for data transfer over the Physical link, many data link protocols do not provide acknowledgments of a successful receipt and acceptance of a frame. Some data link protocols do not even provide for a checksum to detect errors in transmission. In these cases, frames received depend on higher-level protocols for frame flow control, acknowledgments, retransmission, and error checking.

The IEEE 802.3 reference model divides the OSI model's Data Link layer into two sublayers, the Logical Link Control sublayer and the Media Access Control sublayer. The Logical Link Control sublayer resides in the upper layer of the OSI Data

POP QUIZ

Into which two sublayers of the IEEE 802 reference model is the OSI reference model Data Link layer divided?

Link layer, whereas the Media Access Control sublayer is in the lower portion and provides the interface to the Physical layer.

6.3.1 Logical Link Control

The IEEE 802 standard for the Logical Link Control resides in the upper portion of the OSI reference model's Data Link layer and provides the same functions no matter what media is being used. The Physical layer can be Ethernet, Token Ring, or wireless LAN, of which the Logical Link Control sublayer is primarily concerned with providing flow control, error control, and what multiplexing protocols are being used over the Media Access Control sublayer.

Logical Link Control flow control manages the data transmission rate between two network nodes to prevent one node sending faster than the speed of the receiving node. If one node is receiving data from multiple network nodes, it may not be able to receive as quickly as the sending node would like to transmit. Flow control depends on feedback from the receiving node to the sending node signaling possible congestion and its inability to receive data at higher speeds. In an Ethernet network, a receiving node that is unable to keep up with a sending node will transmit a PAUSE frame to halt transmission for a given period of time. The PAUSE frame for flow control can be used only on network segments that are running at full-duplex.[13]

POP QUIZ

With which functions is the Logical Link Control sublayer mainly concerned?

6.3.2 Media Access Control

The Media Access Control sublayer provides the interface between the Physical layer and the Logical Link Control sublayer. The Media Access Control sublayer is responsible for data encapsulation and frame assembly for sending frames, and de-encapsulation and error checking of received frames. It also provides addressing and a channel access control mechanism, which allows multiple nodes on a local area network to communicate.

The Media Access Control address, or the physical address of the node device, is commonly referred to as the *MAC address*. It is an industry standardized unique address assigned to each network adapter at the time of manufacture. Although highly unlikely, there is a possibility of duplicate

[13]We previously defined full duplex as the capability to send and receive simultaneously. It is logical that if a half-duplex node is currently receiving, it is unable to transmit until all the data is received. This makes a PAUSE frame unusable in half-duplex network segment.

MAC addresses on a network segment due to the capability to overwrite a manufacturer's previously assigned MAC addresses.

HELPFUL HINT

Although I have seen only one case of a duplicate MAC address on a LAN segment, I know it is possible. Depending on the network size, it can be a real nightmare. (Unfortunately, for the case I worked, it was a large network.)

For whatever reason, the site in the case I worked decided that they would assign their own MAC addresses for every device in their network. Although they had full control and well-documented logging of MAC addresses, it took a while to find the offending node.

Ultimately, knowing the MAC address of the device that was being adversely affected was helpful. Using a process of elimination that allowed for a digit being entered into a MAC address incorrectly aided in locating the culprit. If the site had not properly documented their MAC addresses and where they were assigned, the other option would have been to assign a new MAC address (which they preferred not doing) to the device that had not been previously assigned.

I am sure they had good reasons to use their own MAC address scheme, and they attempted to document it well, which is a major plus. However, it is best to leave well alone and use the already assigned MAC address to identify the device on the LAN segment.

Because Ethernet is a CSMA/CD (Carrier Sense Multiple Access with Collision Detection) network protocol, not only are all the network nodes on a network segment required to have unique physical hardware addresses, but there must be a provision for the control of the multiple access of more than one node at a time. The Media Access Control sublayer provides channel access control to allow multiple access.[14]

When multiple network nodes are connected to the same physical media, there is a high likelihood of collisions occurring. The multiple access protocol is used to detect and avoid packet collisions where multiple nodes contend for access to the same physical media. Ethernet and IEEE 802.3 are the most common standards used for CSMA/CD networks.

CSMA/CD utilizes a carrier-sensing scheme. If a transmitting node detects another signal on the media while it is transmitting a frame, it ceases transmittal of that frame and immediately transmits a jam signal onto the media. All nodes on the network are aware a collision on the media has taken place and will

[14]Multiple access allows more than one data stream to share the same Physical layer media. Examples of shared media networks are bus topology networks, ring topology networks, wireless networks, and Ethernet point-to-point links running at half duplex.

back off and not transmit for a period of time, which is calculated using a back-off delay algorithm. After the back-off delay has elapsed, the node will attempt to retransmit the frame, giving it a higher probability of success.

The methods used for collision detection depend on the media being used. On a wired Ethernet bus, it is accomplished by comparing the transmitted data with the data being received off the wire. If it is determined that they differ, the transmitting station on that node recognizes that another node is transmitting at the same time and a collision has occurred. All transmitting nodes then cease transmission and use the calculated back-off interval before attempting to transmit again. The back-off algorithm is a calculation that randomizes the back-off interval for each transmitting node so that the probability of another collision is very low.

POP QUIZ

When a collision occurs on the media, what does the transmitting network node do?

HELPFUL HINT

CSMA/CD is required in a half-duplex network environment. Although the protocol works well if all network node devices remain well behaved, a single "chattering" network node can cause all data flow on a network segment to cease. Of course, this is a malfunction, but it is within the realm of possibility. A quick sniffer trace[15] of that network segment should out the culprit pretty quickly.

With the movement to higher-speed full-duplex Ethernet devices, the need for CSMA/CD is diminishing, although it must be maintained for legacy network segments and devices.

6.4 Ethernet Frame Format

Figure 6-8 illustrates the basic Ethernet frame format.

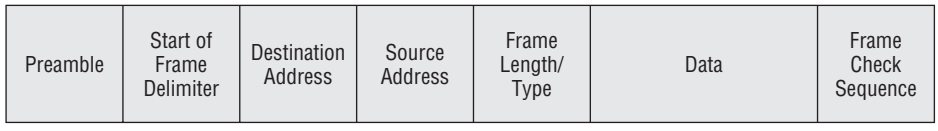
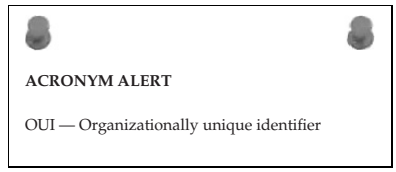


Figure 6-8 The basic Ethernet frame format

[15]Sniffer trace is a technical colloquialism referring to a packet capture. There are dedicated pieces of equipment to capture and display packets or you can load packet-capture software on a laptop. The sniffer trace will permit you to see the traffic that is on a network segment.

The basic frame format illustrated in Figure 6-8 is required for all MAC implementations of the IEEE 802.3 standard. Some additional optional formats also are used to widen the basic capability of the protocol. Following is a list of the basic frame fields:



- **Preamble** — A 7-byte field consisting of alternating 1s and 0s to alert a receiving station that a frame is being received. It is a method used to aid synchronization between the Physical layer receiving circuits and the incoming data stream.
- **Start of Frame Delimiter** — A 1-byte field consisting of a field of alternating 1s and 0s ending with two consecutive 1 bits to signal that the next bit is the leftmost bit in the leftmost byte of the destination address.
- **Destination Address** — A 6-byte field that contains the address of the node that is to receive the frame. The leftmost bit in this field indicates if the frame is destined for a individual node address (0) or a group address (1). The second from the leftmost bit is an indicator if the address is a globally assigned address[16] (0) or a locally administered address[17] (1). The remaining 46 bits of this field contain the address value of the unique node address, a group of network nodes, or all nodes on the network.
- **Source Address** — A 6-byte field that contains the hardware address of the transmitting node, which is always a unique individual address where the leftmost bit of the field is always set to 0.
- **Frame Length/Type** — A 2-byte field that indicates either the number of bytes contained within the Data field of the frame or an alternate frame format type. If the Frame Length/Type has a value of 1500 or less, this value indicates the number of bytes contained within the frame's Data field. If the field value is 1536 or greater, it is used to indicate the

[16] A globally assigned address is the address assigned to the network interface at time of manufacture. These addresses are assigned in blocks to manufacturers and can be used to distinguish which device is from which manufacturer by the hardware address used on that network segment. This can be a valuable troubleshooting tool where large network installations are concerned.

[17] A locally administered address is a MAC address that has been locally assigned by a network administrator. It overrides the default MAC address assigned to the network interface by the manufacturer. Without extreme care, there is a distinct possibility that duplicate addresses could appear on the local network. Duplicate addresses are a big no-no in the networking world. So, if you need to do this, be very careful or you could be in a lot of hot water.

alternate frame type that is being used for either a received or transmitted frame. Table 6-4 lists a handful of the common frame types.

Table 6-4 A Few Common Frame Types

FRAME TYPE	PROTOCOL
0x0800	Internet Protocol Version 4 (IPv4)
0x0806	Address Resolution Protocol (ARP)
0x8035	Reverse Address Resolution Protocol (RARP)
0x809b	AppleTalk
0x80f3	AppleTalk Address Resolution Protocol (AARP)
0x8100	IEEE 802.1Q Tagged Frame
0x8137	Novell IPX
0x86dd	Internet Protocol Version 6 (IPv6)

- **Data** — This field contains the data that is being sent within the frame. It can be any number of bytes of information up to and equaling the maximum number of 1500 bytes that is allowed for this field. However, if the number of bytes is less than 46, a number of bytes must be added to pad the field to reach its minimum length of 46 bytes. The minimum frame size, per the IEEE 802.3 standard, which does not include the preamble, is 64 bytes. Frames of less than 64 bytes are discarded as frames from collisions, faulty NICs, or software-caused under-runs.

RANDOM BONUS DEFINITION

network management — The process of configuring, monitoring, controlling, and administering a network's operation.

- **Frame Check Sequence** — A 4-byte field that contains a 32-bit CRC (cyclical redundancy check) checksum value, which is calculated and inserted by the sending network node and used by the receiving network node to validate the received frame. Both the sending and receiving nodes calculate the CRC value by using the data contained within the Destination Address, Source Address, Frame Length/Type, and Data fields.

POP QUIZ

What is the maximum number of bytes that can be contained in the Data field of an Ethernet frame?

6.4.1 Transmitting a Frame

When a frame request is received by the Media Access Control sublayer from the Logical Link Control sublayer, it is accompanied by the data to be sent and the destination address where the data is to be delivered. The Media Access Control sublayer starts the transmission process by loading the data and address information into the frame buffer. The preamble of alternating ones and zeros, along with the start of frame delimiter, are inserted into their appropriate fields. Destination address and source address information is then added to the fields to which it is assigned. The data bytes received from the Logical Link Control sublayer are counted, and the number of bytes to be contained within the Data field is added to the Frame Length/Type field. The data from the Logical Link Control sublayer is inserted into the Data field, and, if the total number of data bytes is less than 46, a number of pad bytes are added until the number of data bytes is equal to 46. A CRC calculation is performed on the data contained within the Destination Address, Source Address, Frame Length/Type, and Data fields, and then appended to the end of the Data field.

POP QUIZ

What does the Frame Check Sequence field of an Ethernet frame contain?

Once the whole frame is assembled and ready for transmission, the Media Access Control sublayer's next operation depends on whether it is operating in half-duplex mode or full-duplex mode. If it is operating in half-duplex mode, it cannot transmit and receive simultaneously. Since IEEE 802.3 requires that all Ethernet Media Access Control sublayers support half-duplex, if the Media Access Control sublayer is operating in that mode, it is unable to transmit until any incoming frame is completely received. In full-duplex mode, this is not an issue, and the frame can be transmitted immediately.

6.4.1.1 Half-Duplex Transmission

With the development of the CSMA/CD protocol, multiple network nodes are able to share a common media without the need for a centrally located bus arbiter, tokens, or dedicated transmission time slots to determine when they would be allowed to transmit on the media.

NOTE Time division multiplexing (TDM) is a form of digital multiplexing where two or more bit streams are transmitted on a common communications medium. Although it appears as if they are simultaneous, they are actually sharing the time domain. The time domain is divided into a number of fixed time slots. Each data

stream is dedicated to a fixed time slot or channel. Although the same media is being shared, it is not the most efficient use of the media if all or some of the channels are not transmitting. If no data is being streamed on a channel for a particular time slot, it is still using up part of the bandwidth dedicated to it and cannot be used by other channels.

Each portion of the CSMA/CD protocol can be summarized as follows:

- **Carrier Sense** — All network nodes continuously listen on the network media to determine if there are gaps in frame transmission on the media.
- **Multiple Access** — All network nodes are able to transmit anytime they determine that the network media is quiet.
- **Collision Detection** — When two network nodes transmit at the same time, the data streams from both nodes will interfere and a collision occurs. The network nodes involved must be capable of detecting that a collision has occurred while they were attempting to transmit a frame. Upon detecting that the collision has occurred, both nodes cease transmission of the frame and wait a period of time determined by the back-off algorithm before again attempting to transmit the frame.

POP QUIZ

What is the name of the transmission mode that allows either transmitting or receiving at different time intervals but never within the same time interval?

Although bit signals are propagated on a shared network medium at the same rate, the amount of time it takes to transmit a whole frame is inversely proportional to the speed the interface is capable of transmitting it. This means that the time it takes to actually transmit a frame onto the network medium is less. By analyzing this, you can see that a worst-case scenario would be if two network nodes were at two extreme ends of the network media. Electrical signals travel at the same rate, but the amount of time to put a whole frame on the media is much less at higher interface speeds. In order to detect that a collision has taken place, time is needed to travel to the far end of the network segment and back. To allow collision detection to occur within the transmission window of a sending network node, limitations were established for cable lengths and minimum frame length as higher interface speeds were developed. Table 6-5 lists these limitations.

Table 6-5 Half-Duplex Operational Limitations

PARAMETERS	10 MBPS	100 MBPS	1000 MBPS
Minimum frame size	64 bytes	64 bytes	520 bytes
Maximum collision diameter[18] UTP cable	100 meters	100 meters	100 meters
Maximum collision diameter with Repeaters	2500 meters	205 meters	200 meters
Maximum number of repeaters in network path	5	2	1

6.4.1.1.1 Gigabit Ethernet Considerations

Although the Gigabit Ethernet frame is similar to the standard Ethernet frame, it is slightly different in minimum frame length. As you can see in Table 6-5, the minimum frame size expanded from 64 bytes to 520 bytes for a 1000BASE-T frame. The Gigabit Ethernet[19] frame is illustrated in Figure 6-9.

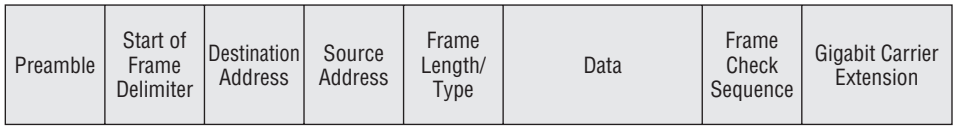


Figure 6-9 The Gigabit Ethernet frame

In order to maintain the same collision domain diameter, the developers opted to increase the minimum frame length to 520 bytes. The longer frame was obtained by adding an extension to the frame after the Frame Check Sequence field. The Carrier Extension field is automatically removed by the receiving network node. The added frame length makes it possible for a frame collision to be detected because of the added time it takes to transmit a minimum-sized gigabit frame

ACRONYM ALERT

SRB — Source routing bridging

[18]Maximum collision diameter refers to the network media length from one transmitting network node to a receiving network node. Worst case is that each node is at the extreme end of a network segment. In wired network media, this equates to cable length and is linear, whereas in a wireless environment, it truly can represent a circle, where the diameter is the maximum distance from transmitter to receiver.

[19]Gigabit per second capability is the capability to pass a billion bits per second on an interface. Remember, a bit is either a single binary 0 or a 1. Whatever the bit value is, there is a lot of stuff coming at you all at once.

onto the network media. The time is close to that of a 64-byte minimum-sized frame being transmitted on the network medium by a 10/100 half-duplex NIC.

The standard for CSMA/CD Gigabit Ethernet added *frame bursting*, the capability of a Gigabit Ethernet NIC's Media Access Control sublayer to transmit a burst of frames without releasing the access to the network media. This is possible since the time needed to place a minimum-sized frame on the network media is much less than the total propagation delay round-trip time of the frame traveling over the network media.

Bursting is accomplished by allowing the transmission of a burst of frames within a time interval slightly greater than that needed for transmitting five maximum-sized frames. The media is kept occupied for the transmitting node by inserting frame carrier extension bits between the frames in the burst. Figure 6-10 illustrates a burst frame sequence.

POP QUIZ

What name is applied to the transmission mode that allows multiple frames to be sent without the need to release the network media between frames?

In Figure 6-10 you will notice that the first frame may have a carrier extension added to it if it does not meet with the minimum frame size of 520 bytes. Between frames or the frame gap periods, the network media is kept busy with a continuous carrier by inserting carrier extension bits. For subsequent frames within a frame burst that do not meet the minimum frame size, a Frame Carrier Extension field is not needed since the frame gaps are being filled with extension bits while in the frame burst transmission mode. Frames will continue to be sent in burst mode until the burst frame limit has been reached. If there is a frame in the process of transmission when the burst frame limit has been reached, the frame is allowed to complete its transmission before the transmitting node releases the network media. Burst frame mode is only supported in Gigabit Ethernet.

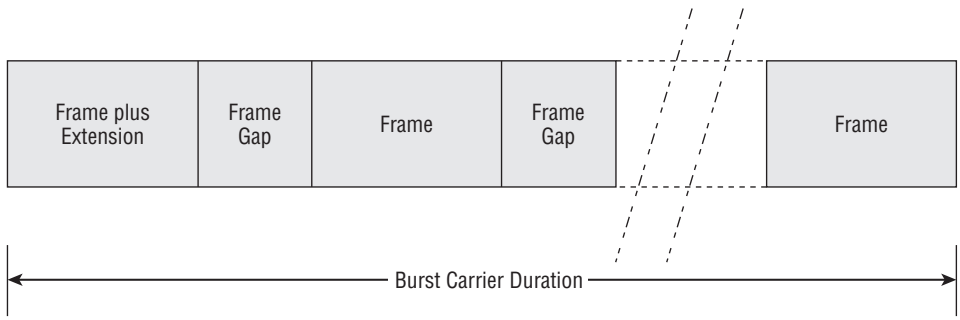


Figure 6-10 The Gigabit Ethernet burst frame sequence

HELPFUL HINT

Since frame burst mode is not supported in 10 Mbps or 100 Mbps Ethernet, it is not a good idea to add these types of network devices to a network segment that is running at gigabit speeds. If you need to mix these devices on the same network segment, you should not use burst mode on that network segment.

6.4.1.2 Full-Duplex Transmission

Full-duplex transmission is the capability of a network node to transmit and receive simultaneously. It is a simpler method of communications than half-duplex since the need for collision detection is eliminated. However, it can only be attained in UTP networks or fiber optic networks, where transmit and receive circuits remain separated. The capability to send and receive at the same time effectually doubles the bandwidth of the network link between network nodes.

The first cabling used for Ethernet networks was coaxial. Because this wired medium was being used for both transmission and reception, the CSMA/CD protocol was developed to permit a sending and receiving network node to communicate over the same cable. Moving from the coaxial wire network media to the UTP cable media, the half-duplicity of the coaxial cable was maintained with the use of hubs that simulated the coaxial cable. So the need to maintain the CSMA/CD protocol was carried forward from the coaxial wire network environment to the UTP cable environment using a half-duplex mode of communications.

Full-duplex is a point-to-point method of communication, where the transmit circuit of one network node is directly connected to the receive circuit of another node, and vice versa. This is fine in a network where two network devices are connected directly to each other, but this is far from the capability to connect many network nodes together over a LAN. If hubs force network nodes into using half-duplex communications, how does one build a multinode network where the devices communicate using a full-duplex communications method?

With the advent of Layer 2 network switches, full-duplex communications are possible on a multinode network. There is a difference between a ''dumb''[20] hub and an ''intelligent'' switch. Hubs are actually considered part of the Physical layer because they are not decision-making devices. They basically provide the interconnectivity on the physical level for network nodes.

[20]Hubs are sometimes called dumb or passive since they do not have any intrinsic intelligence to make a decision on how two nodes are to connect. They are *always* connected in half-duplex mode.

HELPFUL HINT

Do not confuse terms such as *switching hub* or *intelligent hub* with true Layer 2 network switches. What is often being referred to in those terms for a hub is the capability to sense the pins for transmit and receive signals and configure the hub accordingly to accommodate the cable connecting the network node to the hub. Once the hub is configured, it still supports half-duplex communications. To run full-duplex on your local network segment, make sure the device you have selected is a true Layer 2 network switch. Layer 2 switches are more expensive than hubs, so there is a cost consideration.

The name *Layer 2 switch* means exactly what it implies: it is a network device that operates within the first two layers of the OSI reference model. Of course, Layer 1 is the Physical layer, which implies that the construction of the ports of an Ethernet Layer 2 switch is designed with sockets that will accommodate UTP cables terminated with RJ-45 plugs. This physical attribute is no different from that of an Ethernet hub's; they look almost alike but operate very differently. As the name implies, the Layer 2 portion is the Data Link layer of the OSI model, and that is the major difference between a hub and a switch. Hubs do not know or care about the hardware addresses of the devices that are connected to them. In a hub-interconnected network, the endpoint network nodes are responsible for knowing and deciphering the messages on the network media to determine if a frame is addressed to them. The Layer 2 switch uses this very information to electronically interconnect the ports that are connected to it using hardware source and destination addresses. The Layer 2 switch is not concerned with any other aspect of the frame other than being able to direct it to a port that corresponds to the hardware address of the device connected to it. In setting up this connection, the switch is able to maintain the network nodes connected to it to be able to communicate in full-duplex mode.

POP QUIZ

What does the term *full-duplex* mean?

In full-duplex mode, a frame can be transmitted as soon as it is assembled. However, there is a requirement that the gap between successive transmitted frames be long enough for frame synchronization. Each transmitted frame that is transmitted must still adhere to Ethernet framing standards.

6.4.1.2.1 Full-Duplex Flow Control

In the half-duplex mode of operation, a network node does not transmit unless the network medium is silent. It then transmits and while doing so attempts

to detect any network collisions that may have occurred within its transmit interval. Since in full-duplex mode the transmit circuit is separate from the receive circuit, there is no need for collision detection. But how will a transmitting network node know when there is a need for a delay in transmission?

A method of signaling between Media Access Control sublayers was devised to allow a receiving network to signal a transmitting network node that there is network congestion and to cease frame transmission for a period of time. This is referred to as *flow control*. To cause the cessation of frame transmission from a transmitting network node, the receiving network sends a PAUSE frame with a set delay time for the transmitting network node to wait before transmitting the next frame.

If congestion is relieved after a PAUSE frame with a set interval is sent, the receiving network node may transmit another PAUSE frame with the time-to-wait value set to zero. Upon receiving this PAUSE frame, the transmitting network node may begin transmission once again.

PAUSE frames are Media Access Control sublayer frames that have the Frame Length/Type field set to 0x0001 hexadecimal. The destination MAC address that is contained within the transmitted PAUSE frame is set to 01-80-C2-00-00-01. This reserved multicast[21] address is a signal to the receiving switch that the frame is a PAUSE frame for a particular port and will not forward the frame to the other ports that are on the switch. A network node receiving a PAUSE frame will not pass the frame beyond the Media Access Control sublayer.

The time-to-wait interval within a PAUSE frame is contained within a 2-byte unsigned integer with a value between zero and all bits of the 2 bytes set to ones.[22] Each unit of delay is equivalent to 512 bit times. In a 10 Mbps network, the bit time is equivalent to 0.0000001 seconds or a tenth of a microsecond. You can imagine how small these times are by factors of 10 in 100 Mbps and Gigabit Ethernet networks. In

[21]Multicast is the capability to transmit a frame to all network nodes on the network. Upon seeing that the address is set for a multicast broadcast, a node on the network will receive the frame since it was intended to be received by all network nodes on the network.

[22]Two bytes or 16 bits of ones are represented by 1111111111111111 binary, FFFF hexadecimal, or 65,535 decimal. These are all equivalents. However, there will be times in networking or digital circuits where the bit position carries a different connotation than simply a value. Usually these values are represented by a binary bit stream and are more an indication of position or time than just a value.

a 10 Mbps network, the minimum delay would be 51.2 microseconds, which is quicker than you can blink an eye. So you can see that for major congestion, the wait to send delay will have a greater value than the minimum of one.

HELPFUL HINT

Full-duplex and flow control are available for all network speeds of 10 Mbps, 100 Mbps and 1 Gbps. However, on any one particular link between a network node device and a switch, the transmission speed, duplex mode, and flow control all need to match. This is on a link-per-link basis. so it is possible that there can be links of various speeds, duplex, and flow control on differing ports within the same switch. Unless you are certain you know the configuration on a switch, it is not a good idea to swap ports blindly unless you are certain the ports are set identically. If switch ports are set to autonegotiate, they *should* be able to self-configure and settle on the method of communication to be used over the network link.

6.4.1.3 Autonegotiation

Autonegotiation is the capability of a NIC to negotiate the communication parameters that are to be used between it and the port it's connected to. The negotiation between peers only happens on a direct link between the two network nodes. The two devices can have different capabilities but will negotiate upon the duplex and the highest transmission speed the two network interfaces are capable of. Devices of 10 Mbps, 100 Mbps and gigabit speed can be matched on the same network link if needed.

The maximum speed that can be attained on any one network link would be the maximum speed of the slowest network interface. An example of this would be if a 10 Mbs interface set to half-duplex is plugged into a switch port that is set to autonegotiate. Assuming that the switch has the capability to perform at 100 Mbs at full-duplex, it would negotiate the port settings down to 10 Mbs at half-duplex, which is below its rated capability. This allows for flexibility within the network environment where the switch has been placed, but is not really beneficial for network performance. Autonegotiation has its place and at times can be very beneficial, so that network administrators do not have to configure each port every time they want to swap a port.

Another example would be if one end of a network link has a 100 Mbps network interface and the other end has a gigabit interface connected to it. If both interfaces were set to autonegotiate, they would ideally settle

RANDOM BONUS DEFINITION

Physical layer — The lowest layer of the seven-layer OSI model.

upon 100 Mbps at full duplex. However, this is assuming that the two network node devices play nice and can settle on that speed and duplex. Depending on manufacturer and the network interface being used, a link may need to be set permanently to a speed and duplex due to the inability of the two devices to negotiate a speed and duplex that works for both of them.

There may be instances where both interfaces do negotiate a speed but for some reason one interface settles upon half-duplex while the other settles upon full-duplex. On the surface everything may appear to be working as planned. However, performance over the link may be affected and communications seem slow. Mismatch in duplex is not uncommon and at times goes unnoticed until major network degradation is noted.

It is possible when two network node ports are interconnected that it appears that one network interface may have failed. The two devices will not bring up the link. There are a couple of ways to attack this problem. One is to hard-set both ends to a speed and duplex that you know they are capable of and see if you can send data across the link. The other method is to have a third network node device that you know is reliable connect to each to see if the link will come up with either device connected. This test is not conclusive, but if both devices can link with the known device, the culprit may be that autonegotiation between the two network interfaces is not working.

There is a possibility that two network node interfaces may appear to autonegotiate properly and can operate for an extended period of time without any problems. Then it is noticed that some network performance problems have arisen. Traffic over a particular link seems to degrade, comes back, and then degrades again. This can be an indication that the autonegotiation between the two network node interfaces may be flapping.[23] If these network ports are set to autonegotiate, it would be best to manually configure them for the highest common speed and duplex and then monitor the link to see if performance picks up. If not, it can be an indication of bad cabling or possibly one network node interface may be having problems.

POP QUIZ

What is autonegotiation?

[23]Flapping (or flopping or flipping) generally describes an unstable network interface link. This is perhaps an offshoot of the old digital design days when flip-flops were used to maintain a particular state. Flip-flopping has wiggled its way into our society to mean something that is either indecisive or changes state whimsically. A good example of this would be today's politicians.

HELPFUL HINT

Some devices indicate link status and/or speed, but few indicate whether the link is running half- or full-duplex. You may want to become familiar with the network devices being used in that network segment. This will allow you to use monitoring tools to determine if speed and duplex for the link are set properly for the two network node interfaces that are connected on the link. Many network node devices do provide software tools for monitoring and measuring performance of the ports on the device. These software tools are usually a part of the software suite that came with the network device and can be used not only for configuration but also for troubleshooting.

6.4.2 Receiving a Frame

The receiving of a frame is the same no matter what type of network interface is in use. The electrical signals are received from the network media and loaded into a frame receive buffer. The major difference is between half-duplex network interfaces and full-duplex interfaces. A network interface that is strictly a half-duplex interface can use the same frame buffer for both transmitting and receiving a frame. However, full-duplex interfaces need to be capable of both transmitting and receiving at the same time, so a receive frame buffer is needed as well as a transmit frame buffer.

When a frame is received by a network interface, it is loaded into the receive frame buffer and the destination address is compared to see if it matches the unique MAC address of the network interface or network group address or if the frame is a broadcast frame. If there is an address match, the frame length is checked along with the Frame Check Sequence field. The Frame Check Sequence field is checked against the checksum, which was calculated as the frame was received from the Physical layer. If this matches, the Frame Length/Type field is checked to determine the frame type of the frame that was received so it can be properly be parsed and passed to the appropriate upper layer.

Once the frame has been unloaded from the receive buffer and passed up the ISO reference model to the upper layers, the network interface is then ready to receive another frame from the Physical layer. If a frame does not pass the proper framing criteria, it is discarded and the interface is readied to receive the next network frame.

POP QUIZ

When a frame is received, what is the first criteria that is checked?

6.5 Traffic Optimization

What exactly is traffic optimization? It connotes a lot of various things, but the gist of the term is overall improvement in network performance. In the earlier sections of this chapter, we discussed speed and duplex and how they can affect the performance on a particular network link. We can see that there are advantages of having certain network paths being faster than others. Links going between devices that aggregate numerous network nodes need to be faster and more reliable than those of a single workstation to a hub or network switch. Figure 6-11 illustrates a network consisting of many user network nodes interconnected with high-speed switches that have high-speed gigabit interfaces between them.

The high-speed switches in this figure are to aggregate the multiple workstations and allow them to stream network data unimpeded by congestion caused if the data links between the switches were of the same speed as those between the workstation and the switches. In this example, the workstations are connected to the switches using a 100 Mbps full-duplex link. The switches are interconnected with high-speed gigabit full-duplex links and provide a redundant path if needed.

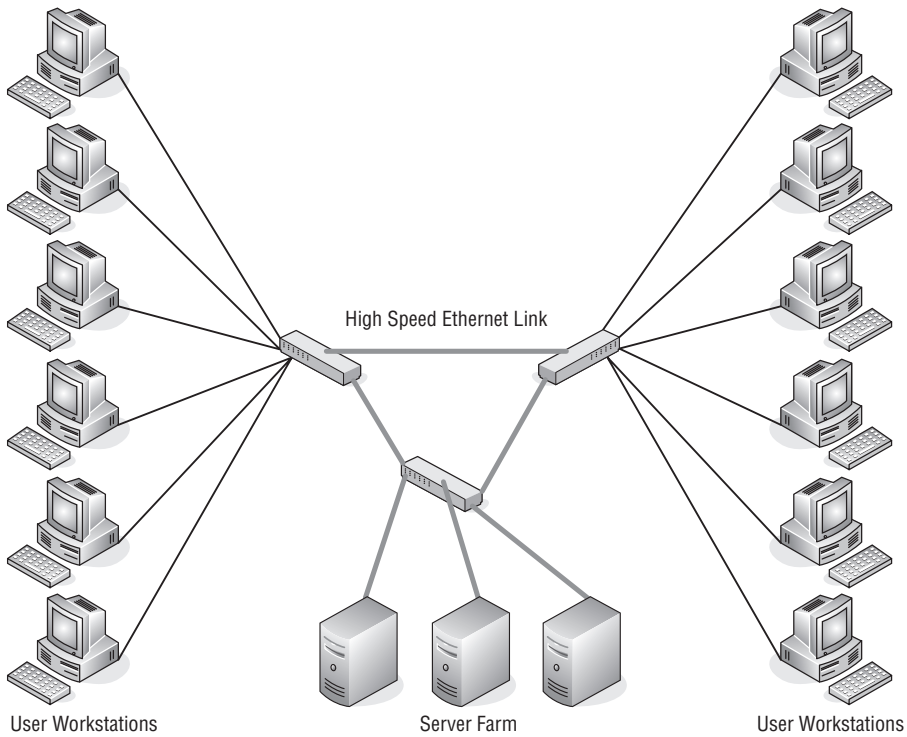
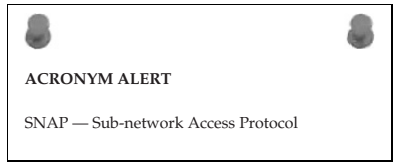


Figure 6-11 A network segment with high-speed links



The redundant[24] path shown in this figure allows for any of the high-speed links to go down and still have workstations on both network switches to which they are connected be able to access the server farm. These servers can provide various services such as e-mail, mass data storage, and client/server applications The servers are interconnected over a high-speed data link with a gigabit NIC to eliminate congestion on any one server. This increases the likelihood that there would be less congestion on these data links but does not totally eliminate the possibility that congestion could occur.

When administering large network installations it is important to understand the traffic patterns that are present on the network. Network efficiency can be increased where needed. The idea is to balance the need versus what it will cost since there can be areas of overkill where the investment in network resources is underutilized and thus is not a wise decision. Careful planning can greatly aid in determining where more network resources are required and limit the amount of waste of underutilized network segments. Know the business environment in which the network you are administering is installed. A carefully thought-out network is easier to install, maintain, and troubleshoot and runs efficiently with higher reliability.

6.5.1 Traffic Shaping

In the previous section, we discussed planning where high-speed links would be required. This approach is best-effort, and there is no differentiation of the type of traffic or if it is more important traffic than that of another transmitting network node. With real-time applications such as Voice over IP and videoconferencing, there is a need to give priority to these frames so they can be delivered in a timely fashion.

What if there was a way to tag a frame so it would be given a priority over another frame that need not be delivered as quickly? If frames are marked, they can be queued so the frames with priority will be forwarded on to the next segment. A simplified diagram illustrates this in Figure 6-12.

[24]Redundant path or redundancy in a network is the capability to provide multiple paths to various network resources to add fault tolerance. If one or more high-speed links go down, the network will either be unaffected or, at worst, be partially affected. It may not be able to have all of the network resources available to all of network users, but there will be areas of unimpeded network operation.

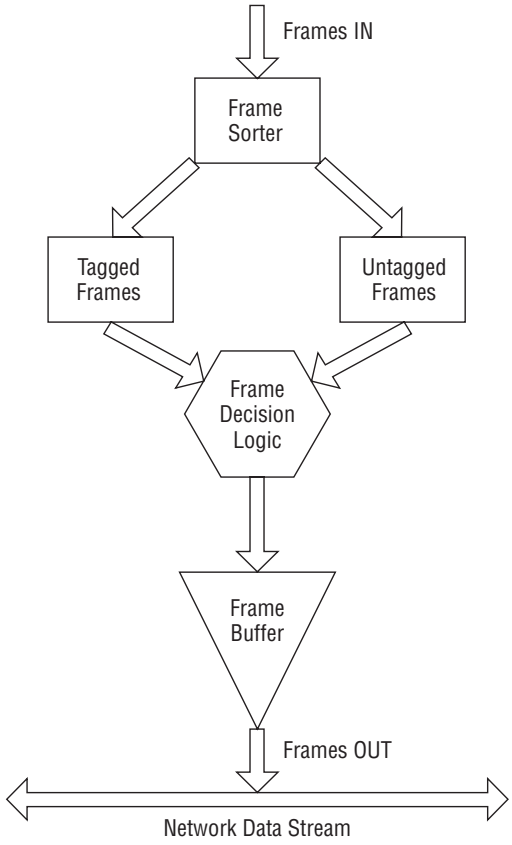


Figure 6-12 Frame prioritization

Frames are tagged to identify them as frames that should be transmitted over the network with priority. As frames enter into a network node that is to transmit tagged frames with priority, they are checked for a priority tag. A queuing system is used to keep both tagged and untagged frames in the same order as they are received. When the network node device is ready to transmit the next frame, a check is made by frame decision logic to see if there are any tagged frames to be sent with priority. If there are tagged frames, they will continue to be transmitted until there are no remaining tagged frames that need to be transmitted. When the tagged frames bin is empty, untagged frames will be transmitted until the next tagged frame arrives in the tagged frame bin. All frames are sent in the order they are received, with the tagged frames being transmitted before any untagged frames.

ACRONYM ALERT

PDU — Protocol data unit

POP QUIZ

How is a frame given priority?

We have discussed the Layer 2 switch, but tagging requires a higher level than that. Routers are capable of operating at Layer 3 and can make decisions on tagged packets. However, there is a more recent development in the switching area — the Layer 3 switch (sometimes called the *routing switch*). Routing switches perform many of the same functions as routers, except they operate much faster. Conventional routers depend on software for the routing protocols and decision making. Routing switches implement the routing decision process in hardware, allowing higher throughput of frames. These network devices may be faster than routers as far as forwarding frames, but they are not as flexible or as programmable as a conventional router.

6.5.1.1 VLAN Tagging

VLAN[25] tagging was standardized in IEEE 802.1Q. The standard allows for 4 bytes used for tagging purposes to be inserted between the Source MAC Address and the Frame Length/Type fields. Any modification of a frame will destroy the Frame Check Sequence checksum, so after the frame is assembled with the 802.1Q tagging the checksum is recalculated and placed in the Frame Check Sequence field. Figure 6-13 illustrates the 802.1Q VLAN header.

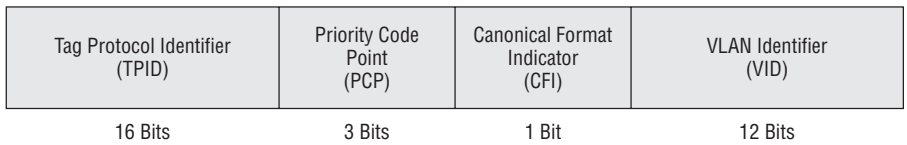


Figure 6-13 The IEEE 802.1Q VLAN header

- **TPID** — The Tag Protocol Identifier is a 16-bit field containing the hexadecimal value of 0x8100 as an indicator that the frame is an 802.1Q tagged frame.
- **PCP** — The Priority Code Point is a 3-bit field[26] that contains a value from 0 to 7 and is used to indicate the priority level of the frame. Zero is the lowest priority and 7 is the highest.

[25]VLAN is an acronym for virtual local area network. Normally, a LAN is localized within a network segment. However, in a switched network environment, the member network nodes of a VLAN do not need to be located within the same local vicinity. They are identified as a group belonging to a particular VLAN.

[26]The maximum value of 3 binary bits is 7: 111(binary) = 7 (decimal). The binary value positions are 4+2+1, which equals 7. This little exercise is for those readers who may find themselves ''base-2 challenged.''

- **CFI** — The Canonical Format Indicator is a 1-bit field when set to the value 0 to indicate that the MAC address is in canonical format, which is always set to 0 for Ethernet switches. If a frame is received with the CFI set to the value 1, it should not be bridged to an untagged port.

RANDOM BONUS DEFINITION

1BASE5 — A baseband Ethernet system operating at 1 Mbps over one pair of UTP cable. Also known as *StarLAN*.

- **VID** — The VLAN Identifier is a 12-bit field that specifies which VLAN the incoming frame belongs to. If this field is set to the value of 0, it indicates that the frame does not belong to a VLAN and that it is only a priority tag.

The advantage of having network node devices that are part of a VLAN group equipped with VLAN tagging is primarily the capability to tag outgoing frames with a priority. This means that frames that require timely delivery are expedited over the network before less critical or best delivery frames. Another advantage is that network node devices can be grouped and are allowed to communicate across multiple LAN networks as if they were all on a single LAN network. The destination address is filtered by the switches and bridges in the network path and only forwards the frames to the ports that service the VLAN the frame belongs to. Because of the configurability of these switches, network management is made simpler, allowing for easy addition, removal, movement, or other configuration changes required on a VLAN port.

POP QUIZ

What does the acronym VLAN stand for?

HELPFUL HINT

Layer 3 (or routing) switches seem so easy to manage and configure. We will again caution about the need for documenting your network well, unless you prefer to go through a multitude of switch configurations, port by port. It is even more imperative because of configurations where ports can be moved and juggled without physically going out and moving a cable on a port. Switch networking issues can be daunting on a large network, so there is no substitute for good network documentation.

(continued)

HELPFUL HINT *(continued)*

If you need to call for support on a problem, remember that the support engineer does not have a crystal ball[27] to look into your network. He is going to rely on your ability to know your network and know it well. Support engineers do not like playing guessing games. It is a waste of their time and will add to your frustration level as your boss blows his hot breath on the back of your neck.

Want to be a good network administrator? Document, document, document!

6.6 Chapter Exercises

1. What does the acronym CSMA/CD stand for?
2. What form of communications eliminates the need for collision detection?
3. When you choose not to configure an Ethernet port for speed and duplex mode, what are you relying on?
4. What is needed when setting up VLAN networking?
5. What is a source address? What is a destination address?
6. What is the maximum number of bytes the Data field can contain in an Ethernet frame? What is the minimum number of data bytes?

6.7 Pop Quiz Answers

1. What was the first type of cable used to form an Ethernet network?

   Coaxial cable

2. An Ethernet network device that forwards data on the network would be considered what type of Ethernet device?

   DCE (data communications equipment)

[27] A crystal ball is a device a network administrator hopes the support engineer at the other end of the hotline has when he frantically calls for support. Alas, he does not possess one, so drop to your knees and start praying. Or you can take the easy way out and start documenting your network from initial installation through configuration changes, additions, and anything that modifies the network.

3. If a cable is wired such that one plug is a T568A and the other is a T568B, it would commonly be referred to as _____ cable.

   Crossover

4. You are interconnecting two Ethernet devices, but neither device is showing a link light on the assigned port. List in order of likelihood where the problem might be.

   - Cable type
   - Defective cable
   - Bad network interface

5. Into which two sublayers of the IEEE 802 reference model is the OSI reference model Data Link layer divided?

   - LLC (Logical Link Control)
   - MAC (Media Access Control)

6. With which functions is the Logical Link Control sublayer mainly concerned?

   - Flow control
   - Error control
   - Multiplexing protocols

7. When a collision occurs on the media, what does the transmitting network node do?

   Stops transmitting

8. What is the maximum number of bytes that can be contained in the Data field of an Ethernet frame?

   1500 bytes

9. What does the Frame Check Sequence field of an Ethernet frame contain?

   CRC calculation using the bytes of the Destination Address, Source Address, Frame Length/Type, and Data fields.

10. What is the name of the transmission mode that allows either transmitting or receiving at different time intervals but never within the same time interval?

    Half-duplex

11. What name is applied to the transmission mode that allows multiple frames to be sent without the need to release the network media between frames?

    Burst mode

12. What does the term *full-duplex* mean?

    The capability to transmit and receive at the same time.

13. What is flow control used for?

    To stop a transmitting node from sending when congestion is detected.

14. What is autonegotiation?

    The capability of two network node peers to negotiate the speed and duplex used on the link they are connected to.

15. When a frame is received, what is the first criteria that is checked?

    Destination address

16. What is the first step you should perform before implementing a network?

    Carefully plan out the network.

17. How is a frame given priority?

    Tagging

18. What does the acronym VLAN stand for?

    Virtual local area network