7.46 With the aid of Figure 7.36(a), explain the role of the following protocols:

(i) signaling data link,
(ii) signaling link,
(iii) signaling network.

Hence, with the aid of the message format shown in Figure 7.36(b), describe the role of the following fields:

(i) FSN and BSN,
(ii) SID and SIF.

7.47 With the aid of Figures 7.26 and 7.37, produce a diagram that shows the exchange of the SS7 network signaling messages to set up a standard telephone channel/connection initiated using analog access signaling.

Enterprise networks

8.1 Introduction

When a person is at home, all calls relating to interpersonal and interactive applications must be made using a PSTN, an ISDN, or a cable distribution network. The calls are then charged at a rate determined by the call duration and the distance involved. In the case of a person in a business or large enterprise, however, the majority of the calls made are to other members of the same business/enterprise and only a small percentage are to people outside of the enterprise. Hence for all but the smallest businesses, in order to reduce call charges, most enterprises install their own private networks to handle those calls that are internal to the enterprise. Normally, the network comprises both a private branch exchange (PBX) and a local area network (LAN) and, collectively, these support all the interpersonal and interactive communications within the enterprise.

For an enterprise that occupies just a single site/establishment – for example, a small to medium-sized company, a hospital complex, a university campus – the PBX and LAN at the establishment handle all internal calls and only those calls that are external to the site are made using an appropriate public network such as a PSTN, an ISDN, or an Internet service provider

(ISP) network. For an enterprise that operates over multiple sites, however, when there is a significant proportion of intersite calls, an alternative solution is to extend the private facilities associated with each site to embrace all sites. This involves linking the sites together using high bit rate transmission lines that are leased from a (national) network provider. The resulting network is then known as a multisite **enterprise network** or, if the sites are located in different countries around the world, a **global enterprise network.**

Normally, leased lines are charged for on an annual rather than a per-call basis and hence this approach is only justified when the annual (public) call charges to other sites exceeds the cost of leasing the lines. An added benefit of creating a private network, however, is that it is often easier to offer more sophisticated services since the utilization of the bandwidth of the intersite leased lines is under the control of the enterprise network manager. Also, because the total network is private – apart from the transmission lines, of course – a private network is considered to be more secure than a public one.

As we explained in the last chapter, a PBX operates in a similar way to a local exchange/end office in a public network with the terminal equipment connected to the PBX using either analog or digital lines. In this chapter, we shall focus on the operation of the different types of LAN and the various approaches that are used to create multisite enterprise networks.

8.2 LANs

As we have just indicated, LANs are used to interconnect distributed communities of end systems – often referred to as **stations** in the context of LANs – including multimedia PCs, workstations, servers, and so on. Typically, these are distributed around an office, a single building, or a localized group of buildings, all of which belong to a single enterprise.

The early LANs – many of which are still in existence – operate using a shared, high bit rate, transmission medium to which all the stations are attached and the information frames relating to all calls are transmitted. To ensure the transmission bandwidth is shared fairly between all of the attached stations, a number of different medium access control (MAC) methods are used. These include **carrier-sense multiple-access with collision detection** (**CSMA/CD**) and **token ring**, both of which have a defined maximum number of attached stations and length of transmission medium associated with them. As we shall see, in practice the maximum distance is relatively small and hence most LANs of this type comprise multiple (LAN) **segments** that are interconnected together using either **repeaters** or devices known as **bridges** and a high bit rate (site-wide) **backbone subnetwork.** We explain the operation of Ethernet/IEEE802.3 LANs – which are based on the CSMA/CD MAC method – in Section 8.3 and token ring LANs in Section 8.4. The operation of bridges is explained in Section 8.5 and, as an example of a backbone network, the **fiber distributed interface** (**FDDI**) in Section 8.6.

More recently, higher bit rate versions of the older LAN types – now known as **legacy LANs** – have become available. To obtain the higher network throughputs that are required with multimedia applications, the central **hubs** associated with the earlier LANs have been upgraded to operate at much higher bit rates. Also, as we shall explain, the older hubs operate in a half-duplex mode and support only a single frame transfer at a time. Hence the newer hubs operate in a duplex mode and allow the frames relating to multiple calls to be transmitted concurrently. Examples include **fast Ethernet** hubs and **Ethernet switching** hubs, both of which we describe in Section 8.7.

In terms of the link layer protocol associated with LANs, the various LAN types all use a standard LC sublayer and there is a different MAC sublayer for each of the LAN types. We describe the structure and the user services offered by each sublayer in Section 8.8.

In multisite enterprise networks, the LANs associated with the different sites are interconnected together using various methods determined by the volume of intersite traffic involved. If this is relatively low, then switched ISDN connections can be used; if it is high, then high bit rate (digital) leased lines are used. Normally, these are leased from the operator of a national public circuit-switched network and are the same as those we described in Section 7.2.3. In both cases, however, a **gateway** is connected to the LAN at each site and this manages all intersite frame transfers. We explain the operation of some of the different types of technology that are used to interconnect the LANs at different sites in Section 8.9.

8.3 Ethernet/IEEE802.3

Ethernet networks – and the more recent derivative IEEE802.3 – are used extensively in technical and office environments. As we shall see, Ethernet has gone through many phases of development since its first introduction but, in general, the same basic mode of operation is still used. All frame transmissions between all the stations that are attached to the LAN take place over a shared 10 Mbps bus and the CSMA/CD MAC method is then used to share the use of the bus in an equitable way. We shall explain the principle of operation of this type of MAC method first and then other selected aspects in the following subsections.

CSMA/CD

Since all the stations are attached directly to the same cable/bus, it is said to operate in a **multiple access (MA) mode**. To transmit a block of data, the source station first encapsulates the data in a frame with the address of the destination station and its own address in the frame header and an FCS field at the tail of the frame. The bus operates in the **broadcast mode** which means that every frame transmitted is received by all the other stations that are attached to the bus. Hence as each of the other stations receives the frame, it

first checks the frame is free of errors using the FCS and, if it is, it compares the destination address in the header with its own address. If they are different, the station simply discards the frame; if they are the same, the frame contents are passed up to the LC sublayer for processing together with the address of the source station.

With this mode of operation, two (or more) stations may attempt to transmit a frame over the bus at the same time. Because of the broadcast mode, this will result in the contents of the two (or more) frames being corrupted and a **collision** is said to have occurred. Hence in order to reduce the possibility of a collision, prior to sending a frame the source station first determines whether a signal/frame is currently being transmitted on the bus. If a signal – known as the **carrier** – is **sensed** (**CS**), the station defers its own transmission until the current frame transmission is complete and only then does it attempt to send its own frame. Even so, in the event of two (or more) stations waiting to send a frame, both will start to transmit their frame simultaneously on detecting that the transmission of the current frame is complete. When this happens, however, it is necessary for the two (or more) stations involved, to detect a collision has occurred before each has finished transmitting its own frame. In practice, because of the possibly large signal propagation delay of the bus and the high transmission bit rate used (10 Mbps), this is not as straightforward as it might seem.

A station detects that a collision has occurred by simultaneously monitoring the signal that is present on the cable all the time it is transmitting a frame. Then, if the transmitted and monitored signals are different, a collision is assumed to have occurred – **collision detected** (**CD**). As we show in Figure 8.1, however, a station can experience a collision not just at the start of a frame but after it has transmitted a number of bits. The worst-case time delay – and hence maximum number of bits that have been transmitted – before detecting that a collision has taken place is known as the **collision window** and occurs when the two colliding stations are attached to opposite extremities of the bus, as we show in the figure.

In the figure, station *A* has determined that no transmission is in progress and hence starts to transmit a frame – part (i). As we explained in Section 6.2.8, irrespective of the bit rate being used, the first bit of the frame will take a small but finite time to propagate over the transmission medium determined by the length of the cable, l, and the signal propagation velocity, v. The maximum length of cable is set at 2.5 km. Hence, assuming a v of 2×10^8 ms^{-1}, the worst-case signal propagation delay time, T_p, going from one end of the cable to the other, is given by:

$$T_p = l/v = 2.5 \times 10^3 / 2 \times 10^8 = 12.5 \text{ microseconds}$$

Now assume that, just prior to the first bit of the frame arriving at its interface, station *B* determines the transmission medium is free and starts to transmit a frame – part (ii).

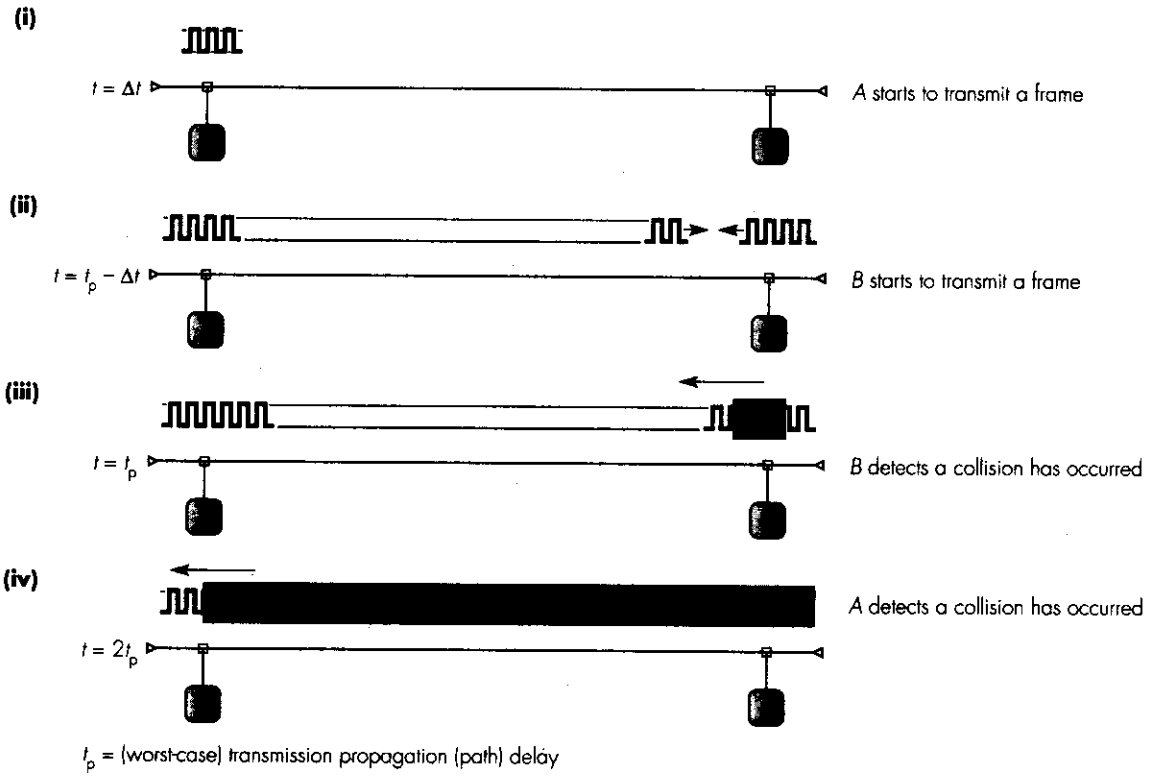


Figure 8.1 CSMA/CD worst-case collision detection.

As we show, after *B* has transmitted just a few bits, the two signals collide – part (iii) – and the collision signal then continues to propagate back to station *A* – part (iv). Hence the worst-case time before station *A* detects that a collision has occurred, $2T_p$, is 25 microseconds. In addition, as we shall expand upon later, in order to transmit the signal over this length of cable, the cable is made up of five 500 m *segments*, all interconnected together by means of four devices called **repeaters**. Each repeater introduces a delay of a few microseconds in order to synchronize to each new frame. Hence the total worst-case time is set at 50 microseconds or, assuming a bit rate of 10 Mbps, after *A* has transmitted:

$$10 \times 10^6 \times 50 \times 10^{-6} = 500 \text{ bits}$$

A safety margin of 12 bits is then added to this and the minimum frame size is set at 512 bits or 64 bytes/octets. This is called the **slot time** (in bits) and ensures that station *A* will have detected a collision before it has transmitted its smallest frame. Also, to ensure that the collision signal persists for

sufficient time for it to be detected by *A*, on detecting the collision, *B* continues to send a random bit pattern for a short period. This is known as the **jam sequence** and is equal to 32 bits.

After detecting a collision, the two (or more) stations involved then wait for a further random time interval before trying to retransmit their corrupted frames. As we shall explain later, the maximum frame size including a four-byte CRC is set at 1518 bytes and hence a collision will occur if two (or more) stations create a frame to send during the time another station is currently transmitting a maximum sized frame. This is equal to a time interval of:

$$1518 \times 8/10 \times 10^6 = 1.2144 \text{ milliseconds}$$

Clearly, the probability of this occurring increases with the level of traffic (number of frames) being generated and the maximum throughput of the LAN occurs when this limit is reached. Hence if a second collision should occur when a station is trying to send a frame, this is taken as a sign that the cable is currently overloaded. To avoid further loading the cable, therefore, the time interval between trying to retransmit a frame is increased exponentially after each new attempt is made using a process known as **truncated binary exponential backoff**. The actual time is a function of the slot time. The number of slot times before the *N*th retransmission attempt is chosen as a uniformly distributed random integer R in the range $0 \leq R < 2^K$, where $K = \min$ (N, backoff limit). In the standard, the **backoff limit** is set to 10.

Wiring configurations

There are a number of different types of cable that can be used with Ethernet. These include:

- 10Base2: thin-wire (0.25 inch diameter) coaxial cable with a maximum segment length of 200 m;
- 10Base5: thick-wire (0.5 inch diameter) coaxial cable with a maximum segment length of 500 m.
- 10BaseT: hub (star) topology with twisted-pair drop cables of up to 100 m;
- 10BaseF: hub (star) topology with optical fiber drop cables of up to 1.5 km.

Although different types of cable are used, they all operate using the same CSMA/CD MAC method.

At the time the first Ethernet installations were carried out, the only transmission medium available that could operate at 10 Mbps was coaxial cable. Initially, thick-wire coaxial cable was installed since this can be used in relatively long lengths of up to 500 m before the transmitted/broadcast signal needs to be repeated. As we explained in Section 6.3.1, this involves the attenuated signal received at the extremity of the cable segment being amplified and restored to its original form before it is retransmitted out onto the next

cable segment. Up to five cable segments – and hence four repeaters – can be used in this way. Hence the maximum length of cable the signal propagates is 2.5 km plus 4 repeaters which is the origin of the slot-time figure used in the standard.

The disadvantage of thick-wire coax is that it is relatively difficult to bend and hence install. To overcome this, thin-wire coax was used but, because of the increased (electrical) resistance associated with it, the maximum length of cable for each segment is reduced to 200 m.

More recently, as we explained in Section 6.2.2, with the arrival of inexpensive adaptive crosstalk canceler circuits to overcome near-end crosstalk (NEXT), it is possible to obtain bit rates of tens of Mbps over twisted-pair cable of up to 100 m in length. Also, it was found that, in a vast majority of offices, the maximum length of cable used for telephony to reach each desktop from the wiring closet was less than 100 m. Hence unshielded twisted-pair (UTP) cable – as used for telephony – has rapidly become the standard for use with Ethernet. The configuration used for each segment is shown in Figure 8.2(a).

Since the cable forms a physical bus, both thick and thin wire coaxial cable installations involve the cable passing near to each attached station. As we can see in the figure, however, with twisted-pair cable a star configuration is used with the hub located in the wiring closet and each station connected to it by means of twisted-pair drop cables. Normally, category three (CAT3) UTP cable is used as for telephony. Each cable contains four separate twisted-pairs. In the case of Ethernet, just two pairs are used: one pair for transmissions from the station to the hub and the second pair for transmissions in the reverse direction.

To emulate the broadcast mode of working associated with CSMA/CD, as we show in Figure 8.2(b), the repeater electronics within the hub repeats and broadcasts out the signal received from each of the input pairs onto all of the other output pairs. Hence the signal output by any of the stations is received by all the other stations and, as a result, the carrier sense function simply involves the MAC unit within each station determining whether a signal is currently being received on its input pair. Similarly, the collision detection function involves the station determining if a signal arrives on its input pair while it is transmitting a frame on the output pair.

Because of their mode of operation, this type of hub is called a **repeater hub** and typical numbers of attached stations – and hence sockets – are from 8 through to 16. Above this number multiple hubs are stacked together and are connected by repeaters or, as we shall explain in Section 8.5, bridges. In the case of repeaters, the maximum length of cable between any two stations – including the 100 m drop cables – must not exceed 1.5 km. To achieve this coverage/distance, however, normally it is necessary to use a central hub to which each twisted-pair hub is connected by means of optical fiber cables.

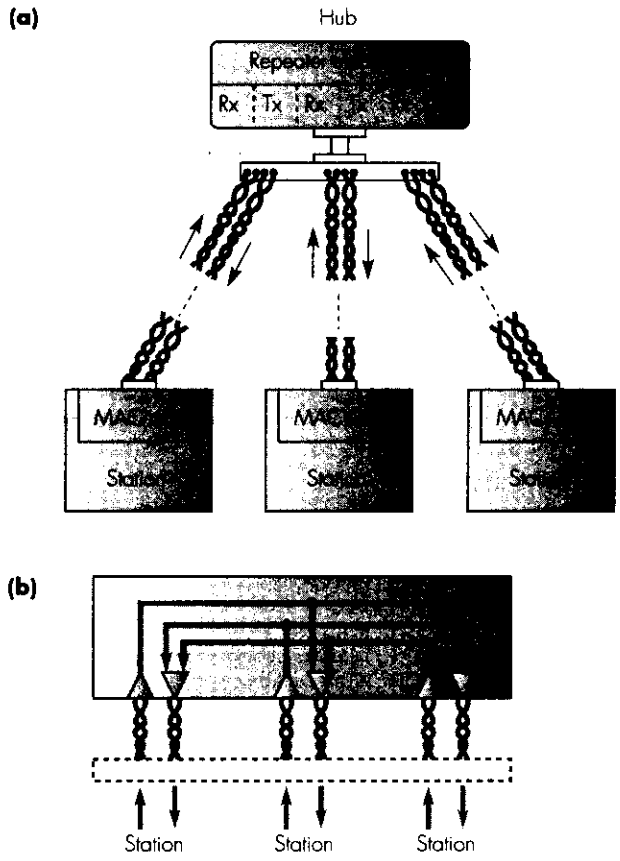
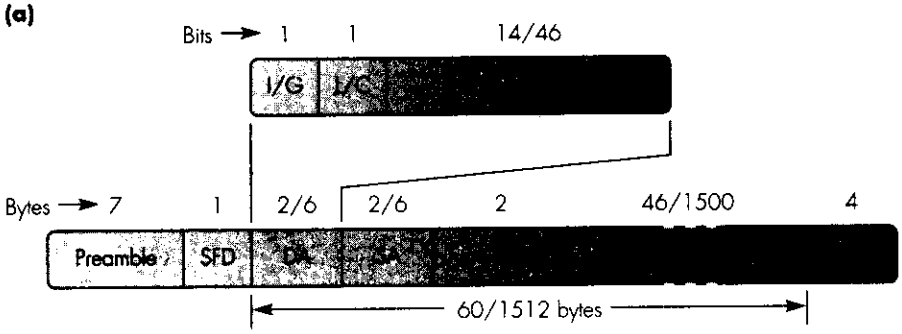


Figure 8.2 Hub configuration principles: (a) topology; (b) repeater schematic.

Frame format and operational parameters

The format of a frame and the operational parameters of a CSMA/CD network are shown in Figure 8.3. The *preamble* field is sent at the head of all frames. Its function is to allow the receiving electronics in each MAC unit and repeater to achieve bit synchronization before the actual frame contents are received. The preamble pattern is a sequence of seven bytes, each equal to the binary pattern 10101010. All frames are transmitted on the cable using Manchester encoding. Hence, as we explained in Section 6.5.1, the preamble results in a periodic waveform being received by the receiver electronics in each DTE which acts as a reference clock. The *start-of-frame delimiter* (*SFD*) is the single byte 10101011 which immediately follows the preamble and signals the start of a valid frame to the receiver.

(a)



SFD = start-of-frame delimiter
DA/SA = source/destination address
I/G = individual (=0)/group (=1) address
FCS = frame check sequence
L/C = locally administered (=1)/ centrally administered (=0)

(b)

Bit rate	10 Mbps (Manchester encoded)
Slot time	512 bit times
Interframe gap	9.6 microseconds
Attempt limit	16
Back off limit	10
Jam size	32 bits
Maximum frame size (including FCS)	1518 bytes
Minimum frame size (including FCS)	512 bits

Figure 8.3 Ethernet/IEEE802.3 characteristics: (a) frame format; (b) operational parameters.

The *destination* and *source addresses* – also known as **MAC addresses** because they are used by the MAC sublayer – specify the identity of the hardware interface of both the intended destination station(s) and the originating station, respectively. Each address field can be either 16 or 48 bits, but for any particular LAN installation the size must be the same for all stations. The first bit in the destination address field specifies whether the address is an **individual address** (=0) or a **group address** (=1). If an individual address is specified, the transmitted frame is intended for a single destination. If a group address is specified, the frame is intended either for a logically related group of stations (group address) or for all other stations connected to the network (**broadcast** or **global address**). In the latter case, the address field is set to all binary 1s and, for a group address, the address specifies a previously agreed group of stations. The type of grouping is specified in the second bit and can be locally administered (=1) or centrally administered (=0). Group addresses are used for multicasting and the MAC unit/circuit associated with each station in the multicast group is then programmed to read all frames with this group address at its head.

With the original Ethernet standard, the two-byte *type* field immediately follows the address fields and indicates the network layer protocol that created the information in the data field. With the more recent IEEE802.3 format, the next two bytes are used as a *length indicator* which indicates the number of bytes in the data field. If this value is less than the minimum number required for a valid frame (minimum frame size), a sequence of bytes is added, known as **padding**. The maximum size of the data field – normally referred to as the **maximum transmission unit** (**MTU**) – is 1500 bytes. Finally, the *frame check sequence* (*FCS*) field contains a four-byte (32-bit) CRC value that is used for error detection. Note that with the original Ethernet standard, the end of a frame is detected when signal transitions end.

Frame transmission and reception

The frame transmission sequence is summariized in Figure 8.4(a). When a frame is to be transmitted, the frame contents are first encapsulated by the MAC unit into the format shown in Figure 8.3(a). To avoid contention with other transmissions on the medium, the MAC unit first monitors the carrier sense signal and, if necessary, defers to any passing frame. After a short additional delay (known as the **interframe gap**) to allow the passing frame to be received and processed by the addressed station(s), transmission of the frame is initiated.

As the bitstream is transmitted, the transmitter simultaneously monitors the received signal to detect whether a collision has occurred. Assuming a collision has not been detected, the complete frame is transmitted and, after the FCS field has been sent, the MAC unit awaits the arrival of a new frame, either from the cable or from the LC sublayer within the station. If a collision is detected, the transmitter immediately turns on the collision detect signal and enforces the collision by transmitting the jam sequence to ensure that the collision is detected by all other stations involved in the collision. After the jam sequence has been sent, the MAC unit terminates the transmission of the frame and schedules a retransmission attempt after a short randomly-computed interval.

Figure 8.4(b) summarizes the frame reception sequence. The MAC unit first detects the presence of an incoming signal and switches on the carrier sense signal to inhibit any new transmissions from this station. The incoming preamble is used to achieve bit synchronization and, when the start-of-frame delimiter has been detected, with an IEEE802.3 LAN, the length indicator is read and used to determine the number of bytes that follow. The frame contents including the destination and source addresses are then received and loaded into a frame buffer to await further processing. The received FCS field is first compared with the computed FCS and, if they are equal, the frame content is further checked to ensure it contains an integral number of bytes and that it is neither too short nor too long. If any of these checks fail then the frame is discarded. If all checks pass, then the destination address is read from the head of the frame and, if the frame is intended for this station – that is, the address of the station is the same as that in the frame or, if it is a group address, the station is a member of the specified group – the frame contents are passed to the LC sublayer for processing.

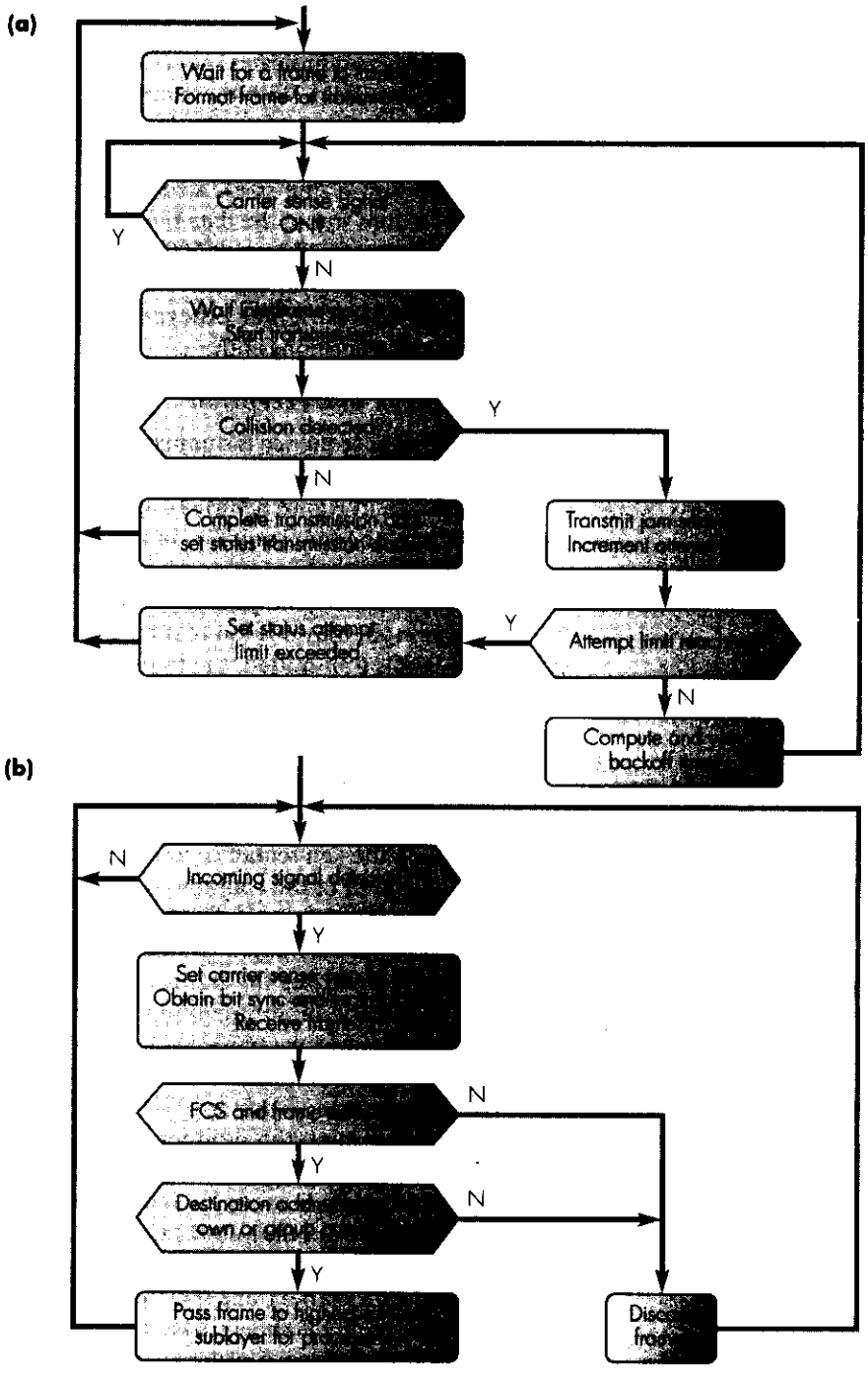


Figure 8.4 CSMA/CD MAC sublayer operation: (a) transmit; (b) receive.

8.4 Token ring

Token ring networks are also used in technical and office environments and, in addition, in industrial environments. As we can deduce from the previous section, with an Ethernet LAN the time to transmit a frame is nondeterministic since, during heavy load conditions when collisions are likely to be frequent, the transmission of a frame may be delayed and, in the limit, unsuccessful. Although in an office environment this is to a degree acceptable, in industrial environments such as manufacturing and process control, it is unacceptable. Hence although token ring LANs also use a high bit rate shared/broadcast transmission medium, in order to provide a deterministic service, they utilize a completely different MAC method.

Also, because the type of information – and hence frame contents – to be transmitted in an industrial application may have different levels of importance, frames can be assigned different priorities. The MAC method, therefore, also contains a priority control algorithm to ensure higher priority frames – for example those containing alarm messages – are transmitted before lower priority frames. We shall discuss this and other issues separately.

Control token

In a token ring LAN, all the stations are connected together by a set of unidirectional links in the form of a ring and all frame transmissions between any of the stations take place over it by circulating the frame around the ring. In its basic form, only one frame transfer can be in progress over the ring at a time. When the ring is first initialized, a single control token (frame) is generated and, in the absence of any frames to transmit, this continuously circulates around the ring. Then, when a station generates a frame to send, the steps taken by the MAC unit within each station are as illustrated in Figure 8.5.

Whenever a station wishes to send a frame, it first waits for the token. On receipt of the token, it initiates transmission of the frame, which includes the address of the intended recipient at its head. The frame is repeated (that is, each bit is received and then retransmitted) by all the stations in the ring until it circulates back to the initiating station, where it is removed. In addition to repeating the frame, each station reads and stores the frame contents. The intended recipient – indicated by the destination address in the frame header – retains a copy of the frame and indicates that it has done this by setting the response bits at the tail of the frame.

A station releases the token in one of two ways depending on the bit rate (speed) of the ring. With slower rings (4 Mbps), the token is released only after the response bits have been received. With higher speed rings (16 Mbps), it is released immediately after transmitting the last bit of the frame. This is known as **early (token) release** and, as we shall explain later, this is done to improve the level of utilization of the ring.

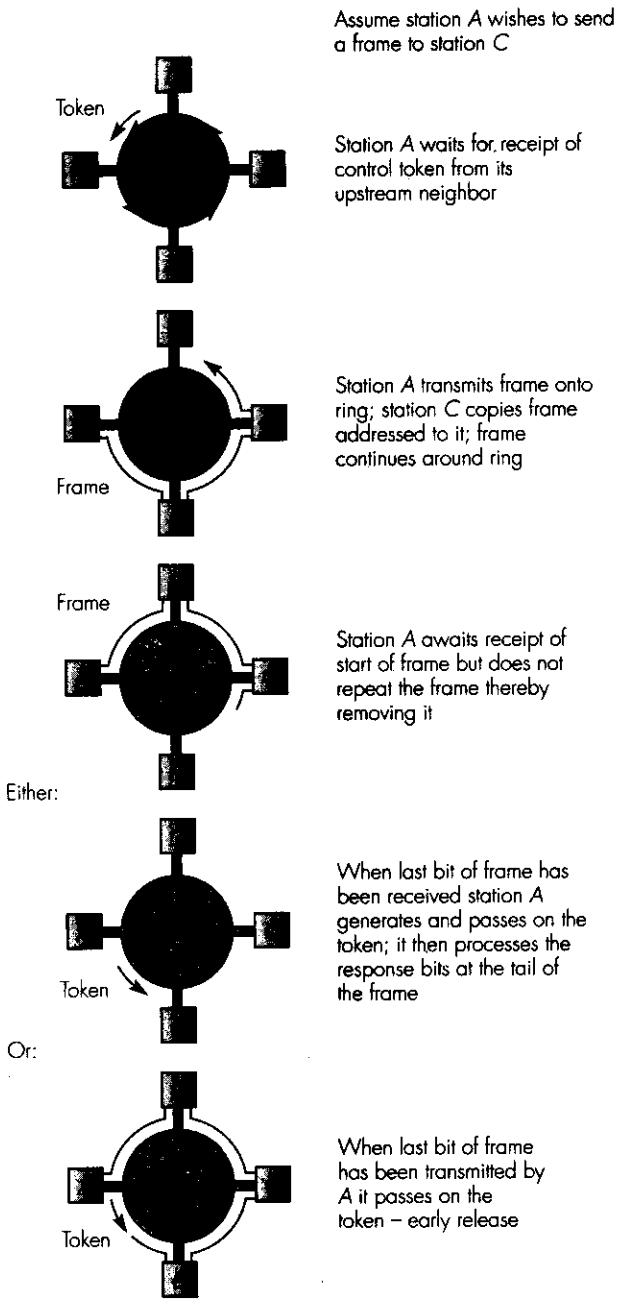


Figure 8.5 Token ring network: principle of operation.

Wiring configurations

A small token ring network is shown in Figure 8.6(a). As we can see, in this configuration is a single hub – also known as a **concentrator** and located in, say, the wiring closet of an office – to which all the stations are attached by a cable containing two twisted-pairs, one for each direction of transmission. Within the hub, associated with each station is a **station coupling unit (SCU)** and, as we show in the figure, these are interconnected so that all the stations are interconnected together to form a unidirectional ring.

An SCU is shown in Figure 8.6(b) and, as we can see, it contains a set of relays and additional electronics to drive and receive signals to and from the cable. The relays are so arranged that whenever a station is switched off, the SCU is in the *bypassed state* and a continuous transmission path through the SCU is maintained. The insertion of a station into the ring is initiated when the station is switched on. A separate pair of wires in the drop cable is used to pass power from the station to the SCU and, when activated, the relays change position so that the station becomes inserted.

When the SCU is in the *inserted state*, all signals from the ring are routed through the MAC unit of the station. The receive/transmit electronics in the MAC unit then either simply read and relay (repeat) the received signal to the transmit side, if this station is not the originator of the frame, or remove the received signal from the ring, if it initiated the transmission.

The use of two pairs of relays connected in this way means that the MAC unit can detect certain open-circuit and short-circuit faults in either the transmit or the receive pair of signal wires. Also, in the bypassed state, the MAC unit can conduct self-test functions, since any data output on the transmit pair is looped back on the receive pair. Each station is connected to the SCU by a shielded twisted-pair (STP) cable containing three twisted-pair wires; one for transmission, the second for reception, and the third to supply power to the SCU.

As we show in Figure 8.6(c), larger configurations are formed by interconnecting multiple nodes/concentrators together by means of either an STP or optical fiber trunk cable. In this case, associated with each hub is a second relay unit known as a **trunk coupling unit (TCU)**. This has the same function as an SCU except that the power to activate the relays – and hence insert the TCU and its attached stations into the ring – is from the power supply of the hub/concentrator. In this way, should a hub fail, the remainder of the ring will continue functioning.

Ring interface

The MAC unit in each station performs such functions as frame encapsulation and de-encapsulation, FCS generation and error detection, and the implementation of the MAC control algorithm. Associated with the ring is a single master station known as the **active ring monitor** that supplies the master clock for the ring. All stations are capable of performing this function and the active monitor is selected using a bidding process involving all the stations that are currently inserted into the ring.

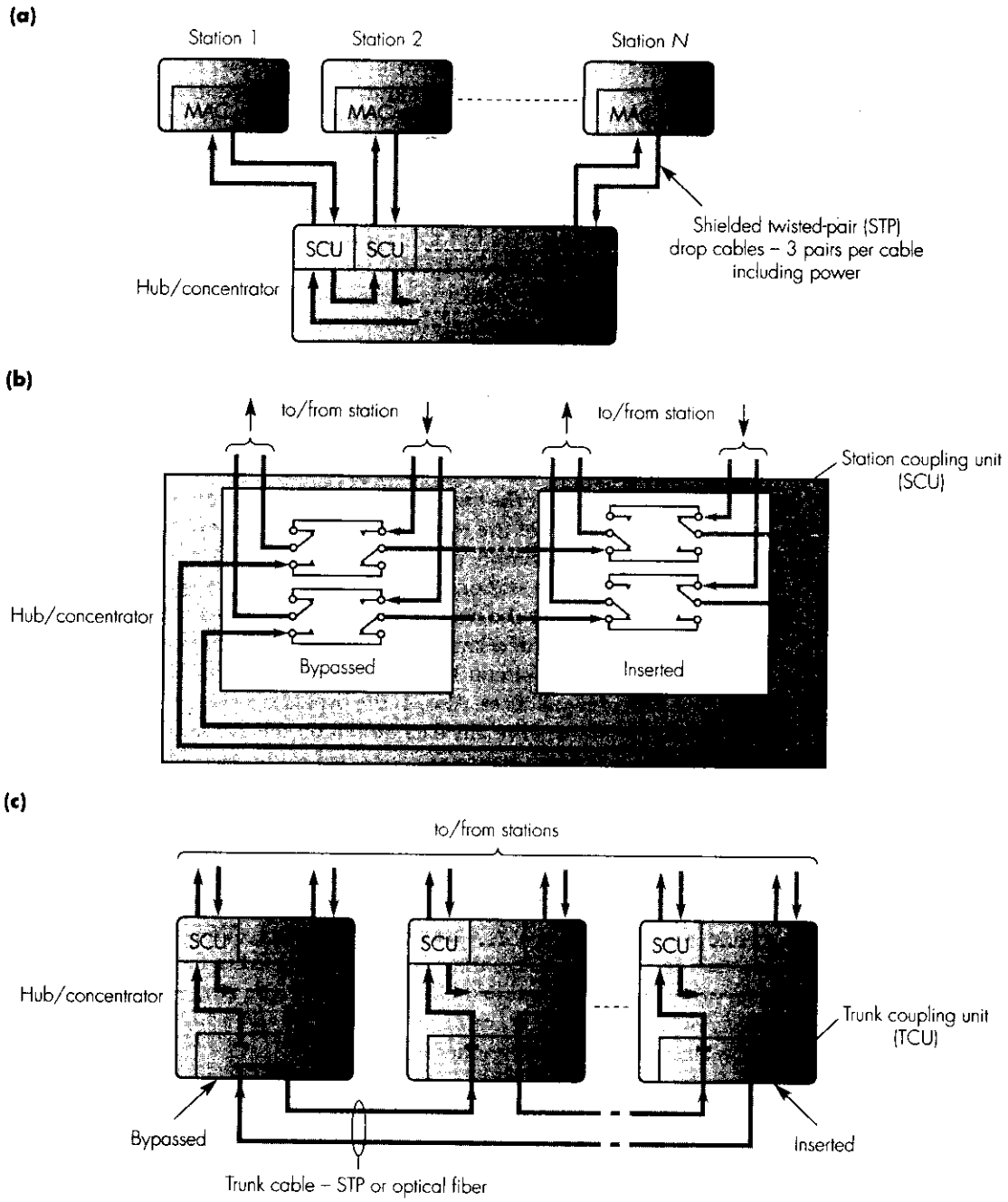


Figure 8.6 Token ring wiring configurations: (a) single hub; (b) station coupling unit; (c) multiple hubs/concentrators.

Each circulating bitstream is differential Manchester encoded by the active ring monitor and all other inserted stations in the ring frequency and phase lock their own clock using a DPLL and the incoming bitstream as we explained in Section 6.5.1. Hence there is only a small fraction of a bit delay in each ring interface as the repeater retransmits each bit as it is being received. In addition, when the station is the active ring monitor, it ensures the ring has a **minimum latency time**. This is the time, measured in bit times at the ring data transmission rate, taken for a signal to propagate once around the ring. The ring latency time includes the signal propagation delay through the ring transmission medium together with the sum of the delays through each MAC unit. For the control token to circulate continuously around the ring when none of the stations requires to use the ring (that is, all stations are simply in the repeat mode), the ring must have a minimum latency time of at least the number of bits in the token sequence to ensure that the token is not corrupted.

The token is 24 bits long, so when a station is the active ring monitor, its MAC unit provides a fixed 24-bit buffer, which effectively becomes part of the ring to ensure its correct operation under all conditions. Although the mean line signaling rate around the ring is controlled by a single master clock in the active monitor, the use of a separate DPLL circuit in each MAC unit means that the actual signaling rate may vary slightly around the ring. The worst-case variation is when the maximum number of stations (250) are all active, which is equivalent to plus or minus three bits. Unless the latency of the ring remains constant, however, bits will be corrupted as the latency decreases, or additional bits will be added as the latency increases. To maintain a constant ring latency, an additional **elastic (variable) buffer** with a length of six bits is added to the fixed 24-bit buffer. The resulting 30-bit buffer is initialized to 27 bits. If the received signal at the master MAC unit is faster than the master oscillator, the buffer is expanded by a single bit. Alternatively, if the received signal is slower, the buffer is reduced by a single bit. In this way the ring always comprises sufficient bits to allow the token to circulate continuously around the ring in the quiescent (idle) state.

Frame formats

In addition to the control token and information frames, additional frames are used for various ring management functions including the selection of an active ring monitor. Collectively, these are known as MAC frames.

Two basic formats are used in token rings: one for the control token and the other for normal frames. The control token is the means by which the right to transmit (as opposed to the normal process of repeating) is passed from one station to another, whereas a normal frame is used by a station to send either data or MAC information around the ring. The format of the two types of frame are given in Figure 8.7 together with the bit sequence used for each field.

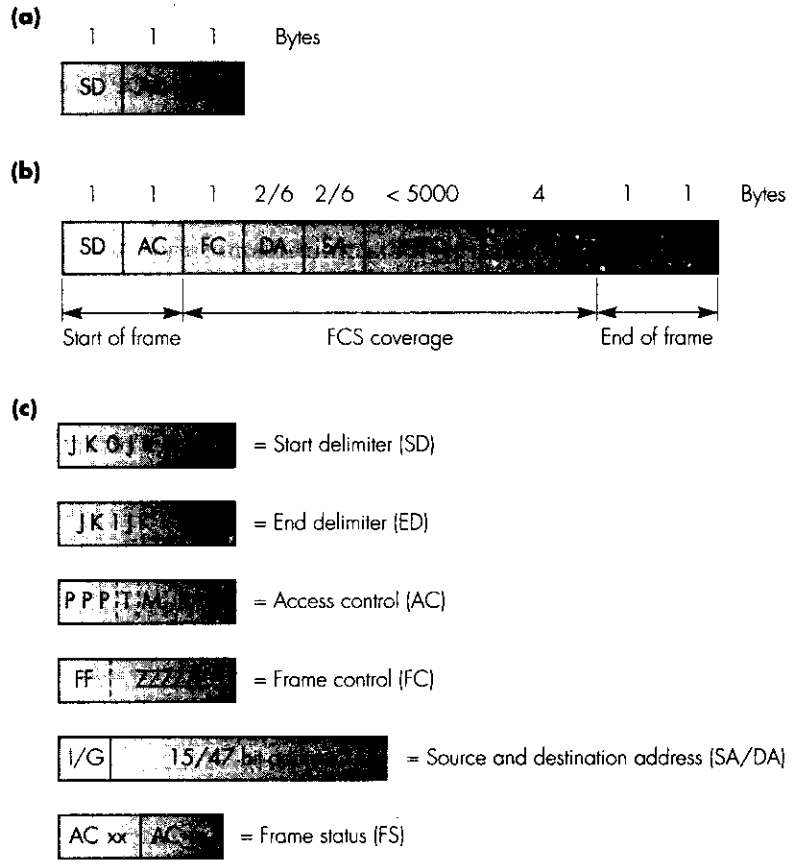


Figure 8.7 Token ring network frame formats and field descriptions: (a) token format; (b) frame format; (c) field descriptions.

The *start delimiter* (*SD*) and *end delimiter* (*ED*) fields are special bit sequences used to achieve data transparency. They exploit the symbol encoding method used on the cable medium: all information bits transmitted on the medium are differential Manchester encoded, except for selected bits in the SD and ED fields. In contrast, the J and K symbols depart from the normal encoding rules, being used instead to represent constant levels for the complete bit cell period. The J symbol has the same polarity as the preceding symbol, whereas a K symbol has the opposite polarity to the preceding symbol. In this way the receiver can reliably detect the start and end of each transmitted token or frame irrespective of its contents or length. Note, however, that only the first six symbols (JK1JK1 in Figure 8.7(c)) are used to indicate a valid end of frame. The other two bits, I and E, have other functions:

- In a token, both the I- and E-bits are 0.
- In a normal frame, the I-bit is used to indicate whether the frame is the first (or an intermediate) frame in a sequence (I = 1) or the last (or only) frame (I = 0).
- The E-bit is used for error detection. It is set to 0 by the originating station but, if any station detects an error while receiving or repeating the frame (for example, FCS error), it sets the E-bit to 1 to signal to the originating station that an error has been detected.

The *access control* (*AC*) field comprises the priority bits, the token and monitor bits, and the reservation bits. As its name implies, the AC field is used to control access to the ring. When it is part of the token, the priority bits (P) indicate the priority of the token and hence which frames a station may transmit on receipt of the token. The token bit (T) discriminates between a token and an ordinary frame (0 indicates a token, 1 a frame). The monitor bit (M) is used by the active monitor to prevent a frame from circulating around the ring continuously. Finally, the reservation bits (R) allow stations holding high-priority frames to request (in either repeated frames or tokens) that the next token to be issued is of the requisite priority.

The *frame control* (*FC*) field defines the type of the frame (MAC or information) and certain control functions. If the frame type bits (F) indicate a MAC frame, all stations on the ring interpret and, if necessary, act on the control bits (Z). If it is an I-frame, the control bits are interpreted only by the stations identified in the destination address field. Both the *source* and *destination addresses* (*SA* and *DA*) are standard 16/48 bit MAC addresses.

The *information* (*INFO*) field is used to carry either user data or additional control information when included in a MAC frame. Although no maximum length is specified for the information field, it is limited in practice by the maximum time which a station is allowed to transmit a frame when holding the control token. A typical maximum length is 5000 bytes.

The *frame check sequence* (*FCS*) field is derived from a 32-bit CRC. Finally, the *frame status* (*FS*) field is made up of two fields: the address-recognized bits (A) and the frame-copied bits (C). Both the A- and C-bits are set to 0 by the station originating the frame. If the frame is recognized by one or more stations on the ring, the station(s) sets the A-bits to 1. Also, if it copies the frame, it sets the C-bits to 1. In this way, the originating station can determine whether the addressed station(s) is non-existent or switched off, is active but did not copy the frame, or is active and copied the frame.

Frame transmission

On receipt of a service request to transmit a block of data (which includes the destination address and the priority of the data as a parameter), the data is first encapsulated by the MAC unit into the standard format shown in Figure 8.7. The MAC unit awaits the reception of a token with a priority less than or equal to the priority of the assembled frame. Clearly, in a system that employs

multiple priorities, a procedure must be followed to ensure that all stations have an opportunity to transmit frames in the correct order of priority. This procedure works as follows.

After formatting a frame and prior to receiving an appropriate token (that is, one with a priority less than or equal to the priority of the waiting frame), each time a frame or a token with a higher priority is repeated at the ring interface, the MAC unit reads the value of the reservation bits contained within the AC field. If these are equal to or higher than the priority of the waiting frame, the reservation bits are simply repeated unchanged. If they are lower, the MAC unit replaces the current value with the priority of the waiting frame. Then, assuming there are no other higher priority frames awaiting transmission on the ring, the token is passed on by the current owner (user) with this priority. On receipt of the token, the waiting MAC unit detects that the priority of the token is equal to the priority of the frame it has waiting to be transmitted. It therefore accepts the token by changing the token bit in the AC field to 1, prior to repeating this bit, which effectively converts the token to a start-of-frame sequence for a normal frame. The MAC unit then stops repeating the incoming signal and follows the converted start-of-frame sequence with the preformatted frame contents. While the frame contents are being transmitted, the FCS is computed and subsequently appended after the frame contents, before transmitting the end-of-frame sequence.

Once transmission of the waiting frame(s) has been started, the MAC unit stops repeating, thus removing the transmitted frame(s) after it has circulated the ring. In addition, the MAC unit notes the state of the A- and C-bits in the FS field at the tail of the frame(s) to determine whether the frame(s) has (have) been copied or ignored. It then generates a new token and forwards this on the ring to allow another waiting station to gain access to the ring. More than one frame may be sent on receipt of a usable token providing, firstly, that the priority of the other waiting frame(s) is greater than or equal to the priority of the token and, secondly, that the total time taken to transmit the other frame(s) is within a defined limit known as the **token holding time**. The default setting for the latter is 10 ms. The frame transmission operation is illustrated in Figure 8.8(a).

Frame reception

In addition to repeating the incoming signal/bitstream, the MAC unit within each active station on the ring detects the start of each frame by recognizing the special start-of-frame bit sequence. It then determines whether the frame should simply be repeated or copied. If the F-bits indicate that it is a MAC frame, the frame is copied and the C-bits are interpreted and, if necessary, acted upon. However, if the frame is a normal data-carrying frame and the DA matches either the station's individual address or relevant group address, the frame contents are copied into a frame buffer and passed on to the LC sublayer for further processing. In either case, the A- and C-bits in the frame status field at the tail of the frame are set accordingly prior to being repeated. The reception operation is shown in Figure 8.8(b).

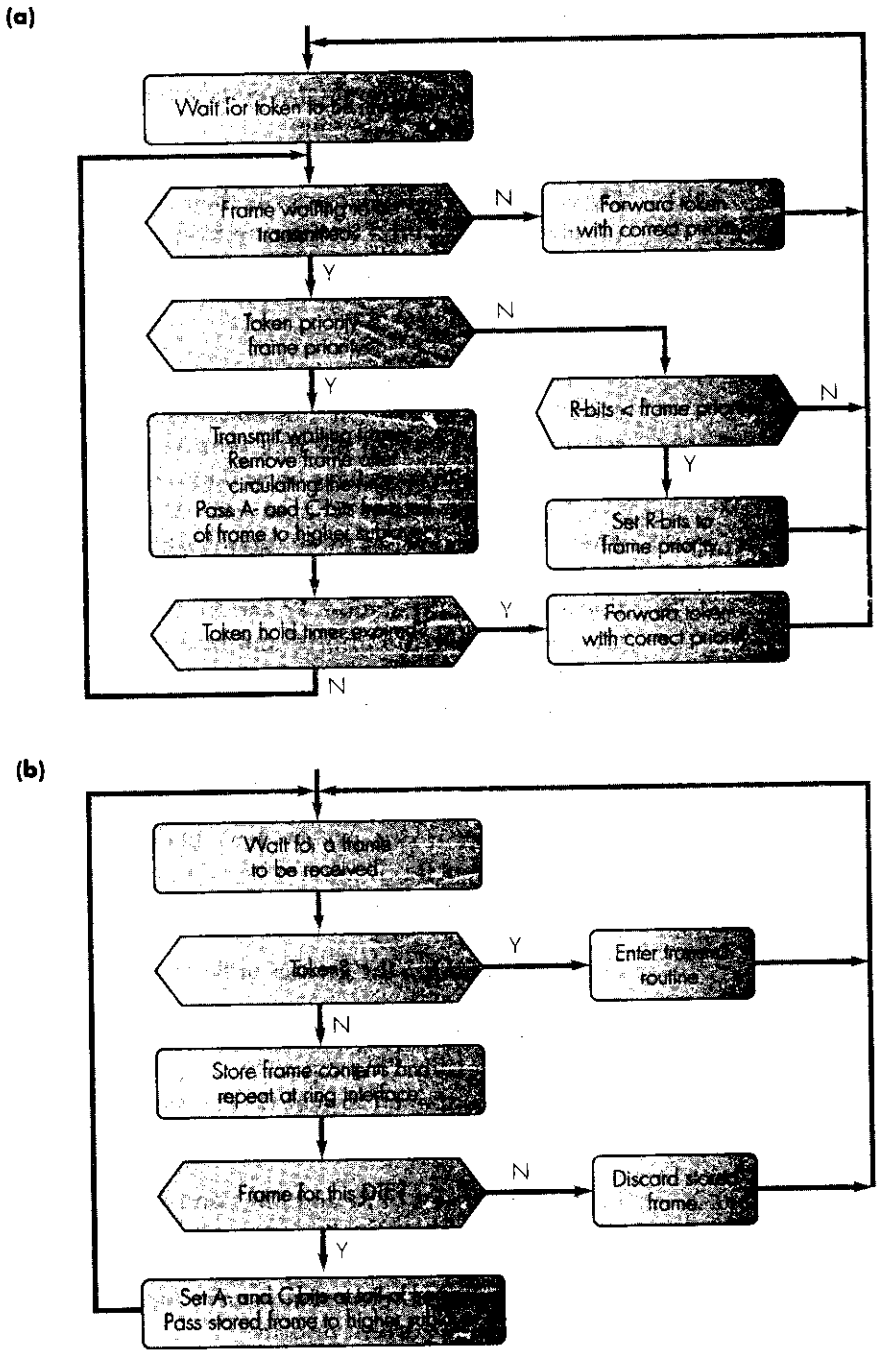


Figure 8.8 Token ring MAC sublayer operation: (a) transmit; (b) receive.

Priority operation

The priority assigned to a token by a MAC unit after it has completed transmitting any waiting frame(s) is determined by a mechanism that endeavors to ensure both of the following:

(1) Frames with a higher priority than the current ring service priority are always transmitted on the ring first.

(2) All stations holding frames with the same priority have equal access rights to the ring.

This is accomplished by using both the P- and the R-bits in the AC field of each frame coupled with a mechanism that ensures a station that raises the service priority level of the ring returns the ring to its original level after the higher priority frames have been transmitted.

To implement this scheme, each MAC unit maintains two sets of values. The first set comprises three variables Pm, Pr, and Rr. Pm specifies the highest priority value contained within any of the frames currently waiting transmission at the station. Pr and Rr are known as **priority registers** and contain, respectively, the priority and reservation values held within the AC field of the most recently repeated token or frame. The second set of values comprises two stacks known as the Sr and Sx stacks which are used as follows.

All frames transmitted by a station, on receiving a usable token, are assigned a priority value in the AC field equal to the present ring service priority Pr, and a reservation value of zero. After all waiting frames at or greater than the current ring priority have been transmitted, or until the transmission of another frame cannot be completed before the token holding time expires, the MAC unit generates a new token with:

(1) P = Pr and R = the greater of Rr and Pm

if the station does not have any more waiting frames with a priority (as contained in register Pm) equal to or greater than the current ring service priority (as contained in register Pr), or does not have a reservation request (as contained in register Rr) greater than the current priority.

(2) P = the greater of Rr and Pm and R = 0

if the station has another waiting frame(s) with a priority (as contained in Pm) greater than the current priority Pr, or if the current contents of Rr are greater than the current priority.

Since in the latter case the station effectively raises the service priority level of the ring, it becomes what is known as a **stacking station** and, as such, stores the value of the old ring service priority (Pr) on stack Sr and the new

ring service priority (P) on stack Sx. These values are saved, as it is the responsibility of the station that becomes the stacking station to lower the ring service priority level when there are no frames ready to transmit, at any point on the ring, with a priority equal to or greater than the P stacked on Sx. Also, a stack is used rather than a single register because a stacking station may need to raise the service priority of the ring more than once before the service priority is returned to a lower priority level. The different values assigned to the P- and R-bits of the token and the actions performed on the two stacks are summarized in Figure 8.9(a).

Having become a stacking station, the MAC unit claims every token that it receives with a priority equal to that stacked on Sx to examine the

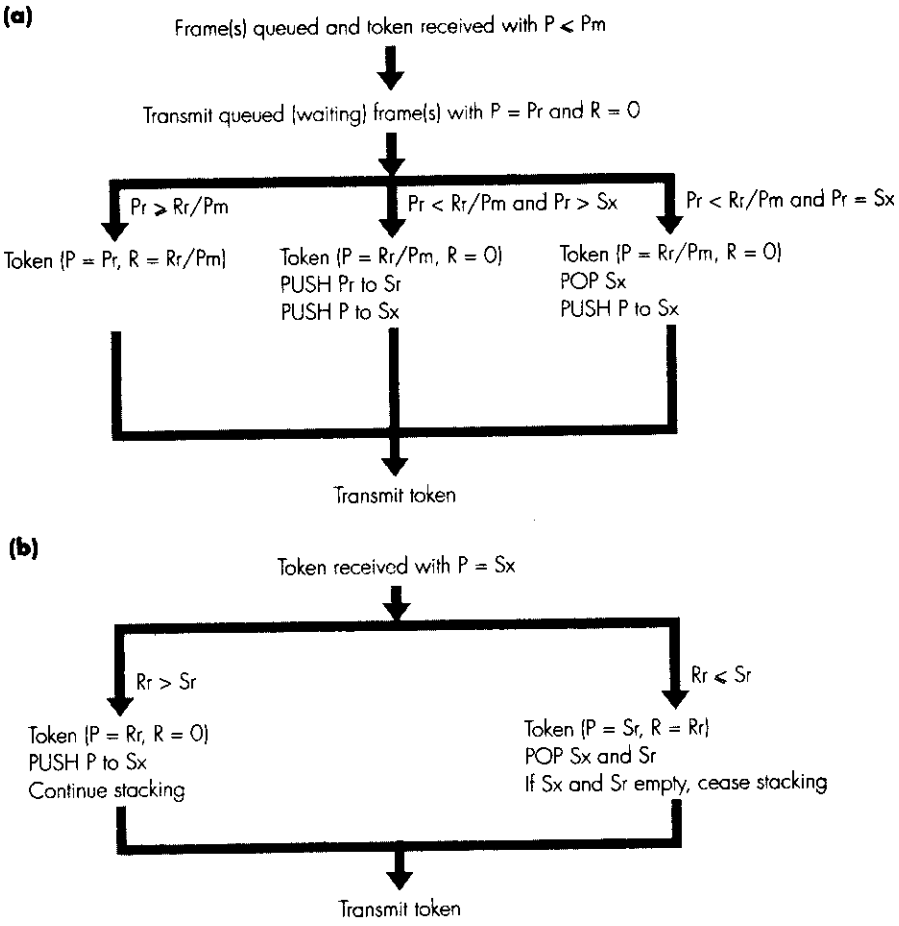


Figure 8.9 Token generation and stack modifications: (a) token generation [Note: Sx = 0 if stack empty]; (b) stack modification.

value in the R-bits of the AC field to determine if the service priority of the ring should be raised, maintained, or lowered. The new token is then transmitted with:

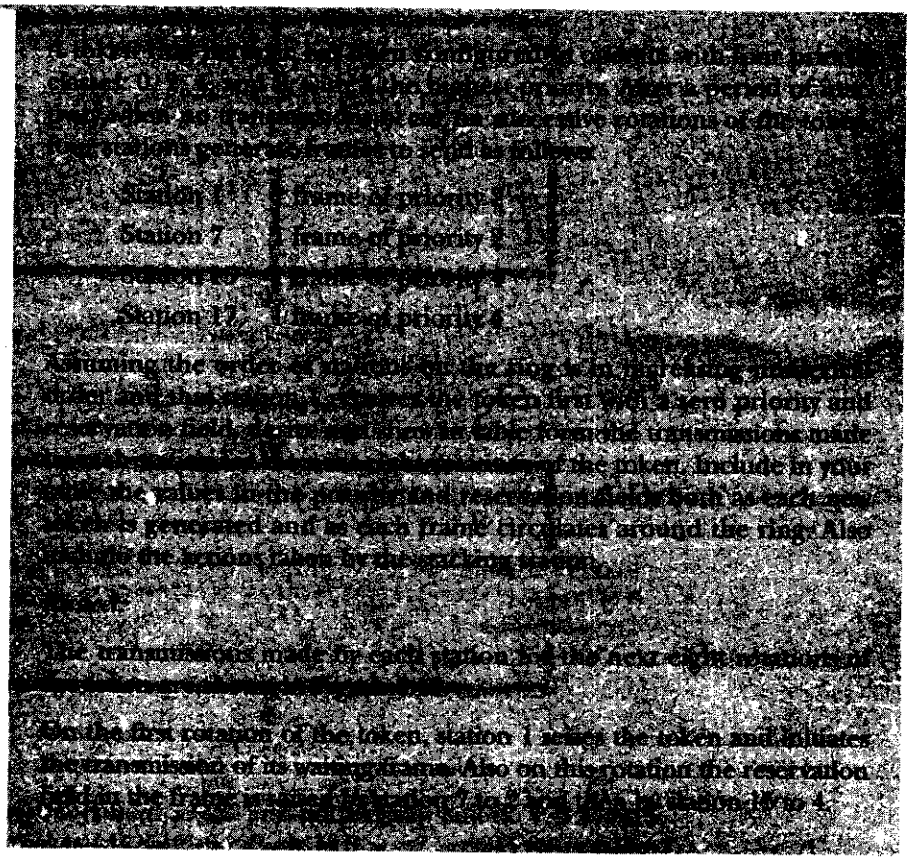
(1) P = Rr and R = 0

if the value contained in the R-bits (the current contents of register Rr) is greater than Sr. The new ring service priority (P) is stacked (PUSHed) on to Sx and the station continues its role as a stacking station.

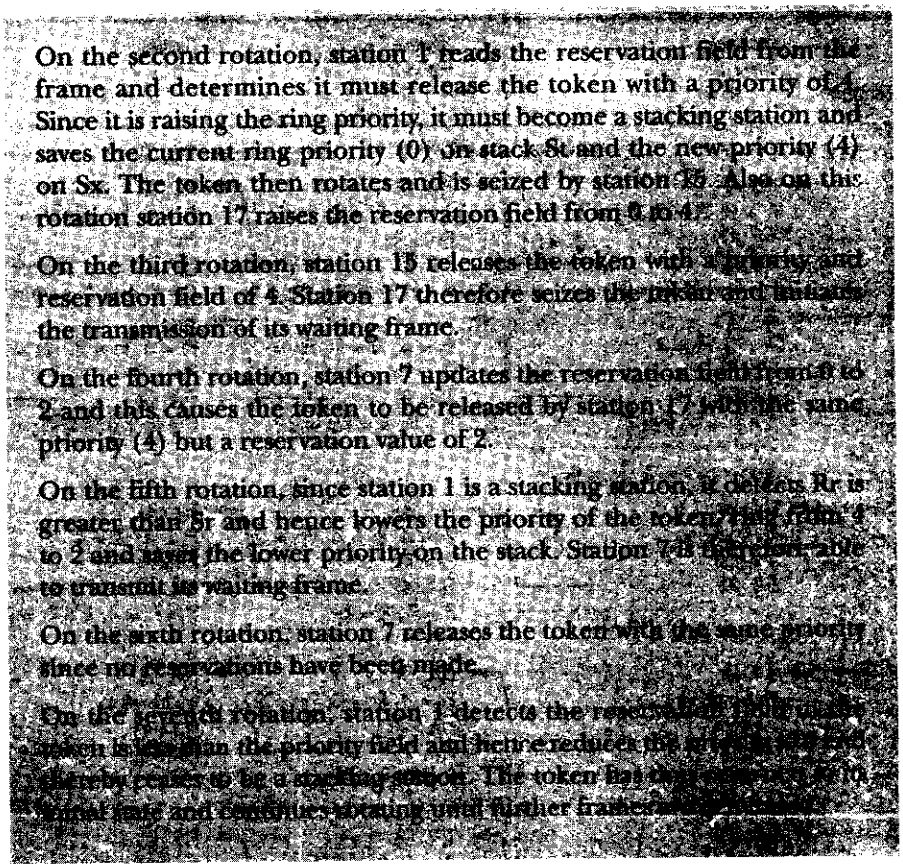
(2) P = Sr and R = Rr (unchanged)

if the value contained in the R-bits is less than or equal to Sr. Both values currently on the top of stacks Sx and Sr are POPped from the stack and, if both stacks are then empty, the station discontinues its role as a stacking station. These two operations are summarized in Figure 8.9(b).

Example 8.1



8.1 Continued



Ring management

We have been primarily concerned with the transmission of frames and tokens during normal operation of the ring. However, the ring must be set up before normal operation can take place. If a station wishes to join an already operational ring, the station must first go through an initialization procedure to ensure that it does not interfere with the correct functioning of the established ring. In addition, during normal operation it is necessary for each active station on the ring to monitor continuously its correct operation and, if a fault develops, to take corrective action to try to re-establish a correctly functioning ring. Collectively, these functions are known as **ring management.** A list of the various MAC frame types associated with these functions is given in Table 8.1 and a summary of the role of each function follows.

Initialization: When a station wishes to become part of the ring after being either switched on or reset, it enters an initialization sequence to ensure that no other stations in the ring are using the same address and to inform its immediate downstream neighbor that it has (re)entered the ring.

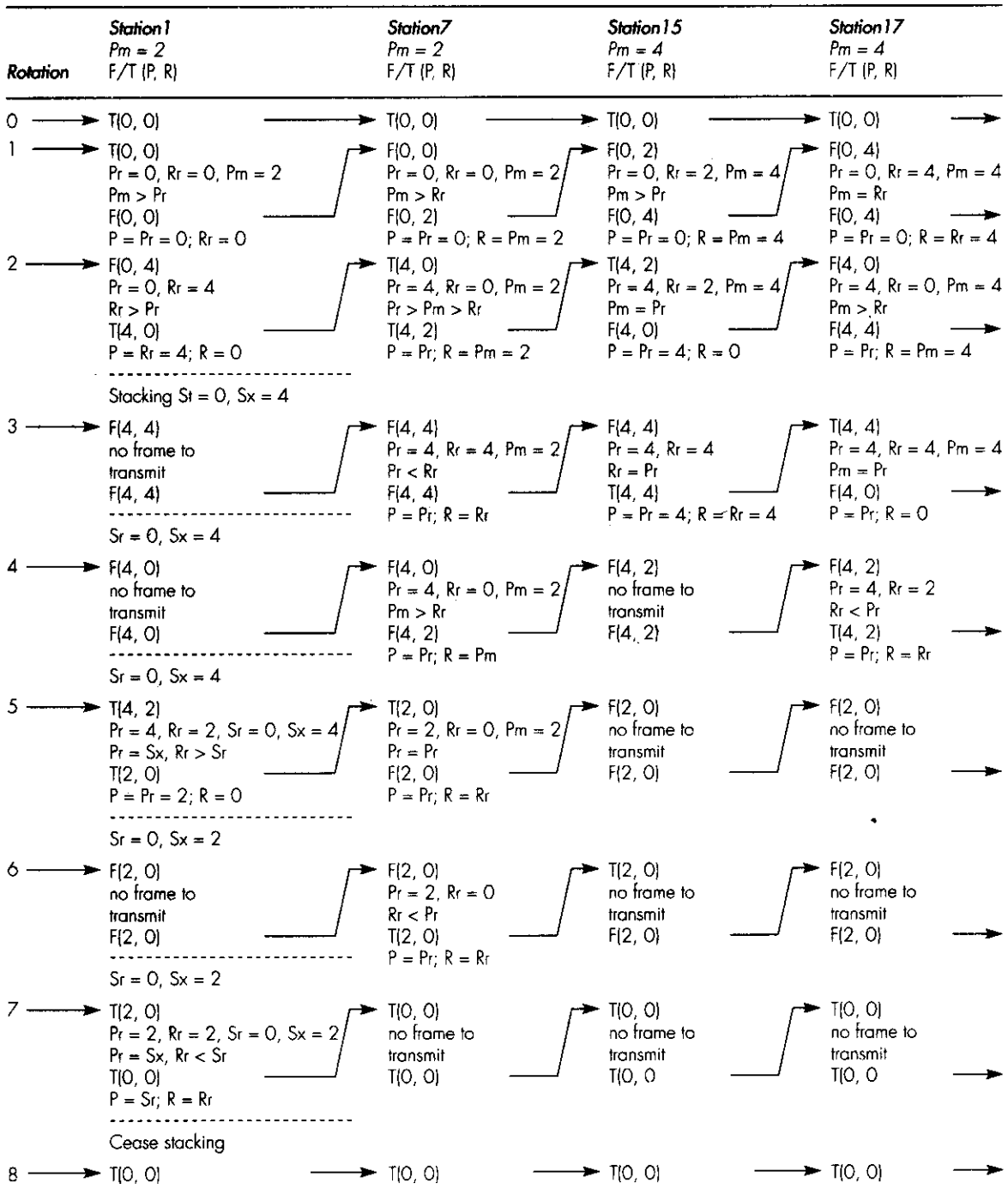


Figure 8.10 Token ring priority example.

Table 8.1 Token ring management: MAC frame types.

Frame type	Function
Duplicate address test (DAT)	Used during the initialization procedure to enable a station to determine that no other stations in the ring are using its own address
Standby monitor present (SMP)	Used in the initialization procedure to enable a station to determine the address of its upstream neighbor (successor) in the ring
Active monitor present (AMP)	These types of frames are transmitted at regular intervals by the currently active monitor and each station monitors their passage
Claim token (CT)	Used in the procedure to determine a new active monitor if the current one fails
Purge (PRG)	Used by a new active monitor to initialize all stations into the idle state
Beacon (BCN)	Used in the beaconing procedure

Standby monitor: Upon completion of the initialization sequence, the station can start to transmit and receive normal frames and tokens. In addition, the station enters the standby monitor state to monitor continuously the correct operation of the ring. It does this by monitoring the passage of tokens and special active monitor present (AMP) MAC frames – which are periodically transmitted by the currently active monitor – as they are repeated at the ring interface.

Active monitor: If the station is successful in its bid to become the new active monitor, it first inserts its latency buffer into the ring and enables its own clock. It then initiates the transmission of a purge (PRG) MAC frame to ensure that there are no other tokens or frames on the ring before it initiates the transmission of a new token.

Beaconing: If a serious failure such as a broken cable arises in the ring, a procedure known as beaconing informs each station on the ring that the token-passing protocol has been suspended (until the affected failure domain has been located and repaired). The failure domain consists of the following:

- the station that reports the failure, which is referred to as the **beaconing station**;
- the station upstream of the beaconing station;
- the ring medium between them.

We can see that the MAC procedures used with a token ring network are quite complicated, certainly compared with a CSMA/CD bus, for example. Remember, however, that most of the procedures are implemented in special integrated circuits within the MAC unit, so their operation is transparent to the user. Moreover, many of these ring management procedures are invoked only when faults develop and so the overheads associated with them are, on the whole, modest.

8.5 Bridges

There are two types of bridge, the ones that are used with Ethernet LANs, known as **transparent bridges**, and the others with token ring **LANs**, known as **source routing bridges**. Before we explain their operation, however, it will be helpful first to review the operation of repeaters since bridges were designed to overcome the limitations that occur when using repeaters.

Repeaters are used to ensure that the electrical signal transmitted by the line drivers within the MAC unit propagate throughout the network. For each LAN segment, in order to limit the signal attenuation to an acceptable level, there is a defined maximum limit set on the physical length of the segment and on the number of (end) stations that may be attached to it. When interconnecting segments, a repeater is used to limit the electrical drive requirements of the line driver circuit to that of a single segment. In this way, the presence of multiple segments (and hence repeaters) in a transmission path is transparent to the source station. The repeater, after achieving clock synchronization, simply regenerates all signals received on one segment and forwards (repeats) them onto the next segment. This form of interconnection is shown in Figure 8.11(a) and, as we can deduce from the figure, all frame transmissions from any station attached to a segment will propagate throughout the total LAN and hence be received by all the other attached stations. This means, therefore, that in terms of available bandwidth, the network behaves like a single segment.

In early LAN installations, because most traffic was text-based email and occasional file transfers, this mode of operation gave an acceptable performance in terms of network access and transfer delays. With the arrival of diskless nodes/stations, however, all disk accesses to the server are via the network and hence the demands on the network bandwidth are substantially higher. In most cases, the server and the set of diskless nodes it serves are all attached to the same LAN segment. Hence there is no necessity for the frames associated with such transfers to be transmitted beyond the segment on which they are generated. Bridges were introduced, therefore, to inhibit the forwarding of such frame transfers and only to forward those frames that are intended for a different segment.

Thus, the function of a bridge is similar to a repeater in that it is used for interconnecting LAN segments. However, when bridges are used, all frames

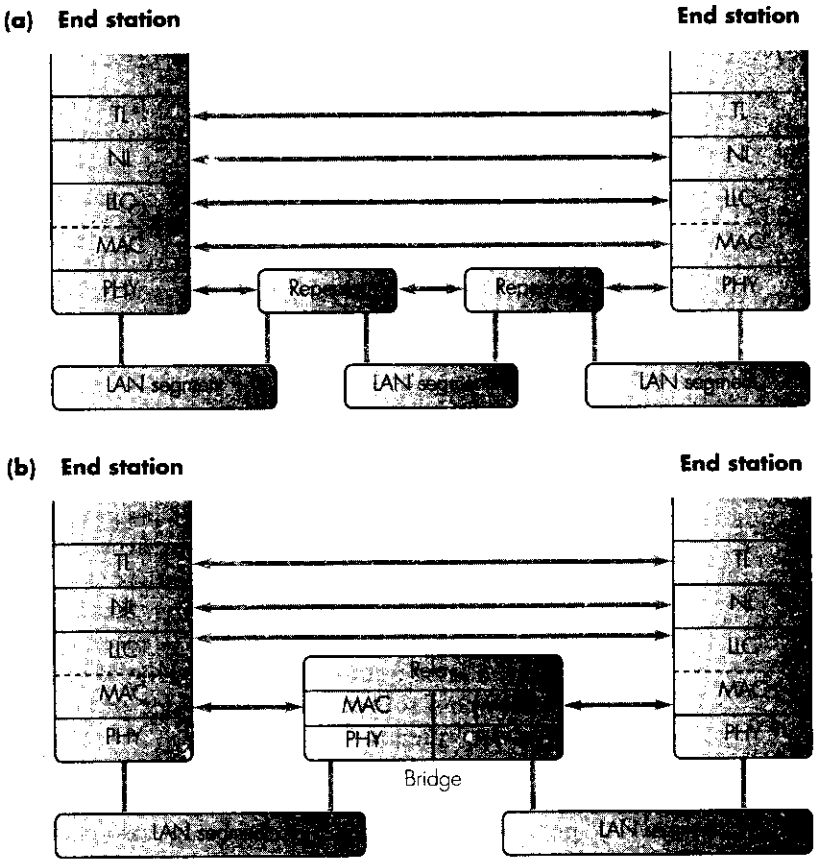


Figure 8.11 LAN interconnection: (a) repeaters; (b) bridges.

received from a segment are buffered (stored) and error checked before they are repeated (forwarded). Moreover, only frames that are free of errors and are addressed to stations on a different segment from the one on which they were received are forwarded. Consequently, all transmissions between stations connected to the same LAN segment are not forwarded and hence do not load the rest of the network. A bridge thus operates at the MAC sublayer in the context of our reference model. This is shown in Figure 8.11(b) and the resulting LAN is then referred to as a **bridged LAN**.

8.5.1 Transparent bridges

With a transparent bridge, as with a repeater, the presence of one (or more) bridges in a route between two communicating stations is transparent to the two stations. All routing decisions are made exclusively by the bridge(s). Moreover, a transparent bridge automatically initializes and configures itself

(in terms of its routing information) in a dynamic way after it has been put into service. A schematic of a bridge is shown in Figure 8.12(a) and a simple bridged LAN in Figure 8.12(b).

A LAN segment is physically connected to a bridge through a **bridge port**. A basic bridge has just two ports whereas a **multiport bridge** has a number of connected ports (and hence segments). In practice, each bridge port comprises the MAC integrated circuit chipset associated with the particular type of LAN segment – CSMA/CD, Ethernet – together with some associated port management software. The software is responsible for initializing the chipset at start-up – chipsets are all programmable devices – and for buffer management. Normally, the available memory is logically divided into a number of fixed-size units known as buffers. Buffer management involves passing a free buffer (pointer) to the chipset for onward transmission (forwarding).

Every bridge operates in the **promiscuous mode** which means it receives and buffers all frames received on each of its ports. When a frame has been received at a port and put into the assigned buffer by the MAC chipset, the port management software prepares the chipset for a new frame and then passes the pointer of the memory buffer containing the received frame to the **bridge protocol entity** for processing. Since two (or more) frames may arrive concurrently at the ports and two or more frames may need to be forwarded from the same output port, the passing of memory pointers between the port management software and the bridge protocol entity software is carried out via a set of queues.

As we shall describe later, each port may be in a number of alternative states and processing of received frames is carried out according to a defined protocol. The function of the bridge protocol entity software is to implement the particular bridge protocol being used.

Frame forwarding (filtering)

A bridge maintains a **forwarding database** (also known as a **routing directory**) that indicates, for each port, the outgoing port (if any) to be used for forwarding each frame received at that port. If a frame is received at a port that is addressed to a station on the segment (and hence port) on which it was received, the frame is discarded; otherwise it is forwarded via the port specified in the forwarding database. The normal routing decision involves a simple look-up operation: the destination address in each received frame is first read and then used to access the corresponding port number from the forwarding database. If this is the same as the port on which it was received, the frame is discarded, else it is queued for forward transmission on the segment associated with the accessed port. This process is also known as **frame filtering**.

Bridge learning

A problem with transparent bridges is the creation of the forwarding database. One approach is for the contents of the forwarding database to be created in advance and held in a fixed memory, such as programmable

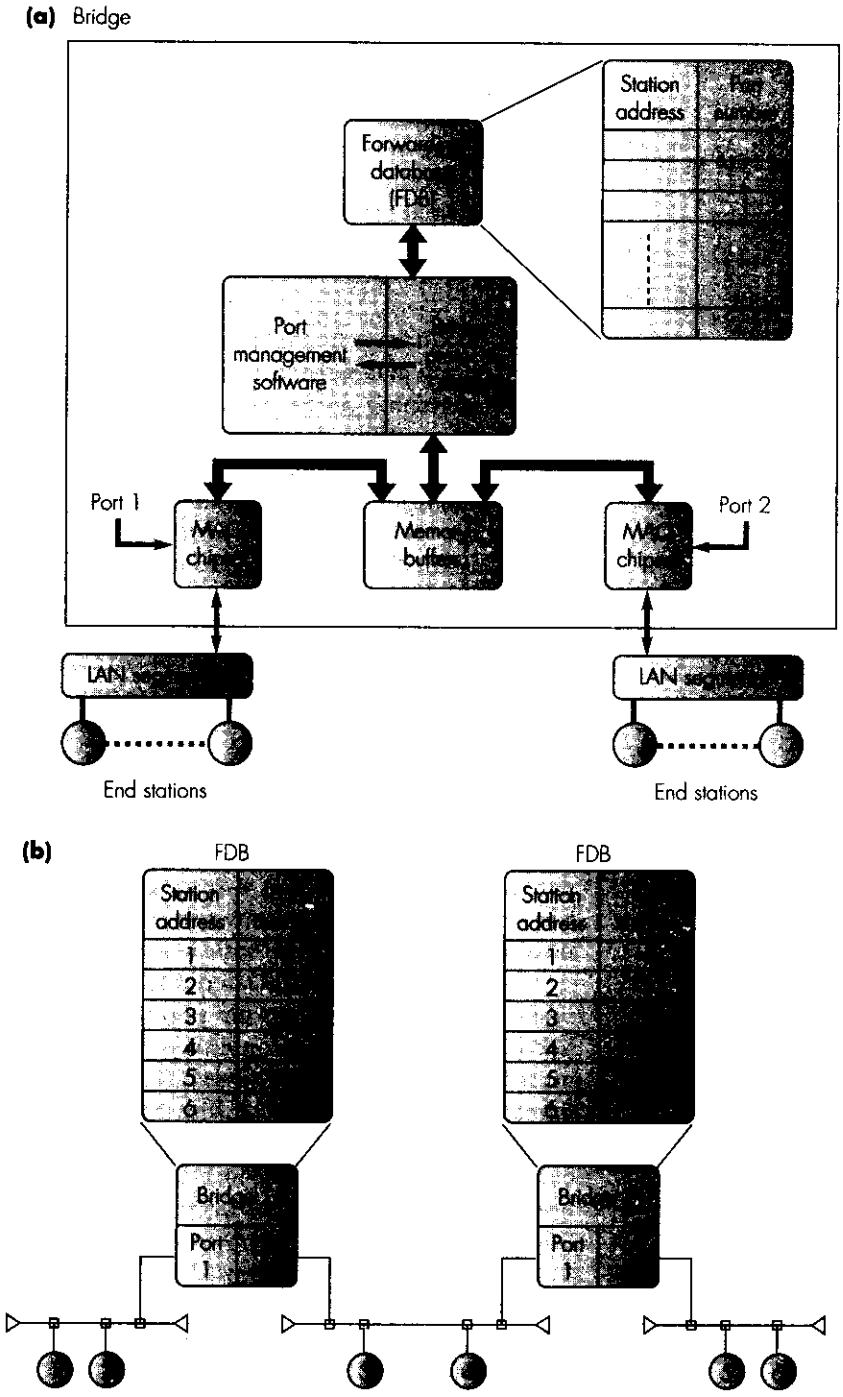


Figure 8.12 Transparent bridge schematic: (a) architecture; (b) application example.

read-only memory (PROM). The disadvantage is that the contents of the forwarding database in all bridges have to be changed whenever the network topology is changed – a new segment added, for example – or when a user changed the point of attachment (and hence segment) of his or her station. To avoid this, in most bridged LANs the contents of the forwarding database are not statically set up but rather are dynamically created and maintained during normal operation of the bridge. This is accomplished using a combination of a learning process and a dialog with other bridges to ascertain the topology of the overall installed LAN. An overview of the learning process is as follows.

When a bridge first comes into service, its forwarding database is initialized to empty. Whenever a frame is received, the *source address* within it is read and the incoming port number on which the frame was received is entered into the forwarding database. In addition, since the forwarding port is not known at this time, a copy of the frame is forwarded on all the other output ports of the bridge. As these frames propagate through the network, this procedure is repeated by each bridge. Firstly, the incoming port number is entered in the forwarding database against the source (station) address and a copy of the frame is forwarded on all the other output ports of the bridge. This action is referred to as **flooding** since it ensures that a copy of each frame transmitted is received on all segments in the total LAN. During the learning phase this procedure is repeated for each frame received by the bridge. In this way, all bridges in the LAN rapidly build up the contents of their forwarding databases.

This procedure works satisfactorily as long as stations are not allowed to migrate around the network (change their point of attachment) and the overall LAN topology is a simple tree structure (that is, there are no duplicate paths (routes) between any two segments). Such a tree structure is known as a **spanning tree**. Since in many networks, especially large networks, both these possibilities may occur, the basic learning operation is refined as follows.

The MAC address associated with a station is fixed at the time of its manufacture. If a user changes the point of attachment to the network of his or her PC/workstation, the contents of the forwarding database in each bridge must be periodically updated to reflect such changes. To accomplish this, an **inactivity timer** is associated with each entry in the database. Whenever a frame is received from a station within the predefined time interval, the timer expires and the entry is removed. Whenever a frame is received from a station for which the entry has been removed, the learning procedure is again followed to update the entry in each bridge with the (possibly new) port number. In this way the forwarding database in a bridge is continuously updated to reflect the current LAN topology and the addresses of the stations that are currently attached to the segments it interconnects. The inactivity timer also limits the size of the database since it contains only those stations that are currently active. This is important since the size of the database influences the speed of the forwarding operation.

The learning process works only if the total bridged LAN has a simple (spanning) tree topology. This means that there is only a single path between any two segments in the network. However, this condition may not always be met since additional bridges may be used to link two segments, for example, to improve reliability, or perhaps by mistake when a LAN is being updated.

Multiple paths between two segments cannot exist with the basic learning algorithm we have outlined since the flooding operation during the learning phase would cause entries in the forwarding database to be continuously overwritten. We can see this by considering the simple LAN topology shown in Figure 8.13. Clearly, if station 10 transmits a frame on segment 1 during the learning phase, then bridges B1 and B2 will both create an entry in their forwarding database – station 10/port 1 – and forward a copy of the frame onto segment 2. Each of these frames will in turn be received by the other bridge, and an entry will be made of station 10/port 2 and a copy of the frame output at port 1. In turn, each of these will be received by the other bridge, resulting in their corresponding entry for port 1 being updated. The frame will thus continuously circulate in a loop with the entries for each port being continuously updated.

Consequently, for topologies which offer multiple paths between stations, we need an additional algorithm to select just a single bridge for forwarding frames between any two segments. As we shall explain, this is done by setting

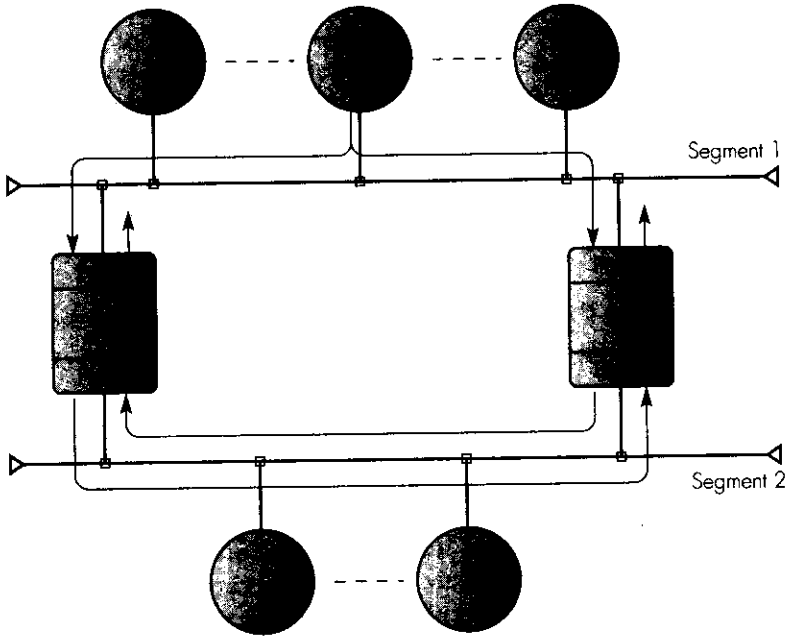


Figure 8.13 Effect of dual paths on learning algorithm.

selected bridge ports into the non-forwarding state and, in the topology in Figure 8.13, port 2 of bridge 2 would be selected. The resulting logical or **active topology** behaves as a single spanning tree and the algorithm is known as the **spanning tree algorithm**. Note that although the algorithm selects only a single port/bridge for connecting two segments – making redundant any alternative bridges that may have been introduced to improve reliability, for example – the algorithm is run at regular intervals and will dynamically select a set of bridges from those currently operational.

Spanning tree algorithm

With the spanning tree algorithm, all the bridges regularly exchange special frames (messages), known as **bridge protocol data units (BPDUs)**. Each bridge has a priority value and a unique identifier. For the total bridged LAN, a single bridge is dynamically chosen by the spanning tree algorithm to be the **root bridge**. This is the bridge with the highest priority and the smallest identifier. It is determined/confirmed at regular intervals.

All the bridges in a LAN have a unique MAC group address which is used for sending all BPDUs between bridges. Also, the path cost associated with each bridge port is determined by the bit rate – known as the **designated cost** – of the segment to which it is attached; the higher the bit rate, the smaller the designated cost. All bridges know the designated cost of the segments to which they are attached.

When a bridge is first brought into service, it assumes it is the root bridge. A bridge that believes it is the root (all initially), initiates the transmission of a **configuration BPDU** on all of its ports (and hence the segments connected to it) at regular time intervals known as the **hello time**.

Each configuration BPDU contains a number of fields including:

- the identifier of the bridge which the bridge transmitting the BPDU believes to be the root (itself initially);
- the path cost to the root from the bridge port on which the BPDU was received (zero initially);
- the identifier of the bridge transmitting the BPDU;
- the identifier of the bridge port from which the BPDU was transmitted.

On receipt of a configuration BPDU, each bridge connected to the segment on which it was transmitted can determine, by comparing the identifier contained within it with its own identifier, firstly, whether the bridge has a higher priority or, if the priorities are equal, whether its own identifier is less than the identifier from the received frame. If this is the case, the receiving bridge will carry on assuming it is the root and simply discard the received frame.

Alternatively, if the identifier from the received BPDU indicates that it is not the root, the bridge proceeds by adding the path cost associated with the

port on which the BPDU was received to that already within the frame. It then creates a new configuration BPDU containing this information, together with its own identifiers (bridge and port), and forwards a copy on all its other ports. This procedure is repeated by all bridges in the LAN. In this way the configuration BPDUs flood away from the root bridge to the extremities of the network and, at the same time, the path cost associated with each port of all the other bridges back to the root is computed. Thus, in addition to a single root bridge being established, all the other bridges will have determined the path cost associated with each of their ports. This is known as the **root path cost** (**RPC**) and the port that has the smallest RPC is then selected as the **root port** of the bridge. If two ports have the same RPC, the one with the lowest port number is chosen.

Once the root bridge and the root ports for the remaining bridges have been determined, the basis of the spanning tree has been established. The next step is to ensure that there are no loops/connections between the branches of the tree. This is done by selecting only a single bridge (port) to forward frames from each segment. This is known as the **designated bridge**. Its selection is based on the least path cost to the root bridge from the segment under consideration. If two bridge ports connected to a segment have the same path cost, the bridge with the smaller identifier is chosen. The bridge port connecting the segment to its designated bridge is known as the **designated port**. In the case of the root bridge, this is always the designated bridge for all the segments to which it is connected. Hence all its ports are designated ports.

When establishing the designated bridge port to be used with a segment, note that once a bridge port has been selected as a root port, it will not take part in the arbitration procedure to become a designated port. The choice of designated port is thus between the non-root ports connected to the segment under consideration.

The exchange of the configuration BPDUs between the two (or more) bridges involved will allow the two (or more) bridges to make a joint decision as to the port to be selected.

After the root bridge and the root and designated ports of all other bridges have been established, the state of the bridge ports can be set either to **forwarding** or to **blocking**. Initially, since all ports of the root bridge are designated ports, they are set to the forwarding state. For all the other bridges, only the root and designated ports are set to the forwarding state; the others are set to the blocking state. This establishes an active topology equivalent to a spanning tree. This procedure is then repeated at regular intervals (determined by the hello timer) to confirm the active topology or, in the event of a bridge failure, to reestablish a new topology.

Example 8.2

To illustrate how the various elements of the spanning tree algorithm work, consider the bridged LAN shown in Figure 8.14(a). The unique identifier of each bridge is shown inside the box representing the bridge together with the port numbers in the inner boxes connecting the bridge to each segment. Typically, the additional bridges on each segment are added to improve reliability in the event of a bridge failure. Also, assume that the LAN is just being brought into service, all bridges have equal priority, and all segments have the same designated cost (the same) associated with them. Determine the active (spanning tree) topology.

Answer:

(a) First the exchange of configuration BPDUs will establish bridge B1 as the root bridge since this has the lowest identifier.

(b) After the exchange of configuration BPDUs, the root path cost (RPC) of each port will have been computed. These are shown in Figure 8.14(b).

(c) The root port (RP) for each bridge is then chosen as the port with the lowest RPC. For example, in the case of bridge B3, port 1 has an RPC of 1 and port 2 an RPC of 2, so port 1 is chosen. In the case of B2, both ports have the same RPC and hence port 1 is chosen since this has the smaller identifier. The selected RPs are also shown in the figure.

(d) For the root bridge B1 all its ports have a designated port cost (DPC) of 0. Hence they are the designated ports for segments S1, S2, and S3.

(e) For S4, port 1 of B2 is an RP and hence is not involved in selecting the designated port. The two other ports connected to S4 both have a DPC of 1. Hence port 2 of B3 is selected as the designated port because it has a lower identifier than B4.

(f) For B5, the only port connected to it is port 2 of B5 and hence this is selected.

(g) Finally, for S6 both ports have a DPC of 1, so port 2 of B4 is selected rather than port 2 of B6.

The DPCs are shown in Figure 8.14(c) and the resulting active topology is shown in Figure 8.14(d). Note that the DPC of a port is always equal to the RPC of the next port of the bridge.

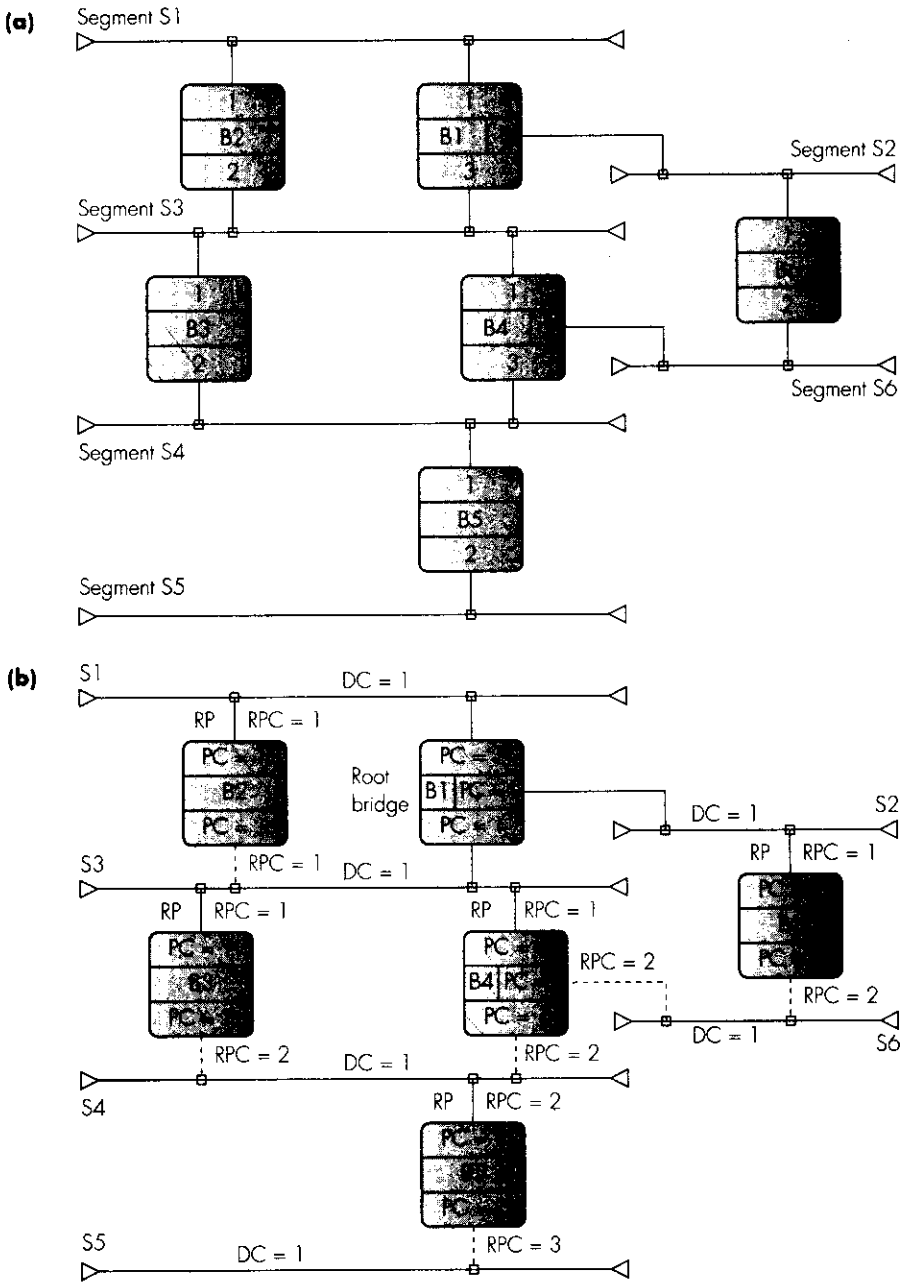


Figure 8.14 Active topology derivation example: (a) LAN topology; (b) root port selection; (c) designated port selection; (d) active topology.

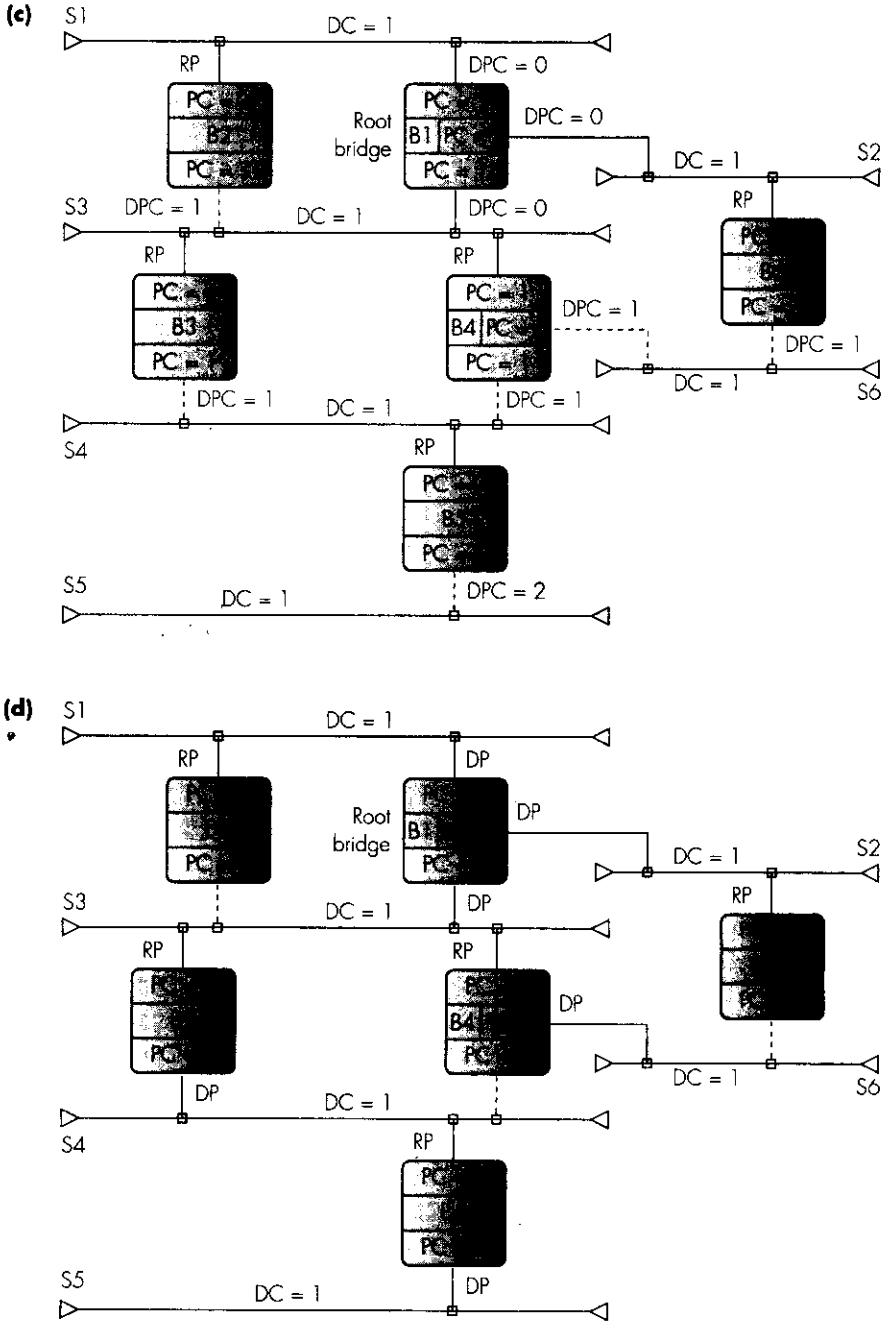


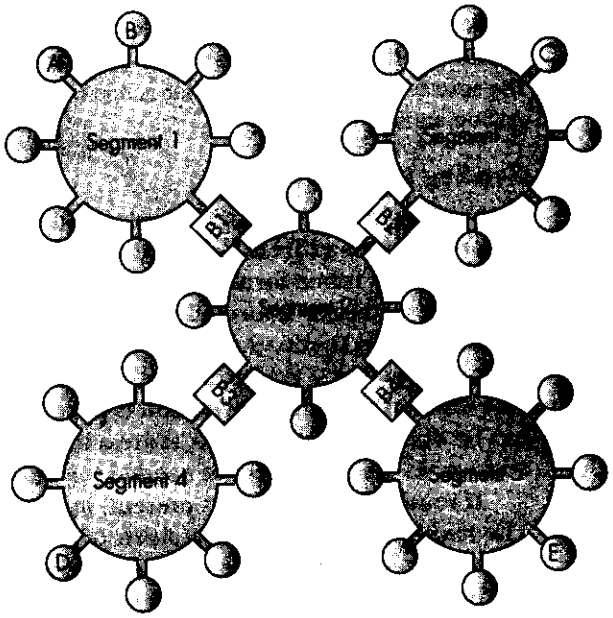
Figure 8.14 Continued.

8.5.2 Source routing bridges

Although we can use **source routing bridges** with any type of LAN segment, we use them primarily for the interconnection of token ring LAN segments. A typical network based on source routing bridges is shown in Figure 8.15(a).

The major difference between a LAN based on source routing bridges and one based on spanning tree bridges is that with the latter the bridges collectively perform the routing operation in a way that is transparent to the end stations. Conversely, with source routing, the end stations perform the routing function.

(a)



(b) Routing table held by A: Destination B = Segment 1 (same segment)
C = Segment 1, B1, Segment 2, B2, Segment 3
E = Segment 1, B1, Segment 2, B4, Segment 5

Routing table held by B: Destination A = Segment 1 (same segment)
D = Segment 1, B1, Segment 2, B3, Segment 4

Routing table held by E: Destination A = Segment 5, B4, Segment 2, B1, Segment 1
C = Segment 5, B4, Segment 2, B2, Segment 3

Figure 8.15 An example source routing bridged LAN: (a) topology; (b) routing table entries.

With source routing, a station ascertains the route to be followed by a frame to each destination before any frames are transmitted. This information is inserted at the head of the frame and is used by each bridge to determine whether a received frame is to be forwarded on another segment or not. The routing information comprises a sequence of segment-bridge, segment-bridge identifiers. Routing tables for selected stations in the example network are shown in Figure 8.15(b).

On receipt of each frame, a bridge needs only to search the routing field at the head of the frame for its own identifier. Only if it is present and followed by the identifier of a segment connected to one of its output ports does it forward the frame on the specified LAN segment. Otherwise it is not forwarded. In either event, the frame is repeated at the ring interface by the bridge and, if forwarded, the address-recognized (A) and frame-copied (C) bits in the frame status (FS) field at the tail of the frame are set to indicate to the source station (bridge) that is has been received (forwarded) by the destination station (bridge).

Routing algorithm

The routing information field contained within each frame immediately follows the source address field at the head of the normal information frame. The modified frame format is thus as shown in Figure 8.16(a).

Since a routing information field is not always required – for example, if the source and destination stations are on the same segment – the first bit of the *source address* – the individual/group (I/G) address bit – is used to indicate whether routing information is present in the frame (1) or not (0). This can be done since the source address in a frame must always be an individual address, so the I/G bit is not needed for this purpose.

If routing information is present, its format is as shown in Figure 8.16(b). The *routing information* field consists of a *routing control* field and one or more *route designator* fields. The routing control field itself comprises three subfields: *frame type, maximum frame size,* and *routing field length.* In addition to normal information frames, two other frame types are associated with the routing algorithm. The frame type indicates the type of the frame.

Source routing bridges can be used for the interconnection of different types of LAN segments in addition to token rings. Since there is a different maximum frame size associated with each segment type, the *maximum frame size* field determines the largest frame size that can be used when transmitting a frame between any two stations connected to the LAN.

Prior to transmitting a route finding frame, a station sets the maximum frame size field to the (known) largest frame size that can be used in the total LAN. Before a bridge forwards the frame on a segment, the bridge checks this field with the (known) maximum frame size of the new segment. If the latter is smaller, the bridge reduces the frame size field to the lower value. In this way, the source station, on receipt of the corresponding route reply frame, can use this information when preparing frames for transmission to that destination.

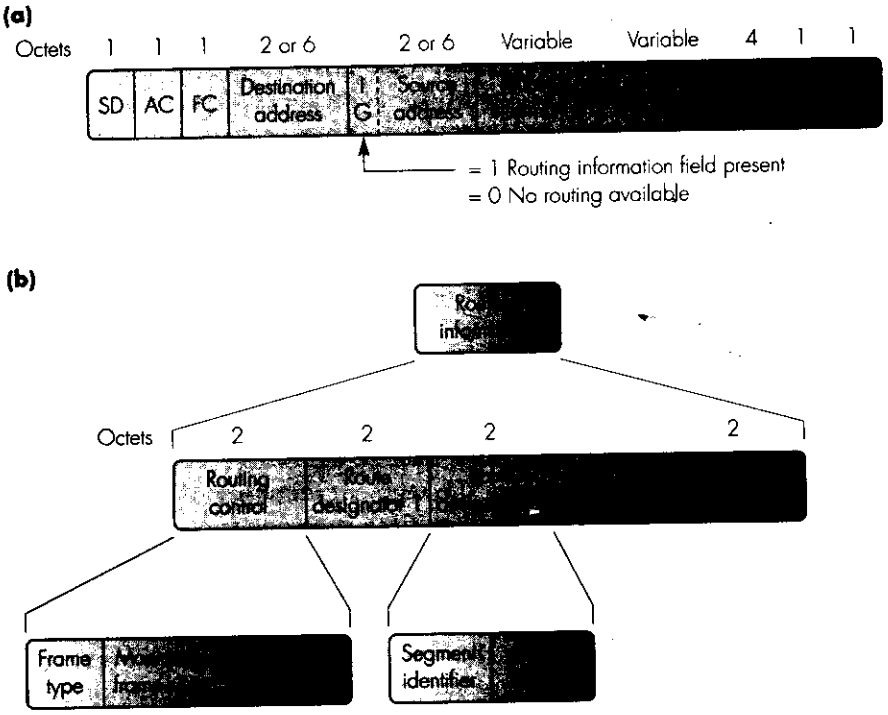


Figure 8.16 Token ring frame format: (a) position of routing information field; (b) structure of routing information field.

Since the number of segments (and bridges) traversed by a frame when going from a source to a destination may vary, the routing field length indicates the number of route designators present in the rest of the routing information field. Each route designator comprises a pair of segment and bridge identifiers.

The two additional frame types associated with the route finding algorithm are the **single-route broadcast frame** and the **all-routes broadcast frame**. To find a route, a station first creates and transmits a single-route broadcast frame with a zero routing field length and the maximum frame size set to the known largest value for the total LAN. As with spanning tree bridges, source routing bridges operate in the promiscuous mode and hence receive and buffer all frames at each of their ports. On receipt of a single-route broadcast frame, a bridge simply broadcasts a copy of the frame on each of the segments connected to its other ports. Since this procedure is repeated by each bridge in the LAN, a copy of the frame propagates throughout the LAN and is thus received by the intended destination station irrespective of the segment on which it is attached.

As we indicated earlier in Figure 8.13, if there are redundant bridges (and hence loops) in the LAN topology, multiple copies of the frame will propagate around the LAN. To prevent this, before any route finding frames are sent, the bridge ports are configured to give a spanning tree active topology. On the surface, this may appear to be the same procedure used with transparent bridges. With source routing bridges, however, the resulting spanning-tree active topology is used only for routing the initial single-route broadcast frames. This ensures that only a single copy of the frame propagates through the network. The spanning tree active topology is not used for routing either normal information frames or the all-routes broadcast frame.

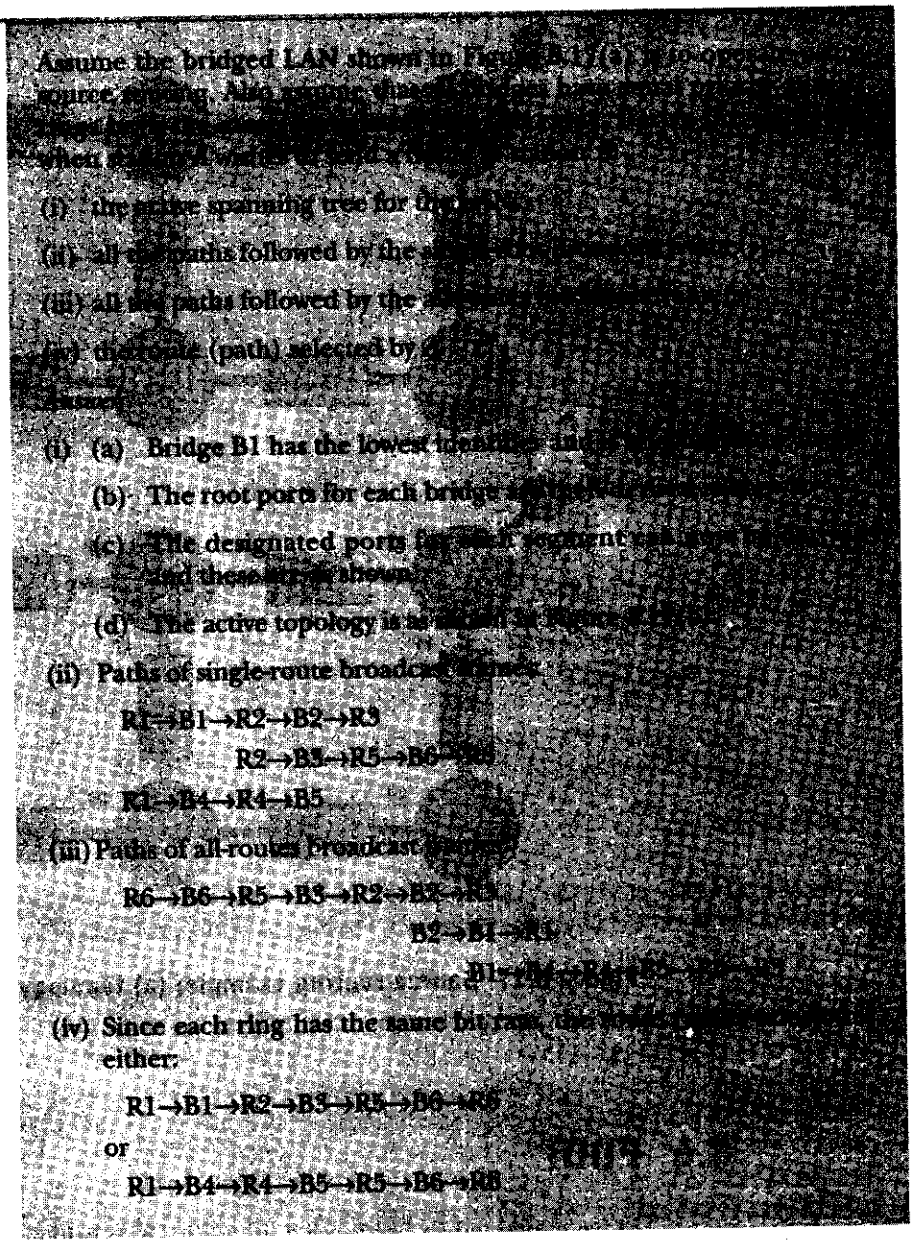
On receipt of a single-route broadcast frame, the required destination station returns an all-routes broadcast frame to the originating station. Unlike the single-route broadcast, this frame is not constrained to follow the spanning tree active topology at each intermediate bridge. Instead, on receipt of such frames, the bridge simply adds a new route designator field (comprising the segment identifier on which the frame was received and its own bridge identifier), increments the routing field length, and then broadcasts a copy of the frame on each of its other port segments.

In this way, one or more copies of the frame will be received by the originating source station via all the possible routes between the two stations. By examining the route designators in their routing control fields, the source station can select the best route for transmitting a frame to that destination. This route is then entered into its routing table and is subsequently used when transmitting any frames to that station.

Since the all-routes broadcast frame is not constrained to follow the spanning tree active topology, on receipt of such frames additional steps must be taken by each bridge to ensure that no frames are simply circulating in loops. Before transmitting a copy of the all-routes broadcast frame on an output segment, each bridge first searches the existing routing information in the frame to determine if the segment identifiers associated with the incoming and outgoing ports are already present together with its own bridge identifier. If they are, a copy of the frame has already been along the route, so this copy of the frame is not transmitted on the segment.

Note that it is not necessary to perform the route finding operation for each frame transmitted. Once a route to an intended destination has been determined and entered (cached) into the routing table of a station, this will be used for the transmission of all subsequent frames to that destination. Moreover, since most stations transmit the majority of their frames to a limited number of destinations, the number of route finding frames is relatively small compared with normal information frames for modest sized LANs.

Example 8.3



Assume the bridged LAN shown in Figure 8.17(a) is to operate using source routing. Also assume that

(i) the active spanning tree for the LAN is

(ii) all single-route broadcast frames follow the paths

(iii) all all-route broadcast frames follow the paths

(iv) the route (path) selected by source routing for frames

Solution

(i) (a) Bridge B1 has the lowest identifier and

(b) The root ports for each bridge are

(c) The designated ports for each segment are

and these are shown

(d) The active topology is as shown in Figure

(ii) Paths of single-route broadcast frames:

R1→B1→R2→B2→R3

R2→B3→R5→B6→R6

R1→B4→R4→B5

(iii) Paths of all-routes broadcast frames:

R6→B6→R5→B3→R2→B2→R3

B2→B1→R1

B1→B4→R4→B5

(iv) Since each ring has the same bit rate, the route is either:

R1→B1→R2→B3→R5→B6→R6

or

R1→B4→R4→B5→R5→B6→R6

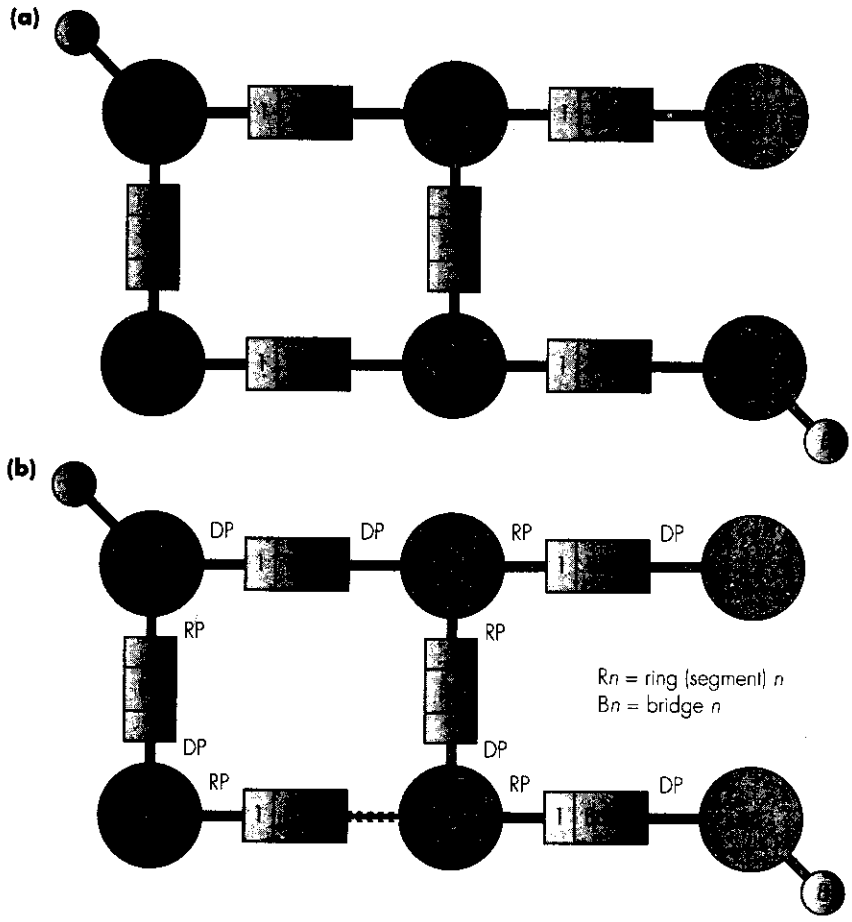


Figure 8.17 Source routing example: (a) topology; (b) spanning tree.

8.6 FDDI

As we explained in Section 8.2, in addition to bridges being introduced as a means of avoiding the traffic that is local to a segment from unnecessarily loading the entire LAN, backbone subnetworks were introduced to ensure the traffic that was forwarded between segments incurred only minimal delays. As an example, a small (but typical) establishmentwide LAN that includes bridges and backbone subnetworks is given in Figure 8.18.

Normally no end systems (workstations, servers, and so on) are connected to a backbone and they are used solely for intersegment traffic. For

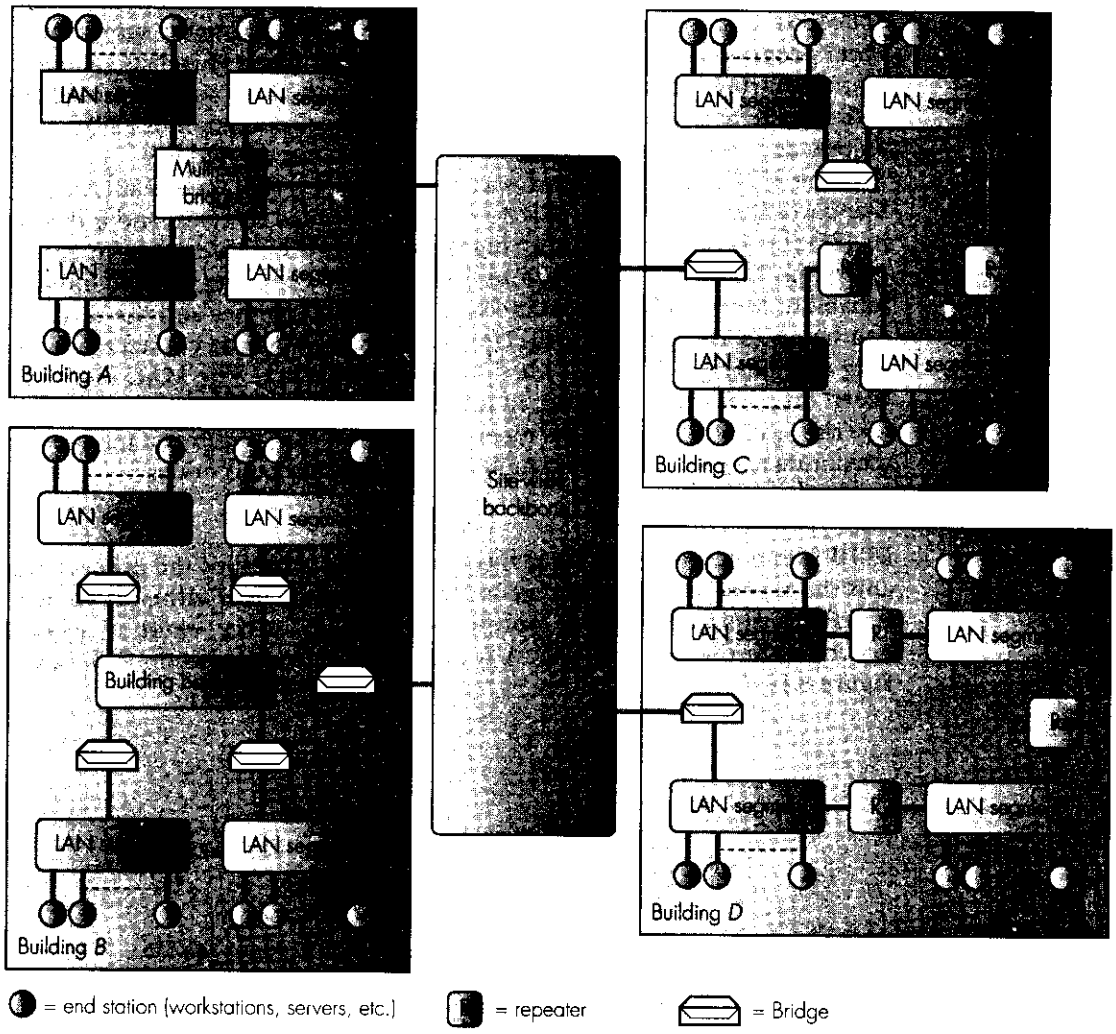


Figure 8.18 Typical establishmentwide LAN

interconnecting only a small number of segments in a single building, backbones of the same type as the interconnected segments (CSMA/CD, or token ring) are used. They are known as **building backbones**.

As the number of interconnected segments increases, there comes a point at which the transmission bandwidth required by the backbone to meet the intersegment traffic starts to exceed that available with the basic LAN types. To overcome this problem, backbones based on newer high-speed LAN types are used. An example is the **fiber distributed data interface (FDDI) LAN**. This is an optical fiber-based ring network that supports a bit rate of

100 Mbps. It can be used for the interconnection of segments spread over a wider geographical area than a single building, such as a university campus or manufacturing plant. The resulting network is then known as an **establishment** or **site backbone**. We explain the principle of operation of FDDI in this section and some other high-speed LANs in Sections 8.7 and 8.8.

The FDDI LAN standard was developed by the American National Standards Institute (ANSI). It is now an international standard and is defined in ISO 9314. It is based on a ring topology and operates at a bit rate of 100 Mbps. Because of its role, it uses dual counter-rotating rings to enhance reliability. Multimode fiber connects each station together and the total ring can be up to 100 km in length. Up to 500 stations can be connected in the ring and hence it forms an ideal backbone network. The MAC method is based on a control token and, in addition to normal data traffic, optionally the ring can also support delay-sensitive traffic that requires a guaranteed maximum access delay; for example, digitized speech and video.

Network configuration

FDDI uses two counter-rotating rings to enhance reliability: the **primary ring** and the **secondary ring**. The secondary ring can be used either as an additional transmission path or purely as a back-up in the event of a break occurring in the primary ring. A typical network configuration is shown in Figure 8.19.

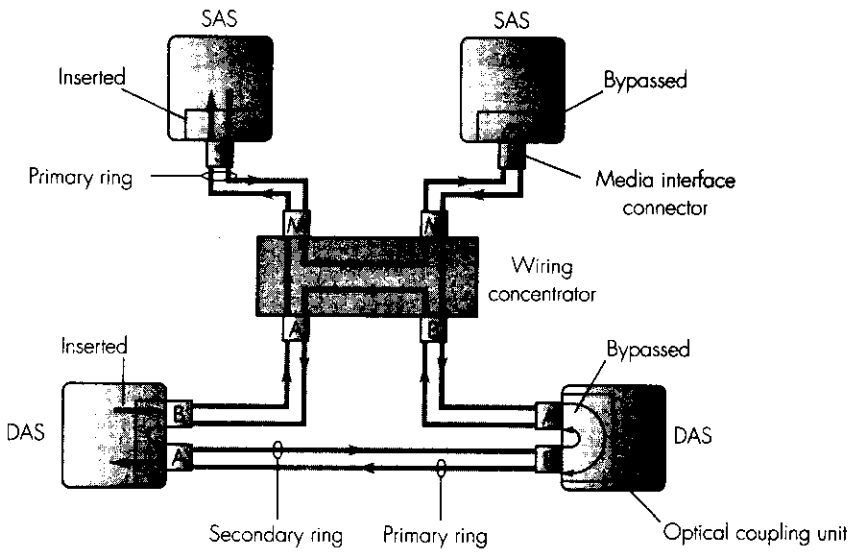


Figure 8.19 FDDI networking components.

As we can see, there are two types of station: a **dual attach station (DAS)** which is connected to both rings and a **single attach station (SAS)** which is attached only to the primary ring. In practice, most user stations are attached to the ring via **wiring concentrators** since then only a single pair of fibers is needed and the connection cost is lower.

If the LAN is being used as a backbone, then most attached stations are bridges. The protocol used to reconfigure the LAN into a single ring in the event of a ring failure is the same as the beaconing procedure associated with a token ring LAN. As we explained in Section 8.4, if a serious failure such as a broken cable arises in the ring, beacon MAC frames are issued by the station that detects the failure. Essentially, these are used to inform all other stations in the ring that the token-passing protocol has been suspended until the affected failure domain has been located and repaired. The failure domain consists of the station that detects the failure, its immediate upstream neighbor, and the ring segment in between them. The failure has deemed to be repaired when the station that issues the beacon frames starts to receive them after rotating around the ring; that is, on receipt of a beacon frame with its own MAC address at the head of the frame.

An example of a failure domain is shown in Figure 8.20(a) and a redundant ring configuration in Figure 8.20(b). In this example we assume a break has occurred in the ring segment between stations *F* and *G*. Hence *G* is the beaconing station and *F* its upstream neighbor. When a redundant ring is being used, the TCU not only supports the functions we explained in Section 8.4, but also the means to bypass a faulty ring segment or station. As an example, we show in Figure 8.20(c) how the faulty ring segment (failure domain) we identified in Figure 8.20(a) is bypassed.

Essentially, once the failure domain has been located and reported, the relays in the TCU of *F* and *G* are activated to (hopefully) re-establish a continuous ring. If isolating the suspected faulty segment does not remove the fault, the next step is to initiate the isolation of station *G* completely, as shown in Figure 8.20(d). Note from these illustrations that the redundant ring does not have a direct path to the MAC unit and simply provides a means of bypassing a section of the ring. The order of the stations in a re-established ring is the same as that in the original ring.

The basic fiber cable is **dual core** with **polarized duplex** (two-position) connectors at each end. This means that each end of the cable has a different physical key so that it can be connected into a matching socket only. This prevents the transmit and receive fibers from becoming inadvertently interchanged and bringing down the total network. As a further precaution, we use different connectors to connect each station type – SAS and DAS. In common with the basic token ring, we use special coupling units to isolate (bypass) a station when its power is lost. With FDDI, these are either active or passive fiber devices.

Although the topology is logically a ring, physically it is normally implemented in the form of a hub/tree structure. An example is shown in Figure 8.21(a). To ensure that changes to the wiring are carried out in a

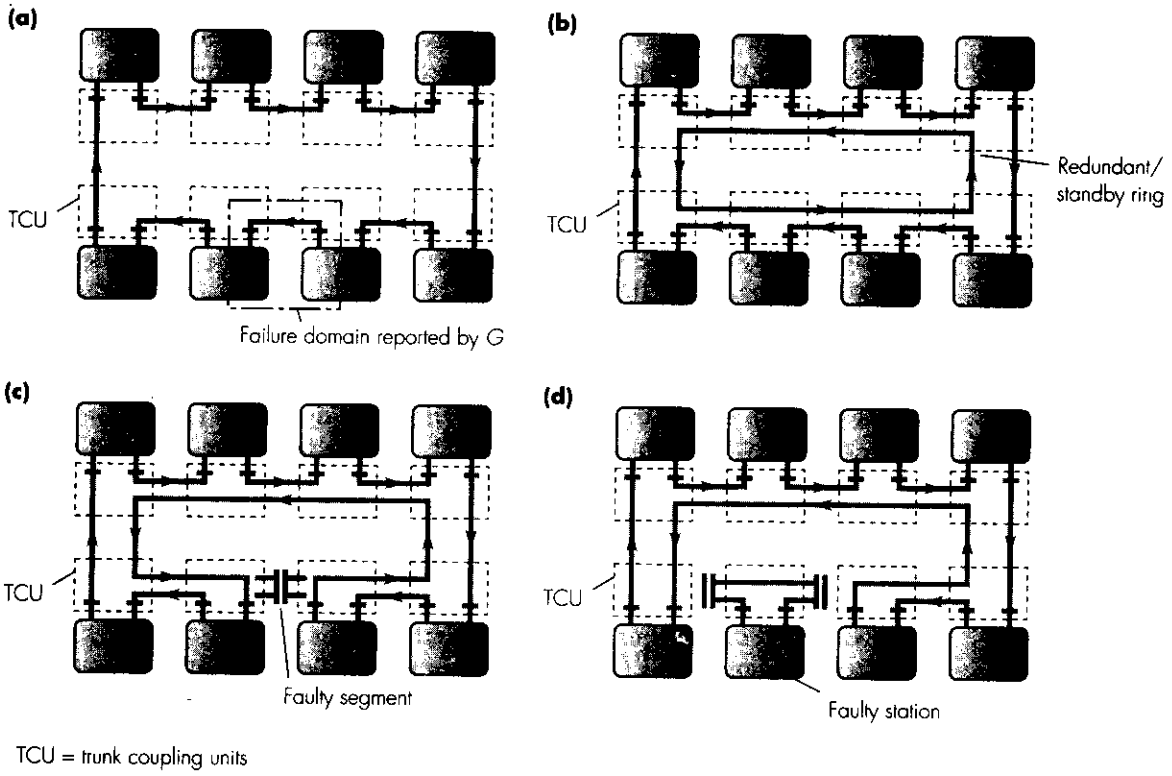


Figure 8.20 Ring fault detection and isolation: (a) failure detection; (b) redundant ring configuration; (c) segment isolation; (d) station isolation.

controlled way, a combination of **patch panels** and **wiring concentrators** are used. Typically, these are located in the wiring closet associated with either a floor (if a local FDDI ring is being used) or a building (if the ring is a backbone). In the latter case, the patch panels are interconnected as shown in Figure 8.21(b) to create a tree structure to connect each station – high-speed servers or bridges – to the ring.

Each patch panel has a number of possible attachment points associated with it. In the absence of a connection at a particular point, the ring is maintained using short **patch cables**, each with the same type of connector. Adding a new station or concentrator simply involves removing a patch cable and replacing it with a corresponding drop cable. This approach is known as **structured wiring**.

Physical interface

The physical interface to the fiber cable is shown in Figure 8.22. In a basic token ring network, at any instant, there is a single active ring monitor which,

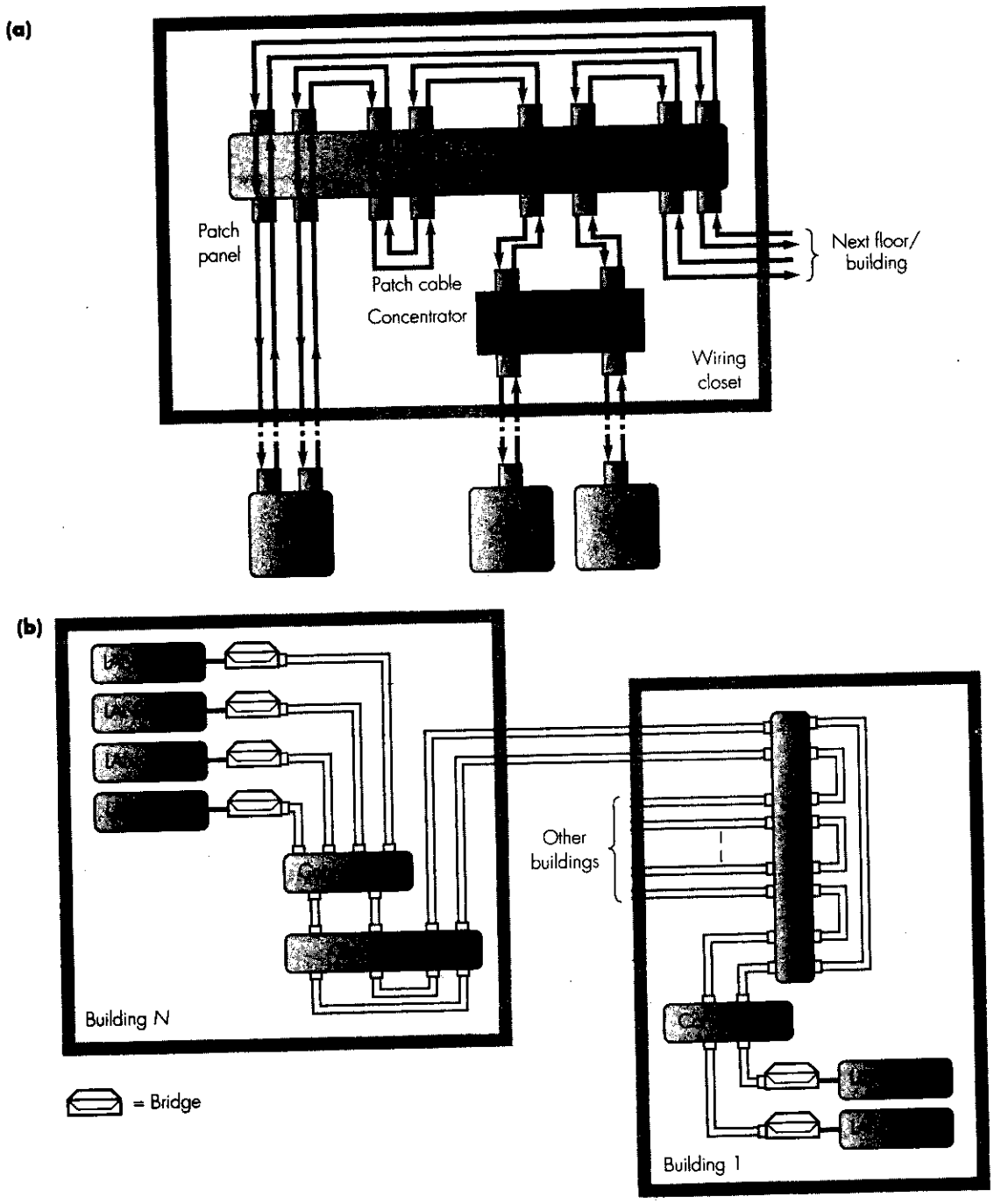


Figure 8.21 FDDI wiring schematic: (a) building; (b) establishment.

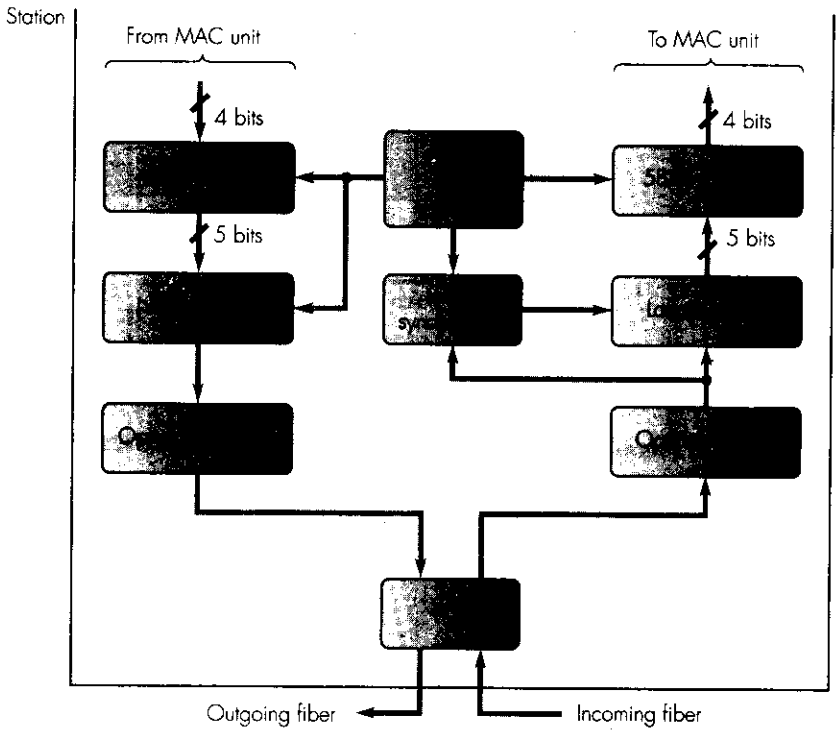


Figure 8.22 FDDI physical interface schematic.

among other things, supplies the master clock for the ring. Each circulating bitstream is encoded by the active ring monitor using differential Manchester encoding. All the other stations in the ring then frequency and phase lock to the clock extracted from this bitstream. However, such an approach is not suitable at the bit rates of an FDDI ring since this would require a signaling rate of 200 Mbaud. Instead, each ring interface has its own local clock. Outgoing data is transmitted using this clock while incoming data is received using a clock that is frequency and phase locked to the transitions in the incoming bitstream. As we shall see, all data is encoded prior to transmission so that there is a guaranteed transition in the bitstream at least every two bit cell periods, ensuring that each received bit is sampled (clocked) very near to the nominal bit cell center.

All data to be transmitted is first encoded, prior to transmission, using a **4 of 5 group code**. This means that for each 4 bits of data to be transmitted, a corresponding 5-bit codeword or symbol is generated by what is known as a **4B5B encoder**. The 5-bit symbols corresponding to each of the sixteen possible 4-bit data groups are shown in Figure 8.23(a). As we can see, there is a maximum of two consecutive zero bits in each symbol. The symbols are then shifted

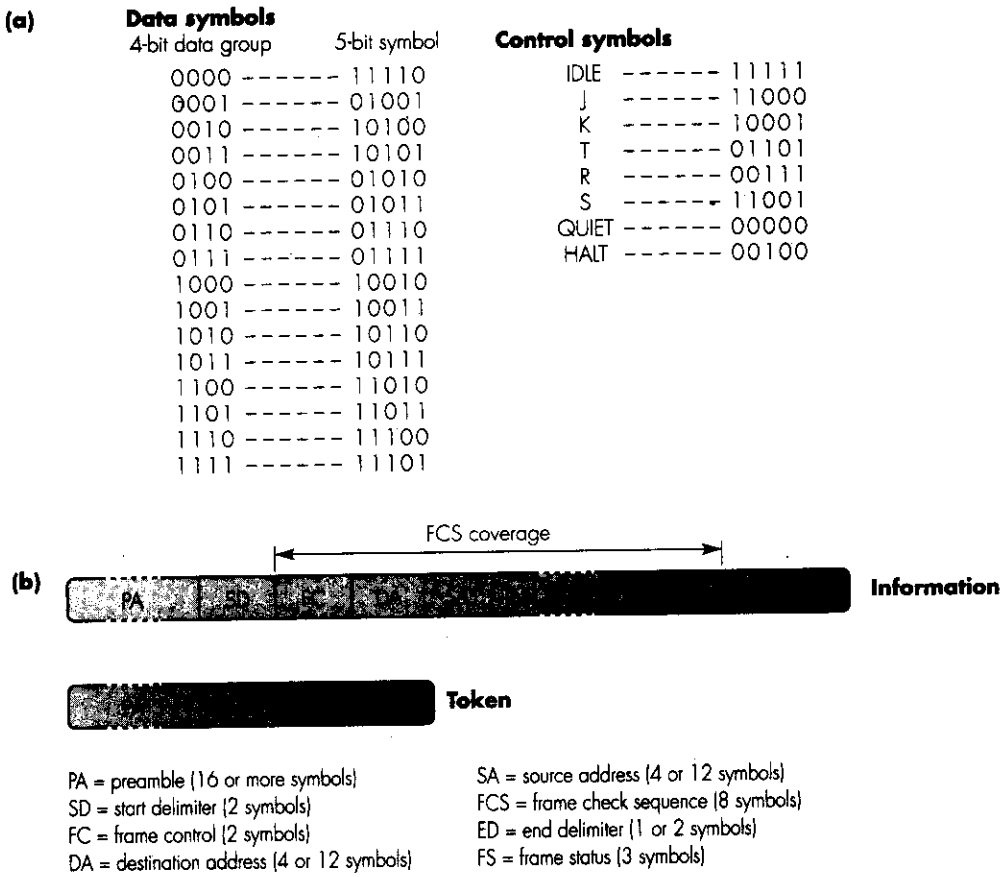


Figure 8.23 FDDI line coding and framing detail: (a) 4B5B codes; (b) frame formats.

out through a further NRZI encoder which produces a signal transition whenever a 1 bit is being transmitted and no transition when a 0 bit is transmitted. In this way, there is a guaranteed signal transition at least every two bits.

The use of 5 bits to represent each of the sixteen 4-bit data groups means that there are a further sixteen unused combinations of the 5 bits. Some of these combinations (symbols) are used for other (link) control functions, such as indicating the start and end of each transmitted frame or token. A list of the link control symbols is shown in Figure 8.23(a) and part (b) shows the format used for frames and tokens. In general, the meaning and use of each field is the same as with the basic token ring but, because symbols are used rather than bits, there are some differences in the structure of each field.

The *preamble* (*PA*) field consists of 16 or more IDLE symbols which, since they each consist of five 1 bits, cause the line signal to change at the maximum frequency. The line signal transitions are used for establishing (and

maintaining) clock synchronization at the receiver. The *start delimiter* (*SD*) field consists of two control symbols (J and K) which enable the receiver to interpret the following frame contents on the correct symbol boundaries. The *FC*, *DA*, and *SA* fields have the same meaning as before, but the (decoded) information field in the data frames can be up to 4500 octets with FDDI. The *end delimiter* (*ED*) field contains one or two control symbols (*T*). Finally, the frame status (*FS*) field, although it has a similar function to the FS field in the basic ring, consists of three symbols that are combinations of the two control symbols R and S.

The local clock used in the physical interface is 125 MHz which, because of 4B5B encoding, yields a data rate of 100 Mbps and a signaling rate of 125 Mbaud. Since all transmissions are encoded into 5-bit symbols, each 5-bit symbol must first be buffered at the receiver before it can be decoded. However, the use of two symbols (J and K) for the SD field to establish correct symbol boundaries means that a 10-bit buffer is used at the receiver. This is known as the **latency** (or elastic) **buffer** since it introduces 10 bits of delay – latency – into the ring. At 125 Mbaud, this is equivalent to a delay of 0.08 μs

Example 8.4