CN UNIT – IV

The Medium Access Control Sublayer

The Channel Allocation Problem:

The central theme of this chapter is how to allocate a single broadcast channel among competing users. The channel might be a portion of the wireless spectrum in a geographic region, or a single wire or optical fiber to which multiple nodes are connected. It does not matter. In both cases, the channel connects each user to all other users and any user who makes full use of the channel interferes with other users who also wish to use the channel.

Static channel allocation:

The traditional way of allocating a single channel, such as a telephone trunk, among multiple competing users is to chop up its capacity by using one of the multiplexing schemes, such as FDM (Frequency Division Multiplexing). If there are N users, the bandwidth is divided into N equal-sized portions, with each user being assigned one portion. Since each user has a private frequency band, there is now no interference among users. When there is only a small and constant number of users, each of which has a steady stream or a heavy load of traffic, this division is a simple and efficient allocation mechanism. A wireless example is FM radio stations. Each station gets a portion of the FM band and uses it most of the time to broadcast its signal. However, when the number of senders is large and varying or the traffic is bursty, FDM presents some problems. If the spectrum is cut up into N regions and fewer than N users are currently interested in communicating, a large piece of valuable spectrum will be wasted. And if more than N users want to communicate, some of them will be denied permission for lack of

bandwidth, even if some of the users who have been assigned a frequency band hardly ever transmit or receive anything. Even assuming that the number of users could somehow be held constant at N, dividing the single available channel into some number of static subchannels is inherently inefficient. The basic problem is that when some users are quiescent, their bandwidth is simply lost. They are not using it, and no one else is allowed to use it either. A static allocation is a poor fit to most computer systems, in which data traffic is extremely bursty, often with peak traffic to mean traffic ratios of 1000:1. Consequently, most of the channels will be idle most of the time. The poor performance of static FDM can easily be seen with a simple queueing theory calculation. Let us start by finding the mean time delay, T, to send a frame onto a channel of capacity C bps. We assume that the frames arrive randomly with an average arrival rate of λ frames/sec, and that the frames vary in length with an average length of 1/μ bits. With these parameters, the service rate of the channel is μC frames/sec. A standard queueing theory result is

$$T = \frac{1}{\mu C - \lambda}$$

(For the curious, this result is for an "M/M/1" queue. It requires that the randomness of the times between frame arrivals and the frame lengths follow an exponential distribution, or equivalently be the result of a Poisson process.) In our example, if C is 100 Mbps, the mean frame length, 1/μ, is 10,000 bits, and the frame arrival rate, λ, is 5000 frames/sec, then T = 200 μsec. Note that if we ignored the queueing delay and just asked how long it takes to send a 10,000bit frame on a 100-Mbps network, we would get the (incorrect) answer of 100 μsec. That result only holds when there is no contention for

the channel. Now let us divide the single channel into N independent subchannels, each with capacity C/Nbps. The mean input rate on each of the subchannels will now be λ/N. Recomputing T, we get

$$T_N = \frac{1}{\mu(C/N) - (\lambda/N)} = \frac{N}{\mu C - \lambda} = NT$$

The mean delay for the divided channel is N times worse than if all the frames were somehow magically arranged orderly in a big central queue. This same result says that a bank lobby full of ATM machines is better off having a single queue feeding all the machines than a separate queue in front of each machine. Precisely the same arguments that apply to FDM also apply to other ways of statically dividing the channel. If we were to use time division multiplexing (TDM) and allocate each user every N^{th} time slot, if a user does not use the allocated slot, it would just lie fallow. The same would hold if we split up the networks physically. Using our previous example again, if we were to replace the 100-Mbps network with 10 networks of 10 Mbps each and statically allocate each user to one of them, the mean delay would jump from 200 μsec to 2 msec. Since none of the traditional static channel allocation methods work well at all with bursty traffic, we will now explore dynamic methods.

Assumptions for Dynamic Allocation:

These are the five key assumptions of dynamic allocation:

1. **Independent Traffic.**
   The model consists of N independent stations (e.g., computers, telephones), each with a program or user that generates frames for transmission. The expected number of frames generated in an interval of length Δt is λΔt, where λ is a constant (the arrival rate of new frames). Once a frame has

been generated, the station is blocked and does nothing until the frame has been successfully transmitted.

2. **Single Channel.**
   A single channel is available for all communication. All stations can transmit on it and all can receive from it. The stations are assumed to be equally capable, though protocols may assign them different roles (e.g., priorities).
3. **Observable Collisions.**
   If two frames are transmitted simultaneously, they overlap in time and the resulting signal is garbled. This event is called a collision. All stations can detect that a collision has occurred. A collided frame must be transmitted again later. No errors other than those generated by collisions occur.
4. **Continuous or Slotted Time.**
   Time may be assumed continuous, in which case frame transmission can begin at any instant. Alternatively, time may be slotted or divided into discrete intervals (called slots). Frame transmissions must then begin at the start of a slot. A slot may contain 0, 1, or more frames, corresponding to an idle slot, a successful transmission, or a collision, respectively.
5. **Carrier Sense or No Carrier Sense.**
   With the carrier sense assumption, stations can tell if the channel is in use before trying to use it. No station will attempt to use the channel while it is sensed as busy. If there is no carrier sense, stations cannot sense the channel before trying to use it. They just go ahead and transmit. Only later can they determine whether the transmission was successful.

Multiple Access Protocols:

MUltipleAccess

• One feature of LAN is that its backbone is a shared channel or transmission link, which provides all user to access to the transmission facilities. It may be possible that two or more stations transmitting simultaneously, causing their signals to interfere and becomes garbled.

•To resolve these conflict, a number of different control mechanisms or access protocol have been devised. In order to handle the burstynature of LAN, traffic asynchronous TOM is used.

• The asynchronous TDM mechanism is further divided into contention methods (random access) and deterministic methods (controlled methods).

Random access techniques are

i) ALOHA

ii) Carrier sense multiple-access (CSMA)

iii) CSMA with collision-detection (CSMA/CD)

iv) Register insertion.

Controlled access to LAN can be performed in two types :

1) **Centralized technique** : In centralized technique master node decides which node is to access the channel at anyone time.

e.g. polling.

2)**Distributed technique** ; In distributed technique each station is given an opportunity to transmit on the channel.

e.g. i) Token passing method ii) Slotted ring method.

• Fig. 5.1.1 illustrates the typical multiple access communications where a number of user stations share a transmission medium.

• This sharing techniques are used in wired communications, and networks based on radio communication.

• In wired communication multidrop cables are used in data networks to connected number of stations to a host computer.

• The host computer broadcasts information to the users on the outbound line.

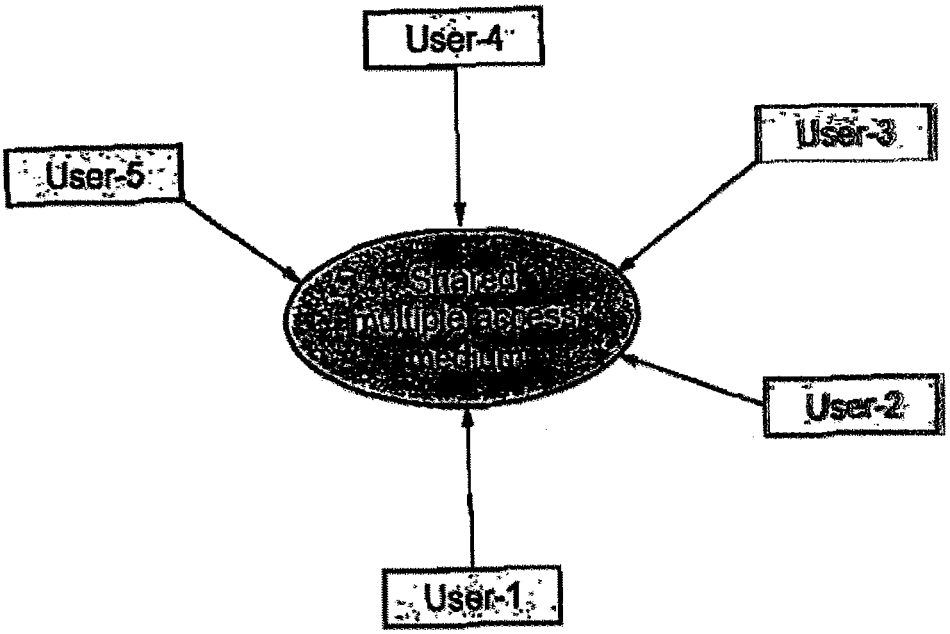


Fig. 5.1.1 Multiple access communication

The stations transmit information to the host using the inbound line. This system is illustrated in Fig. 5.1.2.

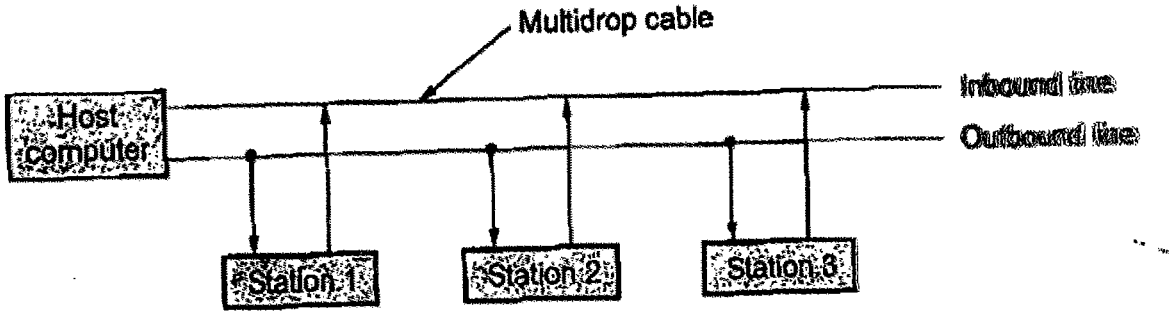
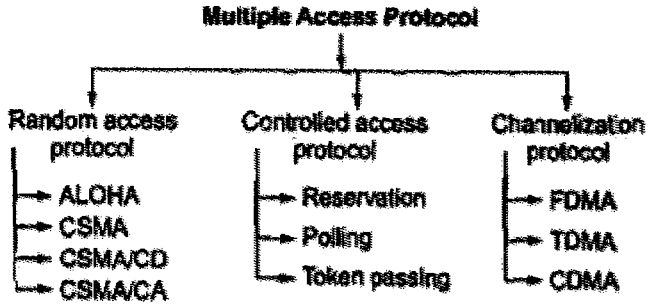


Fig. 5.1.2 Multidrop cable system for access control

- A medium access control (MAC) protocol is developed for this system. Here the host computer issues polling messages to each stations, providing it with permission to transmit on the inbound line.
- In radio communication several stations share two frequency bands, one for transmitting and one for receiving.
- In satellite communications each station is assigned a channel in an uplink frequency band that it uses to transmit to the satellite. The satellite sends back the signals on different frequency band called down link frequency band.



Random Access :

• Access to the medium from many entry points is called contention. It is controlled with a contention protocol.

•In a random access method, each station has the right to the medium without being controlled by any other station. However, if more than one station tries to send, there is an access conflict i.e. collision and the frames will be either destroyed or modified.

ALOHA:

• The ALOHA protocol was developed at the university of Hawaii in the early 19705. ALOHA was developed for packet radio networks. However, it is applicable to any shared transmission medium.

•In a system when multiple users try to send messages to other stations though a common broadcast channel random access or contention techniques are used.

•Random access means there is no definite or scheduled time for any station to transmit. This scheme L'5 simplest possible and it is asynchronous. It is asynchronous because there is no co-ordination among users.

•The basic idea of ALOHA system is applicable to any system in which uncoordinated users are competing for the use of a single shared channel.

•When a station send data, another station may attempt to do so at the same time. The data from the two station collide and become garbled. If two signals collided, so be it. Each station would simply wait a random lime and try again.

.Pure ALOHA :

• The Original ALOHA protocol is called pure ALOHA, The idea is that each station sends a frame whenever it has a frame to send., Since there is only one channel to sharer there is the possibility of collision between frames from different stations.

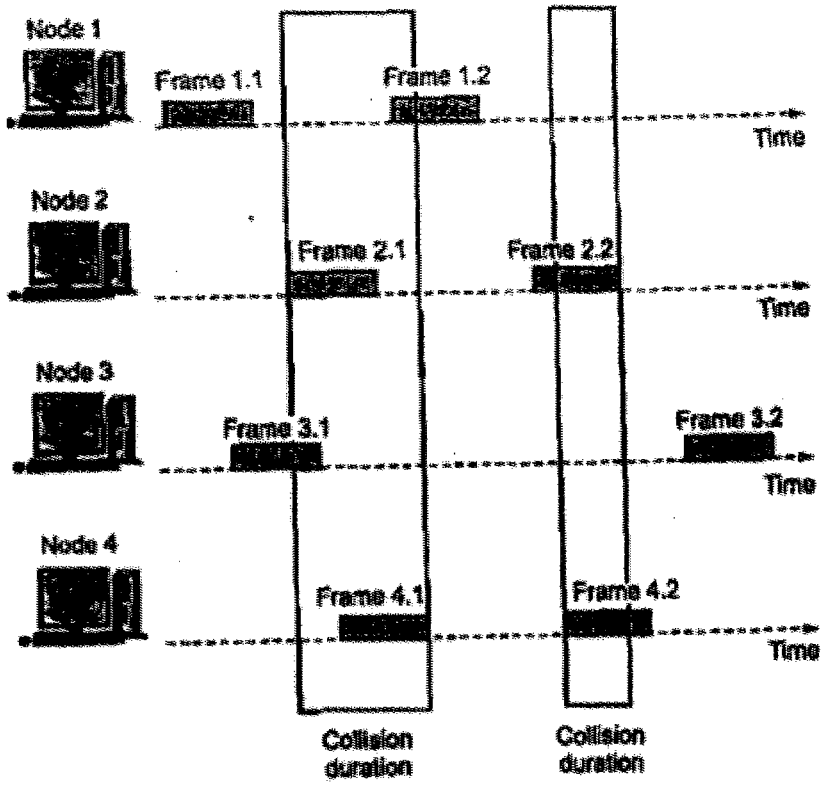


Fig. 5.2.1 Frames in pure ALOHA

- The pure ALOHA protocol relies on acknowledgments from the receiver. When a user sends a frame, it expects the receiver to send an acknowledgement. If the acknowledgement does not arrive after a time out period, the station assumes that the frame has been destroyed and resends the frame.
- Whenever two frames try to occupy the channel at the same time, there will be a collision and both will be garbled. If the first bit of a new frame overlaps with Just the last bit of a frame almost finished, both frames ,will be totally destroyed and both will have to be retransmitted later.
- If all users try to resend their frames after the time-out, the frames will collide again. Pure ALOHA dictates that when the time-out period passes, each user waits .a random amount of time before resending its frame. The randomness will help to avoid more collisions. This time is called as back-off time .

Fig. 5.2.2 shows the procedure for pure ALOHA protocol.

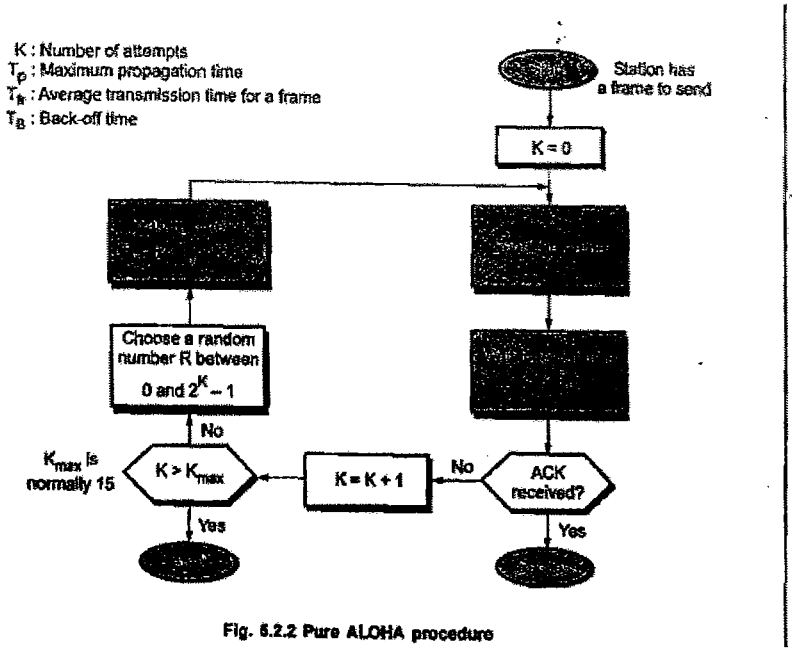


Fig. 6.2.2 Pure ALOHA procedure

- The time out period is equal to the maximum possible round trip propagation delay, which is twice the amount of time required to send a frame between the two most widely separated stations .

• Let all the packets have the same length. And each requires one time unit for transmission (t_p)' Consider any user to send packet A at time 1 , If any other 0 user B has generated a packet between time t_0 and t_0 +t_p, the end of packet B p will collide with the beginning of packet A. Since in pure ALOHA packet, a station does not listen to the channel before transmitting, it has no way of knowing that above frame was already under way.

- Following Fig shows vulnerable periods during which packets can collide

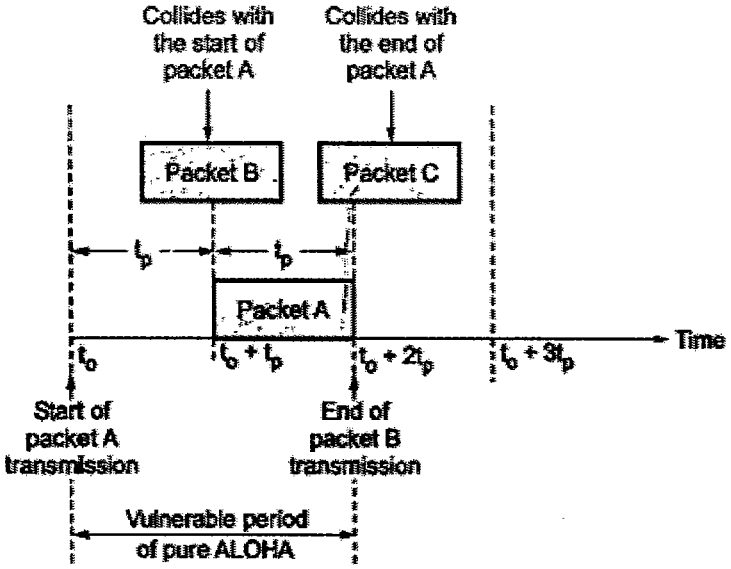


Fig. 5.2.3 Vulnerable period for packet A

Similarly if another user wants to transmit between ($t_o + t_p$) and (to $+2t_p$) i.e. packet C, the beginning of packet C will collide with the end of packet A. Thus if two packets overlap by even the smallest amount in the vulnerable period both packets will be corrupted and need to be retransmitted.

Throughput of pure ALOHA channel :

1) Throughput:

• The throughput S is defined as average successful traffic transmitted between stations per unit time. The unit of time is slot-time, which is the time required to transmit a frame.

•Assuming all packets or frames are of same size. Since only one packet per slot can be transmitted, the maximum value of S is 1. When collision occurs, some of the packets are lost and part of available channel time is wasted. The resulting value of S is lessthan1.

$S = G \times e^{-2G}$

2) **Offered traffic'**

- The offered traffic is the average number of packets per slot time which are presented to the network for transmissions by users. It is denoted by G. The throughput is expressed in terms of offered load or traffic G. Practically G can have any value between 0 to infinity.

3) **Channel capacity :**

- The maximum achievable throughput for a particular type of access scheme is called the capacity of the channel.

- To find the throughput of channel, let us assume that the probability (Pk) that k packets generated during a given slot-time follows a Poisson's distribution with a mean G per packet time is given by

$$P_k = \frac{G^k \cdot e^{-G}}{k!}$$

- The throughput S is then just the offered load G times the probability of a transmission being successful.

$S = Gp_o$

where P_o = Probability that a packet does not suffer a collision

- The probability of no other traffic being initiated during the entire vulnerable period is thus given by

$$p_o = e^{-2G}$$

$$S = G \times e^{-2G}$$

The maximum throughput occurs at S=1/2e =0.184

- This means that the best channel utilization that can be achieved is around 18 % for pure ALOHA method.

- The advantage of pure ALOHA protocol is its Simplicity, which can result in low cost user stations since no synchronization is required between stations in the system, Each station transmits a packet whenever its buffer has one, The disadvantage is somewhat inefficient channel utilization i.e. maximum channel utilization is only 18.4 % of the available capacity.

Slotted ALOHA :

- In slotted ALOHA, the channel time is divided into time slots and the stations are allowed to transmit at specific instance of time. These time slots are exactly equal to the packet transmission time. All users are then synchronized to these time slots, so that whenever a user generates a packet it must synchronize exactly with the next possible channel slot. Consequently the wasted time due to collisions can be reduced to one packet time or vulnerable period is reduced to half.

- Transmission attempts for four network user and random retransmission delays for colliding packets in slotted ALOHA is shown in Fig. 5.2.4

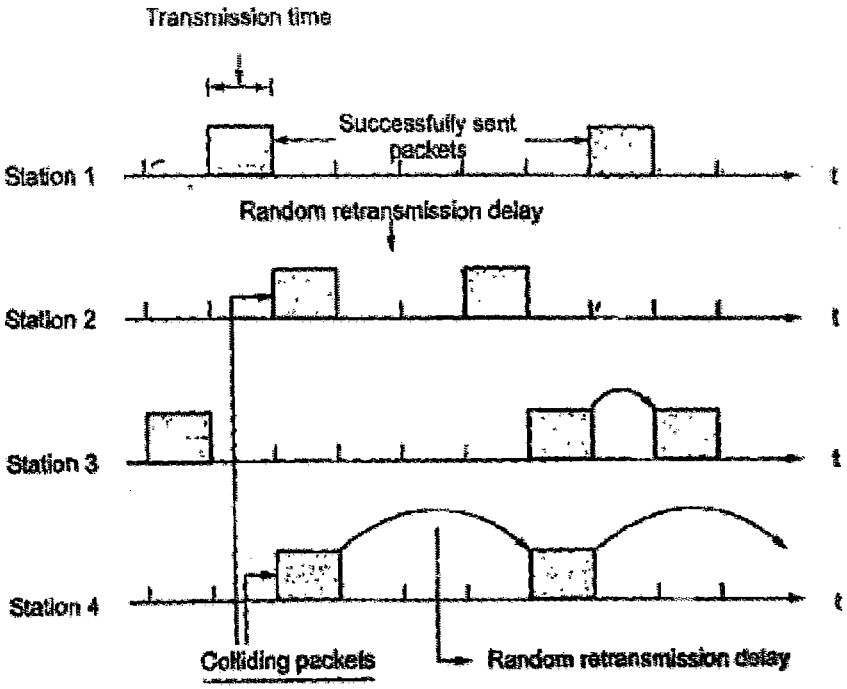


Fig. 5.2.4 Transmission attempts and random retransmission delays for colliding packets in slotted ALOHA

Slotted ALOHA

Assumptions :

1. All frames are of same size.
2. Time is divided into equal sized slots, a slot equals the time to transmit one frame.
3. Nodes start to transmit frames only at beginning of slots.
4. Nodes are synchronized.
5. If two or more nodes transmit in a slot, all nodes detects collision before the slot ends.

Throughput of slotted ALOHA channel

- In slotted ALOHA, the packets arrive in a synchronized fashion.

The probability of single transmission during a slot time is

$p_0 = e^{-2G}$

From equation

$S = G \cdot e^{-G}$

which is twice that 01 pure ALOHA. This means that the best channel utilization that can be achieved is around 37 %.

- The relation between the offered traffic and the throughput is shown in fig

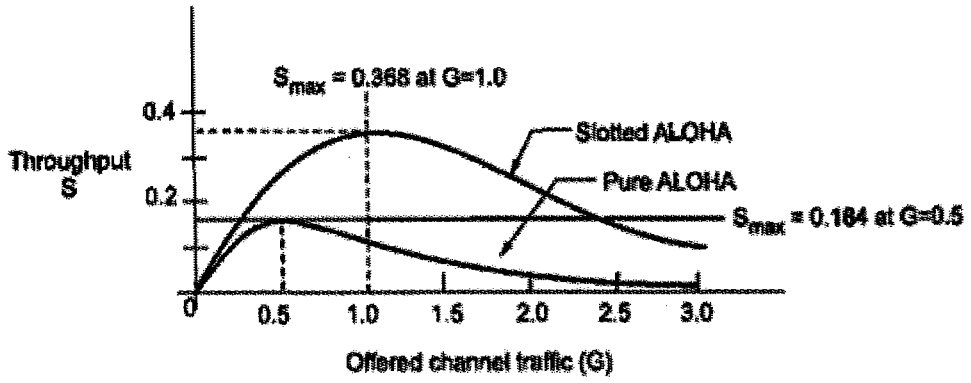


Fig. 5.2.5 Comparison of the throughput as a function of offered load for pure and slotted ALOHA

Pros and Cons of slotted ALOHA :

Pros

I. Single active node can continuously transmit at full rate of channel.

2.Highly decentralized, each node independently decides when to retransmit.

3. Simple to implement.

cons

I. Collisions waste slots.

2. Idle slots.

Carrier Sense Multiple Access (CSMA) :

- The low maximum throughput of the ALOHA schemes is due to the wastage of transmission bandwidth because of frame collisions. This wastage can be reduced by avoiding transmissions that are certain to cause collisions, By sensing the medium for the presence of a carrier signal from other stations, a stations can determine whether there is an on going transmission.

•CSMA requires that each station first listen to the medium before sending. CSMA can reduce the possibility of collision, but it cannot eliminate it.

•The possibility of collision still exists because of propagation delay. A station may sense the medium and find it idle, only because the first bit sent by another station has not yet been received.

Vulnerable Time :

•The vulnerable time for CSMA is the propagation time (Tp). This is the time needed for a signal to propagate from one end of the medium to the other.

• Fig. 5.2.6 shows the vulnerable time for CSMA. Station A sends. frame at time t11 which reaches the rightmost station 0 at time

t1 + T p

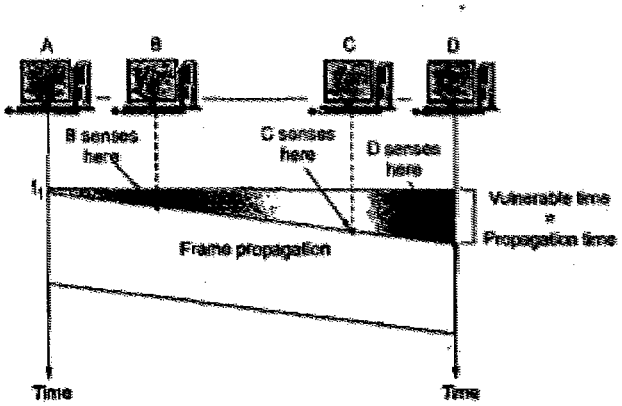


Fig. 5.2.6 Vulnerable period in CSMA

Persistence Methods

•These three protocols are

1) Non-persistent CSMA. 2) I-persistent CSMA. 3) p-persistent CSMA.

All three protocols are differing by the action to be taken by any station after sensing the readiness of the channel.

1) Non-persistent CSMA :

•In non-persistent CSMA, when a station having a packet (frame) to transmit and finds that the channel is busy, it backs off for. fixed interval of time. It then checks the channel .gain and if the channel is free then it transmits. The b.ck-off delay is determined by the transmission time of a frame, propagation lime and ·other system parameters. If the channel is already in use, the station does not continuously senses it for the purpose of seizing it immediately upon detecting the end of the previous transmission. But it waits a random period of time and again checks for activity.

Fig. 5.2.7 shows flow diagram for non-persistent CSMA.

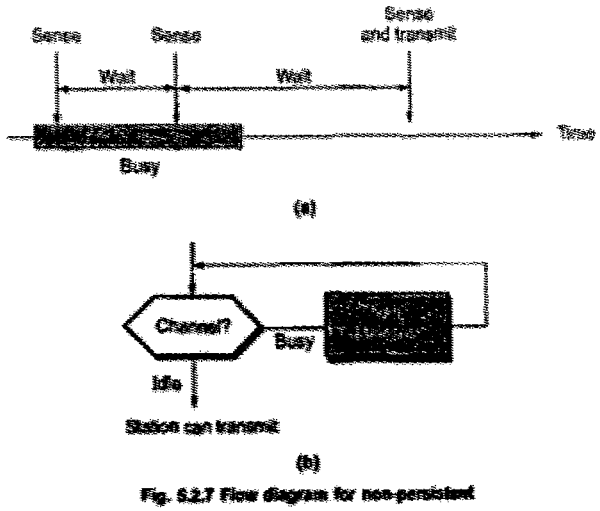


Fig. 5.2.7 Flow diagram for non-persistent

2) 1-persistent CSMA :

• Any station wishing to transmits, monitor the channel continuously until the channel is idle and then transmits immediately with probability one, hence the name l-persistent.

•When two or more stations are waiting to transmit, a collision is guaranteed. Since each station will transmit immediately at the end of busy period. In this case each will wait a random amount of time and will then reattempt to transmit.

• As in the case with non-persistence CSMA, the performance of 1-persistent CSMA protocol depends on the channel delay time.

Fig. 5.2.8 shows the flow diagrams for 1-persistent CSMA.

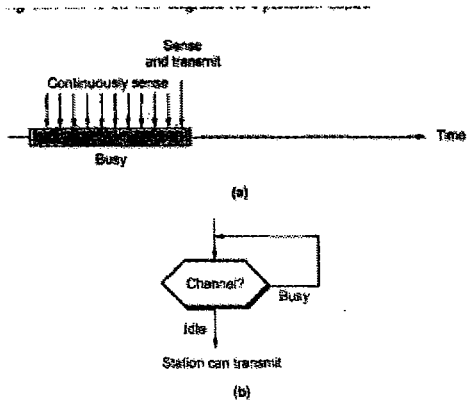


Fig. 5.2.8 Flow diagrams for 1-persistent CSMA

3) **p-persistent CSMA :**

• To reduce the probability of collision in 1-persistent CSMA, not all the waiting stations are allowed to transmit immediately, after the channel is idle.

• When a station becomes ready to send and it senses the channel to be idle, it either transmits with a probability p or it defers transmission by one time slot with a probability q= 1-p. If the deferred slot is also idle, the station either transmits with probability p or defers again with a probability q. This process is repeated until either packets are transmitted or the channel becomes busy.

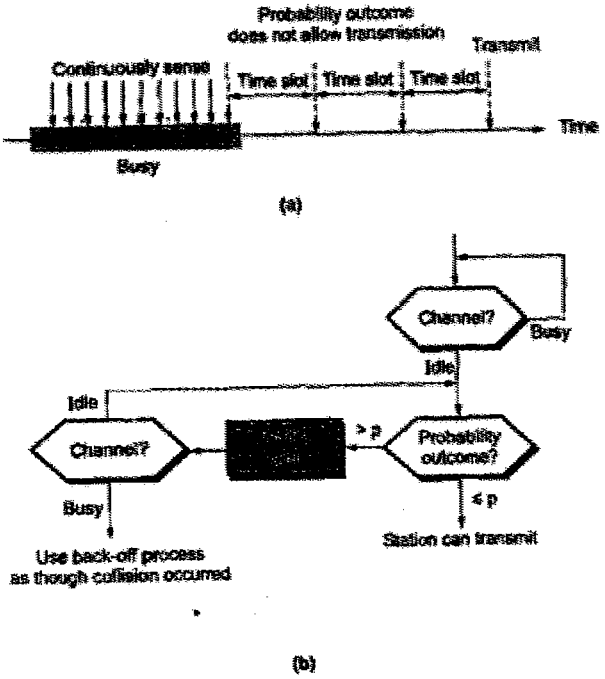


Fig. 5.2.9 Flow diagrams for p-persistent CSMA

Carrier Sense Multiple Access with Collision Detection:

- In both CSMA and ALOHA schemes, collisions involve entire frame transmissions. If a station can determine whether a collision is taking place, then the amount of wasted bandwidth can be reduced by aborting the transmission when a collision is

detected. The carrier sense multiple access with collisiondetection (CSMA/CD) use this approach.

- CSMA/CD is the most commonly used protocol for LANs. CSMA/CD specifications were developed jointly by Digital Equipment Corporation (DEC),Intel, and Xerox. This network is called as Ethernet. The IEEE 802.3 CSMA/CD standard for LAN is based on Ethernet specification.

• The basic protocol is that, a station with a message to send must monitor the channel to see if any other station is sending. If another station is sending, the' second station must wait or defer, until the sending station has finished. Then it may send its message. If no station was sending at the time that it first listened, the station may send its message immediately. The term "carrier sense" indicates this "listening before transmitting" behaviour.

• If two or more stations have messages to send at the same time and they are separated by significant distances on the bus/channel,.each may begin transmitting at roughly the same time without being aware of the other station. The signals from each station will superimpose on the channel and is garbled beyond the decoding ability of the receiving station. This is termed as "collision".

•A protocol is required for transmitting station to monitor the channel while sending each of its messages and to detect such '·collisions".

• When a collision has been detected, each of sending stations must cease transmitting, wait for a random length of time, and then try again. Because of quick termination of transmission time and bandwidth is saved. Therefore CSMA/CD is more efficient than ALOHA, slotted ALOHA and CSMA.

• CSMA/CD networks work best on a bus, multipoint topology with bursty asynchronous transmission. All stations are attached to one path and monitor the signal on the channel through transreceiver attached to the cable.

• CSMA/CD has totally decentralized control and is based on contention access.

Fig.5.2.10 illustrates this technique. Station A and station D are the extreme ends of a bus structure. (Refer Fig. 5.2.10 on next page)

i) Station A listens channel starts transmitting a packet addressing D.

ii) Station B and C are ready for transmission. B senses a transmission on channel so defers. C is unaware of transmission and begins its own transmission.

iii) Station A's transmission reaches C. C detects collision and ceases transmission. Sends jam signal.

iv) Effect of collision propagates back to A, A stops its transmission.

v) A sends jam signal.

vi) No station is transmitting but there are still signals on the bus.

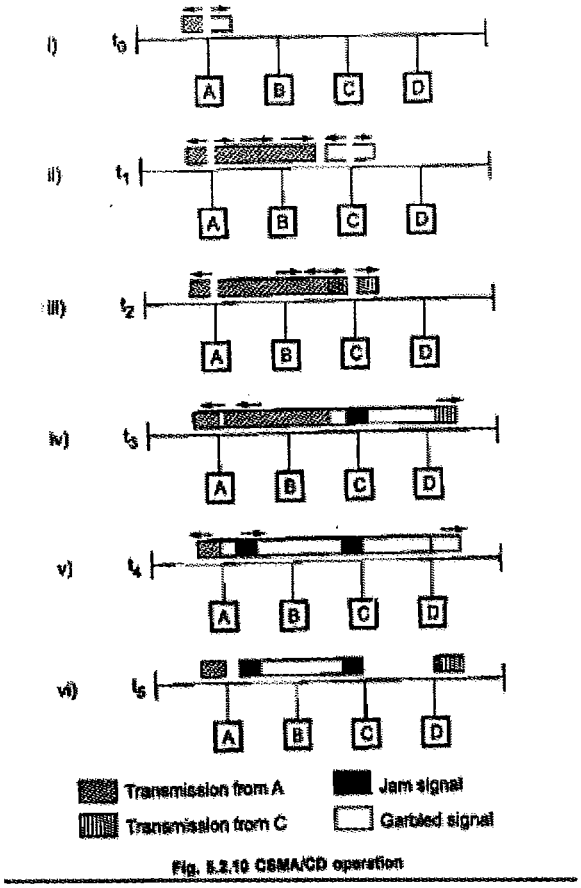


Fig. 5.2.10 CSMA/CD operation

- CSMA/CD supports both baseband and broadband system.

CSMA/CD offers four options in terms of bit rate, signaling method and maximum electrical cable segment length. These are

1) 10BASE5 2) 10BASE2 3) 10BROAD36 4) 1BASES

The numeric field in the beginning indicates the bit rate in Mbps, the middle term indicates type of signaling system i.e. baseband or broadband, the numeric field in the end indicates the electrical cable segment length in X 100 metres.

- Manchester signal code is used at the baseband level 01 transmission. In broadband transmission, differential phase shift keying (DPSK) is used to convert the Manchester encoded signal into analogue form.

Fig. 5.2.11 shows the flowchart for CSMA/CD procedure.

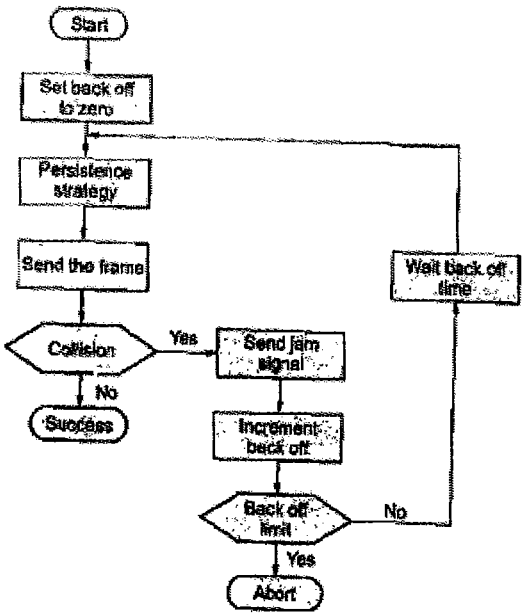


Fig. 5.2.11 Flowchart for CSMA/CD

CSMA/CD throughput

• The throughput of CSMA/CD is greater than that of pure or slotted ALOHA.

•For 1-persistent method, the maximum. throughput is around 50 % when G = 1.

For non-persistent method, the maximum throughput can go upto 90 % when G is between 3 and 8.

Carrier Sense Multiple Access with Collision Avoidance

•Wireless networks cannot use CSMA/CD in the MAC sublayer, since this requires the ability to receive" and transmit at the same time - hence the use of CSMA/CA.

• In a wireless network much of the sent energy is lost in transmission. The received signal has very little energy. Therefore, a collision may add only 5 to 10 percent additional energy. This is not useful for effective collision detection. We need to avoid collision on

wireless networks because they cannot be detected. So CSMA/CA was invented for this network.

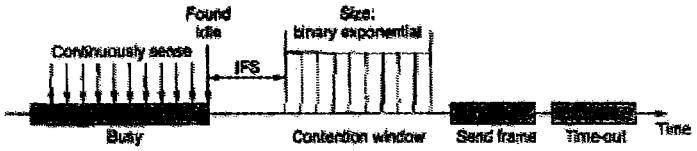
•Collisions are avoided by using three methods.

a. Inter-frame space

b. Contention window'

c. Acknowledgments

•Fig. 5.2.12 shows the all three method 01 CSMA/CA



Inter-frame space

- Collisions are avoided by deferring transmission even if the channel is found idle. .

- When an idle channel is found the station does not send immediately. It waits for a period of time called the inter-frame space (IFS).

- In CSMA/CA, the IFS can also be used to define the priority of a station 01 a frame. A station that is assigned shorter IPS has a higher priority.

Contention window

- Contention windows are an amount of time divided into slots. A station that is ready to send chooses a random number of slots a'i its wait time.

Station set one slot for the first time and then double each time the station cannot detect an idle channel after the IFS time.

•In this method, the station needs to sense the channel after each time slot.

•If the station finds the channel busy, it does not restart the process; it just stops the timer and restarts it when the' channel is sensed as idle.

• This method gives the priority to the station with the longest waiting time.

Acknowledgments

• The data may be corrupted during the transmission. The positive acknowledgment and the time· out can help guarantee that· the receiver has received the frame.

Fig. 5.2.13 shows the flowchart for CSMA/CA.

Hidden Node Problem

In the case of wireless network it is possible that A is sending a message to B, but C is out of its range and hence while "listening" on the network it will find the network to be free and might try to send packets to B at the same time as A. 50, there will be a collision at B.

•The problem can be looked upon as if A and C are hidden from each other. Hence it is called the "hidden node problem".

• Fig. 5.2.14 shows node A is transmitting.

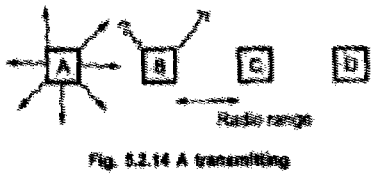


Fig. 5.2.14 A transmitting

Exposed Node Problem

• If C is transmitting a message to D and B wants to transmit a message to A, B will find the network to be busy as B hears C transmitting. Even if B would have transmitted to A, it would not have been a problem at A or D.

• CSMA/CD would not allow it to transmit message to A, while the two transmissions could have gone in parallel.

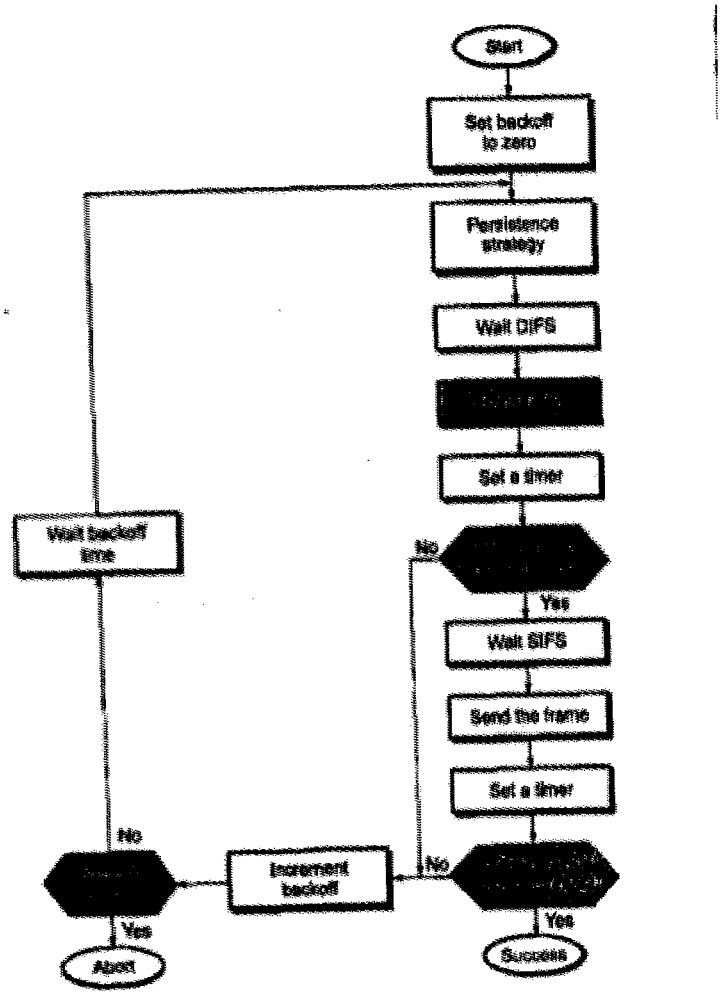


Fig. 5.2.13 Flowchart for CSMA/CA

Fig. 5.2.15 shows node B is transmitting.

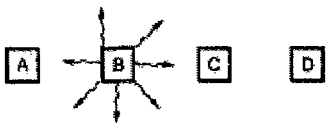


Fig. 5.2.15 B transmitting

Controlled Access

1. In this, the stations consult one another to find which station has the right to send.

2. A station cannot send unless it has been authorized ,by other stations.

3. Controlled access methods are

I. Reservation

ii. Polling

iii. Token passing 011 Reservation

1. Before sending data, station needs to make a reservation.

fig. 5,3.1 shows the reservation access method.

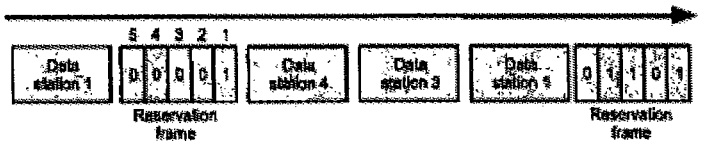


Fig. 6.3.1 Reservation access method

2. Number of reservation are equal to number of stations.

3. Each station have their own mini slot in the reservation frame.

4. When station needs to send a data frame, it makes a reservation in its own mini slot.

5. The stations that have made reservations can send their data frames after the reservation frame.

6. In the first slot, only station I, 3 and 4 have made reservation.

Polling

1. Polling works with topology.

2. One device is designed as primary station and other devices are secondary station.

3. Link control is done by primary device..

4. All data exchange take place through primary device.

5. Primary device decides, which device is allowed to use the channel at a given time.

6. If primary device wants to receive data, it asks the secondary if they have anything to send, this function is called polling.

7. Select mode and poll mode are the two functions of polling.

8. In polling, primary device receive the data.

9. In select mode, primary device sends data to secondary device.

fig. 5.3.2 shows the select mode.

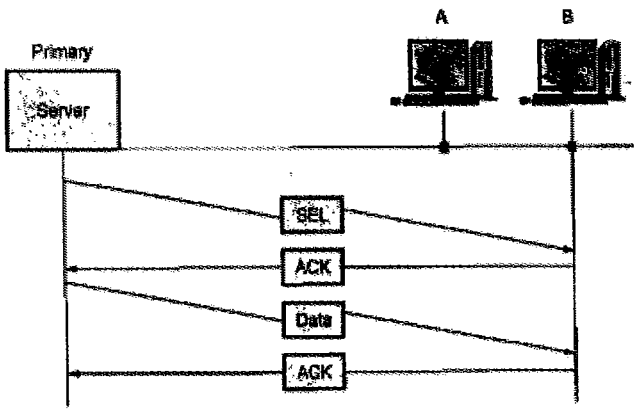


Fig. 5.3.2 Select mode

10. Link is available if primary device is not sending or receiving any data.

11. Before sending data, the primary creates and transmits a select (SEL) frame.

12. SEL frame includes address of the intended secondary device.

Fig. 5.3.3 shows the poll method

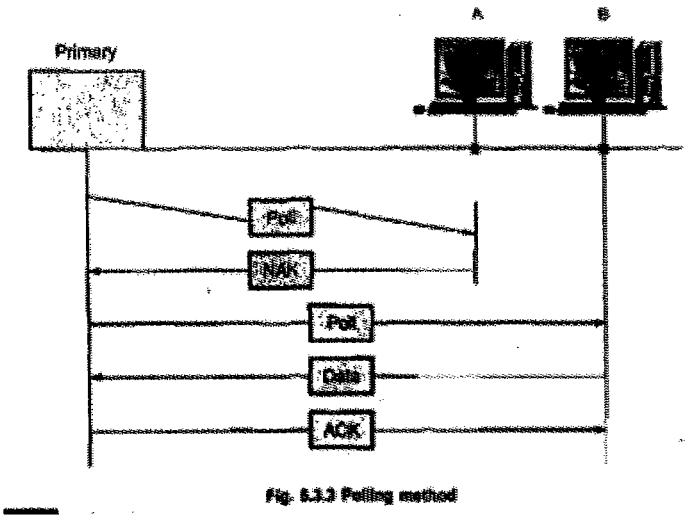


Fig. 5.3.3 Polling method

Token Passing :

1. A station is allowed to send data when it receives a token (special frame).

2. Ring topology is used for connecting devices.

3. Each station has a predecessor and a successor.

4. Frames are coming from predecessor and going to the successor.

Fig. 5.3.4 shows token passing network.

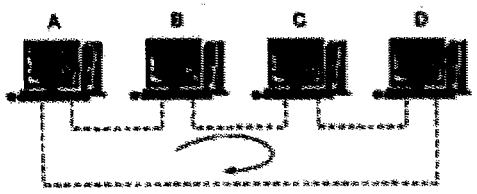


Fig. 5.3.4 Token passing network

5. Token is circulates around the ring.

6. The station captures the token if they want to send data.

7. Flowchart for token passing procedure is shown in Fig. 5.3.5

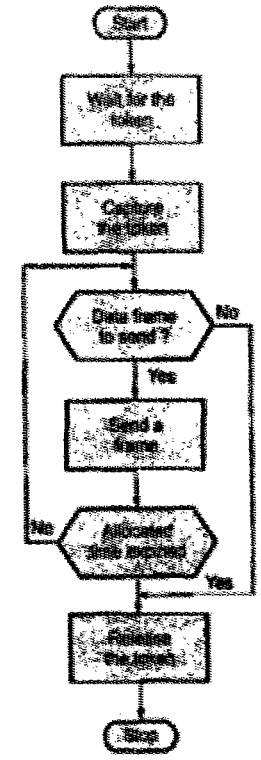


Fig. 5.3.5 Flowchart for token passing

Limited contention protocols:

We have now considered two basic strategies for channel acquisition in a broadcast network: contention, as in CSMA, and collision-free protocols. Each strategy can be rated as to how well it does with respect to the two important performance measures, delay at low load and channel efficiency at high load. Under conditions of light load, contention (i.e., pure or slotted ALOHA) is preferable due to its low delay (since collisions are rare). As the load increases, contention becomes increasingly less attractive because the overhead associated with channel arbitration becomes greater. Just the reverse is true for the collision-free protocols. At low load, they have relatively high delay but as the load increases, the channel efficiency improves (since the overheads are fixed). Obviously, it would be nice if we could combine the best properties of the contention and collision-free protocols, arriving at a new protocol that used

contention at low load to provide low delay, but used a collision-free technique at high load to provide good channel efficiency. Such protocols, which we will call limited-contention protocols, do in fact exist, and will conclude our study of carrier sense networks. Up to now, the only contention protocols we have studied have been symmetric. That is, each station attempts to acquire the channel with some probability, p, with all stations using the same p. Interestingly enough, the overall system performance can sometimes be improved by using a protocol that assigns different probabilities to different stations. Before looking at the asymmetric protocols, let us quickly review the performance of the symmetric case. Suppose that k stations are contending for channel access. Each has a probability p of transmitting during each slot. The probability that some station successfully acquires the channel during a given slot is the probability that any one station transmits, with probability p, and all other k – 1 stations defer, each with probability 1 – p. This value is kp(1 – p)k – 1. To find the optimal value of p, we differentiate with respect to p, set the result to zero, and solve for p. Doing so, we find that the best value of p is 1/k. Substituting p = 1/k, we get

$$\Pr[\text{success with optimal } p] = \left[\frac{k-1}{k}\right]^{k-1}$$

This probability is plotted in Fig. 4-9. For small numbers of stations, the chances of success are good, but as soon as the number of stations reaches even five, the probability has dropped close to its asymptotic value of 1/e. From Fig., it is fairly obvious that the probability of some station acquiring the channel can be increased only by decreasing the amount of competition. The limited-contention protocols do precisely that. They first divide the stations

into (not necessarily disjoint) groups. Only the members of group 0 are permitted

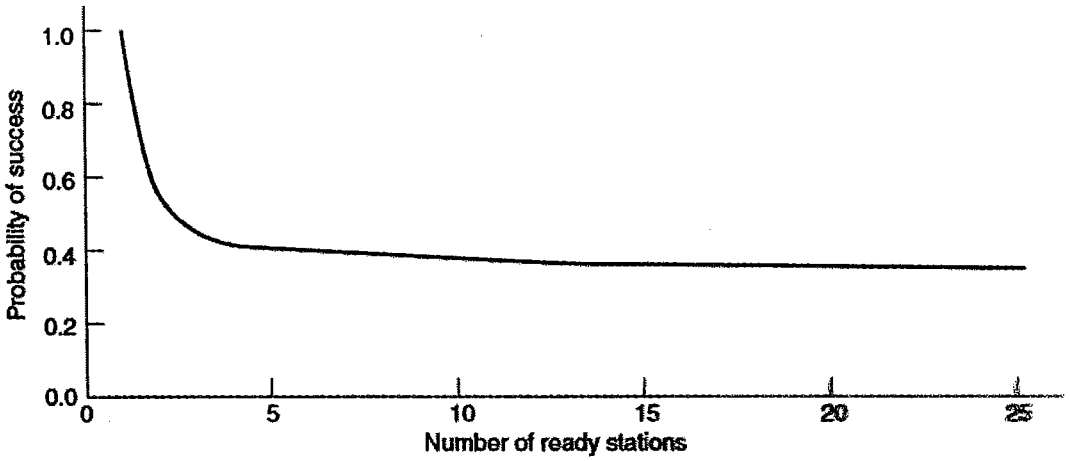


Figure 4-9. Acquisition probability for a symmetric contention channel.

to compete for slot 0. If one of them succeeds, it acquires the channel and transmits its frame. If the slot lies fallow or if there is a collision, the members of group 1 contend for slot 1, etc. By making an appropriate division of stations into groups, the amount of contention for each slot can be reduced, thus operating each slot near the left end of Fig. 4-9. The trick is how to assign stations to slots. Before looking at the general case, let us consider some special cases. At one extreme, each group has but one member. Such an assignment guarantees that there will never be collisions because at most one station is contending for any given slot. We have seen such protocols before (e.g., binary countdown). The next special case is to assign two stations per group. The probability that both will try to transmit during a slot is p2, which for a small p is negligible. As more and more stations are assigned to the same slot, the probability of a collision grows, but the length of the bit-map scan needed to give everyone a chance shrinks. The limiting case is a single group containing all stations (i.e., slotted ALOHA). What we need is a way to assign stations to slots dynamically, with many stations per slot

when the load is low and few (or even just one) station per slot when the load is high.

The Adaptive Tree Walk Protocol:

One particularly simple way of performing the necessary assignment is to use the algorithm devised by the U.S. Army for testing soldiers for syphilis during World War II (Dorfman, 1943). In short, the Army took a blood sample from N soldiers. A portion of each sample was poured into a single test tube. This mixed sample was then tested for antibodies. If none were found, all the soldiers in the group were declared healthy. If antibodies were present, two new mixed sampleswereprepared, one from soldiers 1 through N/2 and one from the rest. The process was repeated recursively until the infected soldiers were determined. For the computerized version of this algorithm (Capetanakis, 1979), it is convenient to think of the stations as the leaves of a binary tree, as illustrated in Fig. 4-10. In the first contention slot following a successful frame transmission, slot 0, all stations are permitted to try to acquire the channel. If one of them does so, fine. If there is a collision, then during slot 1 only those stations falling under node 2 in the tree may compete. If one of them acquires the channel, the slot following the frame is reserved for those stations under node 3. If, on the other hand, two or more stations under node 2 want to transmit, there will be a collision during slot 1, in which case it is node 4's turn during slot 2.

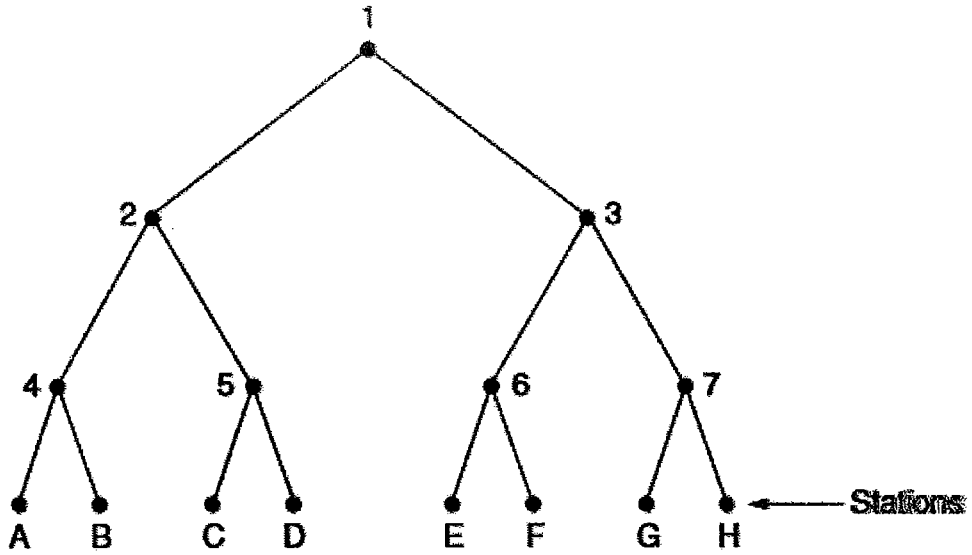


Figure 4-10. The tree for eight stations.

In essence, if a collision occurs during slot 0, the entire tree is searched, depth first, to locate all ready stations. Each bit slot is associated with some particular node in the tree. If a collision occurs, the search continues recursively with the node's left and right children. If a bit slot is idle or if only one station transmits in it, the searching of its node can stop because all ready stations have been located. (Were there more than one, there would have been a collision.) When the load on the system is heavy, it is hardly worth the effort to dedicate slot 0 to node 1 because that makes sense only in the unlikely event that precisely one station has a frame to send. Similarly, one could argue that nodes 2 and 3 should be skipped as well for the same reason. Put in more general terms, at what level in the tree should the search begin? Clearly, the heavier the load, the farther down the tree the search should begin. We will assume that each station has a good estimate of the number of ready stations, q, for example, from monitoring recent traffic. To proceed, let us number the levels of the tree from the top, with node 1 in Fig. 4-10 at level 0, nodes 2 and 3 at level 1, etc. Notice that each node at level

i has a fraction 2−i of the stations below it. If the q ready stations are uniformly distributed, the expected number of them below a specific node at level i is just 2−iq. Intuitively, we would expect the optimal level to begin searching the tree to be the one at which the mean number of contending stations per slot is 1, that is, the level at which 2−iq = 1. Solving this equation, we find that i = log2 q. Numerous improvements to the basic algorithm have been discovered and are discussed in some detail by Bertsekas and Gallager (1992). For example, consider the case of stations G and H being the only ones wanting to transmit. At node 1 a collision will occur, so 2 will be tried and discovered idle. It is pointless to probe node 3 since it is guaranteed to have a collision (we know that two or more stations under 1 are ready and none of them are under 2, so they must all be under 3). The probe of 3 can be skipped and 6 tried next. When this probe also turns up nothing, 7 can be skipped and node G tried next.

Wireless LAN Protocols:

A system of laptop computers that communicate by radio can be regarded as awireless LAN.a LAN is an example of a broadcast channel. It also has somewhat different properties than a wired LAN, which leads to different MAC protocols. In this section, we will examine some of these protocols. In Sec. 4.4, we will look at 802.11 (WiFi) in detail. A common configuration for a wireless LAN is an office building with access points (APs) strategically placed around the building. The APs are wired together using copper or fiber and provide connectivity to the stations that talk to them. If the transmission power of the APs and laptops is adjusted to have a range of tens of meters, nearby rooms become like a single cell and the entire building becomes like the cellular telephony systems we studied in Chap. 2, except that each cell only has one channel. This channel is shared by all the stations in the cell, including the AP. It

typically provides megabit/sec bandwidths, up to 600 Mbps. We have already remarked that wireless systems cannot normally detect a collision while it is occurring. The received signal at a station may be tiny, perhaps a million times fainter than the signal that is being transmitted. Finding it is like looking for a ripple on the ocean. Instead, acknowledgements are used to discover collisions and other errors after the fact. There is an even more important difference between wireless LANs and wired LANs. A station on a wireless LAN may not be able to transmit frames to or receive frames from all other stations because of the limited radio range of the stations. In wired LANs, when one station sends a frame, all other stations receive it. The absence of this property in wireless LANs causes a variety of complications. We will make the simplifying assumption that each radio transmitter has some fixed range, represented by a circular coverage region within which another station can sense and receive the station's transmission. It is important to realize thatin practice coverage regions are not nearly so regular because the propagation of radio signals depends on the environment. Walls and other obstacles that attenuate and reflect signals may cause the range to differ markedly in different directions. But a simple circular model will do for our purposes. A naive approach to using a wireless LAN might be to try CSMA: just listen for other transmissions and only transmit if no one else is doing so. The trouble is, this protocol is not really a good way to think about wireless because what matters for reception is interference at the receiver, not at the sender. To see the nature of the problem, consider Fig. 4-11, where four wireless stations are illustrated. For our purposes, it does not matter which are APs and which are laptops. The radio range is such that A and B are within each other's range and can potentially interfere with one another. C can also potentially interfere with both B and D but not A

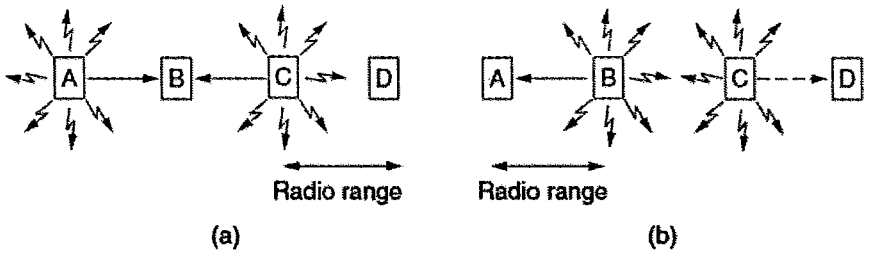


Figure 4-11. A wireless LAN. (a) *A* and *C* are hidden terminals when transmitting to *B*. (b) *B* and *C* are exposed terminals when transmitting to *A* and *D*.

First consider what happens when A and C transmit to B, as depicted in Fig. 4-11(a). If A sends and then C immediately senses the medium, it will not hear A because A is out of range. Thus C will falsely conclude that it can transmit to B. IfC does start transmitting, it will interfere at B, wiping out the frame from A. (We assume here that no CDMA-type scheme is used to provide multiple channels, so collisions garble the signal and destroy both frames.) We want a MAC protocol that will prevent this kind of collision from happening because it wastes bandwidth. The problem of a station not being able to detect a potential competitor for the medium because the competitor is too far away is called the hidden terminal problem. Now let us look at a different situation: B transmitting to A at the same time that C wants to transmit to D, as shown in Fig. 4-11(b). If C senses the medium, it will hear a transmission and falsely conclude that it may not send to D (shown as a dashed line). In fact, such a transmission would cause bad reception only in the zone between B and C, where neither of the intended receivers is located. We want a MAC protocol that prevents this kind of deferral from happening because it wastes bandwidth. The problem is called the exposed terminal problem.

Wireless LANs:

Wireless communication is one of the fastest-growing technologies. The demand for connecting devices without the use of cables is

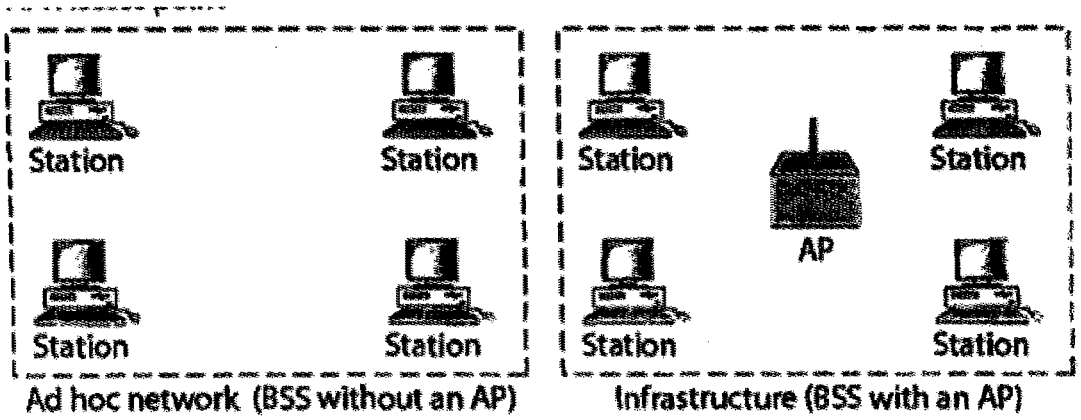
increasing. everywhere. Wireless LANs can be found on college campuses, in office buildings, and in many public areas. IEEE 802.11 IEEE has defined the specifications for a wireless LAN, called IEEE 802.11, which covers the physical and data link layers.

Architecture The standard defines two kinds of services: the basic service set (BSS) and the extended service set (ESS).

Basic Service Set IEEE 802.11 defines the basic service set (BSS) as the building block of a wireless LAN. A basic service set is made of stationary or mobile wireless stations and an optional central base station, known as the access point (AP). In bellow Figure shows two sets in this standard.

The BSS without anAP is a stand-alone network and cannot send data to other BSSs. It is called an ad hoc architecture. In this architecture, stations can form a network without the need of an AP; they can locate one another and agree to be part of a BSS. A BSS with an AP is sometimes referred to as an infrastructure network.

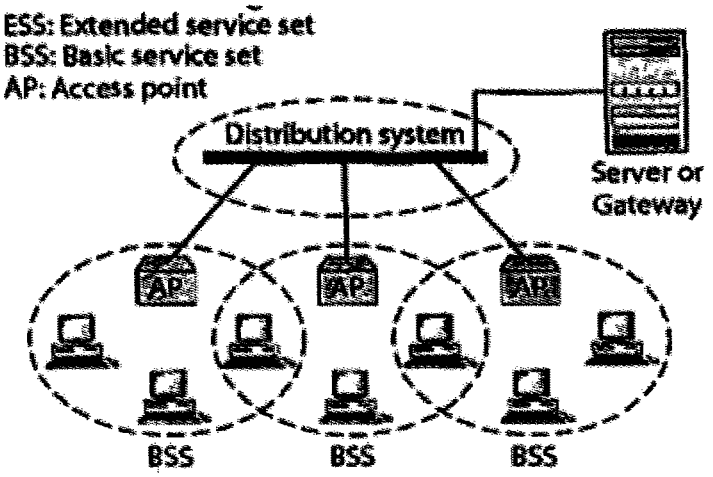
BSS: Basic service set AP: Access point



Extended Service Set

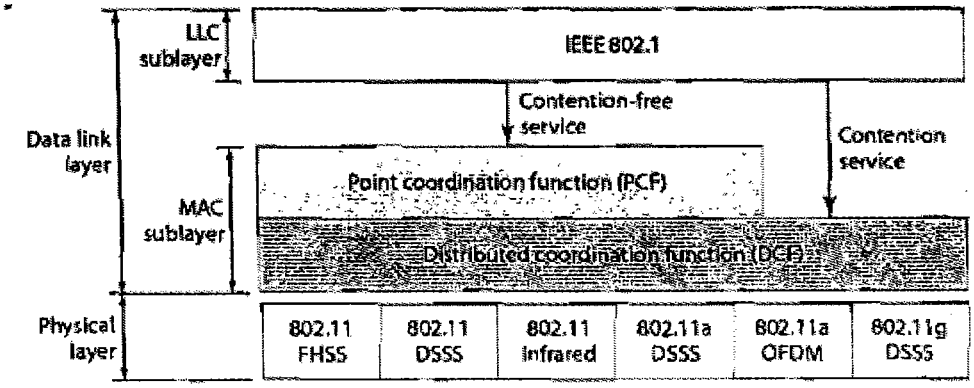
An extended service set (ESS) is made up oftwo or more BSSs with APs. In this case, the BSSs are connected through a distribution

system, which is usually a wired LAN. TIle distribution system COImects the APs in the BSSs. IEEE 802.11 does not restrict the distribution system; it can be any IEEE LAN such as an Ethernet. Note that the extended service set uses two types of stations: mobile and stationary. The mobile stations are normal stations inside a BSS. The stationary stations are AP stations that are part ofa wired LAN. Bellow Figure shows an ESS.



MAC Sublayer

IEEE 802.1 I defines two MAC sublayers: the distributed coordination function (OCF) and point coordination function (PCF). Figure 14.3 shows the relationship between the two MAC sublayers, the LLC sublayer, and the physical layer. We discuss the physical layer implementations later in the chapter and will now concentrate on the MAC sublayer



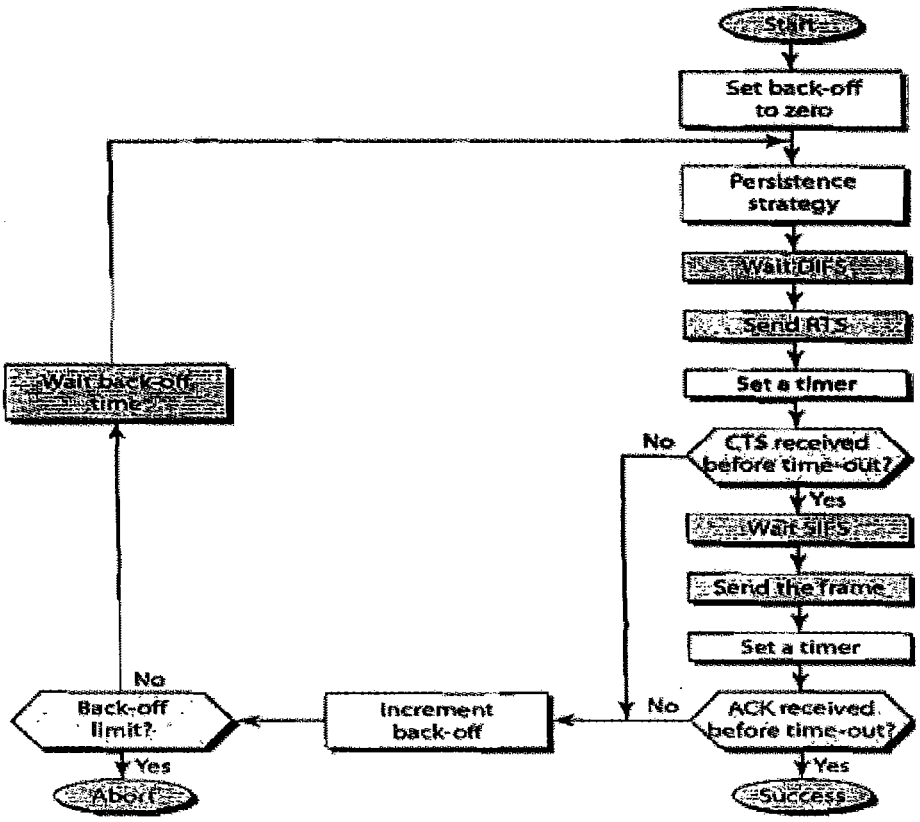
Distributed Coordination Function

One of the two protocols defined by IEEE at the MAC sublayer is called the distributed coordination function (OCF). OCF uses CSMAICA as the access method. Wireless LANs cannot implement CSMAfCO for three reasons:

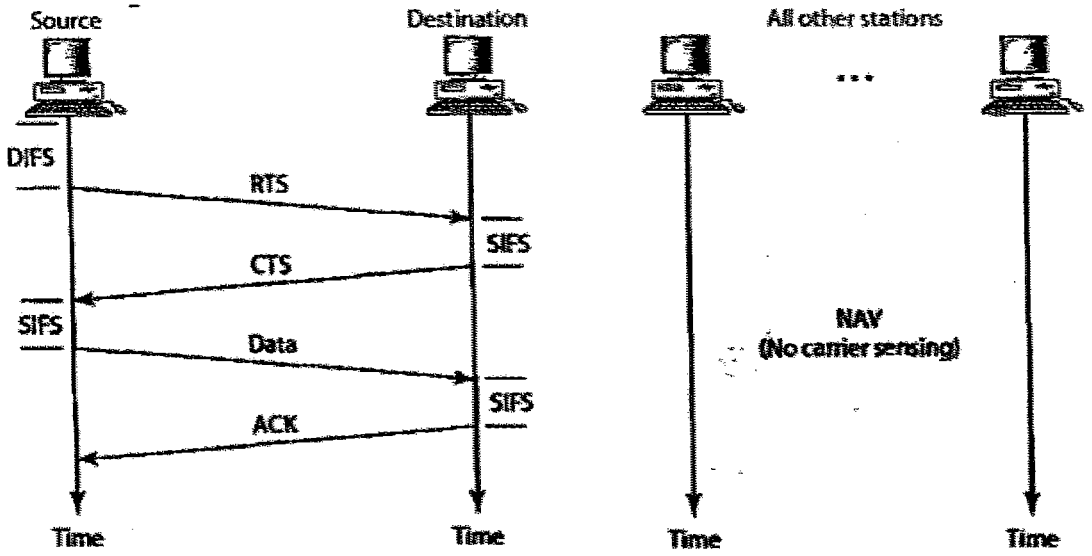
I. For collision detection a station must be able to send data and receive collision signals at the same time. This can mean costly stations and increased bandwidth requirements.

2. Collision may not be detected because ofthe hidden station problem.

3. The distance between stations can be great. Signal fading could prevent a station at one end from hearing a collision at the other end. Frocess Flowchart the process flowchart for CSMAICA as used in wireless LANs. We will explain the steps shortly.



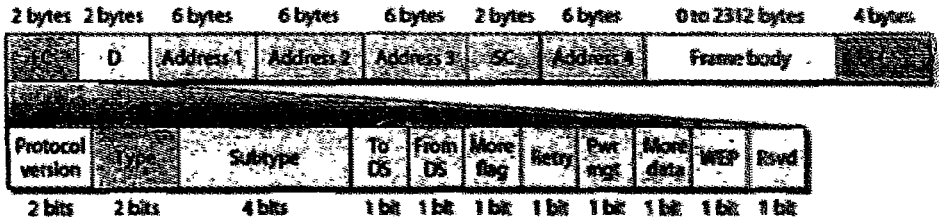
the exchange of data and control frames in time.



I. Before sending a frame, the source station senses the medium by checking the energy level at the carrier frequency. a. The channel uses a persistence strategy with back-off until the channel is idle. b. After the station is found to be idle, the station waits for a period of time called the distributed interframe space (DIFS); then the station sends a control frame called the request to send (RTS). 2. After receiving the RTS and waiting a period of time called the short interframe space (SIFS), the destination station sends -a<control frame, called the clear to send (CTS), to the source station. This control frame indicates that the destination station is ready to receive data. 3. The source station sends data after waiting an amount of time equal to SIFS. 4. The destination station, after waiting an amount of time equal to SIFS, sends an acknowledgment to show that the frame has been received. Acknowledgment is needed in this protocol because the station does not have any means to check for the successful arrival of its data at the destination. On the other hand, the lack of collision in CSMAICD is a kind of indication to the source that data have arrived.

Frame Format

The MAC layer frame consists of nine fields, as shown in Bellow Figure Frame control (Fe). The FC field is 2 bytes long and defines the type of frame and some control information.



NAV Network Allocation vector

Field	*Explanation*
Version	Current version is 0
Type	Type of information: management (00), control (01), or data (10)
Subtype	Subtype of each type (see Table 14.2)
To DS	Defined later
From DS	Defined later
More flag	When set to 1, means more fragments
Retry	When set to 1, means retransmitted frame
Pwr mgt	When set to 1, means station is in power management mode
More data	When set to 1, means station has more data to send
WEP	Wired equivalent privacy (encryption implemented)
Rsvd	Reserved

Subfields in FCfield

- **D. In all frame types except one, this field defines the duration of the transmission that is used to set the value of NAV. In one control frame, this field defines the ID of the frame.**

- **Addresses: There are four address fields, each 6 bytes long. The meaning of each address field depends on the value of the To DS and From DS subfields and will be discussed later.**

- **Sequence control: This field defines the sequence number of the frame to be used in flow control**

•Frame body:This field, which can be between 0 and 2312 bytes, contains information based on the type and the subtype defined in the FC field.

• FCS:The FCS field is 4 bytes long and contains a CRC-32 error detection sequence.

Frame Types A wireless LAN defined by IEEE 802.11 has three **categories of frames**: management frames, control frames, and data frames.

Management Frames Management frames are used for the initial communication between stations and access points.

Control Frames Control frames are used for accessing the channel and cknowledging frames.



For control frames the value of the type field is 0 l; the values of the subtype fields for frames .

Subtype	*Meaning*
1011	Request to send (RTS)
1100	Clear to send (CTS)
1101	Acknowledgment (ACK)

Values of subfields in control frames

Data Frames Data frames are used for carrying data and control information.

Addressing Mechanism

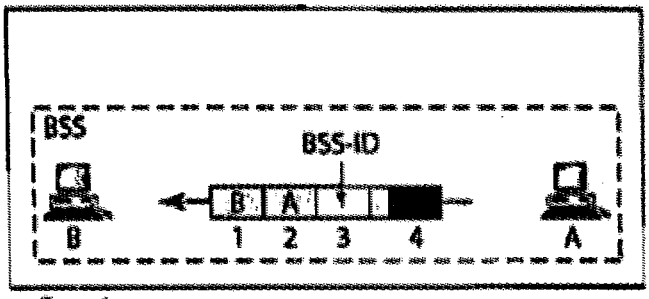
The IEEE 802.11 addressing mechanism specifies four cases, defined by the value of the two flags in the FC field, To DS and From DS. Each flag can be either 0 or I, resulting in four different situations. The interpretation of the four addresses (address I to address 4) in the MAC frame depends on the value of these flags.

To DS	From DS	Address 1	Address 2	Address 3	Address 4
0	0	Destination	Source	BSS ID	N/A
0	1	Destination	Sending AP	Source	N/A
1	0	Receiving AP	Source	Destination	N/A
1	1	Receiving AP	Sending AP	Destination	Source

Addresses

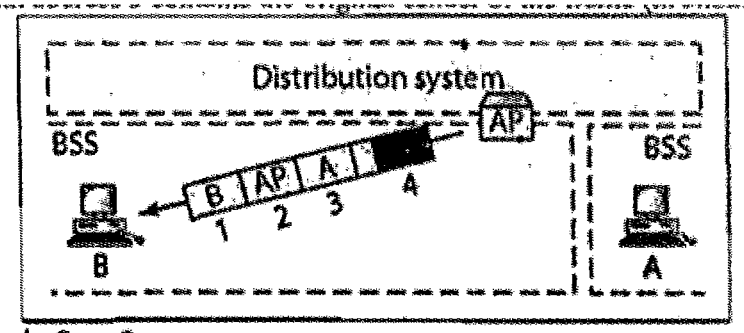
address I is always the address of the next device. Address 2 is always the address of the previous device. Address 3 is the address of the final destination station if it is not defined by address I. Address 4 is the address of the original source station if il is not the same as address 2.

Case 1: 00 In this case, To DS =0 and From DS =O. This means that the frame is not going to a distribution system (To DS = 0) and is not coming from a distribution system(From DS=0)The frame is going from one station in a BSS to another without passing through the distribution system. The ACK frame should be sent to the original sender.



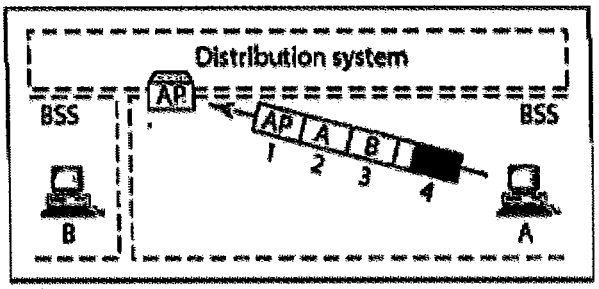
a. Case 1

Case 2: 01 In this case, To DS = 0 and From DS = 1. This means that the frame is coming from a distribution system (From DS = I). The frame is coming from an AP and going to a station. The ACK should be sent to the AP. The addresses are as shown In bellow Figure Note that address 3 contains the original sender ofthe frame (in another BSS).



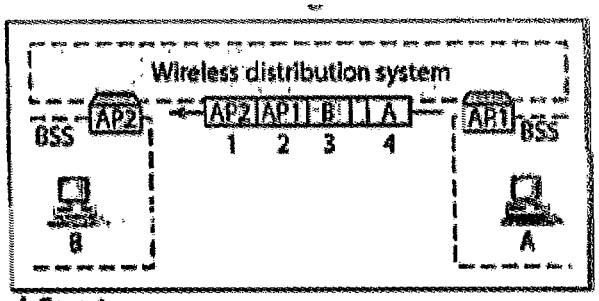
b. Case 2

Case 3: in this case, To DS =1 and From DS =0. This means that the frame is going to a distribution system (To DS = I). The frame is going from a station to an AP. The ACK is sent to the original station. The addresses are as shown in Figure bellow Note that address 3 contains the final destination of the frame (in another BSS).



c. Case 3

Case 4: II In this case, To DS =1 and From DS =1. Thus is the case in which the distribution system is also wireless. The frame is going from one AP to anotherAP in a wireless distribution system. We do not need to define addresses if the distribution system is a wired LAN because the frame in these cases has the format of a wired LAN frame (Ethernet, for example). Here, we need four addresses to define the original sender, the final destination, and two intermediate APs. Figure bellow shows the situation.



d. Case 4

Encoding

- In order to transport digital bits of data across carrier waves, encoding techniques have been developed each with their own merits and demerits.

- Digital signal is a sequence of discrete, discontinuous voltage pulses. Data is represented in binary. These binary data is transmitted by encoding each data bit into signal elements.

Desirable Properties of Encoding Techniques

1. Synchronization capability : The ability to stay synchronized or to get re-synchronized.

2. Error detection capability.

3. Immunity to noise: Ability to separate noise from the transmitted signal.

Manchester Encoding

- In Manchester encoding each bit period has both the high and low voltage values. If the data is a 1, the first half of the bit time period is sent at the positive level, and the second half of the period is at the negative level. For data bit of 0, first a negative signal and then a positive signal. There is a transition which can be used for a synchronization. Sometimes this method is called self

clockingencoding method. Fig. 6.2.4 shows manchester encoding waveform for the 8-bit data stream 11000101.

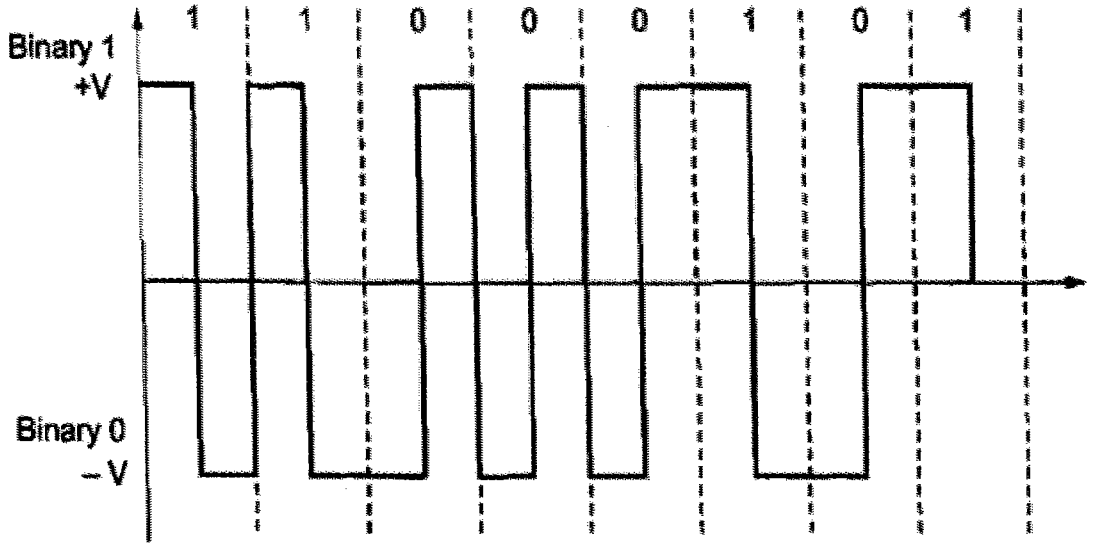


Fig. 6.2.4 Manchester encoding

2) Differential Manchester

- In differential Manchester encoding, a binary 0 is marked by a transition at the beginning of an interval, whereas a 1 is marked by the absence of a transition. In this encoding method, detecting changes is often more reliable, especially when there is a noise in the channel. Fig. 6.2.5 shows the differential manchester encoding for 8 bit data stream 101011100,

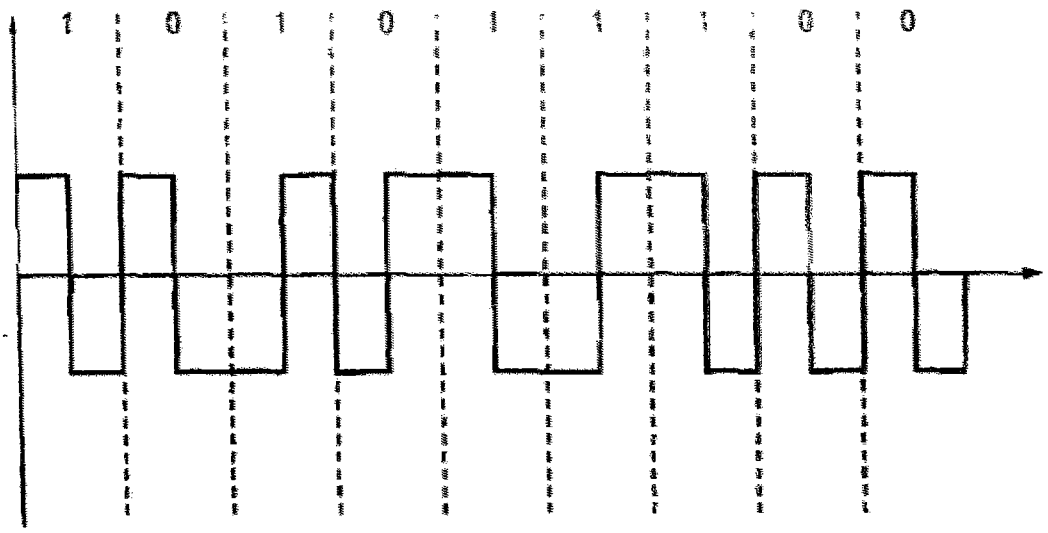


Fig. 6.2.5 Differential manchester

Ethernet:

Ethernet, probably the most ubiquitous kind of computer network in the world. Two kinds of Ethernet exist: classic Ethernet, which solves the multiple access problem using the techniques we have studied in this chapter; and switched Ethernet, in which devices called switches are used to connect different computers. It is important to note that, while they are both referred to as Ethernet, they are quite different. Classic Ethernet is the original form and ran at rates from 3 to 10 Mbps. Switched Ethernet is what Ethernet has become and runs at 100, 1000, and 10,000 Mbps, in forms called fast Ethernet, gigabit Ethernet, and 10 gigabit Ethernet. In practice, only switched Ethernet is used nowadays.

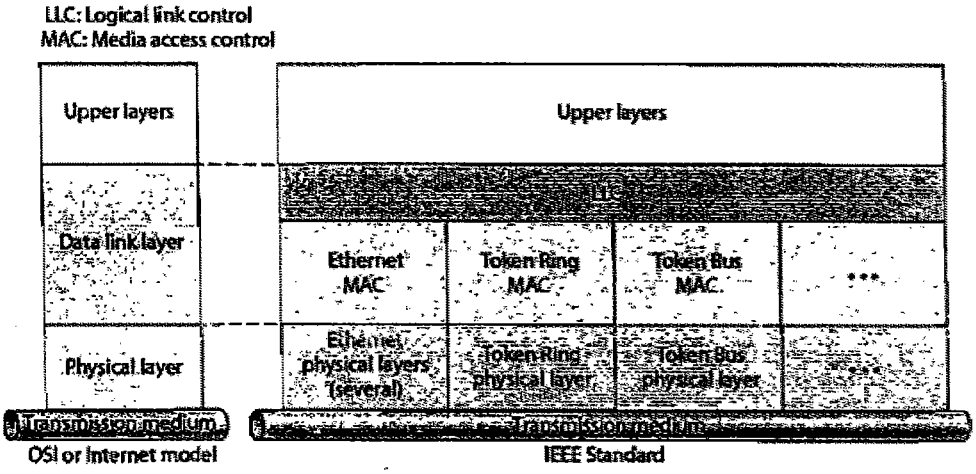
Classic Ethernet physical layer

IEEE STANDARDS

In 1985, the Computer Society of the IEEE started a project, called Project 802, to set standards to enable intercommunication among equipment from a variety of manufacturers. Project 802 is a way of

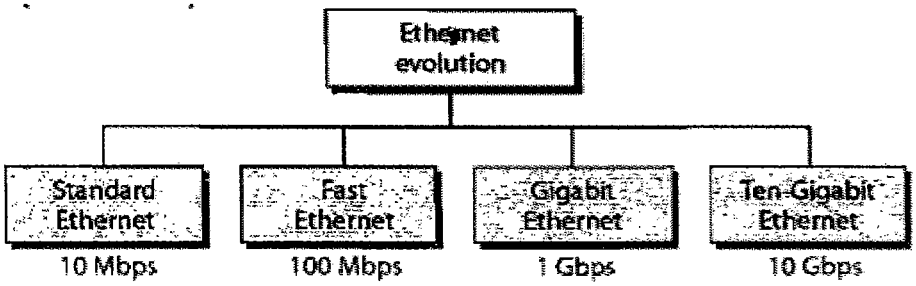
specifying functions of the physical layer and the data link layer of major LAN protocols.

The relationship of the 802 Standard to the traditional OSI model is shown in bellow Figure.The IEEE has subdivided the data link layer into two sublayers: logical link control (LLC) and media access control (MAC). IEEE has also created several physical layer standards for different LAN protocols.



STANDARD ETHERNET

The original Ethernet was created in 1976 at Xerox's Palo Alto Research Center (PARC). Since then, it has gone through four generations: Standard Ethernet (103 Mbps), Fast Ethernet (100 Mbps), Gigabit Ethernet (I Gbps), and Ten-Gigabit Ethernet (10 Gbps), as shown in Figure 13.3. We briefly discuss all these generations starting with the first, Standard (or traditional) Ethernet.

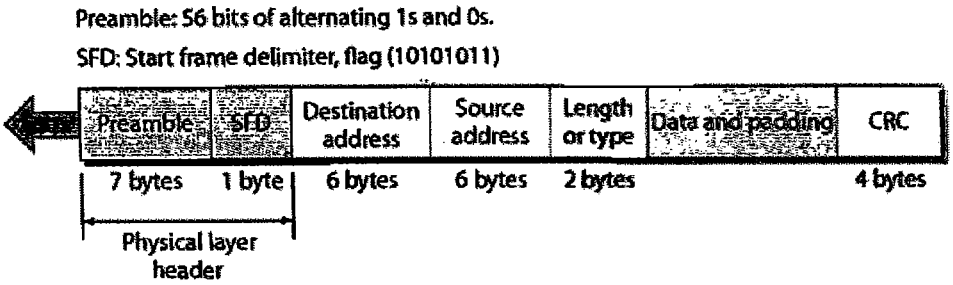


Ethernet evolution through four generations

MAC Sublayer

In Standard Ethernet, the MAC sublayer governs the operation of the access method.It also frames data received from the upper layer and passes them to the physical Layer.

Frame Format The Ethernet frame contains seven fields: preamble, SFD, DA, SA, length or type of protocol data unit. (PDU), upper-layer data, and the CRe. Ethernet does not provide any mechanism for acknowledging received frames, making it what is known as an unreliable medium. Acknowledgments must be implemented at the higher layers.



802.3 MAC frame

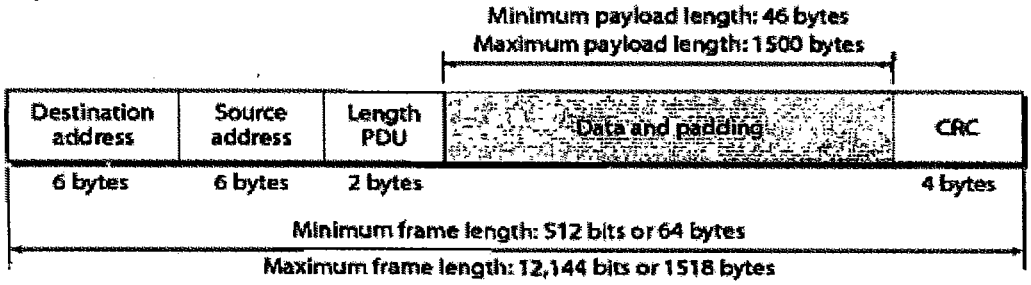
Preamble. The first field of the 802.3 frame contains 7 bytes (56 bits) of alternating as and Is that alerts the receiving system to the coming frame and enables it to synchronize its input timing. The pattern provides only an alert and a timing pulse. The 56-bit pattern allows the stations to miss some bits at the beginning of the frame. The preamble is actually added at the physical layer and is not (formally) part of the frame. Start frame delimiter (SFD). The second field (I byte: 10101011) signals the beginning of the frame. The SPD warns the station or stations that this is the last chance for synchronization. The last 2 bits r 11 and alerts the receiver that the next field is the destination address.

Destination address (DA). The DA field is 6 bytes and contains the physical address of the destination station or stations to receive the

packet. Source address (SA). The SA field is also 6 bytes and contains the physical address of the sender of the packet. . Length or type. This field is defined as a type field or length field. The original Ethernet used this field as the type field to define the upper-layer protocol using the MAC frame. The IEEE standard used it as the length field to define the number of bytes in the data field. Both uses are common today. Data. This field carries data encapsulated from the upper-layer protocols. It is a minimum of 46 and a maximum of 1500 bytes, as we will see later. CRe. The last field contains error detection information, in this case a CRC-32.

Frame Length

Ethernet has imposed restrictions on both the minimum and maximum lengths of a frame,



Frame length: Minimum: 64 bytes (512 bits) Maximum: 1518 bytes (12,144 bits)

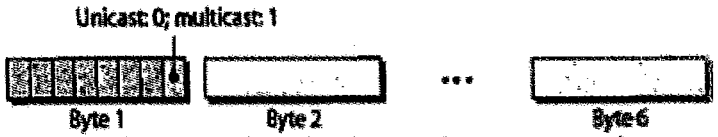
Addressing

Each station on an Ethernet network (such as a PC, workstation, or printer) has its own network interface card (NIC). The NIC fits inside the station and provides the station with a 6-byte physical address. As shown in bellow Figure the Ethernet address is 6 bytes (48 bits), normally written in hexadecimal notation, with a colon between the bytes.

06 : 01 : 02 : 01 : 2C : 4B

6 bytes = 12 hex digits = 48 bits

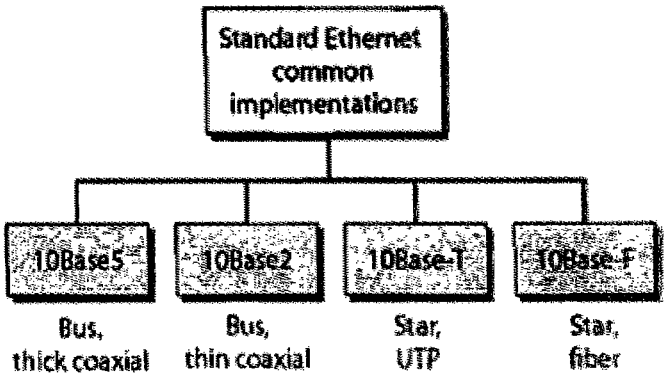
Unicast, Multicast, and Broadcast Addresses A source address is always a unicast address-the frame comes from only one station. The destination address, however,can be unicast, multicast, or broadcast. Figure 13.7 shows how to distinguish a unicast address from a multicast address. If the least significant bit of the first byte in a destination address is 0, the address is unicast; otherwise, it is multicast.



1. The least significant bit of the first byte defines the type of address. If the bit is 0, the address is unicast; otherwise, it is multicast.
2. The broadcast destination address is a special case of the multicast address in which all bits are 1s.

Physical Layer

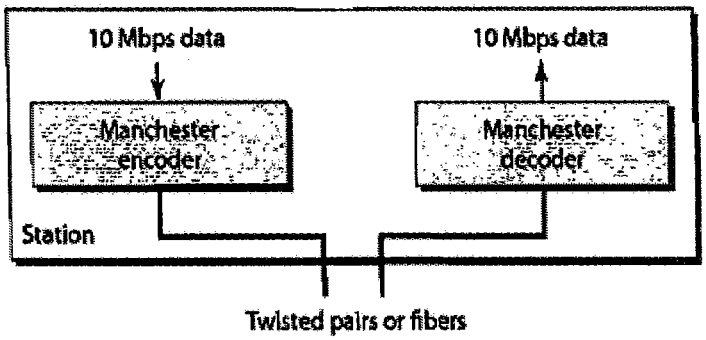
The Standard Ethernet defines several physical layer implementations; four of the most common.



Categories of Standard Ethernet

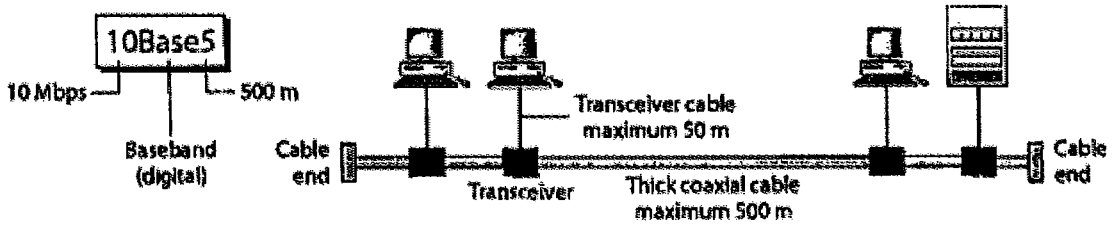
Encoding and Decoding

All standard implementations use digital signaling (baseband) at 10 Mbps. At the sender, data are converted to a digital signal using the Manchester scheme; at the receiver, the received signal is interpreted as Manchester and decoded into data. Manchester encoding is self-synchronous, providing a transition at each bit interval.



Encoding in a Standard Ethernet implementation

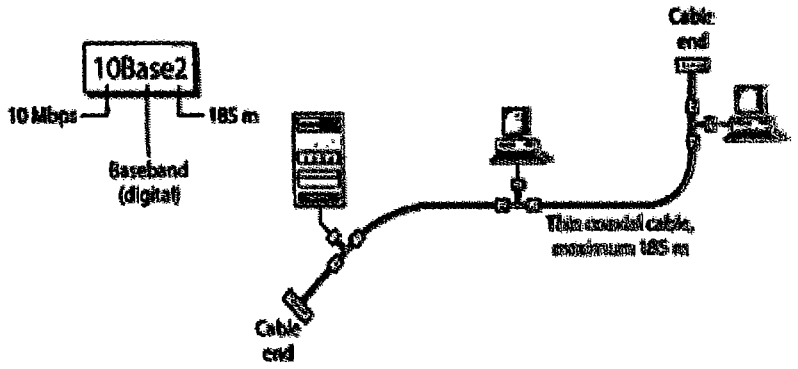
IOBase5: Thick Ethemet; The first implementation is called IOBase5, thick Ethernet, or Thicknet. The nickname derives from the size of the cable, which is roughly the size of a garden hose and too stiff to bend with your hands. IOBase5 was the first Ethernet specification to use a bus topology with an external transceiver(transmitter/receiver) connected via a tap to a thick coaxial cable. Figure 13.10 shows a schematic diagram of a IOBase5 implementation.



10Base5 implementation

The maximum length of the coaxial cable must not exceed 500 m, otherwise, there is excessive degradation of the signal. If a length of more than 500 m is needed, up to five segments, each a maximum of SaO-meter, can be connected using repeaters.

IOBase2: Thin Ethernet The second implementation is called IOBase2, thin Ethernet, or Cheapernet. IOBase2 also uses a bus topology, but the cable is much thinner and more flexible. The cable can be bent to pass very close to the stations. In this case, the transceiver is nonnally part of the network interface card (NIC), which is installed inside the station.

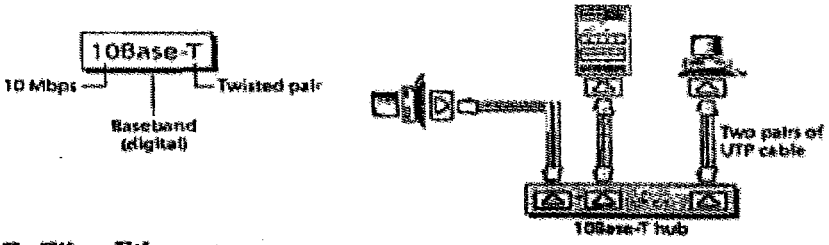


10Base2 implementation

the length of each segment cannot exceed 185 m (close to 200 m) due to the high level of attenuation in thin coaxial cable.

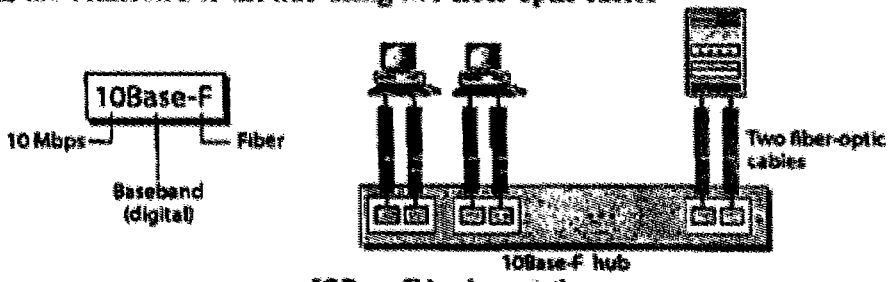
IOBase-T: Twisted-Pair Ethernet

The third implementation is called IOBase-T or twisted-pair Ethernet. IOBase-T uses a physical star topology. The stations are connected to a hub via two pairs of twisted cable, two pairs oftwisted cable create two paths (one for sending and one for receiving) between the station and the hub. Any collision here happens in the hub.Compared to IOBaseTor IOBase2, we can see that the hub actually replaces the coaxial cable as far as a collision is concerned. The maximum length of the twisted cable here is defined as 100 m, to minimize the effect of attenuation in the twisted cable.



10Base-F: Fiber Ethernet

Although there are several types of optical fiber 10-Mbps Ethernet, the most common is called 10Base-F. 10Base-F uses a star topology to connect stations to a hub. The stations are connected to the hub using two fiber-optic cables



10Base-F implementation

Characteristics	*10Base5*	*10Base2*	*10Base-T*	*10Base-F*
Media	Thick coaxial cable	Thin coaxial cable	2 UTP	2 Fiber
Maximum length	500 m	185 m	100 m	2000 m
Line encoding	Manchester	Manchester	Manchester	Manchester

Summary of Standard Ethernet implementations

FAST ETHERNET

Fast Ethernet was designed to compete with LAN protocols such as FDDI or FiberChannel (or Fibre Channel, as it is sometimes spelled). IEEE created Fast Ethernet under the name 802.3u. Fast Ethernet is backward-compatible with Standard Ethernet, but it can transmit data 10 times faster at a rate of 100 Mbps. The goals of Fast Ethernet can be summarized as follows:

1. Upgrade the data rate to 100 Mbps.
2. Make it compatible with Standard Ethernet.
3. Keep the same 48-bit address.
4. Keep the same frame format.

5. Keep the same minimum and maximum frame lengths.

MAC Sublayer

A main consideration in the evolution of Ethernet from 10 to 100 Mbps was to keep the MAC sublayer untouched. However, a decision was _made to drop the bus topologies and keep only the star topology. For the star topo[ogy, there are two choic~aw before: half duplex and full duplex. In the half-duplex approach, the stations are connected via a hub; in the full-duplex approach, the connection is made via a switch with buffers at each port. It was designed particularly for the following purposes:

•To allow incompatible devices to connect to one another. For example, a device with a maximum capacity of 10 Mbps can communicate with a device with a 100 Mbps capacity (but can work at a lower rate).

•To allow one device to have multiple capabilities.

•To allow a station to check a hub's capabilities.

Physical Layer The physical layer in Fast Ethernet is more complicated than the one in Standard Ethernet.

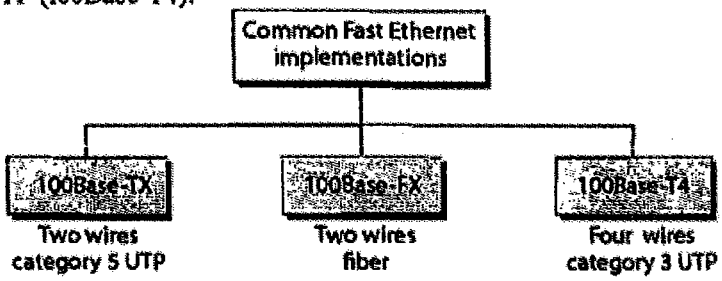
Fast Ethernet implementation at the physical layer can be categorized as either two-wireor four-wire. The two-wire implementation can be either category 5 UTP (IOOBase-TX)

orfiber-optic cable (IOOBase-FX}. The four-wire implementation is designed only for category 3 UTP (IOOBase-T4).

Physical Layer

The physical layer in Fast Ethernet is more complicated than the one in Standard Ethernet.

Fast Ethernet implementation at the physical layer can be categorized as either two-wire or four-wire. The two-wire implementation can be either category 5 UTP (100Base-TX) or fiber-optic cable (100Base-FX). The four-wire implementation is designed only for category 3 UTP (100Base-T4).



Characteristics	100Base-TX	100Base-FX	100Base-T4
Media	Cat 5 UTP or STP	Fiber	Cat 4 UTP
Number of wires	2	2	4
Maximum length	100 m	100 m	100 m
Block encoding	4B/5B	4B/5B	
Line encoding	MLT-3	NRZ-I	8B/6T

Summary of Fast Ethernet implementations

Giga Bit Ethernet:

Gigabit Ethernet supports both copper and fiber cabling, as listed in Fig. 4-21. Signaling at or near 1 Gbps requires encoding and sending a bit every nanosecond. This trick was initially accomplished with short, shielded copper cables (the 1000Base-CX version) and optical fibers. For the optical fibers, two wavelengths are permitted and result in two different versions: 0.85 microns (short, for 1000Base-SX) and 1.3 microns (long, for 1000Base-LX).

Name	Cable	Max. segment	Advantages
1000Base-SX	Fiber optics	550 m	Multimode fiber (50, 62.5 microns)
1000Base-LX	Fiber optics	5000 m	Single (10 μ) or multimode (50, 62.5 μ)
1000Base-CX	2 Pairs of STP	25 m	Shielded twisted pair
1000Base-T	4 Pairs of UTP	100 m	Standard category 5 UTP

Figure 4-21. Gigabit Ethernet cabling.

Ten Giga bit Ethernet:

The main kinds of 10-gigabit Ethernet are listed in Fig. 4-22. Multimode fiber with the 0.85μ (short) wavelength is used for medium distances, and singlemodefiber at 1.3μ (long) and 1.5μ (extended) is used for long distances. 10GBase-ER can run for distances of 40 km, making it suitable for wide area applications. All of these versions send a serial stream of information that is produced by scrambling the data bits, then encoding them with a 64B/66B code. This encoding has less overhead than an 8B/10B code.

Name	Cable	Max. segment	Advantages
10GBase-SR	Fiber optics	Up to 300 m	Multimode fiber (0.85μ)
10GBase-LR	Fiber optics	10 km	Single-mode fiber (1.3μ)
10GBase-ER	Fiber optics	40 km	Single-mode fiber (1.5μ)
10GBase-CX4	4 Pairs of twinax	15 m	Twinaxial copper
10GBase-T	4 Pairs of UTP	100 m	Category 6a UTP

Figure 4-22. 10-Gigabit Ethernet cabling.

802.11 and 805.11 Services:

1. **Association**
2. **Reassociation**
3. **Disassociate**
4. **Authenticate**
5. **WPA2 (WiFi Protected Access 2)**
6. **WEP (Wired Equivalent Privacy)**
7. **Distribution**
8. **Integration**

9. Privacy

10. AES (Advanced Encryption Standard)

11. QoS traffic scheduling service

12. The transmit power control

13. The dynamic frequency selection

The 802.11 standard defines the services that the clients, the access points, and the network connecting them must be a conformant wireless LAN. These services cluster into several groups. The

Association

service is used by mobile stations to connect themselves to APs. Typically, it is used just after a station moves within radio range of the AP. Upon arrival, the station learns the identity and capabilities of the AP, either from beacon frames or by directly asking the AP. The capabilities include the data rates supported, security arrangements, power-saving capabilities, quality of service support, and more. The station sends a request to associate with the AP. The AP may accept or reject the request.

Reassociation lets a station change its preferred AP. This facility is useful for mobile stations moving from one AP to another AP in the same extended 802.11 LAN, like a handover in the cellular network. If it is used correctly, no data will be lost as a consequence of the handover. (But 802.11, like Ethernet, is just a best-effort service.)

Either the station or the AP may also **disassociate**, breaking their relationship. A station should use this service before shutting down or leaving the network. The AP may use it before going down for maintenance.

Stations must also **authenticate** before they can send frames via the AP, but authentication is handled in different ways depending on the choice of security scheme. If the 802.11 network is "open," anyone is allowed to use it. Otherwise, credentials are needed to authenticate.

The recommended scheme, called **WPA2 (WiFi Protected Access 2)**, implements security as defined in the 802.11i standard. (Plain WPA is an interim scheme that implements a subset of 802.11i. We will skip it and go straight to the complete scheme.) With WPA2, the AP can talk to an authentication server that has a username and password database to determine if the station is allowed to access the network. Alternatively a pre-shared key, which is a fancy name for a network password, may be configured. Several frames are exchanged between the station and the AP with a challenge and response that lets the station prove it has the right credentials. This exchange happens after association.

The scheme that was used before WPA is called **WEP (Wired Equivalent Privacy)**. For this scheme, authentication with a preshared key happens before association. However, its use is discouraged because of design flaws that make WEP easy to compromise. The first practical demonstration that WEP was broken came when Adam Stubblefield was a summer intern at AT&T (Stubblefield et al., 2002). He was able to code up and test an attack in one week, much of which was spent getting permission from management to buy the WiFi cards needed for experiments. Software to crack WEP passwords is now freely available.

Once frames reach the AP, the **distribution** service determines how to route them. If the destination is local to the AP, the frames can be

sent out directly over the air. Otherwise, they will have to be forwarded over the wired network.

The **integration** service handles any translation that is needed for a frame to be sent outside the 802.11 LAN, or to arrive from outside the 802.11 LAN. The common case here is connecting the wireless LAN to the Internet. Data transmission is what it is all about, so 802.11 naturally provides a data delivery service. This service lets stations transmit and receive data using the protocols we described earlier in this chapter. Since 802.11 ismodeled on Ethernet and transmission over Ethernet is not guaranteed to be 100% reliable, transmission over 802.11 is not guaranteed to be reliable either. Higher layers must deal with detecting and correcting errors. Wireless is a broadcast signal. For information sent over a wireless LAN to be kept confidential, it must be encrypted.

This goal is accomplished with a **privacy** service that manages the details of encryption and decryption.

The encryption algorithm for WPA2 is based on **AES (Advanced Encryption Standard),** a U.S. government standard approved in 2002. The keys that are used for encryption are determined during the authentication procedure.

To handle traffic with different priorities, there is a **QoS traffic scheduling service**. It uses the protocols we described to give voice and video traffic preferential treatment compared to best-effort and background traffic. A companion service also provides higher-layer timer synchronization. This lets stations coordinate their actions, which may be useful for media processing.

Finally, there are two services that help stations manage their use of the spectrum. **The transmit power control** service gives stations the

information they need to meet regulatory limits on transmit power that vary from region to region.

The dynamic frequency selection service give stations the information they need to avoid transmitting on frequencies in the 5-GHz band that are being used for radar in the proximity. With these services, 802.11 provides a rich set of functionality for connecting nearby mobile clients to the Internet. It has been a huge success, and the standard has repeatedly been amended to add more functionality.