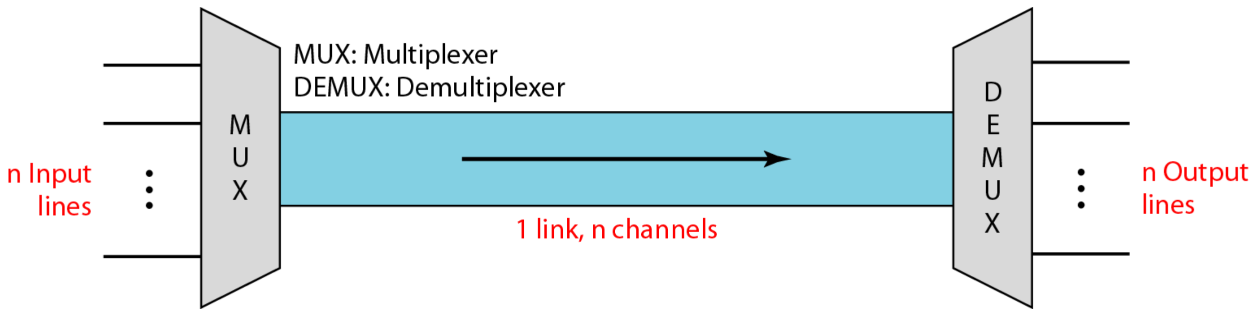
BANDWIDTH UTILIZATION

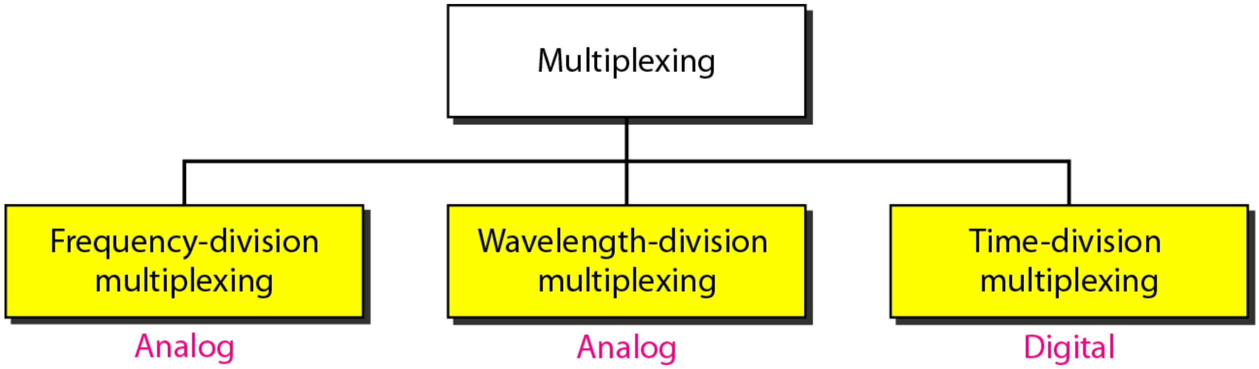
Bandwidth utilization is the wise use of available bandwidth to achieve specific goals. Efficiency can be achieved by multiplexing; privacy and anti jamming can be achieved by spreading.

3.1 Multiplexing

- Multiplexing is the set of techniques that allows the simultaneous transmission of multiple signals across a single data link.
- In a multiplexed system, n lines share the bandwidth of one link.
- Below figure shows the basic format of a multiplexed system.



- The lines on the left direct their transmission streams to a multiplexer (MUX), which combines them into a single stream (many-to one).
- At the receiving end, that stream is fed into a demultiplexer (DEMUX), which separates the stream back into its component transmissions (one-to-many) and directs them to their corresponding lines.
- In the figure, the word **link** refers to the physical path.
- The word **channel** refers to the portion of a link that carries a transmission between a given pair of lines.
- One link can have many (n) channels.
- There are three basic multiplexing techniques: frequency-division multiplexing, wavelength-division multiplexing, and time-division multiplexing.

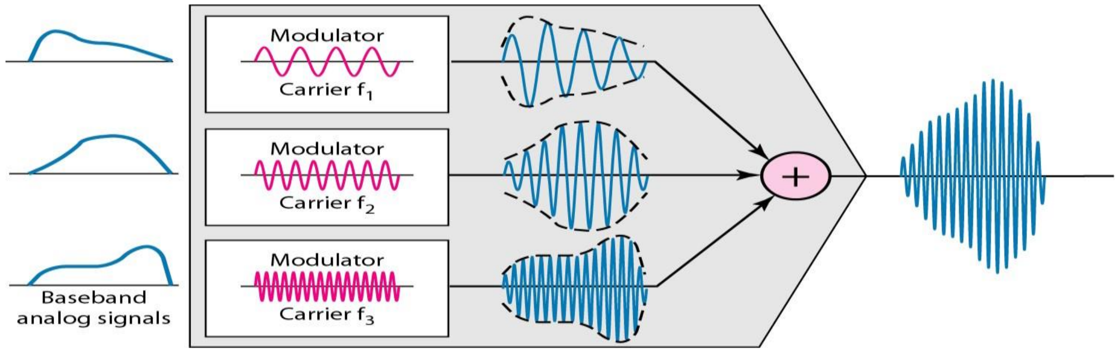


Frequency-Division Multiplexing

- Frequency-division multiplexing (FDM) is an analog technique that can be applied when the bandwidth of a link (in hertz) is greater than the combined bandwidths of the signals to be transmitted.
- In FDM, signals generated by each sending device modulate different carrier frequencies.
- These modulated signals are then combined into a single composite signal that can be transported by the link.
- Carrier frequencies are separated by sufficient bandwidth to accommodate the modulated signal. These bandwidth ranges are the channels through which the various signals travel.
- Channels can be separated by strips of unused bandwidth called **guard bands** - to prevent signals from overlapping.
- In addition, carrier frequencies must not interfere with the original data frequencies.
- Though FDM is an analog multiplexing technique, however it does not mean that FDM cannot be used to combine digital signals. A digital signal has to be converted to analog signals before FDM is used to multiplex them.

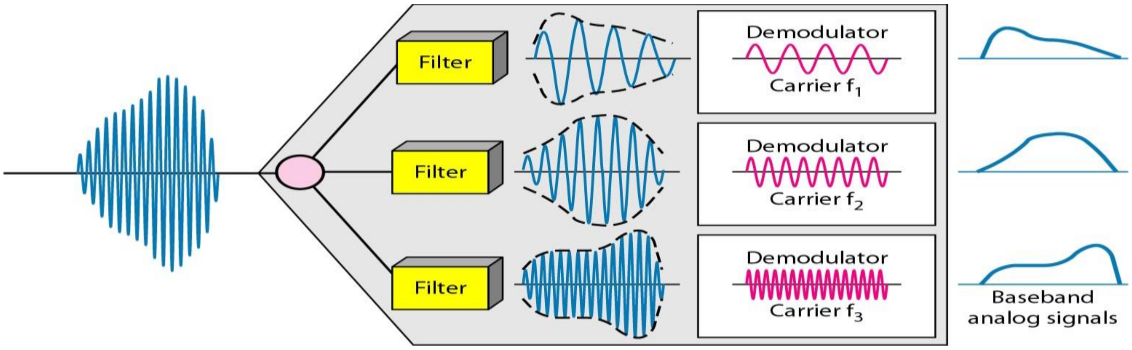
Multiplexing Process

- Each source generates a signal of a similar frequency range.
- Inside the multiplexer, these similar signals modulates different carrier frequencies *(f1,f2 and f3).*
- The resulting modulated signals are then combined into a single composite signal that is sent out over a media link that has enough bandwidth to accommodate it.



Demultiplexing Process

- The demultiplexing process uses filters
- The demultiplexer uses a series of filters to decompose the multiplexed signal into its constituent component signals.
- The individual signals are then passed to a demodulator that separates them from their carriers and passes them to the output lines.

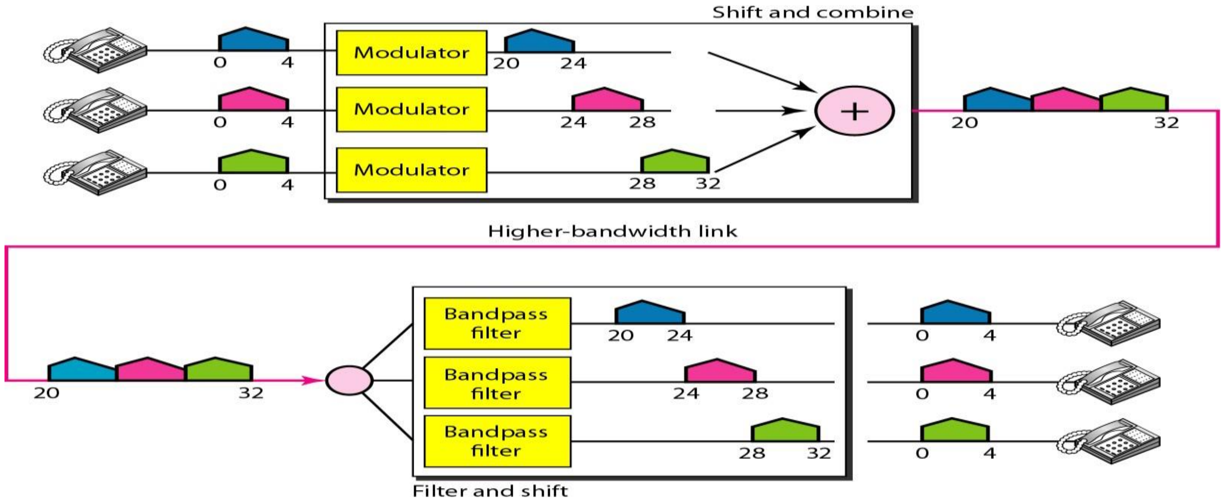


Examples:

1) Assume that a voice channel occupies a bandwidth of 4 kHz. We need to combine three voice channels into a link with a bandwidth of 12 kHz, from 20 to 32 kHz. Show the configuration, using the frequency domain. Assume there are no guard bands.

Solution

We shift (modulate) each of the three voice channels to a different bandwidth, as shown in Figure

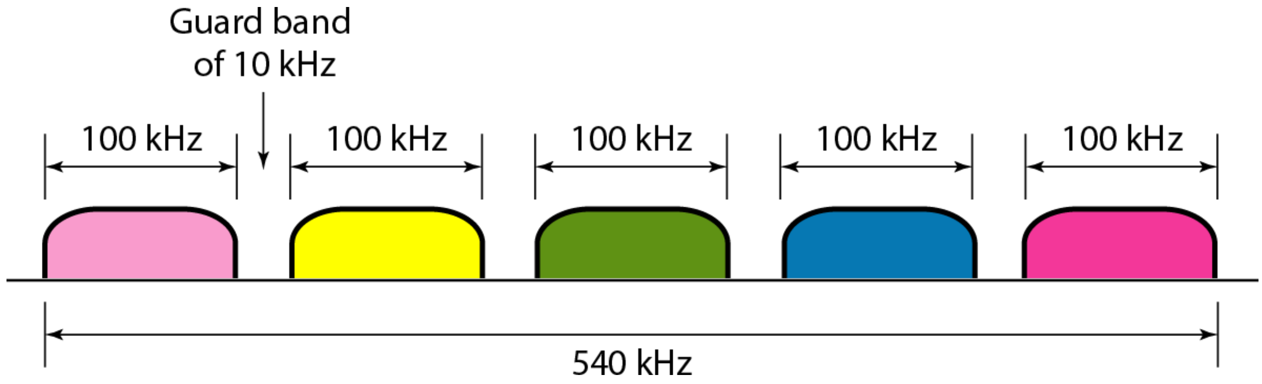


- We use the 20- to 24-kHz bandwidth for the first channel, the 24- to 28-kHz bandwidth for the second channel, and the 28- to 32-kHz bandwidth for the third one.
- Then we combine them. At the receiver, each channel receives the entire signal, using a filter to separate out its own signal.
- The first channel uses a filter that passes frequencies between 20 and 24 kHz and filters out (discards) any other frequencies. The second channel uses a filter that passes frequencies between 24 and 28 kHz, and the third channel uses a filter that passes frequencies between 28 and 32 kHz. Each channel then shifts the frequency to start from zero.

2) Five channels, each with a 100-kHz bandwidth, are to be multiplexed together. What is the minimum bandwidth of the link if there is a need for a guard band of 10kHz between the channels to prevent interference?

Solution:

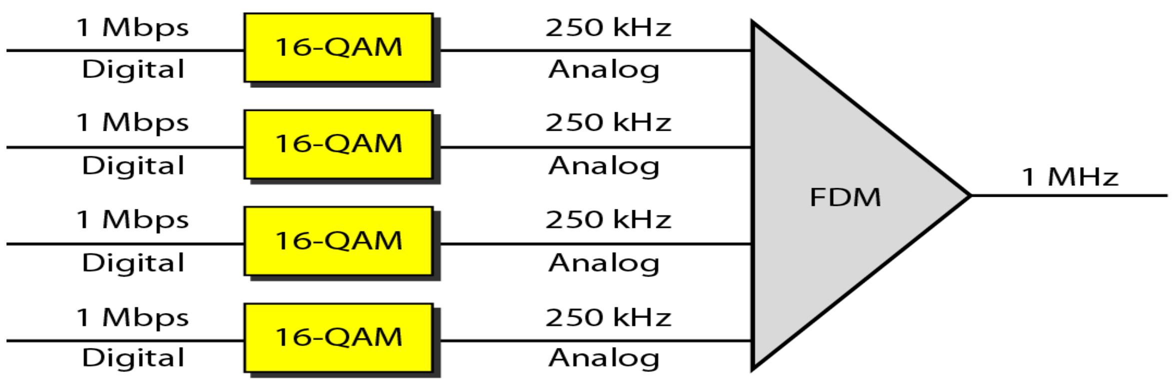
For five channels, we need at least four guard bands. This means that the required bandwidth is at least 5 x 100 + 4 x 10 =540 kHz, as shown in Figure



3) Four data channels (digital), each transmitting at I Mbps, use a satellite channel of I MHz. Design an appropriate configuration, using FDM.

Solution

The satellite channel is analog. We divide it into four channels, each channel having a 250-kHz bandwidth. Each digital channel of 1 Mbps is modulated such that each 4 bits is modulated to 1 Hz. One solution is 16-QAM modulation. Figure shows one possible configuration.

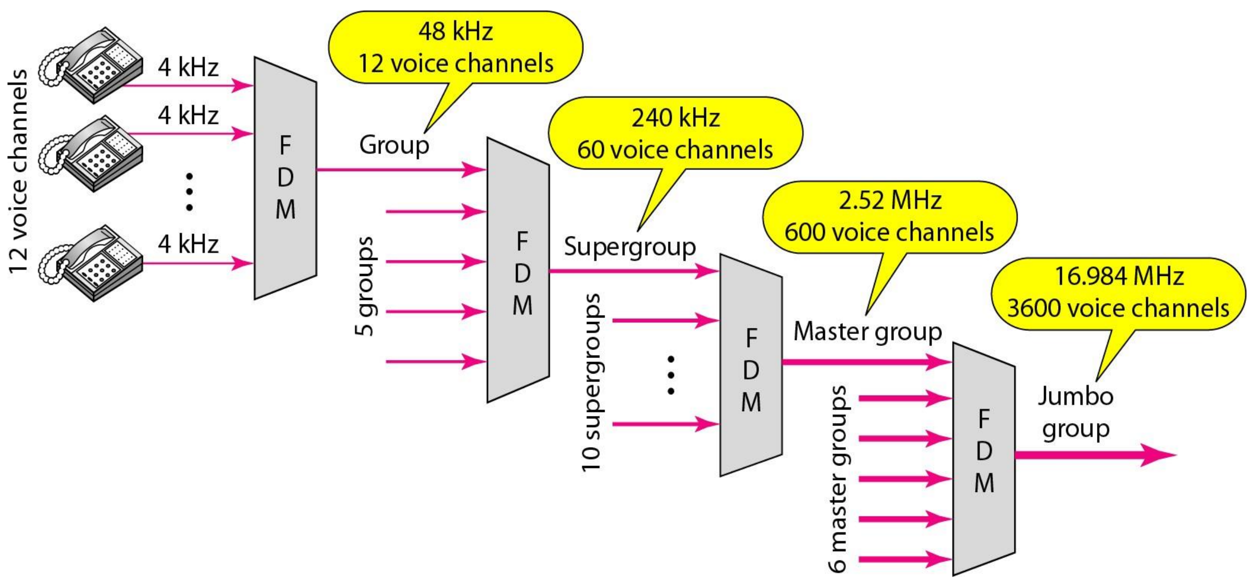


Application:

The Analog Carrier System

- To maximize the efficiency of infrastructure, telephone companies have traditionally multiplexed signals from lower-bandwidth lines onto higher-bandwidth lines. In this way, many switched or leased lines can be combined into fewer but bigger channels. For analog lines, FDM is used.
- One of these hierarchical systems used by AT&T is made up of groups, supergroups, master groups, and jumbo groups.

- In this analog hierarchy, 12 voice channels are multiplexed onto a higher-bandwidth line to create a group. A group has 48 kHz of bandwidth and supports 12 voice channels.
- At the next level, up to five groups can be multiplexed to create a composite signal called a supergroup. A supergroup has a bandwidth of 240 kHz and supports up to 60 voice channels. Supergroups can be made up of either five groups or 60 independent voice channels.
- At the next level, 10 supergroups are multiplexed to create a master group. A master group must have 2.40 MHz of bandwidth, but the need for guard bands between the supergroups increases the necessary bandwidth to 2.52 MHz. Master groups support up to 600 voice channels.
- Finally, six master groups can be combined into a jumbo group. A jumbo group must have 15.12 MHz (6 x 2.52 MHz) but is augmented to 16.984 MHz to allow for guard bands between the master groups.



Other Applications of FDM

- **AM and FM radio broadcasting.**

  Radio uses the air as the transmission medium. A special band from 530 to 1700 kHz is assigned to AM radio. All radio stations need to share this band. Each AM station needs 10 kHz of bandwidth. Each station uses a different carrier frequency, which means it is shifting its signal and multiplexing. The signal that goes to the air is a combination of signals. A receiver receives all these signals, but filters (by tuning) only the one which is desired. Without multiplexing, only one AM station could broadcast to the common link, the air.

However, we need to know that there is physical multiplexer or demultiplexer here. Multiplexing is done at the data link layer.

The situation is similar in FM broadcasting. However, FM has a wider band of 88 to 108 MHz because each station needs a bandwidth of 200 kHz.

- **Television broadcasting.**

  Each TV channel has its own bandwidth of 6 MHz.

- **The first generation of cellular telephones.**

  Each user is assigned two 30-kHz channels, one for sending voice and the other for receiving. The voice signal, which has a bandwidth of 3 kHz (from 300 to 3300 Hz), is modulated by using FM.

Example 4

The Advanced Mobile Phone System (AMPS) uses two bands. The first band of 824 to 849 MHz is used for sending, and 869 to 894 MHz is used for receiving. Each user has a bandwidth of 30 kHz in each direction. The 3-kHz voice is modulated using FM, creating 30 kHz of modulated signal. How many people can use their cellular phones simultaneously?

Solution:

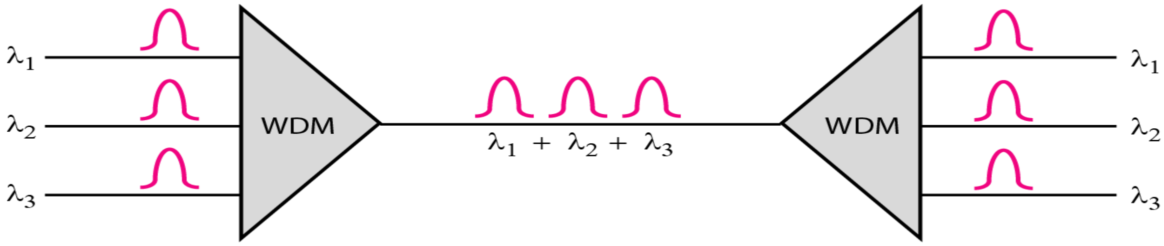
Each band is 25 MHz. If we divide 25 MHz by 30 kHz, we get 833.33.

In reality, the band is divided into 832 channels. Of these, 42 channels are used for control, which means only 790 channels are available for cellular phone users.

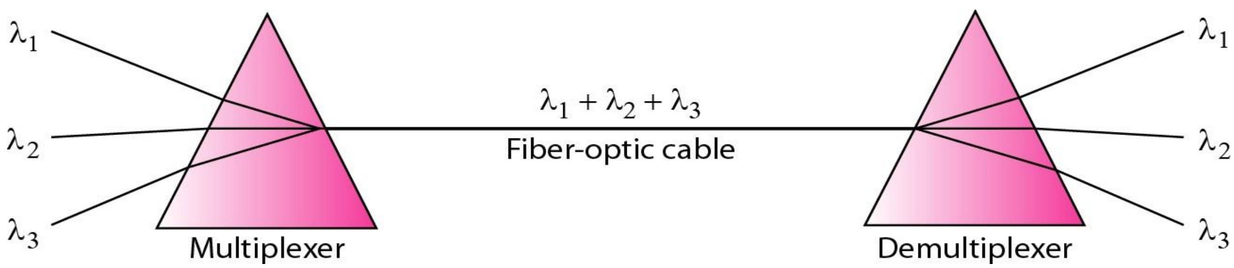
Wavelength-Division Multiplexing

- Wavelength-division multiplexing (WDM) is designed to use the high-data-rate capability of fiber-optic cable.
- The optical fiber data rate is higher than the data rate of metallic transmission cable. Using a fiber-optic cable for one single line wastes the available bandwidth. Multiplexing allows us to combine several lines into one.
- WDM is conceptually the same as FDM, except that the multiplexing and demultiplexing involve optical signals transmitted through fiber-optic channels.

- Very narrow bands of light from different sources are combined to make a wider band of light. At the receiver, the signals are separated by the demultiplexer.



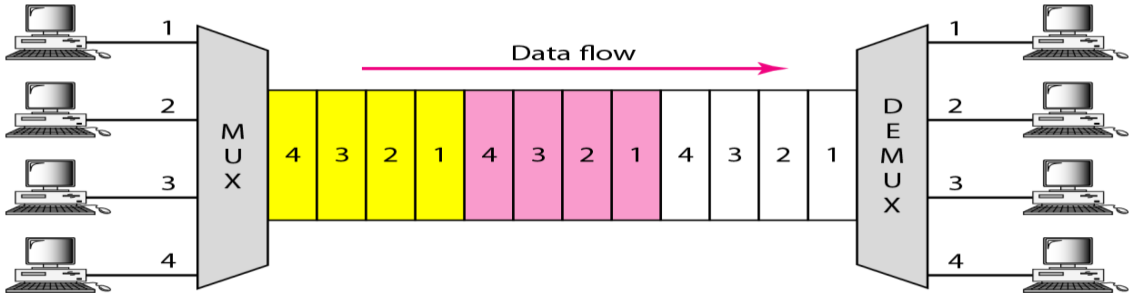
- The combining and splitting of light sources are easily handled by a prism. Prism bends a beam of light based on the angle of incidence and the frequency. Using this technique, a multiplexer can be made to combine several input beams of light, each containing a narrow band of frequencies, into one output beam of a wider band of frequencies. A demultiplexer can also be made to reverse the process.



- One application of WDM is the SONET network in which multiple optical fiber lines are multiplexed and demultiplexed.
- A new method, called dense WDM (DWDM), can multiplex a very large number of channels by spacing channels very close to one another. It achieves even greater efficiency.

Time-Division Multiplexing

- Time-division multiplexing (TDM) is a digital process that allows several connections to share the high bandwidth of a link.
- Instead of sharing a portion of the bandwidth as in FDM, time is shared. Each connection occupies a portion of time in the link.
- TDM is a digital multiplexing technique for combining several low-rate channels into one high-rate one.

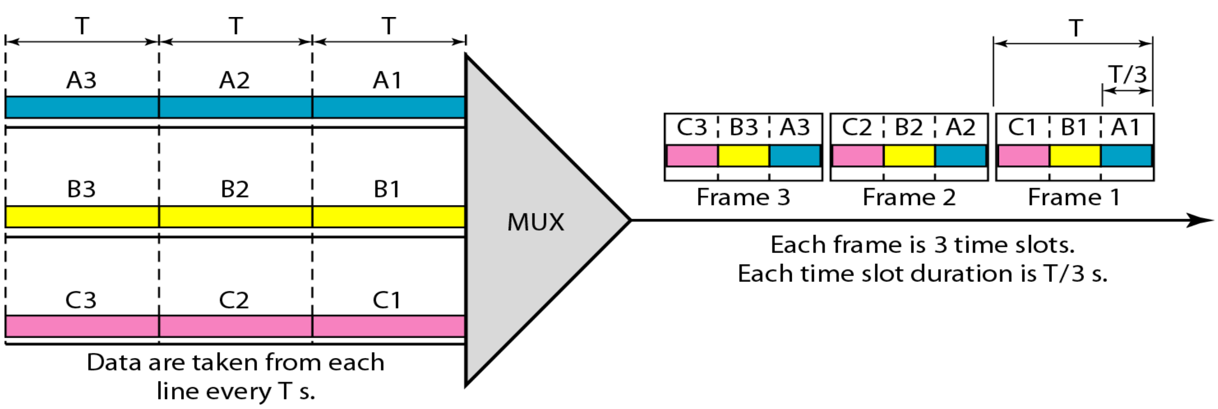


TDM is divided into two different schemes: synchronous and statistical

Synchronous Time-Division Multiplexing

Time Slots and Frames

- In synchronous TDM, the data flow of each input connection is divided into units, where each input occupies one input time slot.
- A unit can be 1 bit, one character, or one block of data.
- Each input unit becomes one output unit and occupies one output time slot.
- However, the duration of an output time slot is *n* times shorter than the duration of an input time slot. If an input time slot is *T* s, the output time slot is *T/n* s, where *n* is the number of connections.

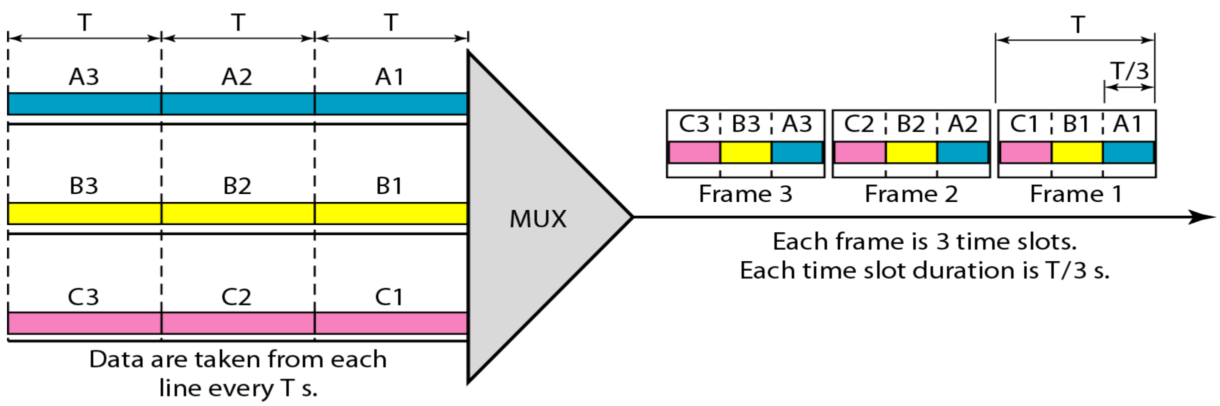


- In synchronous TDM, a round of data units from each input connection is collected into a frame. If we have n connections, a frame is divided into n time slots and one slot is allocated for each unit, one for each input line. If the duration of the input unit is T, the duration of each slot is Tin and the duration of each frame is T.
- **In synchronous TDM, the data rate of the link is n times faster, and the unit duration is n times shorter.**

- Time slots are grouped into frames. A frame consists of one complete cycle of time slots, with one slot dedicated to each sending device. In a system with *n* input lines, each frame has *n* slots, with each slot allocated to carrying data from a specific input line.

Example 5:

In Figure below, the data rate for each one of the 3 input connection is 1 kbps. If 1 bit at a time is multiplexed (a unit is 1 bit), what is the duration of (a) each input slot, (b) each output slot, and (c) each frame?



Solution

We can answer the questions as follows:

a. The data rate of each input connection is 1 kbps. This means that the bit duration is 1/1000 s or 1ms. The duration of the input time slot is 1 ms (same as bit duration).

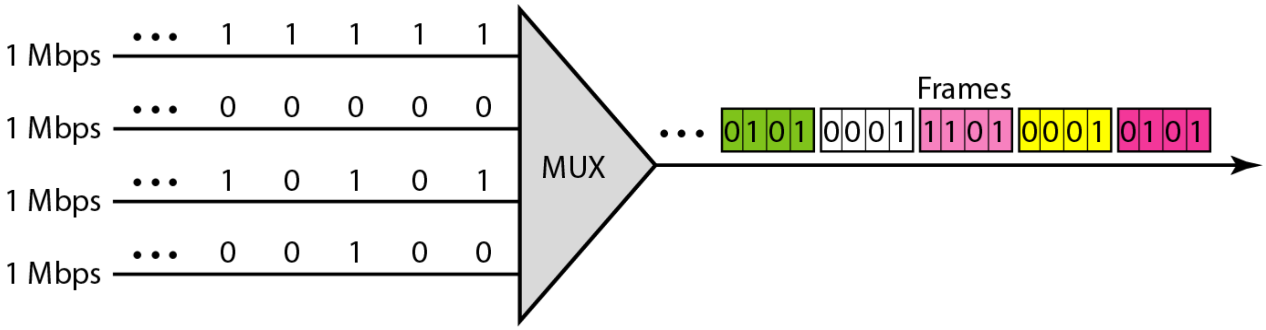
b. The duration of each output time slot is one-third of the input time slot. This means that the duration of the output time slot is 1/3 ms.

c. Each frame carries three output time slots. So the duration of a frame is 3 × 1/3 ms, or 1 ms.

<u>Note</u>: The duration of a frame is the same as the duration of an input unit.

Example 6:

Figure shows synchronous TOM with a data stream for each input and one data stream for the output. The unit of data is 1 bit. Find (a) the input bit duration, (b) the output bit duration, (c) the output bit rate, and (d) the output frame rate.



Solution

a. The input bit duration is the inverse of the bit rate: 1/1 Mbps = 1 μs.

b. The output bit duration is one-fourth of the input bit duration, or 1/4μs.

c. The output bit rate is the inverse of the output bit duration or 1/4 μs, or 4 Mbps. This can also be deduced from the fact that the output rate is 4 times as fast as any input rate; so the output rate =4 x 1 Mbps =4 Mbps.

d. The frame rate is always the same as any input rate. So the frame rate is 1,000,000 frames per second. Because we are sending 4 bits in each frame, we can verify the result of the previous question by multiplying the frame rate by the number of bits per frame.

Example 7:

Four 1-kbps connections are multiplexed together. A unit is 1 bit. Find (a) the duration of 1 bit before multiplexing, (b) the transmission rate of the link, (c) the duration of a time slot, and (d) the duration of a frame

Solution

We can answer the questions as follows:

a. The duration of 1 bit before multiplexing is 1 / 1 kbps, or 0.001 s (1 ms).

b. The rate of the link is 4 times the rate of a connection, or 4 kbps.

c. The duration of each time slot is one-fourth of the duration of each bit before multiplexing, or 1/4 ms or 250 μs. Note that we can also calculate this from the data rate of the link, 4 kbps. The bit duration is the inverse of the data rate, or 1/4 kbps or 250 μs.

d. The duration of a frame is always the same as the duration of a unit before multiplexing, or 1 ms. We can also calculate this in another way. Each frame in this case has four time slots. So the duration of a frame is 4 times 250 μs, or 1 ms.

Interleaving

- TDM can be visualized as two fast-rotating switches, one on the multiplexing side and the other on the demultiplexing side.
- The switches are synchronized and rotate at the same speed, but in opposite directions.
- On the multiplexing side, as the switch opens in front of a connection, that connection has the opportunity to send a unit onto the path. This process is called **interleaving.**
- On the demultiplexing side, as the switch opens in front of a connection, that connection has the opportunity to receive a unit from the path.
- Figure 6.15 shows the interleaving process for the connection shown in Figure for synchronous TDM. In this figure, we assume that no switching is involved and that the data from the first connection at the multiplexer site go to the first connection at the demultiplexer.

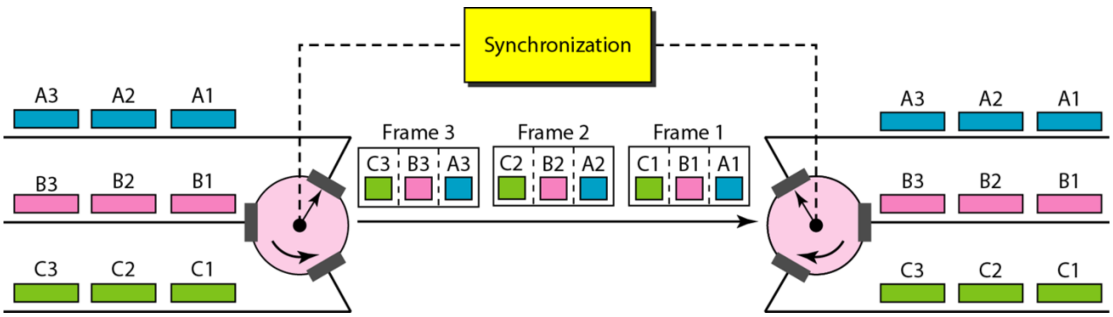


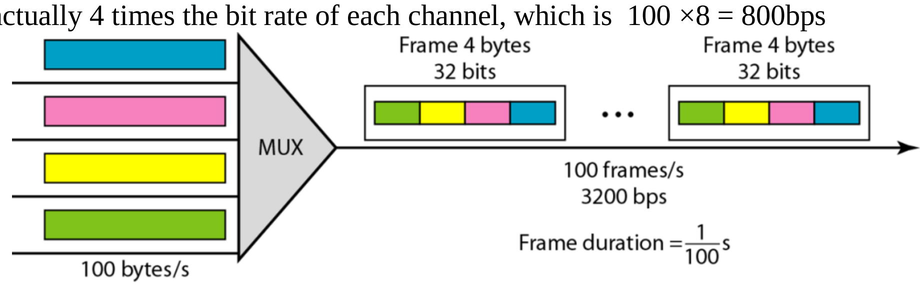
Figure 6.15 Interleaving

Example: 8

Four channels are multiplexed using TDM. If each channel sends 100 bytes /s and we multiplex 1 byte per channel, show the frame traveling on the link, the size of the frame, the duration of a frame, the frame rate, and the bit rate for the link.

Solution

- The multiplexer is shown in Figure 6.16.
- Each frame carries 1 byte from each channel; the size of each frame, therefore, is 4 bytes, or 32 bits. B
- ecause each channel is sending 100 bytes/s and a frame carries 1 byte from each channel, the frame rate must be 100 frames per second.
- The duration of a frame is therefore 1/100s. The link is carrying 100 frames per second, and since the frame contains 32 bits. The bit rate is 100 × 32, or 3200 bps.
- This is actually 4 times the bit rate of each channel, which is 100 ×8 = 800bps

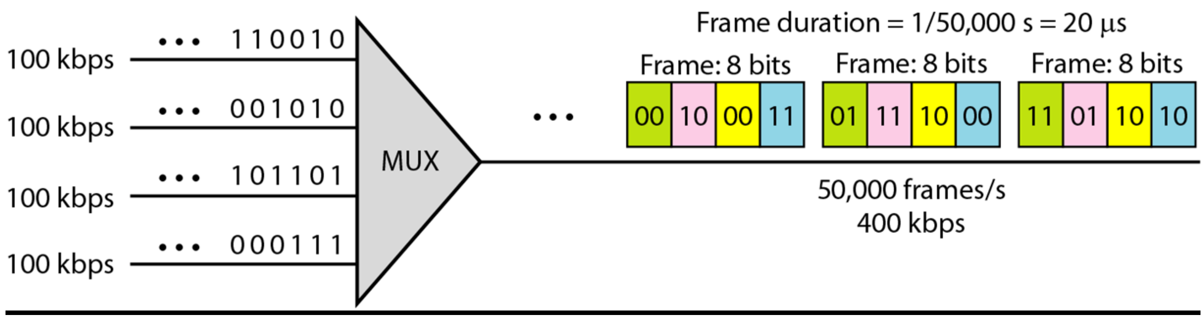


Example 9

A multiplexer combines four 100-kbps channels using a time slot of 2 bits. Show the output with four arbitrary inputs. What is the frame rate? What is the frame duration? What is the bit rate? What is the bit duration?

Solution:

- Figure 6.17 shows the output for four arbitrary inputs.
- The link carries 50,000 frames per second since each frame contains 2 bits per channel.
- The frame duration is therefore 1/50,000 s or 20 μs.
- The frame rate is 50,000 frames per second, and each frame carries 8 bits; the bit rate is 50,000 × 8 = 400,000 bits or 400 kbps. The bit duration is 1/400,000 s, or 2.5 μs.
- Note that the frame duration is 8 times the bit duration because each frame is carrying 8 bits.

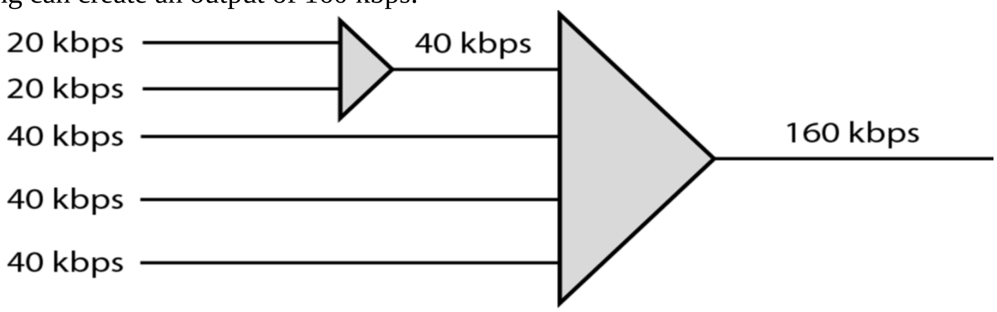


Data Rate Management

- One problem with TDM is how to handle a disparity in the input data rates. In all our discussion so far, we assumed that the data rates of all input lines were the same.
- However, if data rates are not the same, three strategies, or a combination of them, can be used.
- Three different strategies: 1) Multilevel multiplexing 2) Multiple-slot allocation and 3) Pulse stuffing

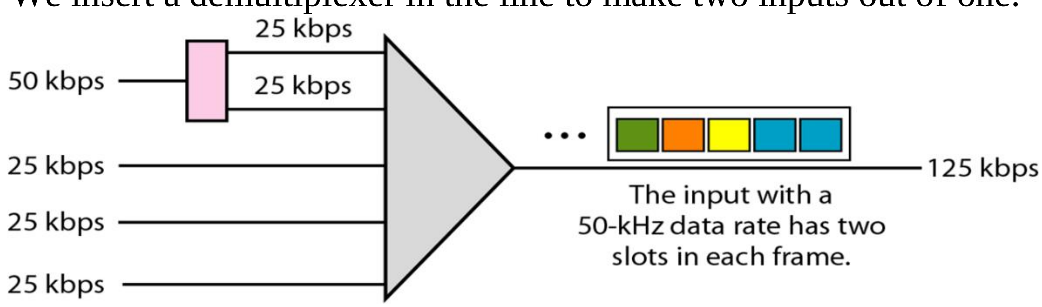
Multilevel Multiplexing

- Multilevel multiplexing is a technique used when the data rate of an input line is a multiple of others.
- For example, in Figure 6.19, we have two inputs of 20 kbps and three inputs of 40 kbps. The first two input lines can be multiplexed together to provide a data rate equal to the last three. A second level of multiplexing can create an output of 160 kbps.



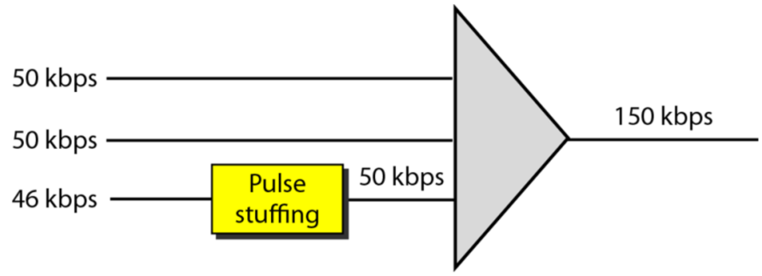
Multiple-Slot Allocation

- Sometimes it is more efficient to allot more than one slot in a frame to a single input line.
- For example, we might have an input line that has a data rate that is a multiple of another input.
- In Figure 6.20, the input line with a 50-kbps data rate can be given two slots in the output. We insert a demultiplexer in the line to make two inputs out of one.

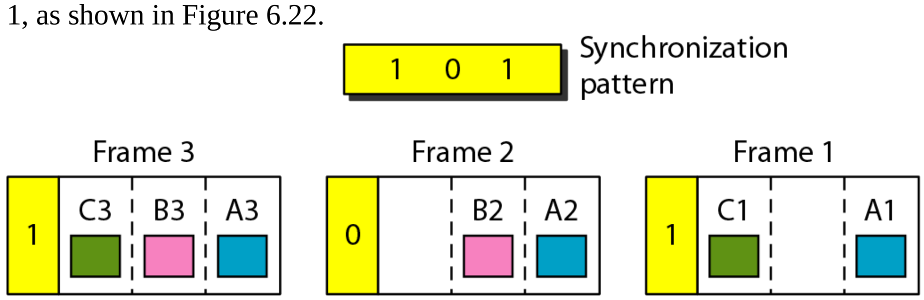


Pulse Stuffing

- Sometimes the bit rates of sources are not multiple integers of each other.
- Therefore, neither of the above two techniques can be applied.
- One solution is to make the highest input data rate the dominant data rate and then add dummy bits to the input lines with lower rates. This will increase their rates. This technique is called pulse stuffing, bit padding, or bit stuffing.
- The idea is shown in Figure 6.21. The input with a data rate of 46 is pulse-stuffed to increase the rate to 50 kbps. Now multiplexing can take place.



Frame Synchronizing

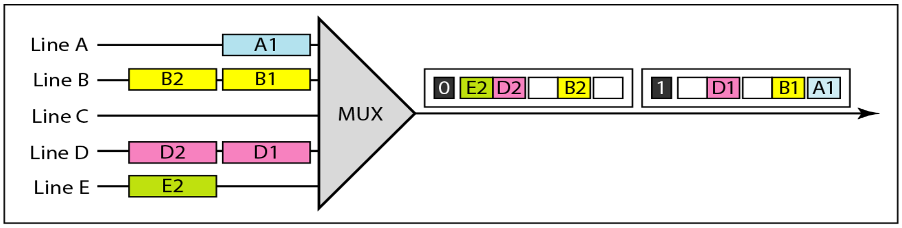
- The implementation of TDM is not as simple as that of FDM.
- Synchronization between the multiplexer and demultiplexer is a major issue.
- If the multiplexer and the demultiplexer are not synchronized, a bit belonging to one channel may be received by the wrong channel.
- For this reason, one or more synchronization bits are usually added to the beginning of each frame. These bits, called framing bits, follow a pattern, frame to frame, that allows the demultiplexer to synchronize with the incoming stream so that it can separate the time slots accurately.
- n most cases, this synchronization information consists of 1 bit per frame, alternating between 0 and 1, as shown in Figure 6.22.
- 

Statistical Time-Division Multiplexing

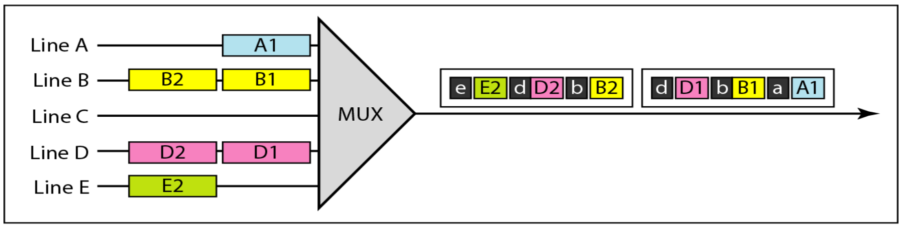
In statistical time-division multiplexing, slots are dynamically allocated to improve bandwidth efficiency.

Only when an input line has a slot's worth of data to send is it given a slot in the output frame. In statistical multiplexing, the number of slots in each frame is less than the number of input lines.

The multiplexer checks each input line in round robin fashion; it allocates a slot for an input line if the line has data to send; otherwise, it skips the line and checks the next line.



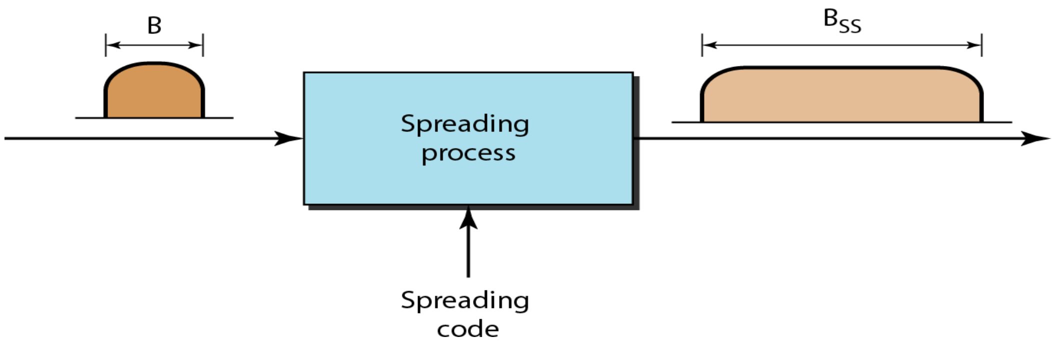
a. Synchronous TDM



b. Statistical TDM

3.2 Spread Spectrum

- Multiplexing combines signals from several sources to achieve bandwidth efficiency; the available bandwidth of a link is divided between the sources.
- In spread spectrum, we also combine signals from different sources to fit into a larger bandwidth.
- Spread spectrum is designed to be used in wireless applications (LANs and WANs).
- In wireless applications, all stations use air (or a vacuum) as the medium for communication. Stations must be able to share this medium without interception by an eavesdropper and without being subject to jamming from a malicious intruder (in military operations, for example).
- If the required bandwidth for each station is B, spread spectrum expands it to Bss such that Bss » B. The expanded bandwidth allows the source to wrap its message in a protective envelope for a more secure transmission.



Spread spectrum achieves its goals through two principles:

1. The bandwidth allocated to each station needs to be, by far, larger than what is needed. This allows redundancy.
2. The expanding of the original bandwidth *B* to the bandwidth Bss must be done by a process that is independent of the original signal. In other words, the spreading process occurs after the signal is created by the source.

After the signal is created by the source, the spreading process uses a spreading code and spreads the bandwidth. The figure shows the original bandwidth B and the spreaded bandwidth Bss. The spreading code is a series of numbers that look random, but are actually a pattern.

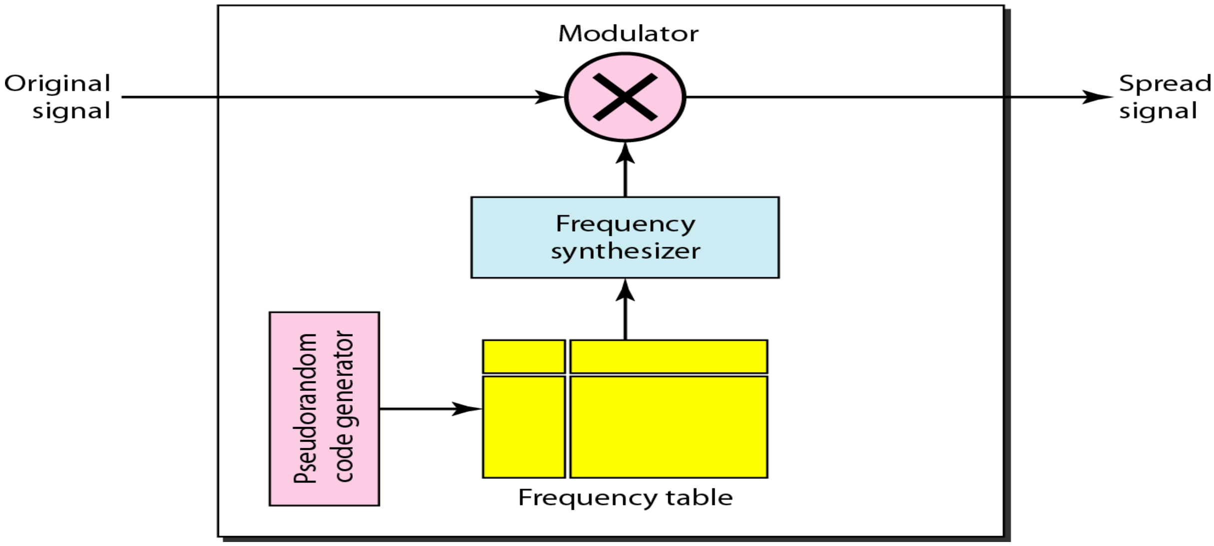
There are two techniques to spread the bandwidth:

1. Frequency hopping spread spectrum (FHSS)
2. Direct sequence spread spectrum (DSSS)

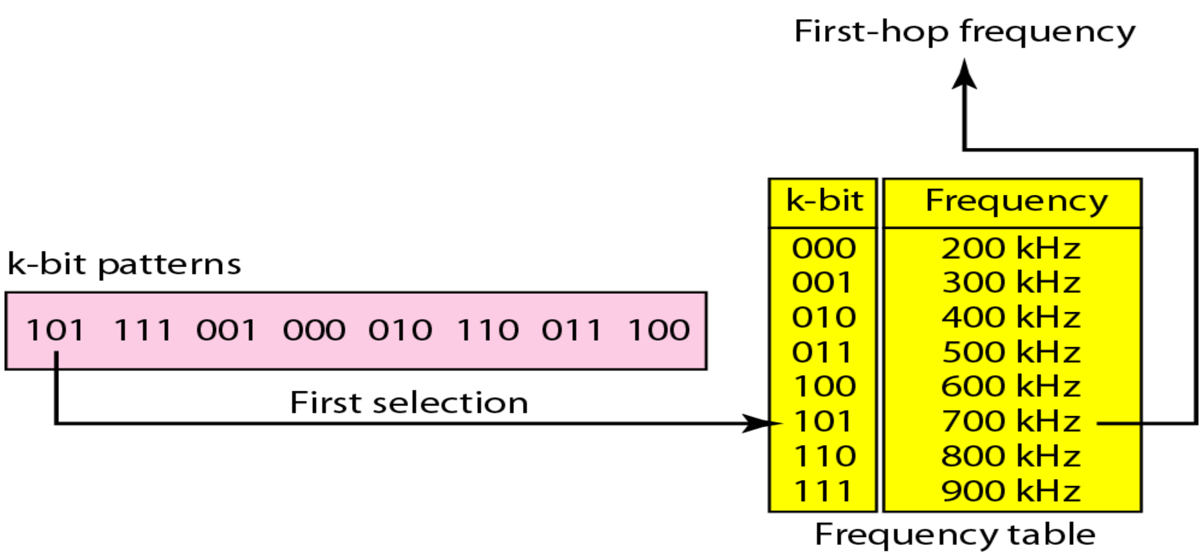
Frequency Hopping Spread Spectrum (FHSS)

- The frequency hopping spread spectrum (FHSS) technique uses *M* different carrier frequencies that are modulated by the source signal.
- At one moment, the signal modulates one carrier frequency; at the next moment, the signal modulates another carrier frequency. Although the modulation is done using one carrier frequency at a time, *M* frequencies are used in the long run.
- The bandwidth occupied by a source after spreading is *BpHSS »B.*

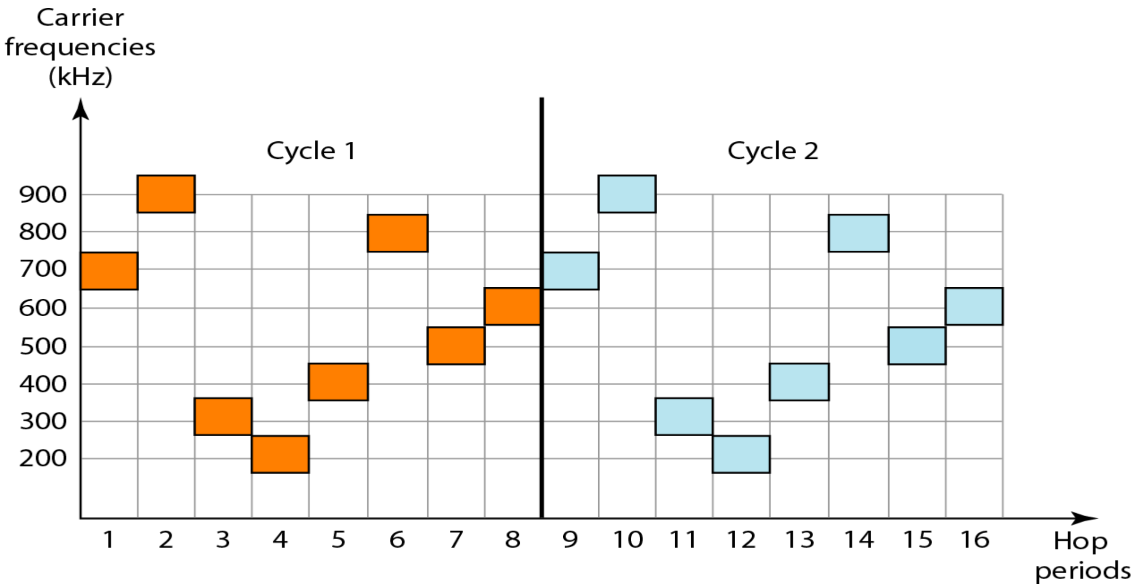
The general layout for FHSS is shown below:



- A pseudorandom code generator, called pseudorandom noise (PN), creates a k-bit pattern for every hopping period *Th.*
- The frequency table uses the pattern to find the frequency to be used for this hopping period and passes it to the frequency synthesizer.
- The frequency synthesizer creates a carrier signal of that frequency, and the source signal modulates the carrier signal.
- Suppose we have decided to have eight hopping frequencies. This is extremely low for real applications and is just for illustration. In this case, M is 8 and k is 3. The pseudorandom code generator will create eight different 3-bit patterns. These are mapped to eight different frequencies in the frequency table.

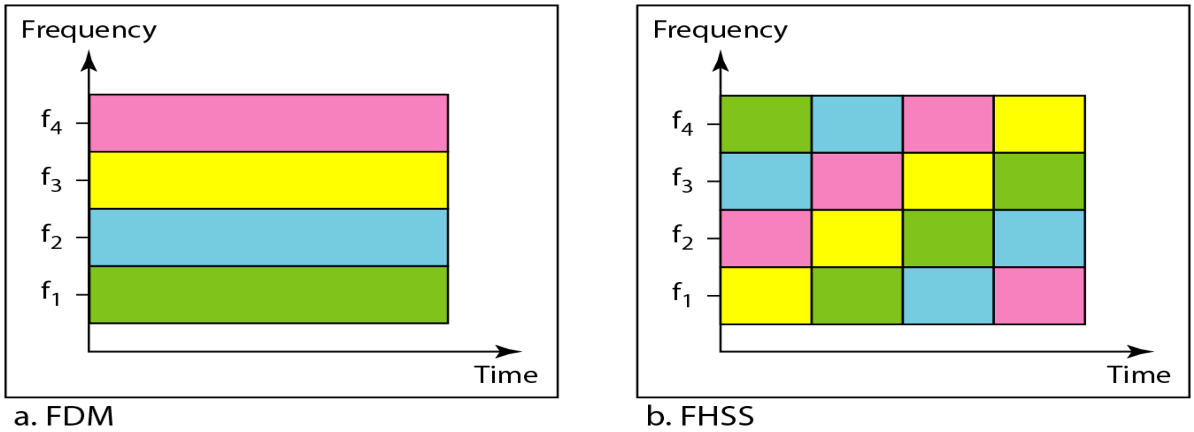


The pattern for this station is 101, 111, 001, 000, 010, all, 100. Note that the pattern is pseudorandom it is repeated after eight hoppings. This means that at hopping period 1, the pattern is 101. The frequency selected is 700 kHz; the source signal modulates this carrier frequency. The second k-bit pattern selected is 111, which selects the 900-kHz carrier; the eighth pattern is 100, the frequency is 600 kHz. After eight hoppings, the pattern repeats, starting from 101 again. Figure shows how the signal hops around from carrier to carrier. We assume the required bandwidth of the original signal is 100 kHz.



Bandwidth Sharing

If the number of hopping frequencies is *M*, we can multiplex *M* channels into one by using the same Bss bandwidth. This is possible because a station uses just one frequency in each hopping period; *M* - 1 other frequencies can be used by other *M* - 1 stations. In other words, *M* different stations can use the same Bss if an appropriate modulation technique such as multiple FSK (MFSK) is used.

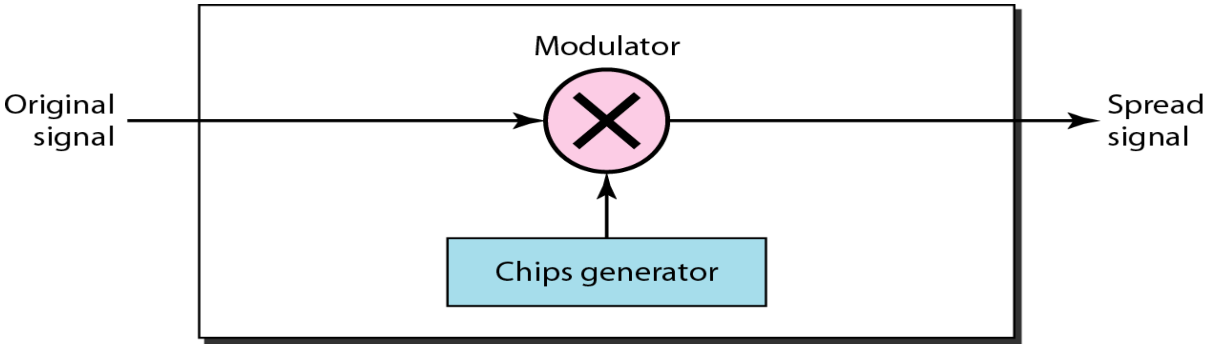


Above Figure shows an example of four channels using FDM and four channels using FHSS. In FDM, each station uses 11M of the bandwidth, but the allocation is fixed; in FHSS, each station uses 11M of the bandwidth, but the allocation changes hop to hop.

Direct Sequence Spread Spectrum

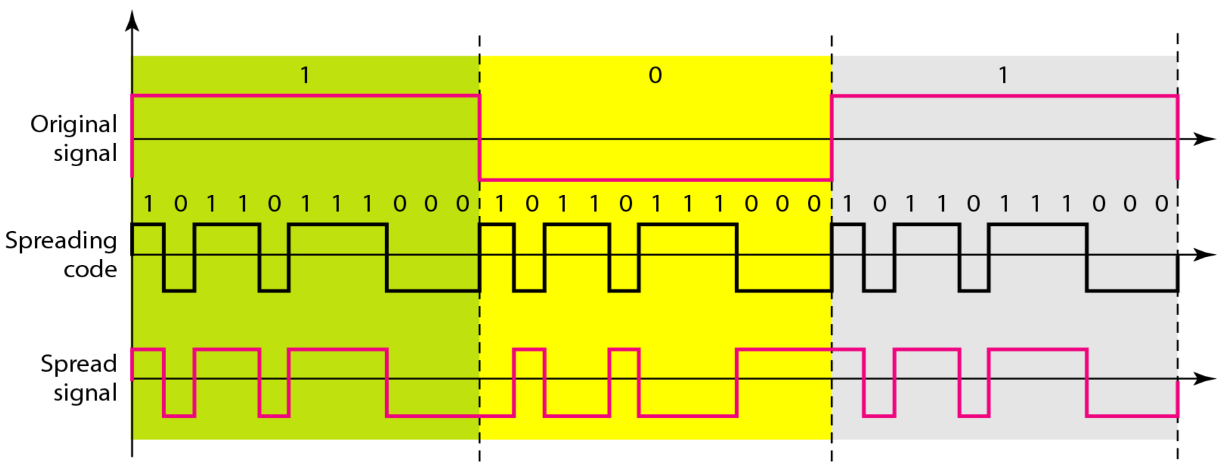
The direct sequence spread spectrum (DSSS) technique also expands the bandwidth of the original signal, but the process is different.

In DSSS, we replace each data bit with 11 bits using a spreading code. In other words, each bit is assigned a code of 11 bits, called **chips**, where the chip rate is 11 times that of the data bit.



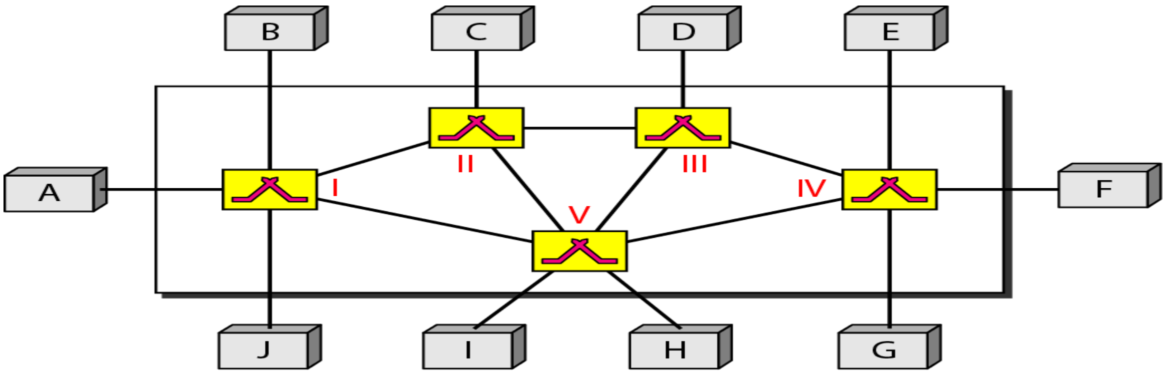
- As an example, let us consider the sequence used in a wireless LAN, the famous Barker sequence where 11 is 11.
- We assume that the original signal and the chips in the chip generator use polar NRZ encoding. Below Figure shows the chips and the result of multiplying the original data by the chips to get the spread signal.

- In Figure, the spreading code is 11 chips having the pattern 10110111000 (in this case). If the original signal rate is N, the rate of the spread signal l1N.
- This means that the required bandwidth for the spread signal is 11 times larger than the bandwidth of the original signal.
- The spread signal can provide privacy if the intruder does not know the code. It can also provide immunity against interference if each station uses a different code.



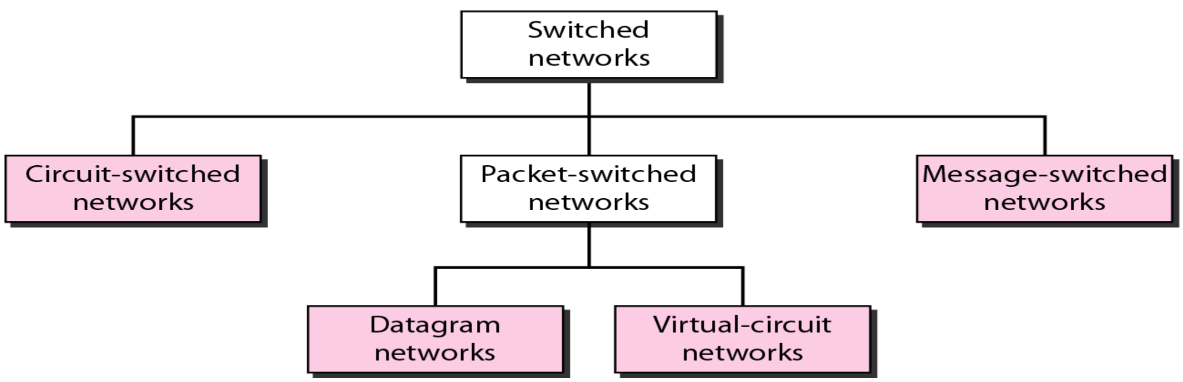
Switching

- A switched network consists of a series of interlinked nodes, called **switches.**
- Switches are devices capable of creating temporary connections between two or more devices linked to the switch.
- In a switched network, some of these nodes are connected to the end systems. Others are used only for routing.



- Switched networks can be divided into three broad categories: circuit-switched networks, packet- switched networks, and message-switched.

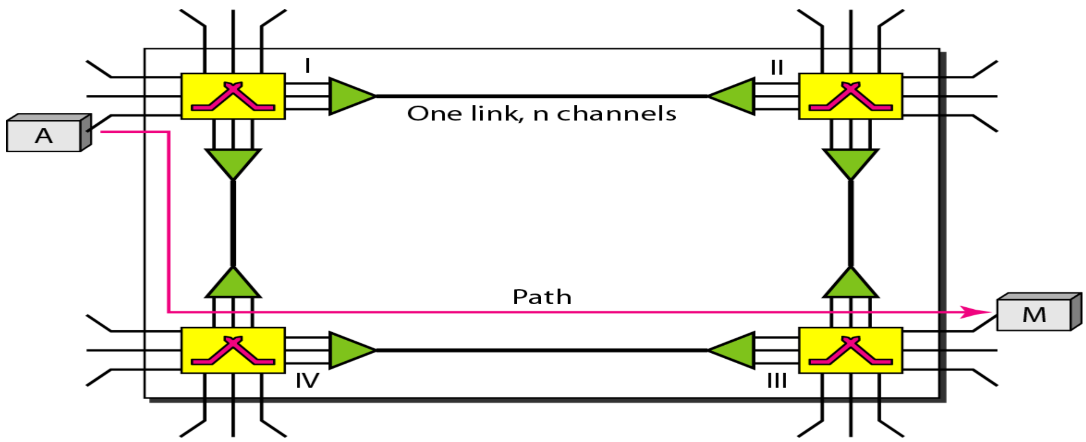
- Packet-switched networks can further be divided into two subcategories-virtual-circuit networks and datagram networks.



3.3 Circuit Switched Networks

- A circuit-switched network consists of a set of switches connected by physical links.
- A connection between two stations is a dedicated path made of one or more links. However, each connection uses only one dedicated channel on each link.

- Each link is normally divided into *n* channels by using FDM or TDM.
- Below Figure shows a trivial circuit-switched network with four switches and four links. Each link is divided into n (n is 3 in the figure) channels by using FDM or TDM.



- Circuit switching takes place at the physical layer.
- Before starting communication, the stations must make a reservation for the resources to be used during the communication. These resources, such as channels (bandwidth in FDM and time slots in TDM), switch buffers, switch processing time, and switch input/output ports, must remain dedicated during the entire duration of data transfer until the teardown phase.
- Data transferred between the two stations are not packetized. The data are a continuous flow sent by the source station and received by the destination station, although there may be periods of silence.
- There is no addressing involved during data transfer. The switches route the data based on their occupied band (FDM) or time slot (TDM).

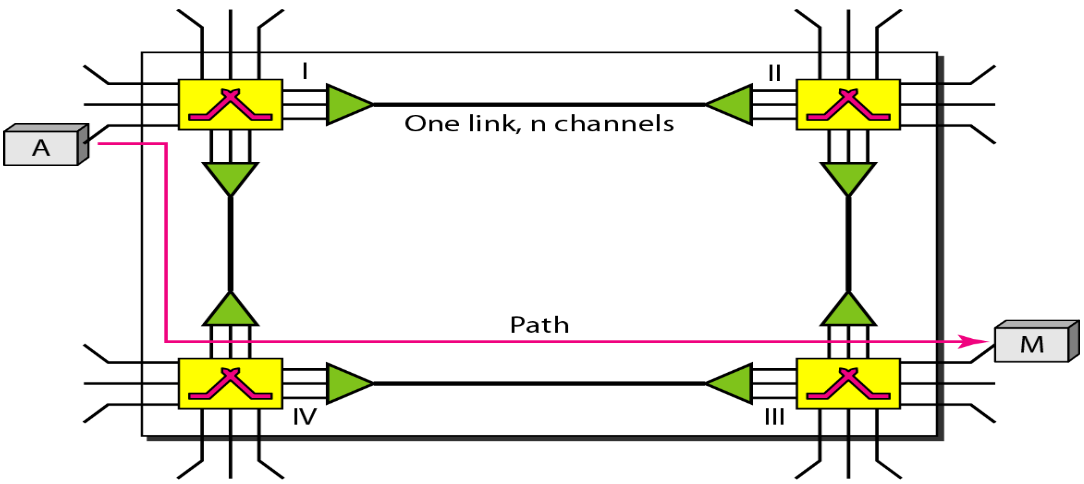
Three Phases

The actual communication in a circuit-switched network requires three phases: connection setup, data transfer, and connection teardown.

***1)* Setup Phase**

- Before the two parties can communicate, a dedicated circuit needs to be established. The end systems are normally connected through dedicated lines to the switches, so connection setup means creating dedicated channels between the switches.

- In the next step to making a connection, an acknowledgment from system M needs to be sent in the opposite direction to system A.
- Only after system A receives this acknowledgment is the connection established.



2) Data Transfer Phase

After the establishment of the dedicated circuit (channels), the two parties can transfer data.

3) Teardown Phase

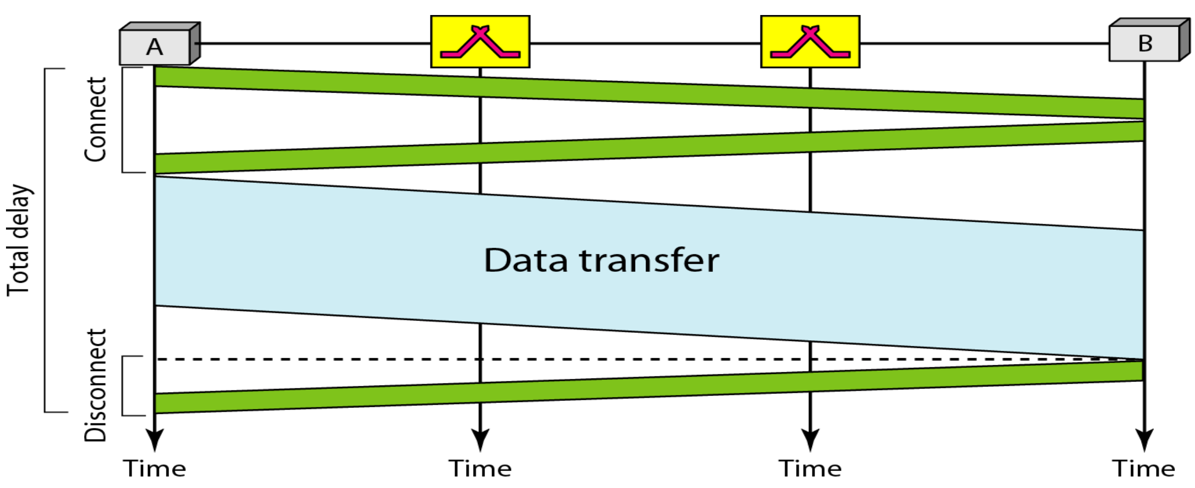
When one of the parties needs to disconnect, a signal is sent to each switch to release the resources.

Efficiency

Circuit-switched networks are not as efficient as the other two types of networks because resources are allocated during the entire duration of the connection. These resources are unavailable to other connections.

Delay

- Although a circuit-switched network normally has low efficiency, the delay in this type of network is minimal.
- During data transfer the data are not delayed at each switch; the resources are allocated for the duration of the connection.



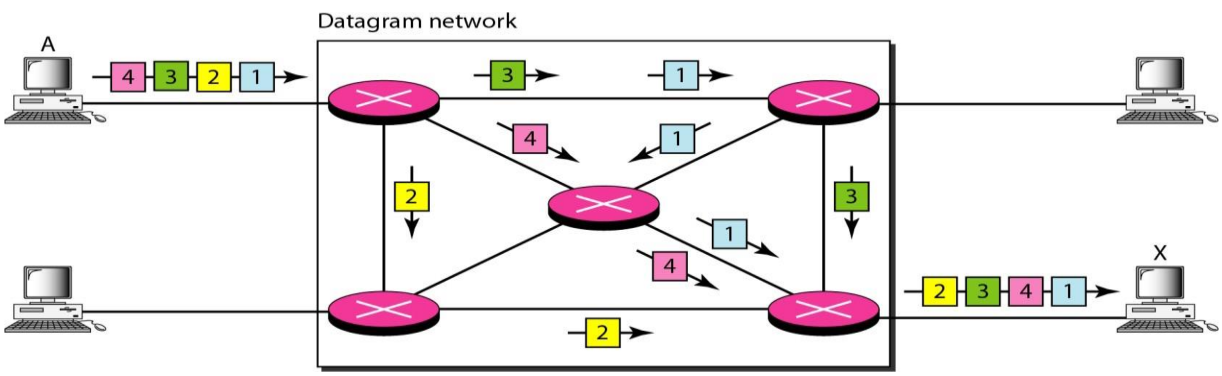
- As Figure shows, there is no waiting time at each switch. The total delay is due to the time needed to create the connection, transfer data, and disconnect the circuit.
- The delay caused by the setup is the sum of four parts: the propagation time of the source computer request, the request signal transfer time, the propagation time of the acknowledgment from the destination computer, and the signal transfer time of the acknowledgment.
- The delay due to data transfer is the sum of two parts: the propagation time and data transfer time, which can be very long.
- The third box shows the time needed to tear down the circuit.

3.4 Datagram Networks

- In data communications, we need to send messages from one end system to another.
- If the message is going to pass through a packet-switched network, it needs to be divided into packets of fixed or variable size.
- The size of the packet is determined by the network and the governing protocol.
- In packet switching, there is no resource allocation for a packet. This means that there is no reserved bandwidth on the links, and there is no scheduled processing time for each packet.
- Resources are allocated on demand. The allocation is done on a first come, first-served basis. When a switch receives a packet, no matter what is the source or destination, the packet must wait if there are other packets being processed.

- In a datagram network, each packet is treated independently of all others. Even if a packet is part of a multipacket transmission, the network treats it as though it existed alone. Packets in this approach are referred to as datagrams.
- Datagram switching is normally done at the network layer.
- The datagram networks are sometimes referred to as connectionless networks.
- The term *connectionless* here means that the switch (packet switch) does not keep information about the connection state. There are no setup or teardown phases. Each packet is treated the same by a switch regardless of its source or destination.

Example: Below Figure shows how the datagram approach is used to deliver four packets from station A to station X. The switches in a datagram network are traditionally referred to as routers.



In this example, all four packets (or datagrams) belong to the same message, but may travel different paths to reach their destination.

This is so because the links may be involved in carrying packets from other sources and do not have the necessary bandwidth available to carry all the packets from A to X.

This approach can cause the datagrams of a transmission to arrive at their destination out of order with different delays between the packets.

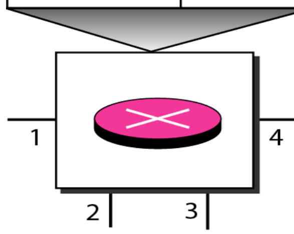
Packets may also be lost or dropped because of a lack of resources.

Routing Table

- Each switch (or packet switch) has a routing table which is based on the destination address. The routing tables are dynamic and are updated periodically.

- The destination addresses and the corresponding forwarding output ports are recorded in the tables.

Destination address	Output port
1232	1
4150	2
⋮	⋮
9130	3



Destination Address

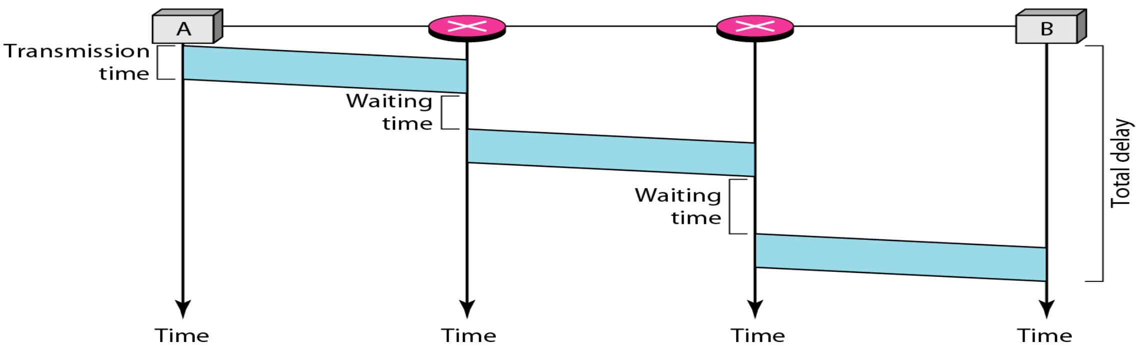
- Every packet in a datagram network carries a header that contains the destination address of the packet.
- When the switch receives the packet, this destination address is examined; the routing table is consulted to find the corresponding port through which the packet should be forwarded.
- The destination address in the header of a packet in a datagram network remains the same during the entire journey of the packet.

Efficiency

- The efficiency of a datagram network is better than that of a circuit-switched network; resources are allocated only when there are packets to be transferred.
- If a source sends a packet and there is a delay of a few minutes before another packet can be sent, the resources can be reallocated during these minutes for other packets from other sources.

Delay

- There may be greater delay in a datagram network than in a virtual-circuit network.
- Although there are no setup and teardown phases, each packet may experience a wait at a switch before it is forwarded.
- In addition, since not all packets in a message necessarily travel through the same switches, the delay is not uniform for the packets of a message.
- Below figure shows an example of delay in a datagram network for one single packet.



- The packet travels through two switches. There are three transmission times (3T), three propagation delays (slopes 3τ of the lines), and two waiting times (W1 + W2)' We ignore the processing time in each switch.

The total delay is

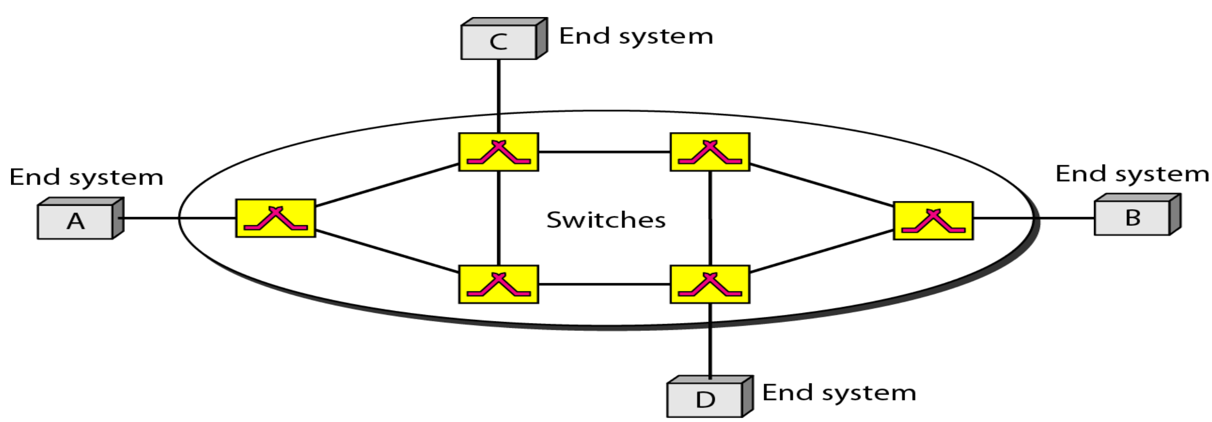
Total delay =3T + 3τ + W1 + W2

3.5 Virtual-Circuit Networks

A virtual-circuit network is a cross between a circuit-switched network and a datagram network. It has some characteristics of both.

1. As in a circuit-switched network, there are setup and teardown phases in addition to the data transfer phase.
2. Resources can be allocated during the setup phase, as in a circuit-switched network, or on demand, as in a datagram network.
3. As in a datagram network, data are packetized and each packet carries an address in the header. However, the address in the header has local jurisdiction, not end-to-end jurisdiction.
4. As in a circuit-switched network, all packets follow the same path established during the connection.
5. A virtual-circuit network is normally implemented in the data link layer, while a circuit-switched network is implemented in the physical layer and a datagram network in the network layer.

Below Figure is an example of a virtual-circuit network. The network has switches that allow traffic from sources to destinations. A source or destination can be a computer, packet switch, bridge, or any other device that connects other networks.



Addressing

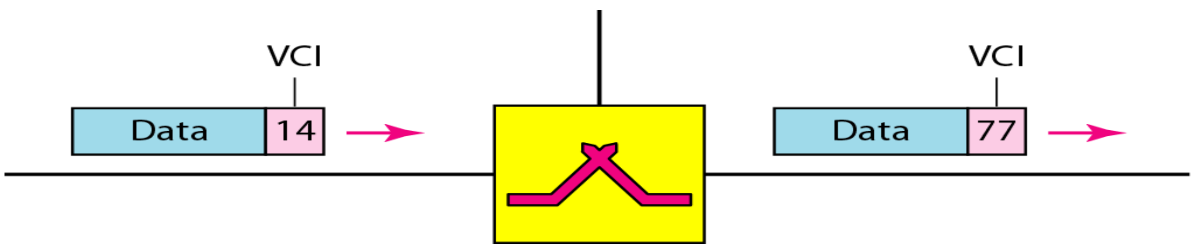
In a virtual-circuit network, two types of addressing are involved: global and local (virtual-circuit identifier).

Global Addressing

A source or a destination needs to have a global address-an address that can be unique in the scope of the network or internationally if the network is part of an international network.

Virtual-Circuit Identifier

- The identifier that is actually used for data transfer is called the virtual-circuit identifier (VCI).
- A VCI is a small number that has only switch scope-it is used by a frame between two switches. When a frame arrives at a switch, it has a VCI; when it leaves, it has a different VCI.
- Below Figure shows how the VCI in a data frame changes from one switch to another.



Three Phases

Data Transfer Phase

- To transfer a frame from a source to its destination, all switches need to have a table entry for this virtual circuit.
- The table, in its simplest form, has four columns.
- This means that the switch holds four pieces of information for each virtual circuit that is already set up.

Below Figure shows such a switch and its corresponding table.

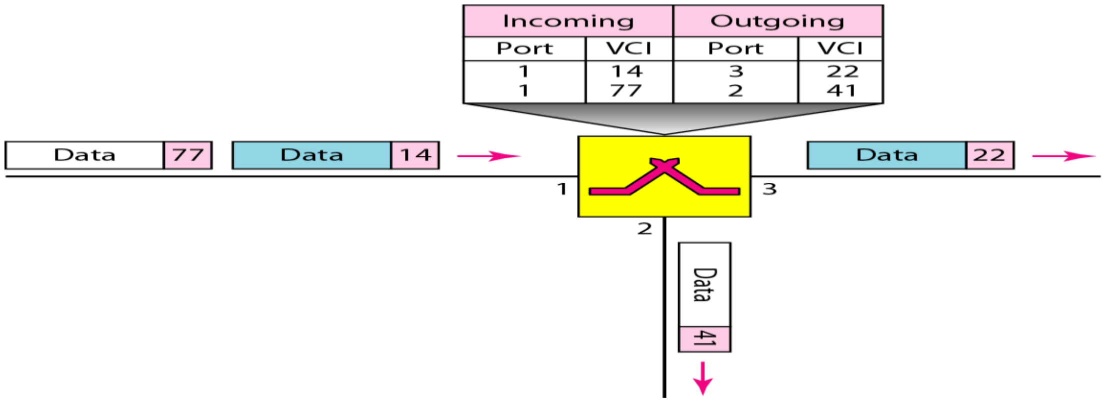
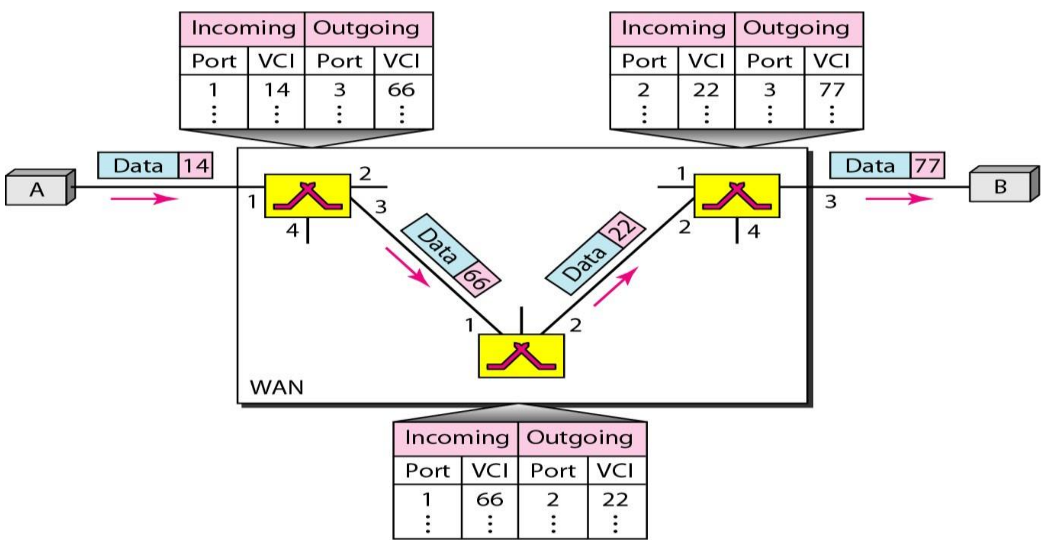


Figure shows a frame arriving at port 1 with a VCI of 14.

- When the frame arrives, the switch looks in its table to find port 1 and a VCI of 14.
- When it is found, the switch knows to change the VCI to 22 and send out the frame from port 3.

Below Figure shows how a frame from source A reaches destination B and how its VCI changes during the trip. Each switch changes the VCI and routes the frame.



- The data transfer phase is active until the source sends all its frames to the destination.
- The procedure at the switch is the same for each frame of a message.
- The process creates a virtual circuit, not a real circuit, between the source and destination.

Setup Phase

In the setup phase, a switch creates an entry for a virtual circuit.

For example, suppose source A needs to create a virtual circuit to B.

Two steps are required: the **setup request** and the **acknowledgment.**

Setup Request: A setup request frame is sent from the source to the destination.

Setup request in a virtual-circuit network

a. Source A sends a setup frame to switch 1.

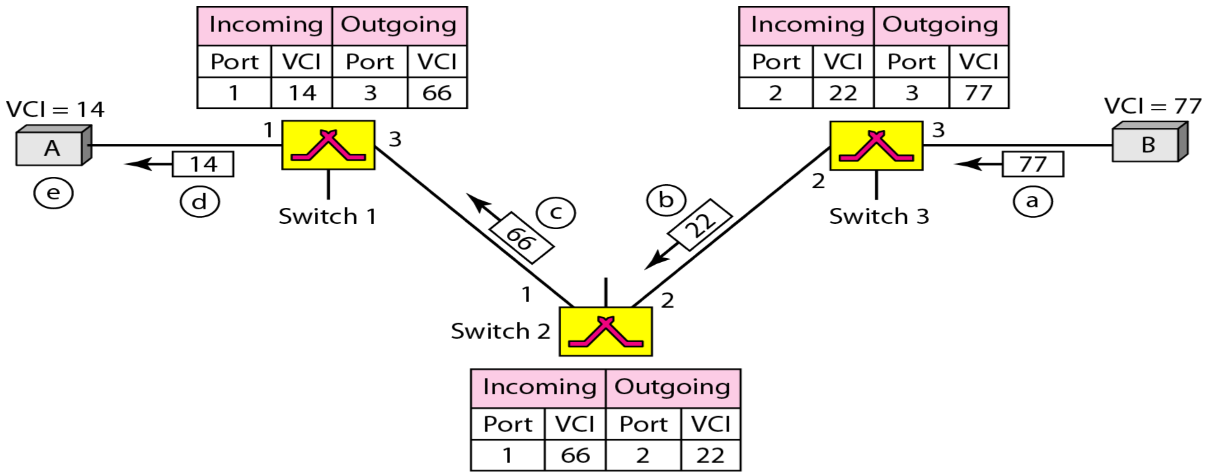
b. Switch 1 receives the setup request frame. It knows that a frame going from A to B goes out through port 3. The switch, in the setup phase, acts as a packet switch; it has a routing table which is different from the switching table. For the moment, assume that it knows the output port. The switch creates an entry in its table for this virtual circuit, but it is only able to fill three of the four columns. The switch assigns the incoming port (1) and chooses an available incoming VCI (14) and the outgoing port (3). It does not yet know the outgoing VCI, which will be found during the acknowledgment step. The switch then forwards the frame through port 3 to switch 2.

c. Switch 2 receives the setup request frame. The same events happen here as at switch 1; three columns of the table are completed: in this case, incoming port (l), incoming VCI (66), and outgoing port (2).

d. Switch 3 receives the setup request frame. Again, three columns are completed: incoming port (2), incoming VCI (22), and outgoing port (3).

e. Destination B receives the setup frame, and if it is ready to receive frames from A, it assigns a VCI to the incoming frames that come from A, in this case 77. This VCI lets the destination know that the frames come from A, and not other sources.

Acknowledgment: A special frame, called the acknowledgment frame, completes the entries in the switching tables.



Setup acknowledgment in a virtual-circuit network

a. The destination sends an acknowledgment to switch 3. The acknowledgment carries the global source and destination addresses so the switch knows which entry in the table is to be completed. The frame also carries VCI 77, chosen by the destination as the incoming VCI for frames from A. Switch 3 uses this VCI to complete the outgoing VCI column for this entry. Note that 77 is the incoming VCI for destination B, but the outgoing VCI for switch 3.

b. Switch 3 sends an acknowledgment to switch 2 that contains its incoming VCI in the table, chosen in the previous step. Switch 2 uses this as the outgoing VCI in the table.

c. Switch 2 sends an acknowledgment to switch 1 that contains its incoming VCI in the table, chosen in the previous step. Switch 1 uses this as the outgoing VCI in the table.

d. Finally switch 1 sends an acknowledgment to source A that contains its incoming VCI in the table, chosen in the previous step.

e. The source uses this as the outgoing VCI for the data frames to be sent to destination B.

Teardown Phase

In this phase, source A, after sending all frames to B, sends a special frame called a *teardown request.* Destination B responds with a teardown confirmation frame.

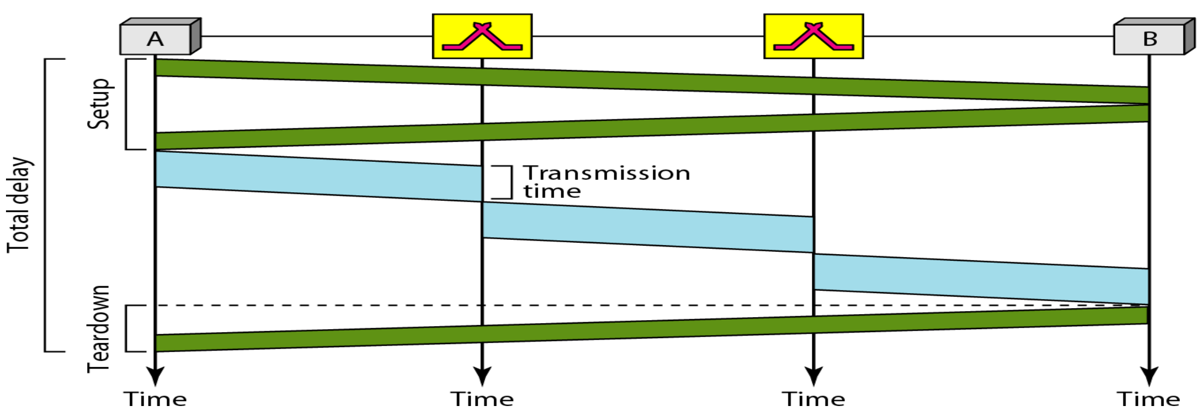
All switches delete the corresponding entry from their tables.

Efficiency

- Resource reservation in a virtual-circuit network can be made during the setup or can be on demand during the data transfer phase.
- In the first case, the delay for each packet is the same; in the second case, each packet may encounter different delays.
- There is one big advantage in a virtual-circuit network even if resource allocation is on demand. The source can check the availability of the resources, without actually reserving it.

Delay in Virtual-Circuit Networks

- In a virtual-circuit network, there is a one-time delay for setup and a one-time delay for teardown.
- If resources are allocated during the setup phase, there is no wait time for individual packets. Below Figure shows the delay for a packet traveling through two switches in a virtual-circuit network.



The packet is traveling through two switches (routers). There are three transmission times (3T), three propagation times (3τ), data transfer depicted by the sloping lines, a setup delay (which includes transmission and propagation in two directions), and a teardown delay (which includes transmission and propagation in one direction). We ignore the processing time in each switch. The total delay time is

$$\textbf{Total delay} = 3T + 3\tau + \textbf{setup delay} + \textbf{teardown delay}$$