5.4.3 Analog Versus Digital Transmission: Which Is Better?

Although analog voice and video can be converted to digital, and digital data can be converted to analog, each format has its own advantages.

Analog : Advantages

Analog transmission offers advantages in the transmission of analog information. Additionally, it is more bandwidth-conservative and is widely available.

Analog Data : Analog has an advantage with respect to the transmission of information that is analog in its native form, such as voice, image and video. The process of transmission of such information is relatively straightforward in an analog format, whereas conversion to a digital bit stream requires conversion equipment. Such equipment adds cost, contributes additional points of failure, and can negatively affect the quality of the signal through the conversion process, itself.

Bandwidth : A raw information stream consumes less bandwidth in analog form than in digital form. This is particularly evident in CATV transmission, where 50 or more analog channels routinely are provided over a single coaxial cable system. Without the application of compression techniques on the same cable system, only a few digital channels could be supported.

Digital : Advantages

Digital transmission offers advantage in the transmission of digital information. Additionally, such data can be compressed effectively and relatively easily. Security of the data can be more readily ensured and the error performances of digital networks are much improved over their analog counterparts.

Finally, the cost-effectiveness of such networks is improved by virtue of the greater bandwidth they provide, especially since they can be more easily upgraded and more effectively managed.

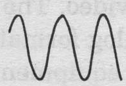
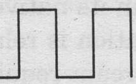
Digital Data : Just as it is better to transmit analog information in an analog format, it is better to transmit digital information in a digital format. Digital transmission certainly has the advantage where binary computer data is being transmitted. The equipment required to convert digital data to an analog format and send the digital bit streams over an analog network can be expensive, susceptible to failure, and can create errors in the information.

Compression : Digital data can be compressed relatively easily, thereby increasing the efficiency of transmission. As a result, substantial volumes of voice, data, video and image information can be transmitted using relatively little raw bandwidth.

Security : Digital systems offer better security. While analog systems offer some measure of security through the *scrambling* or intertwining/mixing of several frequencies, scrambling is fairly simple to defeat. Digital information, on the other hand, can be *encrypted* to create the appearance of a single, pseudo-random bit stream. Thereby, the true meaning of individual bits, sets of bits or the total bit stream cannot be determined without having the key to unlock the encryption algorithm employed.

Quality : Digital transmission offers improved error performance (quality) as compared to analog. This is due to the devices that boost the signal at periodic intervals in the transmission system in order to overcome the effects of *attenuation*. Additionally, digital networks deal more effectively with noise, which always is present in transmission networks.

Summary of comparison of Analog and Digital Transmissions

	Analog	Digital
Data	Continuous (e.g., voice)	Discrete (e.g., text)
Signal	Continuous electromagnetic waves  Used mainly for transmitting data across a network.	Sequence of voltage pulses  Used mainly internally within computers.
Trans-mission	Transmission of analog signals without regards to their content (the data may be analog or binary). The signals become weaker (attenuated) with the distance. Amplifiers may be used to strengthen the signals, but as side effect they also boost the noise. This might not be a problem for analog data, such as voice, but is a problem for digital data.	Transmission that is concerned with the content of the signal. Repeaters are used to overcome attenuation. A repeater recovers the digital pattern from the signal it gets, and resubmits a new signal.

5.4.4 Need of Amplifiers and Repeaters

To understand the need of amplifiers and repeaters first we will discuss two important terms attenuation and noise.

Attenuation

Electromagnetic signals tend to weaken or *attenuate*, over a distance; this is particularly true of electrical signals carried over twisted pair copper wire, due to the level of resistance in the wire. It is also particularly true of microwave radio and other terrestrial radio systems, due to matter in the air. Attenuation is sensitive to carrier frequency, with higher frequency signals attenuating more than lower frequency signals.

Noise

Signals also tend to pick up noise as they transverse the network. Again, this is particularly true of twisted pair, copper wire systems. Such wires tend to act as antennae and, therefore, absorb noise from outside sources of Electromagnetic Interference (EMI). Thus, the quality of the signal degenerates as it is distorted by the noise.

Amplifiers Versus Repeaters

As we have noted, electromagnetic energy attenuates over a distance through either a wire or the air. Therefore, it is necessary to place some sort of device at regular intervals in a network to overcome this phenomenon. These boosting units read a weak incoming signal and create a stronger outgoing signal until the signal reaches another boosting unit, and so on. Analog networks make use of devices known as *amplifiers*, while *repeaters* are employed in digital networks.

Amplifiers (Analog)

The boosting devices in an analog network are known as *amplifiers*. Amplifiers boost, or amplify the weak incoming signal, much like an amplifier in a radio or TV. As it traverses the network the signal accumulates noise. Through every step of the transmission and through each amplifier, the noise is amplified along with the signal, creating the potential for significant accumulated noise at the receiving end of the transmission. The resulting *signal-to-noise* ratio can be unacceptable. Amplifiers typically are spaced every 18,000 feet or so in an analog network.

The impact of amplification on voice communications generally is tolerable, as humans are relatively intelligent receivers who can filter out the noise or, at least adjust to it. In the event of a truly garbled transmission, the human-to-human error detection and correction process simply involves a request for re-transmission. Should the quality of the connection be totally unacceptable, the connection can be terminated and re-established. Computer systems, however, are not so forgiving, and garbled data is of decidedly negative value.

Repeaters (Digital)

In a digital system, periodic amplifiers are replaced by *regenerative repeaters*, which regenerate the signal, rather than simply amplifying it. The repeater guesses the binary value (1 or 0) of the weak incoming signal based on its relative voltage level and regenerates a strong signal of the same value, without the noise. This process immensely enhances the signal quality. Repeaters are spaced at approximately the same intervals as amplifiers, although spacing is sensitive to the carrier frequency, which affects both transmission speed, or bandwidth provided, and the level of attenuation experienced.

The performance advantage of digital networks can be illustrated by comparing the error rate of amplifiers and regenerative repeaters. For example, a twisted-pair, analog network can be expected to yield an error rate on the order of 10^{-5}. In other words, digital data sent across an analog network will suffer 1 errored bit for every 100,000 bits transmitted. The very same twisted-pair network, if digitized and equipped with repeaters, will yield an expected error rate of 10^{-7} or 1 errored bit in every 10,000,000. This is an improvement of two orders of magnitude. Digital fiber optic systems, currently considered to be the ultimate, yield error rates in the range of 10^{-11} to 10^{-14} or an error rate as low as 1 bit for every 100,000,000,000,000 transmitted-virtually perfect !

5.4.5 The Conversion Process: Digital to Analog (D To A) and Analog to Digital (A To D)

Digital to Analog : Modems

As local loops generally are analog, computer communications across such circuits is not possible without the assistance of a device to accomplish the digital-to-analog conversion. Of course, one might gain access to a more expensive digital circuit, by so specifying, if it is available. The device that accomplishes this is known as modem. (Discussed later)

Analog to Digital : Codecs

The reverse conversion process is necessary to send analog information across a digital circuit. This is often the case in the carrier networks, where huge volumes of analog voice are digitized and sent across high capacity, digital circuits. This requirement also exists where high capacity digital circuits connect premise-based, PBX voice systems to central office exchanges or to other PBXs, assuming that the PBXs or COs have not already performed the conversion. As video also is analog in its native form, a similar process must be employed to send video across a digital circuit.

The device that accomplishes the A-to-D conversion is known as a *codec*. Codecs code an analog input into a digital (data) format on the transmit side of the connection, reversing the process or decoding the information, on the receive side, in order to reconstitute the analog signal.

Encoding is the process of converting an analog information stream (e.g., voice or video) into a digital data stream. The voice or video signal is sampled at frequent intervals with each sample of amplitude then being expressed in terms of a binary (computer) value, which is usually a 4-bit or 8-bit byte. The reverse process of *decoding* takes place on the receiving end, resulting in recomposition of the information in its original form or at least a reasonable approximation thereof.

5.5 Modulation

The point to modulation is to take a **message-bearing signal** and superimpose it upon a **carrier signal** for transmission. For ease of transmission carrier signals are generally high frequency for several reasons:

1. For easy (low loss, low dispersion) propagation as electromagnetic waves.
2. So that they may be simultaneously transmitted without interference from other signals.
3. So as to enable the construction of small antennas (a fraction, usually a quarter of the wavelength).
4. So as to be able to multiplex that is to combine multiple signals for transmission at the same time.

In communications, modulation is a process in which some characteristic of a wave (the carrier wave) is made to vary in accordance with an information-bearing signal wave (the modulating wave). Modulation is described as being performed at the physical layer of the OSI Model. The carrier signal can be an electrical current, a radio or microwave frequency or light. The original, unmodulated wave may be of any kind, such as sound or, most often, electromagnetic radiation, including optical waves. In modulation, it is processed in such a way that its amplitude, frequency, or some other property varies.

The process of modulation of the data also requires *demodulation* at the receiving end. Demodulation is the process by which the original signal is recovered from the wave produced by modulation.

Carrier Signal

In information technology, a carrier (or carrier signal) is a transmitted electromagnetic pulse or wave at a steady base frequency of alternation on which information can be imposed by increasing signal strength, varying the base frequency, varying the wave phase, or other means. This variation is called modulation. With the advent of laser transmission over optical fiber media, a carrier can also be a laser-generated light beam on which information is imposed.

Carrier detect (see modem lights) is a control signal between a modem and a computer that indicates that the modem detects a "live" carrier that can be used for sending and receiving information.

Signal modulation can be done in either of two main ways:

1. Analog modulation
2. Digital modulation

In telecommunications, an analog signal is one in which a base carrier's alternating current frequency is modified in some way, such as by amplifying the strength of the signal or varying the frequency, in order to add information to the signal. Broadcast and telephone transmission have conventionally used analog technology.

Types of analog modulation of a carrier include :

- **Amplitude modulation (AM) :** In which the voltage applied to the carrier is varied over time.
- **Frequency modulation (FM) :** In which the frequency of the carrier waveform is varied in small but meaningful amounts.
- **Phase modulation (PM) :** In which the natural flow of the alternating current waveform is delayed temporarily.

In recent years, digital modulation has been getting more common, while analog modulation methods have been used less and less. There are still plenty of analog signals around, however, and they will probably never become totally extinct.

Types of digital modulation include :

- Pulse Modulation Schemes
- Keying Schemes

5.6 Analog Modulation Schemes

The basic idea here is to superimpose the message signal in analog form on a carrier that is a sinusoid of the form

$$A \cos(\omega_c t + \phi)$$

There are three quantities that can be varied in proportion to the modulating signal: the **amplitude**, the **phase** and **frequency**. The first scheme is called Amplitude Modulation and the second two are called **Angle Modulation** schemes.

5.6.1 Amplitude Modulation

It is a method by which radio waves are altered for the transmission of broadcasting signals. AM waves are constant in frequency, but the amplitude of the transmitting wave varies in accordance with the signal being broadcast.

Amplitude

In physics, amplitude is the maximum displacement of an oscillation from the equilibrium position. For a transverse wave motion, as in electromagnetic waves, it is the height of a crest (or the depth of a trough). For a longitudinal wave, such as a sound wave, amplitude is the maximum distance a particle is pushed (due to compression) or pulled (due to rarefaction) from its resting position. A quiet sound has lower amplitude and a loud sound has higher amplitude. For a louder sound, more sound energy enters the ear every second. Amplitude is generally denoted by a.

AM (amplitude modulation) radio broadcasting is a method of transmitting audio-frequency signals that would otherwise travel not very far in space on their own and would have to be transmitted one at a time. The amplitude of the low-frequency audio signal is made to modulate (vary slightly) the amplitude of a continuously transmitted radio **carrier wave** of a higher frequency. In this way the **modulating signal** is imprinted on the carrier wave.

Amplitude modulation (AM) is the modulation method used in the AM radio broadcast band. In this system the intensity, or amplitude, of the carrier wave varies in accordance with the modulating signal. When the carrier is thus modulated, a fraction of the power is converted to sidebands extending above and below the carrier frequency by an amount equal to the highest modulating frequency. If the modulated carrier is rectified and the carrier frequency filtered out, the modulating signal can be recovered. This form of modulation is not a very efficient way to send information; the power required is relatively large because the carrier, which contains no information, is sent along with the information.

In a variant of amplitude modulation, called single sideband modulation (SSB), the modulated signal contains only one sideband and no carrier. The information can be demodulated only if the carrier is used as a reference. This is normally accomplished by generating a wave in the receiver at the carrier frequency. SSB modulation is used for long-distance telephony (such as in the amateur radio bands) and telegraphy over land and submarine cables.

In AM, the carrier itself does not fluctuate in amplitude. Instead, the modulating data appears in the form of signal components at frequencies slightly higher and lower than that of the carrier. These components are called *sidebands*. The lower sideband (LSB) appears at frequencies below the carrier frequency; the upper sideband (USB) appears at frequencies above the carrier frequency. The LSB and USB are essentially "mirror images" of each other in a graph of signal amplitude versus frequency, as shown in the figure 5.5. The sideband power accounts for the variations in the overall amplitude of the signal.

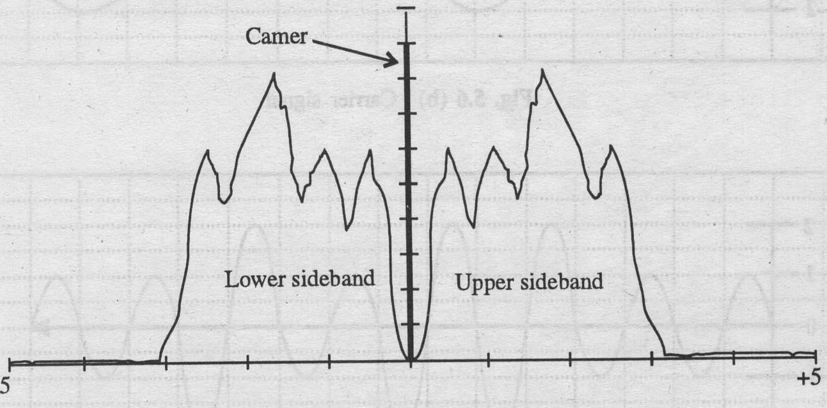


Fig. 5.5 Representation of sidebands

When a carrier is amplitude-modulated with a pure sine wave, up to 1/3 (33 percent) of the overall signal power is contained in the sidebands. The other 2/3 of the signal power is contained in the carrier, which does not contribute to the transfer of data. With a complex modulating signal such as voice, video or music, the sidebands generally contain 20 to 25 percent of the overall signal power; thus the carrier consumes 75 to 80 percent of the power. This makes AM an inefficient mode. If an attempt is made to increase the modulating data input amplitude beyond these limits, the signal will become distorted and will occupy a much greater bandwidth than it should. This is called *overmodulation* and can result in interference to signals on nearby frequencies.

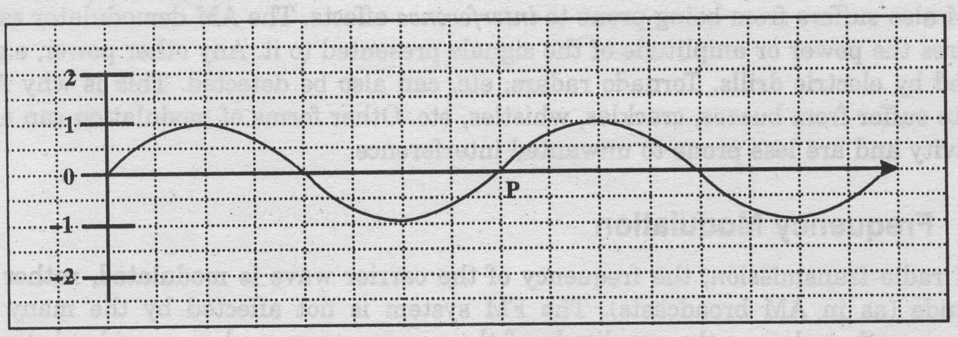


Fig. 5.6 (a) Data carrying signal

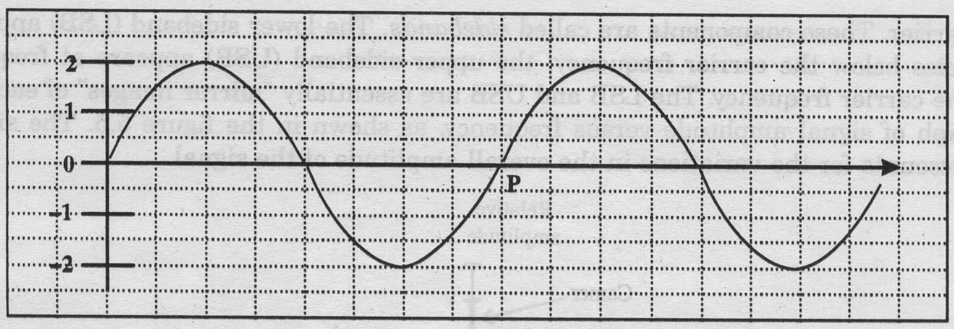


Fig. 5.6 (b) Carrier signal

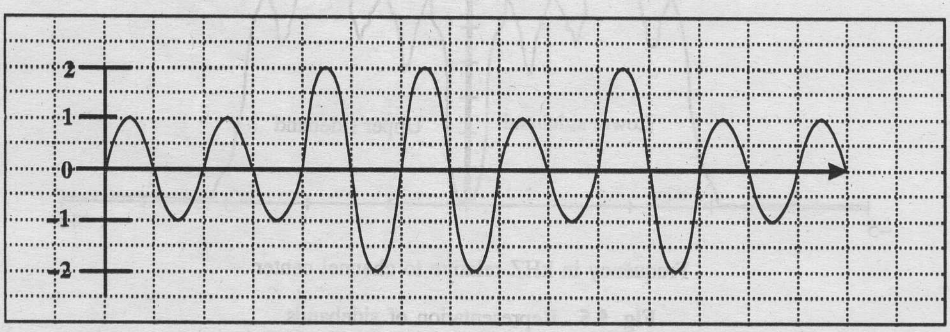


Fig. 5.6 (c) Amplitude modulated wave

The modulated wave is said to consist of three parts :

- The *carrier*, $A_0 \cos(2\pi f_c t)$, which is unaffected by the modulation.
- The *Upper Sideband Components*, $(f_c + f_m)$, which are an 'up converted' copy of the original modulation spectrum.
- The *Lower Sideband Components*, $(f_c - f_m)$, which are a 'mirror image' of the upper sideband spectrum.

AM also suffers from being prone to *interference* effects. The AM demodulator essentially measures the power or amplitude of the signals presented to it. Any other power, e.g., pulses radiated by electric drills, Tornado radars, etc. can also be detected. This is why AM radio tends to suffer from buzzes, crackles, whistles, etc. Other forms of modulation can avoid this sensitivity and are less prone to unwanted interference.

5.6.2 Frequency Modulation

In FM radio transmission, the frequency of the carrier wave is modulated, rather than its amplitude (as in AM broadcasts). The FM system is not affected by the many types of interference that change the amplitude of the carrier wave, and so provides better quality reception than AM broadcasts.

In frequency modulation (FM), the frequency of the carrier wave is varied in such a way that the change in frequency at any instant is proportional to another signal that varies with time. Its principal application is also in radio, where it offers increased noise immunity and decreased distortion over the AM transmissions at the expense of greatly increased bandwidth. The FM band has become the choice of music listeners because of its low-noise, wide-bandwidth qualities; it is also used for the audio portion of a television broadcast.

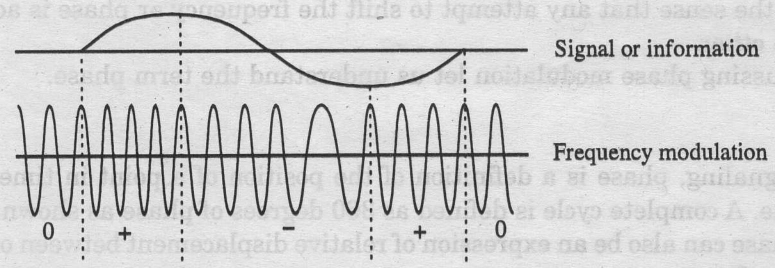


Fig. 5.7 Frequency modulation

This scheme can be used with analog or digital data. In analog FM, the frequency of the AC signal wave, also called the *carrier*, varies in a continuous manner. Thus, there are infinitely many possible carrier frequencies. In **narrowband FM**, commonly used in two-way wireless communications, the instantaneous carrier frequency varies by up to 5 kilohertz (kHz, where 1 kHz = 1000 hertz or alternating cycles per second) above and below the frequency of the carrier with no modulation. In **wideband FM**, used in wireless broadcasting, the instantaneous frequency varies by up to several megahertz (MHz, where 1 MHz = 1,000,000 Hz). When the instantaneous input wave has positive polarity, the carrier frequency shifts in one direction; when the instantaneous input wave has negative polarity, the carrier frequency shifts in the opposite direction. At every instant in time, the extent of carrier-frequency shift (the *deviation*) is directly proportional to the extent to which the signal amplitude is positive or negative.

In digital FM, the carrier frequency shifts abruptly, rather than varying continuously. The number of possible carrier frequency states is usually a power of 2. If there are only two possible frequency states, the mode is called frequency-shift keying (FSK). In more complex modes, there can be four, eight or more different frequency states. Each specific carrier frequency represents a specific digital input data state.

FM is a so-called angle modulation scheme; it was inspired by phase modulation but has proved to be more useful partly for its ease of generation and decoding.

The main advantages of FM over AM are:

1. Improved signal to noise ratio (about 25dB) with respect to man made interference.
2. Smaller geographical interference between neighboring stations.
3. Less radiated power.
4. Well-defined service areas for given transmitter power.

Disadvantages of FM:

1. Much more Bandwidth (as much as 20 times as much).
2. More complicated receiver and transmitter.

5.6.3 Phase Modulation

Phase modulation, like frequency modulation, is a form of angle modulation (so called because the angle of the sine wave carrier is changed by the modulating wave). The two methods are very similar in the sense that any attempt to shift the frequency or phase is accomplished by a change in the other.

Before discussing phase modulation let us understand the term phase.

Phase

In electronic signaling, phase is a definition of the position of a point in time (instant) on a **waveform** cycle. A complete cycle is defined as 360 degrees of phase as shown in upper wave of figure 5.8. Phase can also be an expression of relative displacement between or among waves having the same **frequency.**

Phase difference, also called *phase angle*, in degrees is conventionally defined as a number greater than −180, and less than or equal to +180. *Leading phase* refers to a wave that occurs "ahead" of another wave of the same frequency. *Leading phase* refers to a wave that occurs "behind" another wave of the same frequency. When two signals differ in phase by -90 or +90 degrees, they are said to be in *phase quadrature*. When two waves differ in phase by 180 degrees (−180 is technically the same as +180), the waves are said to be in *phase opposition*. Lower wave in figure 5.8 shows two waves that are in phase quadrature. The wave depicted by the dashed line leads the wave represented by the solid line by 90 degrees.

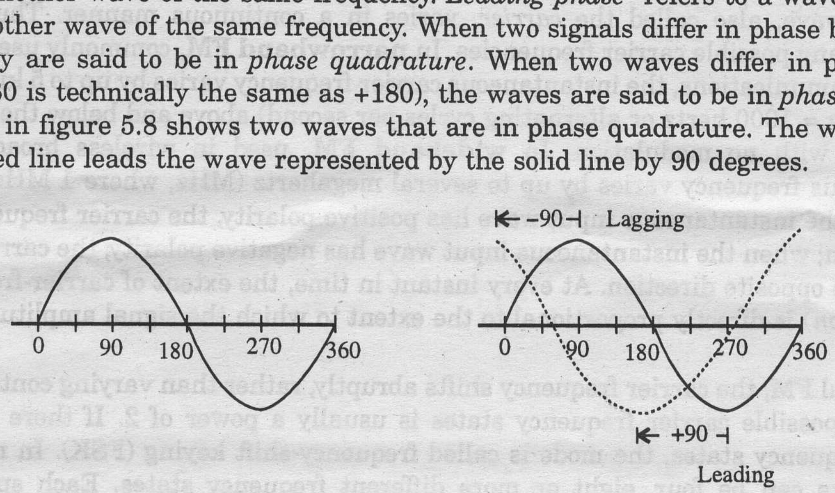


Fig. 5.8 Representation of phase quadrature

Phase is sometimes expressed in radians rather than in degrees. One radian of phase corresponds to approximately 57.3 degrees. Engineers and technicians generally use degrees; physicists more often use radians. The time interval for one degree of phase is inversely proportional to the frequency. If the frequency of a signal (in **hertz**) is given by f, then the time t_{deg} (in seconds) corresponding to one degree of phase is:

$$t_{deg} = \frac{1}{(360f)}$$

The time t_{rad} (in seconds) corresponding to one radian of phase is approximately:

$$t_{rad} = \frac{1}{(6.28f)}$$

Now let us discuss what is phase modulation? Phase modulation (PM) is a method of impressing data onto an alternating-current (AC) waveform by varying the instantaneous phase of the wave. This scheme can be used with analog or digital data.

In analog PM, the phase of the AC signal wave, also called the *carrier*, varies in a continuous manner. Thus, there are infinitely many possible carrier phase states. When the instantaneous data input waveform has positive polarity, the carrier phase shifts in one direction; when the instantaneous data input waveform has negative polarity, the carrier phase shifts in the opposite direction. At every instant in time, the extent of carrier-phase shift (the *phase angle*) is directly proportional to the extent to which the signal amplitude is positive or negative.

In digital PM, the carrier phase shifts abruptly, rather than continuously back and forth. The number of possible carrier phase states is usually a power of 2. If there are only two possible phase states, the mode is called biphase modulation. In more complex modes, there can be four, eight or more different phase states. Each phase angle (that is, each shift from one phase state to another) represents a specific digital input data state.

Phase modulation is similar in practice to frequency modulation (FM). When the instantaneous phase of a carrier is varied, the instantaneous frequency changes as well. The converse also holds: When the instantaneous frequency is varied, the instantaneous phase changes. But PM and FM are not exactly equivalent, especially in analog applications. When an FM receiver is used to demodulate a PM signal, or when an FM signal is intercepted by a receiver designed for PM, the audio is distorted. This is because the relationship between phase and frequency variations is not linear; that is, phase and frequency do not vary in direct proportion.

5.7 Digital Modulation Schemes

In these schemes one is interested in preparing a digital signal of 1's and 0's for transmission. This signal arises from several sources, for example:

1. Encoding type written characters (ASCII).
2. Sampling and quantizing speech.
3. Sampling and quantizing a video signal.

The simplest digital modulation scheme is called On Off Keying (OOK): here a signal is turned on in response to a 1 and off in response to a 0.

The main types of digital modulation schemes include:

- Pulse Modulation Schemes
- Keying Schemes

5.7.1 Pulse Modulation Schemes

Pulse modulation involves modulating a carrier that is a train of regularly recurrent pulses. The modulation might vary the amplitude (PAM or pulse amplitude modulation), the duration (PDM or pulse duration modulation) or the presence of the pulses (PCM or pulse code modulation). PCM can be used to send digital data; audio signals on a compact disc use pulse code modulation. Developed in 1939 by the English inventor Alec H. Reeves, pulse code modulation is the most important form of pulse modulation because it can be used to transmit information over long distances with hardly any interference or distortion; for this reason it has become increasingly important in the transmission of data in the space program and between computers. Although PCM transmits digital instead of analog signals, the modulating wave is continuous.

Pulse modulation techniques include:

- Pulse-code modulation (PCM).
- Pulse-width modulation (PWM).
- Pulse-amplitude modulation (PAM).
- Pulse-position modulation (PPM).

Pulse-code Modulation (PCM)

Pulse-code modulation (PCM) is a modulation technique. It is a digital representation of an analog signal where the magnitude of the signal is sampled regularly at uniform intervals of duration. Every sample is quantized to a series of symbols in a digital code, which is usually a binary code. PCM is used in digital telephone systems and for digital audio recording on compact discs. Morse code is a very simple example of pulse-code modulation.

In telephony, several PCM streams may be multiplexed into a larger aggregate data stream. The signals in PCM are binary; that is, there are only two possible states, represented by logic 1 (high) and logic 0 (low). This is true no matter how complex the analog waveform happens to be.

To obtain PCM from an analog waveform at the source (transmitter end) of a communications circuit, the analog signal amplitude is sampled (measured) at regular time intervals. The sampling rate, or number of samples per second, is several times the maximum frequency of the analog waveform in cycles per second or hertz. The instantaneous amplitude of the analog signal at each sampling is rounded off to the nearest of several specific, predetermined levels. This process is called quantization. The number of levels is always a power of 2, for example, 8, 16, 32 or 64. These numbers can be represented by three, four, five or six binary digits (bits) respectively. The output of a pulse code modulator is thus a series of binary numbers, each represented by some power of at the destination (receiver end) of the communications circuit, a pulse code demodulator converts the binary numbers back into pulses having the same quantum levels as those in the modulator. These pulses are further processed to restore the original analog waveform.

Pulse-code modulation can be either return-to-zero (RZ) or non-return-to-zero (NRZ). For a NRZ system to be synchronized using in-band information, there must not be long sequences of identical symbols, such as ones or zeroes. For binary PCM systems, the density of 1-symbols is called 'ones-density'.

Sampling Theorem

A signal can be reconstructed from a sample taken at

- Regular time intervals.
- A rate higher than twice the highest significant frequency.

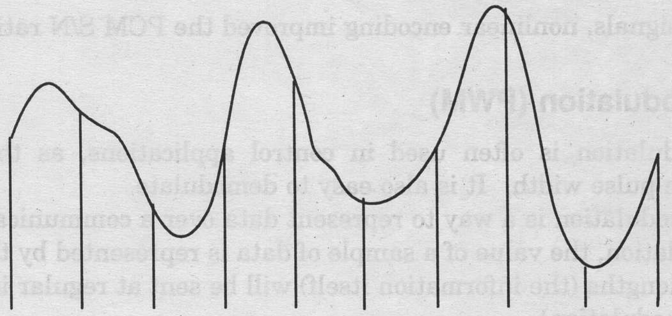


Fig. 5.9 A sample wave

The conversion has three stages :

Sampling of amplitude signals (PAM pulses).
Digitizing of the amplitude signals (PCM pulses).
Encoding of the stream of bits.

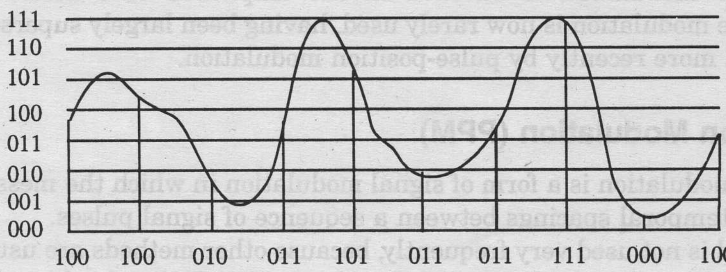


Fig. 5.10 Encoding of the stream of bits

- The quantization of the PAM pulse introduces **quantization error**, not allowing the recovery of the original pulse.
- A reduction in the signal distortion can be obtained by adding bits. The quantizing noise satisfies S/N = 6n + 1.8 dB.

- **A reduction in the signal distortion can be obtained by nonuniform quantization which is finer at low amplitudes.**

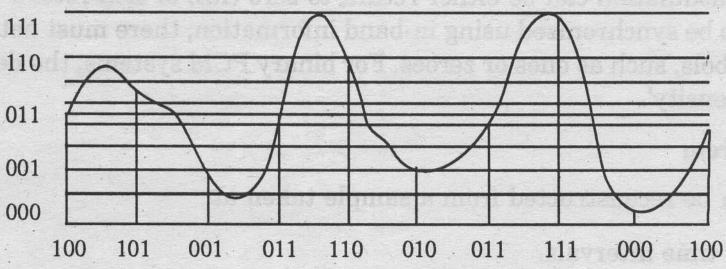


Fig. 5.11 Nonuniform quantization

- For voice signals, nonlinear encoding improved the PCM S/N ratio by 24-30 dB.

Pulse-width Modulation (PWM)

This form of modulation is often used in control applications, as the average value is proportional to the pulse width. It is also easy to demodulate.

Pulse-width modulation is a way to represent data over a communications channel. With pulse-width modulation, the value of a sample of data is represented by the length of a pulse. Pulses of various lengths (the information itself) will be sent at regular intervals (the carrier frequency of the modulation).

PWM is also used to vary the total amount of power delivered to a load without resistive waste. An RC filter can be used to smooth the pulse train into a steady analog voltage. This method is commonly used in DC motor speed control.

Pulse-amplitude Modulation (PAM)

Pulse-amplitude modulation is encoded in the amplitude of a series of signal pulses. Pulse-amplitude modulation is now rarely used, having been largely superseded by pulse-code modulation and more recently by pulse-position modulation.

Pulse-position Modulation (PPM)

Pulse-position modulation is a form of signal modulation in which the message information is encoded in the temporal spacings between a sequence of signal pulses.

This method is not used very frequently, because other methods are usually more suitable for typical data transmission, however one common use of PPM is for the radio control of model aircraft. Here the position of each pulse represents the angular position of an analogue control on the transmitter. The number of pulses per frame gives the number of controllable channels available. The advantage of using PPM for this type of application is that the electronics required to decode the signal is extremely simple, which leads to small, low weight receiver/decoder units, essential in model aircraft where every weight saving is necessary. Servos made for model radio control include some of the electronics required to convert the pulse to the motor position—the receiver is merely required to demultiplex the separate channels and feed the pulses to each servo.

5.7.2 Keying Schemes

In these schemes either the frequency or phase of a carrier signal is keyed in response to patterns of 1's and 0's. Of course one could also key the amplitude, but this is what we already called PAM. Frequency Shift Keying (FSK) stands for the process of keying between two different frequencies.

Frequency-shift Keying (FSK)

Frequency-shift keying (FSK) is a method of transmitting digital signals. The two binary states, logic 0 (low) and 1 (high), are each represented by an analog waveform. Logic 0 is represented by a wave at a specific frequency, and logic 1 is represented by a wave at a different frequency. A modem converts the binary data from a computer to FSK for transmission over telephone lines, cables, optical fiber, or wireless media. The modem also converts incoming FSK signals to digital low and high states, which the computer can "understand".

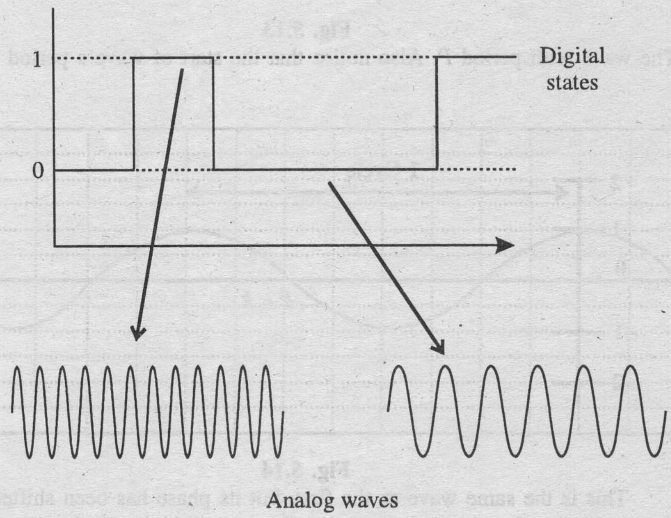


Fig. 5.12 Representation of analog waves corresponding to digital states

The FSK mode was introduced for use with mechanical teleprinters in the mid-1900s. The standard speed of those machines was 45 baud, equivalent to about 45 bits per second. When personal computers became common and networks came into being, this signaling speed was tedious. Transmission of large text documents and programs took hours; image transfer was unknown. During the 1970s, engineers began to develop modems that ran at faster speeds, and the quest for even-greater bandwidth has continued ever since. Today, a standard telephone modem operates at thousands of bits per second. Cable and wireless modems work at more than 1,000,000 bps (one megabit per second or 1 Mbps), and optical fiber modems function at many Mbps. But the basic principle of FSK has not changed in more than half a century.

Phase Shift Keying (PSK)

Phase-shift keying (PSK) is a method of transmitting and receiving digital signals in which the phase of a transmitted signal is varied to convey information.

So it is a technique that shifts the period of a wave.

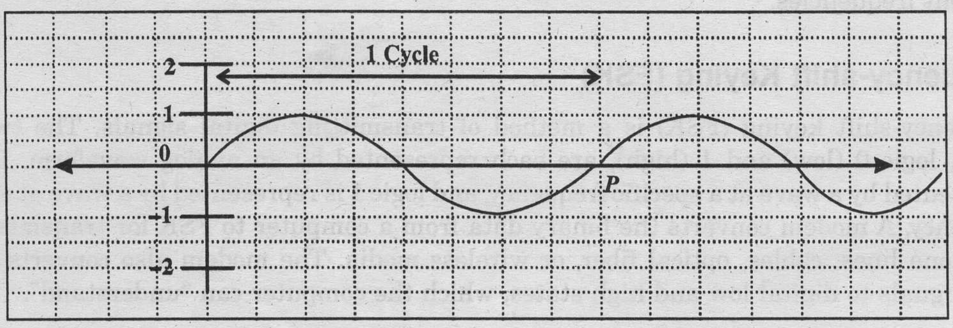


Fig. 5.13

The wave with period *P*. Also notice that the start of wave's period is at 0

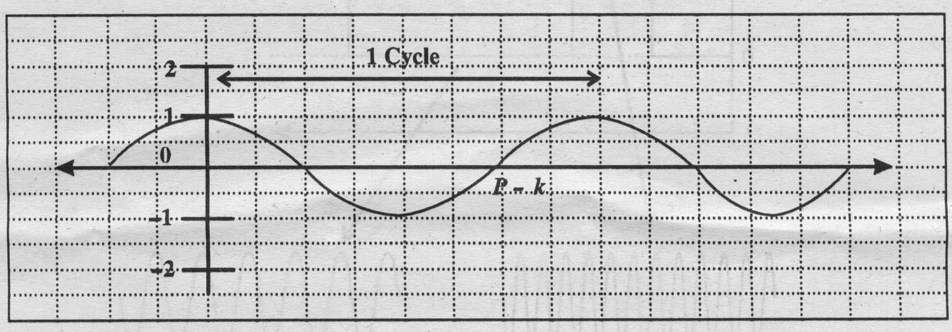


Fig. 5.14

This is the same wave as the first, but its phase has been shifted.
Notice that the period starts at the wave's highest point (1).

So what's the point? It just so happens that we have shifted this wave by **one quarter** of the wave's full period. We can shift it another quarter, if we wanted to, so the original wave would be shifted by **half** its period. And we could do it one more time, so that it would be shifted three **quarters** of its original period.

This means we have 4 separate waves. So why not let each wave stand for some binary value? Since there are 4, we can let each wave signify 2 bits (00, 01, 10, 11):

Bit value	Amount of shift
00	None
01	1/4
10	1/2
11	3/4

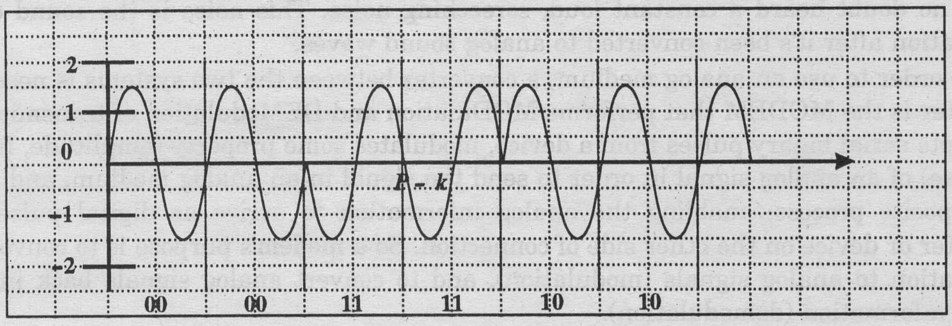


Fig. 5.15 Phase shift keying

This technique of letting each shift of a wave represent some bit value is **phase shift keying**.

It is a technique for switching phases in response to the signal. As before we have quadrature versions of PSK which are referred to as either QPSK or 4-PSK, 8-PSK, etc. The 1200 Baud full duplex, 2400 Baud full duplex and 4800-baud full duplex modems use 4-PSK or 8-PSK.

In channels where there is rapid fading such as mobile radio (wireless) channels it is useful to use PSK modulation, since the decoder does not need to keep track of the intensity of the received signal. For greater noise immunity in these applications, one uses a so-called differential modulation scheme, where one encodes the changes in the group of bits being encoded. Thus, for example digital cellular radio in North America uses a so-called differential PSK-4 scheme.

There are several schemes that can be used to accomplish PSK. The simplest method uses only two signal phases: 0 degrees and 180 degrees. The digital signal is broken up time wise into individual bits (binary digits). The state of each bit is determined according to the state of the preceding bit. If the phase of the wave does not change, then the signal state stays the same (low or high). If the phase of the wave changes by 180 degrees—that is, if the phase reverses—then the signal state changes (from low to high, or from high to low). Because there are two possible wave phases, this form of PSK is sometimes called biphase modulation.

More complex forms of PSK employ four or eight wave phases. This allows binary data to be transmitted at a faster rate per phase change than is possible with biphase modulation. In four-phase modulation, the possible phase angles are 0, +90, -90, and 180 degrees; each phase shift can represent two signal elements. In eight-phase modulation, the possible phase angles are 0, +45, –45, +90, –90, +135, –135, and 180 degrees; each phase shift can represent four signal elements.

5.8 Introduction To Modems

The need to communicate between distant computers led to the use of the existing phone network for data transmission. Most phone lines were designed to transmit analog information-voices, while the computers and their devices work in digital form-pulses.

You can't simply transmit bits directly across telephone lines. It first needs to be converted into sound waves. If you've ever picked up your telephone while you were using your modem, you've no doubt heard a constant loud, screeching noise. This noise is the sound of digital information after it's been converted to analog sound waves.

In order to use an analog medium, a converter between the two systems is needed. This converter is the MODEM that performs MODulation and DEModulation of transmitted data. It accepts serial binary pulses from a device, modulates some property (amplitude, frequency, or phase) of an analog signal in order to send the signal in an analog medium, and performs the opposite process, enabling the analog information to arrive as digital pulses at the computer or device on the other side of connection. So a modem's purpose is to convert digital information to analog signals (modulation), and to convert analog signals back into useful digital information (demodulation).

Modems, in the beginning, were used mainly to communicate between data terminals and a host computer. Later, the use of modems was extended to communicate between end computers. This required more speed and the data rates increased from 300 bps in early days to 28.8 Kbps today. Today, transmission involves data compression techniques that increase the rates, error detection and error correction for more reliability.

Today's modems are used for different functions. They act as textual and voice mail systems, facsimiles, and are connected or integrated into cellular phones and in notebook computers enabling sending data from anywhere. The future might lead to new applications. Modem speeds are not expected to be increased much over today's 28.8 kbps. Further dramatic speed increases will require digital phone technology such as ISDN and fiber optic lines.

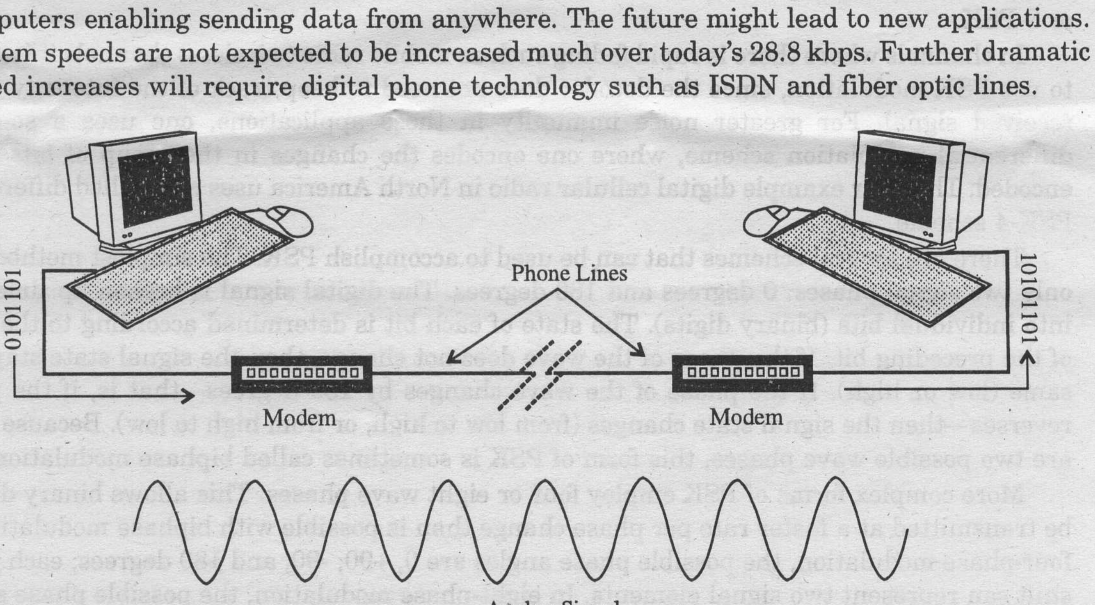


Fig. 5.16 Transmission of data using modems

There are many different types of modems in use operating at different speeds and conforming to different standards. To function, the modems on both sides of the connection must agree to use the same standard. When modems connect, they must negotiate:

- Quality of the line.
- Speed.
- Options such as encoding and compression.
- Disable echo suppression.

This negotiation takes about 10 seconds for a V.34 modem and is usually heard as a series of squeals from your modem. Some systems may use a modem to connect to distant computers over a dedicated or leased line. This line is always connected between the two points. No dialing or hang-up is required. Frequently these modems can operate without the initial parameter negotiation since they can be configured during installation to be compatible.

Aside from the transmission protocols that they support, the following characteristics distinguish one modem from another.

5.8.1 Characteristics of Modems

1. Speed : The speed at which the modem can send data is measured in bps (bits per second). Typical modem speeds are: 300, 600, 1200, 2400, 4800, 9600, 14.4K, 19.2K, 28.8K bps.

bps : It measures how fast the modem can transmit and receive data. At slow rates, modems are measured in terms of baud rates. The carrier signal is characterized by the number of signal intervals, or pulses, that are transmitted per second. Each pulse is called a **baud**.

The slowest rate is 300 baud (about 25 cps). At higher speeds, modems are measured in terms of bits per second (bps).

Bps stands for **b**its **p**er **s**econd. Bps is a measure of how many bits can be transmitted during one pulse (one baud). So,

bps = baud × number of bits per baud

The fastest modems run at 57,600 bps, although they can achieve even higher data transfer rates by compressing the data. Obviously, the faster the transmission rate, the faster you can send and receive data. Note, however, that you cannot receive data any faster than it is being sent. If, for example, the device sending data to your computer is sending it at 2,400 bps, you must receive it at 2,400 bps. It does not always pay, therefore, to have a very fast modem. In addition, some telephone lines are unable to transmit data reliably at very high rates.

A modem operating at 9600 bps is *still* only transmitting at 1200 baud. But it is "packing" 8 bits into each baud:

9600 bps = 1200 baud × 8 bits per baud

2. Voice/data : Many modems support a switch to change between voice and data modes. In data mode, the modem acts like a regular modem. In voice mode, the modem acts like a regular telephone. Modems that support a voice/data switch have a built-in loudspeaker and microphone for voice communication.

3. Auto-answer : An auto-answer modem enables your computer to receive calls in your absence. This is only necessary if you are offering some type of computer service that people can call in to use.

4. Data compression : Some modems perform data compression, which enables them to send data at faster rates. However, the modem at the receiving end must be able to decompress the data using the same compression technique.

5. Flash memory : Some modems come with *flash memory* rather than conventional ROM, which means that the communications protocols can be easily updated if necessary.

6. Fax capability : Most modern modems are fax modems, which means that they can send and receive faxes.

7. Self Testing : Newer modems have self-testing features. They can test the digital connection to the terminal/computer and the analog connection to a remote modem. They can also check the modem's internal electronics.

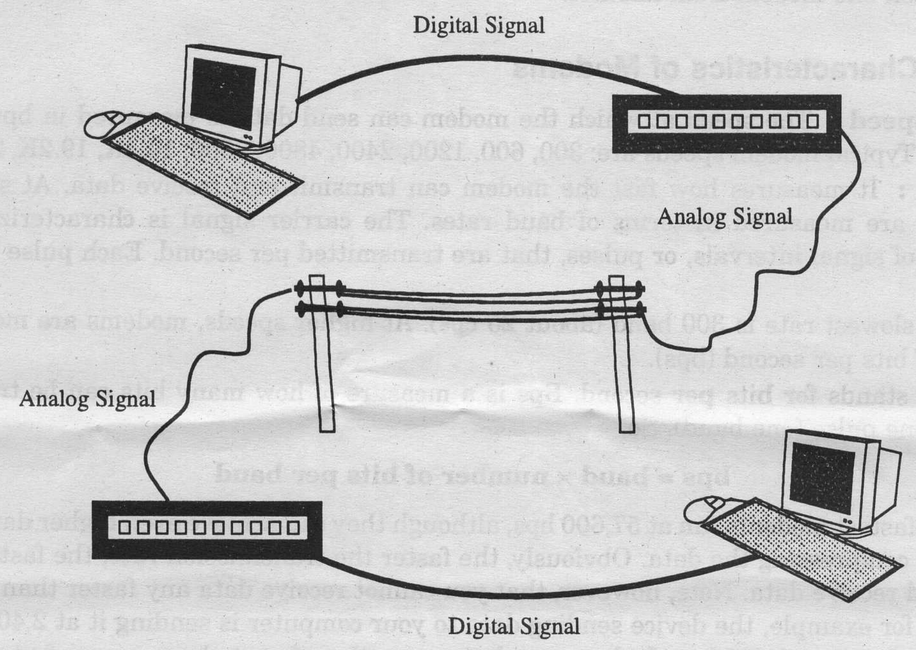


Fig. 5.17 Transmission of signals

Buying and using a modem used to be relatively easy. Not too long ago, almost all modems transferred data at a rate of 2400 Bps (bits per second). Today, modems not only run faster, they are also loaded with features like error control and data compression. In addition, modems also act like traffic cops, monitoring and regulating the flow of information. That way one computer doesn't send information until the receiving computer is ready for it. Each of these features—modulation, error control, and data compression—requires a separate kind of protocol. That's what some of the terms you see like **V.32**, **V.32** bis, **V.42** bis and **MNP5** refer to.

5.8.2 Standards and Protocols

Communication between two devices might work only when the interface is defined and agreed. In order to enable modems of various types and different manufacture to communicate, interface standards were developed by some standard organizations. For modems, the standards define techniques used for modulation, for error correction, for data compression, and other attributes. There are some standard organizations for the development of interface standards: The ITU—International Telecommunications Union an agency of the United Nations (Geneva, Switzerland), ISO—International Standards Organization, and CCITT—International Telegraph and Telephone Consultative Committee a group of ITU.

While the modem interfaces are standardized, a number of different protocols for formatting data to be transmitted over telephone lines exist. Some, like CCITT V.34, are official standards, while others have been developed by private companies. Most modems have built-in support for the more common protocols—at slow data transmission speeds at least, most modems can communicate with each other. At high transmission speeds, however, the protocols are less standardized.

5.8.3 Types of Modems

There are two types of modems : **internal modem** and **external modem.**

Internal modem

A modem that resides on an expansion board that plugs into a computer. An internal modem is a plug-in circuit board that sits inside the computer. It incorporates the serial port on-board. They are less expensive than external modems because they do not require a case, power supply and serial cable. They appear to the communication programs as if they were an external modem for all practical purposes.

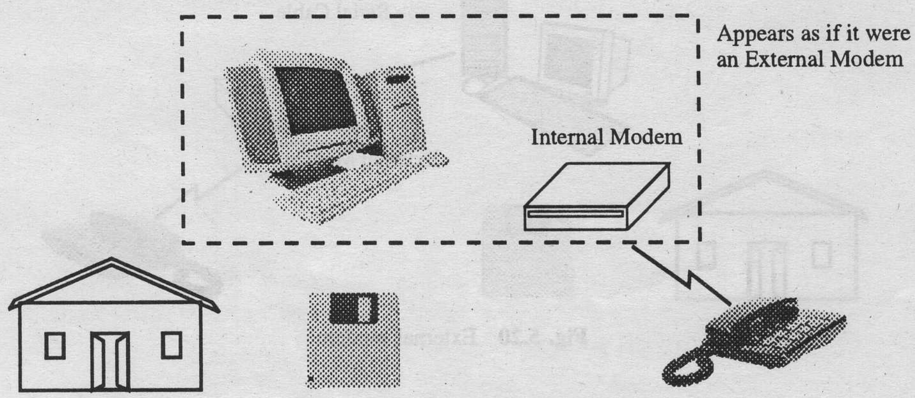


Fig. 5.18 Internal modem

External modem

External modems are separate boxes, and you have to have a cable to connect to the serial port that is on the back of your computer (most, but not all computers have a serial port available).

It is a modem that resides in a self-contained box outside the computer system. Contrast with an internal modem, which resides on a printed circuit board inserted into the computer.

External modems tend to be slightly more expensive than internal modems. Many experts consider them superior because they contain lights that indicate how the modem is functioning. In addition, they can easily be moved from one computer to another. However, they do use up one COM port.

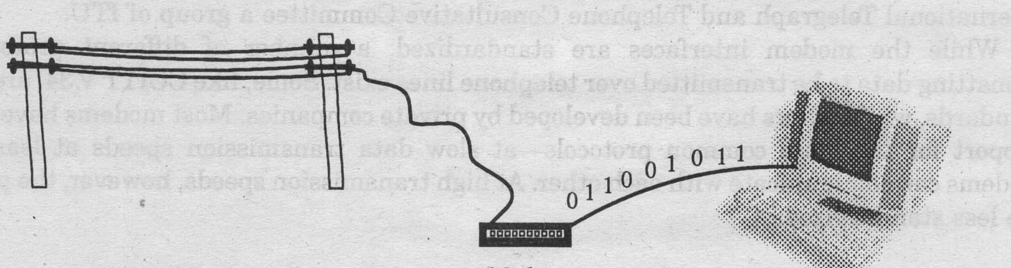


Fig. 5.19 Transmission of data using external modem

Fortunately, there is one standard interface for connecting external modems to computers called *RS*-232. Consequently, any external modem can be attached to any computer that has an RS-232 port, which almost all personal computers have.

External modems sit next to the computer and connect to the serial port using a straight-through serial cable.

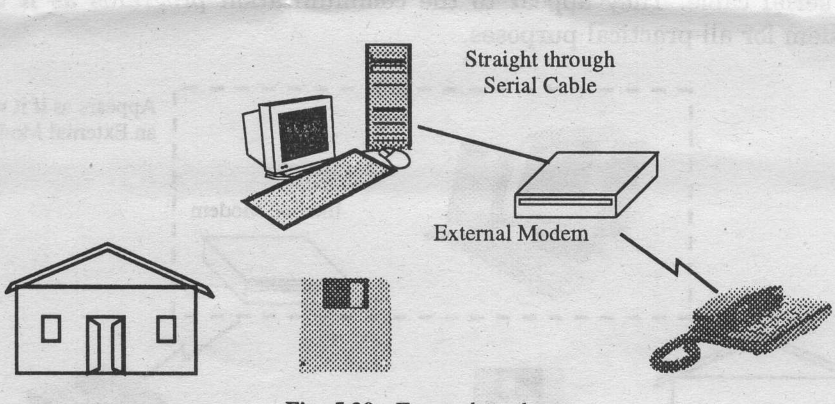


Fig. 5.20 External modem

5.8.4 Classification of Modems

The modems can be classified according to their characteristics:

- Range
  - ☐ Short Haul
  - ☐ Voice Grade (VG)
  - ☐ Wideband
- Line Type
  - ☐ Dial-up
  - ☐ Leased
  - ☐ Private
- Operation Mode
  - ☐ Half Duplex
  - ☐ Full Duplex
  - ☐ Simplex
- Synchronization
  - ☐ Asynchronous
  - ☐ Synchronous
- Modulation
  - ☐ AM
  - ☐ FM/FSK
  - ☐ PM
  - ☐ TCM

Classifying Modems According to Range

Short Haul

Short haul modems are cheap solutions to systems of short ranges (up to 15 km), which use private lines and are not part of a public system. Short haul modems can also be used, even if the end-to-end length of the direct connection is longer than 15 km, when both ends of the line are served by the same central office in the telephone system. These lines are called "local loops". Short haul modems are distance-sensitive, because signal attenuation occurs as the signal travels through the line. The transmission rate must be lowered to ensure consistent and error-free transmission on longer distances.

Short haul modems tend to be cheaper than other modems for two reasons:

(1) No circuitry is included in them to correct for differences between the carrier frequency of the demodulator and the frequency of the modulator.

(2) Generally no circuitry is included to reduce/correct for noise rejection, which is less of a problem over short distances than over long distances.

Voice Grade (VG)

Voice-grade modems are used for unlimited destination, using a moderate to high data rate. These modems are expensive and their maintenance and tuning are sophisticated. Communication channels are leased lines and dial-up.

Voice-band telephone network is used for data transmission. A user-to-user connection may be either dedicated or dialed. The links in the connection are the same in the two cases, and the only difference for the user is that for some impairments (particularly attenuation and delay distortion), a dedicated (*private* or *leased*) line is guaranteed to meet certain specifications, whereas a dialed connection can only be described statistically.

Wideband

Wideband modems are used in large-volume telephone-line multiplexing, dedicated computer-to-computer links. These modems exceed high data rates.

Classifying Modems According To—Line Type

Leased, Private

Leased, private or dedicated lines (usually 4-wire) are for the exclusive use of "leased-line" modems—either pair (in a simple point-to-point connection) or several (on a multidrop network for a polling or a contention system). If the medium is the telephone network, their transmission characteristics are usually guaranteed to meet certain specifications, but if the link includes any radio transmission, the quality of it may be as variable as that of a switched (i.e. nondedicated) line.

Dial up

Dial-up modems can establish point-to-point connections on the PSTN by any combination of manual or automatic dialing or answering. The quality of the circuit is not guaranteed, but all phone companies establish objectives. The links established are almost always 2-wire because 4-wire dialing is tedious and expensive.

*** Two and Four-Wires Lines**

A four-wire (4W) line is a pair of two-wire (2W) lines, one for transmitting and one for receiving, in which the signals in the two directions are to be kept totally separate. Perfect separation can be maintained only if the four-wire configuration is sustained from transmitter to receiver. The lines may be combined in a 4W/2W network (often called a *hybrid* or a *hybrid transformer*) at any point in the signal path. In this case impedance mismatches will cause reflections and interference between the two signals.

Classifying Modems According To—Operation Mode

Half Duplex

Half duplex means that signals can be passed in either direction, but not in both simultaneously. A telephone channel often includes an echo-suppressor, allowing transmission in only one direction, this renders the channel half-duplex. Echo suppressors are slowly being replaced by echo cancellers, which are theoretically full-duplex devices.

When a modem is connected to a two-wire line, its output impedance cannot be matched exactly to the input impedance of the line, and some part of its transmitted signal (usually badly distorted) will always be reflected back. For this reason half-duplex receivers are disabled (received data is clamped) when their local transmitter is operative. Half-duplex modems can work in full-duplex mode.

Full Duplex

Full duplex means that signals can be passed in either direction, simultaneously. Full duplex operation on a two-wire line requires the ability to separate a receive signal from the reflection of the transmitted signal. This is accomplished by either FDM (frequency division multiplexing) in which the signals in the two directions occupy different frequency bands and are separated by filtering, or by Echo Canceling (EC).

The implication of the term full duplex is usually that the modem can transmit and receive simultaneously at full speed. Modems that provide a low-speed reverse channel are sometimes called split-speed or asymmetric modems. Full duplex modems will not work on half-duplex channels.

Simplex

Simplex means that signals can be passed in one direction only. A remote modem for a telemetering system might be simplex and a 2-wire line with a common unidirectional amplifier is simplex.

Classifying Modems According To—Synchronization

Asynchronous Modems

Most of the modems that operate in slow and moderate rates, up to 1800 bps, are asynchronous (using asynchronous data). Asynchronous modems operate in FSK modulation and use two frequencies for transmission and another two for receiving. Asynchronous modems can be connected in different options to the communication media:

- Using 2-wire or 4-wire interface.
- Using switched lines or leased lines.
- Using interface to call unit/automatic answer, when dialing-up.

Asynchronous data

Asynchronous data is not accompanied by any clock, and the transmitting and receiving modems know only the nominal data rate. To prevent slipping of the data relative to the modems' clocks, this data is always grouped in very short blocks (characters) with framing bits (start and stop bits). The most common code used for this is the seven-bit ASCII code with even parity.

Synchronous Modems

Synchronous modems operate in the audio domain, at rates up to 28800 bps in audio lines, used in telephones systems (using synchronous data). The usual modulation methods are the phase modulation and integrated phase and amplitude (at higher rates than 4800 bps).

In synchronous modems, equalizers are used, in order to offset the misfit of the telephone lines. These equalizers are inserted in addition to the equalizers that sometimes already exist in the telephone lines.

These equalizers can be classified into three main groups:

1. *Fixed/statistical equalizer*—these equalizers offset the signal according to the average of the known attenuation in each frequency. Tuning the equalizer is sometimes done in the factory and stays fixed; usually they are used to operate at low rates in a dial up line.
2. *Manually adjusted equalizer*—these equalizers can be tuned to optimal performance to a given line. These equalizers should be re-tuned when the line is replaced and periodically. Specially, it should be tuned frequently when the line is of a low quality and its parameters are changed frequently. Tuning is done using a button inside the modem (or on the external board).
3. *Automatic equalizer*—these equalizers are tuned automatically when the connection is established. Depending on the line quality in a specific moment, in a process of about 15ms to 25ms, after the first tuning, the equalizer samples the line continually and adjusts itself to the changed conditions, so the modem operates at each moment under optimal conditions. The fitness process operates, in some modems, at rates of 2400 times in a second.

Synchronous modems operate in the same manner as asynchronous modems. However, synchronous modems operate at higher rates and since the requirements to transmit at these rates is increasing, most of the innovations are implemented for synchronous modems.

In synchronous modems the channel can be split for several consumers at various speeds. Modems who have this ability are called SSM - Split System Modem. These modems can use a simple split or a split using multipoint connection.

Synchronous data is accompanied by a clock signal. Synchronous data is almost always grouped in blocks, and it is the responsibility of the data source to assemble those blocks with framing codes and any extra bits needed for error detecting and/or correcting according to one of many different protocols (BISYNC, SDLC, HDLC, etc.). The data source and destination expect the modem to be transparent to this type of data, conversely, the modem can ignore the blocking of the data.

Classifying Modems According To—Modulation

Communication channels like telephone lines are usually analog media. Analog media is a bandwidth-limited channel. In the case of telephone lines the usable bandwidth frequencies is in the range of 300 Hz to 3300 Hz.

Data communication means moving digital information from one place to another through communication channels. These digital information signals have the shape of square waves and the meaning of "0" and "1".

If such digital signals were transmitted on analog media the square waves of the digital signals would be distorted by the analog media. The receiver that receives these distorted signals will be unable to interpret accurately the incoming signals. These digital signals must be converted into analog signals so that the communication channels can carry the information from one place to another. The technique that enables this conversion is called modulation.

Modulation is a technique of modifying some basic analog signal in a known way in order to encode information in that basic signal. Any measurable property of an analog signal can be used to transmit information by changing this property in some known manner and then detecting those changes at the receiver end. The signal that is modulated is called the carrier signal because it carries the digital information from one end of the communication channel to the other end. The device that changes the signal at the transmitting end of the communication channel is called the modulator. The device at the receiving end of the channel, which detects the digital information from the modulated signal, is called the demodulator. There are three modulation techniques; (Discussed before) each of them changes one of the properties of the basic analog signal.

5.8.5 A Look At The Future

So what's in the future for today's modems? Not much, unfortunately. The available bandwidth in today's phone lines has reached its limit; the only way modem companies are going to be able to compete is to develop better and faster compression algorithms.

But all hope is not lost! Cable modems may be the "next big thing"!

5.8.6 Cable Modems

The term "Cable Modem" is quite new and refers to a modem that operates over the ordinary cable TV network cables. Basically you just connect the Cable Modem to the TV outlet for your cable TV, and the cable TV operator connects a Cable Modem Termination System (CMTS) in his end (the Head-End).Actually the term "Cable Modem" is a bit misleading, as a Cable Modem works more like a Local Area Network (LAN) interface than as a modem.

Cable modems don't connect to your phone line; they connect to the same coaxial cable that you connect your television to. The idea behind cable modems to allow faster direct access to the Internet via your cable company, which, because of recent legislation, may enter the telecommunications market.

With speeds of up to 36 Mbps, cable modems download data in seconds that might take fifty times longer with a dial-up connection. Best of all, it's always on, so there is no need to connect and no more busy signals! This service is now available in parts of the United States, Europe and Asia.

5.9 Codec

Codec is basically "**coder/decoder**", which describes a device or program capable of performing transformations on a data stream or signal. Codecs can both put the stream or signal into an encoded form (often for transmission, storage or encryption) and retrieve, or decode that form for viewing or manipulation in a format more appropriate for these operations. Codecs are often used in videoconferencing and streaming media solutions.

For instance, many multimedia data streams need to contain both audio and video data, and often some form of metadata that permits synchronization of the audio and video. Each of these three pieces of data may be handled by different programs, processes or hardware; but for the multimedia data stream to be useful in stored or transmitted form, they must be encapsulated together. The raw encoded form of audio and video data is often called essence, to distinguish it from the metadata information that together make up the information content of the stream and any "wrapper" data that is then added to aid access to or improve the robustness of the stream. File formats like ".ogg", ".mpg", ".avi", ".mov", etc. are used to store information encoded by a codec.

● ● ●

6

Switching Principles

Introduction

Telecommunication Networks can be mainly classified into two groups based on the criteria of who has made the decision of which nodes are not going to receive the transmitted information. When the network takes the responsibility of this decision, we have a switching network. When this decision is left to the end-nodes, we have a broadcast network that can be divided in packet radio networks, satellite networks and local area networks.

Communication Units

The two important communication units are discussed below:

Message

The unit of communication from the programmer's perspective. Its size is limited only by the user memory space.

Packet

Fixed-size smallest unit of communication containing *routing information* (e.g., a destination address) and **sequencing** information in its **header**. Its size is of order hundreds or thousands of bytes or words.

6.1 Switching Mechanism

It determines how network resources are allocated for data transmission, i.e., **how** and **when** the input channel is connected to the output channel selected by the routing algorithm. It is the actual mechanism that removes data from input channels and places them on output channels.

Switching is concerned with connecting any communicating end-points in a network without the need of having direct, dedicated connections between all the possible end-points in the network. Its purpose is to efficiently determine the transmission links that are used in transporting information between two end-points. It is essentially a Physical Layer (Layer 1) and Data-Link Layer (Layer 2) function, and although its description will involve some description of routing type problems, it should not be confused with Network level (Layer 3) routing functions.

Switching Principles

Historically, most communications systems have started with point-to-point links which directly connect together the users wishing to communicate using a dedicated communications circuit. As the distance between users increases beyond the length of the cable (e.g., the length of a telegraph wire), the connection between the users was formed by a number of sections that were connected end-to-end in series to form the circuit. The connection between the users (*A* and *D*) in the figure below (i.e., *A* and *D*) is represented by a series of links (*AB*, *BC* and *CD*) each link connects two entities known as nodes. For a point-to-point circuit, (also known as a permanent circuit) the nodes are patch panels that provide a simple connection between the two links (i.e., the two transmission circuits).

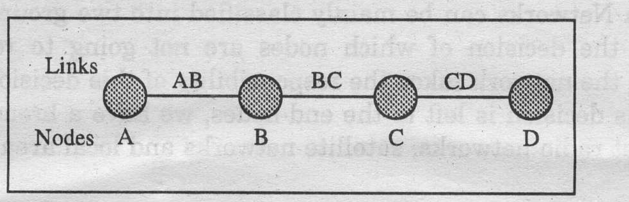


Fig. 6.1 A connection between two systems *A* and *D* formed from 3 links

As the number of connected users increased, it has become infeasible to provide a circuit that connects every user to every other user, and some sharing of the transmission circuits (known as "switching") has become necessary. To accomplish this goal, the data communications network has evolved. A network is a set of nodes that are interconnected to permit the exchange of information.

In this section, we will begin to discuss how various types of "switching" work. Three switching techniques have been proposed for building networks:

- Circuit switching.
- Message switching.
- Packet switching (both virtual circuit and datagram).

Each allows sharing communication facilities among multiple users (end systems) and each uses equipment located at the nodes (intermediate systems) to replace the patch-panels used in a point-to-point connection. Packet switching is most often used for data communication. Most networks consist of many links (see the figure 6.2) that allow more than one path through the network between nodes. A data communications network must be able to select an appropriate path for each required connection.

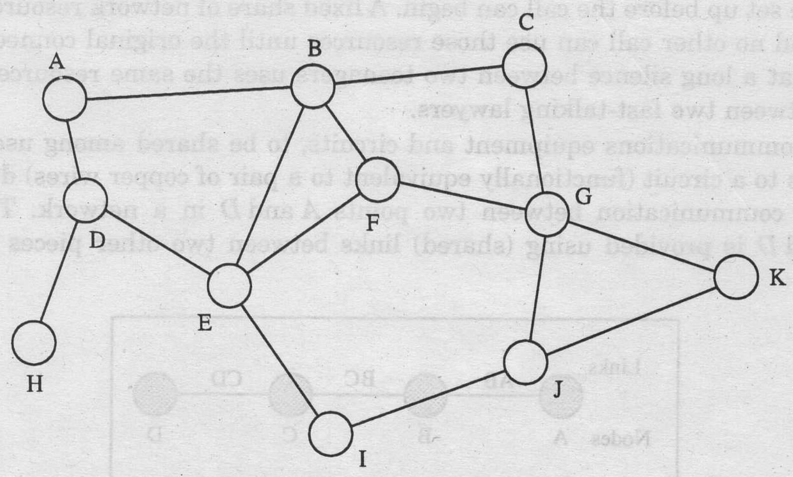


Fig. 6.2 A mesh of network nodes (A—K) connected by links

Any of the three approaches (circuit switching, message switching, and packet switching) could yield minimum delay in a particular situation, though situations where message switching yields minimum delay are rare. The relative performance of circuit switching and packet switching depends strongly on the speed and "cost" of establishing a connection.

6.2 Circuit Switching

In a network like the phone system, communication requires the establishment of a dedicated path between the communicating parties.

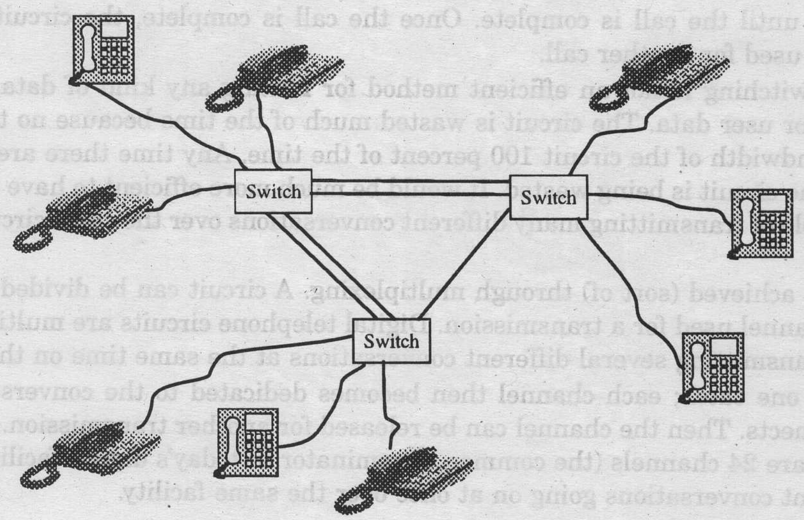
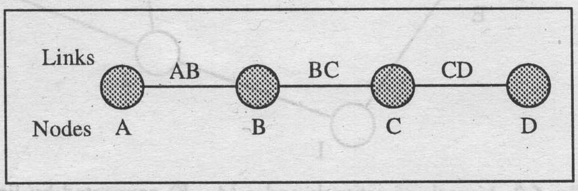


Fig. 6.3 Circuit switcing

Circuit switching is the most familiar technique used to build a communications network. It is used for ordinary telephone calls. Phone networks use circuit switching: an end-to-end circuit must be set up before the call can begin. A fixed share of network resources is reserved for the call, and no other call can use those resources until the original connection is closed. This means that a long silence between two teenagers uses the same resources as an active negotiation between two fast-talking lawyers.

It allows communications equipment and circuits, to be shared among users. Each user has sole access to a circuit (functionally equivalent to a pair of copper wires) during network use. Consider communication between two points *A* and *D* in a network. The connection between *A* and *D* is provided using (shared) links between two other pieces of equipment, *B* and *C*.



In a circuit-switched network, a dedicated circuit must first be connected. Once the circuit has been "nailed up", transmission can begin. When the transmission is complete, the circuit is released for the next transmission.

Let us look at a simple telephone call. When you remove the receiver from the telephone and dial a telephone number, the telephone company searches its database to determine which circuit should be used to deliver the telephone call. If it is a long distance call, the switch knows it must connect to another telephone company office, where a switch called a tandem is located. The tandem switch will then use a circuit that connects it to another office, the toll office switch.

This process continues until there are circuits connected from the originator to the destination. These circuits cannot be used for any other telephone call; they are dedicated to this one call until the call is complete. Once the call is complete, the circuits can then be released and used for another call.

Circuit switching is not an efficient method for routing any kind of data, whether it is digital voice or user data. The circuit is wasted much of the time because no transmission is using the bandwidth of the circuit 100 percent of the time. Any time there are idle period on the circuit, the circuit is being wasted. It would be much more efficient to have a transmission facility capable of transmitting many different conversations over the same circuit at the same time.

This was achieved (sort of) through multiplexing. A circuit can be divided into channels, with each channel used for a transmission. Digital telephone circuits are multiplexed and are capable of transmitting several different conversations at the same time on the same circuit.

There is one catch; each channel then becomes dedicated to the conversation until the caller disconnects. Then the channel can be released for another transmission. So, in the case where there are 24 channels (the common denominator in today's digital facilities), there can be 24 different conversations going on at once over the same facility.

This is better than wasting the circuit for one transmission, but it could still be better. Imagine having no channels. Transmissions are sent over the same circuit as needed, but there are no limits to the number of conversations that can be sent over the same facility at the same time.

Network use is initiated by a connection phase, during which a circuit is set up between source and destination, and terminated by a disconnect phase. These phases, with associated timings, are illustrated in the figure below.

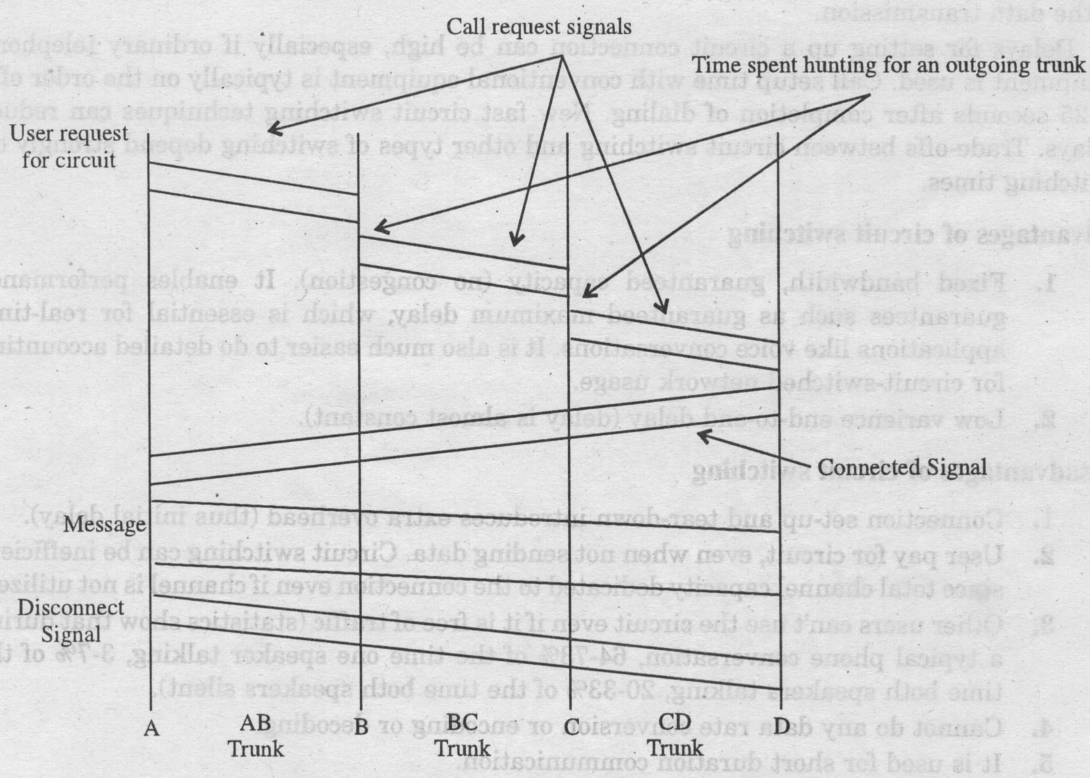


Fig. 6.4 Circuit switched connection between *A* and *D*

In figure 6.4, the most important events in the life of a connection in a four-node circuit-switching network are shown. When a connection is established, the origin-node identifies the first intermediate node (node *A*) in the path to the end-node and sends it a communication request signal. After a user requests a circuit, the desired destination address must be communicated to the local switching node (*B*). In a telephony network, this is achieved by dialing the number.

Node *B* receives the connection request and identifies a path to the destination (*D*) via an intermediate node (*C*). This is followed by a circuit connection phase handled by the switching nodes and initiated by allocating a free circuit to *C* (link *BC*), followed by transmission of a call request signal from node *B* to node *C*. In turn, node C allocates a link (*CD*) and the request is then passed to node *D* after a similar delay.

The circuit is then established and may be used. While it is available for use, resources (i.e., in the intermediate equipment at *B* and *C*) and capacity on the links between the equipment are dedicated to the use of the circuit.

After completion of the connection, a signal confirming circuit establishment (a connect signal in the diagram) is returned; this flows directly back to node *A* with no search delays since the circuit has been established. Transfer of the data in the message then begins. After data transfer, the circuit is disconnected; a simple disconnect phase is included after the end of the data transmission.

Delays for setting up a circuit connection can be high, especially if ordinary telephone equipment is used. Call setup time with conventional equipment is typically on the order of 5 to 25 seconds after completion of dialing. New fast circuit switching techniques can reduce delays. Trade-offs between circuit switching and other types of switching depend strongly on switching times.

Advantages of circuit switching

1. Fixed bandwidth, guaranteed capacity (no congestion). It enables performance guarantees such as guaranteed maximum delay, which is essential for real-time applications like voice conversations. It is also much easier to do detailed accounting for circuit-switched network usage.
2. Low varience end-to-end delay (delay is almost constant).

Disadvantages of circuit switching

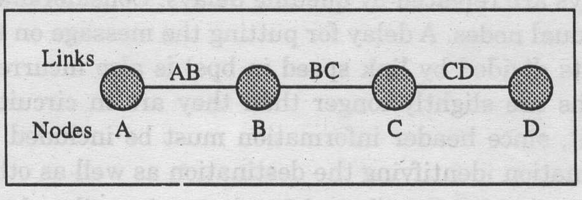
1. Connection set-up and tear-down introduces extra overhead (thus initial delay).
2. User pay for circuit, even when not sending data. Circuit switching can be inefficient since total channel capacity dedicated to the connection even if channel is not utilized.
3. Other users can't use the circuit even if it is free of traffic (statistics show that during a typical phone conversation, 64-73% of the time one speaker talking, 3-7% of the time both speakers talking, 20-33% of the time both speakers silent).
4. Cannot do any data rate conversion or encoding or decoding.
5. It is used for short duration communication.

6.3 Message Switching

Sometimes there is no need for a circuit to be established all the way from the source to the destination. In a computer network, computer's themselves can act as switches.

In a computer network, it is reasonable to imitate the telegraph system rather than the phone company. All of the digital data is sent from the source to the destination as a unit. When there are intermediate nodes between the source and destination, each intermediate node must receive the entire message before sending it on to the next intermediate or final destination. This is called "store and forward" transmission. The intermediate nodes may have to make a decision as to which route the message will be sent. A header is attached to the beginning of the message to identify the destination.

Consider a connection between the users (*A* and *D*) in the figure below (i.e., *A* and *D*) is represented by a series of links (*AB*, *BC* and *CD*).



For instance, when a telex (or email) message is sent from *A* to *D*, it first passes over a local connection (*AB*). It is then passed at some later time to *C* (via link *BC*), and from there to the destination (via link *CD*). At each message switch, the received message is stored, and a connection is subsequently made to deliver the message to the neighboring message switch. Message switching is also known as store-and-forward switching since the messages are stored at intermediate nodes through their route to their destinations.

Figure 6.5 shows life connection events for a message-switching network. When a connection is established, the origin-node identifies the first intermediate node in the path to the end-node and sends it the whole message. After receiving and storing this message, the first intermediate node (node A) identifies the second one (node B) and, when the transmission line is not busy, the former sends the whole message (store-and-forward philosophy). This process is repeated up to the end-node. As can be seen in figure 6.5 no communication release or establishment is needed.

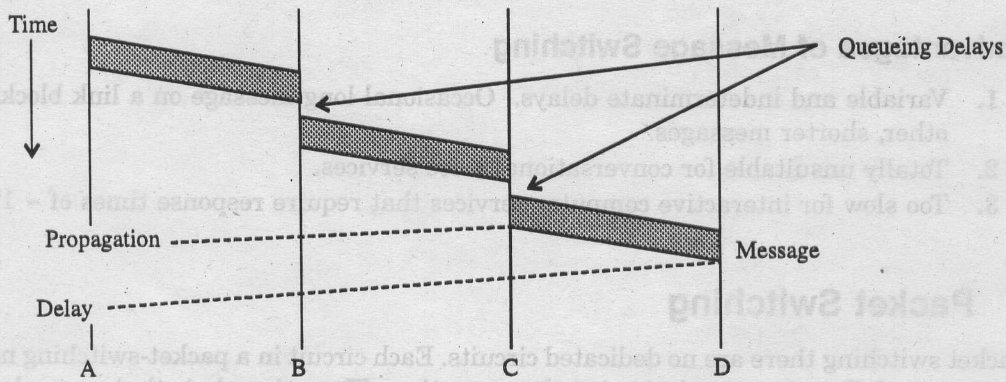


Fig. 6.5 The use of message switching to communicate between *A* and *D*

The figure illustrates message switching; transmission of only one message is illustrated for simplicity. As the figure indicates, a complete message is sent from node *A* to node *B* when the link interconnecting them becomes available. Since the message may be competing with other messages for access to facilities, a queuing delay may be incurred while waiting for the link to become available. The message is stored at *B* until the next link becomes available, with another queuing delay before it can be forwarded. It repeats this process until it reaches its destination.

Circuit setup delays are replaced by queuing delays. Considerable extra delay may result from storage at individual nodes. A delay for putting the message on the communications link (message length in bits divided by link speed in bps) is also incurred at each node through route. Message lengths are slightly longer than they are in circuit switching, after establishment of the circuit, since header information must be included with each message; the header includes information identifying the destination as well as other types of information.

Although message switching is still used for electronic mail and telex transmission, it has largely been replaced by packet switching (in fact, most electronic mail is carried using message switching with the links between message switches provided by packet or circuit-switched networks).

Advantages of Message Switching

1. Data channels are shared among communication devices improving the use of bandwidth.
2. Messages can be stored temporarily at message switches, when network congestion becomes a problem.
3. Priorities may be used to manage network traffic.
4. Broadcast addressing uses bandwidth more efficiently because messages are delivered to multiple destinations.

Disadvantages of Message Switching

1. Variable and indeterminate delays. Occasional long message on a link blocks it for other, shorter messages.
2. Totally unsuitable for conversational voice services.
3. Too slow for interactive computer services that require response times of ~ 1's.

6.4 Packet Switching

In packet switching there are no dedicated circuits. Each circuit in a packet-switching network carries many different transmissions at the same time. The only rule is that every data unit sent through a packet-switching network must have enough information in the header that the nodes in the network can determine how to route the data unit. This tends to add overhead to the data unit, but the trade-off is well invested.

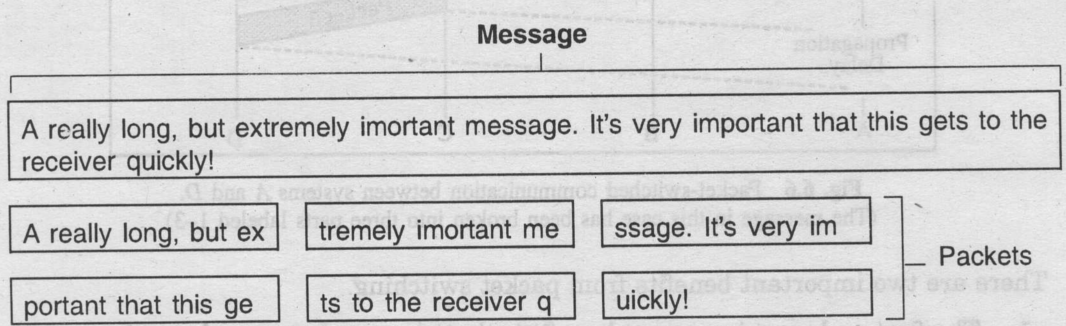
Message switching becomes packet switching if "logical" messages are broken up into smaller units. Any message exceeding a network-defined maximum length is broken up into shorter units, known as packets, for transmission; the packets, each with an associated header, are then transmitted individually through the network.

Definition

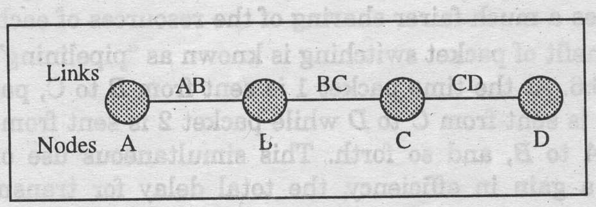
> *"In packet-switching networks, a transmission unit of fixed maximum size that consists of binary digits representing both data and a header containing an identification number, source and destination addresses, and sometimes error-control data."*

The packets can be variable sized or (more often) fixed sized. The size of a packet is usually much smaller than the total data size. Packet sizes range from 48 bytes for ATM to 1500 bytes for Ethernet to 8K bytes for frame relay. Intermediate nodes must receive an entire packet before sending on towards the destination, but they do not have to receive the entire message. Each packet needs a header to identify its destination.

Another advantage of packet switching is the ability to route data units over any route, rather than a fixed route. For example, if I have a lot of data to send, the data will have to be divided into many different data units. These data units do not have to follow the same route in a packet-switching network.



The trick is being able to place the data units in the proper order when they are received. If data units are routed over different paths, it is highly likely that the first data unit may be received after subsequent data units, which means the order of transmission is now mixed up. The protocols used in packet-switching networks have the ability to reassemble the data units into their proper order.



The fundamental difference in packet communication is that the data is formed into packets with a pre-defined header format (i.e., PCI), and well known "idle" patterns that are used to occupy the link when there is no data to be communicated.

A packet network equipment discards the "idle" patterns between packets and processes the entire packet as one piece of data. The equipment examines the packet header information (PCI) and then either removes the header (in an end system) or forwards the packet to another system. If the out-going link is not available, then the packet is placed in a queue until the link becomes free. A packet network is formed by links that connect packet network equipment.

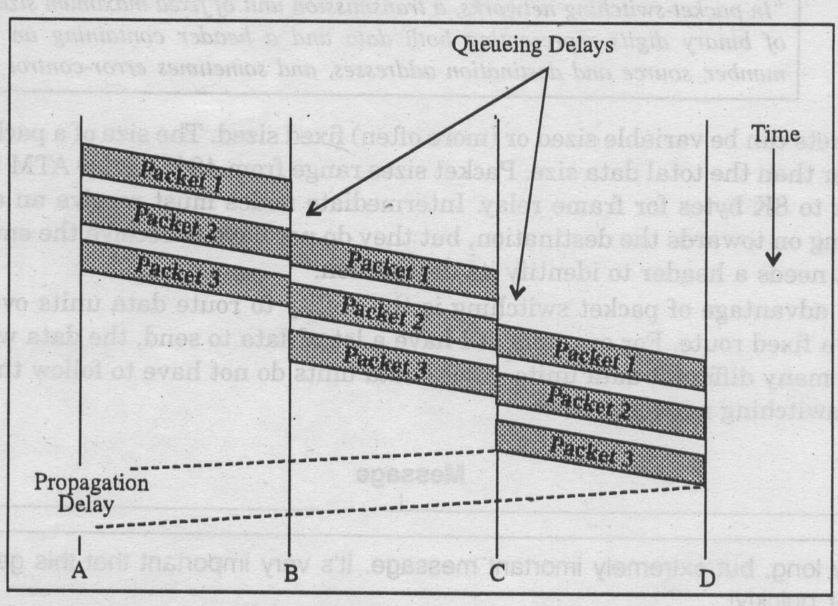


Fig. 6.6 Packet-switched communication between systems *A* and *D*. (The message in this case has been broken into three parts labeled 1–3)

There are two important benefits from packet switching.

1. The first and most important benefit is that since packets are short, the communication links between the nodes are only allocated to transferring a single message for a short period of time while transmitting each packet. Longer messages require a series of packets to be sent, but do not require the link to be dedicated between the transmission of each packet. The implication is that packets belonging to other messages may be sent between the packets of the message being sent from *A* to *D*. This provides a much fairer sharing of the resources of each of the links.
2. Another benefit of packet switching is known as "pipelining". Pipelining is visible in the figure 6.6. At the time packet 1 is sent from *B* to C, packet 2 is sent from *A* to *B*; packet 1 is sent from *C* to *D* while packet 2 is sent from *B* to C, and packet 3 is sent from *A* to *B*, and so forth. This simultaneous use of communications links represents a gain in efficiency, the total delay for transmission across a packet network may be considerably less than for message switching, despite the inclusion of a header in each packet rather than in each message.

Types of packet switched network

Two basic approaches to packet switching are common which are as follows:

1. Datagram Packet Networks.
2. Virtual Circuit Packet Networks.

6.4.1 Datagram Packet Networks

The most common is datagram switching (also known as a "best-effort network" or a network supporting the connection-less network service). This is what is used in the network layer of the Internet.

Datagram transmission uses a different scheme to determine the route through the network of links. Using datagram transmission, each packet is treated as a separate entity and contains a header with the full information about the intended recipient. The intermediate nodes examine the header of a packet and select an appropriate link to an intermediate node that is nearer the destination. In this system, the packets do not follow a pre-established route, and the intermediate nodes (usually known as "routers") do not require prior knowledge of the routes that will be used.

A datagram network is analogous to sending a message as a series of postcards through the postal system. Each card is independently sent to the final destination (using the postal system). To receive the whole message, the receiver must collect all the postcards and sort them into the original order. Not all postcards need be delivered by the postal system, and not all take the same length of time to arrive.

Packet switching based on datagram

The most important events in the life of a communication in a datagram switching network are shown in figure 6.7. The origin-node identifies the first intermediate node in the path and begins to send packets. Each packet carries an origin-node and end-node identifier. The first intermediate node (node A) begins to send packets, without storing the whole message, to the following intermediate node. This process is repeated up to the end-node. As there is neither connection establishment nor connection release, the path follow for each packet from the origin-node to the end-node can be different and therefore, as a consequence of different propagation delays, they can arrive disordered.

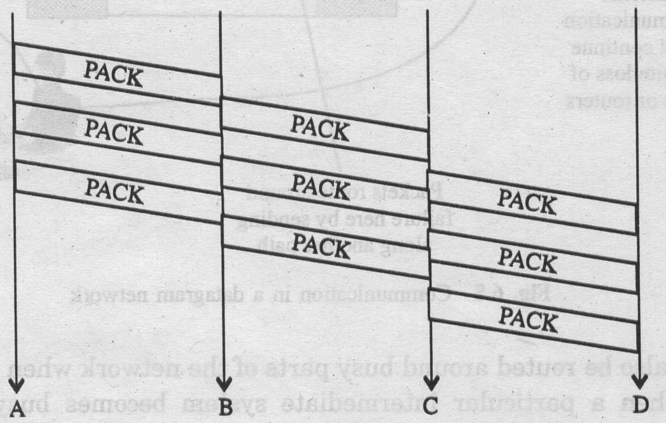


Fig. 6.7
The use of datagram switching to communicate between A and D

In a datagram network delivery is not guaranteed (although they are usually reliably sent). Enhancements, if required, to the basic service (e.g., reliable delivery) must be provided by the end systems (i.e., user's computers) using additional software. The most common datagram network is the Internet, which uses the IP network protocol. Applications that do not require more than a best effort service can be supported by direct use of packets in a datagram network (using the User Datagram Protocol (UDP) transport protocol). Such applications include Internet Video, Voice Communication, messages notifying a user that she/he has received new email, etc. Most Internet applications need additional functions to provide reliable communication (such as end-to-end error and sequence control). Examples include sending email, browsing a web site, or sending a file using the file transfer protocol (FTP). This reliability ensures all the data is received in the correct order with no duplication or omissions. It is provided by additional layers of software algorithms implemented in the End Systems (*A*, *D*). Two examples of this are the Transmission Control Protocol (TCP), and the Trivial File Transfer Protocol (TFTP) that uses UDP.

One merit of the datagram approach is that not all packets need to follow the same path (route) through the network (although frequently packets do follow the same route). This removes the need to set-up and tear-down the path, reducing the processing overhead, and a need for Intermediate Systems to execute an additional protocol.

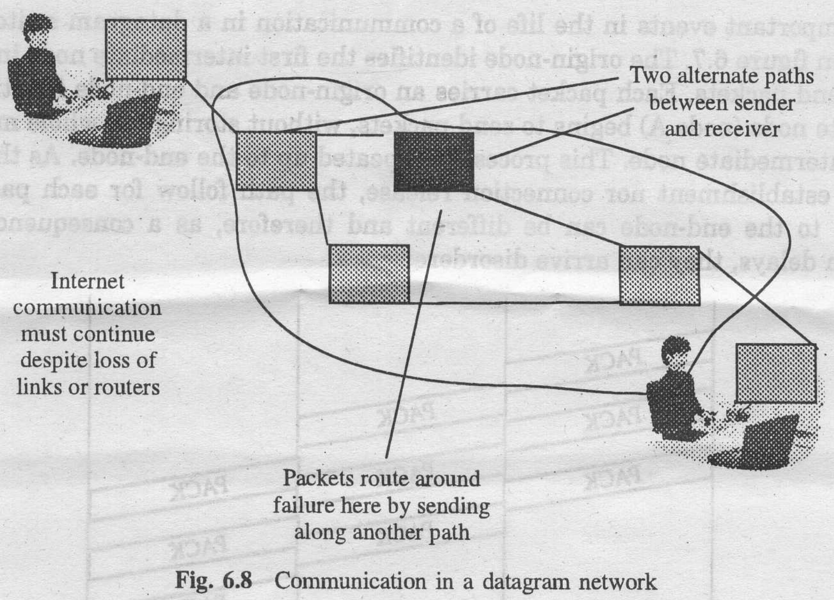


Fig. 6.8 Communication in a datagram network

Packets may also be routed around busy parts of the network when alternate paths exist. This is useful when a particular intermediate system becomes busy or overloaded with excessive volumes of packets to send. It can also provide a high degree of fault tolerance, when an individual intermediate system or communication circuit fails. As long as a route exists through the network between two end systems, they are able to communicate. Only if there is no possible way to send the packets, will the packets be discarded and not delivered. The fate (success/failure) of an application therefore depends only on existence of an actual path between the two End Systems. This is known as "fate sharing"—since the application shares the "fate" of the network.

6.4.2 Virtual Circuit Packet Networks

In virtual circuit packet switching, an initial setup phase is used to set up a fixed route between the intermediate nodes for all packets, which are exchanged during a session between the end nodes (analogous to the circuit-switched telephony network). At each intermediate node, an entry is made in a table to indicate the route for the connection that has been set up. Packets can then use short headers, since only identification of the virtual circuit rather than complete destination address is needed. The intermediate nodes (B, C) process each packet according to the information that was stored in the node when the connection was established.

Packet switching based on virtual circuit

Figure 6.9 shows the events for a virtual circuit (packet) switching network. When a connection is established, the origin-node identifies the first intermediate node (node A) in the path to the end-node and sends it a communication request packet. This process is repeated as many times as needed to reach. Then, the end-node sends a communication acknowledge packet to the origin-node through the intermediate nodes (A, B, C and D) that have been traversed in

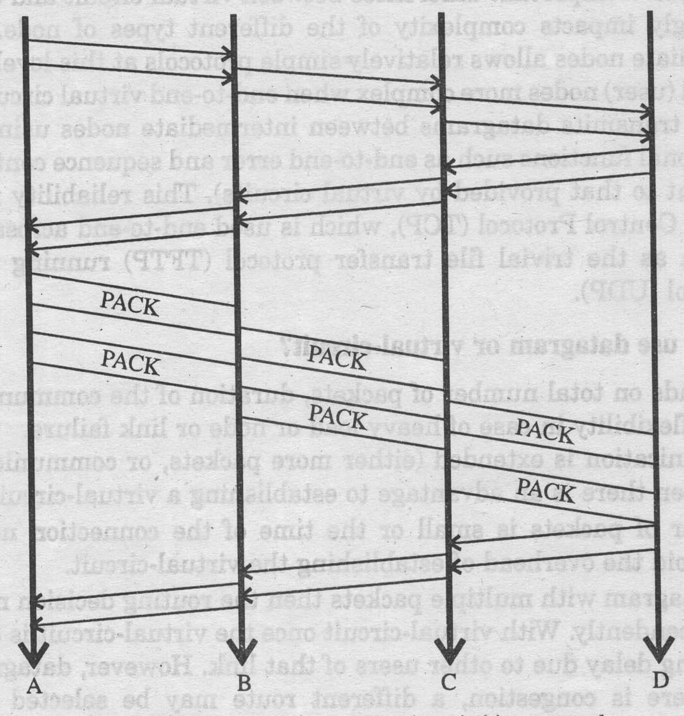


Fig. 6.9 Events for virtual circuit switching network

the communication request. The virtual circuit established on this way will be kept for the whole communication. Once a virtual circuit has been established, the origin-node begins to send packets (each of them has a virtual circuit identifier) to the first intermediate node. Then, the first intermediate node (node A) begins to send packets to the following node in the virtual circuit without waiting to store all message packets received from the origin-node. This process is repeated until all message packets arrive to the end-node. In the communication release, when the origin-node sends to the end-node a communication end packet, the latter answers with an acknowledge packet. There are two possibilities to release a connection:

- No trace of the virtual circuit information is left, so every communication is set-up as if it were the first one.
- The virtual circuit information is kept for future connections.

Enhancements to provide reliability may also be provided. Delivery of packets in proper sequence and with essentially no errors is guaranteed, and congestion control to minimize queuing is common. Delays are more variable than they are with a dedicated circuit, however, since several virtual circuits may compete for the same resources. An initial connection setup phase and a disconnect phase at the end of data transfer are required (as in the circuit-switched network). The most common form of virtual circuit network was ATM and X.25, which for a while were commonly used for public packet data networks.

Differences between datagram and virtual circuit networks

There are a number of important differences between virtual circuit and datagram networks. The choice strongly impacts complexity of the different types of node. Use of datagrams between intermediate nodes allows relatively simple protocols at this level, but at the expense of making the end (user) nodes more complex when end-to-end virtual circuit service is desired.

The Internet transmits datagrams between intermediate nodes using IP. Most Internet users need additional functions such as end-to-end error and sequence control to give a reliable service (equivalent to that provided by virtual circuits). This reliability may be provided by the Transmission Control Protocol (TCP), which is used end-to-end across the Internet, or by applications such as the trivial file transfer protocol (TFTP) running on top of the User Datagram Protocol (UDP).

When would you use datagram or virtual-circuit?

This choice depends on total number of packets, duration of the communication, traffic load, and also desired flexibility in case of heavy load or node or link failure.

If the communication is extended (either more packets, or communication over a longer period of time) then there is an advantage to establishing a virtual-circuit.

If the number of packets is small or the time of the connection needed is small, use datagram and avoid the overhead of establishing the virtual-circuit.

If you use datagram with multiple packets then the routing decision needs to be made for each packet independently. With virtual-circuit once the virtual-circuit is established the only delay is a queueing delay due to other users of that link. However, datagram is more flexible in that when there is congestion, a different route may be selected and so the overall throughput may be higher under heavy traffic load. The same applies in the case of a link or node failure—if the virtual-circuit was used the effect is more pronounced than if datagram were used. Also with virtual-circuit there is a possibility of providing end-to-end sequence and error control if needed at the network layer rather than at the transport layer.

Advantages of Packet Switching

1. Bandwidth optimized by enabling many devices to use the same communication channel.
2. Routes may be changed on the fly, achieving the best efficiency possible.
3. Entire messages not stored at switches reducing delays.
4. Line efficiency is greater since link is shared by packets from multiple connections.
5. Data rate conversion between hybrid links possible; in contrast circuit switching would revert to lowest common data rate.
6. Less hard drive space needed but more actual memory needed.
7. It is more flexible (i.e. doesn't care too much what's been sent, as long as it can be packetized).

Disadvantages of Packet Switching

1. More processing power required due to complex routing protocols. (Like recognizing when a packet has been lost).
2. No guarantee in delay.
3. Algorithms are more complicated.

6.5 Comparison Of Various Techniques

Circuit Switching vs. Packet Switching

Packet switching is favored over circuit switching for many different reasons. It is more reliable than circuit switching because if a particular circuit in the network should fail, the routers in the network simply route data units over different circuits, taking a different route altogether. In a circuit-switched network, this is not possible. If a circuit fails in the middle of a transmission, the entire connection must be released and a new one established, which means the conversation must start over again (think of being disconnected from a telephone call; the whole process of connecting must be repeated).

In principle, circuit switching and packet switching both are used in high-capacity networks. In circuit-switched networks, network resources are static, set in "copper" if you will, from the sender to receiver before the start of the transfer, thus creating a "circuit". The resources remain dedicated to the circuit during the entire transfer and the entire message follows the same path. In packet-switched networks, the message is broken into packets, each of which can take a different route to the destination where the packets are recompiled into the original message.

All the above can be handled by a router or a switch but much of IT today is going toward flat-switched networks. So when we're talking about circuit switching or packet switching, we are more and more talking about doing it on a switch.

Circuit-switched circuits are transparent to the user but there is no data rate matching/ any form of encoding or decoding, with dissimilar links. So this may be a disadvantage. The advantage is no delays at intermediate nodes.

Packet switching could have variable delays. Packets could arrive out of sequence. Packets must be digitized. Will work with mismatched links and will perform the data rate conversion.

Circuit switching

- Used in the telephone system: Network resources in the telephone system are reserved from your phone to the phone you call when you place the call; they're released when you hang up.

Packet switching

- Internet is packet switched: Data you wish to send is put into packets, which are then sent through the network. Network resources are only used during the time it takes to transmit each packet.

Feature	Circuit Switching	Packet Switching
Data sent as packets?	No	Yes
Packets follow same route?	N/A	May or may not (may for VC)
Resources reserved in network?	Yes	May or may not (not for DG)
Data send can experience variable latency	No	Yes
Connection establishment done?	Yes	VC:yes DG: no
State information stored at network nodes	?	VC: yes (tables to route virtual circuits) DG: no
Good for connection-less service	No	VC: yes DG: no
Impact of node/switch crash	All circuits through switch fail	VC: all virtual circuits through node fail DG: only packets at node are lost
Addressing info needed?	Only when call is set up	Every packet needs... VC: virtual circuit number DG: full source, destination address
Congestion control	Unnecessary	VC: easy if sufficient buffers allocated DG: hard

6.6 Cell Relay

Cell Relay is a newer approach to data networks. The cell is the data unit. In order to support voice, the data units must be small so that they can be processed quickly and sent through the network with minimal delay. This is not true with data, which favors large data units. Voice requires small data units, and data favors large data units.

In a cell relay network, the facility is used when needed. Whenever there is information to be transmitted, the switch simply sends the data units. There is no need to negotiate for a connection (as is the case in circuit switching), there is no need for a channel to be allocated (there are no channels in ATM), and as long as there is enough bandwidth to support it, there can be unlimited transmissions over the same facility.

6.7 Applications Of Switching Techniques

There's much talk about the coming mobile Internet, about how people will have a wireless, always on connection to the web. How will that come about? In two words, packet switching, a fundamental, elemental change between how wireless was delivered in the past and how it will be presented in the future.

Conventional cellular radio and landline telephony use circuit switching. A service like Cellular Digital Packet Data or CDPD, by contrast, employs packet switching. Wireless services now developing such as General Packet Radio Service or GPRS, Bluetooth, and 3G, will use packet switching as well.

Circuit switching dominates the public switched telephone network or PSTN. Network resources set up calls over the most efficient route, even if that means a call to New York from San Francisco goes through switching centers in San Diego, Chicago, and Saint Louis But no matter how convoluted the route, that path or circuit stays the same throughout the call. It's like having a dedicated railroad track with only one train, your call, permitted on the track at a time.

Packet switching dominates data networks like the Internet. A data call or communication from San Francisco to New York is handled much differently than with circuit switching. With circuit, all packets go directly to the receiver in an orderly fashion, one after another on a single track, like the train, hauling one boxcar after another. With packet switching routers determine a path for each packet or boxcar on the fly, dynamically, ordering them about to use any railroad track available to get to the destination. Other packets from other calls race upon these circuits as well, making the most use of each track or path, quite unlike the circuit switched calls that occupy a single path to the exclusion of all others.

Upon getting to their destination, the individual packets get put back into order by a packet assembler. That's because the different routes practically ensures that packets will arrive at different times. This approach is acceptable when calling up a web page or downloading a file, since a tiny delay is hardly noticed. But one notices even the tiniest delay with voice. This point is really important. Circuit switching guarantees the best sounding call because all packets go in order and there is no delay. Delays in packet switching for voice cause voice quality to fall apart, as anyone who has talked over the Internet can tell you.

As technology gets better with time, voice over packet switched networks will get better. Packet switched networks exist for the data communication needs of education, business, and government. These networks rely on telephone lines, of course, but the circuits are so arranged that they retain a permanent connection with their customers. The Public Data Network or Packet Switched Network stands as the data counterpart to the Public Switched Telephone Network. Unlike circuit switching, no one call takes up an entire channel for an entire session. Bits get sent only when traffic goes on, when people actually speak. During pauses in a conversation a channel gets filled with pieces of other conversations. Because your call doesn't hog an entire circuit the telephone system can permit an always-on connection. You might pay a flat monthly charge or by the bandwidth or bits you actually use. Whether wireless operators can afford to do so is difficult to decide. Too many customers' means building many more expensive cell sites. Even if technology permits we may stay with a per minute charge.

If packet switching is so efficient, why hasn't the landline public switched telephone network converted to it? The answer is time and money. Replacing circuit switched switches with packet switches across the country would be a monumental task, requiring billions of dollars over years and years. The legacy of circuit switching will be around for quite a long time, following us far into the new century. Still, traffic engineers must think about changing, with lengthy dial up calls to the internet placing huge demands on switches that were never planned for, circuits now tied up longer than ever imagined. But change has to come at some point, and the Internet's traffic now motivates engineers to move toward a unified switching method in the PSTN.

DSL and ASDL and cable modem connections will either speed or retard this transition; a local telephone company directs this broadband traffic to a packet switch, bypassing the existing local, circuit-based switch. As broadband users increase call holding times should decrease, as dial up modems are taken out of service. The local switch should not be as overwhelmed as many currently are. A Telco may then decide to delay a transition to packet switching.

• • •

7

Multiplexing Techniques

7.1 Telecommunication Fundamentals

Figure 7.1 shows a switching network. Lines are the media links. Ovals are called network nodes. Media links simply carry data from one point to other. Nodes take the incoming data and route them to an output port.

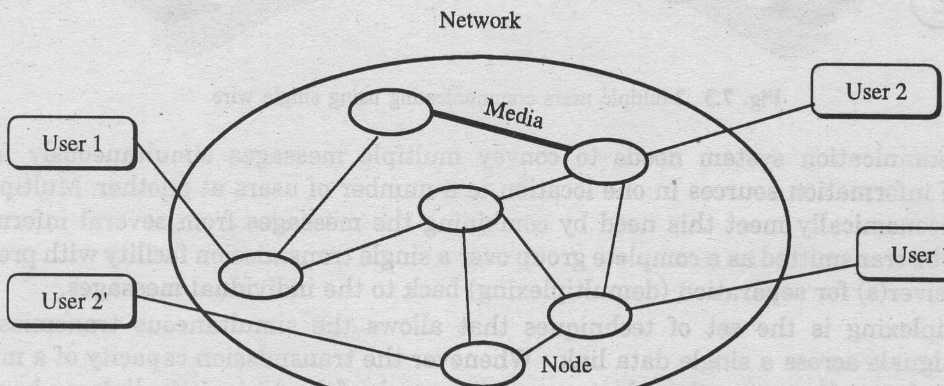


Fig. 7.1 Switching network

If two different communication paths intersect through this network they have to share some resources. Two paths can share a media link or a network node. Next sections describe these sharing techniques.

Media Sharing Techniques

Media sharing occurs when two communication channels use the same media.

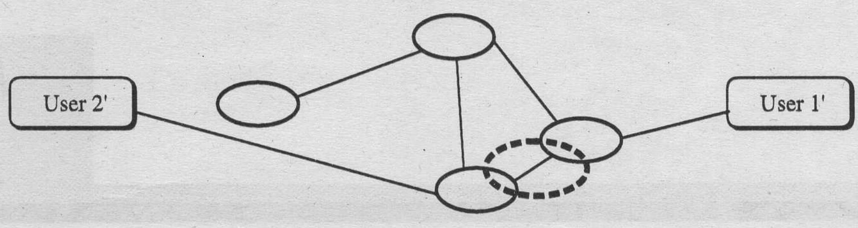


Fig. 7.2 Media sharing

This section presents how some communication channels can use the same media link without architecture considerations.

7.2 Information On Multiplexing

Sometimes, multiple users want to communicate using just one wire/fiber.

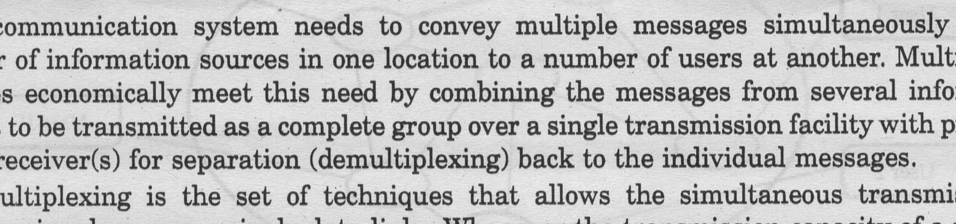


Fig. 7.3 Multiple users communicating using single wire

A communication system needs to convey multiple messages simultaneously from a number of information sources in one location to a number of users at another. Multiplexing schemes economically meet this need by combining the messages from several information sources to be transmitted as a complete group over a single transmission facility with provision at the receiver(s) for separation (demultiplexing) back to the individual messages.

Multiplexing is the set of techniques that allows the simultaneous transmission of multiple signals across a single data link. Whenever the transmission capacity of a medium linking two devices is greater than the transmission needs of the devices, the link can be shared in order to maximize the utilization of the link, much as one cable can carry a hundred channels of TV.

More information can be conveyed in a given amount of time by dividing the bandwidth of a signal carrier so that more than one modulated signal is sent on the same carrier. The carrier is sometimes referred to as a channel and each separate signal carried on it is called a subchannel. (In some usages, each subchannel is known as a channel.) The device that puts the separate signals on the carrier and takes them off of received transmissions is a multiplexer.

Definition

In telecommunications, multiplexing (MUXing) is the combining of two or more information channels onto a common transmission medium using multiplexer or (MUX). A demultiplexer performs the opposite function to a multiplexer in which a multiplexed channel is broken out into its constituent individual channels.

Types of Multiplexing

In electrical communications, the two basic forms of multiplexing are time-division multiplexing (TDM) and frequency-division multiplexing (FDM). In optical communications, the analog of FDM is referred to as wavelength division multiplexing (WDM).

FDM is usually used for analog communication and TDM is used for digital communication.

The following are several examples of different multiplexing methods:

- **Time Division Multiplexing (TDM) :** A type of multiplexing that combines data streams by assigning each stream a different time slot in a set. TDM repeatedly transmits a fixed sequence of time slots over a single transmission channel.
- **Frequency Division Multiplexing (FDM) :** A multiplexing technique that uses different frequencies to combine multiple streams of data for transmission over a communications medium. FDM assigns a discrete carrier frequency to each data stream and then combines many modulated carrier frequencies for transmission. For example, television transmitters use FDM to broadcast several channels at once.
- **Statistical Time Division Multiplexing (STDM) :** Time slots are assigned to signals dynamically to make better use of bandwidth
- **Wavelength Division Multiplexing (WDM) :** Each signal is assigned a particular wavelength; a type of multiplexing developed for use on optical fiber. WDM modulates each of several data streams onto a different part of the light spectrum. WDM is the optical equivalent of FDM.

7.3 Time Division Multiplexing (TDM)

A type of multiplexing where two or more channels of information are transmitted over the same link by allocating a different time interval ("slot" or "slice") for the transmission of each channel, i.e., the channels take turns to use the link. Some kind of periodic synchronizing signal or distinguishing identifier is usually required so that the receiver can tell which channel is which.

Time-division multiplexing (TDM) is a method of putting multiple data streams in a single signal by separating the signal into many segments, each having a very short duration. Each individual data stream is reassembled at the receiving end based on the timing.

The circuit that combines signals at the source (transmitting) end of a communications link is known as a multiplexer. It accepts the input from each individual end user, breaks each signal into segments, and assigns the segments to the composite signal in a rotating, repeating sequence. The composite signal thus contains data from multiple senders. At the other end of the long-distance cable, the individual signals are separated out by means of a circuit called a demultiplexer, and routed to the proper end users. A two-way communications circuit requires a multiplexer/demultiplexer at each end of the long-distance, high-bandwidth cable.

If many signals must be sent along a single long-distance line, careful engineering is required to ensure that the system will perform properly. An asset of TDM is its flexibility. The scheme allows for variation in the number of signals being sent along the line, and constantly adjusts the time intervals to make optimum use of the available bandwidth. The Internet is a classic example of a communications network in which the volume of traffic can change drastically from hour to hour.

TDM becomes inefficient when traffic is intermittent because the time slot is still allocated even when the channel has no data to transmit. Statistical time division multiplexing was developed to overcome this problem.

7.3.1 Theory of Time Division Multiplexing

Time Division Multiplexing operates by dividing the network bandwidth into fixed bandwidth segments. Each segment or channel is assigned to a user and is given its own time slot for using the network. First, information from channel *A* is transmitted, then information from channel *B*, and so on in a regular sequence, cycling back to channel *A* and continuing. Bandwidth allocation is static, that is, each channel has a fixed predetermined bandwidth. Because of this multiplexing technique, TDM is protocol insensitive and is capable of combining various protocols onto a single high speed transmission link.

Conventional TDM systems usually employ either Bit-Interleaved or Byte-Interleaved multiplexing schemes. Each time slot accommodates either a bit (1 or 0) or a byte (usually 8 bits long to represent a character, number, or symbol).

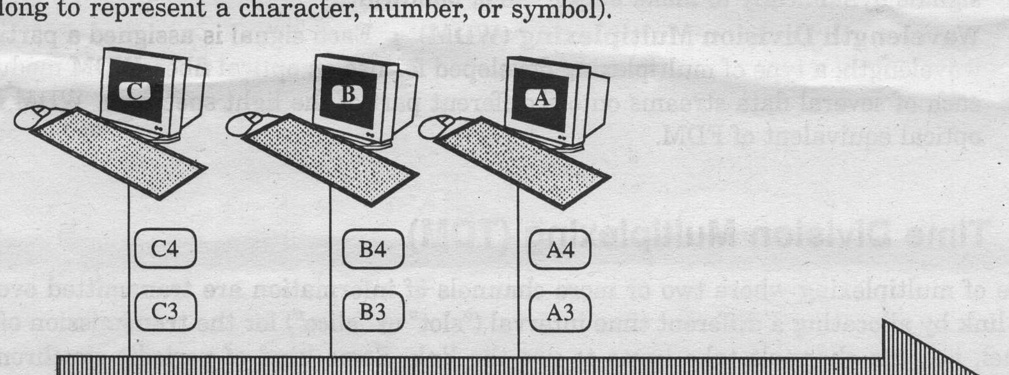


Fig. 7.4 High-bit-rate channel divided into series of time slots

It's often practical to combine a set of low-bit-rate streams, each with a fixed and pre-defined bit rate, into a single high-speed bit stream that can be transmitted over a single channel. This technique is called time division multiplexing (TDM) and has many applications,