including wireline telephone systems and some cellular telephone systems. The main reason to use TDM is to take advantage of existing transmission lines. It would be very expensive if each low-bit-rate stream were assigned a costly physical channel (say, an entire fiber optic line) that extended over a long distance.

Consider, for instance, a channel capable of transmitting 192 kbit/sec from Chicago to New York. Suppose that three sources, all located in Chicago, each have 64 kbit/sec of data that they want to transmit to individual users in New York. As shown in figure 7.5 the high-bit-rate channel can be divided into a series of time slots, and the time slots can be alternately used by the three sources. The three sources are thus capable of transmitting all of their data across the single, shared channel. Clearly, at the other end of the channel (in this case, in New York), the process must be reversed (i.e., the system must divide the 192 kbit/sec multiplexed data stream back into the original three 64 kbit/sec data streams, which are then provided to three

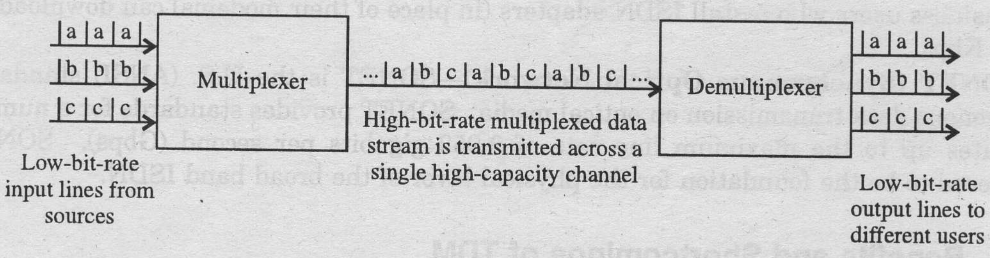


Fig. 7.5 Time division multiplexing

different users). This reverse process is called demultiplexing.

Choosing the proper size for the time slots involves a trade-off between efficiency and delay. If the time slots are too small (say, one bit long) then the multiplexer must be fast enough and powerful enough to be constantly switching between sources (and the demultiplexer must be fast enough and powerful enough to be constantly switching between users). If the time slots are larger than one bit, data from each source must be stored (buffered) while other sources are using the channel. This storage will produce delay. If the time slots are too large, then a significant delay will be introduced between each source and its user. Some applications, such as teleconferencing and videoconferencing, cannot tolerate long delays.

So far we have considered a form of TDM that is based on fixed slot assignments to each of the low-bit-rate data streams. In other words, each stream has predefined slot positions in the combined stream, and the receiver must be aware which slots belong to which input stream. Both transmission ends, the transmitter and the receiver, must be perfectly synchronized to the slot period. For this reason, the technique is usually called synchronous TDM.

There is another important version of TDM, usually referred to as **statistical TDM.** Statistical TDM is useful for applications in which the low-bit-rate streams have speeds that vary in time. For example, a low-bit-rate stream to a single terminal in a computer network may fluctuate between 2 kbit/sec and 50 kbit/sec during an active connection session (we've all seen variable speeds during Internet connections, for instance). If we assign the stream enough slots for its peak rate (that is, for 50 kbit/sec), then we will be wasting slots when the rate drops well below the peak value. This waste can be especially significant if the system has many variable-speed low-bit-rate streams.

7.3.2 Some examples of TDM based services

Public telephone networks

T1 : The T-carrier system is entirely digital, using pulse code modulation and time division multiplexing.

E1 : (or E-1) is a European digital transmission format and is the equivalent of the North American T-1 format.

Narrow-band ISDN (Integrated Services Digital Network)—is a set of standards for digital transmission over ordinary telephone copper wire as well as over other media. Home and business users who install ISDN adapters (in place of their modems) can download at up to 128 Kbps.

SONET (Synchronous Optical Network)—SONET is the U.S. (ANSI) standard for synchronous data transmission on optical media. SONET provides standards for a number of line rates up to the maximum line rate of 9.953 gigabits per second (Gbps). SONET is considered to be the foundation for the physical layer of the broad band ISDN.

7.3.3 Benefits and Shortcomings of TDM

The relative simplicity and low cost of TDM hardware devices makes it possible to build cost-effective networks of widely varying node count, geographic extent, capacity, and speed.

However, TDM is best utilized if everyone always has something to send, otherwise, it wastes time. Time Division Multiplexing operates with low network efficiency. For example, if one of the voice channels on our connection is inactive, the bandwidth dedicated to the channel is unavailable to the other users; therefore, the allotted time slot remains idle.

Another deficiency is TDM's inability to transport a dynamically varying combination of voice, fax, and LAN traffic. Because the bandwidth is statically allocated, the momentarily high bandwidth requirements of one channel on a network cannot be met. There is no way for one TDM channel to grab all of the network bandwidth, even for a short time, to expedite the transfer of a large amount of data.

7.3.4 Statistical Time Division Multiplexing

(STDM, StatMUX) A system developed to overcome some inefficiencies of standard time division multiplexing, where time slices are still allocated to channels, even if they have no information to transmit.

STDM uses a variable time slot length and by allowing channels to compete for any free slot space. It employs a buffer memory that temporarily stores the data during periods of peak traffic. This scheme allows STDM to waste no high-speed line time with inactive channels. STDM requires each transmission to carry identification information (i.e., a channel identifier). To reduce the cost of this overhead, a number of characters for each channel are grouped together for transmission.

Statistical TDM works by calculating the average transmission rates of the streams to be combined, and then uses a high-speed multiplexing link with a transmission rate that is equal to (or slightly greater than) the statistical average of the combined streams. Since the transmission rates from each source are variable, we no longer assign a fixed number of time slots to each data stream. Rather, we dynamically assign the appropriate number of slots to accommodate the current transmission rates from each stream. Because the combined rate of all the streams will also fluctuate in time between two extreme values, we need to buffer the output of the low-bit-rate streams when the combined rate exceeds the transmission rate of the high-speed link.

With statistical TDM, we are no longer relying on synchronized time slots with fixed assignments for each input stream, as we did with synchronous TDM. So how does the demultiplexer in statistical TDM know which of the received bits belongs to which data stream? Prior to transmission, we divide each stream of bits coming from a source into fixed-size blocks. We then add a small group of bits called a header to each block, with the header containing the addresses of the source and intended user for that block. The block and the header are then transmitted together across the channel. Combined, the block and header are called a packet.

Actually, the header may contain other information besides the source and user addresses, such as extra bits for error control or additional bits for link control (used, for example, to indicate the position of a particular block in a sequence of blocks coming from the same user, or to indicate priority level for a particular message). Extra bits can also be added to the beginning and end of a block for synchronization; a particular pattern of bits, called a start flag, can be used in the header to mark the start of a block, and another particular pattern of bits, called an end flag, can be used to conclude the block. Each block transmitted across the channel thus contains a group of information bits that the user wants, plus additional bits needed by the system to ensure proper transmission. These additional bits, while necessary to system operation, reduce the effective transmission rate on the channel. Figures 7.6 and 7.7 present the statistical TDM technique and the structure of a typical packet.

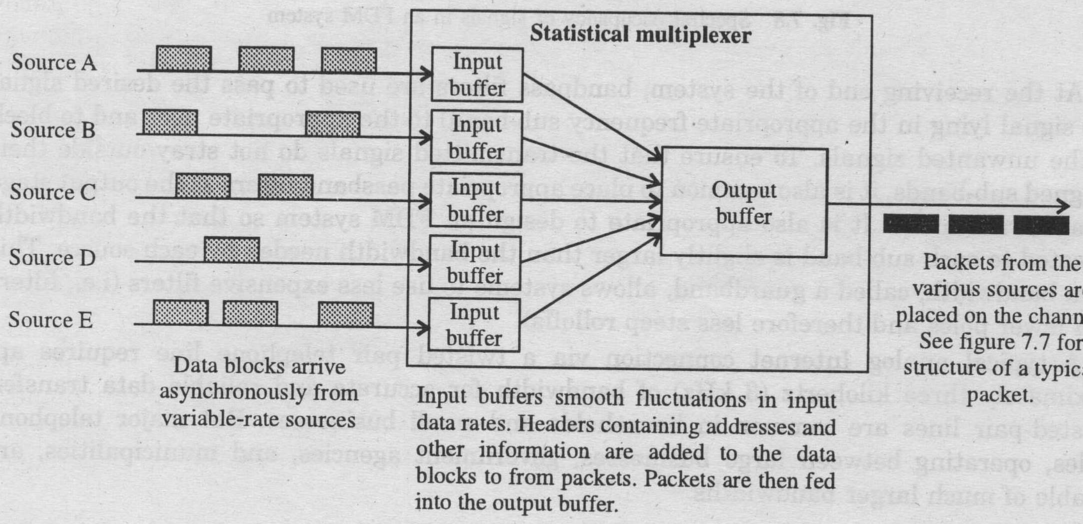


Fig. 7.6 Statistical TDM

Start flag	Address field	Control field	Information bits	Error control	End flag

Fig. 7.7 Structure of a typical statistical TDM packet

7.4 Frequency Division Multiplexing (FDM)

Frequency-division multiplexing (FDM) is a scheme in which numerous signals are combined for transmission on a single communications line or channel. Each signal is assigned a different frequency (subchannel) within the main channel.

FDM means that the total bandwidth available to the system is divided into a series of non-overlapping frequency sub-bands that are then assigned to each communicating source and user pair. Figure 7.8 show how this division is accomplished for a case of three sources at one end of a system that are communicating with three separate users at the other end. Note that each transmitter modulates its source's information into a signal that lies in a different frequency sub-band (Transmitter 1 generates a signal in the frequency sub-band between 92.0 MHz and 92.2 MHz, Transmitter 2 generates a signal in the sub-band between 92.2 MHz and 92.4 MHz, and Transmitter 3 generates a signal in the sub-band between 92.4 MHz and 92.6 MHz). The signals are then transmitted across a common channel.

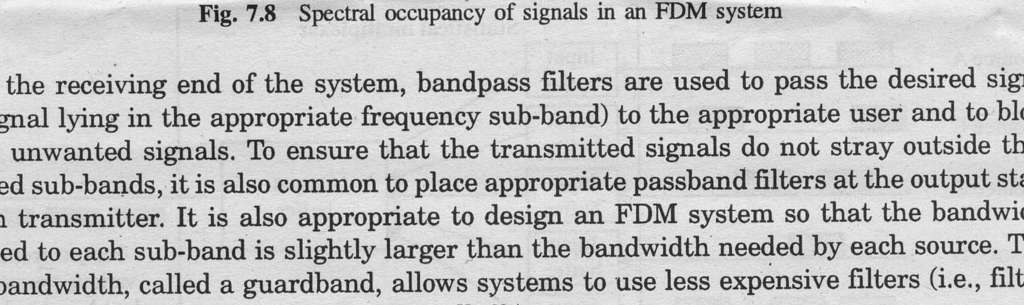


Fig. 7.8 Spectral occupancy of signals in an FDM system

At the receiving end of the system, bandpass filters are used to pass the desired signal (the signal lying in the appropriate frequency sub-band) to the appropriate user and to block all the unwanted signals. To ensure that the transmitted signals do not stray outside their assigned sub-bands, it is also common to place appropriate passband filters at the output stage of each transmitter. It is also appropriate to design an FDM system so that the bandwidth allocated to each sub-band is slightly larger than the bandwidth needed by each source. This extra bandwidth, called a guardband, allows systems to use less expensive filters (i.e., filters with fewer poles and therefore less steep rolloffs).

A typical analog Internet connection via a twisted pair telephone line requires approximately three kilohertz (3 kHz) of bandwidth for accurate and reliable data transfer. Twisted-pair lines are common in households and small businesses. But major telephone cables, operating between large businesses, government agencies, and municipalities, are capable of much larger bandwidths.

Suppose a long-distance cable is available with a bandwidth allotment of three megahertz (3 MHz). This is 3,000 kHz, so in theory, it is possible to place 1,000 signals, each 3 kHz wide, into the long-distance channel. The circuit that does this is known as a multiplexer. It accepts the input from each individual end user, and generates a signal on a different frequency for each of the inputs. This results in a high-bandwidth, complex signal containing data from all the end users. 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.

When FDM is used in a communications network, each input signal is sent and received at maximum speed at all times. This is its chief asset. However, if many signals must be sent along a single long-distance line, the necessary bandwidth is large, and careful engineering is required to ensure that the system will perform properly. In some systems, a different scheme, known as time-division multiplexing, is used instead.

Where frequency division multiplexing is used as to allow multiple users to share a physical communications channel, it is called frequency-division multiple access (**FDMA**).

FDMA is the traditional way of separating radio signals from different transmitters.

7.4.1 Theory of Frequency Division Multiplexing

In many communication systems, a single, large frequency band is assigned to the system and is shared among a group of users. Examples of this type of system include:

Examples of this type of system include:

1.	A microwave transmission line connecting two sites over a long distance. Each site has a number of sources generating independent data streams that are transmitted simultaneously over the microwave link.
2.	AM or FM radio broadcast bands, which are divided among many channels or stations. The stations are selected with the radio dial by tuning a variable-frequency filter.
3.	A satellite system providing communication between a large numbers of ground stations that are separated geographically but that need to communicate at the same time. The total bandwidth assigned to the satellite system must be divided among the ground stations.
4.	A cellular radio system that operates in full-duplex mode over a given frequency band. The earlier cellular telephone system used analog communication methods. The bandwidth for these systems was divided into a large number of channels. Each pair of channels was assigned to two communicating end-users for full-duplex communications.

The figure 7.9 illustrates the FDM multiplexing process. The first PC terminal is sending "1010" whereas the second terminal is sending "0110". The multiplexing process starts by applying amplitude modulation into each signal by using different carrier frequencies as f_1 and f_2. Then both signals are combined and sent.

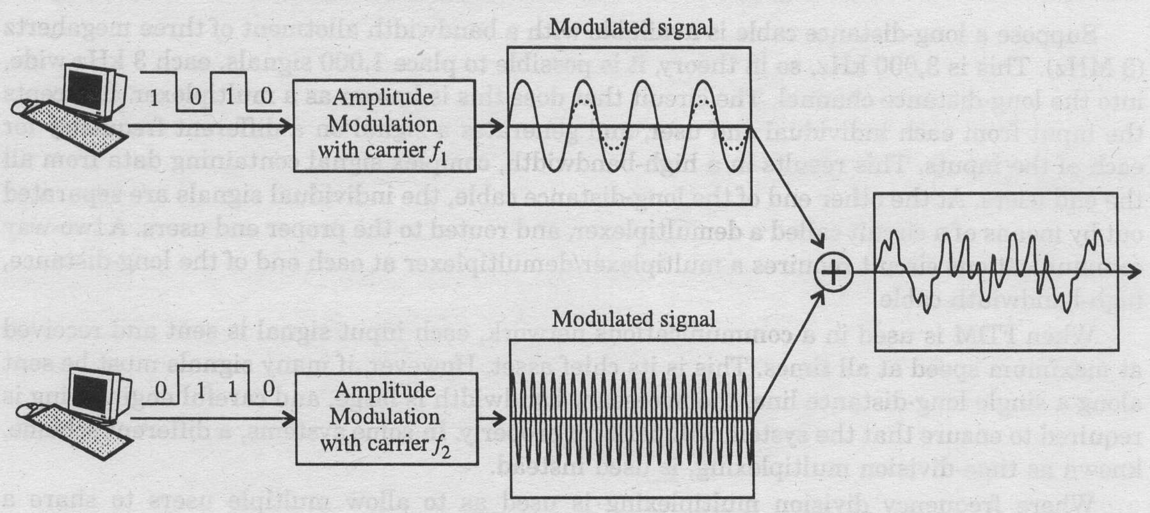


Fig. 7.9 FDM multiplexing process

In demultiplexing process, we use filters to decompose the multiplexed signal into its constituent component signals. Then each signal is passed to an amplitude demodulation process to separate the carrier signal from the message signal. Then, the message signal is sent to the waiting receiver.

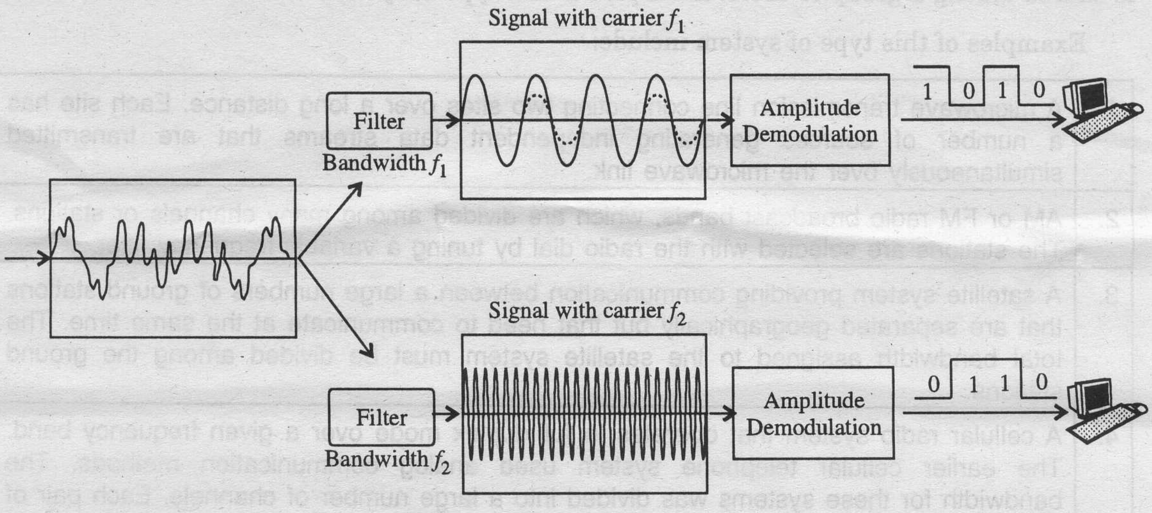


Fig. 7.10 FDM demultiplexing process

7.4.2 Advantages and Disadvantages of FDM relative to TDM

The main advantage is that unlike TDM, FDM is not sensitive to propagation delays. Channel equalization techniques needed for FDM systems are therefore not as complex as those for TDM systems.

Disadvantages of FDM include the need for bandpass filters, which are relatively expensive and complicated to construct and design (remember that these filters are usually used in the transmitters as well as the receivers). TDM, on the other hand, uses relatively simple and less costly digital logic circuits. Another disadvantage of FDM is that in many practical communication systems, the power amplifier in the transmitter has nonlinear characteristics (linear amplifiers are more complex to build), and nonlinear amplification leads to the creation of out-of-band spectral components that may interfere with other FDM channels. Thus, it is necessary to use more complex linear amplifiers in FDM systems.

7.4.3 OFDM (Orthogonal Frequency Division Multiplexing)

OFDM is a technique for increasing the amount of information that can be carried over a wireless network.

In frequency-division multiplexing, multiple signals, or carriers, are sent simultaneously over different frequencies between two points. However, FDM has an inherent problem: Wireless signals can travel multiple paths from transmitter to receiver (by bouncing off buildings, mountains and even passing airplanes); receivers can have trouble sorting all the resulting data out.

Orthogonal FDM deals with this multipath problem by splitting carriers into smaller subcarriers, and then broadcasting those simultaneously. Orthogonal FDM's (OFDM) spread spectrum technique distributes the data over a large number of carriers that are spaced apart at precise frequencies. This spacing provides the "orthogonality" in this technique, which prevents the demodulators from seeing frequencies other than their own. The benefits of OFDM are high spectral efficiency, resiliency to RF interference, and lower multi-path distortion. This is useful because in a typical terrestrial broadcasting scenario there are multipath-channels (i.e., the transmitted signal arrives at the receiver using various paths of different length). Since multiple versions of the signal interfere with each other (inter symbol interference (ISI)) it becomes very hard to extract the original information.

OFDM is sometimes called multi-carrier or discrete multi-tone modulation. It is the modulation technique used for digital TV in Europe, Japan and Australia.

Uses

DAB—OFDM forms the basis for the Digital Audio Broadcasting (DAB) standard in the European market.

ADSL—OFDM forms the basis for the global ADSL (asymmetric digital subscriber line) standard.

Wireless Local Area Networks—Development is ongoing for wireless point-to-point and point-to-multipoint configurations using OFDM technology.

In a supplement to the IEEE 802.11 standard, the IEEE 802.11 working group published IEEE 802.11a, which outlines the use of OFDM in the 5.8-GHz band.

7.5 Wavelength Division Multiplexing (WDM)

The analog of frequency division multiplexing in the optical domain is known as wavelength division multiplexing. In telecommunications wavelength division multiplexing (WDM) is a technology which multiplexes several optical carrier signals on a single optical fibre by using different wavelengths (colours) of laser light to carry different signals.

Note that this term applies to an optical carrier (which is typically described by its wavelength), whereas frequency division multiplexing typically applies to a radio carrier (which is more often described by frequency). However, since wavelength and frequency are inversely proportional, and since radio and light are both forms of electromagnetic radiation, the distinction is somewhat arbitrary.

The technique relies on the fact that a laser can be designed to emit monochromatic light. Each signal to be transmitted is attached to a laser that emits a different color light beam, all the light beams are sent at the same time and a receiving device splits the colors into the original signals again. The device that joins the signals together is known as a multiplexer, and the one that splits them apart is a demultiplexer. With the right type of fibre you can have a device that does both at once, and can function as an optical add-drop multiplexer and that ought to be called a modem but isn't.

7.5.1 WDM in Practice

In practice, WDM systems are built by combining signals from multiple different single-wavelength end devices onto a single fibre.The first WDM systems combined two signals and appeared around 1985. Modern systems can handle up to 160 signals and can expand a basic 10 Gbit/s fibre system to a theoretical total capacity of over 1.6 Tbit/s over a single fiber pair.

WDM systems are popular with telecommunications companies because they allow them to expand the capacity of their fibre networks without digging up roads again more than necessary which is extremely costly. By using WDM and optical amplifiers, they can accommodate several generations of technology development in their optical infrastructure without having to overhaul the backbone network. All they have to do is to upgrade the multiplexers and demultiplexers at each end.

This is often done by using optical-to-electrical-to-optical translation at the very edge of the transport network, thus permitting interoperation with existing equipment with optical interfaces.

Under WDM, the optical transmission spectrum is carved up into a number of non-overlapping wavelength (or frequency) bands, with each wavelength supporting a single communication channel operating at whatever rate one desires, e.g., peak electronic speed. Thus, by allowing multiple WDM channels to coexist on a single fiber, one can tap into the huge fiber bandwidth, with the corresponding challenges being the design and development of appropriate network architectures, protocols, and algorithms. Also, WDM devices are easier to implement since, generally, all components in a WDM device need to operate only at electronic speed; as a result, several WDM devices are available in the marketplace today, and more are emerging.

Current development activities indicate that this sort of WDM network will be deployed mainly as a backbone network for large regions, e.g., for nationwide or global coverage. End-users—to whom the architecture and operation of the backbone will be transparent except for significantly improved response times—will attach to the network through a wavelength-sensitive switching/routing node. An end-user in this context need not necessarily be terminal equipment, but the aggregate activity from a collection of terminals—including those that may possibly be feeding in from other regional and/or local subnetworks—so that the end-user's aggregate activity on any of its transmitters is close to the peak electronic transmission rate.

Dense and Coarse WDM

Early WDM systems were expensive and complicated to run. However, recent standardization and better understanding of the dynamics of WDM systems have made WDM much cheaper to deploy. The market has segmented into two parts, "dense" and "coarse" WDM.

Dense WDM (DWDM) is generally held to be WDM with more than 8 active wavelengths per fibre and with systems with fewer active wavelengths being classified as coarse WDM (CWDM).

7.5.2 Basic Operation

As explained before, WDM enables the utilization of a significant portion of the available fiber bandwidth by allowing many independent signals to be transmitted simultaneously on one fiber, with each signal located at a different wavelength. Routing and detection of these signals can be accomplished independently, with the wavelength determining the communication path by acting as the signature address of the origin, destination or routing. Components are therefore required that are wavelength selective, allowing for the transmission, recovery, or routing of specific wavelengths.

In a simple WDM system (Figure 7.11), each laser must emit light at a different wavelength, with all the lasers' light multiplexed together onto a single optical fiber. After

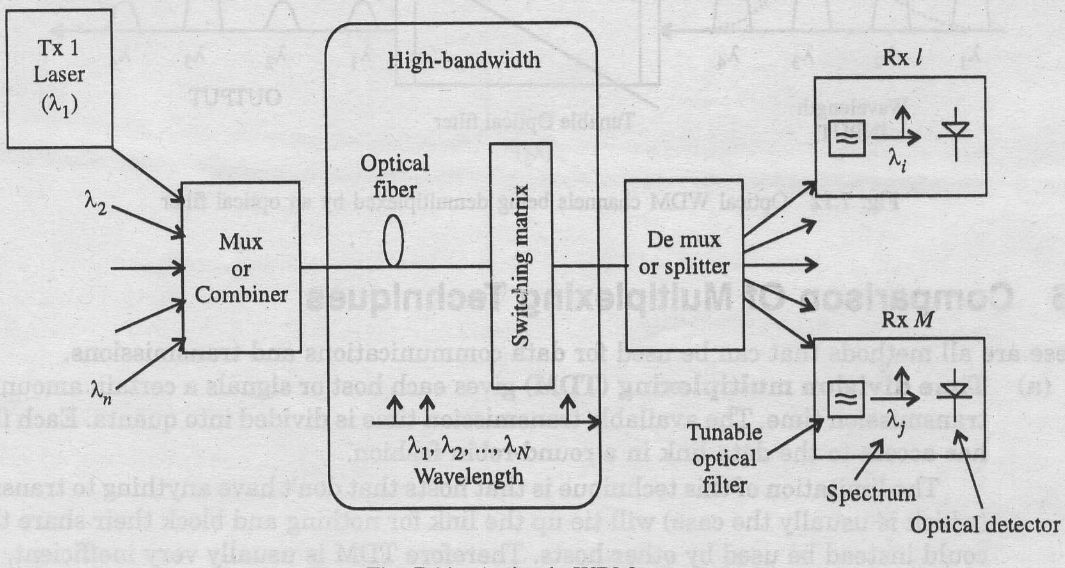


Fig. 7.11 A simple WDM system

being transmitted through a high-bandwidth optical fiber, the combined optical signals must be demultiplexed at the receiving end by distributing the total optical power to each output port and then requiring that each receiver selectively recover only one wavelength by using a tunable optical filter. Each laser is modulated at a given speed, and the total aggregate capacity being transmitted along the high-bandwidth fiber is the sum total of the bit rates of the individual lasers. An example of the system capacity enhancement is the situation in which ten 2.5-Gbps signals can be transmitted on one fiber, producing a system capacity of 25 Gbps. This wavelength-parallelism circumvents the problem of typical optoelectronic devices, which do not have bandwidths exceeding a few gigahertz unless they are exotic and expensive. The speed requirements for the individual optoelectronic components are, therefore, relaxed, even though a significant amount of total fiber bandwidth is still being utilized.

The concept of wavelength demultiplexing using an optical filter is illustrated in Figure 7.12. In the figure, four channels are input to an optical filter that has a nonideal transmission filtering function. The filter transmission peak is centered over the desired channel, in this case thereby transmitting that channel and blocking all other channels. Because of the nonideal filter transmission function, some optical energy of the neighboring channels leaks through the filter, causing interchannel, interwavelength cross-talk. This cross-talk has the effect of reducing the selected signal's contrast ratio and can be minimized by increasing the spectral separation between channels. Although there is no set definition, a nonstandardized convention exists for defining optical WDM as encompassing a system for which the channel spacing is approximately 10 nm.

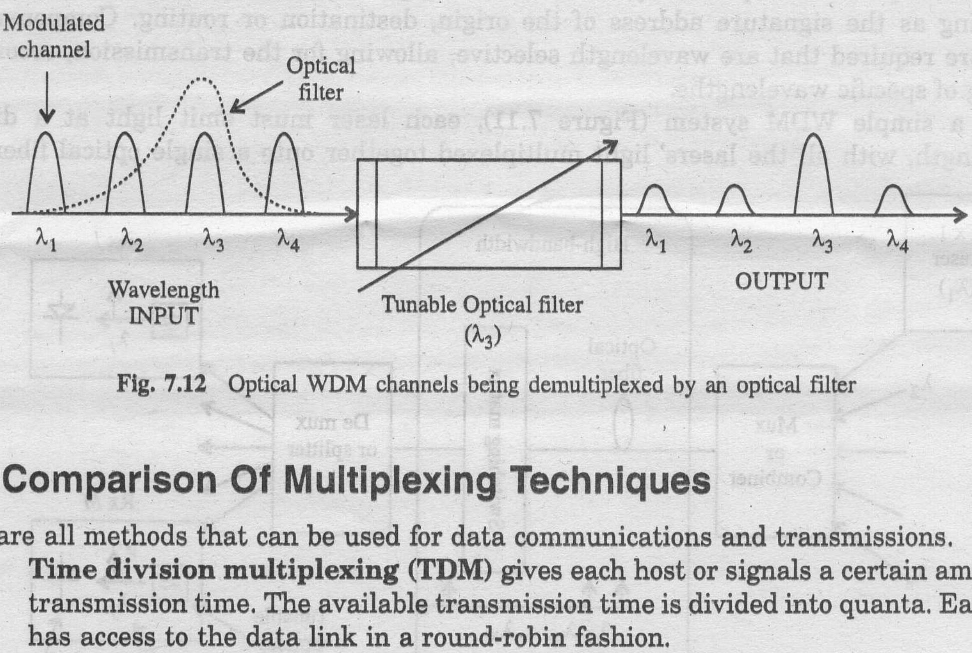


Fig. 7.12 Optical WDM channels being demultiplexed by an optical filter

7.6 Comparison Of Multiplexing Techniques

These are all methods that can be used for data communications and transmissions.

(a) **Time division multiplexing (TDM)** gives each host or signals a certain amount of transmission time. The available transmission time is divided into quanta. Each flow has access to the data link in a round-robin fashion.

The limitation of this technique is that hosts that don't have anything to transmit (which is usually the case) will tie up the link for nothing and block their share that could instead be used by other hosts. Therefore TDM is usually very inefficient.

(b) **Frequency division multiplexing (FDM)** works in a slightly different way. Each flow of data is sent simultaneously over the link at different frequencies (analog TV and radio are examples of FDM). The signals are modulated at the transmission end and demodulated at the receiving end.

The disadvantage is that the number of flows is fixed. Resizing quanta or adding frequencies is not practical.

(c) **Wave division multiplexing (WDM) or wavelength division multiplexing** multiplexes several optical carrier signals onto a single optical fiber. Each signal is transmitted at different wavelengths by using laser light. A multiplexer combines the signal together to be transmitted over the cable, and a demultiplexer separates them once then reaches the other side.

The great advantage that WDM has over the other two technologies is that the bandwidth can be increased without changing the transmission medium or cable. By upgrading the multiplexer and demultiplexer, bandwidth can be increased.

The disadvantage of this system is that the hardware is expensive.

To sum up, TDM is simple but inefficient. FDM allows maximum transmission link usage but cannot evolve easily. WDM is fast and evolutive but is expensive.

• • •

8

Telephone System And ISDN Services

Introduction To PSTN

The Public Switched Telephone Network (PSTN) has been evolving ever since Alexander Graham Bell made the first voice transmission over wire in 1876. This chapter discusses the beginnings of the PSTN and explains why the PSTN exists in its current state.

The oldest and largest telecommunications network in existence is the public switched telephone network (PSTN) that has in excess of 700 million subscribers.

For a long time, the PSTN was the only bearer network available for telephony. Today, many people choose the mobile telephone for their calls. Other bearer networks for voice transmission include integrated service digital network (ISDN), asynchronous transfer mode (ATM), frame relay and the Internet.

The PSTN's primary characteristics are listed below:

- Analog access, 300-3,400 Hz.
- Circuit-switched duplex connection.
- Switched bandwidth, 64 kbit/s, or 300-3,400 Hz for analog exchanges.
- Immobility or, at best, very limited mobility.
- Many functions in common with another bearer network : N-ISDN.

In the course of its long history, which started in 1876, the PSTN has undergone complete technical changes. Even factors such as network structure and network utilization have changed radically. The truly revolutionary changes have come about from the 1960s onwards—innovations such as data communication, telefax, and processor control technique, digital voice transmission, satellite communication, digital switching, optoelectronics, network intelligence structures and the Internet.

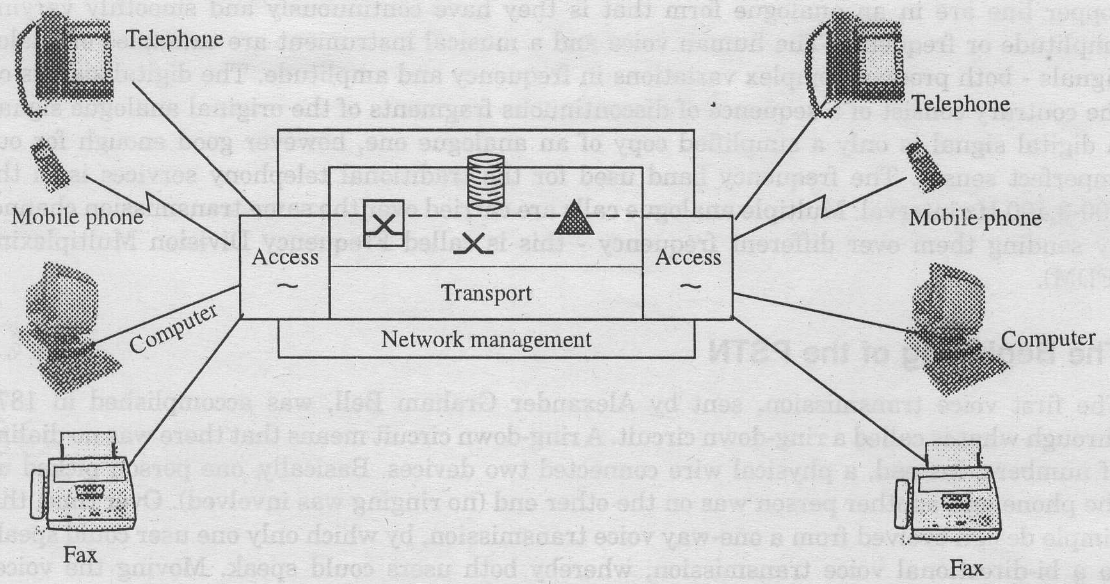


Fig. 8.1 PSTN reference model

To facilitate your understanding of the PSTN and the culture that formed today's telecommunications, we will briefly digress and step back in time to the early 1960s.

The telephone or phone (Greek: tele = far away and phone = voice) is a telecommunications device that transmits sound at great distances. Most telephones operate by means of electric signals. Most modern telephones operate as one part of a complex public switched telephone network of equipment which allows almost any phone user to speak to almost any other. Radiotelephony or wireless telephony transmits messages using radio. Until recently, the word telephone taken alone generally refered to a wired instrument attached to the telephone network. Recently cordless telephones and cell phones have become sufficiently common that no presumption can be made. Wired instruments, or cordless phones attached to wired bases generally provide Plain old telephone service (POTS). Recently Voice over IP has begun to displace POTS in areas where it is offered by Cable Television vendors.

The public switched telephone network (PSTN) is the concatenation of the world's public circuit-switched telephone networks, in much the same way that the Internet is the concatenation of the world's public IP-based packet-switched networks. Originally a network of fixed-line analog telephone systems, the PSTN is now almost entirely digital, and now includes mobile as well as fixed telephones.

The Public Switched Telephone Network (PSTN) is the oldest and largest telecommunications network in existence. PSTN is the telephone network that by default most of the world population is connected to if they have a telephone. In most of the cases a connection to another telephone network, e.g., is done on user request.

The PSTN network is sometimes called an analogue network -- in contrast for example to ISDN and GSM, which are digital networks. The reason is that the signals carried over the copper line are in an analogue form that is they have continuously and smoothly varying amplitude or frequency. The human voice and a musical instrument are examples of analog signals - both produce complex variations in frequency and amplitude. The digital signals on the contrary consist of a sequence of discontinuous fragments of the original analogue signal. A digital signal is only a simplified copy of an analogue one, however good enough for our imperfect senses. The frequency band used for the traditional telephony services is in the 300-3,400 Hz interval. Multiple analogue calls are carried over the same transmission channel by sending them over different frequency - this is called Frequency Division Multiplexing (FDM).

The Beginning of the PSTN

The first voice transmission, sent by Alexander Graham Bell, was accomplished in 1876 through what is called a ring-down circuit. A ring-down circuit means that there was no dialing of numbers; instead, a physical wire connected two devices. Basically, one person picked up the phone and another person was on the other end (no ringing was involved). Over time, this simple design evolved from a one-way voice transmission, by which only one user could speak, to a bi-directional voice transmission, whereby both users could speak. Moving the voices across the wire required a carbon microphone, a battery, an electromagnet, and an iron diaphragm.

It also required a physical cable between each location that the user wanted to call. The concept of dialing a number to reach a destination, however, did not exist at this time. To further illustrate the beginnings of the PSTN, see the basic four-telephone network shown in Figure 8.2. As you can see, a physical cable exists between each location.

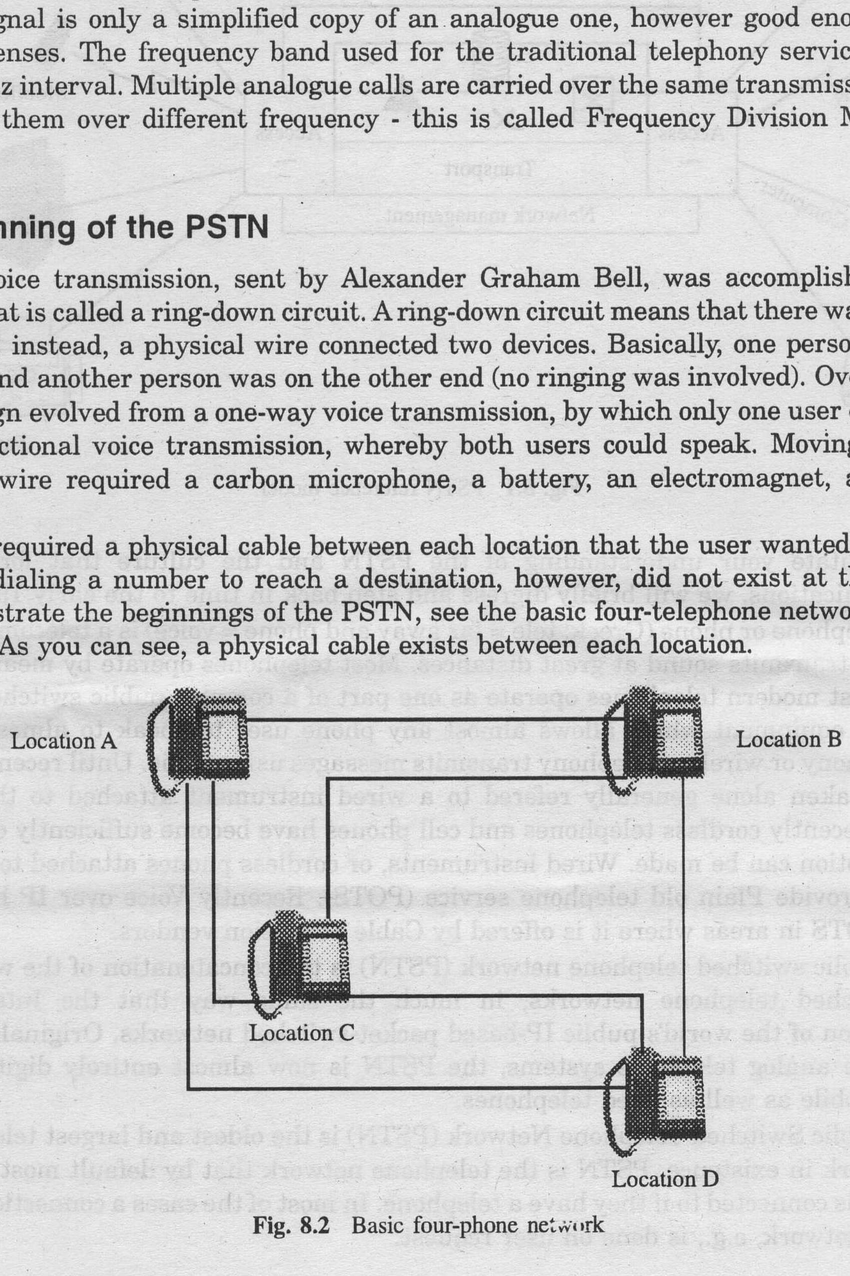


Fig. 8.2 Basic four-phone network

Place a physical cable between every household requiring access to a telephone, however, and you'll see that such a setup is neither cost-effective nor feasible (see Figure 8.3). To determine how many lines you need to your house, think about everyone you call as a value of N and use the following equation: $N \times (N - 1)/2$. As such, if you want to call 10 people, you need 45 pairs of lines running into your house.

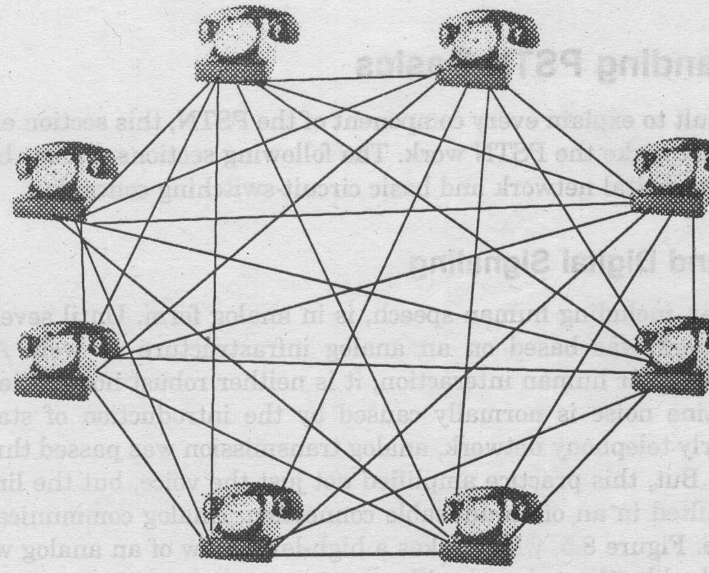


Fig. 8.3 Physical cable between all telephone users

Due to the cost concerns and the impossibility of running a physical cable between everyone on Earth who wanted access to a telephone, another mechanism was developed that could map any phone to another phone. With this device, called a **switch**, the telephone users needed only one cable to the centralized switch office, instead of seven. At first, a telephone

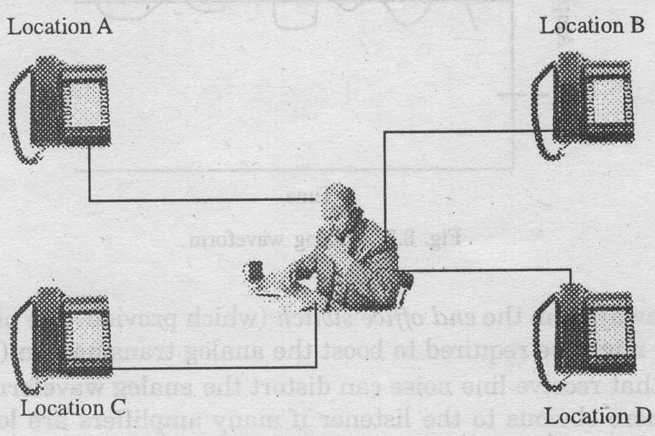


Fig. 8.4 Centralized operator : The human switch

operator acted as the switch. This operator asked callers where they wanted to dial and then manually connected the two voice paths. Figure 8.4 shows how the four-phone network example would look today with a centralized operator to switch the calls.

Now, skip ahead 100 years or so and the human switch is replaced by electronic switches. At this point, you can learn how the modern PSTN network is built.

8.1 Understanding PSTN Basics

Although it is difficult to explain every component of the PSTN, this section explains the most important pieces that make the PSTN work. The following sections discuss how your voice is transmitted across a digital network and basic circuit-switching concepts.

8.1.1 Analog and Digital Signaling

Everything you hear, including human speech, is in analog form. Until several decades ago, the telephony network was based on an analog infrastructure as well. Although analog communication is ideal for human interaction, it is neither robust nor efficient at recovering from line noise. (Line noise is normally caused by the introduction of static into a voice network.) In the early telephony network, analog transmission was passed through amplifiers to boost the signal. But, this practice amplified not just the voice, but the line noise as well. This line noise resulted in an often-unusable connection. Analog communication is a mix of time and amplitude. Figure 8.5, which takes a high-level view of an analog waveform, shows what your voice looks like through an oscilloscope.

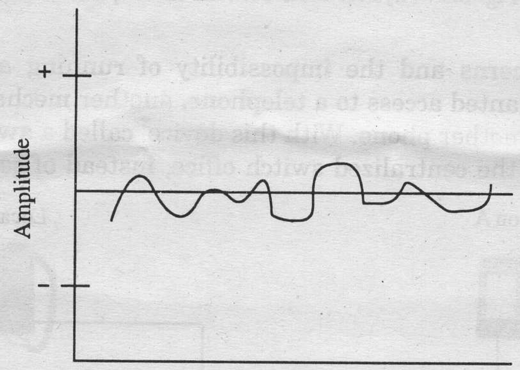


Fig. 8.5 Analog waveform

If you were far away from the *end office switch* (which provides the physical cable to your home), an amplifier might be required to boost the analog transmission (your voice).

Analog signals that receive line noise can distort the analog waveform and cause garbled reception. This is more obvious to the listener if many amplifiers are located between your home and the end office switch. Figure 8.6 show that an amplifier does not clean the signal as it amplifies, but simply amplifies the distorted signal. This process of going through several amplifiers with one voice signal is called accumulated noise.

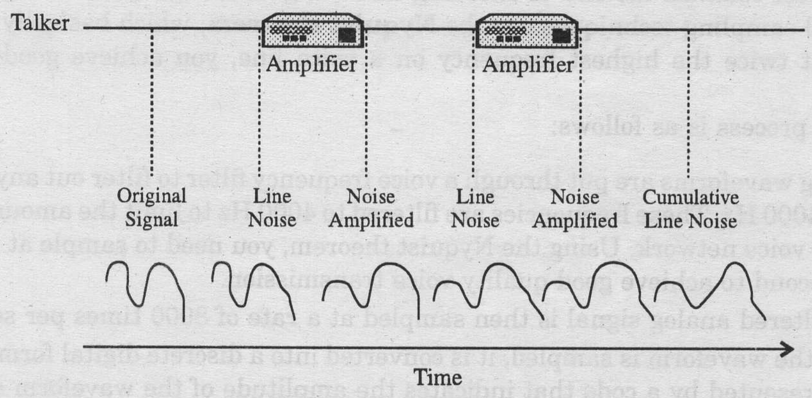


Fig. 8.6 Analog line distortion

In digital networks, line noise is less of an issue because repeaters not only amplify the signal, but also clean it to its original condition. This is possible with digital communication because such communication is based on 1's and 0's. So, as shown in Figure 8.7, the repeater (a digital amplifier) only has to decide whether to regenerate a 1 or a 0.

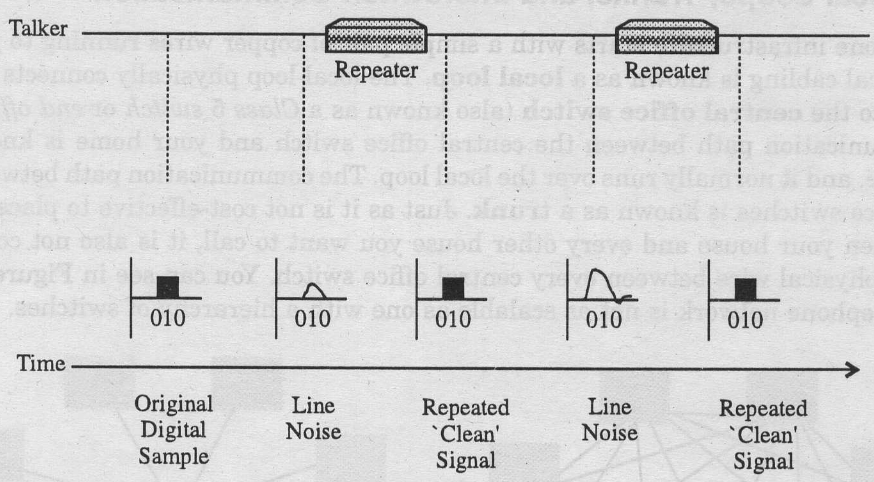


Fig. 8.7 Digital line distortion

Therefore, when signals are repeated, a clean sound is maintained. When the benefits of this digital representation became evident, the telephony network migrated to **Pulse Code Modulation (PCM)**.

Digital Voice Signals

PCM is the most common method of encoding an analog voice signal into a digital stream of 1's and 0's. All sampling techniques use the **Nyquist theorem**, which basically states that if you sample at twice the highest frequency on a voice line, you achieve good-quality voice transmission.

The PCM process is as follows:

- Analog waveforms are put through a voice frequency filter to filter out anything greater than 4000 Hz. These frequencies are filtered to 4000 Hz to limit the amount of crosstalk in the voice network. Using the Nyquist theorem, you need to sample at 8000 samples per second to achieve good-quality voice transmission.
- The filtered analog signal is then sampled at a rate of 8000 times per second.
- After the waveform is sampled, it is converted into a discrete digital form. This sample is represented by a code that indicates the amplitude of the waveform at the instant the sample was taken. The telephony form of PCM uses eight bits for the code and a logarithm compression method assigns more bits to lower-amplitude signals.

If you multiply the eight-bit words by 8000 times per second, you get 64,000 bits per second (bps). The basis for the telephone infrastructure is 64,000 bps (or 64 kbps).

8.1.2 Local Loops, Trunks, and Interswitch Communication

The telephone infrastructure starts with a simple pair of copper wires running to your home. This physical cabling is known as a **local loop**. The local loop physically connects your home telephone to the **central office switch** (also known as a *Class 5 switch* or *end office switch*). The communication path between the central office switch and your home is known as the **phone line**, and it normally runs over the local loop. The communication path between several central office switches is known as a **trunk**. Just as it is not cost-effective to place a physical wire between your house and every other house you want to call, it is also not cost-effective to place a physical wire between every central office switch. You can see in Figure 8.8 that a meshed telephone network is not as scalable as one with a hierarchy of switches.

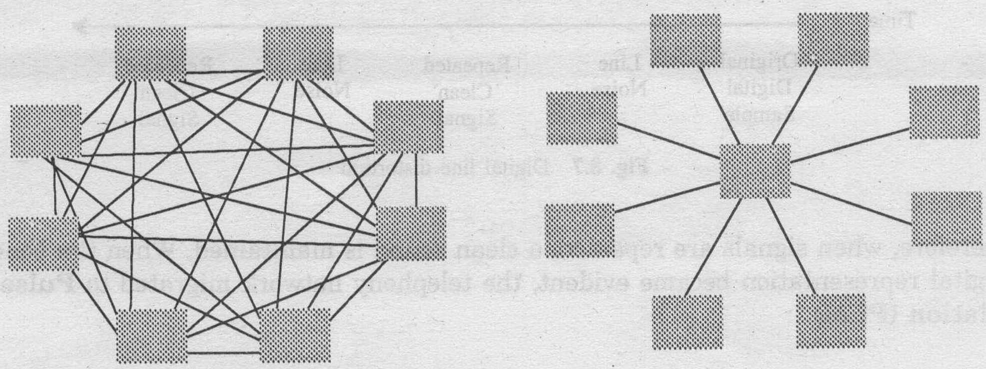


Fig. 8.8 Meshed network versus hierarchical network

Switches are currently deployed in hierarchies. End office switches (or central office switches) interconnect through trunks to **tandem switches** (also referred to as Class 4 switches). Higher layer tandem switches connect local tandem switches. Figure 8.9 shows a typical model of switching hierarchy.

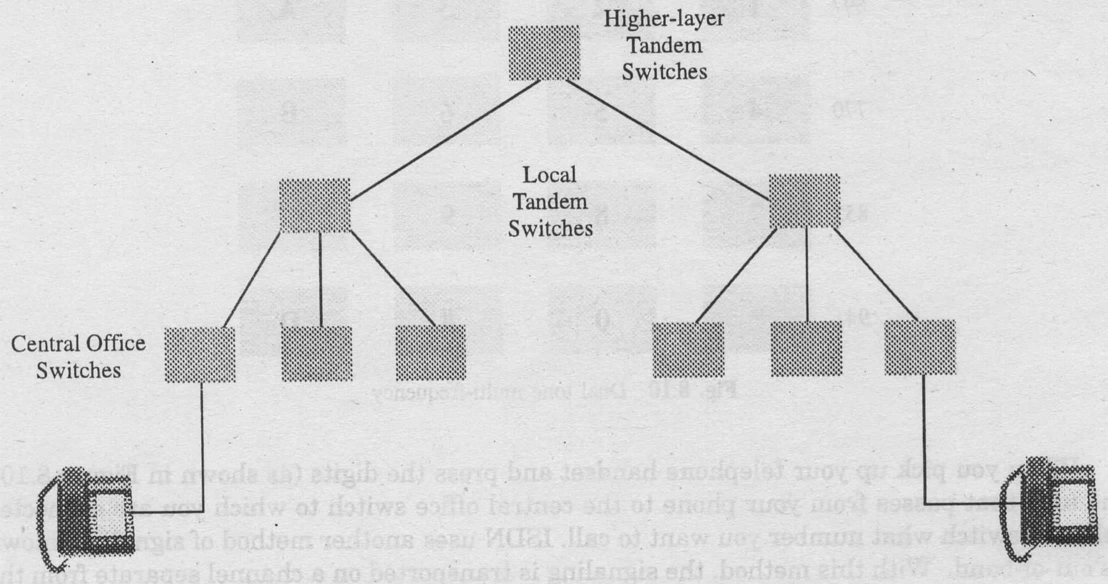


Fig. 8.9 Circuit-switching hierarchy

Central office switches often directly connect to each other. The direct connections that occur between central office switches depends to a great extent on call patterns. If enough traffic occurs between two central office switches, a dedicated circuit is placed between the two switches to offload those calls from the local tandem switches. Some portions of the PSTN use as many as five levels of switching hierarchy.

8.1.3 PSTN Signaling

Generally, two types of signaling methods run over various transmission media. The signaling methods are broken into the following groups:

- **User-to-network signaling :** This is how an end user communicates with the PSTN.
- **Network-to-network signaling :** This is generally how the switches in the PSTN intercommunicate.

User-to-Network Signaling

Generally, when using twisted copper pair as the transport, a user connects to the PSTN through analog, Integrated Services Digital Network (ISDN), or through a T1 carrier. The most common signaling method for user-to-network analog communication is **Dual Tone Multi-**

Frequency (DTMF). DTMF is known as in-band signaling because the tones are carried through the voice path. Figure 8.10 shows how DTMF tones are derived.

	1209	1336	1477	1633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	D

Fig. 8.10 Dual tone multi-frequency

When you pick up your telephone handset and press the digits (as shown in Figure 8.10), the tone that passes from your phone to the central office switch to which you are connected tells the switch what number you want to call. ISDN uses another method of signaling known as out-of-band. With this method, the signaling is transported on a channel separate from the voice. The channel on which the voice is carried is called a bearer (or B channel) and is 64 kbps. The channel on which the signal is carried is called a data channel (D channel) and is 16 kbps.

Out-of-band signaling offers many benefits, including the following:

- Signaling is multiplexed (consolidated) into a common channel.
- Glare is reduced (glare occurs when two people on the same circuit seize opposite ends of that circuit at the same time).
- A lower post dialing delay.
- Additional features, such as higher bandwidth, are realized.
- Because setup messages are not subject to the same line noise as DTMF tones, call completion is greatly increased.

In-band signaling suffers from a few problems, the largest of which is the possibility for lost tones. This occurs when signaling is carried across the voice path and it is a common reason why you can sometimes experience problems remotely accessing your voice mail.

Network-to-Network Signaling

Network-to-network communication is normally carried across the following transmission media:

- T1/E1 carrier over twisted pair.

- T3/E3, T4 carrier over coaxial cable.
- T3, T4 carrier over a microwave link.
- Synchronous Optical Network (SONET) across fiber media.

Network-to-network signaling types include in-band signaling methods such as Multi-Frequency (MF) and Robbed Bit Signaling (RBS). These signaling types can also be used to network signaling methods.

Digital carrier systems (T1, T3) use A and B bits to indicate on/off hook supervision. The A/B bits are set to emulate Single Frequency (SF) tones (SF typically uses the presence or absence of a signal-to-signal A/B bit transitions). These bits might be robbed from the information channel or multiplexed in a common channel (the latter occurs mainly in Europe).

MF is similar to DTMF, but it utilizes a different set of frequencies. As with DTMF, MF tones are sent in-band. But, instead of signaling from a home to an end office switch, MF signals from switch to switch.

Network-to-network signaling also uses an out-of-band signaling method known as Signaling System 7 (SS7) (or C7 in European countries). SS7 is beneficial because it is an out-of-band signaling method and it interconnects to the Intelligent Network (IN). Connection to the IN enables the PSTN to offer Custom Local Area Signaling Services (CLASS) services. SS7 is a method of sending messages between switches for basic call control and for CLASS. These CLASS services still rely on the end-office switches and the SS7 network. SS7 is also used to connect switches and databases for network-based services.

8.1.4 PSTN Services

The two basic services supported over PSTN are Voice and Data. They have different requirements from the point of view of the quality of service that is qualified as "acceptable". For example in the case of voice, some degradation of quality is allowed to level when the speech could be still understood and some noise is tolerable; at the same time delay is not tolerable, we do not want to get the words one by one with variable intervals. The case with data is to the contrary-delay is tolerable—for example, we do not mind if we see the Web site we are viewing coming in portions and during a variable interval of time (as far as we do not get bored), however "noise" is not acceptable because even little noise during data transmission would mean a strange looking web site (if we could understand that this is indeed a web site at all). This special requirement and the nature of the PSTN network itself have made the PSTN very slow when delivering data. The only way to overcome this is to introduce compression techniques that make the data smaller before it is sent over the PSTN.

The introduction of computer-controlled telephone exchanges during the 70s allowed for operators to create new, completely different subscriber services in the network. Today it is much easier for an operator to create a distinctive image for himself and to increase revenue. Some services are charged for, others are free. Many of the free services help to increase the number of successful calls, which in turn generates revenue. These services are normally known under the name Supplementary Services and Value added services.

Different PSTN operators in different countries have chosen to implement different PSTN services which means that not all possible PSTN services could be find in different countries and even within one and the same country depending on the PSTN operator's choice. The following list shows just few of the services that you may have on offer from your operator:

- **Calling line identification presentation (CLIP) :** Also called "caller ID", this service allows a called party to see the telephone number of an incoming call on a display connected to the telephone line. There are two commonly accepted ways of transmitting the CLIP information-via DTMF or Phase-Shift Key (PSK) signalling.
- **Call forwarding :** This service re-routes incoming calls to another number.
- **Call-back (completion of call to a busy subscriber) :** If the called subscriber is busy, the caller can order the call-back service, which means that he is queued for connection to the busy number and when that subscriber gets free the network will connect and notify the caller.
- **Call waiting :** A special signal is generated during a call in progress to indicate that a third party is trying to reach you.

To access and control those services via the telephone, service codes are used (for example the service code 21 is assigned to "call forwarding unconditional"). The subscriber usually activates these services by lifting the handset and then pressing a "star-digits-hash" combination. Also, the register button (R) is used, for example, when alternating between calls and for three-party calls. The system usually confirms the execution of a service by sending a tone or a recorded message.

In addition to offering you a pure voice service, a telephone can be an extremely efficient tool for a subscriber who wishes to send a message, make a financial transaction, or just send a short message (SMS) to another telephone. The PSTN treats such services as any other telephone connection: It sets up the connection to a proper destination after the subscriber has dialled the number of the service. When the call is answered, direct contact with the destination, e.g. a computer or a message gateway is established. The keypad telephone's DTMF signalling is then the tool mostly used for transferring information from the subscriber over the established connection.

DTMF, however, is not capable to provide in all cases all the means needed - for example, if a SMS message is to be sent/received it would require a full set of numbers and letters to be available which DTMF does not provide - instead the already mentioned Phase-Shift Key (PSK) signalling is used.

Of course when using a DECT (Digital Enhanced Cordless Telecommunications) phone and some corded PSTN phones all these may be well hidden from you behind a user friendly menu provided to you by the handset. On the handset display you may see the name of the service and what you may need is pressing a single key to activate it—the telephone itself then will identify the "star-digits-hash" combination needed and will send it to the network.

Connecting Digital devices to PSTN DECT phones, Faxes and Modems are some examples of digital user terminals that can be connected to a PSTN network. In order such devices to operate the digital signal needs to be converted to an analogue signal or as it is normally called "to be modulated". This is needed as well on the Transport network if it is a digital one, e.g., ISDN, ATM, IP. Different modulation techniques may be used. For many decades, A/D conversion of voice has been performed employing the so-called PCM coding. This is pure amplitude coding that results in a bi-directional connection having a bit rate of 64 kbit/s.

8.1.5 Drawbacks to the PSTN

Although the PSTN is effective and does a good job at what it was built to do (that is, switch voice calls), many business drivers are striving to change it to a new network, whereby voice is an application on top of a data network. This is happening for several reasons:

- Data has overtaken voice as the primary traffic on many networks built for voice. Data is now running on top of networks that were built to carry voice efficiently. Data has different characteristics, however, such as a variable use of bandwidth and a need for higher bandwidth. Soon, voice networks will run on top of networks built with a data-centric approach. Traffic will then be differentiated based upon application instead of physical circuits. New technologies (such as Fast Ethernet, Gigabit Ethernet, and Optical Networking) will be used to deploy the high-speed networks that needed to carry all this additional data.
- The PSTN cannot create and deploy features quickly enough. With increased competition due to deregulation in many telecommunications markets, LECs are looking for ways to keep their existing customers. The primary method of keeping customers is by enticing them through new services and applications. The PSTN is built on an infrastructure whereby only the vendors of the equipment develop the applications for that equipment. This means you have one-stop shopping for all your needs. It is very difficult for one company to meet all the needs of a customer. A more open infrastructure, by which many vendors can provide applications, enables more creative solutions and applications to be developed. It is also not possible with the current architecture to enable many vendors to write new applications for the PSTN. Imagine where the world would be today if vendors, such as Microsoft, did not want other vendors to write applications for its software.
- Data/Voice/Video (D/V/V) cannot converge on the PSTN as currently built. With only an analog line to most homes, you cannot have data access (Internet access), phone access, and video access across one 56-kbps modem. High-speed broadband access, such as digital subscriber line (DSL), cable, or wireless, is needed to enable this convergence. After the last bandwidth issues are resolved, the convergence can happen to the home. In the backbone of the PSTN, the convergence has already started.
- The architecture built for voice is not flexible enough to carry data. Because the bearer channels (B channels and T1 circuits), call-control (SS7), and service logic (applications) are tightly bound in one closed platform, it is not possible to make minor changes that might improve audio quality.

It is also important to note that circuit-switched calls require a permanent 64-kbps dedicated circuit between the two telephones. Whether the caller or the person called is talking, the 64-kbps connection cannot be used by any other party. This means that the telephone company cannot use this bandwidth for any other purpose and must bill the parties for consuming its resources. Data networking, on the other hand, has the capability to use bandwidth only when it is required. This difference, although seemingly small, is a major benefit of packet-based voice networking.

8.2 Introduction To ISDN

ISDN (Integrated Services Digital Network) is an all-digital communications network designed to bring the power of the digital network directly to the desktop. It is the ideal communications technology for the information age-perfect for students, cyberspacers, travelers, telecommuters, parents, business people, and corporate executives.

ISDN is often referred as the successor of the existing public networks like the telephone network, data and text networks (Telegraphy). The telephone network was first an analog network. In the early days telephony switching was made by the telephone operator, by plugging both ends of a wire in the right holes of a connection board. In those times the phones had no dialing mechanism. The user had to pick up the phone and tell the operator which connection he desires. After the invention of the dialing disk, switching could be automated, but fast connection set up was possible only after a partially digitization of the telephone net. Today the biggest part of the Telephone system is digital. Only the line from the user to the local exchange is analog. For the connection between two exchange points, digital lines are used (IDN—Integrated Digital Network). A continuation of this development leads to an overall digital network. ISDN is such a network, providing moreover services available in the past—only in other networks and/or with other interfaces, e.g. Teletex, Telefax. ISDN is built on top of IDN. It allows communication, not only with older equipment, for instance an analog telephone connected analog to the local exchange, but also with other networks, e.g., PSPDN (Public Switched Packet Data Network). Thus ISDN can be seen as an evolution rather than a revolution.

In simple terms, ISDN is a replacement for plain old telephone service, which was never designed to meet the needs of the information age. ISDN uses the same wiring that currently serves homes and businesses. You get ISDN service from the same companies who provide telephone service, and you use it to connect telephones, computers, and fax machines. The difference is that you get much faster, much more dependable connections for voice, data, fax, and even video—all through a single line. There is no other technology that comes close to delivering such communications benefits today.

Standard telephone service requires a separate phone line for each device to be used simultaneously. Not only can multiple lines be expensive, but also the amount of information that can be transmitted is limited with analog service; current technology allows 56 kilobits per second (Kbps). ISDN, however, provides multiple channels to operate concurrently on the same pair of wires, and each channel is capable of transmitting at 64 Kbps. Additionally, digital transmissions allow for reduced noise and interference on the carrier channels.

ISDN provides services and capabilities not available through standard telephone service. ISDN furnishes these services through a digital package when a call is initiated. The digital packet includes information regarding : (1) who is calling, (2) the type of call (data/voice/etc...), and (3) the number dialed, if more than one number is used for a single ISDN line. With the information provided, ISDN equipment can determine how to handle a call, based on user-defined preferences. Calls can be accepted, rejected or even rerouted. Data calls can even be routed to an Internet Protocol (IP) address.

Meaning of "ISDN"

"Integrated Services" refers to ISDN's ability to deliver two simultaneous connections, in any combination of data, voice, video, and fax, over a single line. Multiple devices can be attached to the line, and used as needed. That means an ISDN line can take care of most people's complete communications needs, without forcing the purchase of multiple analog phone lines at a much higher transmission rate.

The "Digital" in ISDN refers to its purely digital transmission, as opposed to the analog transmission of plain old telephone service. If you're using a modem for Internet access at this moment, your Internet service provider's modem has converted this site's digital content to analog signals before sending it to you, and your modem converts those signals back to digital when receiving (the same thing happens with every keystroke and mouse click you transmit). When you connect with ISDN, there is no analog conversion. ISDN transmits data digitally, resulting in a very clear transmission quality. There is none of the static and noise of analog transmissions that can slow transmission speed.

"Network" refers to the fact that ISDN is not simply a point-to-point solution like a leased line. ISDN networks extend from the local telephone exchange to the remote user and include all of the telecommunications and switching equipment in between.

Main characteristics

The probably most important keywords concerning ISDN are:

- End to end digital connection.
- Integration of multiple services (voice-, data-, video-, multimedia transmission).
- Standard terminal interface.

Access to ISDN

Two types of ISDN have been specified:

- **Narrowband-ISDN (NISDN).**
- **Broadband-ISDN (BISDN).**

The main difference between NISDN and BISDN is the transmission capacity and the used transfer mode. NISDN can serve with a capacity of up to 2Mbps, while for BISDN exist specifications for 150Mbps and 600Mbps. The term ISDN will basically refer to NISDN. BISDN will be discussed later on.

8.2.1 The Basics of ISDN

ISDN is based on a number of fundamental building blocks.

Channels

A channel is a connection between two users over the network, with a defined capacity of bps. A certain number of channels have been designed for the use of ISDN. There are two types of ISDN "channels" or communication paths:

- **B-channel**
  The Bearer ("B") channel is a 64 kbps channel that can be used for voice, video, data, or multimedia calls. B-channels can be aggregated together for even higher bandwidth applications.
- **D-channel**
  The Delta ("D") channel can be either a 16 kbps or 64 kbps channel used primarily for communications (or signaling) between switching equipment in the ISDN network and the ISDN equipment at your site.

The B channels are primarily for voice and data transmissions, while the D channel is used for signaling and packet transmissions. These ISDN channels are delivered to the user in one of two pre-defined configurations:

- **Basic Rate Interface (BRI)**
  BRI is the ISDN service most people use to connect to the Internet. An ISDN BRI connection supports two 64 kbps B-channels and one 16 kbps *D*-channel over a standard phone line. BRI is often called **"$2B + D$"** referring to its two *B*-channels and one *D*-channel.
- **Primary Rate Interface (PRI)**
  ISDN PRI service is used primarily by large organizations with intensive communications needs. An ISDN PRI connection supports 23 64 kbps B-channels and one 64 kbps *D*-channel (or **$23B + D$**) over a high speed DS1 (or $T-1$) circuit. The European PRI configuration is slightly different, supporting $30B + D$.

BRI is the most common ISDN service for Internet access. A single BRI line can support up to three calls at the same time because it is comprised of three channels ($2B + D$). Two voice, fax or data conversations, and one packet switched data conversation can take place at the same time. Multiple channels or even multiple BRI lines can be combined into a single faster connection depending on the ISDN equipment you have. Channels can be combined as needed for a specific application (a large multimedia file transfer, for example), then broken down and reassembled into individual channels for different applications (normal voice or data transmissions).

BRI Applications

- ☐ Voice Transmissions
- ☐ Interactive Data Exchange
- ☐ LAN-to-LAN connections

- ☐ Bulk data transfer
- ☐ Interactive Video Conferencing
- ☐ Telecommuting
- ☐ Application shared devices and imagingElectronic Funds Transfer
- ☐ Electronic mail
- ☐ Facsimile
- ☐ Graphics
- ☐ Printing
- ☐ Retrieving or sending information from multiple databases

8.2.2 Use of ISDN

ISDN offers the speed and quality that previously was only available to people who bought expensive, point-to-point digital leased lines. Combined with its flexibility as a dial-up service, ISDN has become the service of choice for many communications applications. Popular ISDN applications include:

- Internet access.
- Telecommuting/remote access to corporate computing.
- Video conferencing.
- Small and home office data networking.

Why Should I Use ISDN to Access the Internet ?

More and more people are discovering that ISDN is the right Internet answer. As the Internet becomes more and more information-intensive with graphics, sound, video and multimedia, your ability to take advantage of these new resources depends on the speed of your Internet connection. With ISDN, your Internet access is:

- **Even faster :** By combining your two B-channels you have access to up to 128 kbps —more than four times as fast as a 28.8 kbps modem on a standard phone line. And ISDN's digital technology assures you the cleanest connection to the Internet so you won't be slowed down by re-transmissions because of old analog technology.
- **More efficient and economical :** ISDN brings increased capabilities, reduced costs and improved productivity to organizations both large and small. When you're looking for something on the Internet, you can get there faster. You can be more productive because you aren't waiting as long to get to that next website or download that large file.

ISDN Services

Two different groups of services are supported by ISDN: The bearer services and the teleservices. (See figure 8.11).

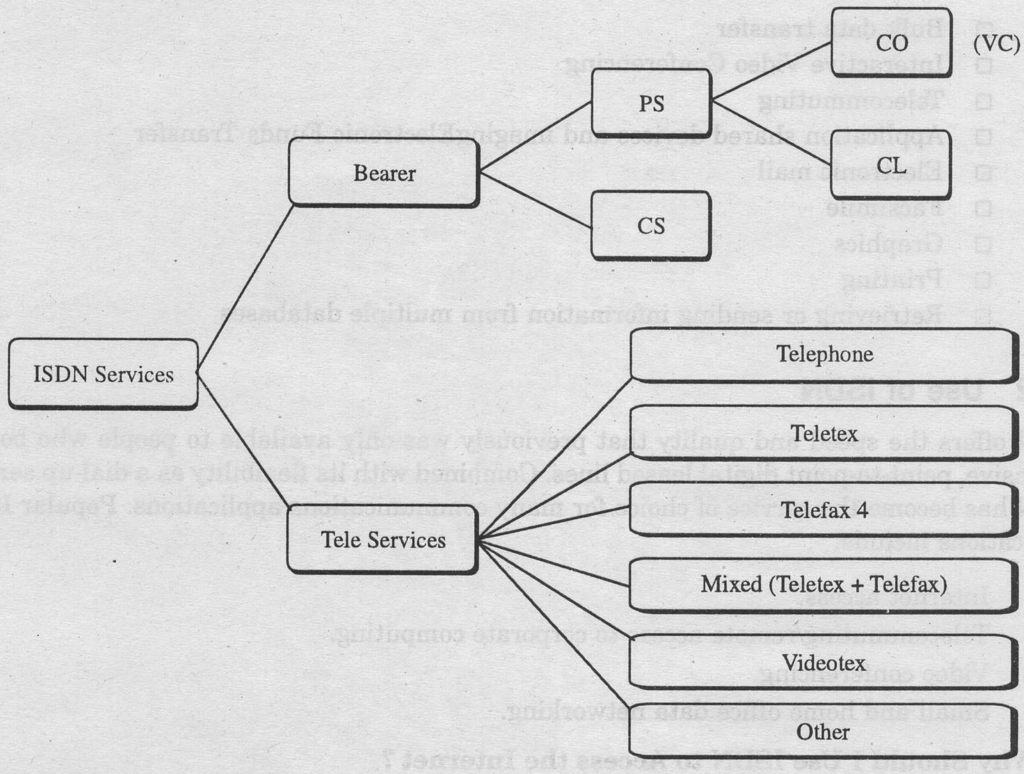


Fig. 8.11 Services of ISDN

Bearer Services

With ISDNs bearer services it is possible to transfer data between two subscribers. The network is acting in this case as a bit pipe. The bearer service can be divided into packet switching (PS), with both connection orientated (CO) and connectionless (CL) modes, and circuit switched connections.

ISDN Teleservices

- **Telephony :** This service enables subscribers to make phone calls using the ISDN. It is a considerable improvement to the (partially) analog telephone-system, as it offers a better signal-noise ratio and the attenuation is unaffected by distance.
- **Teletex :** Text transmission over ISDN is faster than the conventional teletex service over public data networks. It is possible to transmit a page (legal size or OSI A4) in less than 1 second. The connection with tele text terminals in other networks is supported. This includes circuit-switched as well as packet-switched networks.
- **Telefax (telefax 4) :** ISDN supports the pixel-orientated transmission of documents. CCITT has made recommendations for telefax4 and group 4 facsimile machines with a resolution of 300dpi and optionally 400dpi or 1200dpi. One page of legal size (or OSI A4) paper can be transmitted in 15 seconds when using the 400dpi resolution.

- **Still image transfer :** This service offers the possibility of transmitting TV freeze frames over ISDN. A still image sequence can be compressed and send with an update rate between 1 and 10 seconds, depending on the contents of the image.
- **Videophony :** Videophony is used to transmit moving pictures from person to person or person to group. Because of the relatively low transmission rate (64kbps or 2x64kbps) the quality of the video frames are inferior to TV frames (480x240 Pixel) and must be transmitted at a very high compression.
- **Alarm Services :** It is possible to make emergency calls over ISDN, even if the power supply on the user side breaks down and the *B*-channels cannot be used any more. An emergency call is made over the *D*-channel, which is then feeded by the provider.
- **Messaging :** This service provides a mailbox function for the user. Text and voice mail can be deposited in the mailbox, if the user is not available or both of his *B*-channels are busy. The recipient is sent a message from his mailbox over the *D*-Channel. He receives the notification even if his *B*-channels are used at that time.
- **Videotex :** Videotex is a retrieval service that enables the subscriber to view text and graphics based images.
- **Other services :** Some other services are: Telewriting for placing short messages (written with an electronic pen on note pad) in a mailbox. Teleaction is used for controlling installations like heating, gas, and water. Distribution services allow data transmission in a unidirectional way-but this service is much more interesting with the higher transmission rates of BISDN (video and TV transmission).

Benefits of ISDN

While ISDN accommodates telephones and fax machines, its most popular advantage is in computer applications. You can plug an ISDN adapter into a phone jack, like you would an analog modem, and get a much faster connection with no "line noise".

The most common ISDN service, Basic Rate Interface (BRI), provides two 64 Kbps channels per line. When the two channels are bonded in a single connection, you get a speed of 128 Kbps, which is about four times the actual top speed of the fastest analog modems. Compression can increase throughput to around 250 Kbps. ISDN affords many benefits to service providers and customers. The increasing popularity of ISDN allows pricing that continues to fall and compete with standard analog service. Some of the many benefits are:

Internet access is a great application for ISDN. Compared with even the fastest modem access, ISDN makes Web graphics appear almost immediately, and can reduce download times by over 75%. In many markets, it's actually cheaper than an isolated analog business line. Simultaneous audio, video, and data services over a single pair of copper wires reduces infrastructure and maintenance costs for service and subscribers.

ISDN BRI service can use data compression that boosts the 128 Kbps transmission rate to between 256 Kbps and 632 Kbps, depending upon the compression ratio used.

Digital transmissions produce clearer and quieter voice telephone service and more reliable and accurate connectivity than analog technology. Remote computer users benefit from high performance ISDN connections at home or on the road.

The various advantages of ISDN are discussed below:

1. Speed

The modem was a big breakthrough in computer communications. It allowed computers to communicate by converting their digital information into an analog signal to travel through the public phone network. There is an upper limit to the amount of information that an analog telephone line can hold. Currently, it is about 56 kb/s bidirectional. Commonly available modems have a maximum speed of 56 kb/s, but are limited by the quality of the analog connection and routinely go about 45–50 kb/s. Some phone lines do not support 56 kb/s connections at all.

ISDN allows multiple digital channels to be operated simultaneously through the same regular phone wiring used for analog lines. The change comes about when the telephone company's switches can support digital connections. Therefore, the same physical wiring can be used, but a digital signal, instead of an analog signal, is transmitted across the line. This scheme permits a much higher data transfer rate than analog lines. BRI ISDN supports an uncompressed data transfer speed of 128 kb/s, plus bandwidth for overhead and signaling. In addition, the latency, or the amount of time it takes for a communication to begin, on an ISDN line is typically about half that of an analog line. This improves response for interactive applications, such as games.

2. Multiple Devices

Previously, it was necessary to have a separate phone line for each device you wished to use simultaneously. For example, one line each was required for a telephone, fax, computer, bridge/router, and live video conference system. Transferring a file to someone while talking on the phone or seeing their live picture on a video screen would require several potentially expensive phone lines.

ISDN allows multiple devices to share a single line. It is possible to combine many different digital data sources and have the information routed to the proper destination. Since the line is digital, it is easier to keep the noise and interference out while combining these signals. ISDN technically refers to a specific set of digital services provided through a single, standard interface. Without ISDN, distinct interfaces are required instead.

3. Signaling

Instead of the phone company sending a ring voltage signal to ring the bell in your phone ("In-Band signal"), it sends a digital packet on a separate channel ("Out-of-Band signal"). The Out-of-Band signal does not disturb established connections, no bandwidth is taken from the data channels, and call setup time is very fast. For example, a V.90 or V.92 modem typically takes 30–60 seconds to establish a connection; an ISDN call setup usually takes less than 2 seconds.

The signaling also indicates who is calling, what type of call it is (data/voice), and what number was dialed. Available ISDN phone equipment is then capable of making intelligent decisions on how to direct the call.

Applications Of ISDN

ISDN in Business

For business users and even residential subscribers, videoconferencing is the biggest communication advancement that ISDN has to offer. With the simultaneous high-speed transfer of voice and video, ISDN can provide real time video communication on a PC that once was only capable on sophisticated systems costing upwards of $100,000.

A shared electronic chalkboard is another tool available through ISDN. Ideas and illustrations can be distributed in real time to remote locations so people in other cities or other countries can participate in meetings.

Telecommuting is becoming a rule more than an exception; more and more people are working from home. ISDN provides the facilities for users to tap into central network resources from the privacy of their own homes and do so with the functionality of a network node. Node connections are possible with Serial Line Interface Protocol (SLIP) and Point-to-Point Protocol (PPP).

ISDN in Education

Students will also reap the benefits of videoconferencing by relating with other students worldwide. Using the video capabilities of ISDN allows students to see the surroundings of other countries or speak with pen-pals. The value of videoconferencing in educational settings is unlimited.

Computers have become important learning tools for students. Children are introduced to computers and networking at an early age, and ISDN allows the high-speed connections to vast amounts of information and resources.

8.2.3 The Future of ISDN

There is and for years has been a global push to become a digital world. The main evidence of the digital age is the internetworking of computers, providing access to a wealth of resources and information.

ISDN technology, although plagued by political issues, is a cost-effective means to connect communication equipment of all types. As more of the compatibility issues are resolved, ISDN services will increasingly become the service of choice for businesses and residential customers. The rise in popularity should drive down equipment costs and make connection devices readily available off-the-shelf.

Although other digital services are competitive, most require expensive dedicated lines or possess inherent shortcomings like slow speed. ISDN technology is standardized and provides dynamic high-speed end-to-end digital connectivity over the existing worldwide telephone network.

ISDN continues to evolve with broadband ISDN availability in the near future. With ATM-based technology providing B-ISDN transmission speeds up to 600 Mbps or more, one would be hard-pressed to find a viable affordable alternative.

8.3 Broadband ISDN

The original specifications for the integrated services digital network (ISDN) were based around voice and non-voice telephone-type services: telephony, data, telex, facsimile, as it was hoped that the ISDN would evolve from the (then) emerging digital telephone networks. Indeed, this is one of the reasons that the fundamental element of an ISDN link is the 64Kb/s B-Channel. However, the planning for ISDN was started around 1976, and as technology evolved, so did the requirements of the users that wanted to use this technology. In 1988, the CCITT released a document that described a new set of **Broadband ISDN (B-ISDN)** services. To distinguish this new concept from the original ISDN service, we refer to the latter as **Narrowband ISDN (N-ISDN)**.

Since then the CCITT has become the ITU. The B-ISDN work is far from complete, and some of the factors influencing the emergence of the B-ISDN from ITU are:

- **Demand** : Users (both commercial and residential) are showing interest in receiving high speed, reliable services.
- **Technology** : This has been one of the biggest factors. Advances in technology have increased demand, as well as the ability to supply it. As data processing technologies available to the user have become more sophisticated so has his/her demands, while high-speed transmission (based on the use of optical fibre), high speed switching and increased processing power make the realization of these demands possible.
- **Service integration** : There is a need to integrate both circuit switched and packet switched services into one network that can provide interactive and distribution services.
- **Flexibility** : The resulting network must be able to satisfy the needs of the users as well as the network operators in terms of its functionality and usability.

Broadband ISDN services

The need for a Broadband ISDN service sprung from the growing needs of the customers. The planned Broadband ISDN services can broadly be categorized as follows:

- **Interactive services** : These are services allowing information flow between two end users of the network, or between the user and the service provider. Such services can be subdivided:

□ **Conversational services** : These are basically end-to-end, real-time communications, between users or between a user and a service provider, e.g. telephone-like services. Indeed, B-ISDN will support N-ISDN type services. (Note also that the user-to-user signaling, user-to-network signaling, and inter-exchange signaling are also provided) Also the additional bandwidth offered will allow such services as video telephony, video conferencing and high volume, high-speed data transfer.

□ **Messaging services** : This differs from conversational services in that it is mainly a store-and-forward type of service. Applications could include voice and video mail, as well as multi-media mail and traditional electronic mail.

- **Retrieval services :** This service provides access to (public) information stores, and information is sent to the user on demand only. This includes things like tele shopping, videotex services, still and moving pictures, tele software and entertainment.

- **Distribution services :** These are mainly broadcast services and are intended for mainly one way interaction from a service provider to a user:
  - **No user control of presentation :** This would be for instance, a TV broadcast, where the user can choose simply either to view or not. It is expected that cable TV companies will become interested in Broadband ISDN as a carrier for the high definition TV (HDTV) services that are foreseen for the future.
  - **User controlled presentation :** This would apply to broadcast information that the user can partially control, in that the user can decide which part of it he/she accesses, e.g., teletext and news retrieval services.

So we can say that the **B-ISDN needs to provide:**

- Broadband services.
- Narrowband services (for backwards compatibility).
- User-to-network signaling, to allow the B-ISDN user to initiate and control communication.
- Inter-exchange signaling within the network, to allow the network to provide and control resources as requested by the B-ISDN user or by another network exchange.
- User-to-user signaling, to allow B-ISDN users to send control, operation and maintenance information to each other.
- Management facilities for controlling and operating the network.

8.4 Asynchronous Transfer Mode (ATM)

To provide the new B-ISDN services, use of a technology called asynchronous transfer mode (ATM) is specified by ITU. Asynchronous Transfer Mode (ATM) has been accepted universally as the transfer mode of choice for Broadband Integrated Services Digital Networks (BISDN). ATM can handle any kind of information i.e. voice, data, image, text and video in an integrated manner. ATM provides good bandwidth flexibility and can be used efficiently from desktop computers to local area and wide area networks. ATM is a connection-oriented packet switching technique in which all packets are of fixed length i.e. 53 bytes (5 bytes for header and 48 bytes for information). No processing like error control is done on the information field of ATM cells inside the network and it is carried transparently in the network.

ATM meets the following objectives for BISDN networks:

1. Supports all existing services as well as emerging services in the future.
2. Utilizes network resources very efficiently.

3. Minimizes the switching complexity.
4. Minimizes the processing time at the intermediate nodes and supports very high transmission speeds.
5. Minimizes the number of buffers required at the intermediate nodes to bound the delay and the complexity of buffer management.
6. Guarantees performance requirements of existing and emerging applications.

8.4.1 Overview of ATM

Changes in the structure of the telecommunications industry and market conditions have brought new opportunities and challenges for network operators and public service providers. Networks that have been primarily focused on providing better voice services are evolving to meet new multimedia communications challenges and competitive pressures. Services based on asynchronous transfer mode (ATM) and synchronous digital hierarchy (SDH)/synchronous optical network (SONET) architectures provide the flexible infrastructure essential for success in this evolving market.

ATM, which was once envisioned as the technology of future public networks, is now a reality, with service providers around the world introducing and rolling out ATM and ATM-based services. The ability to exploit the benefits of ATM technology within the public network successfully will provide strategic competitive advantage to carriers and enterprises alike.

Definition

Asynchronous transfer mode (ATM) is a high-performance, cell-oriented switching and multiplexing technology that utilizes fixed-length packets to carry different types of traffic. ATM is a technology that will enable carriers to capitalize on a number of revenue opportunities through multiple ATM classes of services; high-speed local-area network (LAN) interconnection; voice, video, and future multimedia applications in business markets in the short term; and in community and residential markets in the longer term.

Asynchronous transfer mode (ATM) is a technology that has its history in the development of broadband ISDN in the 1970s and 1980s. Technically, it can be viewed as an evolution of packet switching. Like packet switching for data (e.g., X.25, frame relay), ATM integrates the multiplexing and switching functions, which is well suited for bursty traffic (in contrast to circuit switching), and allows communications between devices that operate at different speeds. Unlike packet switching, ATM is designed for high-performance multimedia networking. ATM technology has been implemented in a very broad range of networking devices:

- PC, workstation, and server network interface cards
- Switched-Ethernet and token-ring workgroup hubs
- Workgroup and campus ATM switches

- ATM enterprise network switches
- ATM multiplexers
- ATM—edge switches
- ATM—backbone switches

The term **asynchronous transfer mode** needs some explanation. The words **transfer mode** say that this technology is specific way of transmitting and switching through the network. The term **asynchronous** refers to the fact that the packets are transmitted using asynchronous techniques, and the two end-points need not have synchronized clocks. Also, the use and allocation of cells and their subsequent multiplexing and transmission through the network is determined in an asynchronous fashion, e.g. on demand, and is independent of the user. ATM will support both circuit switched and packet switched (sometimes referred to as **circuit mode** and **packet mode**, respectively) services.

ATM is also a capability that can be offered as an end-user service by service providers (as a basis for tariffed services) or as a networking infrastructure for these and other services. The most basic service building block is the ATM virtual circuit, which is an end-to-end connection that has defined end points and routes but does not have bandwidth dedicated to it. Bandwidth is allocated on demand by the network as users have traffic to transmit. ATM also defines various classes of service to meet a broad range of application needs.

8.4.2 ATM Technology

In ATM networks, all information is formatted into fixed-length cells consisting of 48 bytes (8 bits per byte) of payload and 5 bytes of cell header.

Header	Payload
5 bytes	48 bytes

Fig. 8.12 An ATM cell

Bits	4	8	16	3	1	8
	GFC	VPI	VCI	PT	CLP	HEC

User-Network Interface (UNI)

Bits	12	16	3	1	8
	VPI	VCI	PT	CLP	HEC

Network-Node Interface (NNI)

Fig. 8.13 Cell Header

The various fields of ATM Cell Header are discussed below:

- **GFC (Generic Flow Control) :** It permits transmissions of several terminals to be multiplexed on the same user interface—user to network traffic.
- **VPI/VCI (Virtual Path/Circuit Identifiers) :** This field contains 24/28 bits UNI/NNI. VPI field of 8 to 12 bits allows 256 to 4096 virtual paths each of 64,000 VCIs.
- **PT (Payload Type) :** It identifies payload type for OAM (Operations, Administration and Maintenance). Also reserved bits for explicit congestion control.
- **CLP (Cell Loss Priority) :** It is used for buffer management
  - ☐ If set (1) then the cell may be discarded according to network conditions - congestion control.
  - ☐ If not set (0) cell cannot be discarded.
- **HEC (Header Error Control) :** It is used to detect and correct errors in the header
  - ☐ Single bit error—correction is attempted
  - ☐ Multiple bit errors—cell discarded
  - ☐ Calculated using the polynomial $x^8 + x^2 + x + 1$

VPI / VCI

- ATM is connection oriented.
- Connections are identified by the Virtual Channel Identifier.
- A virtual channel is established at connection time and torn down at termination time.
- Multiple VCIs may be used for a multi-component service e.g. sound and video over separate VCIs in video-telephony.
- ATM switches perform switching on a per VCI basis.

The fixed cell size ensures that time-critical information such as voice or video is not adversely affected by long data frames or packets. The header is organized for efficient switching in high-speed hardware implementations and carries payload-type information, virtual-circuit identifiers, and header error check.

ATM is connection oriented. Organizing different streams of traffic in separate calls allows the user to specify the resources required and allows the network to allocate resources based on these needs. Multiplexing multiple streams of traffic on each physical facility (between the end user and the network or between network switches)—combined with the ability to send the streams to many different destinations—enables cost savings through a reduction in the number of interfaces and facilities required to construct a network.

ATM standards defined two types of ATM connections: virtual path connections (VPCs), which contain virtual channel connections (VCCs). **A virtual channel connection** (or virtual circuit) is the basic unit, which carries a single stream of cells, in order, from user to user. A collection of virtual circuits can be bundled together into a virtual path connection. **A virtual path connection** can be created from end-to-end across an ATM network. In this case, the ATM network does not route cells belonging to a particular virtual circuit. All cells belonging to a particular virtual path are routed the same way through the ATM network, thus resulting in faster recovery in case of major failures.

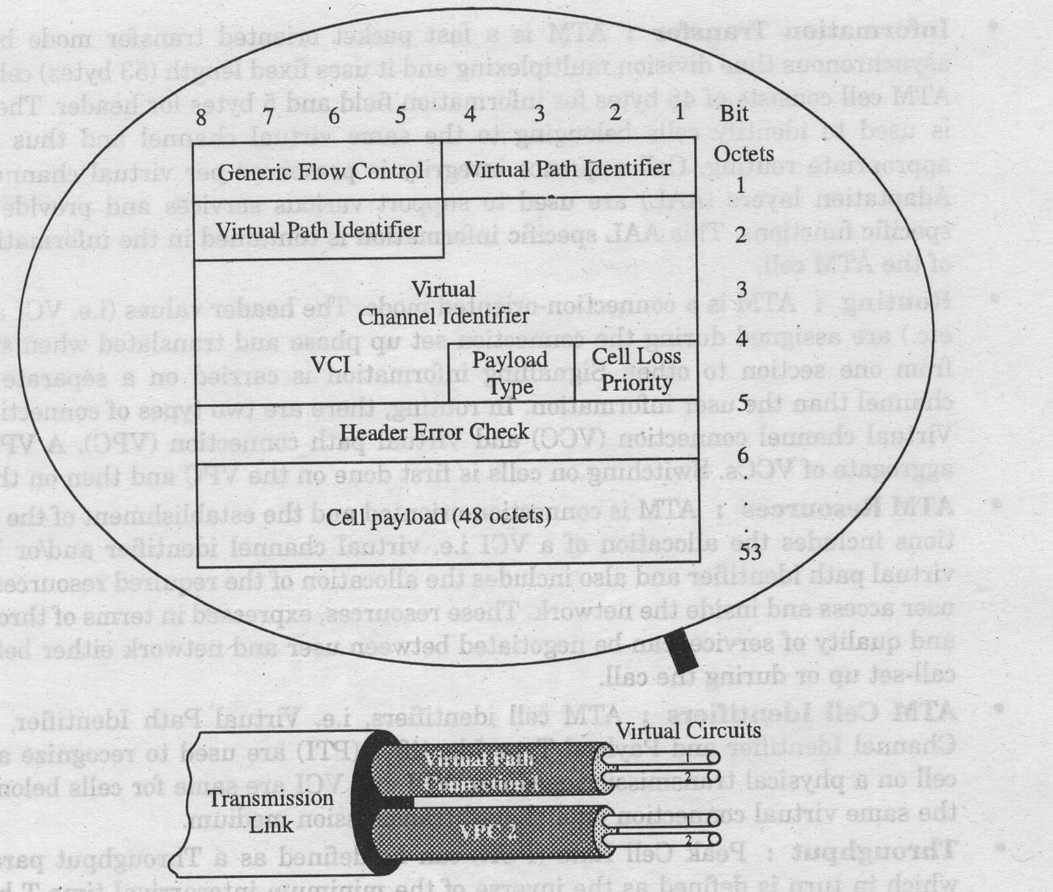


Fig. 8.14 Fixed-length cells

An ATM network also uses virtual paths internally for the purpose of bundling virtual circuits together between switches. Two ATM switches may have many different virtual channel connections between them, belonging to different users. These can be bundled by the two ATM switches into a virtual path connection. This can serve the purpose of a virtual trunk between the two switches. This virtual trunk can then be handled as a single entity by, perhaps, multiple intermediate virtual path cross connects between the two virtual circuit switches.

Virtual circuits can be statically configured as **permanent virtual circuits (PVCs)** or dynamically controlled via signaling as **switched virtual circuits (SVCs)**. They can also be point-to-point or point-to-multipoint, thus providing a rich set of service capabilities. SVCs are the preferred mode of operation because they can be dynamically established, thus minimizing reconfiguration complexity.

8.4.3 Basic Concepts in ATM

The basic concepts in ATM are discussed below:

- **Information Transfer :** ATM is a fast packet oriented transfer mode based on asynchronous time division multiplexing and it uses fixed length (53 bytes) cells. Each ATM cell consists of 48 bytes for information field and 5 bytes for header. The header is used to identify cells belonging to the same virtual channel and thus used in appropriate routing. Cell sequence integrity is preserved per virtual channel. ATM Adaptation layers (AAL) are used to support various services and provide service specific functions. This AAL specific information is contained in the information field of the ATM cell.
- **Routing :** ATM is a connection-oriented mode. The header values (i.e. VCI and VPI etc.) are assigned during the connection set up phase and translated when switched from one section to other. Signalling information is carried on a separate virtual channel than the user information. In routing, there are two types of connections i.e., Virtual channel connection (VCC) and Virtual path connection (VPC). A VPC is an aggregate of VCCs. Switching on cells is first done on the VPC and then on the VCC.
- **ATM Resources :** ATM is connection-oriented and the establishment of the connections includes the allocation of a VCI i.e. virtual channel identifier and/or VPI i.e. virtual path identifier and also includes the allocation of the required resources on the user access and inside the network. These resources, expressed in terms of throughput and quality of service, can be negotiated between user and network either before the call-set up or during the call.
- **ATM Cell Identifiers :** ATM cell identifiers, i.e. Virtual Path Identifier, Virtual Channel Identifier and Payload Type Identifier (PTI) are used to recognize an ATM cell on a physical transmission medium. VPI and VCI are same for cells belonging to the same virtual connection on a shared transmission medium.
- **Throughput :** Peak Cell Rate (PCR) can be defined as a Throughput parameter, which in turn is defined as the inverse of the minimum interarrival time T between two consecutive basic events and T is the peak emission interval of the ATM connection. PCR applies to both constant bit rate (CBR) and variable bit rate (VBR) services for ATM connections. It is an upper bound of the cell rate of an ATM connection and there is another parameter sustainable cell rate (SCR) allows the ATM network to allocate resources more efficiently.
- **Quality Of Service :** Quality of Service (QoS) parameters includes cell loss, the delay and the delay variation incurred by the cells belonging to the connection in an ATM network. QoS parameters can be either specified explicitly by the user or implicitly associated with specific service requests. A limited number of specific QoS classes will be standardized in practice.
- **Usage Parameter Control :** In ATM, excessive reservation of resources by one user affects traffic for other users. So the throughput must be policed at the user-network interface by a Usage Parameter Control function in the network to ensure that the negotiated connection parameters per VCC or VPC between network and subscriber

is maintained by each other user. Traffic parameters describe the desired throughput and QoS in the contract. The traffic parameters are to be monitored in real time at the arrival of each cell. CCITT recommends a check of the peak cell rate (PCR) of the high priority cell flow (CLP = 0) and a check of the aggregate cell flow (CLP = 0 + 1), per virtual connection.

- **Flow Control :** In order to control the flow of traffic on ATM connections from a terminal to the network, a Generic Flow Control (GFC) mechanism is proposed by CCITT at the User to Network Interface (UNI). This function is supported by GFC field in the ATM cell header. Two sets of procedures are associated with the GFC field i.e., Uncontrolled Transmission which is for use in point-to-point configurations and Controlled Transmission which can be used in both point-to-point and shared medium configurations.

8.4.4 ATM Switching

ATM Switching is also known as fast packet switching. ATM switching node transports cells from the incoming links to outgoing links using the routing information contained in the cell header and information stored at each switching node using connection set-up procedure. Two functions at each switching node are performed by a connection set up procedure.

1. A unique connection identifier at the incoming link and the link identifier and a unique connection identifier at the outgoing link are defined for each connection.
2. Routing tables at each switching node are set up to provide an association between the incoming and outgoing links for each connection. VPI and VCI are the two connection identifiers used in ATM cells.

Thus the basic functions of an ATM switch can be stated as follows:

- Routing (space switching) that indicates how the information is internally routed from the inlet to outlet.
- Queueing that is used in solving contention problems if 2 or more logical channels contend for the same output.
- And final function is header translation that all cells, which have a header equal to some value j on incoming link, are switched to outlet and their header is translated to a value k. There are various switching networks existing and available from various manufacturers and research institutes for ATM switch architecture.

Performance Issues

There are 5 parameters that characterize the performance of **ATM** switching systems. They are discussed below:

1. **Throughput :** This can be defined as the rate at which the cells depart the switch measured in the number of cell departures per unit time. It mainly depends on the technology and dimensioning of the ATM switch. By choosing a proper topology of the switch, the throughput can be increased.

2. **Connection Blocking Probability :** Since ATM is connection oriented, there will be a logical connection between the logical inlet and outlet during the connection set up phase. Now the connection blocking probability is defined as the probability that there are not enough resources between inlet and outlet of the switch to assure the quality of all existing as well as new connection.
3. **Cell Loss Probability :** In ATM switches, when more cells than a queue in the switch can handle will compete for this queue, cells will be lost. This cell loss probability has to be kept within limits to ensure high reliability of the switch. In Internally Non-Blocking switches, cells can only be lost at their inlets/outlets. There is also possibility that ATM cells may be internally misrouted and they reach erroneously on another logical channel. This is called Cell Insertion Probability.
4. **Switching Delay :** This is the time to switch an ATM cell through the switch. The typical values of switching delay range between 10 and 1000Micro secs. This delay has two parts. 1.Fixed Switching Delay and it is because of internal cell transfer through the hardware. 2.Queueing delay and this is because of the cells queued up in the buffer of the switch to avoid the cell loss.
5. **Jitter on the Delay :** This is also called as Cell Delay Variation (CDV) and this is denoted as the probability that the delay of the switch will exceed a certain value. This is called a quantile.

8.4.5 Benefits of ATM

The high-level benefits delivered through ATM services deployed on ATM technology using international ATM standards can be summarized as follows:

- High performance via hardware switching with terabit switches on the horizon.
- Dynamic bandwidth for bursty traffic meeting application needs and delivering high utilization of networking resources; most applications are or can be viewed as inherently bursty; data applications are LAN-based and are very bursty, voice is bursty, as both parties are neither speaking at once nor all the time; video is bursty, as the amount of motion and required resolution varies over time.
- Class-of-service support for multimedia traffic allowing applications with varying throughput and latency requirements to be met on a single network.
- Scalability in speed and network size supporting link speeds of T1/E1 to OC-12 (622 Mbps) today and into the multi-Gbps range before the end of the decade; networks that scale to the size of the telephone network (i.e., as required for residential applications) are visualized.
- Common LAN/WAN architecture allowing ATM to be used consistently from one desktop to another; traditionally, LAN and WAN technologies have been very different, with implications for performance and interoperability.
- Opportunities for simplification via switched VC architecture; this is particularly for LAN-based traffic that today is connectionless in nature; the simplification possible through ATM VCs could be in areas such as billing, traffic management, security, and configuration management.

- International standards compliance in central-office and customer-premises environments allowing for multivendor operation.

ATM Applications

There are several practical applications using ATM Technology. ATM is going to be the Backbone Network for many broadband applications including Information SuperHighway. Some of the key applications can be mentioned as follows :

- Video Conferencing
- Desktop Conferencing
- Multimedia Communications
- ATM Over Satellite Communications
- Mobile Computing over ATM for Wire-less Networks

ATM technologies, standards, and services are being applied in a wide range of networking environments, as described briefly below (see Figure 8.15).

- **ATM services :** Service providers globally are introducing or already offering ATM services to their business users.
- **ATM workgroup and campus networks :** Enterprise users are deploying ATM campus networks based on the ATM LANE standards. Workgroup ATM is more of a niche market with the wide acceptance of switched-Ethernet desktop technologies.
- **ATM enterprise network consolidation :** A new class of product has evolved as an ATM multimedia network-consolidation vehicle. It is called an ATM enterprise network switch. A full-featured ATM ENS offers a broad range of in-building (e.g., voice, video, LAN, and ATM) and wide-area interfaces (e.g., leased line, circuit

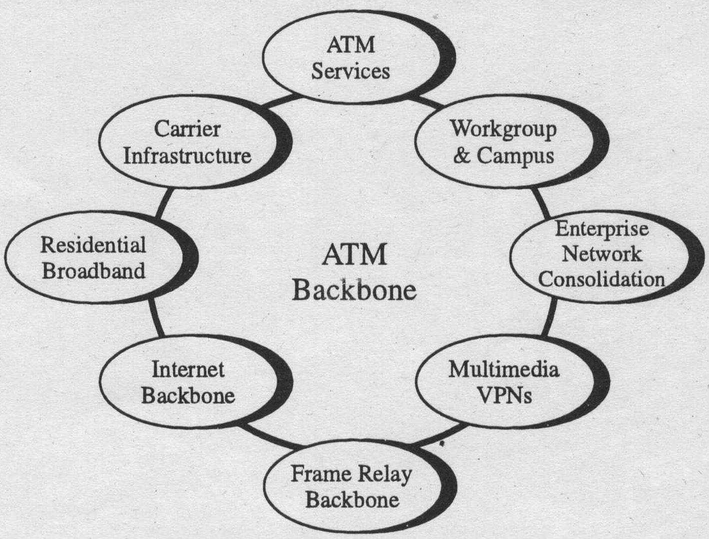


Fig. 8.15 ATM technologies standards and services

switched, frame relay, and ATM at narrowband and broadband speeds) and supports ATM switching, voice networking, frame-relay SVCs, and integrated multiprotocol routing.

- **Multimedia virtual private networks and managed services :** Service providers are building on their ATM networks to offer a broad range of services. Examples include managed ATM, LAN, voice and video services (these being provided on a per-application basis, typically including customer-located equipment and offered on an end-to-end basis), and full-service virtual private-networking capabilities (these including integrated multimedia access and network management).
- **Frame-relay backbones :** Frame-relay service providers are deploying ATM backbones to meet the rapid growth of their frame-relay services to use as a networking infrastructure for a range of data services and to enable frame relay to ATM service interworking services.
- **Internet backbones :** Internet service providers are likewise deploying ATM backbones to meet the rapid growth of their frame-relay services, to use as a networking infrastructure for a range of data services, and to enable Internet class-of-service offerings and virtual private intranet services.
- **Residential broadband networks :** ATM is the networking infrastructure of choice for carriers establishing residential broadband services, driven by the need for highly scalable solutions.
- **Carrier infrastructures for the telephone and private-line networks :** Some carriers have identified opportunities to make more-effective use of their SONET/SDH fiber infrastructures by building an ATM infrastructure to carry their telephony and private-line traffic.

• • •

9

Data Link Layer

Introduction

In this chapter we will study the design of layer 2, the data link layer (also known as the physical link control layer). The purpose of the data link layer is to transfer blocks of data without error between two adjacent devices. Adjacent devices are physically connected by a communication channel such as telephone lines, coaxial cables, optical fibres, or satellites. The implication of such a physical link is that the data bits are delivered in exactly the same order in which they are sent. The physical link has no inherent storage capacity; therefore the delay involved is the propagation delay over the link.

Transmission of data over the link would be very simple indeed if no error ever occurred. Unfortunately, this is not so in a real physical link for a number of reasons:

- Natural phenomena such as noises and interference are introduced into the link causing errors in detecting the data.
- There is a propagation delay in the link.
- There is a finite data processing time required by the transmitting and receiving stations.

A data link protocol thus has to be designed to ensure an error-free transmission and also to achieve an efficiency of the data transfer as high as possible. In this chapter we will discuss the services the datalink layer of the networking stack is required to have and the various existing layer 2 protocols, which cater to these requirements in different ways. The job of the datalink layer is to transfer the bits to the destination machine, which then can be handed over to the network layer for processing.

The datalink layer doesn't worry about getting bits on and off the medium. The design issues at this layer are the low-level concerns of communication: what about errors? What if the sender outstrips the receiver? How do we distinguish packets within a bit stream? How can we make the best use of the channel? What isn't considered at this layer is anything more than a single-hop away, i.e., no networking issues. All communication issues are point-to-point at this layer and the stream of bits is sequential.

One potential point of confusion when studying concepts like error and flow control: these issues arise in multiple layers, depending on the application, the medium, the design of the stack, etc.

Logical-Link

A Logical-link assumes that the sender is directly connected to the receiver.

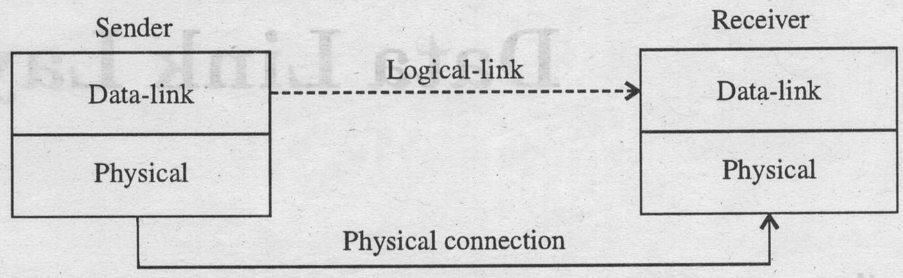


Fig. 9.1 Logical-link shown as dotted lines between sender and receiver

Actually the physical connection may go via any number of intermediate devices, including hubs and bridges.

9.1 Purpose of Data Link Layer

The goal of the data link layer is to provide reliable, efficient communication between adjacent machines connected by a single communication channel. Specifically:

1. Group the physical layer bit stream into units called frames. Note that frames are nothing more than "packets" or "messages". By convention, we'll use the term "frames" when discussing DLL packets.
2. Sender checksums the frame and sends checksum together with data. The checksum allows the receiver to determine when a frame has been damaged in transit.
3. Receiver recomputes the checksum and compares it with the received value. If they differ, an error has occurred and the frame is discarded.
4. Perhaps return a positive or negative acknowledgment to the sender. A positive acknowledgment indicate the frame was received without errors, while a negative acknowledgment indicates the opposite.
5. Flow control—Prevent a fast sender from overwhelming a slower receiver. For example, a supercomputer can easily generate data faster than a PC can consume it.
6. In general, provide service to the network layer. The network layer wants to be able

to send packets to its neighbors without worrying about the details of getting it there in one piece.

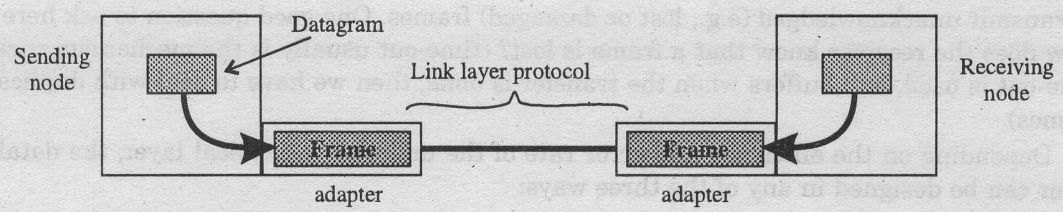


Fig. 9.2 Transmission of frames between sending receiving nodes

9.2 Design Issues

The datalink layer is designed to offer the following functionalities:

1. Services Provided to the Network Layer.
2. Framing.
3. Error Control.
4. Flow Control.

1. Services Provided to the Network Layer

Reliable delivery

Frames are delivered to the receiver reliably and in the same order as generated by the sender.

- The data link layer actually computes and verifies the checksum
- It maintain connection state which keeps track of frame orders and which frames require re-transmission.

Best effort

The receiver may not return acknowledgments to the sender, so the sender has no way of knowing if a frame has been successfully delivered. When would such a service be appropriate?

- When the error rate is very low and higher layer may be able to recover the errors by themselves, the data link layer doesn't have to provide the error recovery service.
- In some real-time applications requiring "better never than later" semantics. Old data may be worse than no data. For example, when a real-time speech is delivered over the network, later arrived speech clips are worse than simply dropping off a few of them.

Acknowledged Delivery

The receiver returns an acknowledgment frame to the sender indicating that a data frame was properly received. The sender keeps connection state, but may not necessarily retransmit unacknowledged frames. The receiver may receive frames out of order. The datalink layer in this case just hands the received frames to the higher layer and let the higher layer decide what to do.

Typically, each frame is assigned a unique sequence number, which the receiver returns in an acknowledgment frame to indicate which frame the ACK refers to. The sender must retransmit unacknowledged (e.g., lost or damaged) frames. One good question to ask here is : How does the receiver know that a frame is lost? (time-out usually is the mechanism used; if time-out is used, and buffers when the transfer is done, then we have to deal with duplicated frames)

Depending on the efficiency and error rate of the underlying physical layer, the datalink layer can be designed in any of the three ways:

(a) **Unacknowledged connectionless service**—This consists of the source host sending independent frames to the destination host without any sort of feedback/acknowledgment mechanism. Frame recovery due to line noise is left to the network layer. This kind of service is appropriate for high speed, low error communication channels such as an optical fiber network.

(b) **Acknowledged connectionless service**—Communication channels that are more error prone and hence require more reliable communication implement a feedback mechanism for each frame sent between two hosts. This enables the sender to know if the frame has arrived correctly or not. This ack sending mechanism is not a requirement but an optimization, because the transport layer can always send a message and wait for it to be acknowledged. However a message, (a unit of data in the transport layer), consists of several frames, (unit of data in the datalink layer), so the re-transmission of each faulty received message would bear a lot of overhead. Such a mechanism is useful on wireless channels that are inherently unreliable.

(c) **Acknowledged connection-oriented service**—This is the most sophisticated service the datalink layer can provide to the network layer. In this service the source and destination hosts establish a connection before any transfer of data takes place. Each frame is received in the same order it is sent in. The service also guarantees that each frame is received only once. Communication between two hosts takes place in three phases. First phase is connection setup, during which each side initializes counters and variables to keep track of frames. Second phase consists of actual frame transmission and the third phase consists of connection release, freeing up the resources

2. Framing

The physical layer deals with bit streams. Everything the datalink layer does is in the context of frames. Frames can be either bit or character oriented. The datalink layer receives a raw bit stream from the underlying physical layer. This bit stream is not guaranteed to be error free. On a noisy communication channel the received number of bits maybe less/more and/or different than the ones transmitted. In order to provide a reliable transfer to the network layer the data link layer breaks the bit stream into frames and computes the checksum for each frame. This checksum is also transmitted along with the frame. The destination host on receiving a frame computes another checksum from its data and compares it to the one transmitted. This enables the datalink layer of the destination host to detect and possibly correct frames, depending upon the method the checksum is computed in.

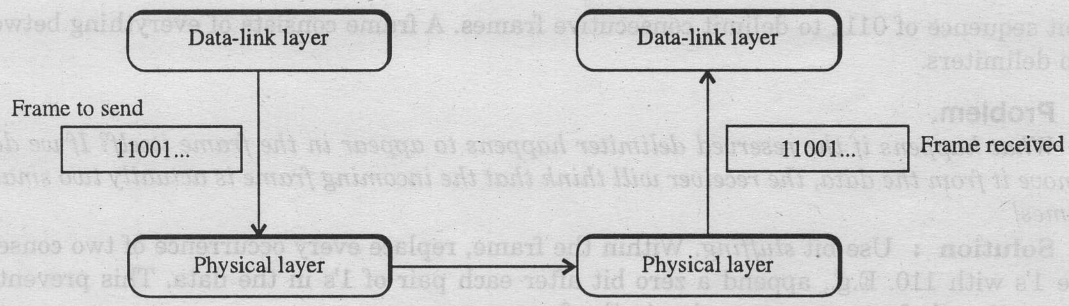
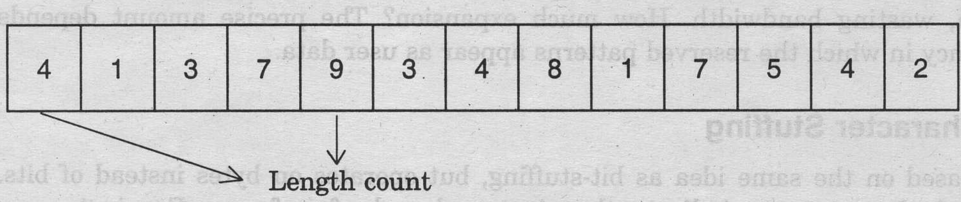


Fig. 9.3 Framing done at data link layer

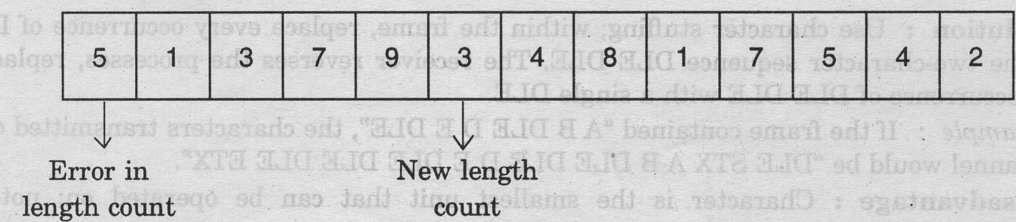
A big question for the receiving datalink layer is how does it convert the stream of bits from the physical layer into a sequence of frames or packets? How can the receiver detect frame boundaries? That is, how can the receiver recognize the start and end of a frame? This issue is more complicated than it might seem. Some possible approaches are as follows:

(a) Length Count

Make the first field in the frame's header be the length of the frame. That way the receiver knows how big the current frame is and can determine where the next frame begins. Example:



Disadvantage : Receiver loses synchronization when bits become garbled. If the bits in the count become corrupted during transmission, the receiver will think that the frame contains fewer (or more) bits than it actually does. Although checksum will detect the incorrect frames, the receiver will have difficulty resynchronizing to the start of a new frame. The above example is repeated with an error in length count.



This technique is not used anymore, since better techniques are available.

(b) Bit Stuffing

It uses reserved bit patterns to indicate the start and end of a frame. For instance, use the

4-bit sequence of 0111 to delimit consecutive frames. A frame consists of everything between two delimiters.

■ **Problem.**

What happens if the reserved delimiter happens to appear in the frame itself? If we don't remove it from the data, the receiver will think that the incoming frame is actually two smaller frames!

Solution : Use bit *stuffing*. Within the frame, replace every occurrence of two consecutive 1's with 110. E.g., append a zero bit after each pair of 1's in the data. This prevents 3 consecutive 1's from ever appearing in the frame.

Likewise, the receiver converts two consecutive 1's followed by a 0 into two 1's, but recognizes the 0111 sequence as the end of the frame.

Example : The frame "1011101" would be transmitted over the physical layer as "0111101101010111".

Note : When using bit stuffing, locating the start/end of a frame is easy, even when frames are damaged. The receiver simply scans arriving data for the reserved patterns. Moreover, the receiver will resynchronize quickly with the sender as to where frames begin and end, even when bits in the frame get garbled.

The main **disadvantage** with bit stuffing is the insertion of additional bits into the data stream, wasting bandwidth. How much expansion? The precise amount depends on the frequency in which the reserved patterns appear as user data.

(c) Character Stuffing

It is based on the same idea as bit-stuffing, but operates on bytes instead of bits. It uses reserved characters to indicate the start and end of a frame. For instance, use the two-character sequence DLE STX (Data-Link Escape, Start of TeXt) to signal the beginning of a frame, and the sequence DLE ETX (End of TeXt) to flag the frame's end.

■ **Problem.**

What happens if the two-character sequence DLE ETX happens to appear in the frame itself?

Solution : Use character stuffing; within the frame, replace every occurrence of DLE with the two-character sequence DLE DLE. The receiver reverses the processes, replacing every occurrence of DLE DLE with a single DLE.

Example : If the frame contained "A B DLE D E DLE", the characters transmitted over the channel would be "DLE STX A B DLE DLE D E DLE DLE DLE ETX".

Disadvantage : Character is the smallest unit that can be operated on; not all architectures are byte oriented.

(d) Signal violations

The last technique is to use a signal for the start/stop frames that is normally not used to transmit data. This only works if there are such "illegal" signals in the coding scheme. For example, a Manchester code works like this:

The transition represents the data. Low-to-high is a 1, high-to-low a 0.

An illegal signal would then be a bit period where no transition takes place (i.e., high-high, or low-low)

The advantage of encoding violations is that no extra bandwidth is required as in bit-stuffing. Synchronization is another big advantage (as there is a always a transition)

The IEEE 802.4 standard uses this approach.

Finally, some systems use a combination of these techniques. IEEE 802.3, for instance, has both a length field and special frame start and frame end patterns.

3. Error Control

A bad communication channel can cause a variety of strange events during communication, like flipping of bits, losing bits from a frame, new bits in the frame, frames completely disappearing, either host going down.

Error control is concerned with ensuring that all frames are eventually delivered (possibly in order) to a destination. **Three items are required for it:**

Acknowledgements

The basic idea for error control is to provide a feedback channel from the receiver to the sender. This sort of feedback is known as acknowledgments, or ACKs.

Typically, reliable delivery is achieved using the "acknowledgments with retransmission" paradigm, whereby the receiver returns a special acknowledgment (ACK) frame to the sender indicating the correct receipt of a frame. In some systems, the receiver also returns a negative acknowledgment (NACK) for incorrectly-received frames. This is nothing more than a hint to the sender so that it can retransmit a frame right away without waiting for a timer to expire.

Timers

One problem that simple ACK/NACK schemes fail to address is recovering from a frame that is lost, and as a result, fails to solicit an ACK or NACK. What happens if an ACK or NACK becomes lost?

Retransmission timers are used to resend frames that don't produce an ACK. When sending a frame, schedule a timer to expire at some time after the ACK should have been returned. If the timer goes off, retransmit the frame.

Sequence Numbers

Retransmissions introduce the possibility of duplicate frames. To suppress duplicates, add sequence numbers to each frame, so that a receiver can distinguish between new frames and old copies.

This whole managing of timers and frame sequencing is an integral part of data link layer design.

4. Flow Control

It basically controls the rate of data transmission between two hosts. This is another of the important issues in the design of the data link layer, how to coordinate the rate of transmission between two hosts. What if the sender is sending data at a much faster rate than the receiver can process/accept, causing dropping of packets at the receiver end which further causes the sender to timeout on the ACK packets it is expecting, causing retransmission, leading to a less efficient network. The sender in such a case is usually throttled by using a feedback mechanism. There are various ways in which the receiver communicates to the sender using control frames about the number/rate of packets it can accept comfortably.

Usually, this is a dynamic process, as the receiving speed depends on such changing factors as the load, and availability of buffer space. One solution is to have the receiver extend credits to the sender. For each *credit*, the sender may send one frame. Thus, the receiver controls the transmission rate by handing out credits.

9.3 Error Detection And Correction

In data communication, line noise is a fact of life (e.g., signal attenuation, natural phenomenon such as lightning, and the telephone repairman). Moreover, noise usually occurs as bursts rather than independent, single bit errors. Data processing and transmission systems use a variety of techniques to detect and correct errors that occur, usually for any of the following reasons:

- Electrostatic interference from nearby machines or circuits.
- Attenuation of the signal caused by a resistance to current in a cable.
- Distortion due to inductance and capacitance.
- Loss in transmission due to leakages.

It has been estimated that an error occurs for every 1 in 200,000 bits. While most LAN technologies and optical cable networks reduce errors considerably, wireless networks and WAN links can have high error rates.

Types of errors

- **Bit errors** are errors that corrupt single bits of a transmission, turning a 1 into a 0, and vice versa. These errors are caused by power surges and other interferences.
- **Packet errors** occur when packets are lost or corrupted. Packet loss can occur during times of network congestion when buffers become full and network devices start discarding packets. Errors and packet loss also occur during network link failures.

Here we consider how to determine whether a logical unit of data has been received correctly (after either data transmission or storage), and if it is determined to be incorrect, how to attempt to correct it. Detecting and correcting errors requires **redundancy**—sending additional information along with the data. There are two types of attacks against errors:

- **Error correction strategy :** Send enough additional information to correct